Addressing MAP-21 Freight Objectives using GPS Data

CFIRE 09-18
December 2016

National Center for Freight & Infrastructure Research & Education
Department of Civil and Environmental Engineering
College of Engineering
University of Wisconsin–Madison

Authors:
Mihalis Golias, Stephanie Ivey (University of Memphis)
MD Sarder, David Holt (University of Southern Mississippi)
Mike Anderson (University of Alabama-Huntsville)
Maria Hart (University of Wisconsin-Madison)

Principal Investigator:
Sabyasachee Mishra, Ph.D., PE
University of Memphis

(Blank page in the template report)

Technical Report Documentation

DISCLAIMER

This research was funded by the National Center for Freight and Infrastructure Research and Education. The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange.

The U.S. Government assumes no liability for the contents or use thereof. The contents do not necessarily reflect the official views of the National Center for Freight and Infrastructure Research and Education, the University of Wisconsin, the Wisconsin Department of Transportation, or the USDOT's RITA at the time of publication. The United States Government assumes no liability for its contents or use thereof. This report does not constitute a standard, specification, or regulation.

The United States Government does not endorse products or manufacturers. Trade and manufacturers names appear in this report only because they are considered essential to the object of the document.

FOREWARD

Freight planning and operation perspectives of MAP-21 includes development of a national freight plan to address freight congestion bottlenecks, connectivity enhancement of major intermodal centers, and determination of barriers to improve freight performance. Currently state DOTs are at a very early stage in their development of freight performance measures (FPMs) to meet MAP-21 objectives. Some states are reviewing methodologies and appropriate data needs to assess their feasibility in developing FPMs. The goal of this project is to develop FPMs to meet MAP-21 objectives and apply the methodology in CFIRE region.

ACKNOWLEDGEMENT

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

Freight transportation plays a significant role in national, state, and local economies, thus the performance of freight networks are of great concern. The projected rise in freight volumes only strengthens these concerns highlighting the need for Freight Performance Measures (FPMs) that can be used to monitor and identify issues within the transportation network.

The current transportation bill, Moving Ahead for Progress in 21st Century (MAP-21), indicates a freight plan to address freight congestion bottlenecks, identify critical major intermodal centers to enhance connectivity, determine barriers to improved freight performance, and explore the critical sections of the transportation network that need prioritization in resource allocation to enhance Freight Performance Measures (FPMs).

GPS technology provides a new avenue for the estimation of FPMs that breaks away from costly data collection methods such as spot count data and roadside interviews. Researchers developed several approaches to analyze truck GPS data and estimate network and freight facility FPMs, but issues such as the device spatial error, identifying stops and trip ends, effect of non-recurring congestion still remain a challenge.

The American Transportation Research Institute (ATRI) in collaboration with the Federal Highway Administration (FHWA) developed the Freight Performance Measures Web-Based (FPMweb) Tool in 2011 to estimate operating speeds in 25 interstate corridors using GPS data. Average speed values can be retrieved for a given state, corridor, year, month, day, and time of the day but the tool cannot be used to forecast truck volumes and speeds or provide any other FPMs. Other FPMs that can be obtained using GPS data are: travel time reliability, connectivity and resiliency of intermodal facilities, short term and long term travel time predictions, and temporal and spatial patterns of travel time/speed/volume variation. These FPMs vary by urban typologies (rural, suburban, and urban), by functional class (freeway, arterials), by trip type (short or long by distance), by origin and destination (II, IE, EI, EE)[1] and by agency (private and public sector).

The scope of this project was to evaluate the applicability of GPS truck data in developing FPMs at the local, regional, and state level using the CFIRE region as a case study. The major goals of the project are to: (a) provide a set of comprehensive FPMs that can provide insight into functioning of the multifaceted freight transportation network, and (b) examine the CFIRE freight network and compute FPMs using truck GPS data to address MAP-21 objectives.

[1] Internal-Internal(II), Internal-External(IE), External Internal (EI), and External External (EE)

1. INTRODUCTION

Freight transportation in the United States is expected to grow 23.5% until 2025 with associated revenues up to 72% according to the American Trucking Associations (ATA) and IHS Global Insight[2]. This increase in freight volume and its impact on the nation's freight network has raised great concern over the anticipated network's performance. The current transportation bill, Moving Ahead for Progress in 21st Century (MAP-21), acknowledges the significance of freight and its impact on national, state, local and regional networks and suggests a national and state strategic freight plan to assess and improve freight corridors' condition and performance.

The goals of this national freight policy are: (i) to invest in infrastructure and operational improvements that strengthen the contribution of the national freight network to the economic competitiveness of the United States; reduce congestion; and increase productivity; (ii) to improve the safety, security, and resilience of freight transportation; (iii) to improve the state of good repair of the national freight network; (iv) to use advanced technology to improve the safety and efficiency of the national freight network; (v) to incorporate concepts of performance, innovation, competition, and accountability into the operation and maintenance of the national freight network; (vi) to improve the economic efficiency of the national freight network; and (vii) to reduce the environmental impacts of freight movement on the national freight network (Moving Ahead for Progress in the 21st Century Act, 2012). MAP-21 indicates a freight plan to address freight congestion bottlenecks, identify critical major intermodal centers to enhance connectivity, determine barriers to improved freight performance, and explore the critical sections of the transportation network that need prioritization in resource allocation to enhance Freight Performance Measures (FPMs).

FPMs estimation in the US has advanced with the utilization of truck GPS data by private and public agencies at the end of the 20th century. Before that, data collection has been a challenging task as it required spot count data and roadside interviews, methods that provided inadequate information and is usually time consuming and costly. Information provided by GPS data includes spatial information (X and Y coordinates), time stamp, heading, spot speed, and a unique truck identifier. Additional information can also be obtained such as weather conditions, distance, fuel consumption etc. Since the GPS utilization is a new concept in FPMs estimation there are still many obstacles to be addressed. Researchers developed several approaches to analyze truck GPS data and estimate network and freight facility FPMs, but issues such as the device spatial error, identifying stops and trip ends, effect of non-recurring congestion still remain a challenge. The American Transportation Research Institute (ATRI) in collaboration with the Federal Highway Administration (FHWA) developed the Freight Performance Measures Web-

[2] http://www.trucking.org/article.aspx?uid=41434598-4c60-444d-bc83-38f06ded539d

Based (FPMweb) Tool in 2011 to estimate operating speeds in 25 interstate corridors using GPS data. Average speed values can be retrieved for a given state, corridor, year, month, day, and time of the day but the tool cannot be used to forecast truck volumes and speeds or provide any other FPMs. Other FPMs that can be obtained using GPS data are: travel time reliability, connectivity and resiliency of intermodal facilities, short term and long term travel time predictions, and temporal and spatial patterns of travel time/speed/volume variation. These FPMs vary by urban typologies (rural, suburban, and urban), by functional class (freeway, arterials), by trip type (short or long by distance), by origin and destination (II, IE, EI, EE)[3] and by agency (private and public sector).

1.1 Project Purpose and Scope

The scope of this project is to evaluate the applicability of GPS truck data in developing FPMs at the local, regional, and state level using the CFIRE region as a case study. The major goals of the project are to: (a) provide a set of comprehensive FPMs that can provide insight into functioning of the multifaceted freight transportation network, and (b) examine the CFIRE freight network and compute FPMs using truck GPS data to address MAP-21 objectives.

1.2 Report Organization

This report is organized as follows: Chapter 2 provides an up-to-date literature review on practices used in freight performance measures in the public and private sectors using truck GPS data. Chapter 3 describes the data collection and methodology used to analyze data in the CFIRE region followed data processing methodology in Chapter 4. Chapter 5 presents the suggested FPMs and how they are computed for the study area. Chapter 6 presents the linkage of the FPMs to workforce development. Chapter 7 presents the conclusions and recommendations for future research.

[3] Internal-Internal(II), Internal-External(IE), External-Internal (EI), and External-External (EE)

2. LITERATURE REVIEW

In the following paragraphs the reader will be introduced in a thorough literature review that has been conducted for each topic this study addresses. The literature review presented in this section summarizes past and contemporary published work and relevant studies conducted for: (i) road network reliability, (ii) truck parking demand analysis, and (iii) freight performance measures using truck GPS data.

2.1 Road Network Reliability Literature

Elementary studies on transportation network reliability appeared as early as the mid-1940s, but the topic attracted greater attention during the 1990s (Murray, et al., 2007). Network reliability can generically be described as the probability that a network will be able to function when specific elements fail. Road network reliability should be treated differently from other networks (e.g., logical or cyber networks), due to the multi-commodity flows and uniqueness of each trip's origin and destination nodes (Bell, et al., 1997) (Iida, 1999). Road networks are by definition both stochastic and dynamic and their functionality depends on a variety of factors that fluctuate with time (e.g., demand and supply, recurrent and non-recurrent events) and can generate instability. Nicholson and Du (Nicholson, et al., 1997) indicate two distinct sources of unreliability in transportation networks: arc flow variations and capacity variations. In the flow case, travel time varies with flow variations (given a constant link capacity), while in the capacity case (given a constant link flow), travel time can vary due to capacity variations. In reality, travel time variations occur due to the combined effects of both sources, but it is difficult to identify the separate effects of each source, per se. In road network reliability analysis, a number of uncertainty factors emerge, including the simplified network representation, the treatment of trips outside the study area, and the non-uniqueness of link reliability definitions (Bell, et al., 1997).

Typically, network reliability analysis examines short-term changes, in demand and capacity (i.e., peak hour demand changes or capacity degradation due to incidents), in contrast to uncertainty network studies that consider long-term changes in demand and supply (e.g., network robustness and vulnerability, respectively). One challenging task (both for researchers and practitioners) has been the development of metrics and evaluation tools that can monitor and support network reliability decision making at all levels (planning, tactical, operational, real time).

2.1.1 Network Reliability basics

Let $G = (V, E)$ be a *stochastic graph* where: V and E are finite sets of nodes and links respectively (Lucet, et al., 1999) (Rebaiaia, et al., 2013). The stochastic graph describes a network whose elements function or fail independently under a specific probability. Network reliability analysis aims to evaluate the global probability of functionality, given the failure/function probabilities of each system element. Let $G' = (V', E')$, such that: $V' \subset V$ *and* $E' \subset (V' \mathrm{x} V') \cap E$ be subgraphs and $G'' = (V, E'')$ such $E'' \subset E$ partial graphs.

The elements of a network can have either a function or fail state only, thus the Boolean cardinality that defines the element states is equal to 2 and there are 2^{n+m} possible network states, ($n = |V|$, $m=|E|$). It should be noted that for typical road network analysis, possible network states are reduced to 2^m as nodes are considered to function at all times (*perfect nodes*). A *path P* can be defined as a chain $\mu = (x_1, \ldots, x_{k+1})$ of links (where $x_1, \ldots, x_{k+1}$, are the *nodes* those links connect) in which the end point of a link *i* is the start point of link *i+1*. Often it is written as $\mu(x_1, x_{k+1})$. To have a connected graph, between any two nodes $x, y \in V$, a chain $\mu(x,y)$ must exist. Similarly, a *cut* set *C* is a set of links such that when they all fail the system fails as well. Accordingly, a *minimal path* or *minpath* can be defined as a path with the minimum required links to keep a pair (or set) of nodes connected (functioning system) and a *minimal cut set* or *mincut* as a set of minimum required links to disconnect a system.

Reliability can be estimated for different numbers of *terminals* (i.e., origin and destination nodes). *K*-terminal reliability is defined as the probability that every node that belongs in $K \subset V$ is connected with all other nodes in *K. All*-terminal reliability is defined as the probability that every node in the network is connected to all other nodes. In general, network reliability evaluation problems, both deterministic and stochastic (Rebaiaia, et al., 2013) (Frank, et al., 1971) (Hwang, et al., 1981) fall under the *NP-complexity* (*hard* or *complete*) category (Rebaiaia, et al., 2013) (Rosenthal, 1974) (Rosenthal, 1977). Even in cases of *K*=2, the problems are considered as *#P-complete* i.e., *numbered P-complete* (Rebaiaia, et al., 2013) (Valiant) (Brecht, et al., 1988).

2.1.2 Main Reliability Definitions

Road network reliability can be generically defined either as “connectivity” or as “travel-time” reliability (Bell, et al., 1997) (Mine, et al., 1982). The major definitions identified in the literature are: i) *Connectivity* (or terminal) *reliability* (Bell, et al., 1997) (Mine, et al., 1982) defined as: “*The probability that there exists at least one path without disruption or heavy delay to a given destination within a given time period”,* ii) *Travel-Time Reliability* (Bell, et al., 1997) (Mine, et al., 1982) defined as: *“The probability that traffic can reach a given destination within a stated time”,* and *iii) Capacity Reliability* (Chen, et al., 1999), defined as: *“The probability that the network can accommodate certain traffic demand on the concept of network reserve capacity”.* Other, less commonly observed, reliability definitions from the literature (Murray, et al., 2007) (Watling, 2008) include: *Behavioral reliability*, *Travel-time budget reliability* (Lo, et al., 2000) (Lo, et al., 2006), *Travel demand satisfaction reliability* (Zhang, et al., 2001), and *Road vulnerability* (Berdica, 2000) (Berdica, 2002).

2.1.3 Connectivity Reliability: Basic Concepts

Link and System Reliability

Reliability r_a of link α, can be defined as the expected value of a state binary (0-1) variable x_α, that equals 1 if link α is not disrupted/congested and zero otherwise (Mine, et al., 1982). System (of links) reliability *R,* can be defined similarly by the expected value of a "structure function" $\varphi(\mathbf{x})$, that can take values of one or zero, if the system is functioning or is congested/disrupted, respectively. Link and system reliability are given by the following equations:

$$r_\alpha = \mathrm{E}\{x_\alpha\}, \tag{1}$$

$$R = \Pr(\varphi(\mathbf{x})=1)=\mathrm{E}\{\varphi(\mathbf{x})\}, \tag{2}$$

In simple link formations (serial or parallel as shown in Figure 1), reliability can be estimated easily, given link function/failure probabilities using equations 3 and 4.

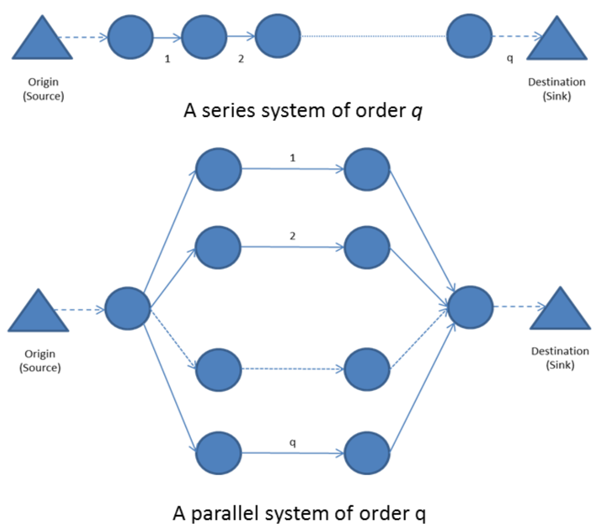


Figure 1 Serial and Parallel System of "q" Links
Source: Bell and Iida (1997)

Series system reliability:

$$R = \mathrm{E}\left\{\prod\nolimits_{i\in S} \mathrm{x}_i\right\} = \prod\nolimits_{i\in S} r_i \tag{3}$$

Parallel system reliability:

$$R = \mathrm{E}\left\{\left(1 - \prod\nolimits_{i\in P}(1-\mathrm{x}_i)\right)\right\} = \left(1 - \prod\nolimits_{i\in P}(1-r_i)\right) \tag{4}$$

Reliability Evaluation

Two main categories of connectivity reliability evaluation approaches have been proposed in literature: topological methods and enumeration methods. The first category is based on topological methods that involve the use of the reduction techniques, the factoring theorem (which is the basis for a class of K-terminal reliability algorithms), and decomposition (Rosenthal, 1977) (Satyanarayana, 1982) (Wood, 1985) (Wood, 1986). Reduction techniques aim to reduce the network size and produce (an easier to evaluate) equivalent network in terms of reliability. Decomposition methods aim to decompose the network into smaller fragments whose reliabilities are integrated into a global network reliability value.

The second category is comprised of enumeration methods and is subcategorized into state and path-cut set enumeration methods. State methods enumerate all the possible stochastic graph states keeping the smaller set between those that allow network functionality and those that lead to failure. Path-cut set methods enumerate either minimal paths or cuts to provide Boolean expressions and then to estimate this expressions probability (Lucet, et al., 1999). To convert Boolean expressions into probabilities one can use the inclusion-exclusion formula (also called Poincaré's Theorem) or the sum of disjoint products. Next we briefly present each method.

2.1.3.1 Exact Methods

Combination Method

The *combination method* (Bell, et al., 1997) is a simplified decomposition procedure, applicable when a system can be expressed as a combination of *series* and *parallel* systems, Global reliability can then be estimated by combining the estimated reliabilities of all subsystems (i.e., using equations 3 and 4). The combination method allows direct estimation of the system reliability but complexity increases if a link appears more than once in the equivalent transformation (use of Boolean algebra is required).

Factoring Methods

Factoring methods (Rubino, 1998) are based on the link *contraction* concept. The idea is to "merge" the two vertices of link *i* and generate a new graph G_i^c (contraction graph) that has one less node and link, than the initial graph *G*. If, on the other hand, link *i* was just deleted, this would result in a (deletion) graph G_i^d, which has the same nodes with initial graph *G*, but one edge less. For *2-terminal* reliability factoring is based on the following equation:

$$R_{s,t}(G) = r_i R_{s,t}(G_i^c) + (1 - r_i) R_{s,t}\left(G_i^d\right) \tag{5}$$

where: r_i is the reliability of link *i* and $R_{s,t}(G)$, $R_{s,t}(G_i^c)$ and $R_{s,t}\left(G_i^d\right)$, are the reliabilities of the initial, contraction and deletion graphs respectively. Equation 5 can be generalized for *K*-terminal reliability, as well.

Decomposition Methods

Decomposition methods (Lucet, et al., 1999) decompose systems to subsystems with reliabilities easier to estimate. Global reliability is found by composing these smaller subsystem reliabilities. The basic principle of decomposition is to split a graph *G* into two subgraphs *L* and *H* separated by an *articulation vertex* (simple case) or a *separating boundary set F*. The *articulation vertex* (or set of vertices *F*), disconnects *L* from *H* and reliability of graph *G* can be estimated as:

$$R(G) = R(H) \cdot R(L) \tag{6}$$

For graphs that *L* is separated from *H* with a boundary set *F*, system reliability can be estimated by the formula:

$$R(G) = \sum_{H_i, L_j \,/\, H(H_i) \cup L(L_j) \text{ is connected}} \Pr(H_i)\Pr(L_j) \tag{7}$$

where: H_i and L_j are the sets of states of subgraphs *H* and *L* respectively

The reader is referred to the literature for a detailed discussion on decomposition algorithms for network reliability estimation (Lucet, et al., 1999) (Rosenthal, 1977) (Rubino, 1998) (Shogan, 1978) (Nakazawa, 1981) (Carlier, et al., 1996).

Enumeration Methods

Enumeration methods (Lucet, et al., 1999) can be categorized into state and paths or cuts (also known as path–and–cut) enumerations, for minimal paths and minimal cut sets, respectively (as described in the Network Reliability Basics section). For more information on enumeration methods and algorithms, than the information given in the next sections, the reader is referred to the literature (Wood, 1986) (Carlier, et al., 1996).

State Enumeration Method

The basic state enumeration principle estimates graph reliability by enumerating all the possible states of a stochastic graph and keep the smaller set between those that allow functionality and those that do not. For a stochastic graph G reliability can be estimated as:

$$\mathrm{R}(G) = \sum_{G(G_i) functions} \Pr(G_i) = 1 - \sum_{G(G_i) fails} \Pr(G_i) \tag{8}$$

For a network comprised of m links, complete state enumeration requires the evaluation of 2^m states.

<u>Path-and-Cut Method</u>
The path-and-cut method is more practical (to state enumeration) as reliability can be estimated from minimal paths (minpaths) or minimal cut sets (mincuts), depending on which are fewer on the corresponding graph, by estimating the probability of Boolean expressions. When selecting to use minpaths, the path enumeration method estimates all minpaths that allow network functionality and reliability is equal to the probability that there exists at least one functioning minimal path. The cut set enumeration method estimates all mincuts that lead to network failure. To better understand the path-and-cuts method a small 2-terminal network reliability example is presented next (Bell, et al., 1997).

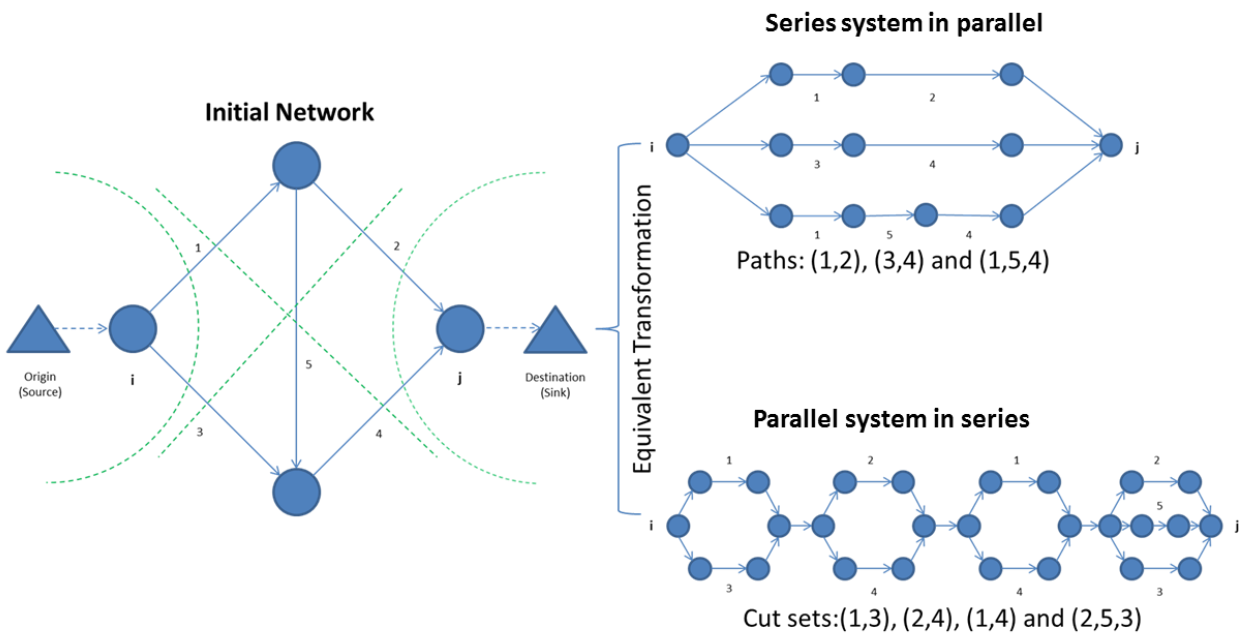


Figure 2 System Equivalent Transformations
Source: Bell and Iida (1997)

The initial network shown in Figure 2 can be transformed (Bell, et al., 1997) (Lucet, et al., 1999) to an equivalent series system in parallel or to a parallel system in series. *Minpath* enumeration will lead to the equivalent series system in parallel. If one of the (enumerated) *minpaths* (1, 2), (3, 4) or (1, 5, 4) is functional, then the whole system is functional as well. *Mincuts* enumeration will lead to the equivalent parallel system in *series*. If one of the four *mincuts* (1, 3), (3, 4), (1, 4) or (2, 5, 3) is not functional, then the whole system is not functional as well. In general , if the total number *minpaths* is *p* and the total *mincuts* number is *c,* we can express *minpaths* as *P(1),P(2),…,P(p)* and *mincuts* as *C(1), C(2),…,C(c)*.

Link α reliability can be defined as the expected value of some random binary state variable x_α (as in equation 1). Furthermore, since the structure function α_s for a specific path *s* corresponds to a series system it can be written as:

$$a_s(\mathbf{x}) = \prod_{i \in P(s)} \mathrm{x}_i \tag{9}$$

where: **x** is a link state vector. In the example network if links 1, and 2 function and 3, 4, 5 fail then the state vector will be: $\mathbf{x}^\mathbf{T}$ = [1 1 0 0 0].

From the equivalent system (*series in parallel*) of parallel *minpaths P(1),P(2),…,P(p),* the system structure function can be estimated by the following equation:

$$\varphi(\mathbf{x}) = 1 - \prod_{s=1\ to\ p} \left(1 - \ a_s(\mathbf{x})\right) = 1 - \prod_{s=1\ to\ p} \left(1 - \prod_{i \in P(s)} \mathrm{x}_i\right) \tag{10}$$

For the equivalent (*series system in parallel)* transformation in the example network the structure function of the initial system can be written as:

$$\varphi(\mathbf{x}) = 1 - (1 - \mathrm{x}_1\mathrm{x}_2)(1 - \mathrm{x}_3\mathrm{x}_4)(1 - \mathrm{x}_1\mathrm{x}_5\mathrm{x}_4)$$

In the same manner and while the structure function β_s of a *mincut C*(*s*) corresponds to an equivalent parallel (sub) system, it can be written as:

$$\beta_s(\mathbf{x}) = 1 - \prod_{i \in C(s)} (1 - \mathrm{x}_i) \tag{11}$$

Similarly, while the system consists of a series of *mincuts C*(1), *C*(2),…,*C*(*c*), the structure function of the system can be written as:

$$\varphi(\mathbf{x}) = \prod_{s=1\ to\ c} \beta_s(\mathbf{x}) = \prod_{s=1\ to\ c} \left(1 - \prod_{i \in C(s)} (1 - \mathrm{x}_i)\right) \tag{12}$$

For the equivalent (*parallel in series*) transformation in the example network, the structure function of the initial system can be written as:

$$\varphi(\mathbf{x}) = [1 - (1 - \mathrm{x}_1)(1 - \mathrm{x}_3)] \times [1 - (1 - \mathrm{x}_2)(1 - \mathrm{x}_4)] \times [1 - (1 - \mathrm{x}_1)(1 - \mathrm{x}_4)] \times$$
$$[1 - (1 - \mathrm{x}_2)(1 - \mathrm{x}_5)(1 - \mathrm{x}_3)]$$

Both equations 10 and 12 will provide the same result for the structure function $\varphi(\mathbf{x})$, i.e., zero or one. In a similar manner, system reliability *R* can be defined as the expected value of the system's structure function φ(**x**) via equation 2. Thus, equations 2 and 10, can lead to a *minpath* based expression of system reliability for a pair of nodes as:

$$R = E\{\varphi(\mathbf{x})\} = E\left\{1 - \prod\nolimits_{s=1\ to\ p}\left(1 - \prod\nolimits_{i \in P(s)} \mathrm{x}_i\right)\right\} \tag{13}$$

and equations 2 and 12, to an equivalent *mincut* based reliability expression for a pair of nodes:

$$R = E\{\varphi(\mathbf{x})\} = E\left\{\prod\nolimits_{s=1\ to\ c}\left(1 - \prod\nolimits_{i \in C(s)} (1 - \mathrm{x}_i)\right)\right\} \tag{14}$$

To transform the Boolean $\varphi(\mathbf{x})$ expressions into probability expressions, one can use the Poincaré's Theorem also known as *inclusion–exclusion* method. This method is presented next.

Inclusion–Exclusion Method

The *inclusion–exclusion method* is a fundamental tool for transforming Boolean into probabilistic expressions, when minimal paths or cut sets are known. It provides a path-based approach for estimating network reliability (Bell, et al., 1997) (Lucet, et al., 1999). After enumeration (e.g., *minpaths* or *mincuts* enumeration) is applied, a Boolean expression $\varphi(G)$ is obtained. The terms of φ are all the *minimal paths* (or *mincuts*) and each term is a product of Boolean variables (*state* variables) that are associated with each element of a specific path (as those of the example given in the path-and-cut method). This can be expressed mathematically as:

$$\varphi(G) = \sum_i P_i \tag{15}$$

where: P_i is the i-th *minimal path*'s Boolean expression such that:

$$P_i = \prod_k \mathrm{x}_{ik} \tag{16}$$

where: k is the number of elements (links) that constitute the minimal path i.

The reliability is then given as:

$$\mathrm{R}(G) = \mathrm{E}\{\varphi(G)\} \tag{17}$$

Poincaré's (*inclusion–exclusion*) formula for s *minpaths* is:

$$E\{\varphi(G)\} = \sum_{1 \le i \le s} E\{P_i\} - \sum_{1 \le i1 < i2 \le s} E\{P_{i_1} \cdot P_{i_2}\} + \cdots + (-1)^{s+1} E\{P_1 \cdot P_2 \ldots P_s\} \tag{18}$$

For example given a graph G that consists of two minimal paths P_1 and P_2 then:

$$\varphi(G) = P_1 + P_2 \text{ and } E\{\varphi(G)\} = E\{P_1 + P_2\} = E\{P_1\} + E\{P_2\} - E\{P_1 \cdot P_2\}$$

Note: In Boolean expressions (of $\varphi(G)$, etc.) the operators $\overline{x}$, "+", "·", stand for Boolean operations of "not", "or" and "and".

An equivalent formulation of the inclusion-exclusion formula that might be more convenient for road network reliability evaluation is also provided (Bell and Iida, 1997). This formulation is presented right next. If E_s is the event that all links in a path *P(s)* are functioning, then reliability can be represented by a probability of the union of such events E_s that belong to the path *P*:

$$\mathrm{R} = \Pr\left\{\bigcup_{s=1\ to\ p} E_s\right\} \tag{19}$$

The inclusion–exclusion formula is:

$$\begin{aligned} R = & \sum_{s=1 to\ p} \Pr\{E_s\} \\ & - \sum_{s=1\ to\ p} \sum_{all\ t\neq s} \Pr\left\{E_s \bigcap E_t\right\} + \sum_{s=1\ to\ p} \sum\nolimits_{t\neq s} \sum_{u\neq s,t} E_s \bigcap E_t \bigcap E_u \\ & + \cdots + (-1)^{p-1} \Pr\left\{\bigcap_{s=1-p} E_s\right\} \end{aligned} \tag{20}$$

The main drawback of the formula is that it contains many pairs of terms which cancel out. Readers interested in the topic may find algorithms that generate only the non-cancelling terms in the literature (Sun, et al., 2012).

Sum of Disjoint Products (Fratta–Montanari) Method

The *Fratta-Montanari* method (Bell, et al., 1997) (Lucet, et al., 1999) (Fratta, et al., 1973) converts Boolean expressions φ into probabilistic expressions as well. It differs from the previous method in that the Boolean expression $\varphi(G)$ is transformed so that one event will not include another event of the sum. i.e., all the product terms will be disjoint (Lucet, et al., 1999). This can be achieved with the following formula:

$$\varphi(G) = \sum_{i=1}^{s}(P_i) = P_1 + \sum_{i=2}^{s}[P_i \cdot \prod\nolimits_{j=1}^{i-1}(\overline{P_j})] \tag{21}$$

where:

$$P_i = \prod\nolimits_{j=1}^{p} x_{ij} \tag{22}$$

and

$$\overline{P_i} = x_{i1} + \sum_{j=2}^{p} [\overline{x_{ij}} \cdot \prod_{k=1}^{j-1} (\overline{x_{ik}})] \tag{23}$$

Reliability can then be estimated as:

$$R = E\{\varphi(G)\} = E\{P_1\} + \sum_{i=2}^{s} E\{[P_i \cdot \prod_{j=1}^{i-1} (\overline{P_j})]\} \tag{24}$$

Bell and Iida (1997), provide an equivalent formulation of equation 24, as well:

$$R = Pr\{ E_1 + \left[not\ E_1 \bigcap (E_2 \bigcup E_3 \bigcup \ldots)\right]\} \tag{25}$$

The first and second terms of equation 25, become exclusive events to each other and this transformation is repeated until the second term becomes an empty event. Both *Fratta-Montanari* and *inclusion-exclusion* methods are applicable only in the case that all relevant paths and cut sets are known.

There is a variety of methods to simplify the calculation of the sum of disjoint products with *Abraham* method being the most popular (Lucet, et al., 1999) (Rebaiaia, et al., 2013) (Abraham, 1973).

2.1.3.2 Heuristic Approaches for Connectivity Reliability Estimation

A variety of heuristic methods has been suggested in the literature to estimate connectivity reliability sufficient accuracy, and reduced computational time and capacity needs (Rebaiaia, et al., 2013) (Colbourn, et al., 1985). These can be classified into two main categories: *Bounding Methods* and *Monte Carlo Sampling Techniques* (Rebaiaia, et al., 2013). In the following sections we introduce the most prevalent of these methods in more detail.

Bounding Methods

All Paths and Cuts Method

The all-paths-and-cuts method uses the steps of the path-and-cut method (Bell, et al., 1997). When all paths are known (p = **p**) we can obtain the exact value of the system reliability:

$$R' = 1 - \prod_{s=1\ to\ \boldsymbol{p}} \left(1 - \prod_{\mathrm{i} \in \mathrm{P(s)}} \mathrm{x}_i\right) \tag{26}$$

Similarly, when all cut sets are known (c = **c**), the exact value of reliability is:

$$R'' = \prod_{s=1\ to\ \boldsymbol{c}} \left(1 - \prod_{\mathrm{i} \in \mathrm{C(s)}} (1 - r_i)\right) \quad (27)$$

If we substitute link state variable xi with link reliability r_i, we will obtain following upper *U* and lower *L* bounds equations of the true system reliability:

$$U = 1 - \prod_{s=1\ to\ \boldsymbol{p}} \left(1 - \prod_{\mathrm{i} \in \mathrm{P(s)}} r_i\right) \quad (28)$$

$$L = \prod_{s=1\ to\ \boldsymbol{c}} \left(1 - \prod_{\mathrm{i} \in \mathrm{C(s)}} (1 - r_i)\right) \quad (29)$$

Equations 28 and 29 correspond to the U and L curves of Figure 3, for increasing values of p and c, respectively. Due to the omission of Boolean algebra, they cannot converge to a single (true) value of reliability R. Instead they provide an approximating interval in which R lies. Equations 28 and 29, for p = p and c = c, will lead to the same result (R'=R''=R) of the true reliability value R. This can be verified from Figure 3, where curve R' (equation 26) is monotonically increasing as p increases and curve R'' (equation 27) is monotonically decreasing as *c* increases, while both curves converge to the point *R*, when *p* = ***p*** and *c* = ***c***.

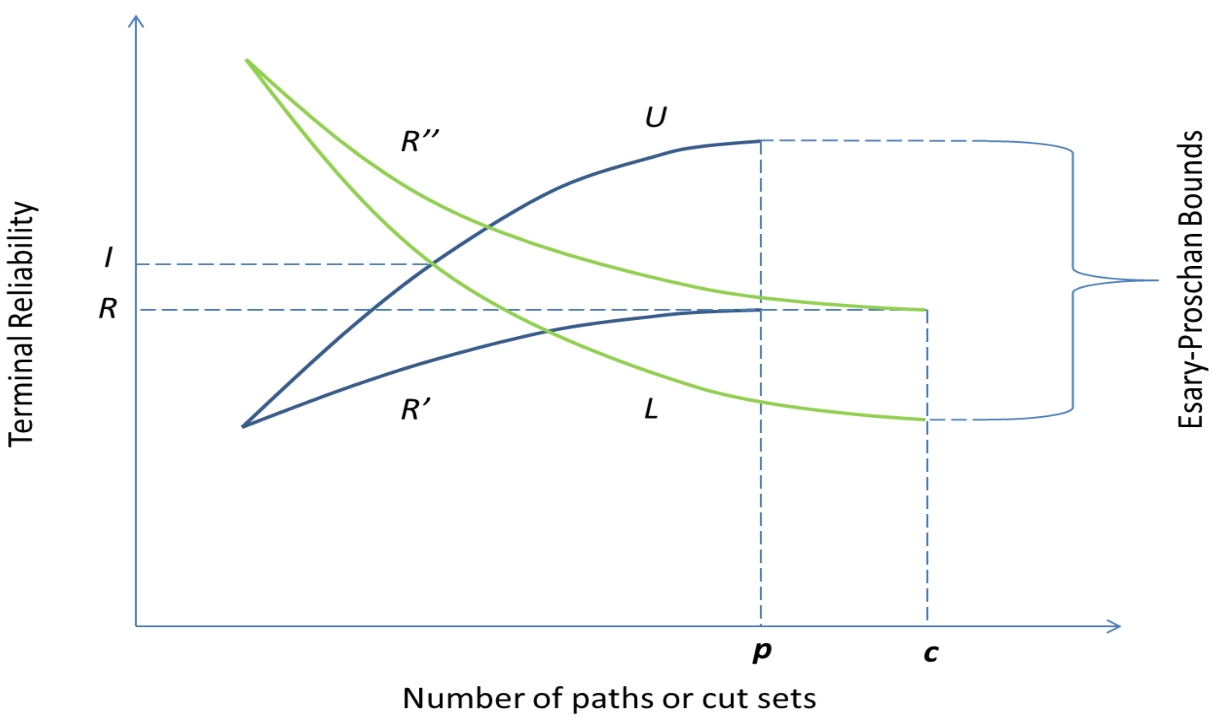


Figure 3 Esary-Proschan Upper and Lower Bounds and Intersection Point
Source: Bell and Iida (1997)

For p = $\boldsymbol{p}$ and c = $\boldsymbol{c}$**,** equations 28 and 29 provide the bounds, within which *R* lies, also known as *Esary-Proschan* bounds of system reliability *R* (Bell, et al., 1997) (Rebaiaia, et al., 2013) (Esary, et al., 1963). The Esary-Proschan bounding method, when enumerating all paths or all cuts, is considered *#P-complete*, thus not efficient computationally. Other bounding methods have been proposed in the literature and we refer to the study of Brecht and Colbourn (Brecht, et al., 1988), where the authors present (and improve) the *Kruskal-Katona* and the *edge-disjoint path* bounding methods for *2-terminal* network reliability, which according to them outperforms the Esary-Proschan bounding method.

Not All Paths and Cuts Method

The not-all-paths-and-cuts method (Bell, et al., 1997) (Wakabayashi, et al., 1992) also employs the same equations as the all-paths-and-cuts method and follows the same procedures as the previously described method, but in this case:

$$p \neq \boldsymbol{p} \text{ and } c \neq \boldsymbol{c} \tag{30}$$

The calculations are simplified, but lack precision. When Boolean algebra is not omitted, *R'* and *R''* values obtained, may be inaccurate and dependable on the number of paths *p* or cut sets *c*, while when Boolean algebra is omitted (Bell, et al., 1997) the values *U* and *L* cannot guarantee bounding of the real value of *R*.

Intersection Method

The intersection method (Bell, et al., 1997) approximates system reliability with the intersection point I, of curves U and L (as shown in Figure 3), which lies between the Esary-Proschan bounds. Assuming that the intersection point I, is a good approximation of the true value of system reliability R, the effort is then, to reduce the number of paths and cut sets (thus, the processing amount) required, in order to obtain this point. An efficient way to achieve that goal, in early stages, is to select the sequence of paths and cut sets “so that curves U and L are as steep as possible” (Bell, et al., 1997). Initially, the path that maximizes U (equation 28) is found, which is equivalent with finding the path s that maximizes the following function:

$$\prod_{i \in P(s)} r_i \tag{31}$$

Noting that, $\ln\left(\prod_{i \in P(s)} r_i\right) = \sum_{i \in P(s)} ln(r_i)$ and ln(r_i)<0, $(0 \leq r_i \leq 1)$, finding path *s* that maximizes *U* reduces to the shortest path problem (Bell, et al., 1997) with –ln(r_i) the length of link *i*. This also implies that maximizing the increase of curve *U* is equivalent with finding *n* shortest paths.

To find the *L*-curve minimum, a *dual graph* is developed (Bell, et al., 1997) (Ray, 2013) based on the initial graph (Figure 4). To obtain the dual graph *G'* from the initial graph *G* we have to draw a new node on each *face/region* of the initial graph. A *face* (or *region*) is characterized by a set of links that form its boundary and there are internal regions, like those defined by the links (1,3,5) and (2,4,5) as well as external regions, like those that the external graph boundaries define, i.e., (Origin,1,2,Destination) and (Origin,3,4,Destination). The new *dual nodes*, are then positioned in each face and connected (only) with those of adjacent faces (Ray, 2013). Each *dual* link's length is equal to the length of the (unique) initial link that it intersects (e.g., the length of dual link 1' is equal to the length of link 1). It should be always verified (as in Figure 4) that each dual link intersects one (and only one) initial link.

The *cut sets* of nodes *i* and *j,* in Figure 4, are equivalent to the paths of nodes *i'* and *j'* of the dual graph *G'*, and minimizing *L*-curve, is similar to finding the cut set *s* that maximizes:

$$\prod_{i \in C(s)} (1 - r_i) \qquad (32)$$

Thus, since $\ln(1-r_i)<0$ and if $-\ln(1-r_i)$ is considered as the *i*-th link's length (in the dual graph), then the procedure of estimating the path *s* that minimizes *L,* is equivalent to the estimation of the shortest path in the dual graph *G'*. Implicitly, maximizing the decrease of curve *L*, is equivalent with finding *n* shortest paths in the dual graph *G'*. We refer to Bell and Iida (Bell, et al., 1997), where an example with 16 links is presented and a very good approximation of reliability is reached, using the intersection method.

It should be noted that:

$$\ln\left(\prod_{i \in P(C)} (r_i)\right) = \sum_{i \in P(C)} \ln(r_i)\,, 0 \le r_i \le 1 \qquad (33a)$$

$$\ln\left(\prod_{i \in P(C)} (1 - r_i)\right) = \sum_{i \in P(C)} \ln(1 - r_i)\,, 0 \le 1 - r_i \le 1 \qquad (33b)$$

The intersection method has many advantages due to its efficiency to provide quickly a good network reliability approximation (Bell, et al., 1997). Furthermore, the paths or cut sets used correspond to realistic paths and cut sets from the transportation planning perspective (zigzag paths and distant detours are considered last). An alternative would be to use the *inclusion-exclusion* formula (as shown previously) and stop the procedure, when a convergence criterion is satisfied (before the method's "natural" ending iteration).

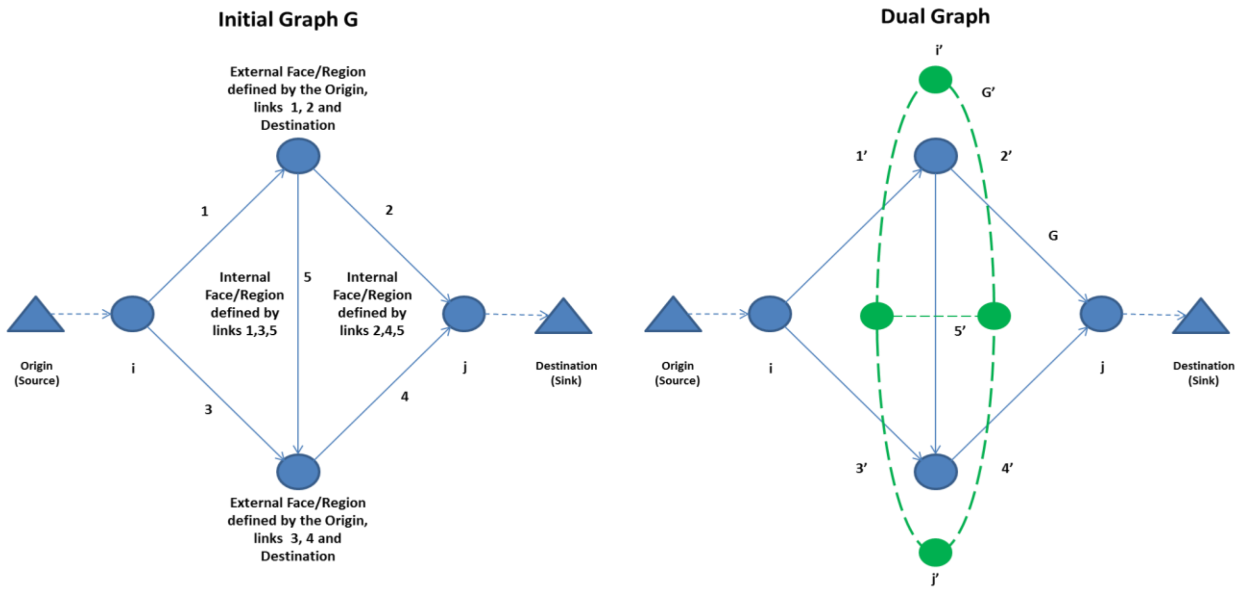


Figure 4 Dual Graph of an Example Network
Source: Bell and Iida (1997)

<u>*Monte Carlo Sampling Techniques*</u>

Monte Carlo techniques typically yield good network reliability estimations, but there is no guarantee of accuracy and/or convergence (Brecht, et al., 1988). The *standard Monte Carlo* technique could be briefly described as an algorithm that generates *n* independent copies/samples of the *stochastic graph* (i.e., with random link states), that produce equivalent system function values. The system reliability is then approximated as the average of those values:

$$R(G) = \frac{\sum_{j=1}^{n} \varphi\left(\mathbf{x}^{(j)}\right)}{n} \tag{34}$$

where: $\mathbf{x}^{(j)}$ is the *j*-th copy of the stochastic graph (or else a stochastic element state vector for all the graph elements).

This reliability estimator is considered unbiased. The law of large numbers indicates that this approximation will almost definitely converge to *R*(*G*) as $n \to \infty$. A confidence interval (with confidence 1–α centered at the half-width equal to $c_\alpha \cdot \sigma/\sqrt{n}$, where: c_α is the (1–α/2)-quantile of the standard normal distribution (i.e., mean=0, var =1) and σ is the standard deviation of $\varphi(\mathbf{x})$. The confidence interval (according to the *Central Limit Theorem*) is approximately valid when $n \cdot R(G)$ is a large number (Asmussen, et al., 2007) (Cancela, et al., 2010). Thus, the relevance of estimation is related to the number of samples and if that number is low, then the approximation may be incorrect, while if high, the computational cost of simulation may approach (or even exceed) that of exact methods (Rebaiaia, et al., 2013).

The structure function $\varphi(\mathbf{x})$, as used in the *standard Monte Carlo* simulation, is a *Bernouli random variable* (i.e., following the *Bernouli* distribution which is a special case of the *Binomial* distribution for n=1), and can be shown (Cancela, et al., 2010) (Gertsbakh, et al., 2015) that the *relative error* produced is:

$$\mathrm{r.e.}\{\mathrm{R}(G)\} = \frac{\sqrt{(1-\mathrm{R}(G))}}{\sqrt{n\cdot\mathrm{R}(G)}} \tag{35}$$

Equation 35, results in two basic drawbacks: i) necessity for high n values to decrease the relative error, and ii) very high relative errors when R values are small (i.e., highly sensitive to rare events (Cancela, et al., 2010). Variance reduction techniques (e.g., common random numbers, antithetic variates, importance sampling and stratified *sampling*) have been proposed in the literature to address these drawbacks (Asmussen, et al., 2007) (Cancela, et al., 2010) (Kumamoto, et al., 1977) (Rubino, et al., 2009).

2.1.4 Travel Time Reliability

Travel-time reliability can be considered a generalization of connectivity reliability where the probability of travel time exceeding a threshold value (representing disconnectedness) is calculated. This definition provides a measure of travel time stability (Bell, et al., 1997). While connectivity reliability was developed to study severe events, travel-time reliability was developed to study more frequent disruptions by less severe (supply or demand) variations that may occur on a daily basis (Watling, 2008).

If we consider a path *s* with *a* links and assume statistical independence of past link flow observations link travel time distribution (usually a normal) can be developed. The mean path travel time *T* (as the summation of normal distributions means) is normally distributed with a mean μ_α and a variance σ_α^2 shown in equation 36 (Bell, et al., 1997) (Iida, 1999):

$$T \sim N\left(\sum_{\alpha\in P(s)} \mu_\alpha\,, \sum_{\alpha\in P(s)} \sigma_\alpha^2\right) \tag{36}$$

By normalization we can define the probability that travel time along a path is less than some threshold value *t* (Iida, 1999):

$$Pr\{T \le t\} = \Phi\left(\left(t - \sum_{\alpha\in P(s)} \mu_\alpha\right) \Big/ \sqrt{\sum_{\alpha\in P(s)} \sigma_\alpha^2}\right) \tag{37}$$

where: *t* is a threshold travel time value.

Travel time reliability can then be determined for individual paths and path-based performance measures can be developed.

2.1.4.1 Travel Time Reliability Performance Measures
Travel time reliability performance measures may be grouped into three broad categories (Lomax et al., 2003): i) statistical range, ii) buffer time, and iii) tardy trip indicators. Statistical range measures typically use standard deviation statistics to form representative estimates of traffic conditions (in terms of travel time). They are typically presented with an average value plus or minus a deviation value. Buffer time measures indicate the amount of additional time needed to allow on-time arrival at a destination for the majority of trips. These measures may represent average trip times or additional time to average trip times to select a departure time that ensures on-time arrival to a destination with a specific confidence level. Tardy trip indicators provide a measure of unaccepted lateness (i.e., frequency of late arrivals) where a threshold value is used to identify acceptable late arrivals. Numerous studies exist on travel-time reliability performance measures (Lomax, et al., 2003) (Rakha, et al., 2006) (Lyman, et al., 2008) (Pu, 2011).

2.1.4.2 Path based Travel Time Reliability
The uncertainties in transportation system such as congestion lead to freight operators facing uncertainties in goods delivery. Travel time reliability is becoming increasing critical to businesses, especially the manufacturing sector as many manufacturers are positioning to adopt “just-in-time” manufacturing processes and other schedule dependent inventory, assembly and distribution logistics (Cambridge Systematics, 2012).

Once segment roadway travel times have been estimated, the next step is to estimate path or trip travel times. Path travel-time reliability is estimated as the probability that the travel time between an origin-destination pair is within a specified range. The key parameters in estimating path travel-time reliability include estimating the path mean travel time and path travel-time variance.

Two approaches have been used in past studies to define reliability for valuation studies: Mean-variance and Schedule Delay. The former approach uses statistical measures to separate out the value of typical/usual travel time (mean or a measure of central tendency) and measures for the dispersion of the travel time distribution, such as the standard deviation whereas the latter approach focuses on the magnitude of the time during early and late arrivals in relation to a pre-determined schedule.

Mean-variance approach is easy to implement in existing analysis frameworks. However, there is concern that the mean value may include a portion of the reliability component, leading to double counting of benefits when analyzing an improvement. Several researchers have indicated their preference for the schedule delay approach on conceptual grounds, but it is difficult to implement for the highway mode where travelers schedules are not known and would vary widely if they were.

In schedule delay approach travelers define their own schedule and adjust their departure times, routes, and modes accordingly. In the scheduling delay approach, early arrivals can be valued differently than late arrivals. Reliability and scheduling are related concepts. The former refers to the disutility of the inconvenience and possible penalties attributed to the unreliability of travel times. The latter refers to the disutility of arriving either too early or too late, when the traveler has time restrictions in terms of flexibility of schedules.

One of the initial studies (Small, 1982) established that scheduling costs play a major role in choice of departure times by defining a variable to measure how early or late the commuter is with respect to the official work start time. Let tw be the official work start time. If a commuter leaves home at time th and the travel time is T, then commuter arrives early if $t_h + T < t_w$. Two components of Schedule delay concept are Schedule Delay Early (SDE), defined as $t_w - (t_h + T)$ and Schedule delay late (SDL) is $(t_h + T) - t_w$. The scheduling cost function is as follows:

$$C_s = \alpha T + \beta(SDE) + \gamma(SDL) + \theta D_L \quad (38)$$

where, α is the cost of travel time β and γ are the costs/min of arriving early and late respectively and θ is an additional discrete lateness penalty. D_L is 1 when SDL>0 and 0 otherwise.

The scheduling cost function (Noland, et al., 1995) to allow for decomposition of morning commute which are the expected cost of schedule delay, lateness and travel time. The modified model is:

$$EC_s = \alpha E(T) + \beta E(SDE) + \gamma E(SDL) + \theta P_L \quad (39)$$

where, PL = E(DL) is the lateness probability.

A recent study (Lyman, et al., 2008) used the standard travel time reliability measures for corridor analysis: 95th percentile TT, Travel Time Index, Buffer Index, Planning time index (PTI), congestion frequency. The study corridor was I-5N, 23.5 miles in length, a freeway in Portland, Oregon. The analysis was carried out using PORTAL's monthly report system which is a collection of all measured corridor travel times, extracted at 5 minute intervals for all of 2005.

The procedure for estimating path travel-time reliability assumes that travel times follow a normal distribution and requires a measure of trip travel-time variance. Past study by

(Rakha, et al., 2006), shows that the assumption of normality is, from a theoretical standpoint, inconsistent with field travel-time observations and that a lognormal distribution is more representative of roadway travel times through goodness-of-fit tests that. However, visual inspection of the data demonstrates that the normality assumption may be sufficient from a practical standpoint due to its computational simplicity. This study also proposes five methods for the estimation of path travel-time variance from its component link travel-time variances as shown in Table 1.

Table 1 Path Travel Time Variability

Method	Equation
1	$\hat{\sigma}_{1t}^2 = \frac{\sum_{i \in V(s)} \left(\sum_{j \in L(s)} t_{ij}^2 \right)}{n} - \sum_{j \in L(s)} \overline{t}_j^2$
2	$\hat{\sigma}_{2t}^2 = \frac{\overline{t}_t^2}{\sum_{j \in L(s)} \overline{t}_j^2} \cdot \left\{ \frac{\sum_{i \in V(s)} \left(\sum_{j \in L(s)} t_{ij}^2 \right)}{n} - \sum_{j \in L(s)} \overline{t}_j^2 \right\}$
3	$\hat{\sigma}_{3t}^2 = \frac{\overline{t}_t^2}{m^2} \left(\sum_{j \in L(s)} \frac{\sigma_j}{\overline{t}_j} \right)^2$
4	$\hat{\sigma}_{4t}^2 = \left\{ \overline{t}_t \cdot \mathrm{Med}_j \left(CV_j \right) \right\}^2$
5	$\hat{\sigma}_{5t}^2 = \frac{\overline{t}_t^2}{4} \left(CV_{\max} - CV_{\min} \right)$

Source: Rakha et al. (2006)

The mean-variance approach allows the estimation of two widely used reliability metrics: value of travel time reliability (VTTR) and Reliability Ratio (RR). VTTR represents the user's monetary weight for improving reliability and RR is defined as ratio of VTTR to VOT. An established RR along with knowledge of the VOT simplifies the task of VTTR estimation. In recent studies (Asensio, et al., 2008) (Brownstone, et al., 2005) (Tilahun, et al., 2010) VTTR and RR are determined to capture the travel time reliability.

A method for synthesizing a distribution of consistent path-dependent O-D travel times from the known distribution of link counts is suggested in the SHRP 2 C04 report (Vovsha, et al., 2013). This method generates of origin– destination (O-D) travel time distribution for the base year, which is needed for calculating travel time reliability measures. These reliability measures are used in travel demand models to explain travel choices along with the average travel time and cost. The method is designed to produce a distribution of travel times for a full regional O-D matrix for a certain time of day, period or hour.

SHRP L35B report (Sadabadi, et al., 2014) uses an instantaneous travel time aggregation method to estimate path travel times based on link travel times. Travel time data used as input in this study are provided by INRIX. In this study, data archived during calendar year 2011 are used at 1-minute resolution on all segments considered. The study shows that as trip length becomes longer, the risk impact of any newly added segment, while still positive, becomes marginal compared with the rest of the path. This phenomenon is reflected by the concavity of reliability ratio (RR) curves. Other elaborate path travel time estimation methods (e.g., trajectory construction based models), will result in more accurate travel time estimates for long distance trips.

SHRP 2 L35B study uses "Real Options Theory" which was first used by the SHRP 2 L11 project (2013) for determining the value Background of travel time reliability by using speed and volume data as input (Sadabadi, et al., 2014). The options-theoretic approach introduced by the SHRP 2 L11 uses an analogy where premiums are set for an insurance policy on guaranteed speed levels. Specifically, the method calculates the dollar value of reliability by multiplying the certainty-equivalent penalty (measured in minutes/mile and obtained by applying the closed form Black-Scholes equation) by the value of time, thus it requires an estimation or adoption of VOTT as input. The SHRP 2 L11 study takes into account heterogeneity of the road users and different trip purposes by applying a separate value of time that corresponds to each user group.

Freight travel time reliability

Past studies on the valuation of freight travel time reliability are limited compared to passenger travel. Most of the studies (Bergkvist, et al., 2001) (Bolis, et al., 1999) (Danielis, et al., 2005) (Wigan, et al., 2000) indicate that the freight value of reliability varies by commodity, with bulk commodities having the lowest value. However, there is little consensus on what the values of VORs or Reliability Ratios should be. If the Reliability Ratios for freight are equivalent to passenger travel, i.e., around 1.0, then VOR for freight will be higher (Cambridge Systematics, 2012).

Path based dynamic travel time

In most of the past studies, it is generally assumed that path travel time is the aggregation of the travel times on its consisting links. However, for a probe-based data collection system in which the number of reports is rather limited, this link-based estimation/prediction might not be reliable. (Chen, et al., 2001), evaluate the performance of dynamic travel time prediction models with real-time data (travel time) collected by probe vehicles on path and its consisting link. In this study "Kalman filtering method" is chosen because it enables the prediction of the state variable (travel time) to be continually updated as new observation becomes available. This approach has been used in the forecasting of traffic volume and real-time demand diversion as well as the estimation of trip-distribution and traffic density. In this study, this technique is used to perform travel time prediction based on real-time information provided by probe vehicles. Specifically, the average travel time of probe vehicles at each time period is used as the real-time observation to predict the travel time in the next (or future) time period.

Preliminary Methodology (Chen, et al., 2001)

Step 1: Collect Link Travel time data at discretized time step (We have that for each link)

Step 2: Define regional O-D

Step 3: Create/build k-shortest paths (with predefined impedance function)

Step 4: Obtain path travel time by aggregating the TT of the links

Step 5: Obtain path travel time reliability measures

Table 2 Path Based Reliability Summarized Literature Review

Literature	Reliability Measure	Notes
Bergkvist and Westin (1996)*	VTTR	Data was collected through computer based SP survey.
Bolis and Maggi *	VTTR (dependent on type of operation such as Just In Time production)	Based on the Leeds Adaptive SP (LASP) survey which provides choice of alternative ways for the freight operators.
Danielis R. et al. *	VTTR	Also determined VOT.
Wigan et al. *	VTTR (dependent on segment type)	Data was collected on three market segments: Inter-capital FTL, Metropolitan FTL and metropolitan multidrop.
Brownstone and Small (2003)	90^{th} -50^{th} percentile and Reliability Ratio= VTTR/VOT	
Asensio and Matas (2008)	Standard deviation using Scheduling approach and Reliability Ratio	
Tilahun and Levinson (2007)	Difference between actual late arrival and usual travel time and Reliability Ratio	

SHRP S2-C04-RW-1	Standard Deviation per unit distance	Used a distribution of path dependent Origin– Destination travel times
SHRP S2-L35B-RW-1	95% TT, TTI, PTI, BI, VTTR, RR	Uses Real Options Theory to improve travel time reliability
Lyman and Bertini (2008)	95% TT, TTI, PTI, BI,	Also provides segment, corridor and network analysis
Rakha et al. (2006)	Path Travel time variability	

* For Freight

2.1.5 Capacity Reliability

Capacity reliability captures user behavior, is a generalization of connectivity reliability, and conceptually, lies between travel-time and behavioral reliability (Watling, 2008). Assuming a link is reliable when it provides a smooth flow, connectivity reliability can be defined as: "*the probability that at least one path without congestion exists"* (Iida, 1999). Connectivity methods as described herein can be used to estimate reliability under capacity degradation (ranging from disasters to less severe events like traffic incidents) but fail to capture user behavior. A number of studies have been conducted to approach reliability, incorporating both capacity changes and user route choices impacts. Next, a brief introduction of such approaches is presented.

2.1.5.1 Degradable Transportation Systems (DTS)

DTS can be defined as road networks where link capacities, hence system state and system performance, can be degraded (various capacity degradation levels) by a variety of events. Transportation degradation was introduced by (Asakura, 1997), who further extended travel time reliability concept for road networks, to consider capacity degradation as a consequence of road network deterioration. Nicholson and Du (Nicholson, et al., 1997) proposed a design and analysis framework for *DTS*. They developed an integrated steady-state equilibrium model for predicting macroscopic traffic behavior. The model assumed that network users choose paths that minimize their generalized travel cost and that the traffic levels between each *O-D* pair are directly related to demand and supply interactions (elastic traffic demand). They use system surplus as a measure to assess the socio-economic impacts of system degradation. The probability, that the flow reduction, due to capacity degradation, is less than a specified threshold value, constitutes the system reliability. They also propose an exact solution algorithm assuming that the state vector space **X** is discrete. Then the reliability of the *k*-th *O-D* sub system can be estimated as:

$$\mathrm{R}_k(\theta_k) = \sum_{s=0}^{w} p_s z_k(\theta_k, \mathbf{x}_s), \qquad k \in \boldsymbol{K} \tag{40}$$

where: $z(\theta,\mathbf{x})$ is a structure function and θ is the maximum acceptable flow decrement rate for the k-th *O-D* sub-system ($0 < \theta < 1$), while $\mathbf{x}$ is a component state vector, comprising the arc capacities, which belongs to a (discrete) component state vector space $\mathbf{X}$ and p_s is the (known) probability of the s-th component capacity degradation and W is the number of component state vectors, such that:

$$z_k\,(\theta_k, \mathbf{x_s}) \;=\; \begin{cases} 1 \; if \; y_k(\mathbf{x}) \le \theta_k \\ 0 \; if \; y_k(\mathbf{x}) > \theta_k \end{cases} \quad k \in \boldsymbol{K} \tag{41}$$

where: y_k is the decrement rate ($0 \le y_k \le 1$) of traffic flow f :

$$y_k\,(\mathbf{x}) \;=\; \frac{f_k\,(\mathbf{x_0}) - \; f_k\,(\mathbf{x}))}{f_k\,(\mathbf{x_0})} \quad k \in \boldsymbol{K} \tag{42}$$

where: $\mathbf{x_0}$ is the non-degraded component state vector.

The system total traffic flow F is:

$$\mathrm{F}(\mathbf{x}) = \sum_{k \in \boldsymbol{K}} f_k\,(\mathbf{x}) \tag{43}$$

and the system decrement rate is:

$$y\,(\mathbf{x}) \;=\; \frac{\mathrm{F}\,(\mathbf{x_0}) - \; \mathrm{F}\,(\mathbf{x}))}{\mathrm{F}\,(\mathbf{x_0})} \quad k \in \boldsymbol{K} \tag{44}$$

Similarly the system structure function will be:

$$z\,(\theta, \mathbf{x}) \;=\; \begin{cases} 1 \; if \; y(\mathbf{x}) \le \theta \\ 0 \; if \; y(\mathbf{x}) > \theta \end{cases} \quad k \in \boldsymbol{K} \tag{45}$$

where:

$$\theta \;=\; \sum_{k \in \boldsymbol{K}} v_k(\mathbf{x_0})\theta_k \quad \text{and } v_k(\mathbf{x_0}) = \frac{f_k(\mathbf{x_0})}{\mathrm{F}(\mathbf{x_0})} \;, \qquad k \in \boldsymbol{K} \tag{46}$$

and the system reliability will be:

$$\mathrm{R}(\Theta) = \sum_{s=0}^{w} p_s z(\theta, \mathbf{x}_s), \tag{47}$$

To estimate these reliabilities (k-th and system) all possible component state vectors $\mathbf{x_s}$ and probabilities p_s must be enumerated. Then traffic flows $f_k(\mathbf{x_s})$ and $F(\mathbf{x_s})$ must be estimated by solving an integrated equilibrium model for each $\mathbf{x_s}$, and then the decrement rates and structure

2.1.5.2 Network Reserve Capacity Approach

According to (Chen, et al., 1999), the method proposed by (Nicholson, et al., 1997), is not suitable to assess the capacity reliability of the network. In their study they introduce the concept of *network reserve capacity* μ for the estimation of a maximum flow capacity that is consistent with users' route choice behavior as defined by Wardrop. Link reliability is defined as a random variable that can take continuous or discrete values between zero and one, which allows modeling for modest events without restricting to severe events that disconnect segments of the network. A side product of this capacity-based reliability is the estimation of travel time reliability when solving the user equilibrium problem in order to obtain the maximum network reserve capacity. The authors modeled this reliability problem as a bi-level problem with maximum capacity problem as the upper level (equations 48, 49) and the *User Equilibrium* (*UE*) on the lower (equations 50-53).

$$\max \mu = \mathrm{g}(c_1, c_2, \dots c_\alpha) \tag{48}$$

$$s.t.$$

$$\upsilon_\alpha(\mu \boldsymbol{q}) \leq c_\alpha, \forall a \in A \tag{49}$$

where: μ (reserve capacity) is an output parameter which can be calculated by the capacities c_α of the links and serves as a maximum *O-D* matrix multiplier, $\upsilon_\alpha(\mu \boldsymbol{q})$ is the equilibrium flow on link α, with demands of all *O-D* pairs being uniformly scaled by μ times the base *O-D* demands $\boldsymbol{q}$, ($\boldsymbol{q}$ is the existing *O-D* demand matrix in vector form). The matrix multiplier μ is obtained by finding the equilibrium flow $\upsilon_\alpha(\mu \boldsymbol{q})$*, by* solving the following *UE* problem:

$$\min Z = \sum_{\alpha \in A} \int_0^{\upsilon_\alpha} t_\alpha(x) dx \tag{50}$$

$$s.t.$$

$$\sum_{\alpha \in A} f_p = \mu q_w \ \forall \ w \in W, \tag{51}$$

$$\upsilon_\alpha = \sum_{\alpha \in A} f_p \, \delta_{\alpha p}, \forall \ \alpha \in A, \tag{52}$$

$$f_p \geq 0, \forall \ p \in P \tag{53}$$

where:

P: set of routes in the network

P_w: set of routes between *O-D* pair w∈W

μ: *O-D* matrix multiplier for the whole network (upper level decision variable)

Z: user-equilibrium objective function

υ_α: flow on arc α∈A

$t_\alpha(\upsilon_\alpha)$: travel time on arc α∈A

q_w: existing demand between *O-D* pair w∈W

$\boldsymbol{q}$: existing *O-D* demand matrix in vector form

f_p: flow on route *p*∈*P*

$\delta_{\alpha p}$: 1 if route *p* uses arc α, 0 otherwise

The lower level takes into account route choice behaviors and congestion impacts, while the upper level determines the maximum *O-D* matrix multiplier *μ*, subject to the capacity constraints. As the scaled demand approaches the network capacity, the equilibrium constraints will have a substantial effect on the distribution of flow as well as on the network reserve capacity.

Reliability for this formulation can be defined as: *"The probability that the network reserve capacity is greater than or equal to the required demand for different levels of capacity degradation*" or $R(\mu_p) = \Pr(\mu \geq \mu_p)$, where: μ_p is a required demand level. When link capacity takes only binary values of zero or one, it reduces to the connectivity reliability measure (as a special case of capacity reliability).

2.1.5.3 Network Robustness Index (NRI)

NRI is a link-based capacity disruption approach (Scott, et al., 2006) (Sullivan, et al., 2010) to quantify how individual links affect the total network travel time and the network robustness, when they are isolated or degraded (at various capacity-disruption levels). If a link's degradation results in a significant increase in the total network travel time (when compared to the removal of other links), then this link is considered critical. In this way, one can rank all the links of a network. Furthermore, with the estimation of the network robustness index (estimated in the second step of the process) the network's reliability over link disruptions can be evaluated. A single link's NRI_α is the total travel-time change over a given time interval as a consequence of traffic being re-routed through other links when a specific link is removed.

A base case network, at equilibrium, is used as a measure of comparison and total travel-time cost T*C* is calculated as the sum of the products of link travel times t_i (in min/trip) multiplied by the respective link flows v_i (all links are fully functional and at user equilibrium):

$$TC = \sum_{i \in I} t_i \, v_i \tag{54}$$

where: The travel-time factor, $t_i v_i$, is the total minutes of travel per time interval on link *i* that belongs to a set of network links *I.*

The *system-wide travel-time cost* TC_a is then estimated for the case when an individual link *a,* is removed or degraded and all traffic on the network is re-assigned.

$$TC_\alpha = \sum_{i \in I/a} t_i^{(\alpha)} \, v_i^{(\alpha)} \tag{55}$$

The *NRI* of link *α* can be calculated as the difference between TC_α and TC.

$$NRI_a = TC_a - TC \tag{56}$$

2.1.6 Behavioral Reliability

Connectivity and travel time reliability adopt a system operator perspective (Watling, 2008) where user satisfaction is implicitly measured by the system performance. Behavioral reliability methods try to incorporate a user perspective. The inclusion of user perception produces a measure of how (un)reliability primarily affects the average user behavior rather than a measure of the system reliability performance. Thus, behavioral reliability could only be a complementary approach to travel time reliability. The reader may refer to a wide range of sources on behavioral reliability methods, like those that use:

i) *fuzzy logic* to capture uncertainty in the travelers perception of reliability of given paths (Chen, et al., 2001), and ii) minimal modifications of conventional methods to incorporate user behavior with a utility that is a linear combination of travel time mean and variance (Van Berkum, et al., 1999) or of mean and standard deviation (Lo, et al., 2006).

2.1.7 Other Reliability Approaches

In this section we briefly describe approaches that cannot be grouped under any of the previous categories or present some interesting/unique modeling attributes that allows separate classification and description. These are *the Game Theoretical, Absorbing Markov Chains* and *Microsimulation* approaches, briefly presented next.

2.1.7.1 Game theoretical Approaches

In a 2000, Bell (Bell, 2000), introduced a game theoretical approach for measuring a transport network's reliability performance. He defined a network as reliable '*if the expected trip costs are acceptable even when users are extremely pessimistic about the state of the network'*. At that time all methods proposed assumed knowledge of link performance frequency distributions (usually delay, travel time or capacity distributions), an information that in most cases is unavailable. Bell, adopted a two-player, non-cooperative, zero-sum game between the network user (aiming to find a path with minimum expected trip cost) and an "*evil entity" or network "spoiler"* that can impose link costs to the user aiming to maximize the expected trip cost. The author claimed that "*the mixed strategy Nash equilibrium for this game offers a useful measure of network reliability, since it yields the expected trip cost when the user is extremely pessimistic about the state of the network*". The game was reformulated into two equivalent linear programs, *dual* to each other, using the *maximin* approach, with path choice probabilities as the primal and link-based scenario probabilities as the dual variables. To avoid path enumeration the author suggested substituting link choice probabilities for path choice probabilities as the primal variables, and imposing node probability conservation constraints on the primal variables. A simple iterative solution scheme based on the simple *Method of Successive Averages* (*MSA*) was proposed.

2.1.7.2 Absorbing Markov Chains Approach

A *Markov chain* is characterized by a transition matrix with elements t_{ij} that define the probability of an entity moving from state *i* to state *j*. In the context of reliability for road networks, *states* represent links/vertices of a network graph and *entities* represent travelers. In the study of Bell and Schmöcker (Bell, et al., 2002), the *states* represent the intermediate links of a graph, the origins (that can be multiple), one destination and one *bin* where trips that encounter link degradation are collected. Estimation of such a transition matrix can be done with *All-or-Nothing* (*AON*), *Stochastic User Equilibrium* (*SUE*) as well as *UE* assignment and the rows of the matrix must sum up to unity (for conservation reasons). The assignment selection (*AON*, *SUE* or *UE*) also defines the type of the transition probability which can be binary (0-1) for *AON* or real (between zero and one) for *SUE* or *UE*. It should be noted that transition probabilities are destination-specific, thus if multiple destinations, have to be considered, they should be encountered one per time. The destination and bin elements are called *absorbing* i.e., the traveler cannot leave them, once entered. Vertex failure probabilities can be used to determine the trip failure probabilities (Bell, et al., 2002). The transition probabilities were calculated according to the least cost paths. The link cost was calculated by the following formula:

$$c_{ij} = d_{ij} - \beta \ln(r_i) \tag{57}$$

where: d_{ij} is the cost for ignoring unreliability, β is the risk averseness factor and r_i is the reliability of node i. If β=0, the link reliability is not considered and if β is large and the user can have information on the reliability r_i, then the reliability is a significant factor in user's cost estimation. The study has shown that the encountered reliability increases as β increases.

2.1.7.3 Microsimulation Approach

The majority of models described previously require large computing capacity and time. Equilibrium based models may be appropriate when assessing the effect of long term degradations (natural disasters), but most degradations are short-term and it is unlikely that an equilibrium will be achieved in such cases. For both short term degradations and rapid demand increases, equilibrium based models will most probably lead to erroneous assumptions (Nicholson, 2003). For such events studies have shown (Berdica, et al., 2003) that microsimulation models are more sensitive.

2.2 Truck Parking Demand Analysis using GPS Data

The freight transportation system in USA makes one of the most valuable contributions to the nation's economy and progress. In this system, truck traffic mode makes most contributions and it is expected to increase by 45% by 2040 (FHWA, Freight Facts and Figures 2013). Long term economic growth shall result in even greater demand for truck traffic transportation mode. Even after such great demand, there is a huge lack of truck parking in many states (Dowling et al., 2014). Also, truck traffic does not get access to all roadways and cities and counties regulate truck traffic by restricting parking, prohibit from certain roads and designate specific routes which leads the truck drivers to search for parking areas for rest and if not available, they tend to park in areas not designated for parking such as ramps and spillover parking which signifies a safety concern for the other forms of traffic. Moreover, truck parking has been indicated as the most influential factor for route selection decisions (Dowling et al., 2014). Lack of truck parking is also indicated as a characteristic of an unreliable route as the truck drivers do not get the required amount of rest or sleep and this may lead to safety concerns during travel for the truck driver as well as other modes.

The review is first categorized by the different states which has concerns regarding truck parking and then on other studies. From the following review, it can be understood that roughly 42% of the literature mentioned about using truck GPS data for evaluating parking demand, about 33% mentioned different methodologies like surveys, to collect truck data and additional data and about more or less 20% mentioned about technological usage of implementing safe and easier truck parking.

2.2.1 Truck Parking in Wisconsin

A study of truck parking issues was conducted along the major state highways in Wisconsin (Adams, et al., 2009). This study determined specific locations in Wisconsin with parking issues and prioritized them based on specific criteria. The methodology that was used included development of a GIS-supported online survey tool to collect information on truck parking issues. Data was collected from 3 groups of participants: truckers/carriers, highway patrol officers and public freight planners. The data collected through the survey was exported to ArcGIS for spatial analysis. Other data that were used include shape files of Interstate and state highways networks and related attribute information, taken from the National Transportation Atlas Database (NTAD) 2006 and the Freight Analysis Framework (FAF). One of the outputs from these analyses was determination of priority of interstate and state corridors and cities in the region suffering from truck parking issues. Figure 5 shows that facilities with different priority levels suffering from capacity issues. The locations of parking facilities were clustered using the Nearest Neighbor Hierarchical (NNH) clustering algorithm in the software tool CrimeStat 3.1.

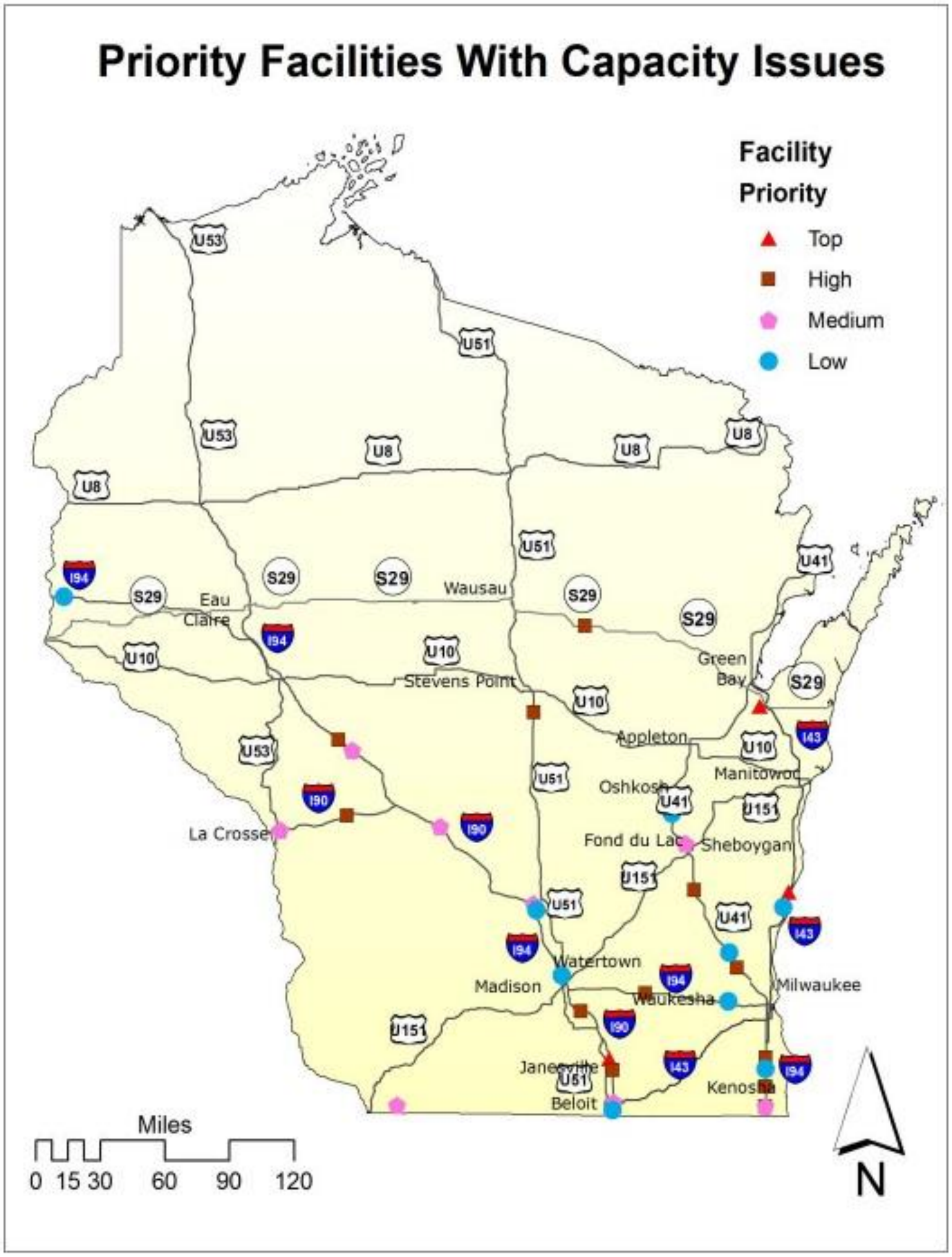


Figure 5 Priority Facilities with Capacity Issues
Source: Adams et al. (2009)

This study found many truck parking problems and the most common parking problem found is related to capacity and ramp parking. There are not enough parking spaces to meet the peak demand during popular hours of use and the overflow trucks park at the ramps. Moreover, parking capacity shortages occur in the early evening or late at night.

2.2.2 Truck Parking in Washington

Washington State Department of Transportation (WSDOT) performed a study to determine if there is shortage of truck parking at public rest areas (PRAs) and commercial truck stops (CTSs) and identify strategies to increase the amount of truck parking in future (Parametrix, 2005). The study area was I-5, I-90 and I-82. It was found that PRAs are over capacity by 8% and CTSs are underutilized by 13%. Truck parking data was collected by telephone survey for the PRAs and CTSs at the corridor, segment and facility levels. Truck parking data were also collected at other locations along the study corridors, such as weigh stations, on- and off-ramps, shoulders, and chain-up areas and these were collectively referred to as illegal truck parking in this study. Interestingly, in this study, the highest truck parking demand for both CTSs and PRAs occurred between 6 PM and 6 AM and this was defined as the peak period. Existing truck parking demand was calculated by recording the number of trucks parked at these locations in terms of volume and location. Future truck parking demand was estimated for 2030 by multiplying the existing demand by a growth factor that was developed for the study corridors. The growth factors were estimated based on:

Washington State annual truck growth rates observed in WSDOT historical traffic volume data

The Strategic Freight Transportation Analysis (2003) and Eastern Washington Intermodal Transportation Study (1993) truck volume databases.

WSDOT's Weigh-In-Motion recorders for truck traffic volumes.

Freight forecast estimates for the Port of Seattle and Port of Tacoma.

The Federal Highway Administration (FHWA) Study of Adequacy of Commercial Truck Parking Facilities (FHWA June 2002).

The draft Freight Report for the 2005 Washington Transportation Plan Update (WSDOT 2005)

The study recommended several improvement strategies and options to increase the truck parking capacity at PRAs and CTSs.

2.2.3 Truck Parking in Virginia

(Garber, et al., 2002) developed a methodology to determine the supply and demand for heavy truck parking. I-81 was used as the study area. Supply was defined as the number of parking spaces available for large truck parking and demand was defined as the sum of the parking accumulation and illegal parking at a given time. Parking duration and accumulation data was obtained for different times along with location, number and types of parking spaces, and availability of other facilities of each truck stop and rest area. Figure 6 and 7 show the accumulation and duration of trucks at different time of day.

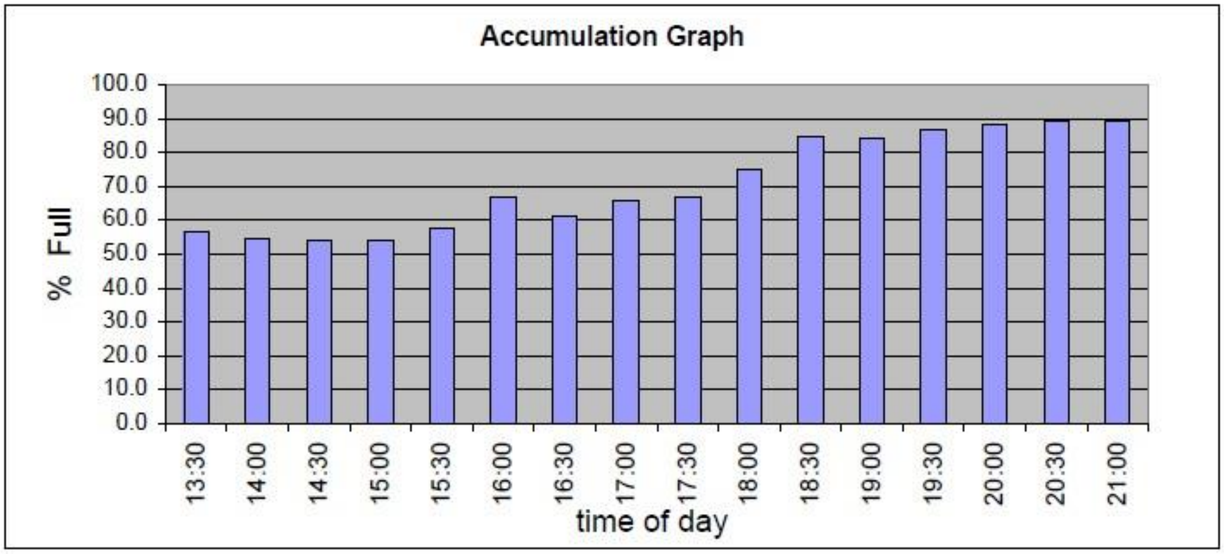


Figure 6 Accumulation vs Time of Day
Source: Garber et al. (2002)

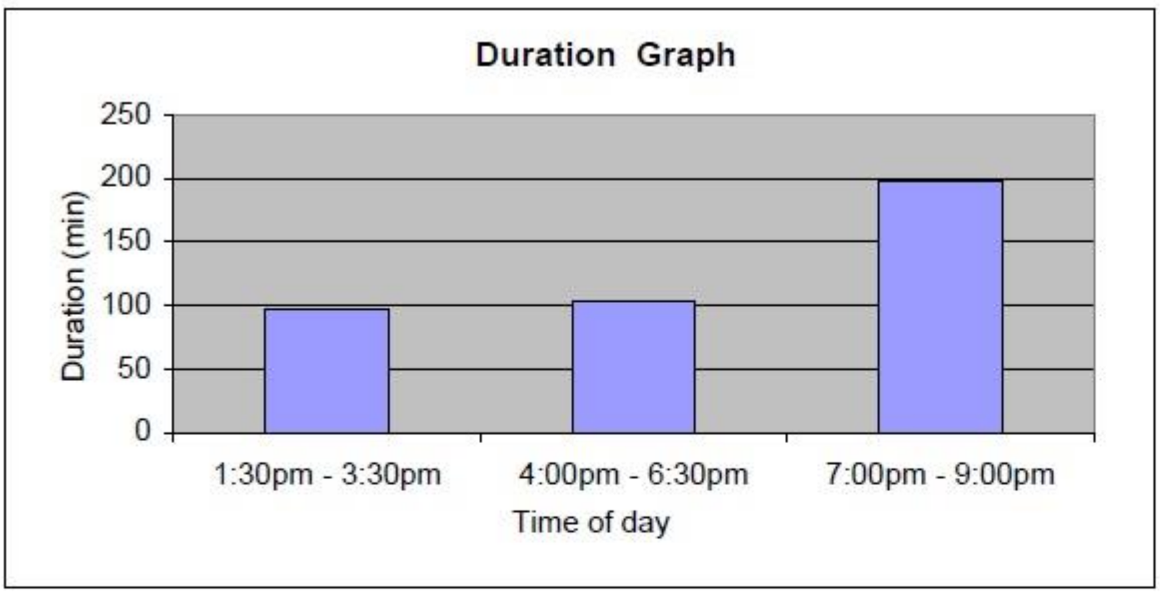


Figure 7 Average Duration vs Time of Day
Source: Garber et al. (2002)

Survey data was also collected from the truck drivers. Using the dataset, stepwise regression analysis was used to develop demand models. The variables used for the model were:

TotalTruck: Total numbers of trucks on I-81 near a truck stop in half hour intervals.

PercentTruck: Percentage of trucks in the traffic stream in half-hour intervals.

Duration: Duration at a truck stop in half-hour intervals.

Dist_81: Distance from a truck stop to I-81.

Dist_TS: Distance from a truck stop to the nearest other truck stop.

Dist_RA: Distance from a truck stop to nearest rest area.

SERVICE: Dummy variable for measuring the difference of services between large and small truck stops. (Number of spaces > 60, SERVICE = 1.)

The estimated coefficients for the truck parking model are given in Table 3 as follows:

Table 3 Estimated Coefficients for Truck Parking Model

Independent Variable	Model 1	Model 2	Sign
Intercept	-1586.89	-1475.79	-
Percent of truck	1.41039	1.5478	+
Parking Duration	0.15563	0.13912	+
Total truck volume	0.06955	0.05898	+
Distance to I-81	-123.293	-114.328	-
Distance to Nearest Truck Stop	111.9563	103.7537	+
Distance to Nearest Rest Area	14.22398	13.80663	+
Service Provided	988.9973	919.6157	+

Using the estimated model, demand was forecasted for 10 and 20 years. It was found that there is a deficiency of 309 spaces at present.

2.2.4 Truck Parking in Florida

(Bayraktar, et al., 2014) made an attempt to determine the supply and demand characteristics for commercial truck parking in Florida and explored technology that can be used to improve parking management in order to better utilize the truck parking spaces at public rest areas. This research was conducted in two phases. Phase one included collection of rest areas in Florida and observation of truck parking facilities and determine shortfalls in parking supply and determine illegal parking. Phase two consisted of implementation of a smart parking management system for trucks. The study area included all of the public rest areas along the I-10, I-75, and I-95 corridors. The data collected was the number of truck parking spaces at each location and total truck parking utilization. Total parking utilization is the percentage of trucks parked both legally in the parking spaces and illegally elsewhere at the facility with respect to the available capacity. The data was analyzed and report was made on each rest area. A sample report is shown in figure 8.

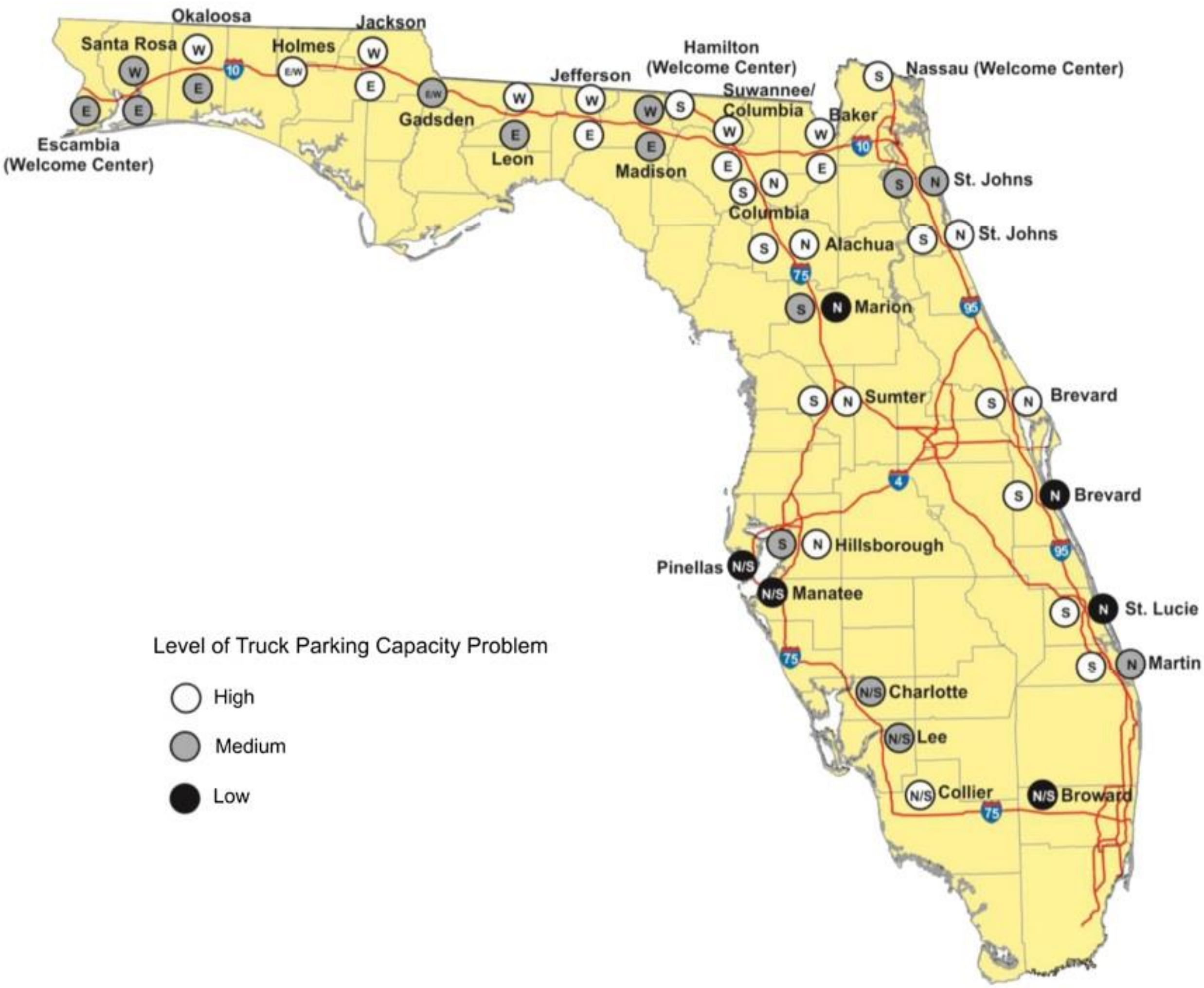


Figure 8 Truck Parking Problem Map for Florida
Source: Bayraktar et al. (2014)

2.2.5 Truck Parking in Tennessee

(Chatterjee, et al., 2000) presented a survey based study of truck parking in public rest areas along Tennessee's interstate highways. The study area included all public rest areas in Tennessee. The data was collected in survey form and these were the occupancy of each space from 10 PM to 6 AM along with some identifying information about the trucks, like company name, color, and the configuration of the truck / trailer. The data was analyzed and reported as shown in Figure 9.

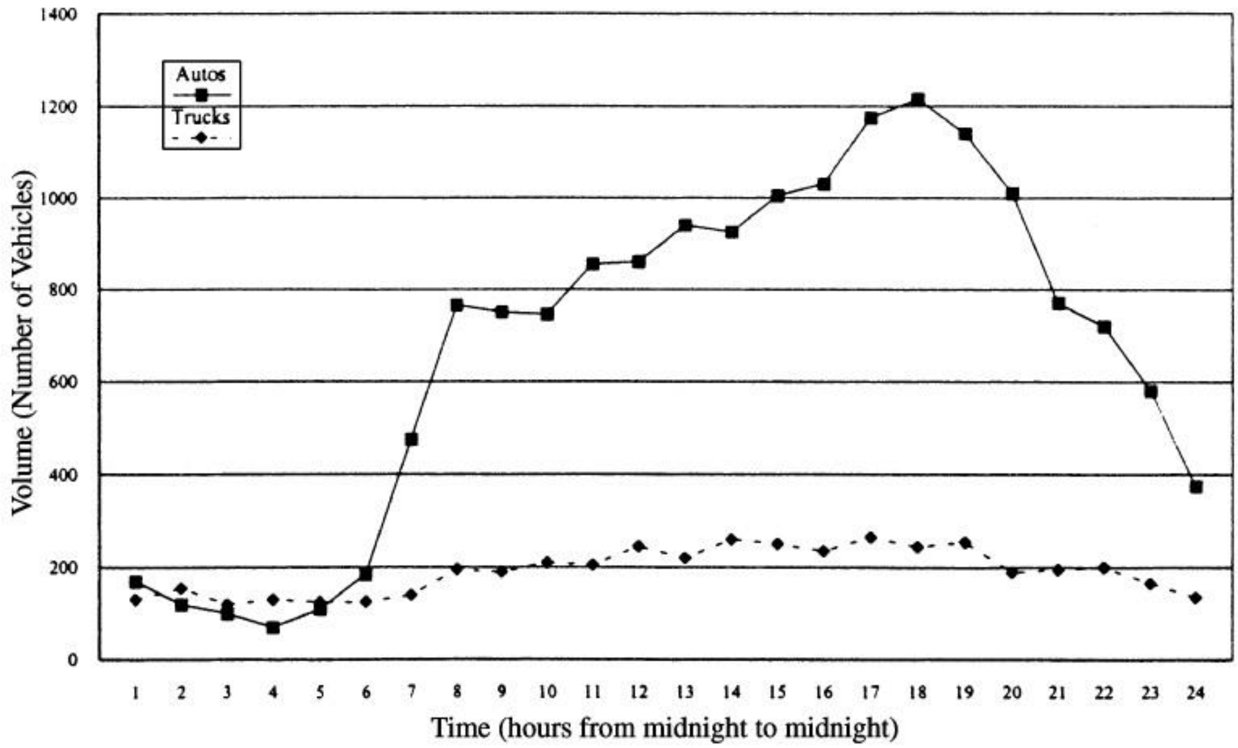


Figure 9 Truck Parking Problem Map for Florida
Source: Chatterjee and Wegmann (2000)

Various findings were reported from the data statistics. It is mentioned that rest areas are more heavily used on Mondays through Thursdays with Monday and Tuesday being the busiest of days. Also, it is mentioned that among the trucks parked inside the rest area, nearly 75% occupy a space for more than 4 hours.

2.2.6 Truck Parking in Minnesota

Minnesota Interstate truck parking study was done for developing information for supporting future truck parking decisions (Maze, 2008). The study examined the supply and demand of public and private commercial vehicle parking. The study area was Minnesota's three primary interstate corridors, I-90, I-35 and I-94. The study was conducted through 3 phases. Phase one consisted of collection of data regarding truck parking demand by time of day. Aerial photographs, State DOT maps and google maps were used to obtain the parking supply information and site characteristics. Phase two consisted of truck parking demand analysis. The data collected in phase one were compiled and field records were supplemented with truck parking capacity usage database. This data was summarized and a measure was developed to identify over capacity facilities. Table 4 shows an example of capacity constraints on I-90.

Table 4 Capacity Constraint on I-90

Interstate 90 Rest Areas	Truck Parking Capacity				Truck Stalls
	Adequate	15%	25%	50%	
Beaver Creek (E.B.) Exit 0	x				16
Adrian (E.B.) Exit 25		x			6
Adrian (W.B.) Exit 26			x		7
Clear Lake (E.B.) Exit 69			x		10
Des Moines River (W.B.) Exit 72	x				9
Blue Earth (E.B.) Exit 118		x			10
Blue Earth (W.B.) Exit 119	x				11
Hayward (E.B.) Exit 161		x			10
Oakland Woods (W.B.) Exit 171			x		10
Highforest (E.B.) Exit 202			x		6
Marion (W.B.) Exit 220				x	20
Enterprise (E.B.) Exit 244	x				11
Dresbach TIC (W.B.) Exit 275	x				5

Source: Maze (2008)

Phase three consisted of conducting survey on trucking companies within 48 hours of vehicle observation to find out information about the attitude and behavior of drivers. 433 motor carriers were given 89 survey questions to which only 41% responded. All these information was processed and demand and supply maps for the state of Minnesota was created. Based on the demand and supply scenarios, it was found that five rest area facilities are at or over capacity 50% of the time. Several other rest areas are over capacity at least 25% of the time. The authors recommended immediate investment to these rest areas.

2.2.7 Other Studies

Many other studies have been conducted regarding truck parking. Moreover, several other studies have been conducted regarding truck GPS data and truck trips which are summarized in Table 5. (Davis , 1997) performed an empirical research at the state and national level to express the concern for additional truck parking space along U.S. interstate highways. Survey was done to measure truck parking supply and demand through peak-period (late night and early morning) at four public rest areas and three private truck stops along a 200 mile segment of I-81. Total number of available parking spaces were counted and legal and illegal space utilization was monitored on an hourly basis. Capacity and facility characteristics were also determined. It was found that large numbers of trucks were parked illegally on shoulders and ramps of rest areas, often before the corridor reached capacity and even when legal parking spaces were available at a rest area.

(Fleger, et al., 2002) investigated the adequacy of commercial truck parking facilities that serves the National Highway System (NHS). The study involved multiple tasks including national survey of truck drivers, develop an inventory of public and private rest areas and developing a truck parking demand model. Nationwide survey of parking spaces at PRAs was done to find the number and characteristics of Government owned spaces for trucks. Also, an inventory of CTSs was created and maintained by Interstate America. Using these data, truck parking demand model was estimated on a highway segment considering the daily truck volume across the segment and other parameters. The simplified demand model is described in the following:

$$D = THT\ .\ P_{avg} \tag{58}$$

Where, D is the demand along a highway segment

THT is the total truck hours of travel per day

P_{avg} is the average parking time per truck-hour of travel.

THT is calculate by:

$$THT = P_t . AADT . \frac{L}{S} \tag{59}$$

Where P_t is the percent of vehicles consisting of trucks

AADT is the average daily traffic

L is the length of segment

S is the speed limit or average truck speed

Table 2 shows the model parameters were used to adjust the truck volume estimate in equation (59) and Pavg is equation (58).

Table 5 Demand Model Parameters

Parameter	**Description**	**Value**
F_s	Seasonal peaking factor	1.15
SH/LH	Short-haul to long-haul ratio	0.36/0.64, 0.07/0.93
D_{ST}	Short-term parking duration per hour traveled	5 min/h
T_{DRVING}	Time driving for long-haul drivers	70 h/8 days

T_{HOME}	Time at home for long-haul drivers	42 h/8 days
$T_{LOAD/UNLOAD}$	Time loading and unloading for long-haul drivers	15 h/8 days
$T_{SHIPPER/RECEIVER}$	Time at shipper/receiver for long-haul drivers	16 h/8 days
P_{RA}, P_{TS}	Portion of demand for public rest areas/commercial truck stops	0.23, 0.77
PPF_{SH}	Peak-parking factor for short-haul trucks	0.02
PPF_{LH}	Peak-parking factor for long-haul trucks	0.09
PR_{LH}	Long-haul parking ratio	0.7833

Then the parking demand and supply of a highway segment was compared to find if there is a shortage or surplus. The two most important factors that contribute to the demand for truck parking are the need to comply with Federal Hours of Service (HOS) rules and the need for drivers to perform certain non-driving activities like eating, fueling.

Table 6 Collected Studies Overview

Serial No.	Literature	Objective	Type of study	Study Area	Data Used	Tools/Models developed
1	(Adams, et al., 2009)	Truck parking issues	Survey based	Interstate 43, U.S. Highways 8, 10, 41, 51, 53 and 151, and State Trunk Highway 29	online and paper-based surveys, in-person and telephone interviews	GIS survey tool
2	(Bayraktar, et al., 2014)	Truck parking supply trends	Survey based	public rest areas along the I-10, I-75, and I-95 corridors	Truck parking characteristics, shortfalls in the spaces, illegal parking	Occupancy prediction model

3	(Chatterjee, et al., 2000)	Parking occupancy characteristics	Survey based	Rest areas in TN	Survey data and truck characteristics	N/A
4	(Davis , 1997)	Truck parking space shortfall	Survey based	4 public and 3 private rest area along I-81 corridor	Truck parking space utilization on hourly basis	N/A
5	(Garber, et al., 2002)	Truck parking supply and demand estimation	Inventory and survey based	Rest areas and truck stops at I-81	Parking duration and accumulation data	Stepwise regression model for demand
6	(Parametrix, 2005)	Find truck parking shortage	Inventory and survey based	I-5, I-82 and I-90	Parking demand data by recordings	Growth factor for future demand forecasting
7	(Fleger, et al., 2002)	Investigate adequacy of truck parking facilities	Inventory and survey based	Truck parking facilities serving the National Highway System (NHS)	Interstate America database of commercial truck stops	Simplifies demand model
8	(Maze, 2008)	Examined truck parking supply and demand	Inventory and survey based	I-90, I-35, I-94	Truck parking count data by time of day, site information	N/A
9	(Fallon, et al., 2011)	Determine truck parking availability	Technology based	Truck stop at US 1 and	Truck counts	Magnetometer device for

		using magnetometer		public rest stop on I-95		truck counting
10	(Fischer, et al., 2006)	Develop performance measures to evaluate strategies for reducing truck trips	Technology based	Ports of Long Beach and Los Angeles	Port and cargo characteristics	QuickTrip truck trip generation model
11	(Gaber, et al., 2005)	Review various literature and methodologies to assess truck parking availability	Review and methodology based	Nebraska Interstate corridors	Survey data from stakeholder focus group	N/A
12	(Garber, et al., 2004)	Proposed methodology for improving truck parking information system	Methodology and technology based	I-81	Truck driver survey, truck crash data, truck AADT, traffic facilities	Prototype of truck parking information system
13	(Gentler, et al., 2011)	Field operational test of parking monitoring	Technology based	Charlton Westbound Service Center on I-90	N/A	N/A
14	(Haghani, et al., 2013)	Improve truck parking safety using technology	Technology based	Truck parking facility at I-95 northbound prior to MD 32	N/A	automated real-time parking information system

15	(Heinitz, et al., 2009)	Proposed demand modeling approach for scarce truck parking facility	Application based	Germany	HGV inflow or time-variation curves of road freight transport demand	Car park choice model
16	(Kawamura, et al., 2014)	Identify factors for truck parking violation	Inventory based	Chicago urban area	parking citations for 12 month period	Simple regression model
17	(Mbiydzeny uy, et al., 2012)	Developing concept of intelligent truck parking	Technology based	Sweden	Parking occupancy and vehicle location data	N/A
18	(Pecheux, et al., 2002)	Estimate the distribution of truck parking demand and supply along the NHS	Application based	29 highways segments on I-81	Truck AADT, % of trucks, length of segment, speed limit or average truck speed	N/A
19	(Rodier, et al., 2007)	Explore the truck parking problems and solutions	Review and Survey based	California	Trucker survey data	N/A

2.2.8 Literature Gap

A thorough literature review on truck parking has shown that the prediction of parking demand in the rest areas is dependent on several factors and many of them have not been addressed. Moreover, none of the research methodologies insisted on using truck GPS data for estimating demand and supply. Research to estimate the parking demand using truck GPS data should be taken into consideration, and identify factors affecting parking demand for design purposes and also identify locations which requires rest areas or truck parking so that the truck drivers get ample rest thus increasing safety and efficiency.

The state of Tennessee acquired few approaches to address the growing demand for truck parking subsequent to the 1996 Study. The University of Tennessee led nighttime observational studies at all public rest areas in Tennessee to learn about the parking space occupancy characteristics of trucks. They examined the availability of space in private truck stops near interchanges. Their results showed that the rest areas were swarming with trucks at night, since a lot of trucks were found parked along the shoulders of highway exit and entrance ramps, as well as on interchange ramps. On the other hand, around 30% of the private truck parking spaces remained vacant (Pecheux et al., 2002). Interview was also held to understand why some truck drivers parked along the highway when there were available private parking spaces.

However, it was a preliminary study and does not explain the demand or shortage during the different time periods like peak and off peak period. Also, nothing in the literature was found regarding parking supply. In order to get an accurate estimation of the supply, a thorough analysis must be done to minimize the truck parking shortage. This involves analyzing various other factors that may affect the truck parking.

2.3 Freight Performance Measures (FPMs) Using Truck GPS Data

The evaluation of corridors' performance is essential in identifying bottlenecks and determining network sections that need to be improved. Past practices include travel diaries and traffic counts but these practices tend to be time consuming and with low accuracy. Since every truck in the U.S. is equipped with a GPS device researchers have explored possibilities of using information from these devices not only to calculate FPMs but also to define travel patterns and make prediction models. The performance measure used mainly in studies is travel time (TT), hence this part of the literature review is classified in three categories based on how TT is computed: i) link TT (LTT), where travel time is computed for a link; ii) trip TT (TTT), where travel time is calculated for a trip or tour; iii) miscellaneous – different from i and ii.

2.3.1 LTT Focus

Quiroga & Bullock (Quiroga, et al., 1998) proposed a methodology to perform studies for estimating TT of roadway segments using GPS and Geographic Information System (GIS) technologies. GPS data were collected from three metropolitan areas in Louisiana, LA (i.e., Baton Rouge, Shreveport, and New Orleans). Average TT and travel speed (TS) values were computed for all highway segments. A length of segment comprised 0.2-0.5 miles. GIS was utilized to process queries, produce reports and colored-theme maps, depicting TT by link. Results showed that shorter GPS sampling periods (1 to 2 seconds) decreased errors in TS estimation. The authors underlined that median speed was a more accurate measure of the central tendency than mean speed as the latter was affected by incidents occurred during peak hours. Quiroga (Quiroga , 2000) conducted a similar study for the LA transportation network (Baton Rouge). Highways were separated into segments, and LTT was calculated for each segment. The author also provided a procedure for estimating several other performance measures (acceptable TT, segment TS, travel rate, delay, total delay, delay rate, and relative delay rate) that could be used for quantifying congestion.

Storey & Holtom (Storey , et al., 2003) used GPS data to compute link TS (LTS) and LTT at West Midlands highways in the UK. The GPS device provided information every 60 seconds, while a vehicle ignition was being on. Around 20% of the data were discarded, as they provided coordinates (latitude and longitude) that didn't belong to the road network. Links of the considered highways were separated into 50 m segments, and the average TS was calculated for each segment. It was assumed that segments between two GPS data points had the same average speeds. The journey times at the link level, estimated using GPS data, were calibrated, and results demonstrated an acceptable accuracy of the proposed approach. The analysis of journey speeds indicated the existence of congestion issues at major junctions of links, leading to the city center.

Jones et al. (Jones, et al., 2005) presented a methodology that could be applied to measure performance of busy freight corridors. The procedure was separated in 4 steps: 1) identification of freight corridors, 2) review of data collection technologies, 3) System Alpha Test, and 4) System Beta Test. Top ten US cities with the highest truck volumes were identified using American Transportation Research Institute (ATRI) satellite position reports. The busiest freight corridors were determined for each of those cities based on the data, provided by Cambridge Systematics. Different methods of data collection were described: satellite-based systems, terrestrial wireless systems, hybrid systems, on-board systems, and fixed site systems. GPS was found to be efficient for the analysis. The Alpha Test was performed to associate a vehicle ID with a highway segment geo-position, to calculate the average vehicle TS, and to remove outliers that could affect the accuracy of speed estimation. The main purpose of the Beta Test was to process TT and TS at each segment and to transfer the data to the visualization tool. As a result of the conducted study, the authors created a map, depicting the average TS at the busiest US corridors.

Ando & Taniguchi (Ando, et al., 2006) developed a model for the vehicle routing problem with time windows (VRPTW), minimizing the total cost of LTT uncertainty and penalties due to early arrival/delayed arrival to customers, requesting a particular time window. The information on LTT was collected using sensors, radio beacons, and GPS devices. Truck arrival times were assumed to follow a normal distribution. Statistical TT distributions were obtained for each link and were approximated to triangular distributions. An additional linear regression analysis was performed to quantify relationship between LTT and link distance. The traffic flow simulation was used to estimate TT distribution for each route and determine the optimal visiting order of customers. Results indicated that the proposed approach reduced the total cost by 4.1%, the total cost standard deviation by 75.1%, and mitigated environmental impacts, caused by trucks.

Schofield & Harrison (Schofield, et al., 2007) underlined the importance of FPMs for the US Department of Transportation (DOT), State DOTs, and various transportation agencies. Practices for assessing performance of freight corridors, employed in different states, were described in the report. The study focused on developing appropriate FPMs in the Texas (TX) area. The busiest state highways were identified. GPS records were provided by ATRI for the entire year of 2005. The authors indicated that the location error for each observation could reach up to ¼ mile. The segment length comprised 50 miles. TT, TS, and TT index (TTI) were estimated for each segment. Changes in travel pattern were noticed when the Hurricane Rita notification was announced. The report provided distribution of hourly truck traffic. Future research directions included comparison of the actual speed with the free-flow speed for each segment, estimating FPMs for highway corridors in case of non-recurring congestion, calculating of truck wait time at boarders, consideration of other FPMs, etc.

Liao (Liao, 2008) compared two ATRI FPM database systems: the GIS – based system and the Structured Query Language (SQL) – based system. The second system was able to process truck GPS data without the GIS software. The GIS-based system allowed separation of a highway into segments with minimum size of 10 miles. The minimum segment size for the SQL-based system was 3-miles. It was found that smaller segments improved accuracy of average speed estimation. The author underlined the importance of trip filtering parameters and projection algorithms. The GIS-based system employed a ¼ mile radius search method, while the SQL-based system used more complex snapping algorithm. Several deficiencies of the SQL-based system were mentioned (e.g., duplication of data in tables). According to the report, the ideal FPM system should include the SQL-server, capable to process data from external applications and visualize performance measures using a GIS - based software.

Liao (Liao, 2009) evaluated performance of I-94/I-90 freight corridor between St. Paul, Minnesota (MN), and Chicago, Illinois (IL). GPS data for 12 months (May 2008-April 2009) were provided by ATRI. The raw data were processed in ArcGIS software, GPS points were snapped to the nearest route, and then the average TS was computed for each 3-mile segment. The analysis was performed for the key corridor locations (i.e., St. Paul, O'Hare Airport, I-90 toll highway), including truck speed, volume, TT reliability, truck stops, truck stop duration, etc. Results indicated that average speeds declined in areas approaching Chicago from 55 mph to 40 mph and lower. The westbound traffic between St. Paul and Madison had higher speed standard deviation than the eastbound traffic. A significant speed standard deviation and the average speed drop were observed on I-90 toll highway, leading to Chicago.

McCormack (McCormack, 2009) described how GPS data were used to improve performance of the Washington State (WA) freight network. LTT and its reliability were chosen as performance measures. The data were collected from various vendors. GPS records were received with frequencies, varying from vendor to vendor (every 30 seconds, every half-mile, every 15 min, etc.). ATRI and FWHA developed a program, focusing on performance of interstate corridors. A specific algorithm was developed to define origin and destination of each trip, using stop time, travel distance, GPS signal quality, and location of travel. It was highlighted that some GPS points were removed as they provided erroneous data. In some cases truck information was known only every 15 min. The author concluded that truck GPS data could be very useful for public agencies to evaluate conditions of busy freight corridors and to identify bottlenecks.

The Washington Department of Transportation (WSDOT) outlined the main features of the Truck Performance Measure Program at the Washington State Transportation Commission (WSDOT, 2011). The WSDOT initiated this program in 2007. GPS data process and analysis are similar to the ones, described by McCormack (McCormack, 2009). LTT and its reliability were selected as performance measures. The main objective of the program was to identify and rank bottlenecks at the WA State highways. Four criteria were developed for prioritizing highway segments for further improvements: 1) Truck speed below the congestion threshold (60% of posted speed limit); 2) Average speed; 3) Speed distribution; 4) Truck volume. The authors underlined that the program was efficient, and its future success would be highly dependent on the access to the data, owned by trucking companies. McCormack et al. (McCormack, et al., 2011) and McCormack & Zhao (McCormack, et al., 2011) conducted a similar study, using the same FPMs as McCormack (2009). The authors described the process of bottleneck identification and prioritization in WA. The overall procedure was subdivided into 5 parts: a) Segment the roadway; b) Add attribute information to the segments; c) Geo-locate the truck; d) Locate the bottlenecks; e) Rank the bottlenecks.

Chien et al. (Chien, et al., 2011) estimated link and path TT, variability of TT by departure time of the day and days of the week for 18 New Jersey highway corridors. The data were collected from GPS enabled devices, installed into different vehicles, traveling along considered highways between October 8, 2007 and April 21, 2008 from 6.15 am to 8.15 am during weekdays. The buffer index (BI) and 95th TT percentile were calculated for each route. Results indicated that TT on the most of roads followed a shifted log-normal distribution. The lowest mean TS was found for a segment NJ 208 & NJ 4 (28.3 mph), while the highest one was determined for a segment NJ 24 & I-78 (59.9 mph). The highest TT coefficient of variation (TTCV) was calculated for a segment US 46 & NJ 3 during A.M. peak hour (TTCV=0.4). The lowest TTCV was estimated for US 1 (TTCV=0.09). The scope of research didn't include assessment of incident impacts on link/path TT due to data limitations.

Cortes et al. (Cortes, et al., 2011) used GPS data to evaluate performance of a bus transportation system in Santiago, Chile. Data were collected for 6,178 buses operating over a one week. The authors applied a path rectification procedure to determine paths for each route. The path rectification identified line segments that were located close to GPS points with an acceptable error. Rectified paths were separated for grid elements. An average bus TS was calculated for each grid element. The report presented speed diagrams illustrating bus speeds for each route segment during a given time of day. The proposed methodology was found to be efficient for problem identification in bus operations (e.g., low speeds at certain segments, congestion issues, improper traffic light times, etc.).

The Federal Highway Administration (FHWA) Office of Freight Management and Operations (FHWA, 2011) developed a Freight Performance Measures (FPM) web Tool to evaluate performance of the US freight corridors using truck GPS data. The FPMweb Tool estimates the operating speed of a given segment by averaging over the total number of speed observations. The segment length was assumed to be 3 miles. The tool can process data by time and date for 25 interstate corridors. Several drawbacks of the tool were mentioned: 1) it doesn't provide commodity and origin-destination data; 2) it is not capable to forecast future truck volumes and speeds; 3) it is useful for analysis of average and not individual truck TS.

Figliozzi et al. (Figliozzi, et al., 2011) developed an algorithm for assessing TT reliability of the I-5 interstate in Oregon (OR). GPS data were provided by ATRI. The corridor was separated into particular segments. Traffic flows were estimated for every mile and direction of each segment. Smoothing was performed by averaging counts for 20-miles segments. Volumes were also determined for different seasons of the year. Segments were analyzed based on two factors: a) time of the year and corresponding weather conditions, and b) truck density pattern along the segment. The designed algorithm was able to estimate 95%, 80%, and 50% percentile TT for each segment (if traffic counts were sufficient at considered segment) using GPS data. Minimum and maximum TS limits (10 mph and 80 mph) were set to remove outliers. Results indicated that differences between three types of TT (i.e., 95%, 80%, and 50% percentile TT) were significant for urban areas and relatively small for rural areas. TT costs per mile were calculated and presented in the paper.

Wheeler & Figliozzi (Wheeler, et al., 2011) assessed effects of recurring and non-recurring congestion on freight movement characteristics (LTS, LTT, and TT reliability) at the Oregon I-5 Interstate (the same freeway as studied by Figliozzi et al., 2011). Along with GPS data, the authors used corridor TT loop data and incident data (provided by the Oregon DOT). A specific methodology was developed to identify through trucks (that don't make any stops and provide at least two GPS readings in the beginning and in the end of the corridor). Results of a recurring congestion analysis indicated that the highest TT and TTCV were observed during evening peak. As for non-recurring congestion, it was found that incidents significantly affected truck TS in the incident area throughout the day. Congestion cost estimates indicated that daily delay costs for freight vehicles were 19% higher that free-flow costs without variability consideration (and 22%-31% higher with variability consideration). GPS data were found to be more accurate in estimating TT than the loop sensor data.

Blazquez (Blazquez, 2012) addressed the problem of snapping GPS points to roadways segments. Various techniques, resolving spatial ambiguities, were listed (e.g., semi-deterministic map-matching, probabilistic map-matching, fuzzy logic map-matching, Kalman filter approach, etc.). The author developed a topological map-matching algorithm for snapping GPS points. The algorithm was able to identify a feasibility of the path between two snapped points (by comparing a speed along the path and the average vehicle speed). Numerical experiments were conducted using the data, collected by winter maintenance vehicles in Wisconsin (WI) and Iowa (IA). Preliminary calculations were performed to determine the buffer size. Results demonstrated the efficiency of the presented methodology. It was found that the GPS spatial error decreased the percentage of solved cases on average by 30%. Frequent sampling intervals provided more accurate results. An increasing number of consecutive GPS points improved performance of the algorithm.

Liao (Liao, 2014) used GPS data, provided from ATRI for twelve months in 2012, to estimate FPMs, such as truck mobility, delay, and reliability index, and to identify bottlenecks for 38 key freight corridors in the Twin Cities metropolitan area (TCMA). To validate the methodology the computed average truck speeds and hourly volume percentage at certain locations were compared with the data from weight-in-motion (WIM) sensors and automatic traffic recorders (ATR). Truck bottlenecks were identified and ranked based on hours of delay and number of hours with TS less than the target speeds, set by Minnesota DOT during A.M. and P.M. peak hours. Also the truck congestion cost was estimated for TCMA to be $212 and $286 million annually based on ATRI's truck operation cost and Texas Transportation Institute's (TTI) truck congestion cost respectively. As another part of the study, one month data from FHWA's National Performance Management Research Data Set (NPMRDS) was used to compute freight mobility and speed variations along Minnesota's National Highway System.

Wang et al. (Wang, et al., 2014) suggested naïve and mapping methods to estimate LTT using GPS data. The naïve method computed the average TS and its variability on each link individually. The variability was measured by a standard deviation. The authors presented a mathematical formulation for a mapping method with an objective, minimizing the total difference between the recorded trip times and the estimated trip times for all trips. Both methodologies were tested on the San Antonio corridor (TX) and the Milwaukee highway corridor (WI). The mapping method was found to be more efficient, since it was able to analyze truck trips with large road intervals covering multiple links.

Gong et al. (Gong, et al., 2015) used truck GPS data to estimate link travel times a highway corridor in Wisconsin using a regularized regression model that maximizes the likelihood of obtaining the observed trip travel time while penalizing changes in speeds on adjacent links. Trip travel time is the duration between two successive timestamps and trip length is obtained as the roadway length traversed. Basic assumption of the model is that travel speed of a trip is constant along a link while a trip traverses several partial/full links. The proposed method results were found to outperformed results obtained from a simple OLS regression and a benchmark method. Namely, one hour traffic data collected from double loop detectors was used for validation and it was found that regularized regression method improves the travel time allocation results from the benchmark method, trip travel time allocation errors decrease as link speeds grow and travel time allocation error increases as variation of speed within link grows.

Mishra et al. (Mishra, et al., 2015) used truck GPS data provided by ATRI to calculate link based FPMs on Tennessee freight network. The study provides a guideline on how GPS data should be preprocessed and pinpoints possible problems researchers may face with this type of data. Besides estimating link FPMs the GPS data was used to develop turn times regression models for different types of freight facilities, calculate occupancy and entry/exit volumes. The researchers also developed two algorithms to analyze truck trips. The first one identifies intercity truck trips having as input the TN TAZs while the second one detects inter and intracity trips and their characteristics (dwell times, traffic light stops etc).

2.3.2 TTT Focus

McCormack & Hallenbeck (McCormack, et al., 2005) suggested two data collection methodologies to evaluate truck movements along particular roadway corridors in WA and to measure performance of freight mobility improvement projects against benchmarks. The first approach was based on implementation of Commercial Vehicle Information System and Networks (CVISN) electronic truck transponders, which were installed on the windshields of approximately 20,000 trucks. A specific program was designed to estimate TTT using the data, provided by transponders. Another technology employed GPS devices that transmitted truck movement records every 5 seconds. The information, collected using CVISN and GPS, was processed to identify congested segments, TTT, and TT reliability. It was highlighted that both techniques might be efficient for analysis of truck trip patterns. However, selection of a methodology should depend on the data required for a particular benchmark project.

Greaves & Figliozzi (Greaves, et al., 2008) processed passive GPS data from 30 trucks to identify characteristics of freight movements in the Greater Melbourne region, Australia. The authors underlined difficulties of getting GPS data from trucking companies. The GPS device was installed into each truck and provided second-by-second information. The trip identification algorithm was developed to determine trip ends. Around 5% of records were inaccurate due to loss of satellite signal and were excluded. The final output of the processed data included a summary for all truck trips and tours. The average number of stops per tour was found to be 12.2 stops. The lowest average TS were observed for morning and evening peak hours. A trip length distribution was presented in the paper. It was mentioned that GPS data didn't provide additional information about driver behavioral features (respond to weather, empty/loaded vehicle, type of commodity, etc.) that might be useful for the analysis.

NCHRP Report 008 (2010) highlighted the importance of truck GPS data for evaluation of freight corridors performance. The study was conducted for the following metropolitan areas: Los Angeles (California CA), Chicago (IL), Phoenix (Arizona AZ), and Baltimore (Maryland MD). GPS records were used to identify the number of stops during the trip, distance between stops, stop purpose, stop location, TT between stops, etc. It was found that likelihood of making trip in the tour depended both on the truck trip purpose in the current and subsequent stops. Besides, the information about trip origin, origin land use, and trip destination could be used to predict the destination land use. The highest percent of stops in industrial land use (27%) was observed in Chicago. Retail and commercial land use stops were more common in Los Angeles (31%). The most of residential land use stops occurred in Phoenix (31%).

Bassok et al. (Bassok, et al., 2011) demonstrated how truck GPS data, collected from the device vendors, could be used for the analysis of freight movements in the WA area. The authors developed an algorithm for identifying trip ends. Truck stops for refueling, rest and delivery were filtered out (dwell time threshold comprised 180 sec, which is a common standard in WA). A threshold speed limit of 5 mph was set to determine trip ends. The analysis was performed for 91 days in the Puget Sound region (WA), when 2,400 trucks made 22,000 tours and 215,000 individual trips. Results indicated that each truck made on average 9 tours and 10 trips per tour. Besides, around 2 truck trips at each tour were made to grocery stores. Areas with higher population density produced more truck trips.

Golias et al. (Golias, et al., 2012) used truck GPS data to analyze freight movements within the Greater Memphis area in TN. Available data provided information about truck trips from September 1, 2011 to October 30, 2011. The highest truck volumes on I-40 were observed during evening peak hour between 4 pm and 5 pm. Trip durations were increasing for a period since 10 pm until 8 am. This was explained by the fact that most of truck drivers stopped for rest during that time interval. Truck turn times were considered for 4 types of facilities: public warehouses, private warehouses, distribution centers, and intermodal facilities. The authors developed regression models predicting facility turn times depending on the truck volume per time interval and facility type. The overall fit of proposed models was found to be low due to small sample size. Intermodal facilities and private warehouses demonstrated the best fit. The scope of research included truck stop and rest stop demand analysis. All truck stops with duration from eight to twelve hours were considered. The authors provided frequency of truck stops based on the time of the day for major TN rest stop areas.

Pinjari et al. (Pinjari, et al., 2012) (Pinjari, et al., 2012) (Pinjari, et al., 2013) investigated how GPS data, provided by ATRI, could be used for assessing performance of freight corridors and transportation planning in Florida (FL). The study was directed to identify FPMs for state highways, build a truck-trip database to understand truck travel patterns, and derive truck trip O-D tables for the Florida Statewide Model. Several FPMs were suggested, such as average trip TS (TTS), reliability measures (TTI and Planning Time Index PTI), analysis of chokepoints, truck flow analysis, etc. Truck flows were estimated by month of the year and by day of the week. It was found that seasonal variations of truck speeds were not significant. However, travel patterns during weekdays were different as compared to weekend travel patterns. Trip Origin Destination Identification algorithm was designed to define O-Ds. The procedure was validated based on comparison with Google Earth and discussions with ATRI and FDOT. Trip length and trip duration distributions were provided in the report.

You (You, 2012) studied tour-based models for drayage trucks at San Pedro Bay Ports in Southern California area. The main objective was to develop a methodology, which could help to alleviate congestion of trucks at the gates, reduce truck turn times at the ports, and mitigate environmental impacts. A tour-based approach was found to be more efficient for modeling behavior of drayage trucks than a single trip-based approach. GPS data for 545 drayage trucks was provided by the ports of Los Angeles and Long Beach. The collected data were processed to identify closed and open tours. It was observed that each truck made on average 1.7 tours and 6.2 stops per day. A typical tour TT lied between 3 and 9 hours. The author suggested two approaches to analyze trip-chaining behavior of drayage truck movements: 1) A disaggregate level tour-based model based on Sequential Selective Vehicle Routing Problem (SSVRP); 2) An aggregate level tour-based model based on Entropy Maximization Algorithm (EMA). It was underlined that the SSVRP was more realistic approach for modeling drayage truck tours.

Bierlaire et al. (Bierlaire, et al., 2013) used GPS data, generated by smartphone Nokia N95, for route choice modeling in the Lausanne area, Switzerland. The authors listed advantages (short warm-up time, full track of trips) and disadvantages (weak signals, not accurate data points in some cases, high energy consumption) of GPS capable phones.

A probabilistic map matching method was developed to estimate the likelihood of choosing a particular path based on the smartphone GPS data. A path with a higher log-likelihood was more preferable among all alternative paths. Speed distributions were generated from the observed speed data. Data points with speeds less than 8 km/h were filtered out. Results obtained by the suggested approach were close to the ones, provided by the Mobility Meter (dedicated GPS device, carried by the person along with smartphone).

Carrion & Levinson (Carrion, et al., 2013) assessed the effect of converting I-394 (between Minneapolis and St. Paul, MN) High Occupancy Vehicle (HOV) lanes to High Occupancy Toll (HOT) lanes. The main objective was to determine a traveler's respond to increasing TT reliability on HOT lanes. The GPS devices were installed in 54 vehicles to collect the detailed trip information. A 20-meter buffer was used for all roads. GPS points, located outside the buffer area were excluded. The authors developed an algorithm to identify the commute trips (from origin to home location, from destination to work location and vice versa). The preference of travelers for choosing tolled or non-tolled routes was analyzed using discrete choice models. The utility function included TT measures, travel cost, and socio-demographic factors. TT reliability was measured by standard deviation, shortened right range, and interquartile range. Results of study indicated that the desire of travelers to pay tolls for reliable routes was dependent on how they perceived reliability savings.

Golias & Mishra (Golias, et al., 2013) used truck GPS data, provided by ATRI for the months of September and October 2011, to evaluate the impact of the new Hours of Service (HOS) rule for Commercial Motor Vehicles (CMV) drivers on traffic conditions using as case study a part of I-40 network between Memphis and Nashville, TN. Existing truck TTT and volume by time of day on a daily and weekly basis were computed by statistically analyzing the provided data, while future conditions were estimated for the shifted truck trips which had to be identified based on the new working hours. The Level of Service (LOS) for both cases was calculated based on the methodology suggested in Highway Capacity Manual with some adjustments because of the low percentage of data used. By comparing LOS in both cases it was found that the new HOS would worsen LOS, as truck volumes would increase at certain routes after each rest period, which might cause delays.

Kuppam et al. (Kuppam, et al., 2014) demonstrated how truck GPS data could be used for Tour-Based Truck Travel Demand Modeling. The study was conducted based on GPS data for 22,657 trucks and 58,637 tours, purchased from ATRI. The number of tours for each truck was determined using the information about truck coordinates, changes in TT and TS. The accuracy of vehicle stops was checked using highway maps and Google Earth. The following Tour-Based Truck Models were developed for the Phoenix region

(AZ): tour generation, stop generation, tour completion, stop purpose, stop location, stop time of day choice. It was found that construction tours had lower tendency to making stops, while government-related tours were dedicated to making more stops. An increasing number of stops caused incompletion of tours for the majority of trucks. The purpose of the previous stop influenced duration of the next stop.

2.3.3 Miscellaneous

Fisher et al. (Fisher, et al., 2005) proposed a modeling framework to evaluate the Los-Angeles County (CA) freight transportation network performance. The framework combined characteristics of logistics chain and tour-based models. Logistics chain models were found to be useful for cases, when particular types of goods were transported from the production points to the assigned destinations. Those models combined information from three layers: economic, logistics, and transport. Tour-based models were efficient to determine vehicle tours and trips without focusing on commodity type. Those models provided the following information: generation of tours by zone, number of stops during the tour, stop purpose, stop time, stop location, number of trips during the tour, etc. The suggested integrated framework was found to be promising for analysis of freight movements.

Cambridge Systematics (Cambridge Systematics, 2007) indicated that GPS devices could be effectively employed along with travel diary surveys for data collection and understanding truck traveling patterns in urban areas. Several disadvantages of using diaries were mentioned: 1) process of data depends on willingness of drivers to complete the form, 2) lack of the contact information, 3) some vehicles may not be registered in the study area, 4) low response rates due to confidentiality issues, etc. GPS devices, installed into trucks, might be utilized to validate the data, collected from driver diaries (e.g., trip origin, trip destination, routing, speeds at particular road segments). However, GPS data don't provide any information regarding commodity hauled, size of shipment, and type of carrier operation (e.g., truckload, LTL, private). Besides, high cost of GPS devices was found as a major implementation issue.

NCHRP Report 818 (2008) suggested a set of performance measures that can be used to evaluate highway conditions. Performance measures were classified into two categories: individual measures (related to an individual traveler) and area measures (related to the area, region or corridor). Delay per traveler, TT, TTI, BI, and PTI were referred to individual measures. Area measures included total delay, congested travel, percentage of congested travel, congested roadway, and accessibility. The report also distinguished between the performance measures as primary and secondary depending on the analysis area.

Dong & Mahmassani (Dong, et al., 2009) developed a methodology for estimating TT reliability. TT reliability was associated with traffic flow breakdowns and delays. A probability distribution function for pre-breakdown flow rate was calibrated using field data, from I-405 Irvine freeway in CA. The normal distribution was the most suitable for the Jeffrey section of the freeway, while the Weibull distribution provided the best fit for the Red Hill section. The authors assumed a linear relationship between breakdown and pre-breakdown flow rates. The delay was estimated based on TTI and flow rate values. Numerical experiments were performed for I-405, and results indicated that the proposed concept was efficient for relieving congestion and TT delays.

The Memphis Urban Area MPO (2013) conducted a Freight Peer to Peer Program meeting to exchange the best practices between regional freight industry stakeholders from public and private sectors, and also various transportation agencies. Establishment of performance measures for freight transport was found to be a very important aspect in prioritizing highway improvement projects. It was underlined that performance measures should be set at state level with assistance of regional agencies if necessary. Performance measures should take into consideration interests of both private and public sectors.

Pinjari et al (Pinjari, et al., 2015) used GPS data provided by the American Transportation Research Institute (ATRI) to compute FPMs and develop algorithms that estimate truck trips and Origin-Destination (OD) matrices. Data consists of trucks traveled across Florida in a 4 month period. Based on truck id, GPS data for these trucks was also extracted from ATRI database for the rest of North America in order to track flows in and out of the state. GPS records information provide x, y coordinates, time and date, truck id and distance to the closest interstate, while a subset of the trucks had also spot speed information. A GIS polygon shape file with major truck stops (rest areas, weight stations etc.) was also used in this study. The developed algorithm identified potential stops (origin or destination) based on spatial movement, time gap and speed between consecutive observations for the same truck and eliminated possible stops less than dwell-time buffer, combined small trips (less than 1 mile) and discarded incomplete trips or trips with large time gap between observations. Then, it eliminated trip ends in major truck stops and breaks circuitous trips (ratio between air distance and cumulative geodetic distance from origin to destination less than a predefined value) into multiple ones. In the results trips were categorized in 3 types: all trips (including trips outside Florida), FL-link trips (at least one end in Florida), and FL-only trips (both origin and destination in Florida).

Besides truck trip characteristics, OD matrices for a part of the 6000 Florida TAZs were calculated and compared to travel times used in the Florida Statewide Model (FLSWM) and google maps. It has been found that factors that may affect the calculated travel distance and travel time are route choice, GPS data ping rate (time gap between two consecutive GPS records), TAZ size, time of day and number of trips between ODs. Based on the results travel distances from ATRI data were smaller compared to FLSWM and google maps mainly because of the straight line distance approximation between two GPS points. The computed travel times were found to be higher when compared to google maps results but smaller compared to those extracted from FLSWM.

Another part of this study was to examine the extent to which these trips capture observed traffic flows in Florida. Focus of the study was the truck type composition, the proportion of truck traffic flows captured by the GPS data and geographical differences in the data. Trucks that did not make at least one trip of 100 miles and trucks with more than 5 trips per day were classified as medium trucks and removed from the database. To determine the proportion of heavy truck traffic flows captured in ATRI data in Florida they were compared with observed truck traffic volumes from Telemetered Traffic Monitoring (TTM) sites in Florida. It was found that this coverage is 10% for heavy trucks, information used later to compute the seed matrix in the origin-destination matrix estimation (ODME) model. Other input to the model was a highway network of the study area, observed traffic flows on various links and OD matrices for travel volumes other than freight truck extracted from the FLSWM. Also, cells with zero flows in the seed matrix that was expected to have flows were corrected after aggregating from TAZs to county level. The ODME was evaluated for different assumptions (upper/lower bounds on trip number) and the results for one set of assumptions are presented in this study with acceptable validation results.

Lee and Ross (Lee, et al., 2015) studied how truck GPS data can be utilized for freight demand forecast at the state and regional levels. ATRI GPS data for Atlanta and Birmingham was collected for eight weeks and used to develop a tour-based freight demand model at the state/regional level in conjunction with existing data sources, employment data and transport network. The model was divided in 7 sections. First, the Tour Generation Model produces truck tours in each TAZ based on zonal characteristics. This output was used to scale GPS data. Next, the Tour Main Destination Model calculates the probability of each zone being a primary destination for tours originating from all other zones and the Intermediate Stop Model calculates how many intermediate stops there are for each tour, if any, using a multinomial logit model and identifies destination zones for each intermediate zone. The Time of Day model splits tours into different time periods and the Trip Accumulator breaks tours into truck trips that are used as inputs to the Traffic Assignment model. Link volumes from the developed model were compared to Atlanta Regional Commission (ARC's) trip based model and it was found that the new model assignment was closer to the reported traffic counts for the examined period.

Bernardin et al. (Bernardin, et al., 2015) used ATRI's truck GPS data in Iowa and Tennessee to identify possible biases and calculate ODMEs. Data used consists of 8 week truck observations for each quarter in 2012 and are already processed giving information for begin and end TAZs between two consecutive observations, distance, time, speed and status (moving/stopped). Further process of the data was needed to identify ODs for each sequence of moving records and discard bad data (GPS positional errors, partial trips, intrazonal trips greater than 30 miles). ODME algorithms applied use truck counts on the network and scaled raw ATRI trip table to represent the proper amount of VMT extracted from iTRAM. The results were analyze to evaluate if there were any biases on geographic regions or trip length. It was found that for Iowa there were no geographic biases but there was evidence of bias towards longer haul trips. Also, it was found that ATRI's ODME trip table had a smaller RMSE when compared to iTRAM results which indicated that this data can be used to produce a better model than the existed.

2.3.4 Summary

The following FPMs were identified as a result of conducted literature review:

1. 90^{th} and 95^{th} percentile travel time ($tp_{90\%}$ and $tp_{95\%}$)
2. Buffer index $BI = \frac{tp_{95\%} - \bar{x}}{\bar{x}}$

   where $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$ - mean travel time; x_i - travel time for the observation *i*; N – number of observations
3. Buffer travel time $BTT = tp_{95\%} - \bar{x}$ (minutes, hours)
4. Planning travel time $PTT = tp_{95\%}$ (minutes, hours)
5. Planning travel time index $PTTI = \frac{tp_{95\%}}{x_{FFS}}$

   where x_{FFS} – free flow speed travel time
6. Travel time index $TTI = \frac{x}{x_{FFS}}$
7. Travel time standard deviation $\sigma = \sqrt{\frac{(\sum_{i=1}^{N} x_i - \bar{x})^2}{N-1}}$
8. Travel time coefficient of variation $CV = \frac{\sigma}{\bar{x}}$
9. Travel time range $Range = x_{max} - x_{min}$
10. Ratio of mean travel time to median travel time $r = \frac{\bar{x}}{\hat{x}}$

    where $\hat{x}$ - median travel time
11. Total segment delay $TSD = (tp_{95\%} - x_{FFS}) \times V$ (vehicles-minutes)

    where V – volume of vehicles at the segment
12. Congested travel $CT = \sum ConLength \times V$ (vehicles-miles)

    where $ConLength$ – congested segment length
13. Congested roadway $CR = \sum ConLength$ (miles)

A few studies computed the average travel cost along with FPMs for considered highway corridors. Ando & Taniguchi (Ando, et al., 2006) estimated the total cost of link TT uncertainty and penalties due to early arrival/delayed arrival to customers, requesting a particular time window. Wheeler & Figliozzi (Wheeler, et al., 2011) and Figliozzi et al. (Figliozzi, et al., 2011) included TT, cost of traveling, and TT variability into the cost function. Several researches also assessed environmental impacts and emissions, produced by vehicles. Emissions were estimated based on the vehicle travel distance and the vehicle TS (see Ando & Taniguchi, 2006; Wheeler & Figliozzi, 2011).

3. DATA DESCRIPTION

In this chapter we present the data collection and methodology used to analyze data in the CFIRE region.

3.1 General Statistics by States

The GPS data used in this study is provided by ATRI and consists of trucks traveling within three states: Tennessee, Mississippi and Alabama. This particular dataset description provides information on the data lying in the State of Tennessee. A total of 55,798,200 truck observations were provided in the database with 835,183 unique trucks and include truck GPS data for 24 weeks for the years 2011-2014. The 24 weeks consists of two weeks for months of March, July and October for four years. Observations are not equally distributed between these 4 years. The majority of observations (34.62%) is for 2014 (19,316,100 GPS points). Years 2013, 2012, and 2011 have 14,782,100, 12,715,700 and 8,984,300 observations respectively. Figure 10 shows the percentages of data for individual years.

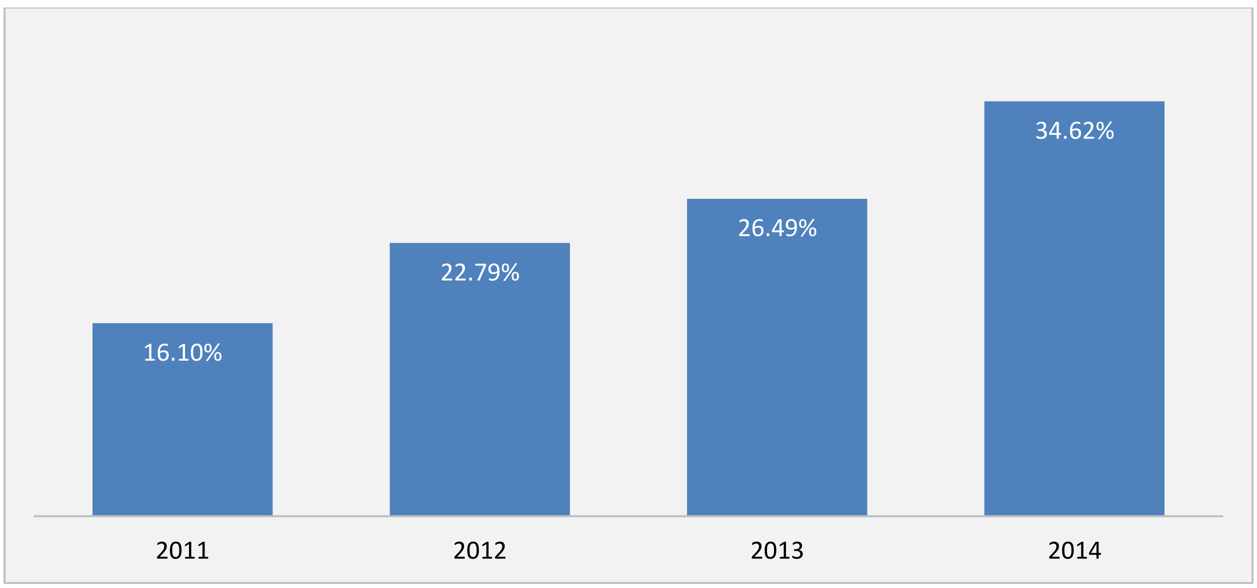


Figure 10 Percentage of Observations per Year

Further analysis revealed that observations were available for 3 months of each year: March, June and October with the latter having the most observations (19,959,700) followed by June (18,748,000 observations) and March (17,090,500 observations). The percentage split is presented in Figure 11. Data distribution by day of the week is shown in Figure 12. The majority (52.23%) of truck GPS data are observed between Tuesday and Thursday. The least number of observations are on Sunday (3,898,400) and Saturday (5,757,100) while Monday with 8,192,300 observations accounts for 14.68% of the data and Friday with 8,807,400 observations for 15.78%.

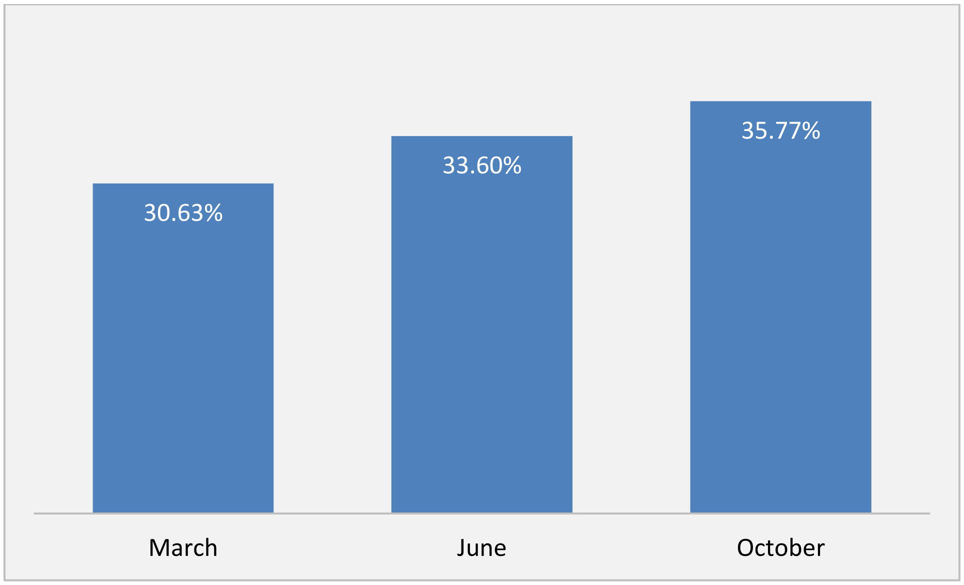


Figure 11 Percentage of Observations per Month

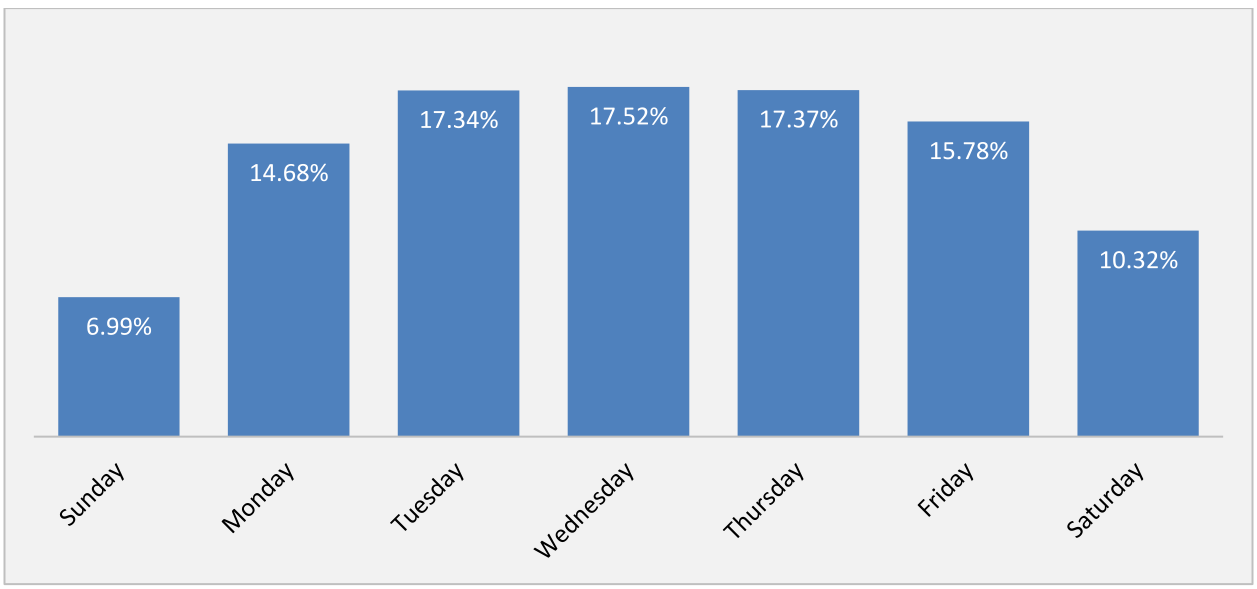


Figure 12 Percentage of Observations per Day of the Week

According to Memphis MPO a day can be divided in four different time periods based on the level of congestion. The AM peak period lasts from 6 am to 9 am, Midday period from 9 am to 2 pm, PM peak from 2 pm to 6 pm, and Off-peak from 12am to 6am and from 6 pm to 12pm. Figure 13 presents the data distribution over these peak periods. As expected Off-peak (20,728,600 observations) and Midday (15,765,600 observations) have the most truck observations. AM peak and PM peak data account for 7,697,700 and 11,606,300 of observations respectively.

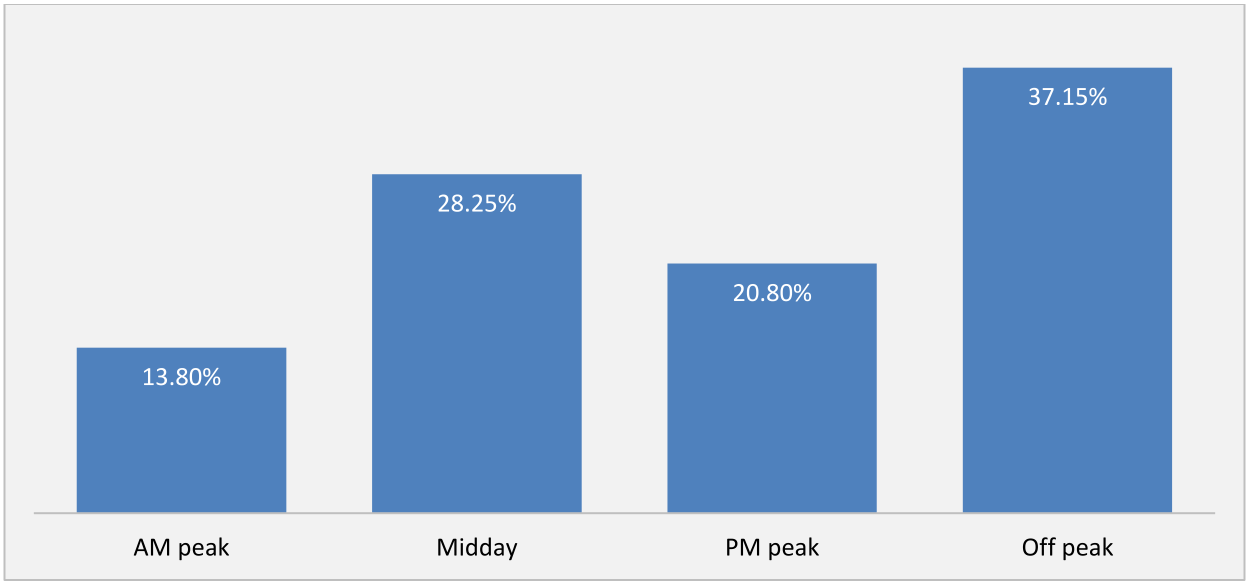


Figure 13 Percentage of Observations per Time of Day

The descriptive statistics presented so far can be obtained from the information provided by the Probe Vehicle GPS Data Processing Tool. Further analysis can be conducted using any database management system to retrieve information and descriptive statistics. In this study PostgreSQL was used to further process the data and conduct the statistical analysis. The total number of observations per truck varied as the provided dataset was a random sample. The maximum number of GPS records for a single truck in a day is 3611.The distribution of daily observations per single truck is shown in Figure 14. The majority of trucks had less than 30 observations in a day.

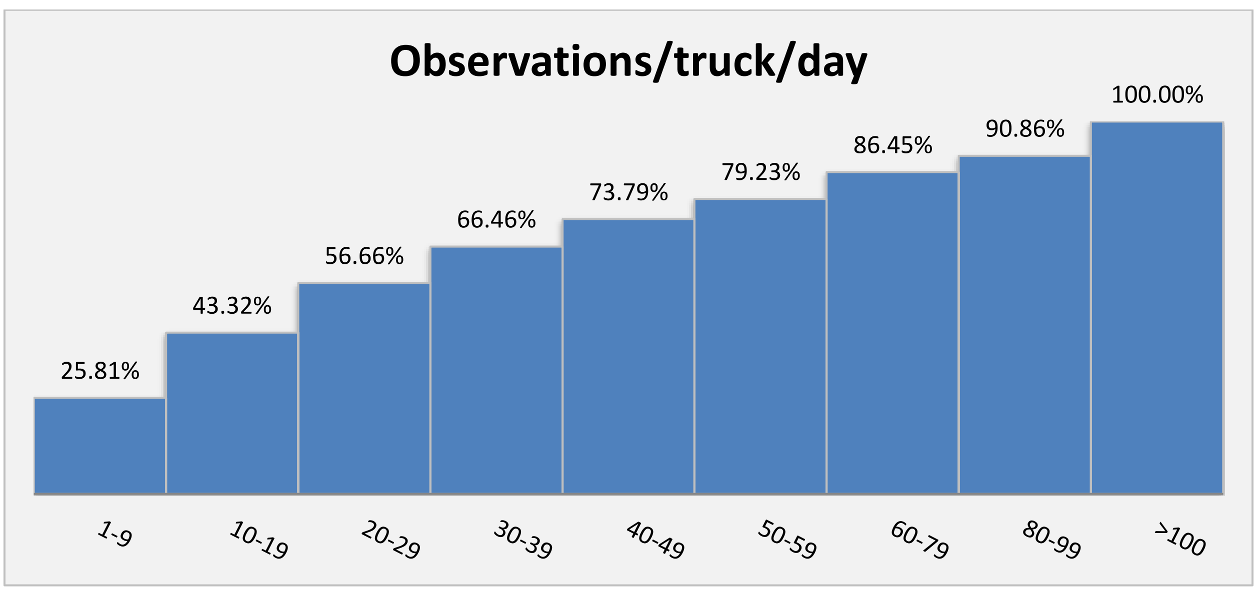


Figure 14 Number of Daily Observations for a Single Truck

The frequency of the transmitted GPS signal was not fixed, while a great percentage of the observations was transmitted from stopped trucks. In this study a truck was assumed stopped, if its spot speed was less than 5 mph. On an hourly basis 98.5% of the trucks had up to 19 observations (Figure 15), which translates to one observation per 3 minutes (stopped trucks included), while approximately 49% had less than 4 observations (i.e., one every fifteen minutes). It is also notable that a significant percentage (23.36%) of trucks had only one observation per hour.

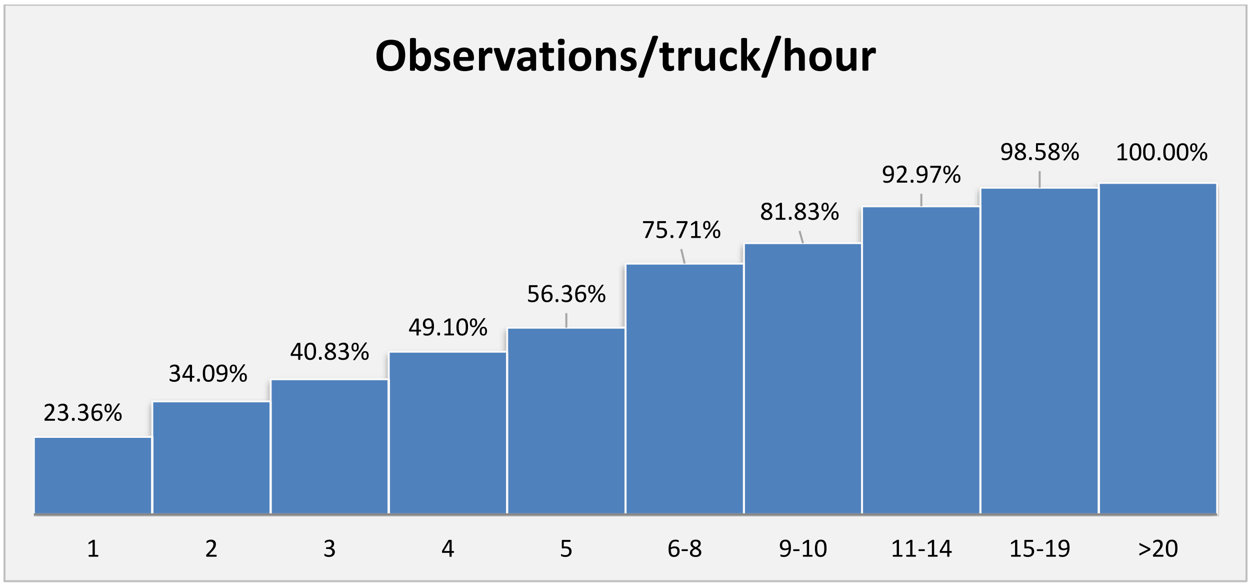


Figure 15 Number of Hourly Observations for a Single Truck

The given data was analyzed for four time periods: i) AM Peak: 6am – 9am, ii) Midday Peak (MD): 9am – 2pm, iii) PM Peak: 2pm – 6pm, and iv) Off-peak (OP): 6pm – 6am. Truck distribution by day of the week and time of the day is presented in Figure 16. The majority of observations were obtained for the OP time period with Midday being the next largest (in means of truck observations) time period. The smallest amount of GPS records were transmitted during the AM peak hours, as it is also the time period with the least duration (3 hours). About 20-22% of observations were observed in the PM peak period.

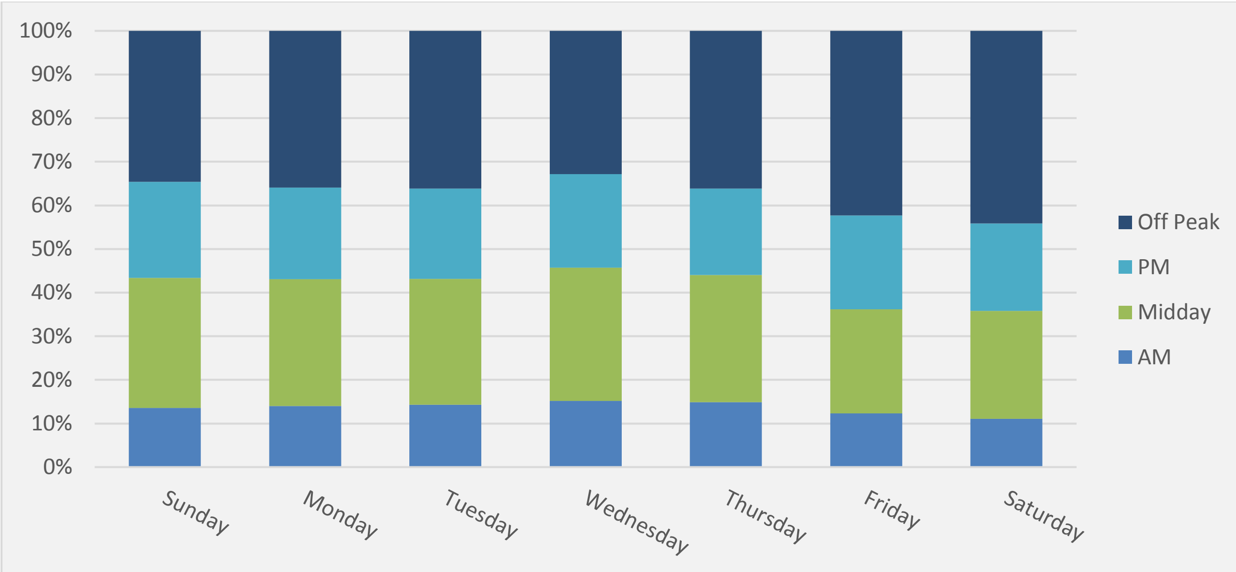


Figure 16 Truck Distribution by Day of the Week and Time of the Day

Figure 17 illustrates the percentage of stopped trucks in the four time periods. The largest percentage of stopped trucks was observed during AM Peak period (36.15%) with Off Peak being next with 35.52% of trucks being stopped. Midday and PM peak periods had the smallest percentages, 32.16% and 31.08% respectively.

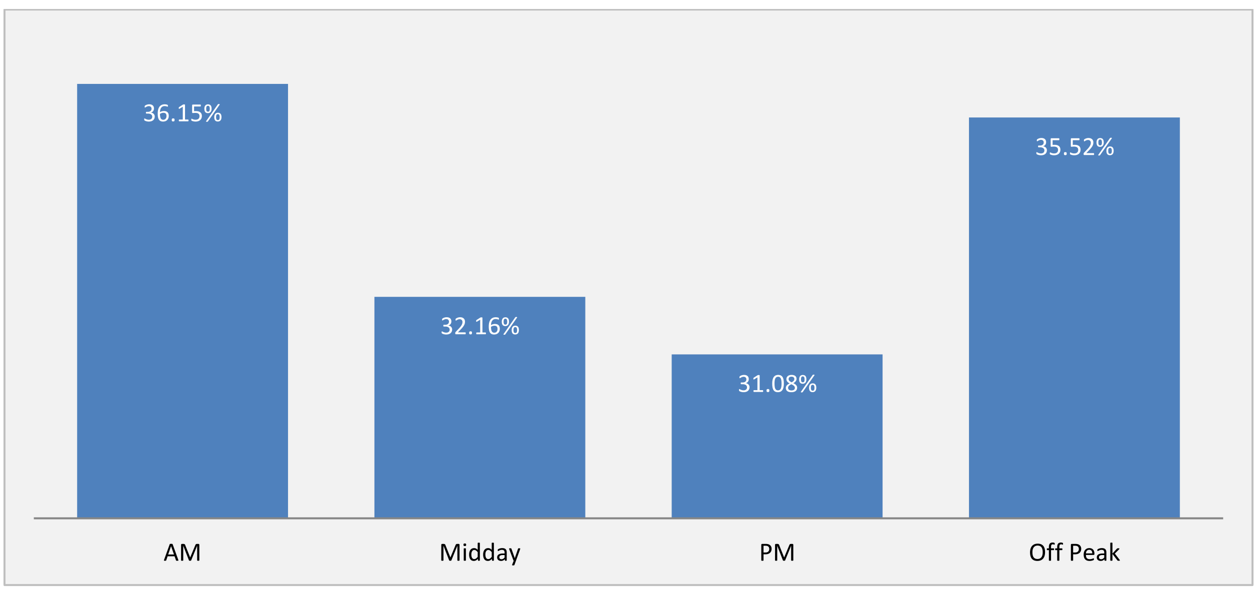


Figure 17 Percentage of Stopped Trucks per Time of the Day

Figure 18 shows that the largest percentage of stopped trucks in a day was observed during the OP time period for each weekday as expected since this period is the largest in a day with the majority of trucks belonging there. This percentage is even larger during Friday and Saturday (46.66% and 48.78%). On the contrary the percentage of stopped trucks for the remaining time periods is decreased these two days compared to Sunday-Thursday percentages.

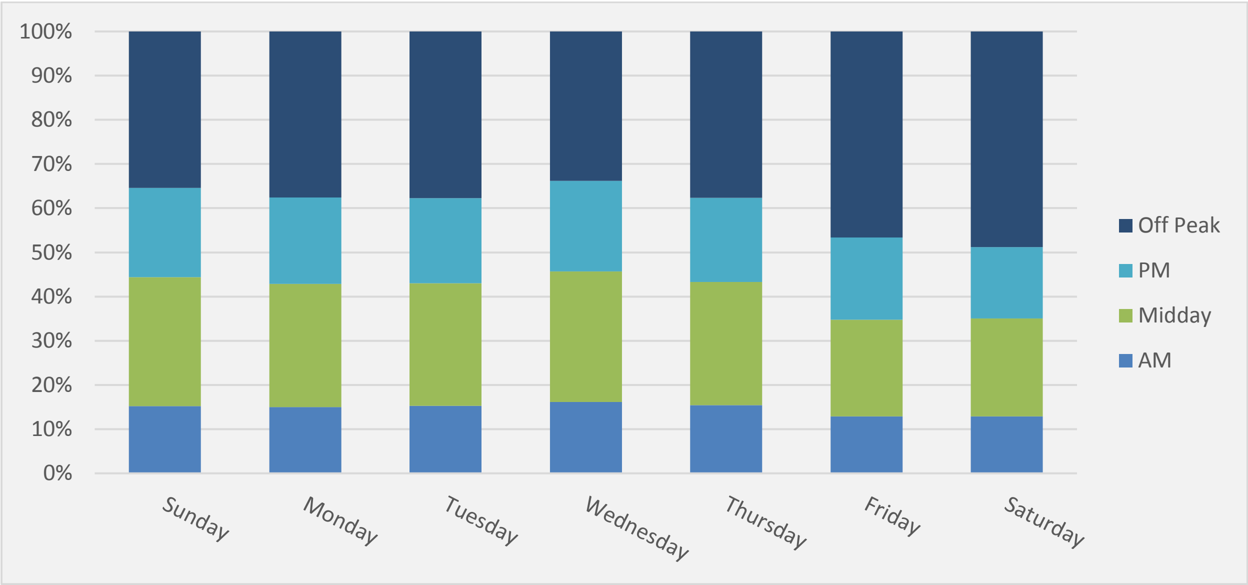


Figure 18 Stopped Truck Distribution by Day of the Week and Time of the Day

3.2 Dataset description

A sample is presented in Figure 19 and used to describe the available data. The following information was provided for each GPS record:

- GPS waypoint (X and Y coordinates)
- Time stamp
- Heading
- Spot speed
- Truck Identifier

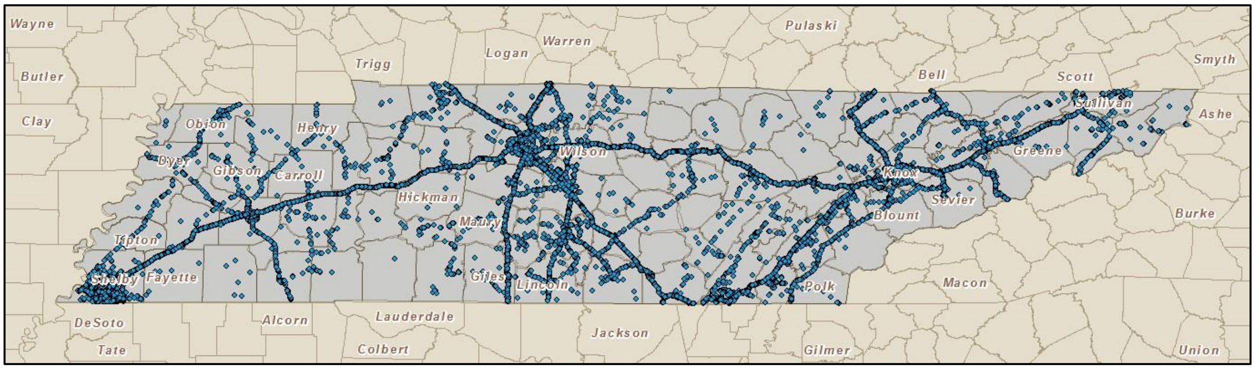


Figure 19 Random Day GPS Data Display

Time stamps were given for Coordinated Universal Time (UTC) zone. The State of TN lies in two time zones: Central Daylight Time (CDT) zone and Eastern Daylight Time (EDT) zone. The local time should be estimated for each GPS point in order to conduct the analysis for specific time periods. The Extract Analysis Toolbox, of ESRI ArcGIS 10.0[4], was used to assign a time zone to each observation based on its spatial disposition (see Figure 20). Once a time zone was determined for a GPS record, a local time was computed based on the difference between the given time zone and UTC zone. A daylight saving time for the year 2012 was considered as well. EDT zone was 4 hours behind UTC, while CDT zone was 5 hours behind UTC between March 11, 2012 and November 4, 2012[5]. For the rest of the year EDT zone was 5 hours behind UTC, while CDT zone was 6 hours behind UTC.

Figure 20 CDT and EDT Zones in TN

One of eight possible headings was recorded for each observation: E, W, N, NE, NW, SE, and SW. A unique identifier was assigned to each truck as most trucking companies are not willing to share any information regarding their vehicles and type of commodity transported (Greaves & Figliozzi, 2008; McCormack et al., 2011; McCormack & Zhao, 2011).

[4] www.esri.com

[5] http://www.timeanddate.com/time/dst/2012.html

4. DATA PROCESSING METHODOLOGY

4.1 Preprocessing Tool

The GPS Data Processing and Extracting Tool, a standalone application, was developed to import and process raw truck GPS data, and extract data for certain time periods specified by the user. In this tool, the time stamp for each truck was changed to the local time based on the time zone (the GPS timestamp was in Coordinated Universal Time - UTC). The tool also filters data based on user input and exports processed data into shape file(s) and .csv file(s). The fields of time zones and local date/time are also appended to the output file(s). The tool accepts .csv and .txt files as input files. The tool can produce two output types of files: shape files and csv files.

4.2 Associating GPS Records with Network/Zone

In order to associate (or snap) GPS points on the network, the Proximity Analysis Toolbox, of ESRI ArcGIS 10.0, was used. In this study the Freight Analysis Framework (FAF) transportation network for the State of TN was evaluated. The FAF network includes 3,393 road segments with average link length of 2.66 miles. Since truck GPS data did not include any information on the accuracy of the GPS devices, the worst case scenario of a quarter mile (as reported in the literature, see Jones et al., 2005; Schofield & Harrison, 2007), was assumed. In theory, the search radius for snapping observations should be equal to sum of the device spatial error and the positional error of the used network. In FAF network this can be up to ±260 feet (FHWA, 2014). GPS records lying outside the search radius were discarded.

4.3 Direction and Outlier Identification Algorithm (DOI)

DOI algorithm was developed to address the issue of multiple directions of GPS truck records, associated with the same link. Figure 21 illustrates this issue with 17 observations, snapped to link, having a total of six unique headings: E, N, NE, SW, SE, and W. These GPS records should be separated in two groups: 1) trucks moving from the link start point (with coordinates [x_{st}, y_{st}]) to the link end point (with coordinates [x_{end}, y_{end}]), and 2) trucks moving from the link end point to the link start point. Based on the link's geometry those groups should be either NE or SW directions respectively. The major steps of DOI are as follows:

DOI Steps

Step 1: Load GPS data for a given day/time period

Step 2: Associate each GPS record with a link (based on a predefined radius around each record)

Step 3: Remove outliers[6] based on speed (if speed threshold is known)

Step 4: For each link

Step 4.1: Identify the number of unique truck headings

Step 4.2: Separate observations in two groups based on the link spatial disposition (see Figure 22)

Step 4.3: Remove additional outliers based on the Chauvenet's criterion (optional)

Next we present a small example to showcase how DOI is implemented.

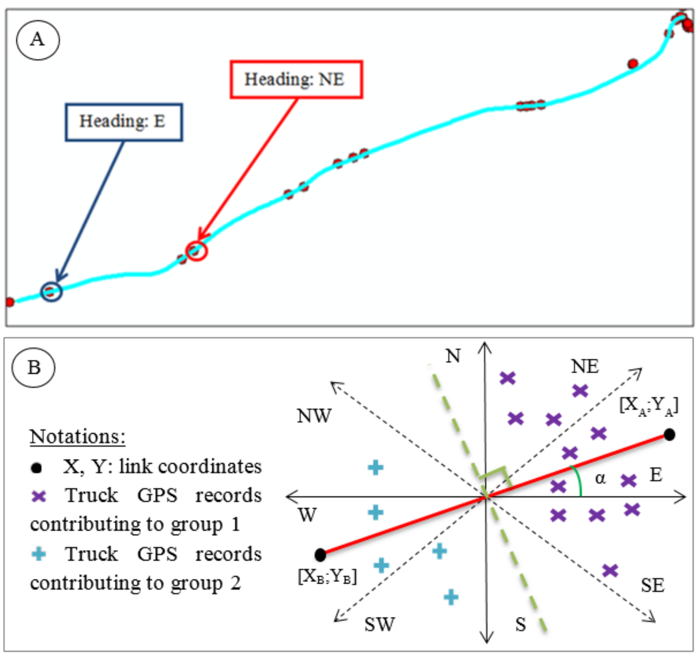


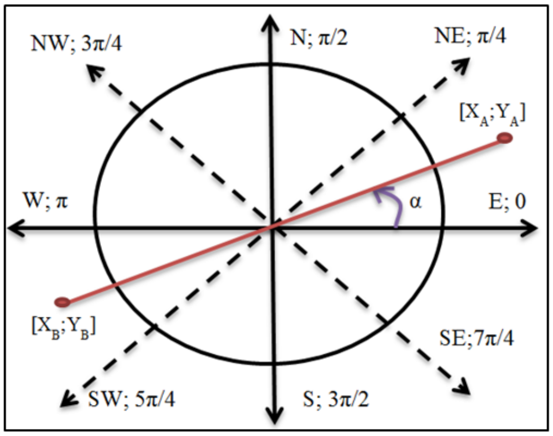
Figure 21 DOI for Resolving the Problem with Headings

4.3.1 DOI Example

Figure 9B provides an example of step 4.2. for a fictitious link. First, the start and end point coordinates for the given link are calculated. The link is then approximated by a straight line, connecting the start and end points. The next step calculates the angle (α), between the E-W axis and the straight line representing the link. The value of α can be estimated using line coordinates and trigonometric functions (e.g., arccosine, arcsine, arctangent, etc.). In the given example (see Figure 22) angle α lies between 0 and π/4, hence trucks with headings E, N, NE or SE will be assigned to the direction from B to A (BA) and trucks with headings W, S, SW or NW to the direction from A to B (AB). Groups

[6] Observations with spot speeds less than 5 mph are considered as outliers.

of headings, contributing to BA and AB directions, for every possible angle *α* are presented in Figure 22 DOI Heading Assignment.



Angle, α	Headings assigned by DOI	
	BA	AB
$0<\alpha<\pi/4$	E, NE, N, SE	W, SW, S, NW
$\pi/4<\alpha<\pi/2$	E, NE, N, NW	W, SW, S, SE
$\pi/2<\alpha<3\pi/4$	N, NW, W, NE	S, SE, E, SW
$3\pi/4<\alpha<\pi$	N, NW, W, SW	S, SE, E, NE
$\pi<\alpha<5\pi/4$	S, SW, W, NW	N, NE, E, SE
$5\pi/4<\alpha<3\pi/2$	S, SW, W, SE	N, NE, E, NW
$3\pi/2<\alpha<7\pi/4$	S, SE, E, SW	N, NW, W, NE
$7\pi/4<\alpha<2\pi$	S, SE, E, NE	N, NW, W, SW

Figure 22 DOI Heading Assignment

4.3.2 Outlier Detection: Chauvenet's Criterion

Detection and removal of outlier GPS truck records is a crucial component of the analysis if accurate FPMs are to be calculated. Removal of outliers based on predetermined thresholds (e.g., 10 mph) may result in high misclassification of records during different time periods of the day (e.g., 10 mph may not be an outlier for peak periods). To escape the use of predetermined speed thresholds the Chauvenet's criterion was adopted (Chauvenet, 1960). The criterion assumes that speeds follow a Normal Distribution, and observations are considered as outliers, if the probability of obtaining their deviation from the mean is less than $1/(2N)$, where N is the number of observations.

4.3.3 FPM Calculation

Once GPS records are associated with links, direction of truck movement has been assigned, and outliers have been detected and removed, preferred FPMs can be calculated using DOI. The list of FPMs, calculated in this study, includes TS (in each direction), TT, and TT reliability measures (90_{th} percentile TT, 95_{th} percentile TT, buffer TT or BTT, BTT index or BI, TT standard deviation or TTSD, TTCV, TT range, mean to median TT ratio). Average TS were computed based on spot speeds available from GPS truck data. This approach was chosen as most of consecutive GPS points for a given truck belong (for the majority of the trucks) to different links (i.e., link length and the mean time interval between observations cannot be used to calculate average TS). Once FPMs are calculated for all links, it will be possible to identify areas, where bottlenecks occur for a given time period.

4.3.4 DOI Validation

DOI was validated on the FAF network with LTS obtained from the FPMweb Tool. Data for the I-40 section in TN was retrieved from the FPMweb Tool for 36 days (3 consecutive weekdays for each month of 2012). Average LTS over 3 days of each month were computed for four time periods (see section 3.1.1). Then average LTS were estimated using DOI for the same links and time periods. Results of a comparative analysis indicated that the differences between LTS, provided by the FPMweb Tool, with the ones, calculated by DOI, were not significant (less than 5% on average). Differences were mostly observed on short links (< 3 mi) and could be possibly caused by snapping errors. Note that DOI can be applied to any network (not only FAF), and its accuracy will depend on length and shape of each roadway segment.

4.4 Origin-Destination Identification Algorithm (ODIA)

ODIA was developed to estimate the number of truck trips between traffic analysis zones (TAZs) in the State of TN. Along with truck trips additional information can be retrieved (e.g., start trip time, end trip time, trip duration, etc.). Once GPS records are loaded, ODIA filters out observations with spot speeds greater than a set value (=5 mph), and leaves for analysis only those observations (with spot speeds ≤5 mph), which can be potentially either origins or destinations. Then the algorithm sorts all trucks by IDs and observations for each truck by their time stamps in the ascending order. Next ODIA starts an iterative process, which consists in checking TAZ for each observation of a given truck. If TAZ_p and TAZ_s[7] for two consecutive GPS records are the same, it is more likely that no trips were made by the truck. When two consecutive observations have different TAZ_p and TAZ_s, ODIA marks the preceding record as "ORIGIN", while for the succeeding record the algorithm checks if it is a genuine destination. If there is only one consecutive observation with TAZ_s, ODIA marks that observation as "DESTINATION". If there is a group of GPS records with the same TAZs as TAZ_s, the algorithm calculates the total travel distance between those observations. If the distance does not exceed ¼ mile (GPS spatial error), ODIA marks the earliest observation of this group as "DESTINATION". Otherwise (the distance >¼ mile), the truck was most probably still traveling (e.g., traffic light stop). Note that travel distance between two consecutive GPS points was computed based on their coordinates. The procedure continues until all observations for all trucks are analyzed. Final ODIA output also contains full Origin-Destination (OD) matrix. The main ODIA steps are outlined next.

<u>ODIA Steps</u>

Step 0: Initialize origin-destination matrix $OD = \emptyset$

Step 1: Load GPS data for a given day/time period

[7] TAZp denotes TAZ for a preceding observation; TAZs denotes TAZ for a succeeding observation

Step 2: Remove observations with spot speeds greater than a set value (=5 mph)

Step 3: Sort GPS data based on truck IDs and time stamps (in the ascending order)

Step 4.0: For each truck t set observation i=0

Step 4.1: Select observation i=i+1

Step 4.2: Does the next observation (i.e., i+1) have the same TAZ

If YES - go to Step 4.1

Else go to Step 4.3

Step 4.3: Flag observation i as “ORIGIN”, record trip start time

Step 4.4: Count the total number of observations j≥i+1 with the same TAZ as i+1 and denote it as Q

Step 4.5: Is Q greater than 1

If YES – go to Step 4.6

Else flag observation i+1 as “DESTINATION”, record trip end time, count trip (, TAZ_{i+1}) = $OD(TAZ_i , TAZ_{i+1})$ + 1 and go to Step 4.8

Step 4.6: Compute the total travel distance between consecutive observations i+1, i+2,…,i+Q and denote it as D

Step 4.7: Is D greater than a set value (=¼ mile)

If YES – go to Step 4.8

Else flag observation i+1 as “DESTINATION”, record trip end time, count trip (, TAZ_{i+1}) = $OD(TAZ_i , TAZ_{i+1})$ + 1 and go to Step 4.8

Step 4.8: Is i+Q the last observation for truck t

If YES – go to Step 4.9

Else go to Step 4.1

Step 4.9: It truck t the last

If YES – go to Step 5

Else go to Step 4.0 and set t=t+1

Step 5: Retrieve necessary truck trip data

4.5 Hours of Delay Methodology

The following methodology has been used to calculate the Hours of Delay (HOD). First, the Link Travel time (LTT) of trucks in each link is calculated for the three time periods (AM, MD and PM) using the following formulation:

$$LTT = \frac{Length\ of\ Link}{Link\ Speed}$$

Then the free flow travel time (FFTT) of trucks in each link is calculated using the following formulations:

$$FFTT = \frac{Length\ of\ Link}{Free\ flow\ Speed}$$

Once both the travel times are obtained, finally, the hours of delay are calculated by multiplying the truck volume on that particular link with the difference of the two speeds using the following formulation:

$$HOD = Volume \times (LTT - FFTT)$$

4.6 Congested Lane Miles Methodology

The congested lane mile (CLM) is defined as the total number of lane miles with congested travel. The following notations clearly defines the congested lane mile (CLM):

$$CLM = \begin{cases} \sum_{i=1}^{n} L_i \ \ \exists \ \ S(L_i) < FFS(L_i) \\ 0, \qquad otherwise \end{cases}$$

$$where,$$

$$CLM = Congested\ Lane\ Mile$$

$$L_i = Length\ of\ Congested\ Link\ i\ (in\ miles)$$

$$S(L_i) = Truck\ speed\ at\ link\ i$$

$$FFS(L_i) = Free\ Flow\ speed\ at\ link\ i$$

$$n = number\ of\ congested\ links$$

The following methodology has been used to calculate the congested lane mile (CLM). First, the average speed of trucks in each link has been identified for the three time periods (AM, MD and PM). Then the obtained link speed is compared with the free flow speed of trucks. In this study, the free flow speed is assumed to be the link speed during off-peak period which is from 6 PM to 6 AM. If the link speed is found to be greater than the free flow speed, then it is certain that the mentioned link is congested at that time period.

5. FREIGHT PERFORMANCE MEASUERS

The methodology described in the previous section (for estimating link based freight performance measures) was applied to the FAF network in the State of TN using truck GPS data (provided by ATRI) for selected weekdays of each month over the years of 2011 to 2014. The research team developed a geodatabase (that is available upon request) containing all the link-based performance indicators described in the methodology section. Figure 23 through Figure 44 show examples of maps that can be produced using the developed geodatabase. Next we present a case study that will showcase how the developed database can be used in practice.

5.1 Case Study

A case study was conducted for the State of TN using the truck GPS data, provided by ATRI, for Mondays-Wednesdays-Fridays of March 2014. The GPS data were retrieved using the GPS Data Processing and Extracting Tool. A total of 2,920,043 GPS records were available for 24,039 trucks. Next we present several examples of FPMs that can be produced using the developed "*GPS-based FPMs Estimation*" toolbox.

5.1.1 Link-based FPMs

The first analysis aimed to estimate link-based FPMs for the roadway segments in TN. The Freight Analysis Framework (FAF) was used as a transportation network. The FAF network has 3,393 roadway segments with the average link length of 2.66 miles. Two time periods were considered (*22, 28*): AM peak period (6 am – 9 am) and PM peak period (2 pm – 6 pm). Scripts "*LinkFPMs*", "*CLM*", and "*HOD*" of the developed ArcGIS toolbox were executed using the retrieved GPS data for the AM and PM peak periods to calculate the average travel speeds, congested lane miles, and hours of delay respectively. Results are presented in Figure 45 and Figure 46 for all roadway segments of the FAF network in TN. Link-based FPMs were not calculated for 32.4% and 26.9% of links for AM and PM peak periods respectively, as the GPS data were not available (those links are colored in black, see Figure 45). We observe that traveling speeds at the major TN freight corridors (i.e., I-24, I-40, I-65, and I-81) are greater than 60 mph and are close to FFTS. Furthermore, hours of delay for the most of links do not exceed 2.00 vehicles·hours. However, reduction in travel speeds and increase in hours of delay are observed at links in the vicinity of large metropolitan areas (i.e., Memphis, Nashville, Knoxville, and Chattanooga).

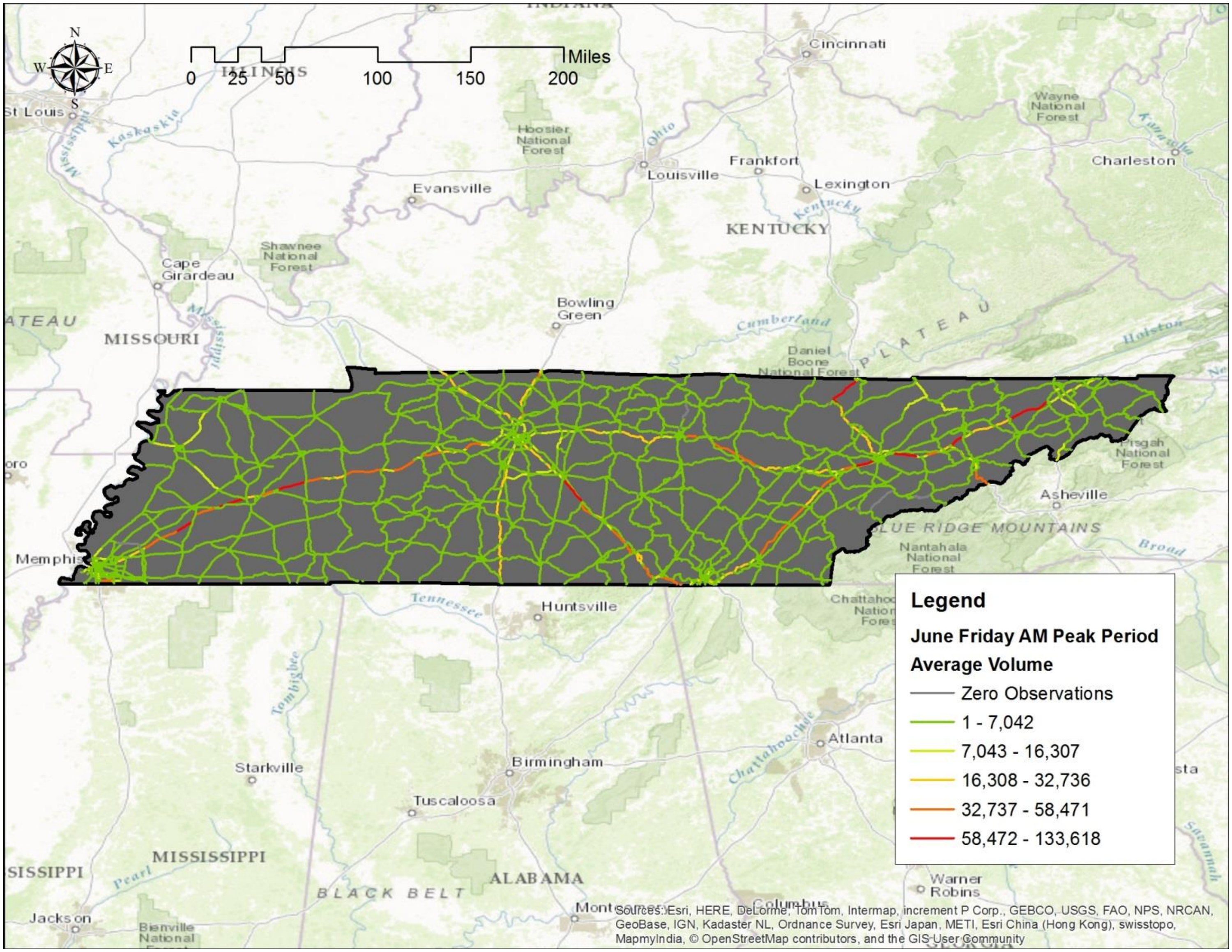


Figure 23 Average Truck Volume for AM Peak Period Friday June 2011-2014

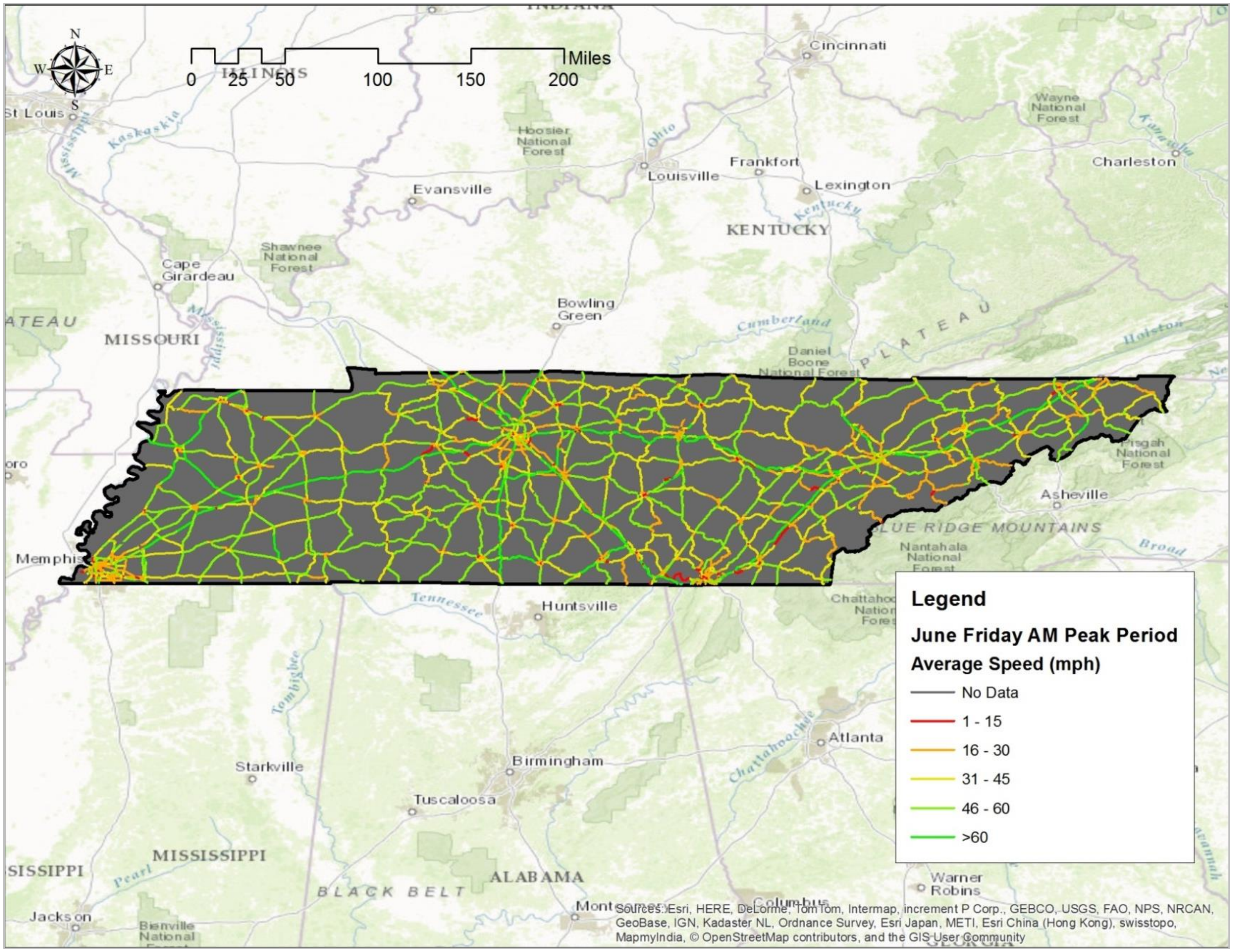


Figure 24 Average Speed for AM Peak Period Friday June 2011-2014

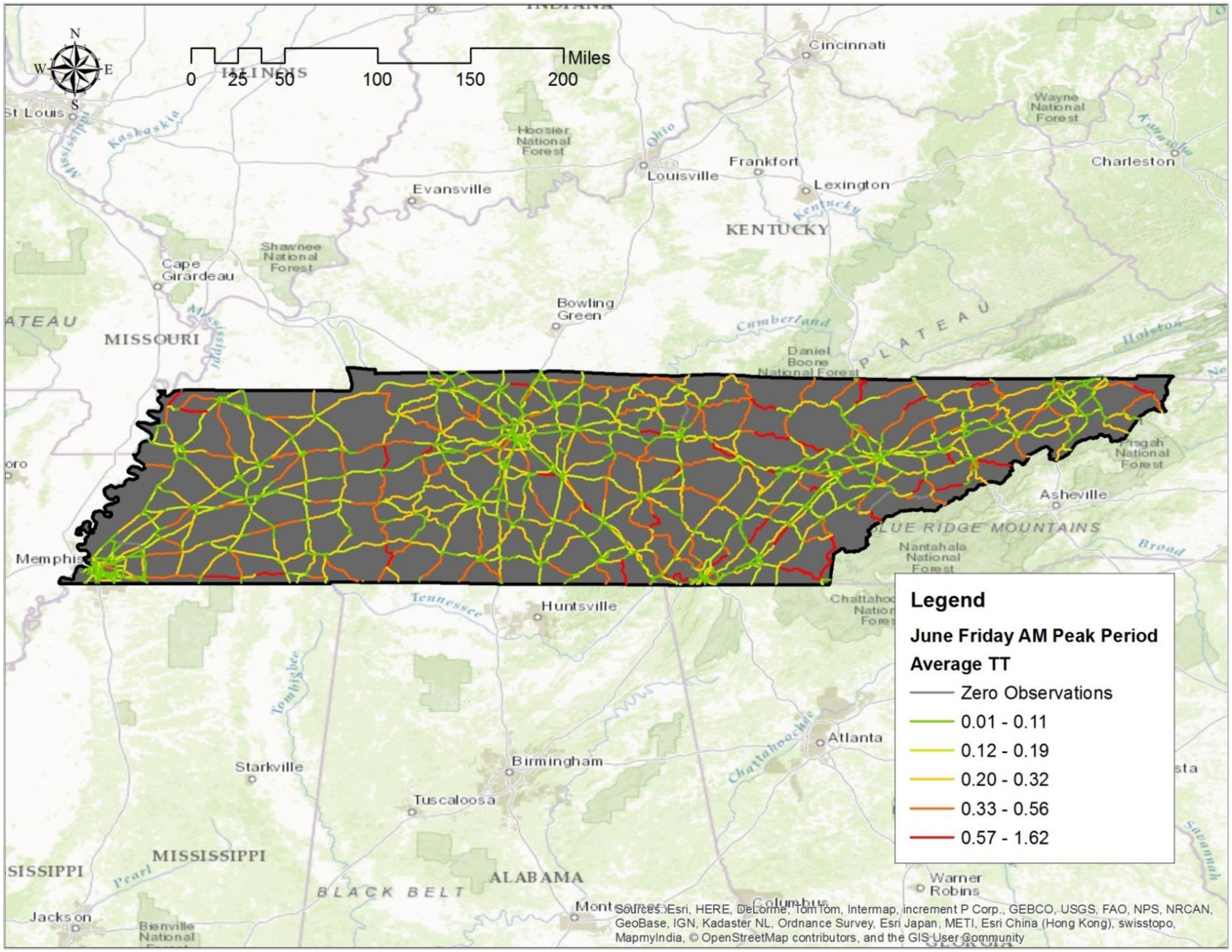


Figure 25 Average Travel Time for AM Peak Period Friday June 2011-2014

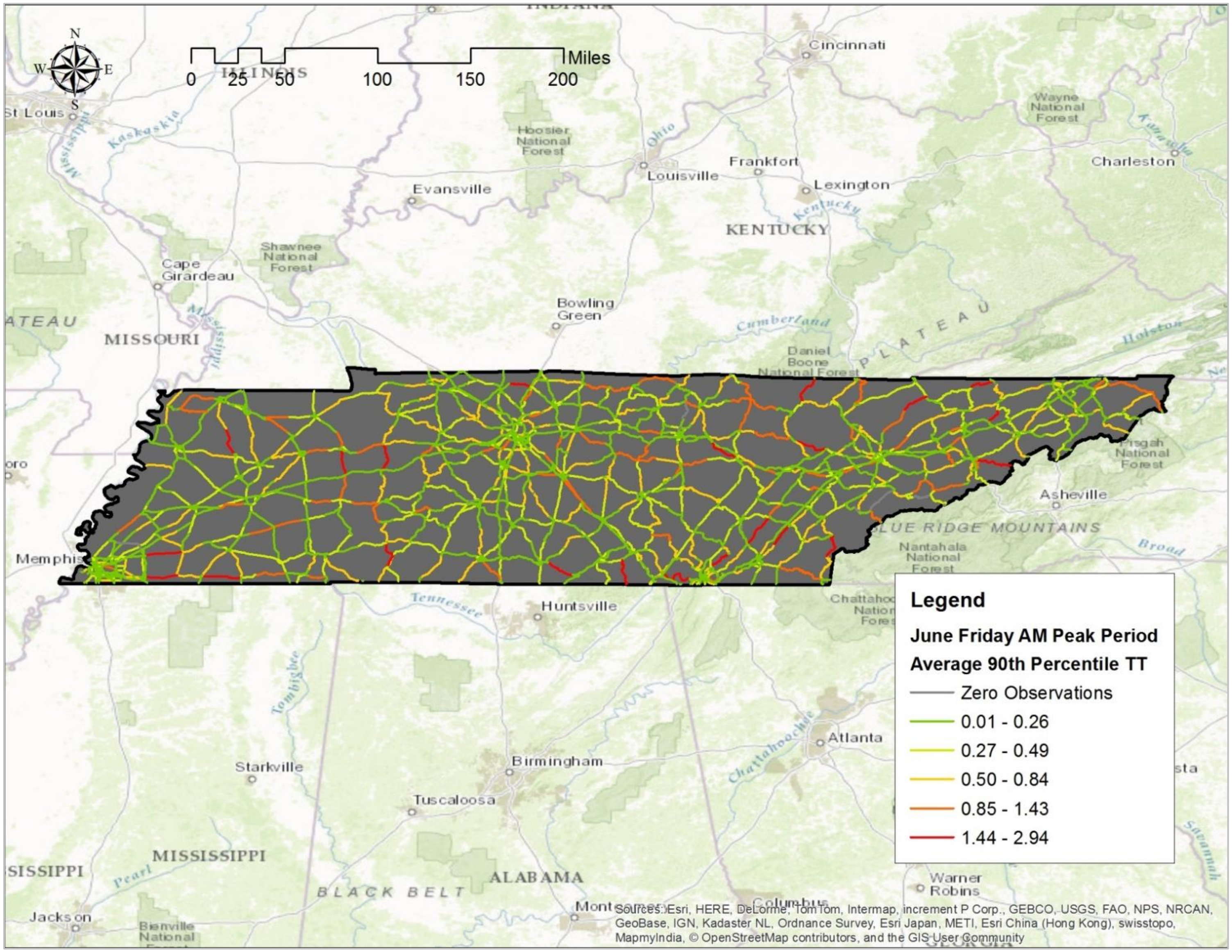


Figure 26 Average 90th Percentile Travel Time for AM Peak Period Friday June 2011-2014

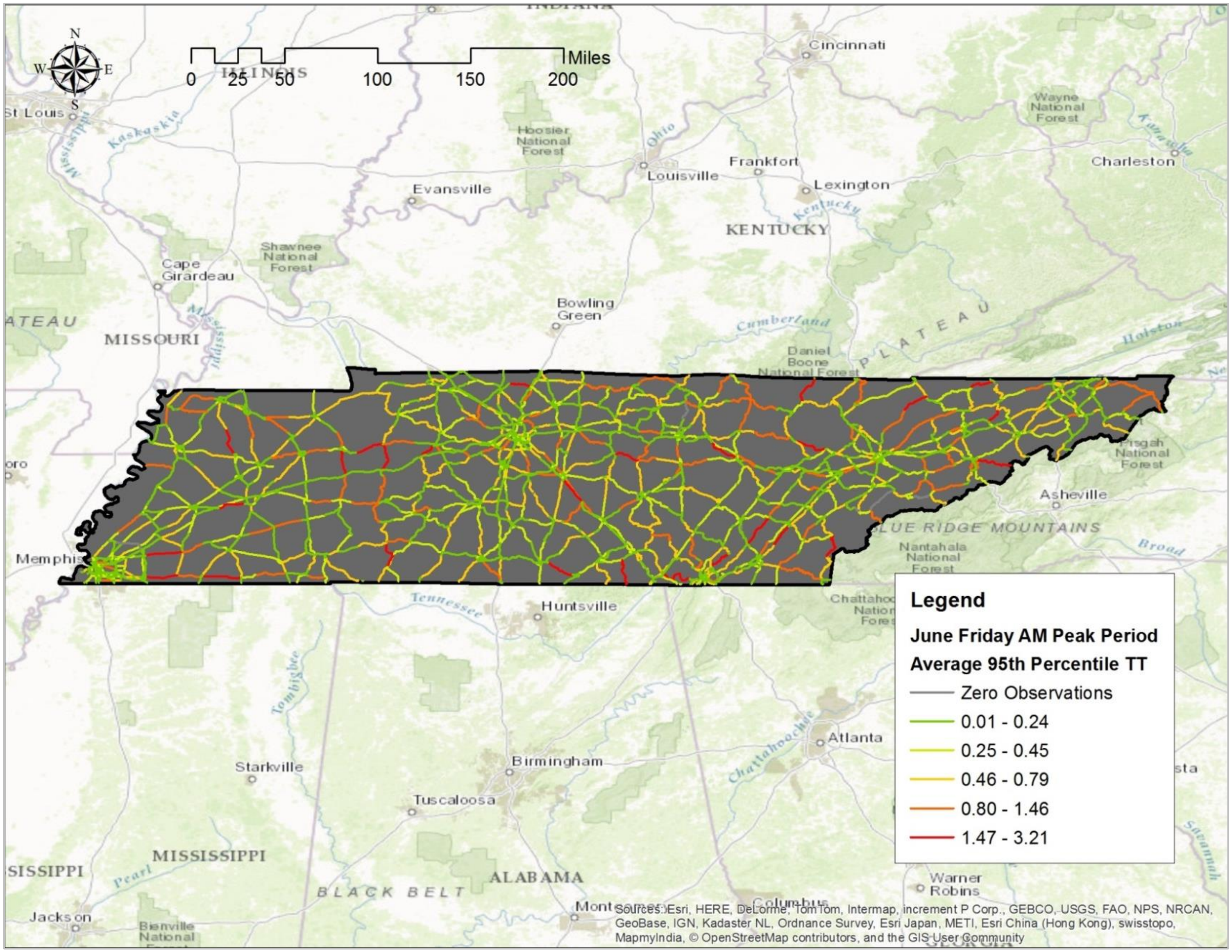


Figure 27 Average 95th Percentile Travel Time for AM Peak Period Friday June 2011-2014

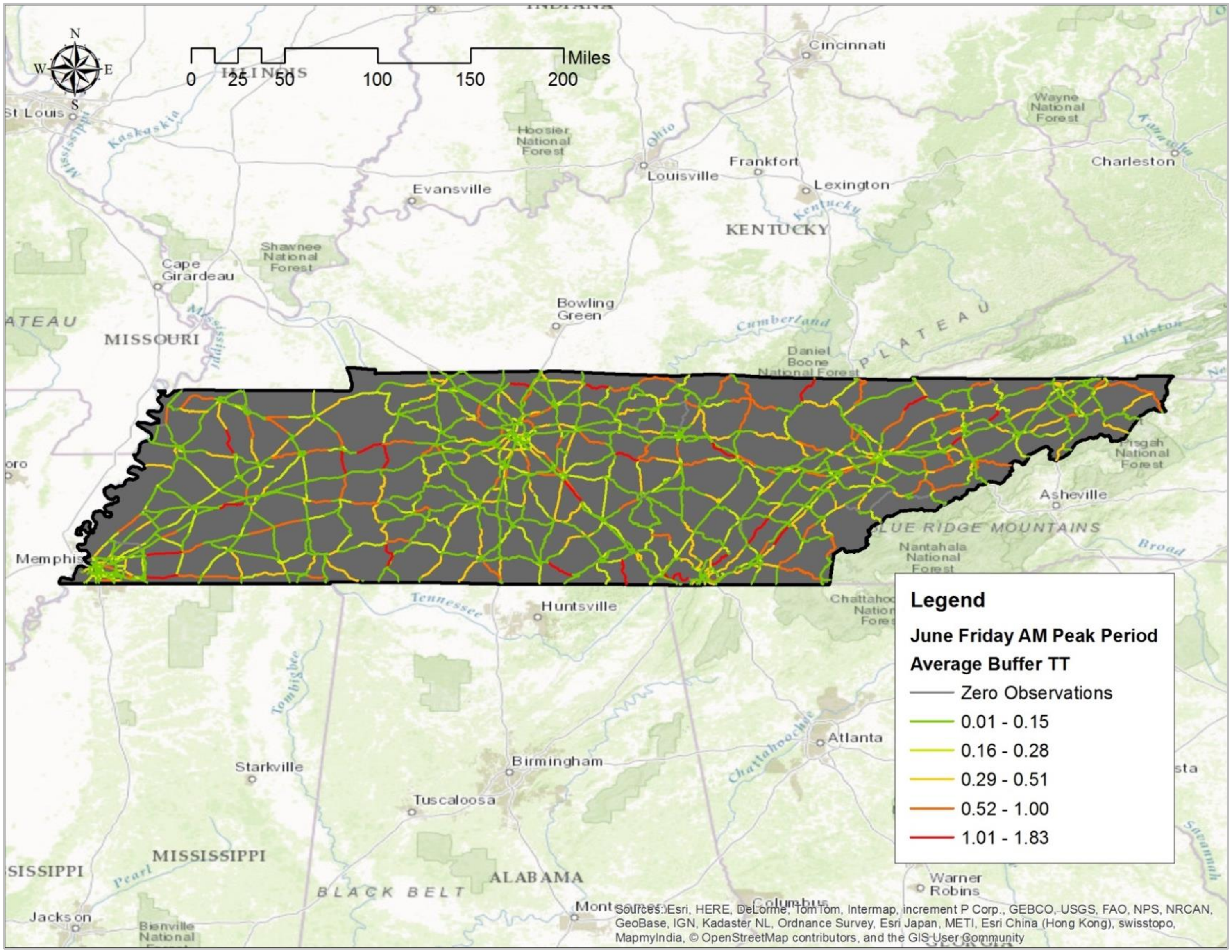


Figure 28 Average Buffer Travel Time for AM Peak Period Friday June 2011-2014

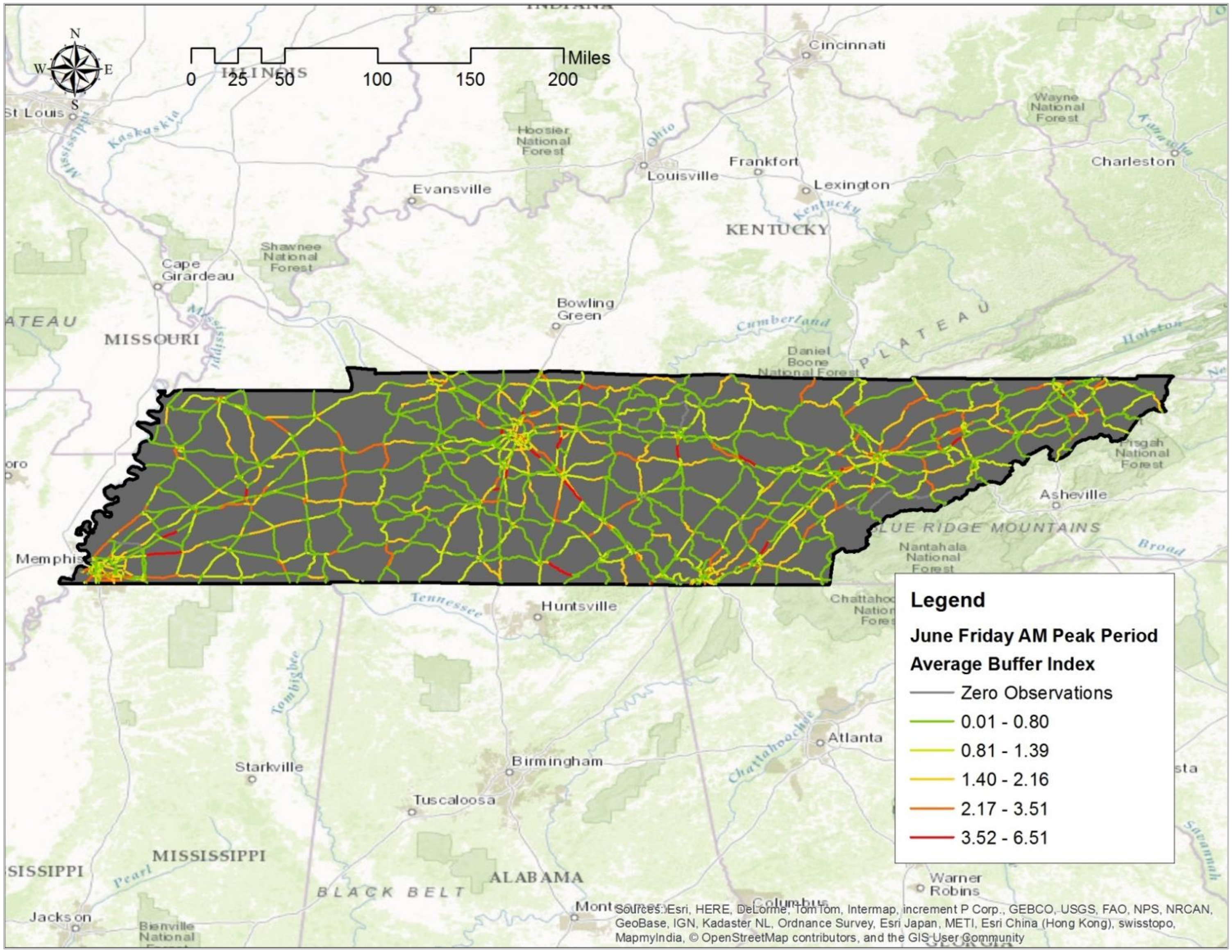


Figure 29 Average Buffer Index for AM Peak Period Friday June 2011-2014

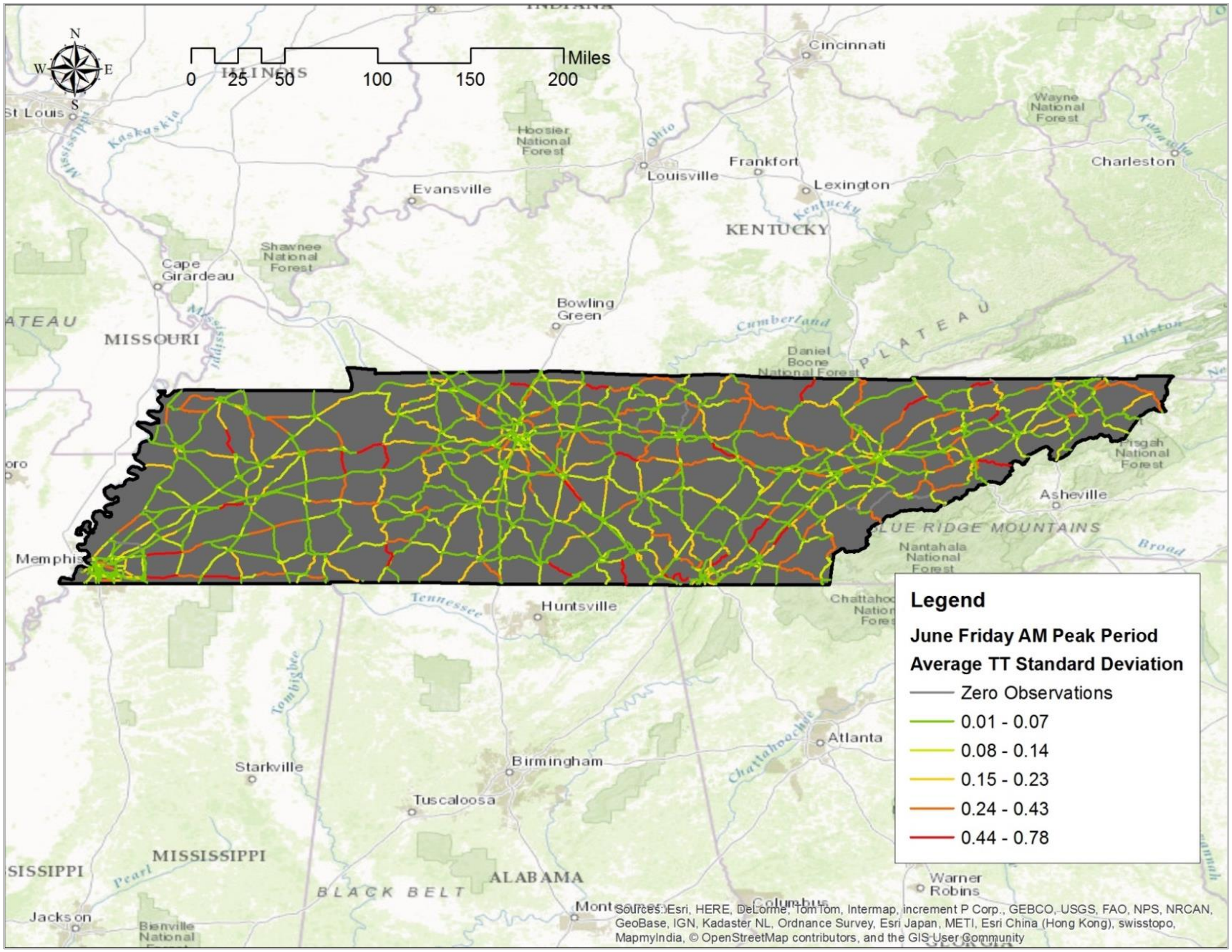


Figure 30 Average Travel Time Standard Deviation for AM Peak Period Friday June 2011-2014

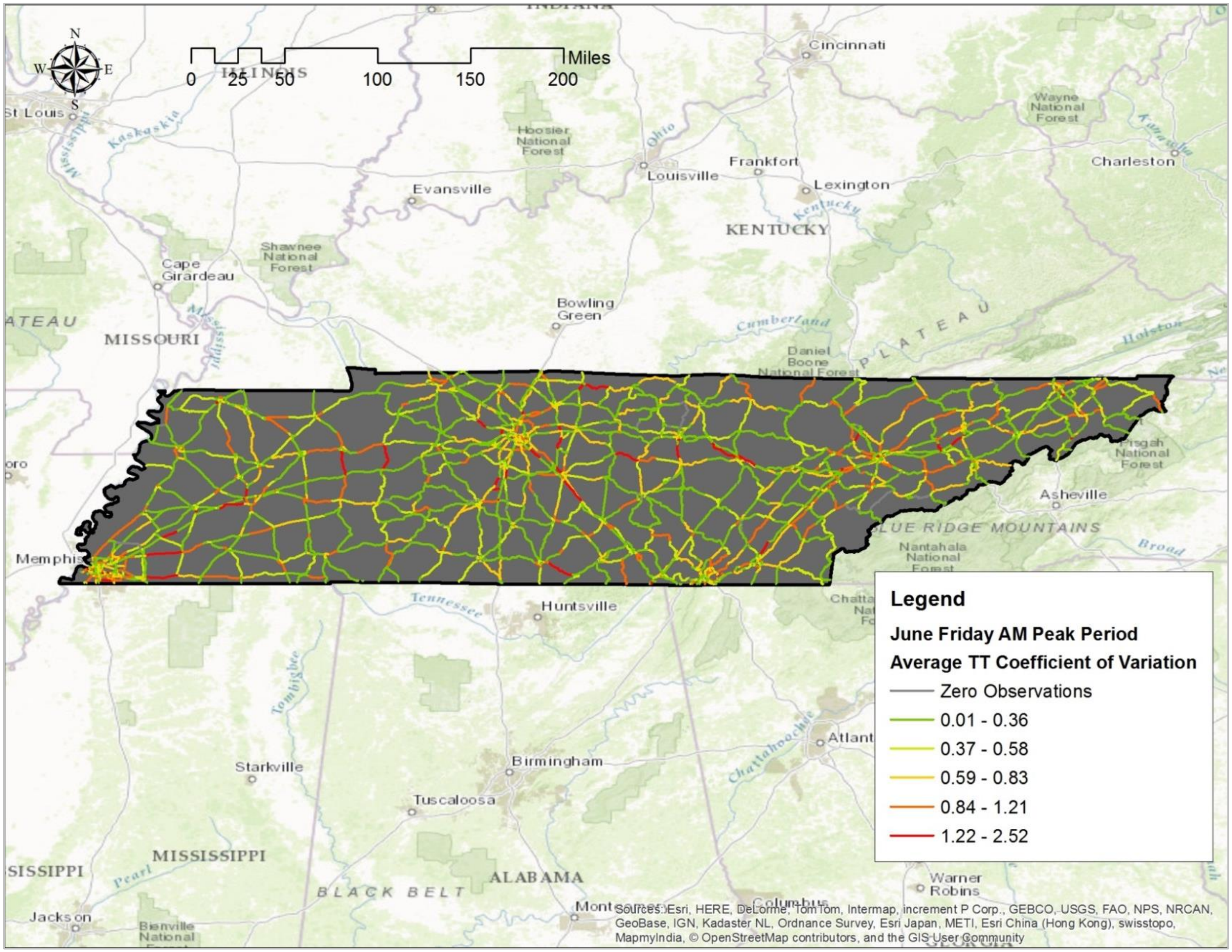


Figure 31 Average Travel Time Coefficient of Variation for AM Peak Period Friday June 2011-2014

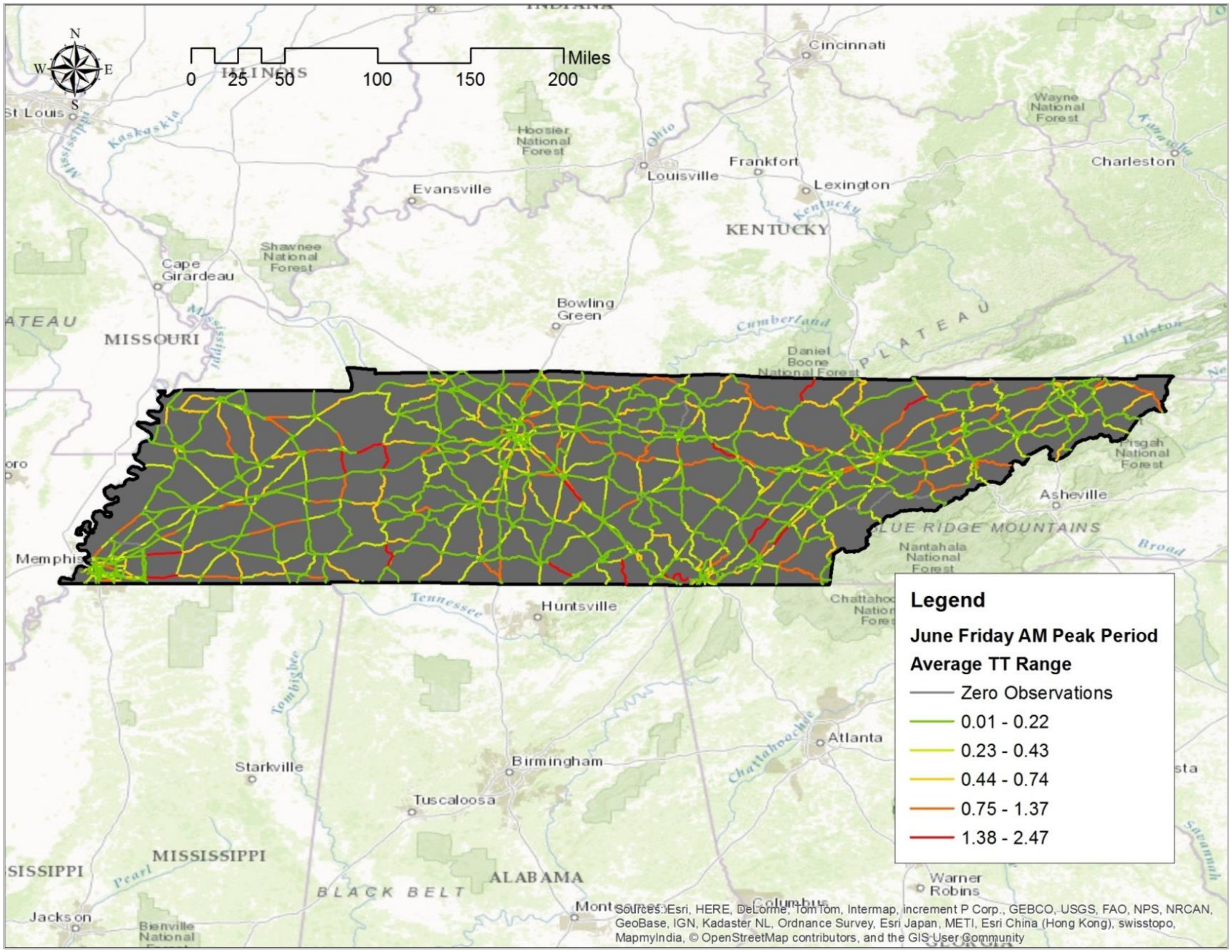


Figure 32 Average Travel Time Range for AM Peak Period Friday June 2011-2014

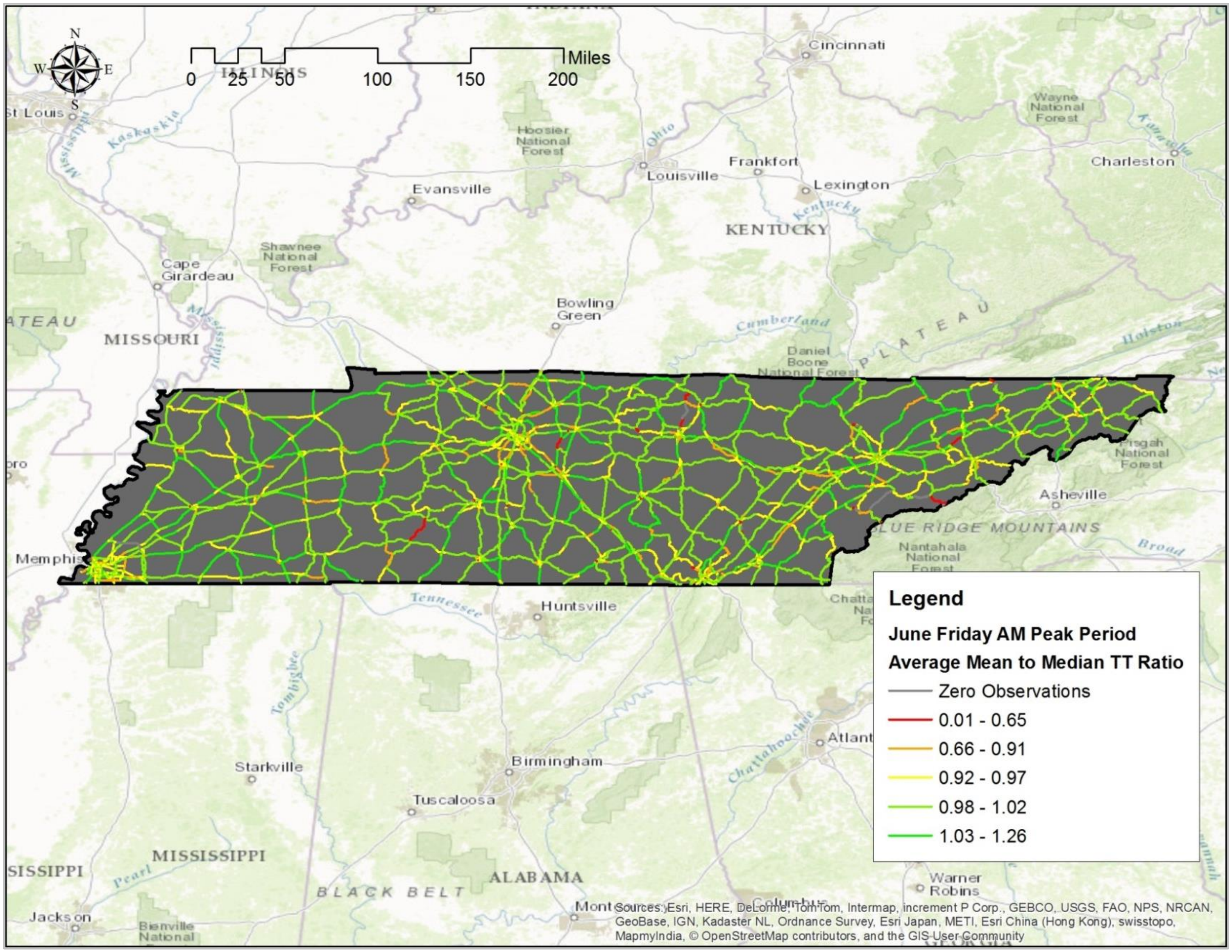


Figure 33 Average Mean to Median Travel Time Ratio for AM Peak Period Friday June 2011-2014

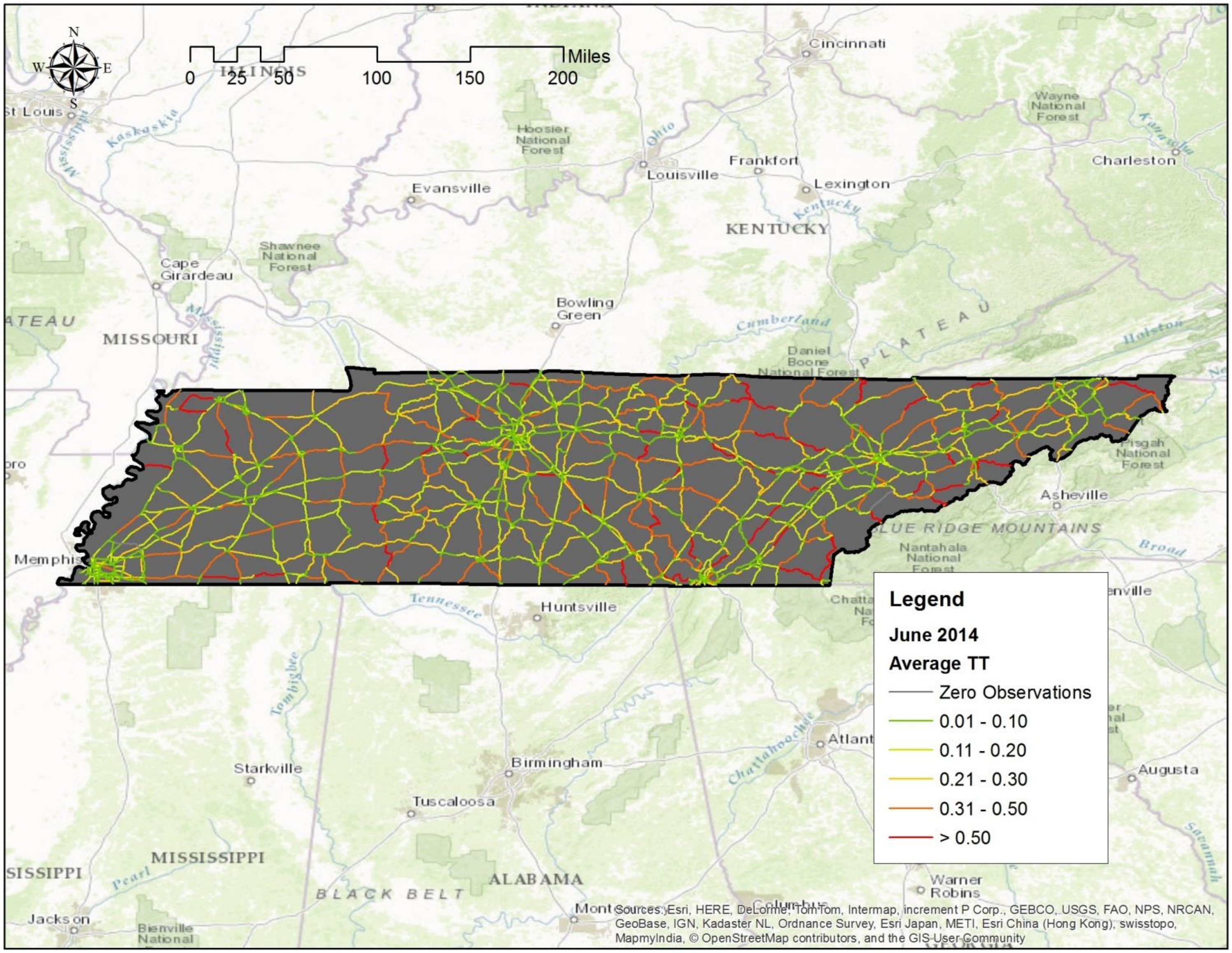


Figure 34 Average Travel Time for June 2014

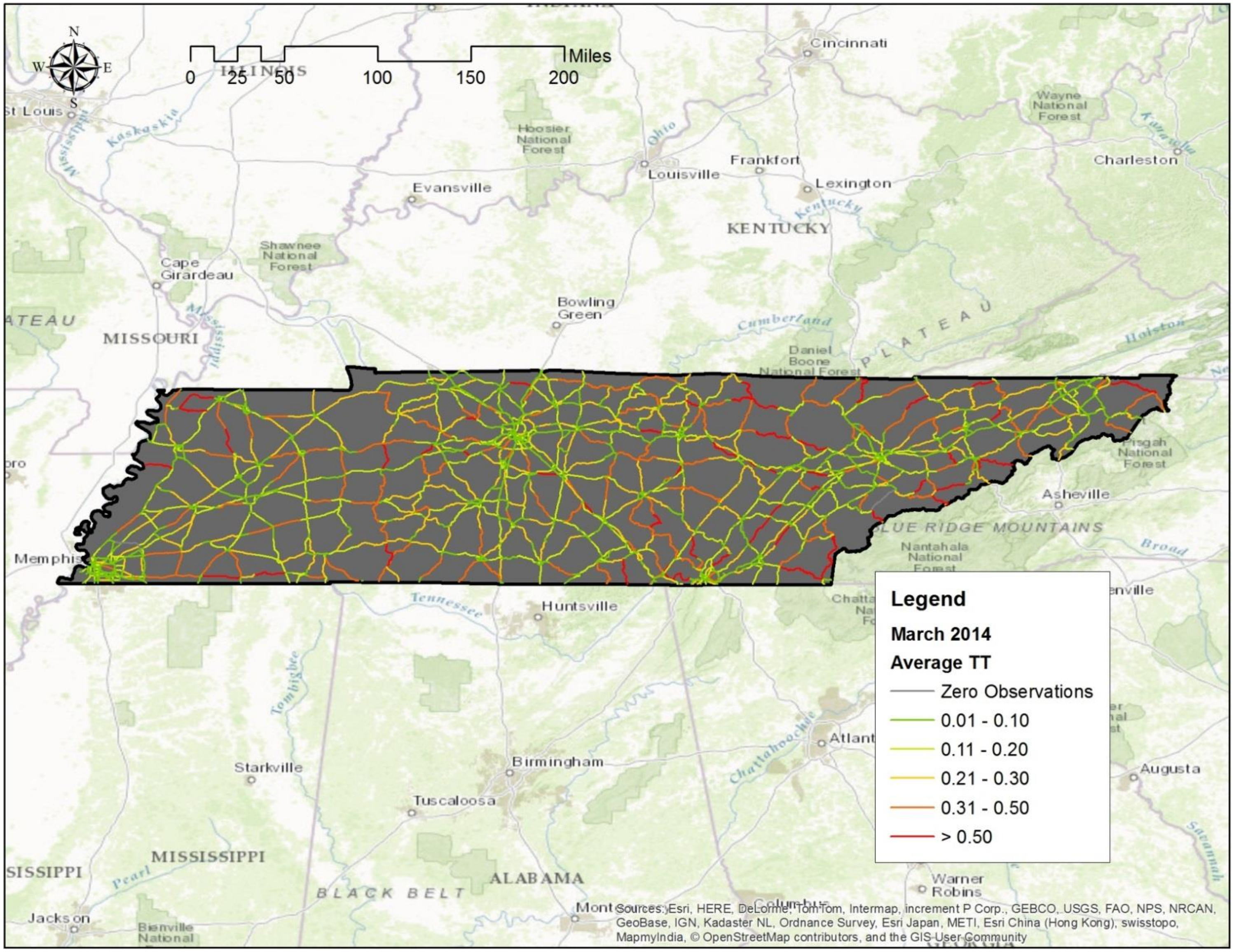


Figure 35 Average Travel Time for March 2014

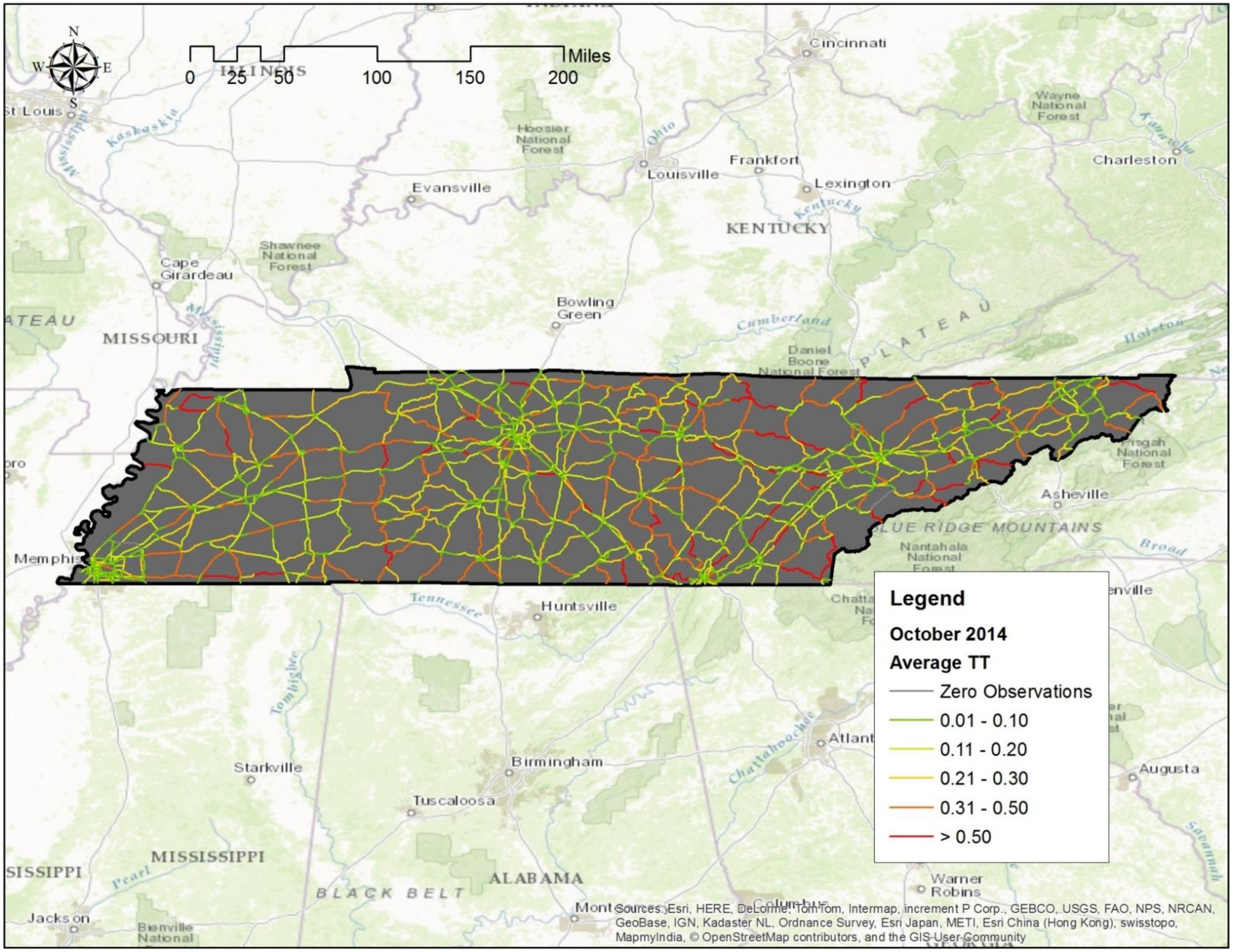


Figure 36 Average Travel Time for October 2014

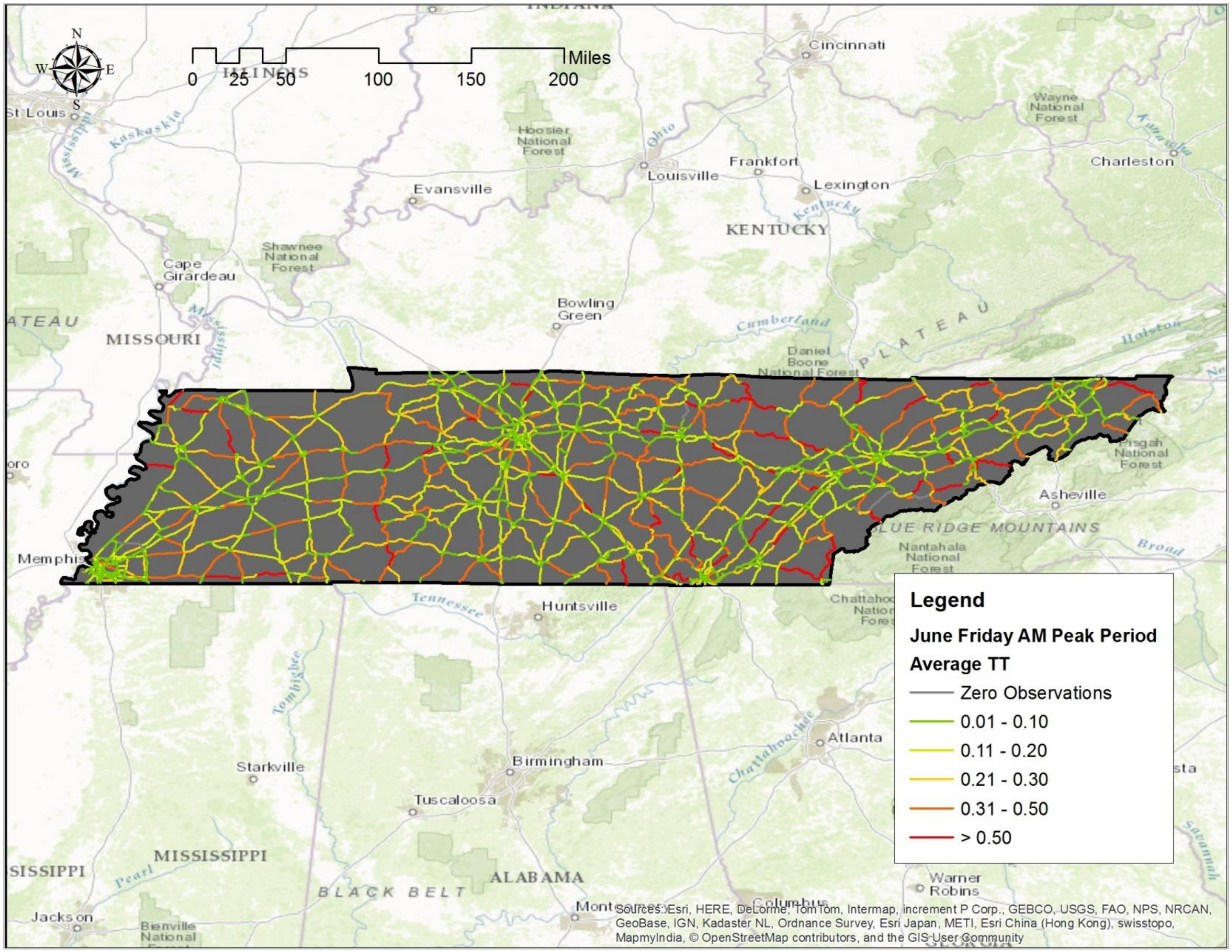


Figure 37 Average Travel Time for AM Peak Period Friday June 2011-2014

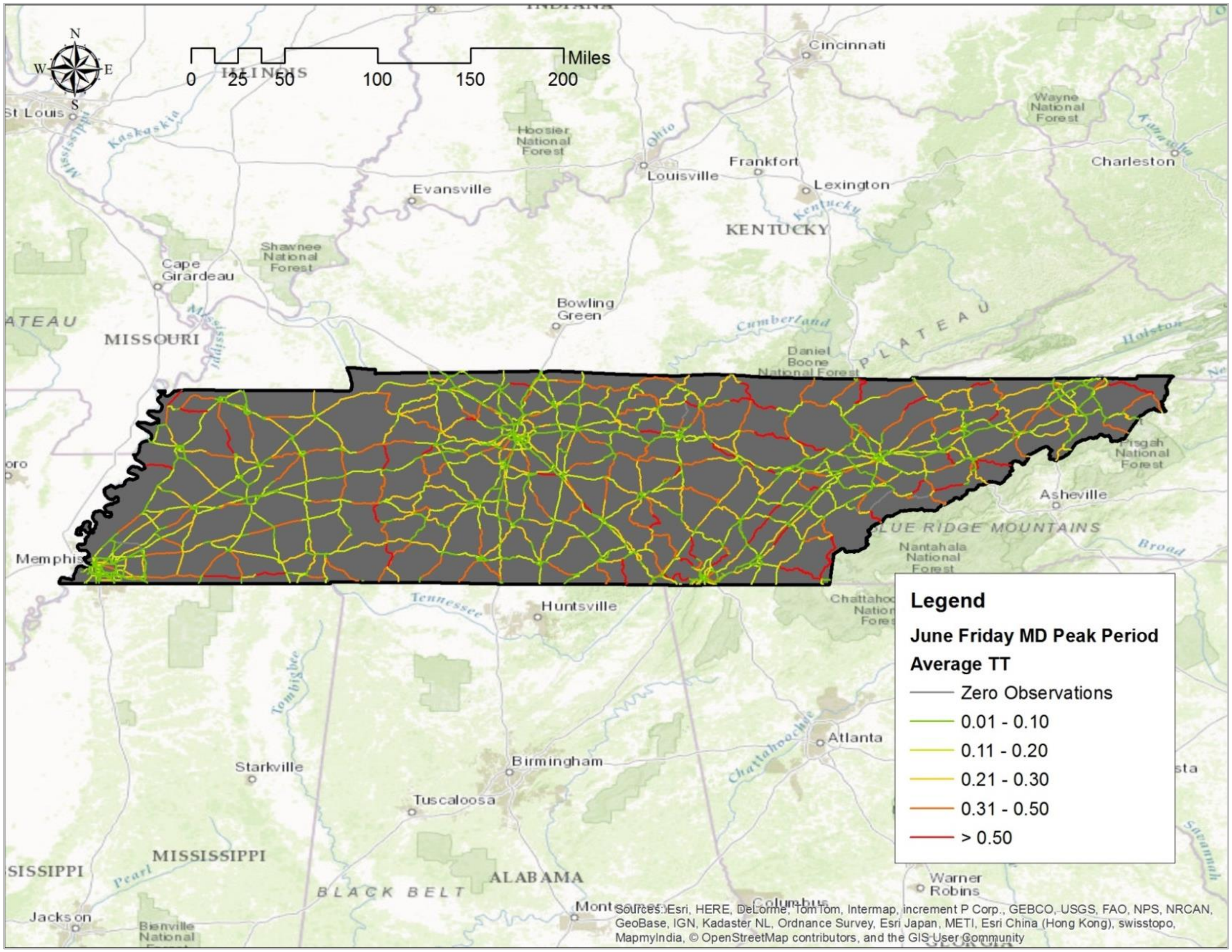


Figure 38 Average Travel Time for MD Peak Period Friday June 2011-2014

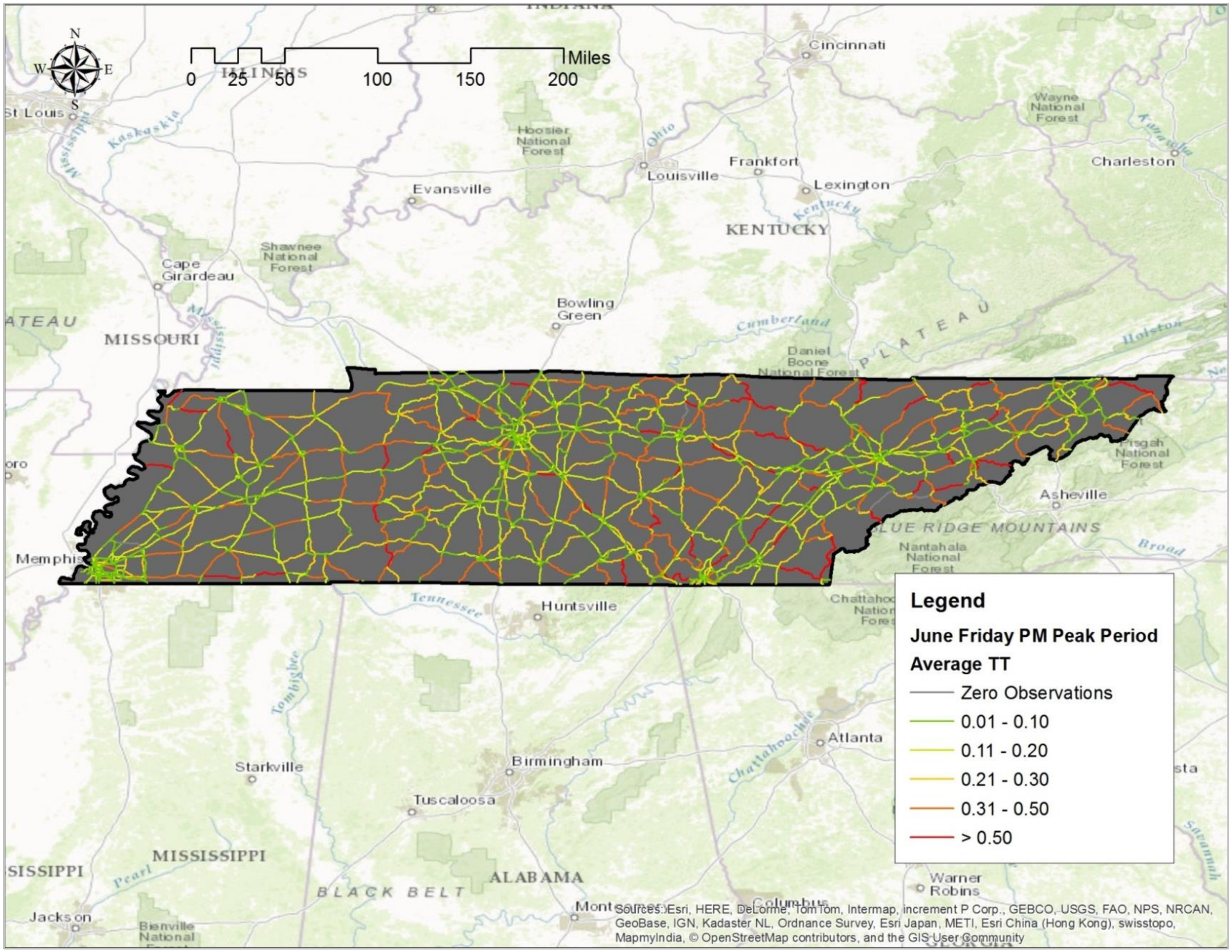


Figure 39 Average Travel Time for PM Peak Period Friday June 2011-2014

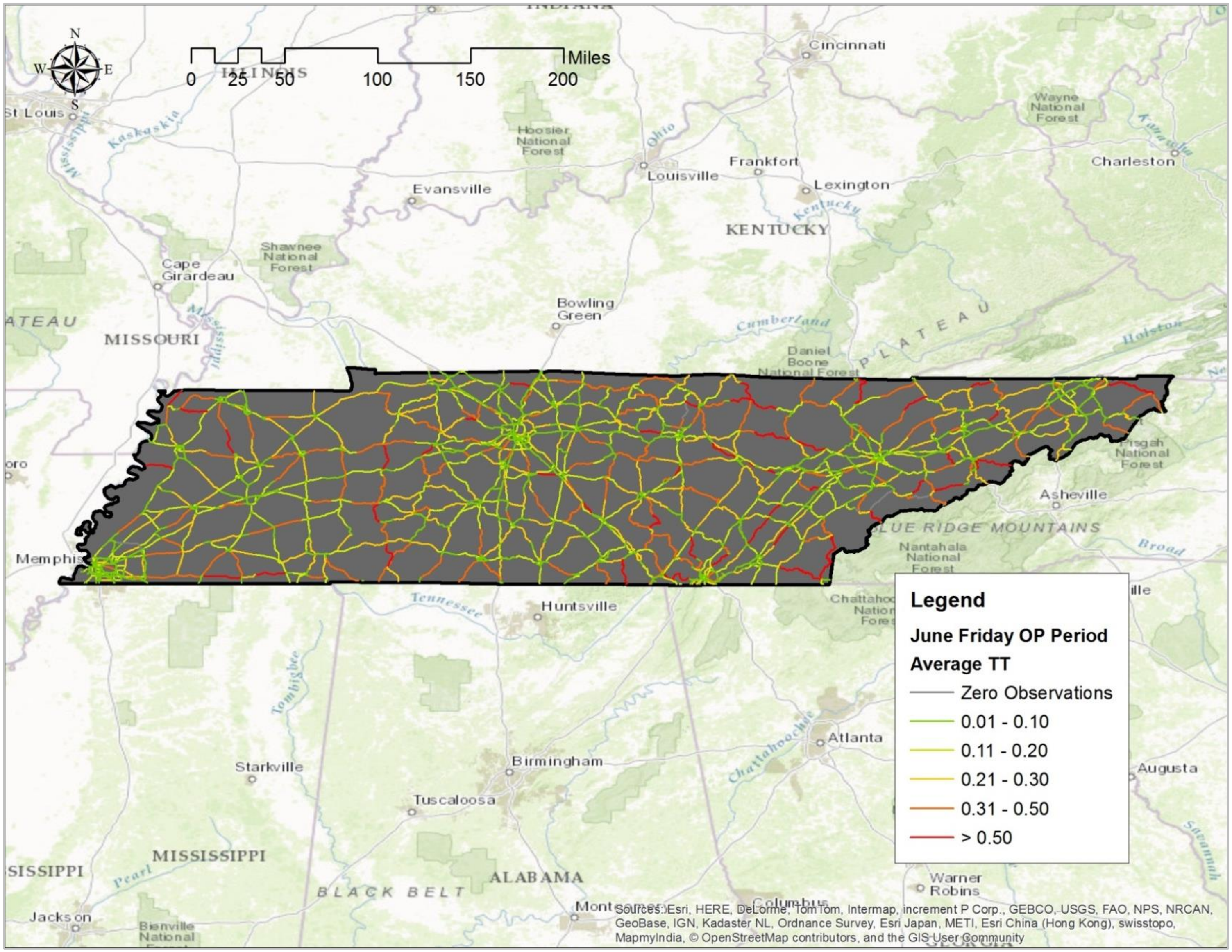


Figure 40 Average Travel Time for Off-Peak Period Friday June 2011-2014

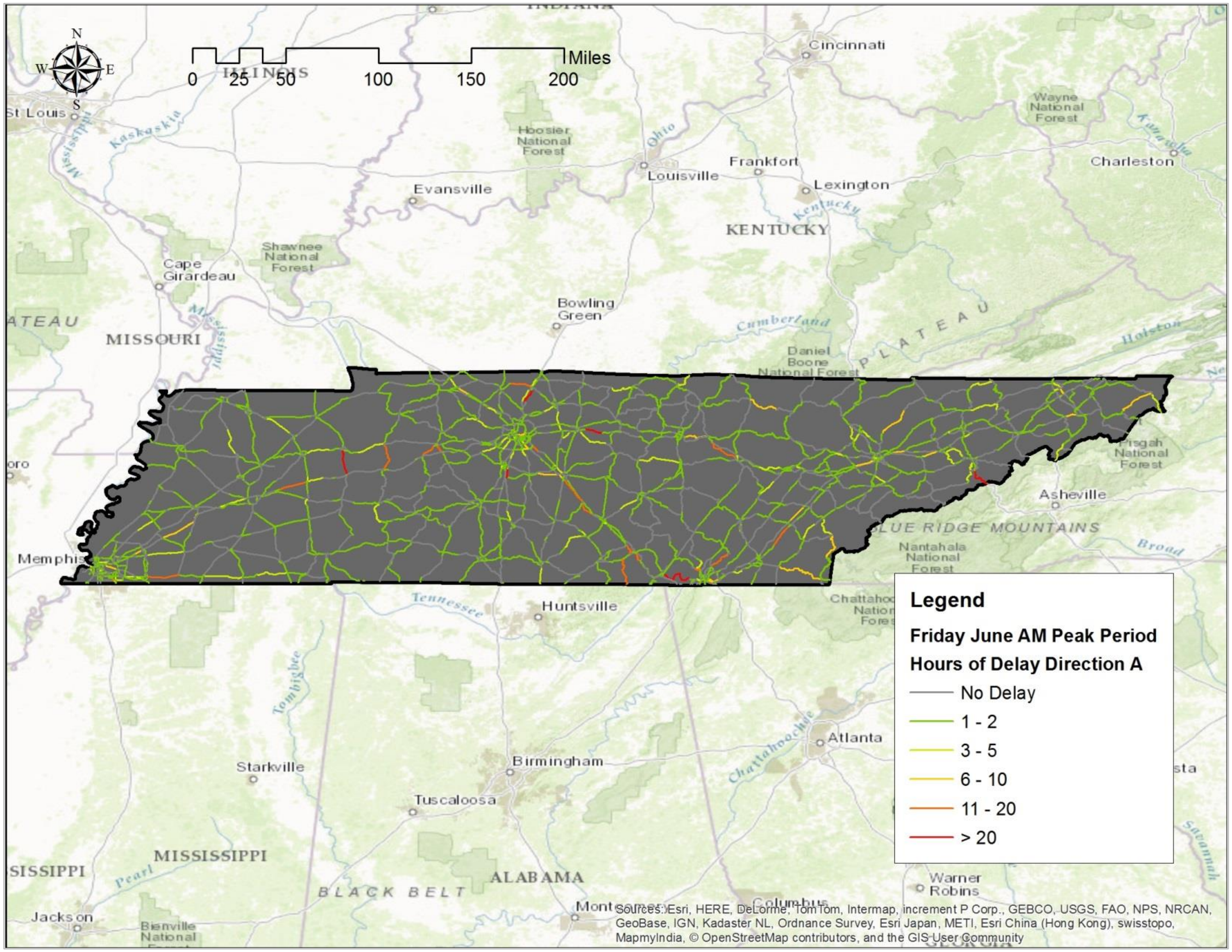


Figure 41 Hours of Delay in Direction A for AM Peak Period Friday June 2011-2014

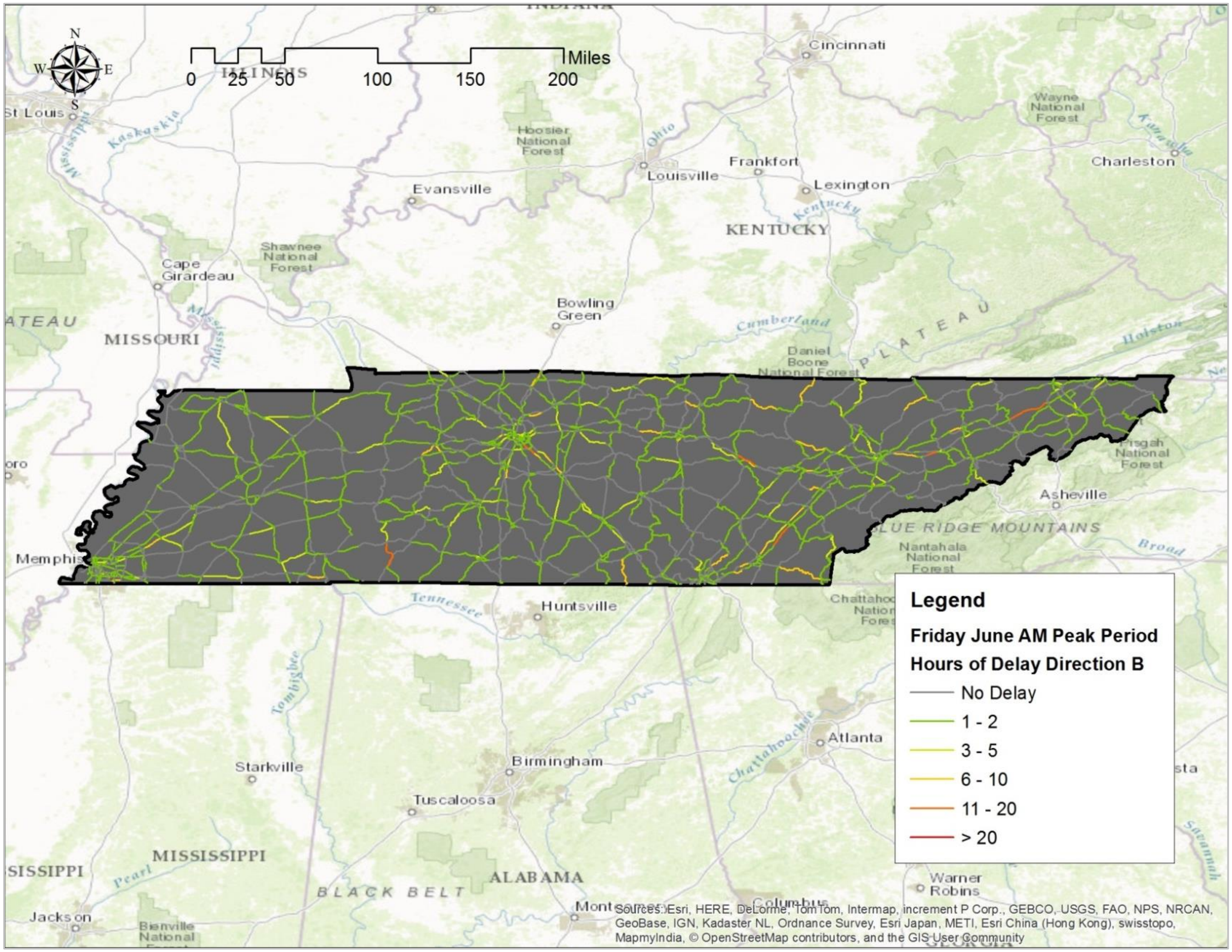


Figure 42 Hours of Delay in Direction B for AM Peak Period Friday June 2011-2014

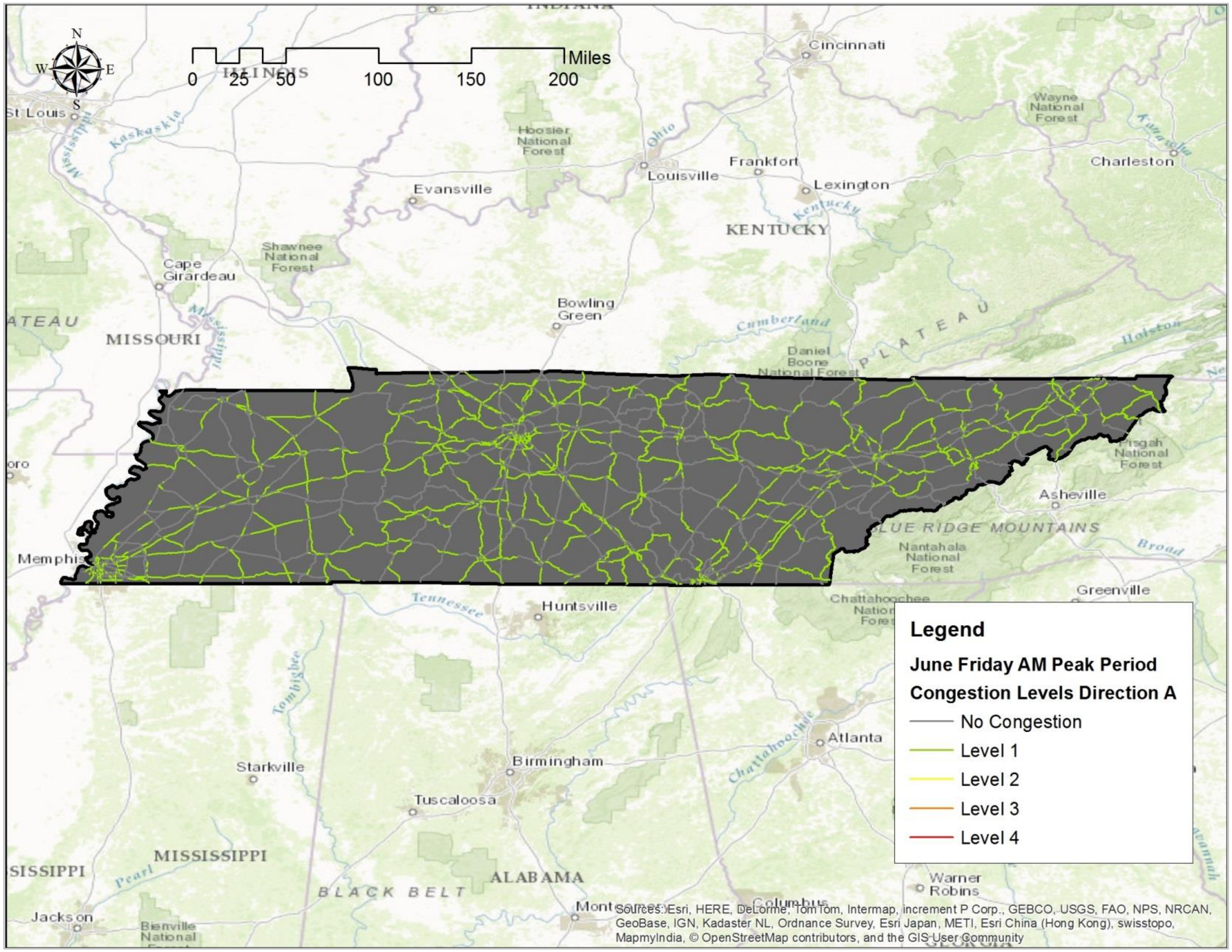


Figure 43 Levels of Congestion in Direction A for AM Peak Period Friday June 2011-2014

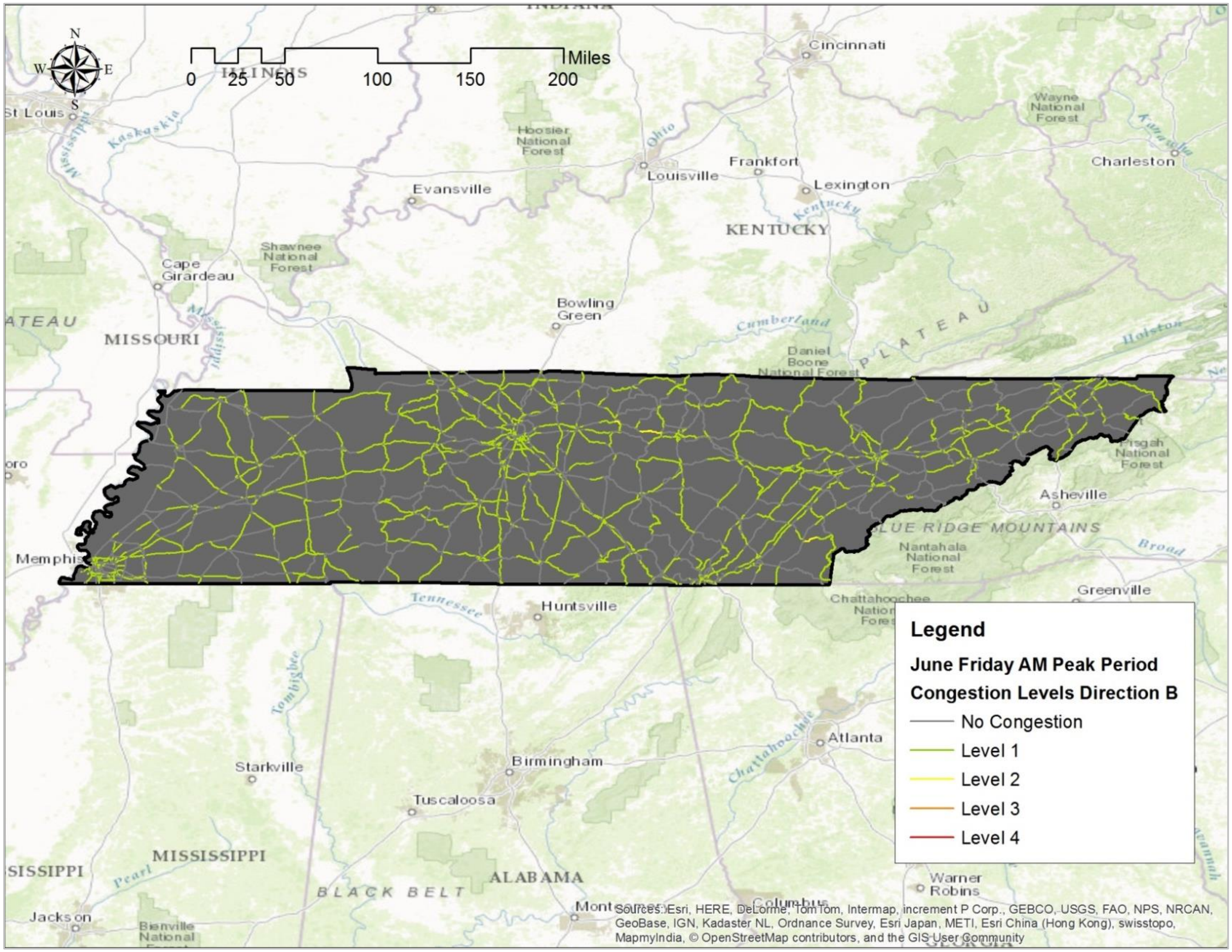


Figure 44 Levels of Congestion in Direction B for AM Peak Period Friday June 2011-2014

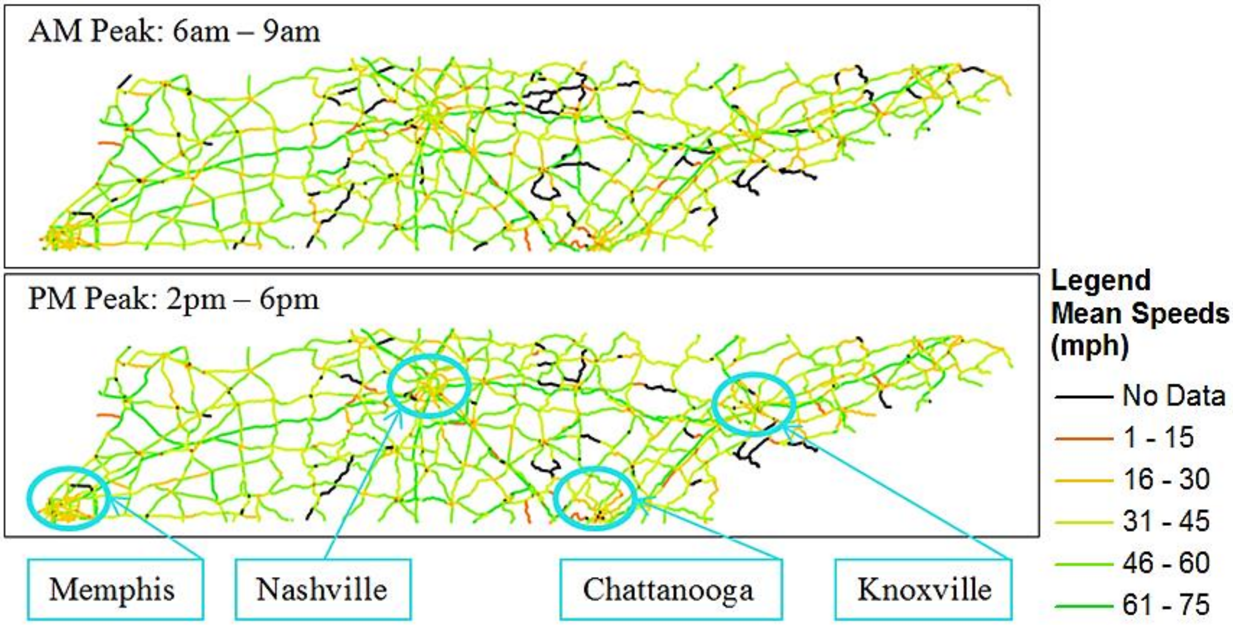


Figure 45 Average TS for AM and PM Peak Periods

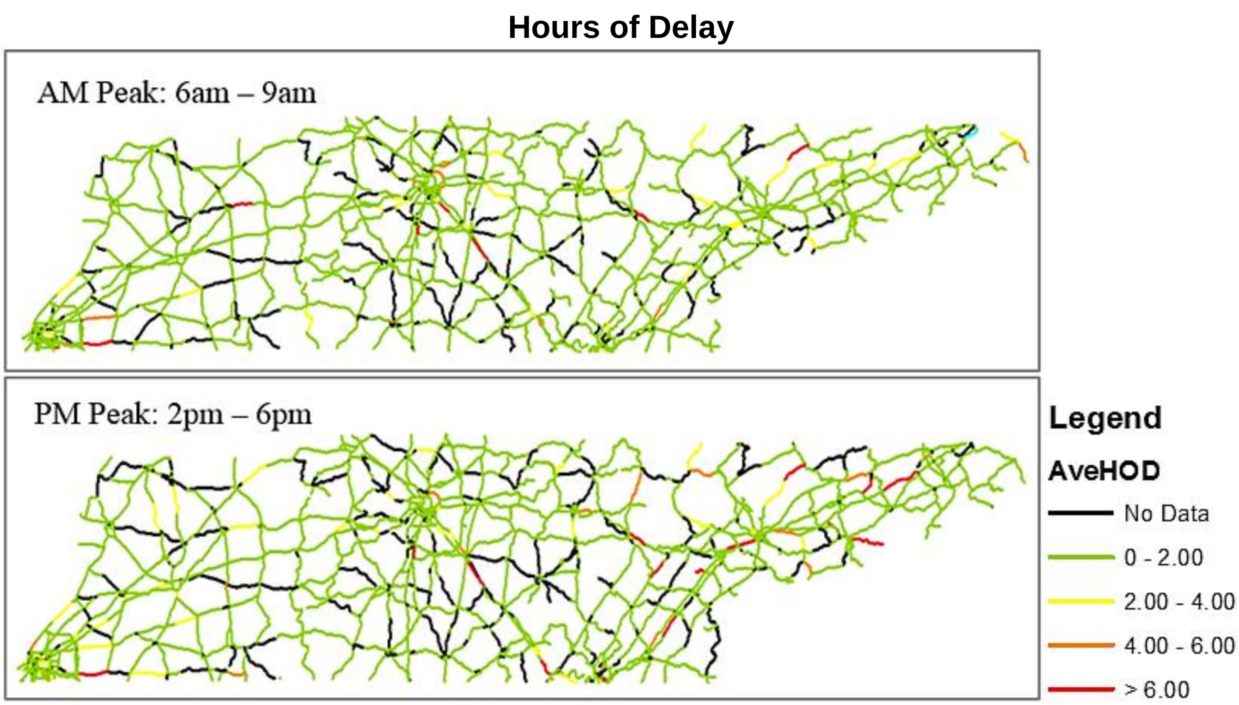


Figure 46 Hours of Delay for AM and PM Peak Periods

Now, we present the freight performance measures, such as average travel time (ATT), average 95th percentile travel time 9 (ATT 95%), average buffer index (ABI) and average travel time standard deviation (ATT-STD), Hours of delay (HOD) and contested lane miles (CLM), for each functional class (FC). These freight performance measures are averaged for each facility type to determine seasonal variation, peak and off-peak travel time variation travel time reliability, level of congestion etc. The result is presented for four years 2011-2014, three months -March, June and October, days of week categorized into Monday, Tuesday-Thursday, Friday, Weekends (Saturday-Sunday), and four time periods (AM, MD, PM, OP). The results are then compared to each other to identify any possible pattern/trend as shown in Table 7 - Table 14. Figure 47 breaks down the case study network into different functional classes. It can be noticed that several facility types such as rural major collector, urban minor arterial, collector and local have very few sample size (i.e. 1, 30, 12 and 7 segments respectively). Hence, the performance indicators obtained for these facility types may not be very reliable.

Table 7 FPMs by Month and Functional Class

	ATT (hrs)			ATT 95th Percentile (hrs)			ABI (hrs)			ATT-STD (hrs)		
FC	June	March	October	June	March	October	June	March	October	June	March	October
0	0.13	0.13	0.13	0.22	0.22	0.22	0.59	0.59	0.60	0.05	0.06	0.06
1	0.05	0.04	0.04	0.10	0.09	0.09	0.83	0.79	0.78	0.02	0.02	0.02
2	0.11	0.11	0.11	0.23	0.23	0.23	0.81	0.78	0.78	0.06	0.06	0.06
6	0.17	0.17	0.17	0.32	0.32	0.32	0.71	0.70	0.70	0.08	0.08	0.08
7	0.26	0.26	0.24	0.43	0.37	0.35	1.83	0.81	0.60	0.17	0.11	0.10
11	0.02	0.02	0.02	0.03	0.03	0.03	1.03	1.00	1.00	0.01	0.01	0.01
12	0.05	0.05	0.05	0.11	0.11	0.11	1.08	1.04	1.06	0.03	0.03	0.03
14	0.07	0.07	0.07	0.16	0.15	0.16	0.97	0.94	0.95	0.04	0.04	0.04
16	0.04	0.05	0.04	0.09	0.09	0.09	0.87	0.84	0.84	0.02	0.02	0.02
17	0.04	0.04	0.04	0.10	0.10	0.10	1.53	1.48	1.49	0.03	0.03	0.03
19	0.03	0.03	0.03	0.05	0.05	0.05	0.66	0.70	0.69	0.01	0.01	0.01

Table 8 By month comparison

	June vs. October				June vs. March				March vs. October			
FC	ATT (hrs)	ATT 95% (hrs)	ABI (hrs)	ATT-STD (hrs)	ATT (hrs)	ATT 95% (hrs)	ABI (hrs)	ATT-STD (hrs)	ATT (hrs)	ATT 95% (hrs)	ABI (hrs)	ATT-STD (hrs)
0	-1.18%	-3.05%	-2.41%	-4.84%	-2.44%	-2.78%	-0.15%	-3.35%	1.23%	-0.27%	-2.25%	-1.44%
1	1.77%	6.08%	6.13%	9.27%	1.11%	4.33%	5.15%	6.83%	0.67%	1.83%	1.03%	2.62%
2	0.65%	1.65%	3.41%	1.90%	0.28%	1.39%	3.87%	1.21%	0.37%	0.26%	-0.48%	0.70%
6	1.06%	0.31%	1.76%	1.22%	0.88%	1.09%	2.59%	1.34%	0.18%	-0.79%	-0.85%	-0.12%
7	6.56%	17.06%	66.94%	39.59%	-1.63%	12.22%	55.52%	35.37%	8.06%	5.52%	25.68%	6.54%
11	0.66%	1.44%	3.16%	2.47%	0.66%	2.02%	3.67%	2.47%	0.00%	-0.59%	-0.53%	0.00%
12	0.63%	0.00%	2.20%	0.69%	-0.21%	-0.09%	3.33%	0.35%	0.84%	0.09%	-1.17%	0.35%
14	0.00%	0.13%	1.81%	0.45%	-0.15%	0.84%	3.01%	1.13%	0.15%	-0.71%	-1.24%	-0.69%
16	0.22%	0.23%	3.76%	1.33%	-0.22%	0.70%	3.95%	1.77%	0.44%	-0.47%	-0.20%	-0.45%
17	0.56%	0.87%	3.08%	1.60%	0.00%	0.58%	3.24%	0.96%	0.56%	0.29%	-0.16%	0.65%
19	-0.32%	-4.43%	-3.42%	-8.65%	-2.57%	-5.84%	-4.68%	-10.58%	2.19%	1.33%	1.21%	1.74%

Table 9 FPMs by Days of week and Functional Class

	ATT (hrs)				ATT 95th Percentile (hrs)				ABI (hrs)				ATT-STD (hrs)			
FC	**Mon**	**Tue-Thu**	**Fri**	**Sat-Sun**	**Mon**	**Tue-Thu**	**Fri**	**Sat-Sun**	**Mon**	**Tue-Thu**	**Fri**	**Sat-Sun**	**Mon**	**Tue-Thu**	**Fri**	**Sat-Sun**
0	0.12	0.13	0.13	0.13	0.22	0.23	0.23	0.22	0.62	0.64	0.63	0.60	0.06	0.06	0.06	0.06
1	0.04	0.04	0.04	0.04	0.09	0.09	0.09	0.09	0.79	0.81	0.81	0.78	0.02	0.02	0.02	0.02
2	0.11	0.11	0.11	0.11	0.23	0.23	0.23	0.23	0.85	0.83	0.83	0.78	0.06	0.06	0.06	0.06
6	0.17	0.17	0.17	0.17	0.33	0.32	0.32	0.32	0.73	0.74	0.73	0.70	0.08	0.08	0.08	0.08
7	0.24	0.25	0.25	0.24	0.45	0.37	0.36	0.35	0.89	0.61	0.61	0.60	0.15	0.11	0.11	0.10
11	0.02	0.02	0.02	0.02	0.03	0.04	0.04	0.03	1.04	1.09	1.08	1.00	0.01	0.01	0.01	0.01
12	0.05	0.05	0.05	0.05	0.11	0.11	0.11	0.11	1.13	1.12	1.12	1.06	0.03	0.03	0.03	0.03
14	0.07	0.07	0.07	0.07	0.16	0.16	0.16	0.16	1.01	1.00	0.99	0.95	0.05	0.05	0.05	0.04
16	0.04	0.04	0.04	0.04	0.08	0.09	0.09	0.09	0.85	0.87	0.87	0.84	0.02	0.02	0.02	0.02
17	0.04	0.03	0.03	0.04	0.10	0.10	0.10	0.10	1.56	1.58	1.58	1.49	0.03	0.03	0.03	0.03
19	0.03	0.03	0.03	0.03	0.05	0.05	0.05	0.05	0.71	0.72	0.72	0.69	0.01	0.01	0.01	0.01

Table 10 By Days of week comparison

	Mon vs. Tue-Thu				Mon vs. Fri				Mon vs. Sat-Sun			
FC	**ATT (hrs)**	**ATT 95% (hrs)**	**ABI (hrs)**	**ATT-STD (hrs)**	**ATT (hrs)**	**ATT 95% (hrs)**	**ABI (hrs)**	**ATT-STD (hrs)**	**ATT (hrs)**	**ATT 95% (hrs)**	**ABI (hrs)**	**ATT-STD (hrs)**
0	4.55%	5.14%	2.41%	3.14%	4.96%	4.31%	0.34%	1.05%	4.39%	2.25%	-3.80%	-1.75%
1	-0.45%	-0.77%	3.56%	-0.54%	-0.23%	0.00%	3.36%	0.54%	0.23%	0.22%	-0.65%	1.09%
2	-0.75%	-1.07%	-1.87%	-2.17%	-0.56%	-1.11%	-2.27%	-2.17%	0.00%	-3.17%	-7.53%	-5.18%
6	-4.44%	-1.86%	0.97%	-1.88%	-4.09%	-1.98%	0.16%	-2.36%	-3.00%	-2.68%	-3.87%	-4.36%
7	7.03%	-16.55%	-31.38%	-31.16%	4.45%	-18.77%	-31.38%	-31.16%	1.99%	-20.86%	-32.35%	-34.75%
11	0.00%	1.72%	4.47%	2.44%	0.67%	2.59%	4.06%	2.44%	0.00%	-1.72%	-3.69%	-3.66%
12	-2.66%	-2.89%	-0.92%	-3.57%	-2.86%	-3.15%	-1.58%	-4.22%	-2.86%	-5.16%	-6.93%	-6.82%
14	-1.45%	-0.75%	-0.59%	-1.31%	-1.45%	-0.88%	-1.20%	-1.53%	-1.01%	-2.63%	-5.66%	-3.72%
16	-0.89%	1.65%	1.99%	3.67%	-0.89%	1.30%	1.78%	3.21%	-0.22%	0.94%	-1.40%	2.29%
17	-2.54%	0.58%	1.21%	0.98%	-2.26%	0.39%	1.42%	0.98%	0.28%	-0.10%	-4.58%	0.00%
19	-6.01%	-3.83%	2.37%	-2.50%	-6.31%	-4.19%	1.56%	-3.33%	-6.31%	-5.46%	-2.62%	-5.83%

Table 10 By Days of week comparison (Continued)

	Tue-Thu vs. Fri				Tue-Thu vs. Sat-Sun				Fri vs. Sat-Sun			
FC	ATT (hrs)	ATT 95% (hrs)	ABI (hrs)	ATT-STD (hrs)	ATT (hrs)	ATT 95% (hrs)	ABI (hrs)	ATT-STD (hrs)	ATT (hrs)	ATT 95% (hrs)	ABI (hrs)	ATT-STD (hrs)
0	0.39%	-0.79%	-2.02%	-2.03%	-0.16%	-2.75%	-6.07%	-4.74%	-0.54%	-1.98%	-4.13%	-2.76%
1	0.23%	0.78%	-0.20%	1.09%	0.68%	1.00%	-4.06%	1.64%	0.45%	0.22%	-3.88%	0.54%
2	0.19%	-0.04%	-0.41%	0.00%	0.76%	-2.12%	-5.77%	-3.08%	0.57%	-2.08%	-5.38%	-3.08%
6	0.36%	-0.12%	-0.80%	-0.48%	1.51%	-0.84%	-4.80%	-2.52%	1.14%	-0.72%	-4.03%	-2.05%
7	-2.41%	-2.66%	0.00%	0.00%	-4.71%	-5.16%	-1.42%	-5.21%	-2.35%	-2.57%	-1.42%	-5.21%
11	0.67%	0.85%	-0.40%	0.00%	0.00%	-3.39%	-7.81%	-5.95%	-0.66%	-4.20%	-7.45%	-5.95%
12	-0.21%	-0.27%	-0.67%	-0.67%	-0.21%	-2.34%	-6.07%	-3.37%	0.00%	-2.08%	-5.44%	-2.71%
14	0.00%	-0.13%	-0.62%	-0.22%	0.44%	-1.89%	-5.10%	-2.44%	0.44%	-1.77%	-4.51%	-2.22%
16	0.00%	-0.35%	-0.21%	-0.44%	0.67%	-0.70%	-3.32%	-1.33%	0.67%	-0.35%	-3.12%	-0.89%
17	0.29%	-0.19%	0.21%	0.00%	2.90%	-0.67%	-5.72%	-0.97%	2.60%	-0.48%	-5.92%	-0.97%
19	-0.32%	-0.38%	-0.79%	-0.85%	-0.32%	-1.70%	-4.88%	-3.42%	0.00%	-1.33%	-4.12%	-2.59%

Table 11 FPMs by Days of week and Functional Class

	ATT (hrs)				ATT 95th Percentile (hrs)				ABI (hrs)				ATT-STD (hrs)			
FC	2011	2012	2013	2014	2011	2012	2013	2014	2011	2012	2013	2014	2011	2012	2013	2014
0	0.13	0.13	0.13	0.13	0.20	0.21	0.22	0.22	0.49	0.54	0.57	0.60	0.05	0.05	0.05	0.06
1	0.04	0.04	0.04	0.04	0.09	0.09	0.09	0.09	0.70	0.74	0.77	0.78	0.02	0.02	0.02	0.02
2	0.11	0.11	0.11	0.11	0.22	0.22	0.22	0.23	0.68	0.72	0.74	0.78	0.05	0.05	0.06	0.06
6	0.17	0.17	0.17	0.17	0.31	0.31	0.32	0.32	0.62	0.65	0.67	0.70	0.08	0.08	0.08	0.08
7	0.23	0.27	0.25	0.24	0.37	0.37	0.35	0.35	0.64	0.61	0.61	0.60	0.10	0.11	0.11	0.10
11	0.01	0.01	0.02	0.02	0.03	0.03	0.03	0.03	0.86	0.91	0.94	1.00	0.01	0.01	0.01	0.01
12	0.05	0.05	0.05	0.05	0.10	0.10	0.11	0.11	0.88	0.94	0.99	1.06	0.03	0.03	0.03	0.03
14	0.07	0.07	0.07	0.07	0.15	0.15	0.15	0.16	0.80	0.86	0.89	0.95	0.04	0.04	0.04	0.04
16	0.05	0.04	0.04	0.04	0.08	0.08	0.08	0.09	0.69	0.74	0.78	0.84	0.02	0.02	0.02	0.02
17	0.05	0.04	0.04	0.04	0.10	0.10	0.10	0.10	1.10	1.22	1.35	1.49	0.03	0.03	0.03	0.03
19	0.04	0.03	0.03	0.03	0.05	0.05	0.05	0.05	0.51	0.58	0.64	0.69	0.01	0.01	0.01	0.01

Table 12 By year comparison

	2011 vs. 2012				2011 vs. 2013				2011 vs. 2014			
FC	ATT (hrs)	ATT 95% (hrs)	ABI (hrs)	ATT-STD (hrs)	ATT (hrs)	ATT 95% (hrs)	ABI (hrs)	ATT-STD (hrs)	ATT (hrs)	ATT 95% (hrs)	ABI (hrs)	ATT-STD (hrs)
0	1.03%	2.80%	9.97%	-0.78%	2.23%	8.26%	16.28%	6.63%	2.15%	9.54%	22.43%	9.75%
1	0.90%	4.48%	6.12%	7.65%	0.68%	5.06%	10.10%	9.41%	0.45%	4.83%	11.95%	9.41%
2	-0.73%	1.53%	5.88%	1.11%	-1.19%	2.83%	8.94%	2.22%	-2.11%	4.97%	15.49%	5.00%
6	-0.06%	1.46%	4.75%	0.00%	0.06%	2.60%	8.62%	0.50%	-0.71%	3.67%	13.76%	1.37%
7	18.18%	-0.24%	-4.12%	6.13%	12.09%	-3.80%	-4.12%	6.13%	7.02%	-4.29%	-5.48%	0.60%
11	1.36%	0.30%	4.81%	-1.30%	2.04%	0.91%	8.79%	0.00%	2.04%	3.64%	15.97%	2.60%
12	-0.21%	2.18%	7.03%	2.63%	0.21%	5.56%	13.54%	6.39%	-0.63%	7.65%	20.47%	7.89%
14	-0.85%	3.29%	7.42%	2.90%	-1.98%	4.38%	11.45%	4.11%	-3.26%	6.37%	18.20%	6.28%
16	-4.89%	1.49%	7.52%	5.45%	-5.32%	1.87%	12.98%	7.43%	-4.68%	6.34%	21.24%	10.40%
17	-6.86%	0.99%	11.21%	2.45%	-15.49%	2.07%	23.52%	6.99%	-21.46%	1.68%	35.60%	7.34%
19	-4.44%	-0.77%	13.85%	7.84%	-9.44%	-1.35%	26.23%	6.86%	-13.33%	0.19%	35.64%	10.78%

Table 12 By year comparison (continued)

	2012 vs. 2013				2012 vs. 2014				2013 vs. 2014			
FC	ATT (hrs)	ATT 95% (hrs)	ABI (hrs)	ATT-STD (hrs)	ATT (hrs)	ATT 95% (hrs)	ABI (hrs)	ATT-STD (hrs)	ATT (hrs)	ATT 95% (hrs)	ABI (hrs)	ATT-STD (hrs)
0	1.18%	5.31%	5.74%	7.47%	1.10%	6.55%	11.33%	10.61%	-0.08%	1.18%	5.29%	2.93%
1	-0.22%	0.55%	3.75%	1.64%	-0.45%	0.33%	5.49%	1.64%	-0.22%	-0.22%	1.68%	0.00%
2	-0.46%	1.28%	2.89%	1.10%	-1.39%	3.38%	9.08%	3.85%	-0.93%	2.08%	6.02%	2.72%
6	0.12%	1.12%	3.70%	0.50%	-0.65%	2.18%	8.60%	1.37%	-0.77%	1.05%	4.73%	0.87%
7	-5.15%	-3.56%	0.00%	0.00%	-9.44%	-4.05%	-1.42%	-5.21%	-4.52%	-0.51%	-1.42%	-5.21%
11	0.67%	0.60%	3.80%	1.32%	0.67%	3.32%	10.65%	3.95%	0.00%	2.70%	6.60%	2.60%
12	0.42%	3.30%	6.08%	3.66%	-0.42%	5.34%	12.55%	5.13%	-0.84%	1.98%	6.10%	1.41%
14	-1.14%	1.06%	3.75%	1.17%	-2.43%	2.98%	10.03%	3.29%	-1.30%	1.90%	6.06%	2.09%
16	-0.45%	0.37%	5.07%	1.88%	0.22%	4.78%	12.76%	4.69%	0.67%	4.40%	7.31%	2.76%
17	-9.26%	1.07%	11.07%	4.44%	-15.68%	0.68%	21.93%	4.78%	-7.07%	-0.39%	9.79%	0.33%
19	-5.23%	-0.58%	10.88%	-0.91%	-9.30%	0.97%	19.14%	2.73%	-4.29%	1.57%	7.45%	3.67%

Table 13 FPMs by Time of day and Functional Class

	ATT (hrs)				ATT 95th Percentile (hrs)				ABI (hrs)				ATT-STD (hrs)				HOD (veh-hrs)			CLM		
FC	AM	MD	PM	OP	AM	MD	PM	OP	AM	MD	PM	OP	AM	MD	PM	OP	AM	MD	PM	AM	MD	PM
0	0.13	0.13	0.13	0.13	0.21	0.23	0.22	0.22	0.47	0.62	0.60	0.61	0.05	0.06	0.06	0.06	2.33	1.92	1.93	0.74	0.66	0.62
1	0.04	0.04	0.04	0.04	0.09	0.09	0.09	0.09	0.70	0.78	0.78	0.79	0.02	0.02	0.02	0.02	5.61	11.38	25.34	0.72	0.70	0.74
2	0.11	0.11	0.11	0.11	0.23	0.24	0.23	0.23	0.69	0.86	0.78	0.80	0.06	0.06	0.06	0.06	3.29	3.62	3.32	0.79	0.75	0.69
6	0.19	0.17	0.17	0.17	0.31	0.33	0.32	0.32	0.58	0.76	0.70	0.72	0.08	0.09	0.08	0.08	4.07	3.05	2.87	0.81	0.73	0.67
7	0.06	0.26	0.24	0.25	0.06	0.40	0.35	0.39	0.00	0.55	0.60	0.55	0.00	0.11	0.10	0.10	0.00	0.19	0.00	0.00	1.00	1.00
11	0.02	0.01	0.02	0.01	0.04	0.04	0.03	0.03	0.93	1.02	1.00	0.96	0.01	0.01	0.01	0.01	5.30	4.98	8.59	0.94	0.85	0.86
12	0.05	0.05	0.05	0.05	0.11	0.11	0.11	0.11	0.89	1.12	1.06	1.07	0.03	0.03	0.03	0.03	3.77	2.87	3.10	0.89	0.76	0.70
14	0.07	0.07	0.07	0.07	0.15	0.16	0.16	0.16	0.78	1.00	0.95	0.97	0.04	0.05	0.04	0.04	3.20	3.58	3.07	0.81	0.85	0.75
16	0.05	0.04	0.04	0.05	0.08	0.09	0.09	0.09	0.71	0.86	0.84	0.86	0.02	0.02	0.02	0.02	1.95	0.76	5.78	0.90	0.67	0.50
17	0.04	0.03	0.04	0.04	0.10	0.10	0.10	0.10	1.22	1.56	1.49	1.47	0.03	0.03	0.03	0.03	4.90	1.07	0.17	0.67	0.50	0.33
19	0.03	0.03	0.03	0.03	0.05	0.05	0.05	0.05	0.54	0.67	0.69	0.67	0.01	0.01	0.01	0.01	0.91	1.35	0.39	0.86	1.00	0.71

Table 14 By time of day comparison

	MD vs. AM						PM vs. AM						OP vs. AM			
FC	ATT (hrs)	ATT 95% (hrs)	ABI (hrs)	ATT-STD (hrs)	HOD (veh-hrs)	CLM	ATT (hrs)	ATT 95% (hrs)	ABI (hrs)	ATT-STD (hrs)	HOD (veh-hrs)	CLM	ATT (hrs)	ATT 95% (hrs)	ABI (hrs)	ATT-STD (hrs)
0	-3.91%	9.52%	32.84%	19.35%	-17.55%	-10.39%	-3.31%	7.68%	27.78%	14.66%	-17.06%	-16.45%	-3.76%	7.15%	30.51%	15.07%
1	-0.67%	2.71%	11.50%	5.82%	102.71%	-2.77%	-0.89%	-1.30%	11.10%	-1.59%	351.26%	3.24%	-1.56%	-2.27%	12.67%	-3.17%
2	-5.89%	6.41%	25.45%	7.03%	9.98%	-4.84%	-6.24%	-0.04%	13.81%	-2.74%	0.91%	-12.22%	-6.33%	0.97%	16.52%	-1.20%
6	-9.48%	6.35%	31.08%	12.98%	-25.04%	-10.21%	-9.42%	1.72%	20.91%	4.37%	-29.40%	-17.48%	-9.32%	2.46%	23.39%	5.40%
7	314.45%	542.70%	-	-	-	-	286.52%	466.45%	-	-	-	-	302.57%	522.47%	-	-
11	-6.88%	-6.12%	9.72%	-8.70%	-6.17%	-9.71%	-6.25%	-9.04%	8.13%	-14.13%	61.93%	-8.58%	-10.00%	-13.83%	3.78%	-19.57%
12	-6.05%	7.08%	26.74%	11.19%	-23.87%	-14.46%	-7.23%	2.26%	19.09%	3.61%	-17.57%	-21.69%	-7.03%	3.58%	21.24%	6.14%
14	-5.49%	8.13%	28.02%	9.79%	12.09%	5.05%	-6.18%	4.37%	21.17%	5.01%	-3.90%	-7.02%	-6.59%	5.17%	24.34%	5.97%
16	-6.13%	5.53%	21.82%	8.21%	-61.03%	-25.92%	-5.29%	5.04%	18.73%	7.73%	196.34%	-44.44%	-4.23%	6.14%	22.07%	9.66%
17	-7.71%	3.28%	27.26%	2.29%	-78.26%	-25.00%	-5.59%	2.38%	21.43%	0.33%	-96.58%	-50.01%	-3.72%	2.58%	20.40%	-0.33%
19	0.00%	6.05%	23.73%	11.88%	47.63%	16.67%	-2.50%	4.64%	26.82%	11.88%	-57.67%	-16.66%	0.31%	5.04%	24.64%	8.91%

Table 14 By time of day comparison (continued)

	MD vs. PM						MD vs. OP				PM vs. OP			
FC	ATT (hrs)	ATT 95% (hrs)	ABI (hrs)	ATT-STD (hrs)	HOD (veh-hrs)	CLM	ATT (hrs)	ATT 95% (hrs)	ABI (hrs)	ATT-STD (hrs)	ATT (hrs)	ATT 95% (hrs)	ABI (hrs)	ATT-STD (hrs)
0	0.63%	-1.68%	-3.80%	-3.92%	0.59%	-6.77%	0.16%	-2.16%	-1.75%	-3.58%	-0.47%	-0.49%	2.14%	0.36%
1	-0.22%	-3.90%	-0.36%	-7.00%	122.61%	6.18%	-0.90%	-4.85%	1.05%	-8.50%	-0.68%	-0.99%	1.41%	-1.61%
2	-0.37%	-6.07%	-9.28%	-9.13%	-8.25%	-7.75%	-0.47%	-5.11%	-7.12%	-7.69%	-0.09%	1.02%	2.38%	1.59%
6	0.06%	-4.35%	-7.76%	-7.62%	-5.82%	-8.09%	0.18%	-3.66%	-5.87%	-6.71%	0.12%	0.72%	2.05%	0.99%
7	-6.74%	-11.86%	9.56%	-7.06%	-100.00%	0.00%	-2.87%	-3.15%	-0.87%	-3.71%	4.15%	9.89%	-9.52%	3.60%
11	0.67%	-3.12%	-1.45%	-5.95%	72.57%	1.25%	-3.36%	-8.22%	-5.41%	-11.90%	-4.00%	-5.26%	-4.02%	-6.33%
12	-1.25%	-4.49%	-6.04%	-6.82%	8.27%	-8.45%	-1.04%	-3.26%	-4.34%	-4.55%	0.21%	1.29%	1.80%	2.44%
14	-0.73%	-3.48%	-5.35%	-4.35%	-14.26%	-11.49%	-1.16%	-2.73%	-2.87%	-3.48%	-0.44%	0.77%	2.61%	0.91%
16	0.90%	-0.47%	-2.54%	-0.45%	660.38%	-25.00%	2.03%	0.58%	0.21%	1.34%	1.12%	1.05%	2.82%	1.79%
17	2.31%	-0.87%	-4.59%	-1.92%	-84.29%	-33.34%	4.32%	-0.67%	-5.40%	-2.56%	1.97%	0.19%	-0.85%	-0.65%
19	-2.50%	-1.33%	2.49%	0.00%	-71.32%	-28.57%	0.31%	-0.95%	0.73%	-2.65%	2.88%	0.39%	-1.72%	-2.65%

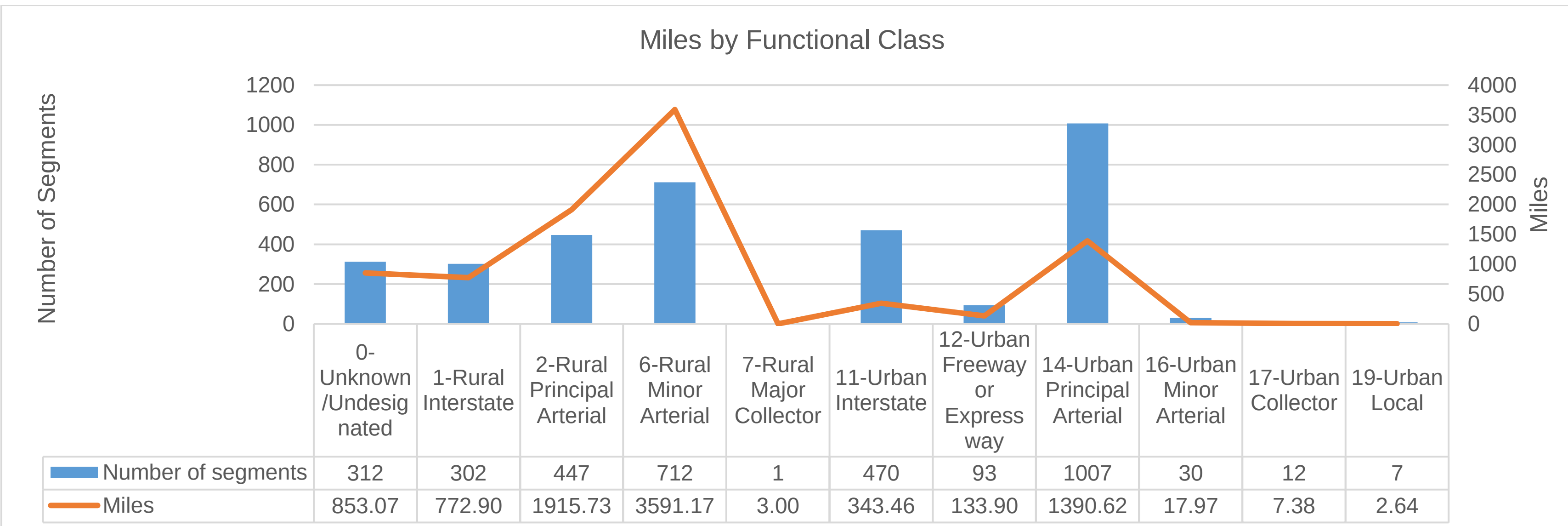


Figure 47 Case study network by functional class

5.2 O-D Based Reliability

Reliability has become a significant performance measure of the transportation system in a region. Travel time reliability, presented in the form of descriptive statistics derived from the distribution of travel times is a critical indication of the operating conditions of any network. Even though many past studies have tried to measure behavioral response to travel time reliability, their application in a transportation-planning context is limited. In the planning stage, agencies often would like to utilize available resources to estimate travel time reliability using existing tools; hence, a framework to measure path-based reliability to calculate network-wide reliability using available data will be very useful, and is currently lacking in the literature. The objective of this study is to model O-D based freight reliability using truck GPS data considering ideal, recurring and non-recurring travel conditions to assist short term transportation planning and operations decision making. This study contributes to the freight literature by (1) measuring travel time reliability for each O-D pairs, and (2) assessing reliability variation in the event of non-recurring congestion.

5.2.1 Methodology

The methodological framework for computing and predicting path-based truck travel time reliability is shown in Figure 48. The first task is to collect GPS data for multiple days for the study area. Typical truck GPS data consists of latitude, longitude of the truck, time stamp, speed, and heading (direction). The next task is to attach the GPS data to the network. Once the GPS data is attached to the network the next task is to determine the shortest path for each O-D pair and corresponding links associated with the shortest path. The shortest path is calculated using free flow travel time. Since the GPS data includes speed of the truck, travel times for each link can be computed. The path travel time is obtained by aggregating the link travel time over the links used in traversing the shortest path. Since travel times will be affected by recurring and by non-recurring congestion, travel times need to be separated by each type. Travel time reliability measures for each path can be determined by replicating the procedure and collecting data for multiple days. Reliability measures including but not limited to 95th percentile travel time, standard deviation, and coefficient of variation can be determined for each path. Path based reliability can be helpful in number of ways such as incorporating travel time reliability in travel demand models, and short term travel time predictions.

In this study, travel time reliability is compared for three different travel conditions: ideal, recurring and non-recurring congestion. Ideal travel conditions consider the free flow travel time and when compared with other conditions help to identify the variations in travel time. Recurring congestion refers to primarily the travel condition mostly associated with roadway network operating at over-capacity. This type of congestion is not affected by external factors such as inclement weather, crash etc. Non-recurring congestion occurs due to construction, severe weather, crashes, and special events. In this study we only considered crashes as the external factor behind non-recurring congestion. Comparing three different travel conditions will allow us to capture the variability in travel time and identify the O-D pairs, which are mostly affected.

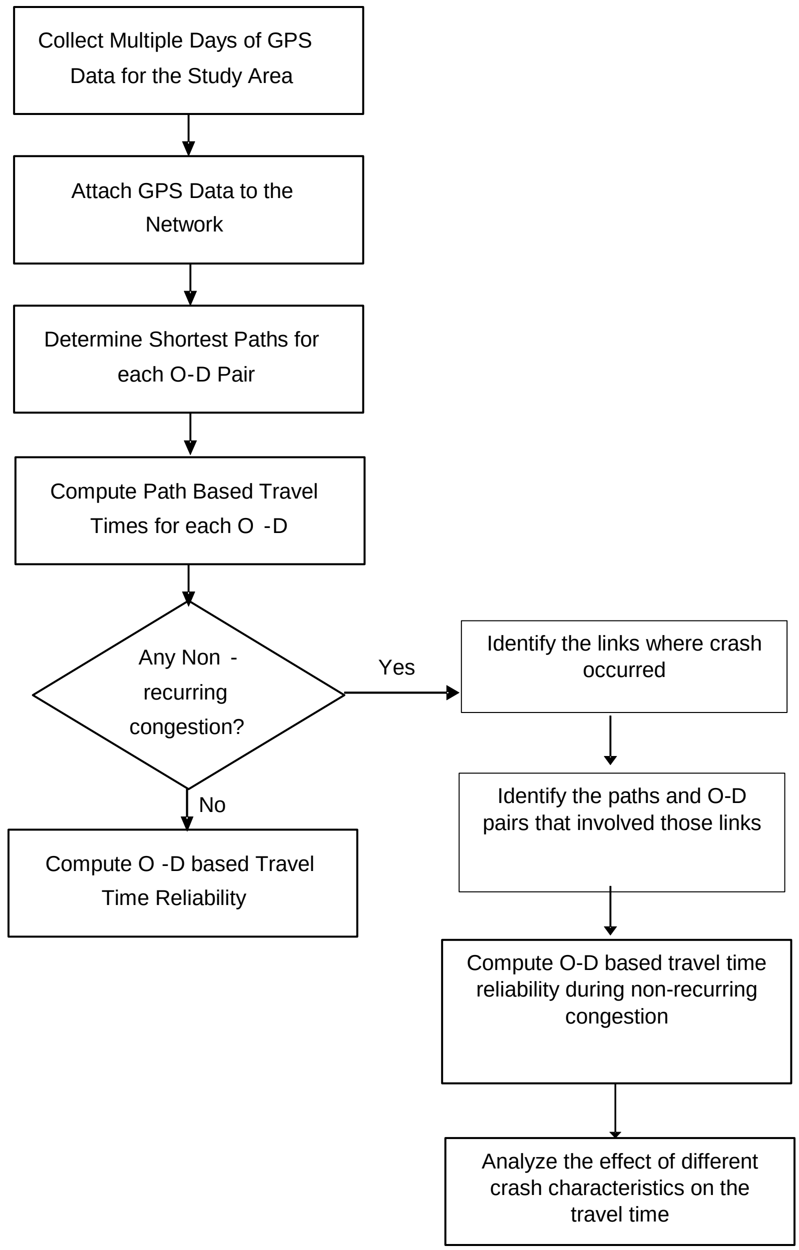


Figure 48: Framework for Computing and Predicting Truck Travel Time Reliability

5.2.2 Study Area

The proposed framework was applied in Shelby County, TN. Figure 49 shows zone, FAF network, and truck GPS data for one day. For computing reliability, the shortest path is considered for each O-D pair. Once path based travel time and travel time reliability were determined a number of combinations of relationship were attempted.

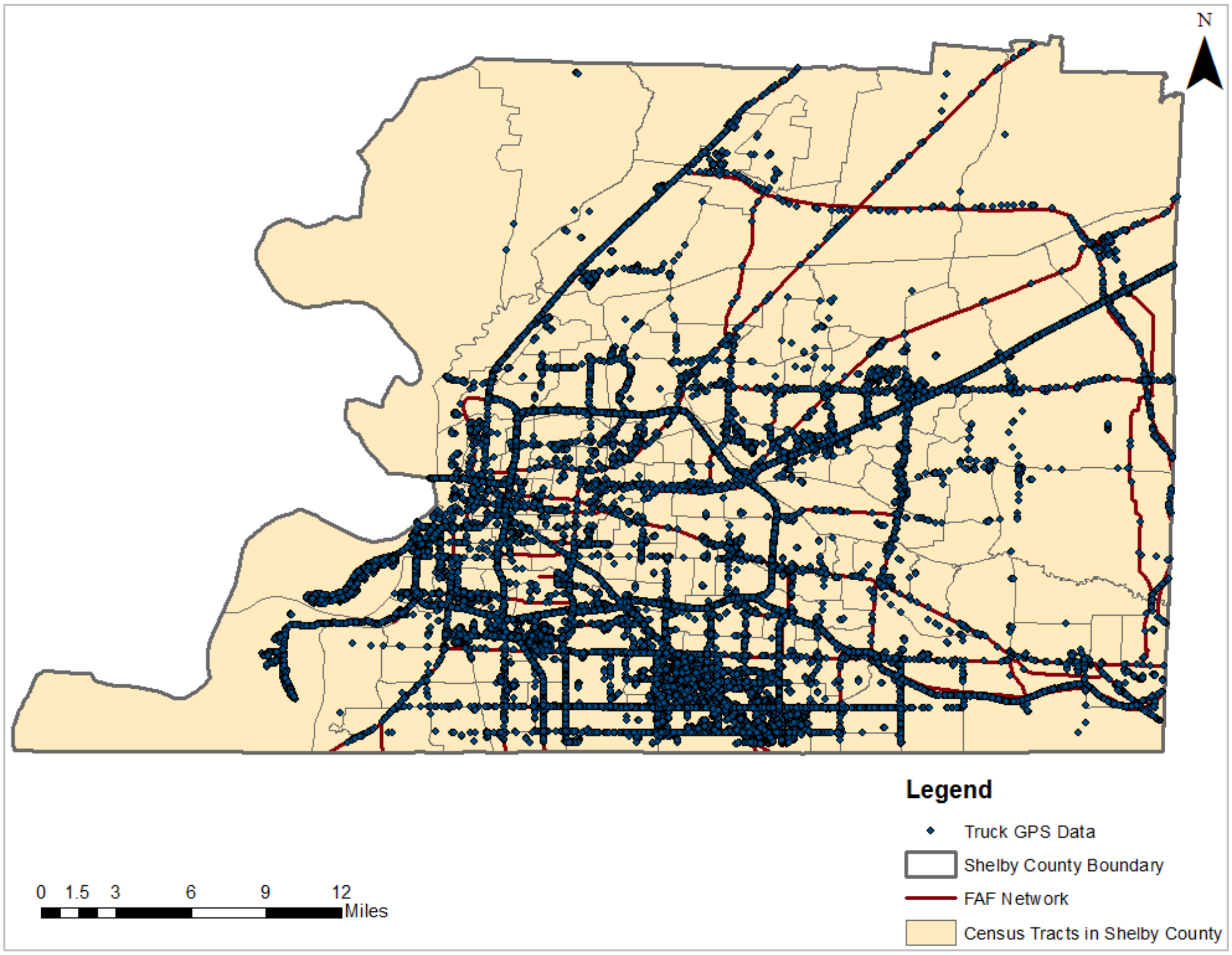


Figure 49 Shelby County with its census tracts, FAF network, and Truck GPS data

5.2.3 Travel time vs. Trip length

The comparison of travel time for different trip length in different travel conditions during weekdays for a particular interval is shown in Figure 50. The maximum distance between an O-D pair was approximately 40 miles. In ideal condition, a 40-mile trip can be covered in 34-36 minutes. If there is recurring congestion same trip length would require 46-48 minutes. During non-recurring congestion due to crash, it can take up to 72 minutes. Comparing different scenarios allowed use to capture the variability in travel time for a given O-D pair. Similarly, Figure 51 shows the comparison of travel time for different trip length in different travel conditions on weekends during a given period of time (7-8 a.m.). The free-flow condition doesn't vary between weekdays and weekends since we consider single value of free-flow time for each link. If there is recurring congestion, same trip length of 40 miles would require 40-42 minutes. During non-recurring congestion due to crash, it can take up to 52-58 minutes. It can be noticed that during the weekends the variability is significantly smaller compared to weekdays. Similar results can be obtained for different time of day and the findings are similar.

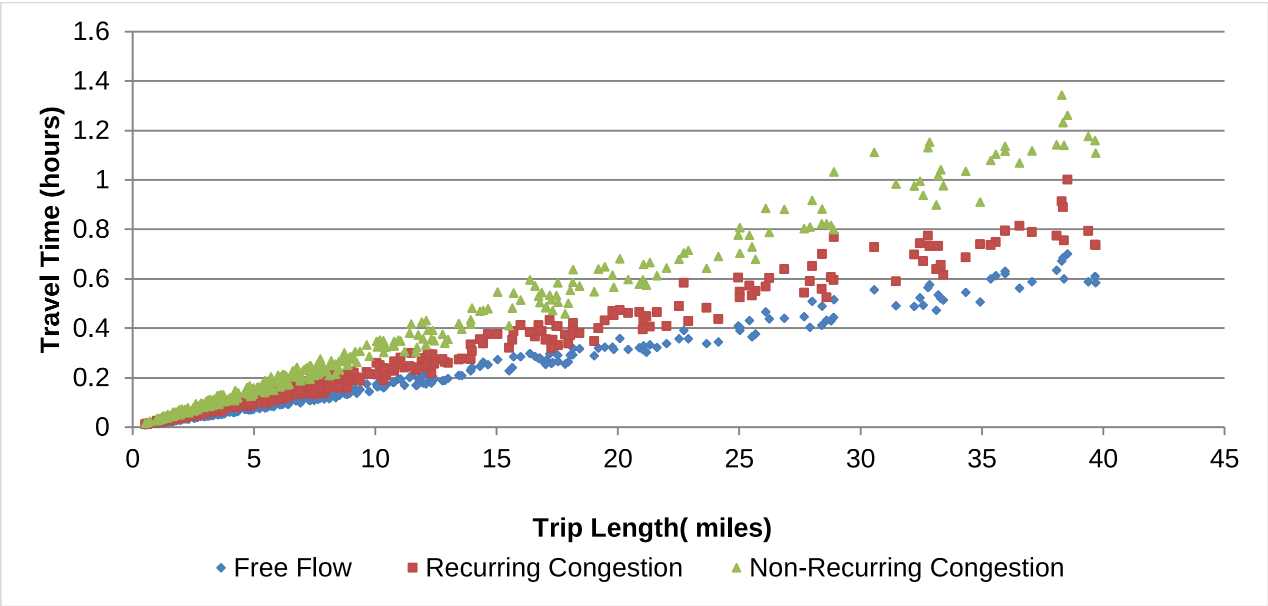


Figure 50 Travel Time vs. Trip Length (on Weekdays) during 7-8 a.m.

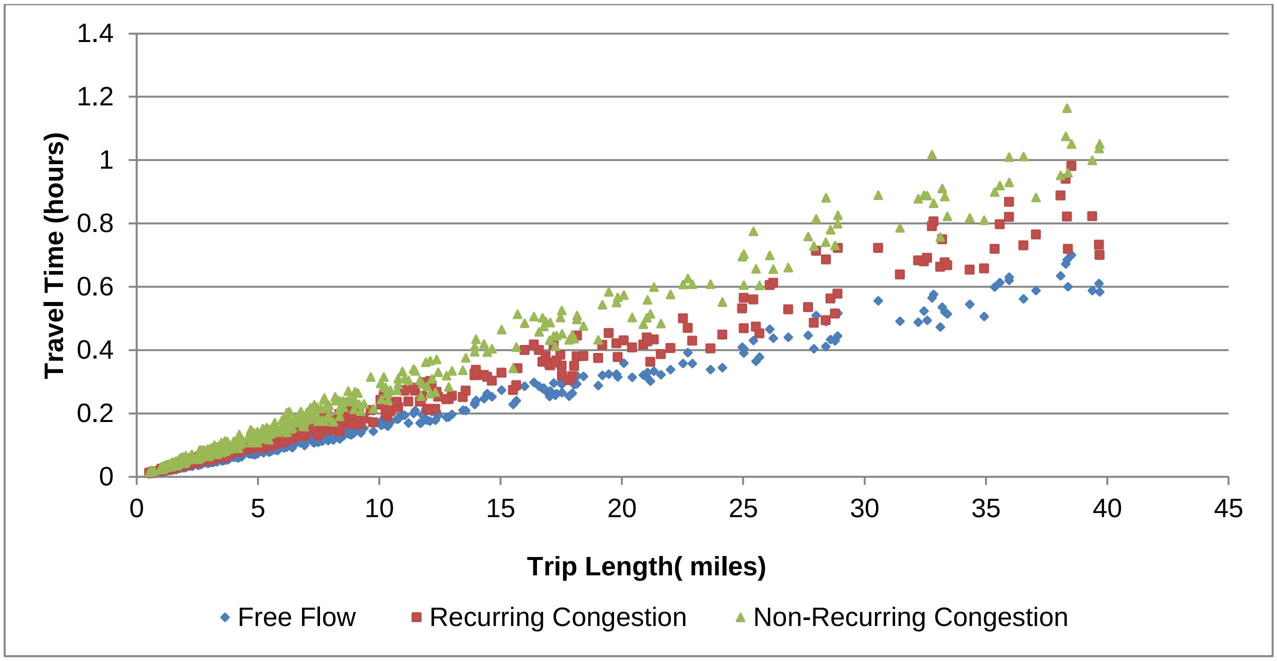


Figure 51 Travel Time vs. Trip Length (on Weekends) during 7-8 a.m.

5.2.4 Standard Deviation vs. Time of Day

To capture the variability over time of day, standard deviation of travel time and 95 percentile travel time is plotted against the time of day. Figure 52 and Figure 53 show the variation due to recurring and non-recurring congestion respectively. It can be seen that there is significantly larger deviation during non-recurring congestion compared to recurring congestion. Two noticeable peaks con be observed suggesting relatively large standard deviation during mid-day (11 am – 12pm) and PM peak hours (5pm-7pm). Standard deviation of 95% travel time also follows the similar pattern but with much larger variation among the observations.

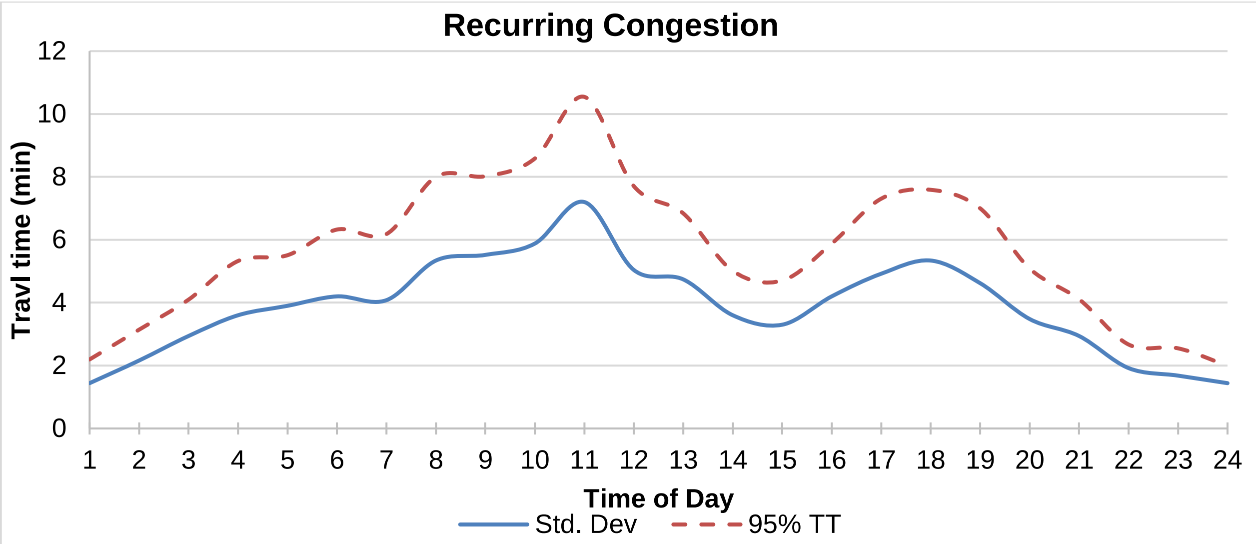


Figure 52 Standard Deviation (Recurring) by Time of Day

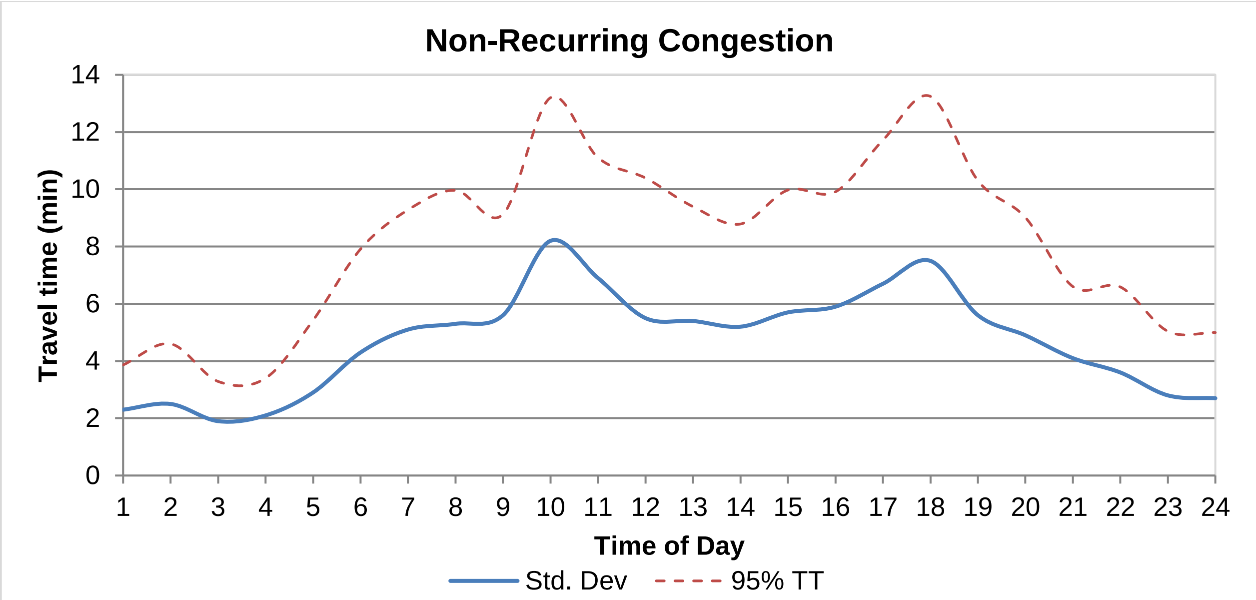


Figure 53 Standard Deviation (Non-recurring) by Time of Day

5.2.5 Effect of Crash Severity on Travel Time

The degree of effect of a crash on non-recurring congestion significantly depends on it severity. For June 2014, there were total 103 crashes that occurred on the given FAF network in Shelby county. 82 were property damage only (PDO), 19 crashes resulted in some injury and only 2 crashes resulted in incapacitating injury. There were no fatal crashes. It should be noted that, a single crash can effect multiple O-D pair as mentioned in the methodology that once a crash occurs on a link, all the O-D pairs which involve that particular link are considered to be affected. From Figure 54, it can be observed that, crashes involving injury have significant effect on travel time comparing to PDO crashes only. It is found that, 53% of the time a crash involving injury would increase the travel time significantly compared to PDO. It should also be noted that effect of crash does not only depend on severity, but it also depends on number of vehicles involved, incident type, clearance time etc.

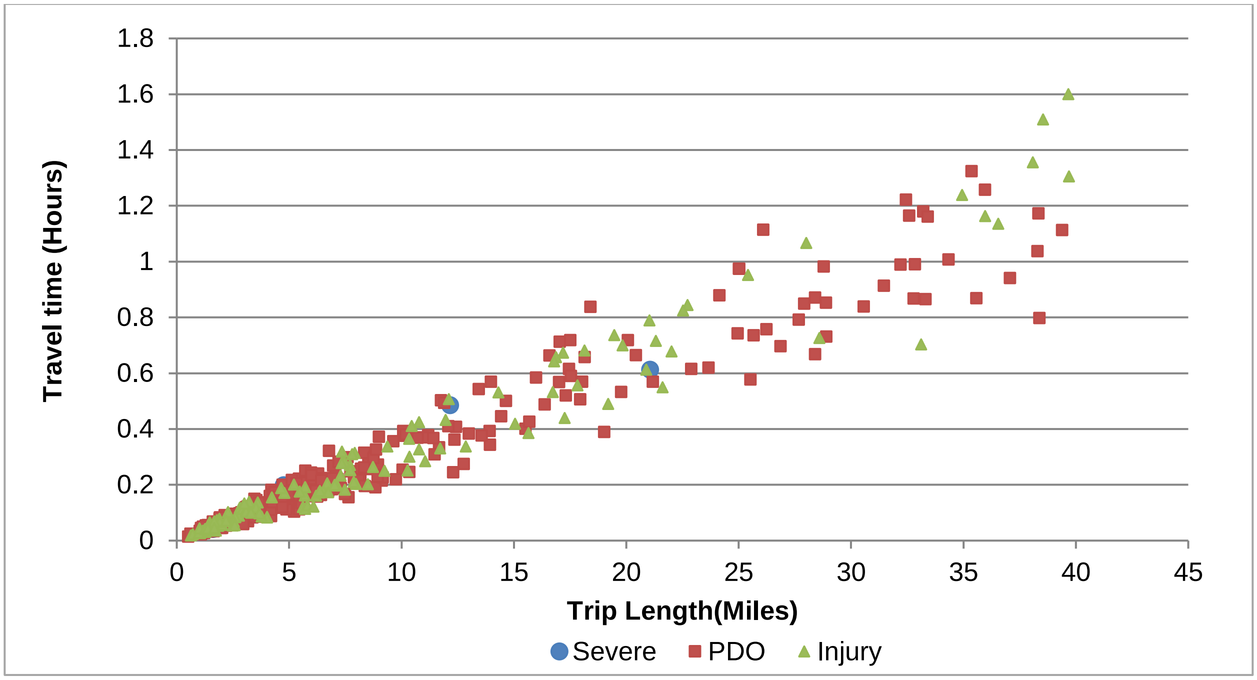


Figure 54 Effect of Crash Severity on Travel Time (Non-Recurring Congestion Only)

5.3 Analysis of Major Stop Locations by States

5.3.1 Data Collection

In this section we present the methodology for the data preparation. First we present the steps to identify the truck parking rest areas for the case study area. Second, we explain the procedure for extracting the polygon area of a rest area sample and how to utilize these polygons to extract the truck parking counts. Finally, we explore the additional dataset that was collected for analysis and model development.

5.3.2 Identification of Rest Areas

The first step is to identify the truck parking areas for the study area. This is done by obtaining the USA rest area map that features the locations of over 2000 highway rest areas across the U.S. The locations are the coordinates for the entrance to the off ramp for the rest areas. From this shape file, the rest areas that belong to the required study area were obtained. A total of 46 rest areas are found within the case study area shown in Figure 55. Once the rest areas were identified, the base map of the U.S. with imagery was loaded and the rest area locations were identified. Next, three types of polygons were created over the rest area of each location which includes parking area, off-ramp and on-ramp (shown in Figure 56).

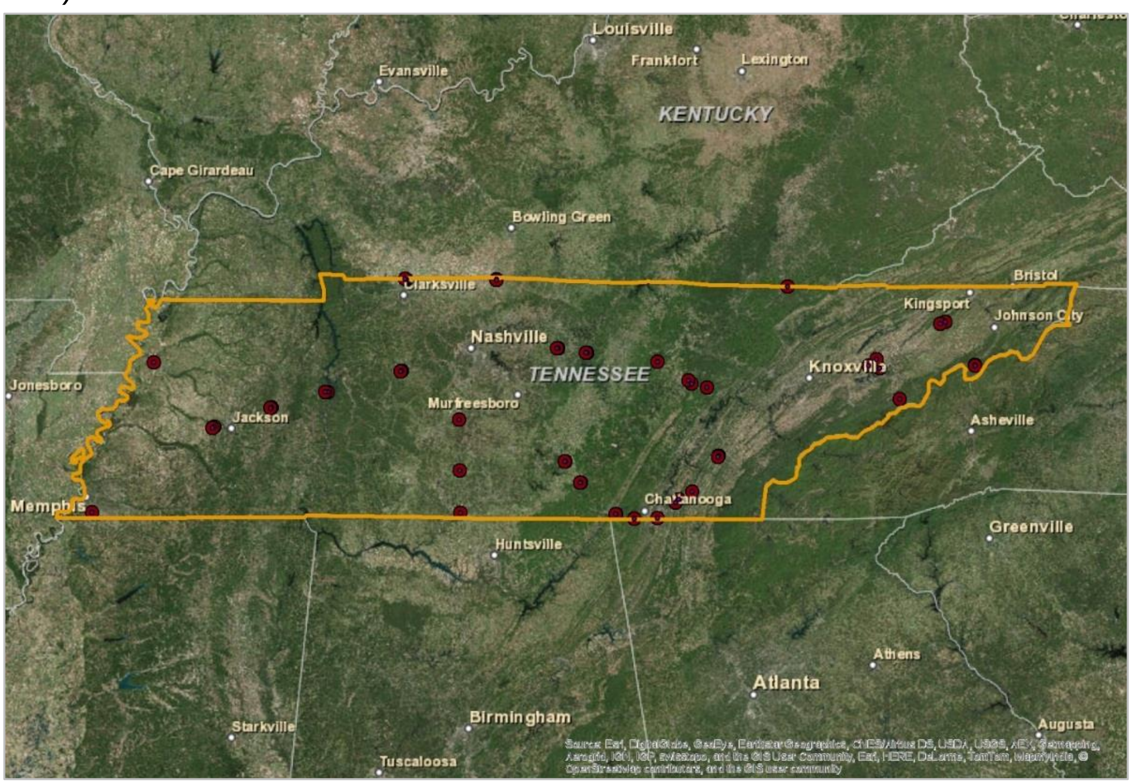
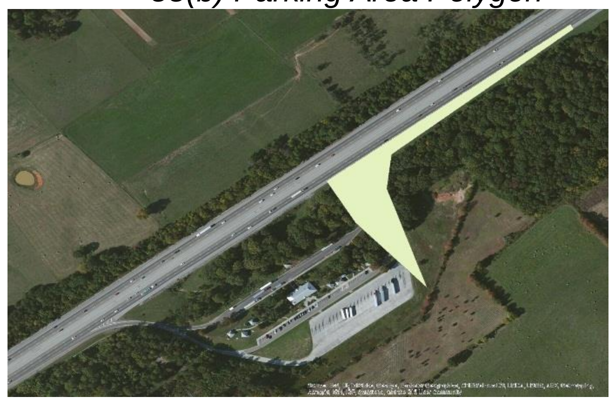


Figure 55 Rest Areas in Tennessee

55(a) Sample Rest Area Location

55(b) Parking Area Polygon

55(c) Off-ramp polygon



55(d) On-ramp Polygon

Figure 56 Polygon Extraction from Rest Area Location

5.3.3 Identification of Parked Trucks

The next step is to determine the number of trucks parked within the three polygons of a particular rest area (shown in Figure 56). The major steps for this procedure are as follows:

Step 1: First, the truck GPS data (shown in Figure 55(a)) from ATRI (American Transportation Research Institute) was processed for the study area using Vehicle Probe GPS Data Processing Tool. This tool provides functions to produce refined set of data from a large data set to be used in other modules.

Step 2: The refined truck dataset was loaded in ArcGIS and trucks with the speed being less than or equal to five miles per hour were identified using SQL tool embedded in the attribute table in ArgGIS.

Step 3: The stopped trucks data were exported in Microsoft Excel and then coordinate pairs (latitude and longitude combinations) of those stopped trucks were identified using a common identification number called TRUCKID.

Step 4: Once the unique stopped trucks were identified using the TRUCKID, their coordinates were projected in ArcGIS.

Step 5: Finally, using the rest area polygons, the number of trucks parked in the respective locations were identified and recorded (shown in figure 55(b).

The trucks parked on the on ramp and off ramp were needed to find the on/off ramp violation criterion which is used as a categorical variable in the model presented. These are indicator variables that indicated whether the truck was parked on ramp or off-ramp at that particular time period. When the truck drivers see trucks parked off-ramp, in order to avoid congestion and save time, sometimes they choose not to park thinking that the parking space might be full. It is also difficult to maneuver trucks when plenty of them are parked off ramp. On the other hand, when the truck drivers park on ramp, it means either the parking space was full or to save time maneuvering from the parking spot, they park on ramps so that they can exit and hit the road as quickly as possible. Moreover, some of the trucks might have met with crash or collision due to which they might park on and off ramp.

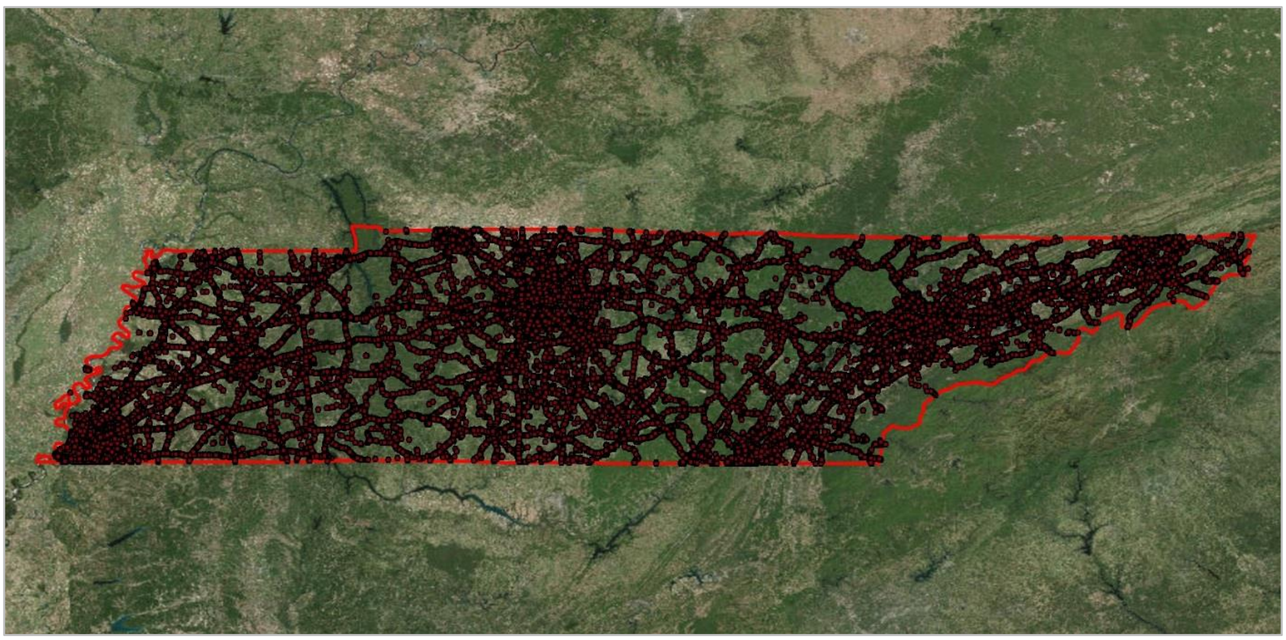


Figure 57 Sample GPS Data

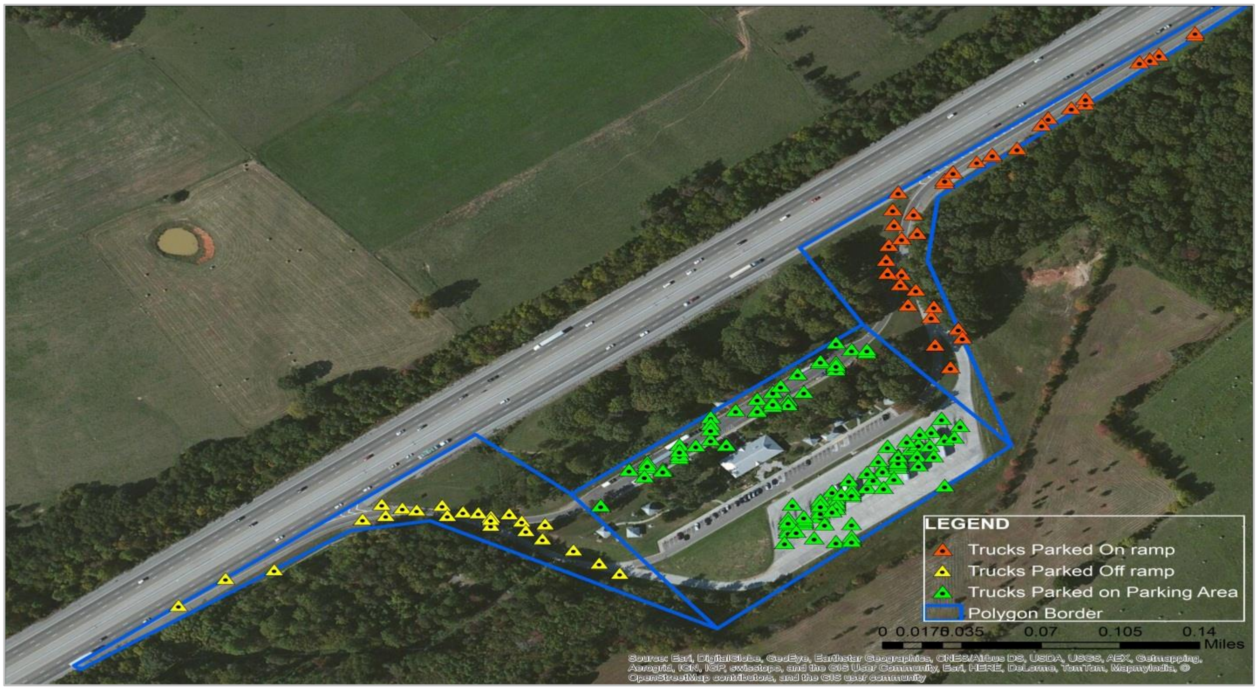


Figure 58 Parked Trucks on Rest Area Location

The ATRI data used for this study consisted of observations for 3 months in the year 2014. Each month comprised of two weeks of truck data. A total of 46368 (i.e. 46x24x42) set of observations were collected, each of the 46 locations with 24 hours and for 42 days. Table 15 shows the frequency distribution of the number of trucks parked on all the 46 rest areas which shows the predominance of zero number of trucks parked at the locations (50%). This is because of the low percentage of observations provided.

Table 15 Frequency Distribution of Truck Utilization

No. of Parked Trucks	Count	(%)
0	23,185	50.00%
1	7,666	16.53%
2	7,659	16.52%
>= 3	7,858	16.95%
Total	46,368	100.00%

5.3.4 Collection of Additional Data

Additional data was collected such as average speed of the truck traffic passing on the adjacent roadway of that particular rest area, the number of lanes of the roadway adjacent to the rest area, and hourly precipitation of the location from the National Climatic Data Center (NCDC). Rest area characteristics such as availability of rest rooms, vending machines, pets' facilities, picnic tables, phone services and handicapped facilities were also collected. These variables were used as categorical variables where unavailability of these features indicated 0 and availability was indicated as 1. The reason why average speed is included is because there might be congestion at the road beside the rest area which might affect the truck parking. If the average speed of the trucks passing the rest area is lower, it indicates congestion and hence might affect truck parking. Hence the effect of average speed was included in the model. The hourly precipitation would help identify if precipitation has any significant effect on the truck utilization, that is, if the truck drivers prefer to park at the rest area during rain or continue to drive. Similarly, the rest area characteristics would also give valuable information as to whether these have any effect on the truck utilization. However, most of these variables were insignificant in the estimated models.

Table 16 and Table 17 shows the descriptive statistics and frequency distribution of the response and explanatory variables used in the model respectively. These are, (a) Speed – whether the average speed of the trucks passing the rest area location is 65 miles per hour or greater. 65 miles per hours was chosen because it is the speed limit of highways in most cases. (b) Number of lanes – whether the number of lanes of the roadway adjacent to the rest area in consideration were 2 or more than 2 (c) On ramp – whether there were any trucks parked on the on ramp of the rest area during that time period (d) Off ramp – whether there were any trucks parked on the off ramp of the rest area during that time period (e) Days of the week when the trucks were parked at the rest area. Table 2b shows that about 99.7 % of the time, the trucks passing the rest area usually travel at an average speed of 65 miles per hour or less. In addition, it can be seen that the on ramp and off ramp violation are evenly distributed and 82.62% of the observations had trucks parked at the rest area adjacent to the 2 lane roadway. Also, about 71.43% of the trucks were parked during weekdays.

Table 16 Descriptive Statistics of Explanatory Variables

Variable	Mean	Std. Dev.	Minimum	Maximum
Number of trucks parked	1.0043	1.1596	0	7
Volume (Vph)	32.0887	18.9001	1	259
Average speed (mph)	23.0612	10.2253	6	84
On ramp	0.5016	0.5	0	1
Off ramp	0.4991	0.5	0	1

Table 17 Frequency Distribution of Explanatory Variables (Categorical)

Explanatory Variable		(%)
Speed	Less than or equal to 65 mph	99.70%
	Greater than 65 mph	0.30%
Number of Lanes	2 Lanes	82.61%
	> 2 Lanes	17.39%
On ramp	No	49.84%
	Yes	50.16%
Off ramp	No	50.09%
	Yes	49.91%
Weekday	Mon-Fri	71.43%
	Sat-Sun	28.57%

5.3.5 Model Estimation Results

In this section, we first discuss the effects of explanatory variables on the number of trucks parked, then the variable effects on the propensity and on the thresholds that affect the translation of propensity to whether or not a truck is parked at any given time. Next we discuss the model fit comparisons and finally discuss the elasticity effects and model validation.

The results section presents the statistically significant explanatory variables along with their estimated coefficients and t-statistics (in parenthesis) for each of the developed models as shown in Table 18. Four models are developed: Poisson (Model 1), negative binomial (NB) (Model 2), Poisson with propensity (Model 3), and Poisson with propensity and threshold specific constant (Model 4). Given that there is no a priori reason for the mean and variance in any practical context to be equal, the use of a NB distribution for Model 2 is an important empirical generalization over the Poisson distribution (Model 1). Model 2 is a regular NB model whereas Model 3 and 4 are Poisson models that include threshold parameters which take heterogeneity across observations into account by allowing some of the parameters to vary across observations. Model 4 is similar to Model 3 except it contains threshold specific constants to allow more flexibility and better predictive accuracy.

By comparing the model results, it can be observed that due to significance of the dispersion parameter, the negative binomial model is more effective in prediction than the Poisson model. However, as the variable *“number of lanes (=2)”* is added to the propensity equation, there is no longer dispersion in the model. Therefore, the negative binomial is collapsing back to Poisson model with Propensity. The dispersion parameter becomes large (implying low dispersion) with the variable *“number of lanes (=2)”* in the propensity. Hence, the model is able to explain variance using explanatory variables without dispersion parameters and therefore Poisson with propensity performs better than the NB model. Hence Model 3 and Model 4 are Poisson with propensity and propensity with threshold respectively.

The variables that have significant effect on truck parking utilization include *truck volume*, *on ramp, off ramp, average speed, number of lanes*, *hour 1*, *hour 2*, *hour 6*, *hour 13*, *hour 19, hour 21,* and *Thursday.* It can be noticed that the mean values of parameter estimates are similar in sign in all the four models. The results indicate that higher truck volume is more likely to increase the truck parking utilization at the rest areas. The *on ramp* and *off ramp* variables indicate that with on ramp and off ramp parking violation, the truck parking utilization will likely decrease. This is intuitive because when truck drivers enter a rest area for parking and when they observe trucks being parked on the ramps, it becomes difficult for them to maneuver the vehicle and they will likely avoid such rest areas. Moreover, they may also assume that the truck parking might be full and continue driving to find the next location. On the other hand, it can also be instinctive that the parking capacity must be full which lead the drivers to park on the ramps. It is also conceivable that the truck drivers park on the on ramps for easy and quick exit on the roadway. The *speed* indicates that with the average speed of trucks passing the rest area being equal to or lower than 65 mph is more likely to decrease the truck parking utilization on the rest area. The *number of lanes* indicate that the roadway adjacent to the rest area having two lanes have more likelihood to increase the truck parking utilization at the rest area because higher number of lanes is usually accompanied with high traffic flow. The positive coefficient of *hour 1* suggests that truck parking utilization increases during the period 12 AM – 1 AM which is intuitive since the truck drivers at this time try to find spots for parking and resting. On the other hand, the results show negative coefficients for *hour 2, hour 6, hour 13, hour 19 and hour 21*. This suggests that during the period 1 AM – 2 AM (*hour 2*), truck parking utilization reduces which is also intuitive since most of the parking spots are usually filled during the previous hour and the truck drivers spent more time searching for a spot. They either park on the ramps or try to move ahead and search a spot at the next location. During the period 5 AM – 6 AM (hour 6), truck parking utilization decreases since most truck drivers start their trip either early or during this period from the parking spot after having a good night’s rest. Similarly, during the period 12 PM – 1 PM (*hour 13*), truck parking utilization reduces because the truck drivers usually stop at a gas station for food and gas since these facilities are usually not available at a rest area. Finally, during the period 6 PM – 7 PM (*hour 19*) and 8 PM – 9 PM (*hour 21*), the utilization reduces which is also intuitive since truck drivers like to drive during these periods and do not stop and park at the rest area unless encountered by an emergency like mechanical fault or accidents because they usually like to rest at late night periods mostly after 12 am as indicated by *hour 1*.

Threshold parameters

The threshold parameters include the threshold specific constants ($\boldsymbol{\alpha_k}$ values), as well as variables associated with *off ramp* and *Number of lanes (=2)* as part of $\boldsymbol{\gamma}$ vector. The thresholds are responsible for the ''instantaneous'' translation of the truck parking utilization propensity to whether or not the truck driver will park at any given time at any location (that is, they determine the mapping of the latent propensity to the observed count outcome). The threshold specific constants ($\boldsymbol{\alpha_k}$) do not have any substantive interpretations. However, their presence provides flexibility in the count model to accommodate high or low probability masses for specific outcomes. The $\boldsymbol{\alpha_k}$ parameters are identified by specifying $\boldsymbol{\alpha_0} = 0$ and $\boldsymbol{\alpha_k} = \boldsymbol{\alpha_k}\ \forall k \geq K$. We, initially set K=4 and with multiple trials K is reduced based on statistical significance and general data fit.

The elements in the $\boldsymbol{\gamma}$ vector are presented next in Table 18. For the other variables, a positive coefficient shifts all the thresholds toward the left of the truck parking utilization propensity scale, which has the effect of reducing the probability of zero trucks parked. The effect of *off ramp* suggests that, given two observations with same truck parking utilization propensity, the segment with off ramp violation is more likely to have a non-zero truck parking utilization occurrence compared to the other. This is an intuitive result since off ramp violation will likely mean truck parking area is full which may not be the case. The effect of *number of lanes (=2)* indicates an increase in non-zero truck parking utilization at rest areas adjacent to two lane roadway, for a given truck parking propensity. That is, the translation of probability into the occurrence of truck parking is elevated for 2 lane roadway adjacent rest areas, most likely because it is easy for truck drivers to enter a rest area adjacent to two lanes where as they may need to change multiple lanes, or maneuver in a different direction before entering a rest area which is difficult and risky.

5.3.6 Model Selection and Statistical fit

The Generalized Poisson count model with propensity and threshold is superior to the other models, as should be clear from the highest log-likelihood value and presence of several additional statistically significant coefficients in Table 18. However, all the models developed in this study were compared formally using the Bayesian Information Criterion (BIC) that penalizes models that attain better fit at the cost of additional parameters. According to the BIC criterion, a model with lower BIC value is preferred. It can be seen from the table that the Generalized Poisson model has the least BIC value of 115,427 among all models suggesting superior data fit. This underscores the importance of using

GORP model structures that provide additional flexibility to standard count models for analyzing count outcomes (for instance, parking utilization in the current empirical context).

Table 18 Model Results

Variables	Poisson (Model 1)	Neg. Bin (Model 2)	Poisson with γ (Model 3)	Poisson with γ and α (Model 4)
	Coefficient (t-stat)	*Coefficient (t-stat)*	*Coefficient (t-stat)*	*Coefficient (t-stat)*
Constant	-0.0684 (-2.818)	-0.0769 (-2.574)	0.0281 -1.539	-0.296 (-2.997)
Log(Volume)	0.0185 -3.194	0.0192 -2.458	0.0081 -2.481	0.0206 -2.763
On Ramp	-0.014 (-1.494)	-0.0139 (-1.193)	-0.0053 (-1.091)	-0.0164 (-1.468)
Off Ramp	-0.0093 (-1.000)		-0.1507 (-6.597)	-0.1064 (-2.108)
Average Speed (<= 65mph)				-0.1684 (-1.762)
Number of Lanes(= 2)	0.0279 -2.258	0.0316 -2.03	1.9014 -48.956	0.0942 (-1.436)
12 am – 1 am (hour 1)	0.0767 -3.421	0.0724 -2.518	0.0305 -2.541	0.0597 -2.193
1 am – 2 am (hour 2)	-0.0793 (-3.284)	-0.0797 (-2.649)	-0.0357 (-2.836)	-0.0875 (-3.025)
5 am – 6 am (hour 6)	-0.0384 (-1.621)	-0.0387 (-1.301)	-0.0176 (-1.415)	-0.0446 (-1.565)
12 pm – 1 pm (hour 13)	-0.0589 (-2.465)	-0.0591 (-1.976)	-0.0267 (-2.132)	-0.0595 (-2.077)
6 pm – 7 pm (hour 19)	-0.0251 (-1.065)		-0.0132 (-1.065)	-0.0358 (-1.264)
8 pm – 9 pm (hour 21)	-0.0366 (-1.546)	-0.0368 (-1.238)	-0.0153 (-1.233)	-0.0443 (-1.558)
Thursday	0.0382 -2.878	0.0349 -2.101	0.0145 -2.093	0.0347 -2.203
Saturday	0.0151			

		-1.131			
Dispersion parameter			1.6815 -42.1429		
β Vector	Off Ramp			0.3357 -7.048	0.0915 -1.94
	Number of Lanes (= 2)			-3.0612 (-43.086)	0.1218 -1.977
Threshold variables	α_1				-0.6194 (-58.338)
	α_3				1.3233 -7.61
Number of Observations		46368	46368	46368	46368
Number of Parameters Estimated		13	11	14	18
Log-composite likelihood at convergence		-1.41783	-1.38229	-1.36687	-1.2426
Log-likelihood		-65,741.94	-64,094.02	-63,379.03	-57,616.88
BIC		131,624	128,306	126,908	115,427

5.3.7 Elasticity Effects

The elasticity computed is a measure of the aggregate percentage change in the response variable due to a change in an exogenous variable *(36)*. By computing the elasticity effects of the exogenous variables, the magnitude of effects of these variables on the truck parking utilization can be determined. In this paper, we computed the percentage change in the expected number of trucks that park in a rest area because of a unit change in each exogenous variable. However, since standard elasticity calculations are not applicable to categorical variables, pseudo-elasticity effects were calculated for such variables. The pseudo-elasticity of an indicator variable essentially represents the average percent change in average truck parking utilization when the value of that particular variable is changed from 0 to 1 for all rest areas.

For brevity, the elasticity effects are only presented for the best model, *i.e.,* Poisson model with propensity and threshold parameters (model 4) (see Table 4). From the Table 19, it can be observed that the elasticity effects are consistent with the coefficient estimates. The first entry in the table indicates 100% increase in truck volume is likely to increase truck parking by 2.001% whereas presence of on ramp and off ramp violations decrease truck parking utilization by 1.545% and 0.85%, respectively. One additional lane to an existing two-lane roadway increases parking utilization in adjacent rest areas by 3.371%. Parking areas that are adjacent to roadways with average truck speeds greater than 60 mph have 14.484% lower utilization than parking areas adjacent to roadways with lower truck speeds. Similarly, parking areas have 5.792% more utilization during the hour past midnight compared to other non-peak rest hours Also, interestingly; parking utilization on Thursdays is 3.39% more than on other days of the week. Other numbers in the table can be interpreted similarly.

Table 19 Elasticity Effects of the Generalized Poisson Model

Variables	Poisson with γ and α (Model 4)
Volume (100% increase)	2.001
On Ramp	-1.545
Off Ramp	-0.85
Average Speed (<= 65mph)	-14.484
Number of Lanes (= 2)	3.371
12 am – 1 am (hour 1)	5.792
1 am – 2 am (hour 2)	-8.041
5 am – 6 am (hour 6)	-4.165
12 pm – 1 pm (hour 13)	-5.525
6 pm – 7 pm (hour 19)	-3.354
8 pm – 9 pm (hour 21)	-4.137
Thursday	3.339

5.3.8 Model Validation

In order to examine the prediction power of the models, a validation exercise was undertaken in which the predicted truck parking counts were compared with the observed count in the data. Then, Absolute Percentage Difference (APD) between predicted and observed counts was calculated. Lastly, Average Absolute Percentage Difference (AAPD) across all truck utilization levels was computed. A model with lower AAPD has better predictive ability than models with higher AAPD values. The results of the prediction analysis are presented in Table 20. It can be seen from the table that the Generalized Poisson model with propensity and threshold parameters (model 4) best has the best predictive performance with an AAPD value of 0.57%. The simple Poisson model that ignores dispersion has a very high AAPD value. Even NB model that accounts for dispersion has higher AAPD value than the Generalized Poisson model. So, Poisson model with propensity and threshold parameters is better suited to capture dispersion in count data than NB model in the context of truck parking utilization.

Table 20 Model Validations

Truck Utilization	Observed Count	Expected Count							
		Poisson (Model 1)		Neg. Bin (Model 2)		Poisson with γ (Model 3)		Poisson with γ and α (Model 4)	
		Count	APD (%)	Count	APD (%)	Count	APD (%)	Count	APD (%)
0	23185	17038	26.5	27124	17	21372	7.82	23307	0.53
1	7666	16997	122	8644	12.8	12434	62.2	7641	0.33
2	7659	8606	12.4	4336	43.4	7107	7.21	7555	1.36
3 or more	7858	3727	52.6	6264	20.3	5455	30.6	7865	0.09
AAPD (%)			53.3		23.4		22.8		0.57

6. WORKFORCE DEVELOPMENT ACTIVITIES

A central issue that MPOs and DOTs face is a lack of stakeholder involvement, understanding, and interest in transportation projects and investments. This poses a challenge for determining how well infrastructure and investments serve a community. Another issue faced by the transportation industry as a whole is the impending retirement of a large percentage of the workforce and the limited pipeline of workers electing to pursue transportation careers. Efforts to educate the public on the importance of the transportation system and the need for increased numbers and diversity of people pursuing careers in transportation are crucial for impacting transportation system and workforce challenges. MAP-21 requirements and freight performance measures provide an excellent opportunity to convey transportation engineering and planning concepts to students and teachers. Thus, for this project, a series of modules was developed to introduce high school students to transportation concepts and to help them better understand the importance and impact of freight and the use of performance measures for assessing the transportation system.

6.1 Curriculum

The curriculum developed through this project was implemented with a group of 20 high school students (10 males, 10 female) as part of the 2016 Transportation Academy hosted at the University of Memphis July 11-15, 2016. This program was a week long, full-day program designed to:

- Expose students to transportation engineering and planning concepts
- Familiarize them with the US transportation system and its impact on our communities and economy
- Provide students with understanding of educational pathways in transportation
- Introduce students to a variety of transportation career opportunities.

The program included mini-lectures on a variety of transportation topics (including focus on freight) followed by brainstorming or problem solving activities, hands-on group challenge projects, industry field trips, classroom visits with industry professionals, and team research projects. A freight story map[8] was developed using ESRI ArcGIS software as part of this project, and provided students with an interactive exploration of the impact of freight on the economy and community in Memphis. This session was followed by a mini-lecture and brainstorming activity on transportation system performance measures.

[8] https://uw-mad.maps.arcgis.com/apps/MapJournal/index.html?appid=510af627bbb448aa9aec9978e28ca2f3

6.2 Assessment

Program entrance (administered on July 11, 2016) and exit surveys (administered on July 15, 2016) assessed student's interests and shifts in perceptions over the course of the week. A parent survey was also emailed to all participants with the program acceptance packets in order to determine parents' perceptions related to the program and goals for their children. Parents were asked to complete and return the survey on the first day of the program. Eighteen parents responded to the survey and indicated that the primary reasons they encouraged their children to attend the program were to help them learn about career opportunities in engineering and science related fields (94%), to keep students engaged in learning activities over the summer (67%) and to expose students to a university environment (56%). Interestingly, only 4 parents (22%) reported that it was their child's idea to attend the program. Parents indicated that they would like their children to pursue an engineering or science career because of career variety (67%), rewarding career opportunities (61%), and because they believed their student would enjoy and has the potential to be successful in a career in engineering or science (55%). Two parents indicated they might not support their child's decision to pursue an engineering or science major because they do not know enough about these disciplines.

Twenty students responded to the program introduction survey and eighteen responded to the exit survey. When asked whether they were interested in majoring in computer/technology, engineering, science or math related majors, response percentages remained the same and generally high between the pre and post surveys, with the exception of computer/technology majors where 50% of students indicated interest prior to the program and 72% expressed interest at the end of the program. Students were also asked to respond to a series of perception questions on both surveys. Students indicated agreement with the following statements (% introduction survey/% exit survey):

- There are good jobs available for people with degrees in STEM (100%/100%)
- Transportation professionals make good money (88%/85%)
- Girls can do just as well as boys in transportation and engineering jobs (75%/89%)
- Being a science or engineering professional in transportation would be a fun job (35%/72%)
- I believe I have the ability to work in a STEM field (75%/89%)

The greatest shift in perceptions was seen related to 'Being a science or engineering professional in transportation would be a fun job.' Additional questions were included on the exit survey to further gauge perceptions of the transportation industry and students' understanding of and interest in transportation career pathways. On the exit survey, students expressed a belief that transportation professionals make a positive impact on society (94%), transportation professionals do important work (94%), transportation professions are challenging (89%), the program helped them develop better problem-solving skills (83%), and the program did a good job of showing how science, math, and engineering are used in transportation (89%). Students also indicated the activities, field trips, and speakers helped them better understand the transportation industry (94%), how to pursue a transportation career (94%) and made them more interested in pursuing a transportation career (72%). These results and shifts in perceptions are important, as 72% of students also reported that programs such as the Transportation Academy are influential factors in determining the college major they will pursue after graduating from high school.

6.3 Dissemination

A LiveBinders site[9] hosts curriculum developed through University of Memphis summer outreach programs since 2011. This site is used for professional development with teachers throughout the West Tennessee region, and the activities are freely available for download and use in K-12 classrooms. The materials developed for the 2016 Transportation Academy were added to the site, and will also be used in a series of professional development engagements for high school teachers in the spring of 2017. The curriculum developed through this project will be further disseminated through the West Tennessee STEM Hub, the Tennessee STEM Innovation Network, and the Southeast Transportation Workforce Center.

[9] **http://www.livebinders.com/play/play?id=2011056**

7. CONCLUSION

Moving Ahead for Progress in 21st Century (MAP-21) recognizes the importance of freight and advocates for a national and state strategic freight plans to assess the condition and performance of the national, state, local and regional freight networks. Adequate disaggregate data and appropriate methods of analysis are paramount to develop such performance measures. With the unavailability of accurate GPS data, or other data sets, literature is limited on developing freight performance measures (FPMs). However, identification and computation of FPMs is one of the imperative goals of MAP-21. MAP-21 identifies a freight plan to address freight congestion bottlenecks, identify critical major intermodal centers to enhance connectivity, determine barriers to improved freight performance, and explore the critical sections of the transportation network that need prioritization in resource allocation to enhance FPMs.

In this study, the researchers demonstrated how truck GPS data can be used to estimate FPMs for transportation networks and freight transportation facilities, evaluate performance of freight corridors, identify inter- and intra-truck trips, and analyze individual truck trip patterns. A number of algorithms were developed to process truck GPS data and develop freight performance indicators. Validation of the algorithms was based on link travel speeds available through the FPMweb Tool, Google maps and satellite images. The algorithms are combined to develop two tools. The first tool allows the user processing raw truck GPS data and extracting the data for certain time periods. The second tool enables the user to load the processed GPS data and calculate a wide range of performance measures within the ArcGIS environment.

The study results showed that accuracy of developed algorithms can be improved if more GPS records are available and more frequent GPS signal is provided. Truck trip analysis also requires additional information (i.e., location of freight facilities, rest stops, pick-up/drop-off business locations, commodity data, etc.). One of the main obstacles of using the available GPS dataset was the large size which prohibited processing of long time periods at a time (e.g. month). The following practical recommendations can be provided to public agencies for processing these large size GPS datasets:

1. Use more advanced CPU (i.e., recent processor, more RAM, multiple cores, etc.)
2. Partition the data in smaller portions based on:
    - time of the day: AM, MD, PM, and OP
    - specific areas of the region under study
    - special characteristics (e.g., freight corridors, major metropolitan areas, etc.)

3. Parallel machine processing – use all available CPUs for processing the given dataset. For example, if there are four CPUs available, each one can be assigned for processing AM, MD, PM, and OP periods of a given day respectively.
4. Gather more recent year truck GPS data, as the representative samples are higher.
5. Use true shape network as opposed to a stick networks used for sketch planning.
6. Adequate caution should be considered when the geographic boundaries consists of multiple time zones.

The truck GPS data used in this project spanned for three states (TN, MS, and AL). FPMs were developed for all three states. The tools developed were tested to ensure that it works under various geographical scales and different network sizes. The proposed methodology and the developed GIS application can be efficient in supporting MAP-21 goals for planning agencies within their short and long range planning efforts by providing network performance measures. Outcomes of this research may be used in development, calibration, and validation of transportation planning and travel demand models as well as assistance in traffic operation and planning for operations.

REFERENCES

1. Ando, N. & Taniguchi, E. Travel Time Reliability in Vehicle Routing and Scheduling with Time Windows. *Netw Spat Econ* (2006) 6: 293–311.

2. Barker, M., and Chen, C. *Freight Planning with TRANSEARCH Data*. Tennessee Model User Group, 2008.

3. Bassok, A. McCormack, E., & Outwater, M. *Use of Truck GPS Data for Freight Forecasting*. Transportation Research Board Annual Meeting, 2011.

4. Battelle. *FAF Freight Traffic Analysis*. Final draft report submitted to Oak Ridge National Laboratory, 2011.

5. Bierlaire, M., Chen, J., & Newman, J. A probabilistic map matching method for smartphone GPS data. *Transportation Research Part C* 26 (2013) 78–98.

6. Blazquez, C. A Decision-Rule Topological Map-Matching Algorithm with Multiple Spatial Data. *Global Navigation Satellite Systems: Signal, Theory and Applications*, Prof. Shuanggen Jin (Ed.), ISBN: 978 - 953-307-843-4, 2012.

7. Cambridge Systematics. *Quick Response Freight Manual II*. Report No. FHWA-HOP-08-010, 2007.

8. Carrion, C., & Levinson, D. Valuation of travel time reliability from a GPS-based experimental design. *Transportation Research Part C 35* (2013) 305–323.

9. Cheaters Spy Shop. *Cheaters CoPilot Real-Time GPS Tracker*. http://www.cheatersspyshop.com/cheaters-cruiser-copilot-real-time-gps-tracker-and-navigator. html. Accessed June 12, 2014.

10. Chien, S., Mouskos, K., Boile, M., Kim, K., & Golias. M. *Variability of travel times on New Jersey highways*. Department of Civil and Environmental Engineering, New Jersey Institute of Technology, 2011.

11. Chauvenet, W. A Manual of Spherical and Practical Astronomy. 5th ed. Dover, N.Y., 1960, pp. 474–566.

12. Cortes, C., Gibson, J., Gschwender, A., Munizaga, M., & Zuniga, M. Commercial bus speed diagnosis based on GPS-monitored data. *Transportation Research Part C* 19 (2011) 695–707.

13. Dong, J., & Mahmassani, H. Flow Breakdown, Travel Reliability and Real-time Information in Route Choice Behavior. *Transportation and Traffic Theory* (2009) 675-695, Springer Science + Business Media.

14. FHWA. (2011). *FPMweb Background and Goals*. https://www.freightperformance.org/ fpmweb/ default.aspx. Accessed 01/12/2014.

15. Figliozzi, M., Walker, L., Sarkar, S., & Rice, D. *Algorithms to Study the Impacts of Travel Time Reliability along Multi- Segment Trucking Freight Corridors*. 90th Annual Meeting of the Transportation Research Board, January 23–27, 2011.

16. Fisher, M., Outwater, M., Cheng, L., Ahanotu, D., & Calix, R. Innovative Framework for Modeling Freight Transportation in Los Angeles County, California. *Transportation Research Record*: Journal of the Transportation Research Board, No. 1906, Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 105–112.

17. Forbes. *The 10 most traffic congested cities in the world*. http://www.forbes.com/pictures/ ehmk45kigm/10-san-francisco-usa-2/. Accessed June 18, 2014.

18. Greaves, S. & Figliozzi, M. Collecting Commercial Vehicle Tour Data with Passive Global Positioning System Technology. Issues and Potential Applications. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2049, Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 158–166. DOI: 10.3141/2049-19.

19. Golias, M., Dobbins, J., Short, J., & Johnson, Z. *GPS/GIS Analysis of Tennessee Truck Trips*. Intermodal Freight Transportation Institute, University of Memphis. Final Report.

20. Golias, M., & Mishra, S. *Evaluating the hours of service rule via GPS/GIS truck trip data. Draft Report*. Intermodal Freight Transportation Institute, University of Memphis, 2013.

21. Huynh, N., & Walton, M. *Methodologies for Reducing Truck Turn Time at Marine Container Terminals*. Research Report 167830-1, 2005.

22. Jones, C., Murray, D., & Short, J. *Methods of Travel Time Measurement in Freight-Significant Corridors*. American Transportation Research Institute. Submitted to Transportation Research Board Annual Meeting, 2005.

23. Kuppam, A., Lemp, J., Beagan, D., Livshits, V., Vallabhaneni, L., & Nippani, S. *Development of a Tour-Based Truck Travel Demand Model using Truck GPS Data*. TRB Annual Meeting, 2014.

24. Liao, C. *Freight Performance Measure Systems (FPMS) System Evaluation and Data Analysis*. Department of Civil Engineering, University of Minnesota. CTS project number 2007076, 2008.

25. Liao, C. *Using Archived Truck GPS Data for Freight Performance Analysis on I-94/I-90 from the Twin Cities to Chicago*. Department of Civil Engineering, University of Minnesota. CTS project number 2009069, 2009.

26. Liao, C. (2014). *Using Truck GPS Data for Freight Performance Analysis in the Twin Cities Metro Area*. Department of Civil Engineering, University of Minnesota. CTS project number 2013035, 2014.

27. Maguire, A., Ivey, S., Golias, M., & Lipinski, M. *Relieving Congestion at Intermodal Marine Container Terminals: Review of Tactical/Operational Strategies*. 88th Annual Meeting of the Transportation Research Board, 2009.

28. McCormack, E. *Collecting GPS Truck Data*. Data and Tools for Linking Goods Movement, Air Quality, and Transportation Infrastructure Decisions, Irvine CA, workshop, 2009.
29. McCormack, E. & Hallenbeck, M. *Options for Benchmarking Performance Improvements Achieved from Construction of Freight Mobility Projects*. Washington State Transportation Center (TRAC), 2005.
30. McCormack, E., Scharnhorst, E., & Zhao, W. Using GPS truck data to identify and rank bottlenecks in Washington State. Washington State Transportation Center (TRAC). University of Washington, Agreement T4118, Task 01, 2011.
31. McCormack, E., & Zhao, W. (2011). *GPS truck data performance measures program in Washington State*. Washington State Transportation Center (TRAC). University of Washington, Agreement T4118, Task 31, 2011.
32. Memphis Urban Area MPO. *Freight Peer to Peer Program*. Memphis, Tennessee February 20th and 21st, 2013.
33. NCHRP Report 618. *Cost-Effective Performance Measures for Travel Time Delay, Variation, and Reliability*. Transportation Research Board, 2008.
34. NCHRP Report 008. *Freight-Demand Modeling to Support Public-Sector Decision Making*. Transportation Research Board, 2010.
35. Ortuzar, J., & Willumsen, L. *Modelling Transport*. Wiley, 4th Edition, 2011.
36. Pinjari, A., Short, J., Pierce, D., Park, L., Murray, D., Mysore, V., Tabatabaee, F., Irmania, A., Zanjani, A., & Koons, J. *Truck GPS Data for Freight Performance Measurement and Planning: Application, Issues, and Opportunities*. Florida Model Task Force, Freight Committee, 2012.
37. Pinjari, A., Short, J., & Tabatabaee, F. *Truck GPS Data for Freight Planning*. Florida Model Task Force, Orlando, 2012.
38. Pinjari, A., Short, J., Pierce, D., Park, L., Murray, D., Mysore, V., Tabatabaee, F., Zanjani, A., & Irmania, A. *Truck GPS Data for Freight Performance Measurement, Modeling and Planning*. Florida Model Task Force, Freight Committee, 2013.
39. Quiroga, C., & Bullock, D. Travel time studies with global positioning and geographic information systems: an integrated methodology. *Transportation Research Part C* 6 (1998) 101-127.
40. Quiroga, C. Performance measures and data requirements for congestion management systems. *Transportation Research Part C* 8 (2000) 287-306.
41. Schofield, M., & Harrison, R. *Developing Appropriate Freight Performance Measures for Emerging Users*. Center for Transportation Research, University of Texas at Austin. Report 473700-00073-1, 2007.
42. Storey, B., & Holtom, R. *The use of historic GPS data in transport and traffic monitoring*. *TEC.11.03*/p. 376-379 IT IS, 2003.

43. US DOT. *America's Freight Transportation Gateways 2009*. Report, November 2009.
44. You, S. *Methodology for Tour-Based Truck Demand Modeling*. Submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy, University of California, Irvine, 2012.
45. 24/7 Wall St. *Ten cities with the worst traffic*. http://247wallst.com/special-report/2013/05/01/ten-cities-with-the-worst-traffic/ Accessed November 4th, 2014.
46. Wang, W., Lin, C., Yin, K., Gong, Q., Adams, T., & Wang B. *Using Sparse GPS Data to Estimate Link Travel Time for Truck Transport*. Transportation Research Board Annual Meeting, 2014.
47. Wheeler, N., & Figliozzi, M. Multicriteria Freeway Performance Measures for Trucking in Congested Corridors. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2224, Transportation Research Board of the National Academies, Washington, D.C., 2011, pp. 82–93. DOI: 10.3141/2224-10
48. WSDOT. *Washington State Truck Freight Performance Measure Program*. Washington State Transportation Commission, March 22, 2011.

APPENDIX- A. MISSISSIPPI CASE STUDY

Data Overview for Mississippi

Building off data supplied by the Department of Transportation, we used global positioning system (GPS) transponders to calculate truck traffic throughout the state of Mississippi. The data released to us included data from 2011, 2012, 2013, and 2014 for the months of March, June, and October. The limitation on months was to reduce the overall data size and use these three months as a sample frame from the annual population of data. We inferred the remaining months for totals.

Our dataset consisted of 26.6 million data points in a comma-separated values (CSV) file. The CSV data contained latitude, longitude, segment, truck identification number, GPS read date, speed, heading, and State name. The data of interest for the Mississippi analysis was the state name of Mississippi, the latitude, longitude, truck identification number, speed, and heading.

Data Mining Tool

Utilizing a custom tool designed for ArcGIS by the University of Memphis Intermodal Freight Transportation Institute called the Direction and Outlier Identification Algorithm (DOI) written in Python language. The DOI looked at bi-directional Freight Performance Measures (FPMs) that addressed:

- VOL - Truck volume
- TS - Travel speed (miles per hour)
- TT - Travel time (hours)
- TTp90 - 90th percentile TT
- TTp95 - 95th percentile TT
- BTT - Buffer TT
- BI - Buffer index
- TTStD - TT standard deviation
- TTCV - TT coefficient of variation
- TTR - TT range
- MMR - mean to median ratio

The data were calculated using:

1. Link/trip/path/tour travel time - TT (min, hrs.)
2. Link/trip/path/tour travel speed - TS (km/hr., mi/hr.)

If more information is provided (i.e., location of freight facilities, purpose of stops, land use at stop location, location of pick up/delivery business locations, location of rest stops, etc.), additional trip/path/tour characteristics can be estimated: dwell time at freight facilities, stop purpose based on purpose of the preceding stop, traffic light stops (GPS signal of a high frequency is required → less than 3 min between consecutive observations), dwell time at rest stops, and others.

TT reliability/variability measures

1. 90th and 95th percentile travel time ($tp_{90\%}$ and $tp_{95\%}$)
2. Buffer index $BI = \frac{tp_{95\%} - \bar{x}}{\bar{x}}$

   where $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$ - mean travel time; x_i - travel time for the observation *i*; N – number of observations
3. Buffer travel time $BTT = tp_{95\%} - \bar{x}$ (minutes, hours)
4. Planning travel time $PTT = tp_{95\%}$ (minutes, hours)
5. Travel time standard deviation $\sigma = \sqrt{\frac{\left(\sum_{i=1}^{N} x_i - \bar{x}\right)^2}{N-1}}$
6. Travel time coefficient of variation $CV = \frac{\sigma}{\bar{x}}$
7. Travel time range $Range = x_{max} - x_{min}$
8. Ratio of mean travel time to median travel time $r = \frac{\bar{x}}{\hat{x}}$

   where $\hat{x}$ - median travel time If free flow travel time (or speed) is provided for links of the considered transportation network, additional FPMs can be computed:
9. Planning travel time index $PTTI = \frac{tp_{95\%}}{x_{FFS}}$

   where x_{FFS} – free flow speed travel time
10. Travel time index $TTI = \frac{x}{x_{FFS}}$
11. Total segment delay $TSD = (tp_{95\%} - x_{FFS}) \times V$ (vehicles-minutes)

    where V – volume of vehicles at the segment
12. Congested travel $CT = \sum ConLength \times V$ (vehicles-miles)

    where $ConLength$ – congested segment length
13. Congested roadway $CR = \sum ConLength$ (miles)

These data were divided into hourly totals, daily totals, and monthly totals. The individual truck identification numbers were interpolated to a median point by hour to eliminate multiple pings from the GPS trackers. These individual truck speeds were averaged by hour.

Data Results

Accounting for individual, non-interpolated truck volume in Mississippi for the years 2011-2014 and in the months of March, June, and October showed that with 4.2 million in 2011, 6.1 million in 2012, 7.3 million in 2013, and 9 million in 2014 (Table 21). The weekly average for all years and months were 1.8 million on Sunday, 4 million on Monday, 4.7 million on Tuesday and Wednesday, 4.6 million on Thursday, 4.3 million on Friday, and 2.5 million on Saturday (Table 22). The hourly averages from transponder GPS unique signals ranged from around 600,000 at 2 AM to 1.6 million at 3 PM (Figure 59). The results by hour showed peak time from 7 AM to 4 PM (Figure 59). We were able to calculate rolling trucks verses stopped (and idling). Stopped trucks peaked at 1 AM and were lowest from 8 AM to 6 PM (Figure 60). This number is reflected in the percentage of trucks rolling with the exception of the lowest percentage being 2 AM, explained by the reduction of rolling trucks at 2 AM (Figure 61). The average rolling speed of trucks throughout the day hovered around 40 mph with slightly higher speeds from 11 AM to 7 PM (Figure 62). Total fuel consumed, based on an average of 6.5 mpg, throughout the day ranged from 101,000 gallons and hour to 253,000 gallons an hour (Figure 63). These totals can be inferred over the entire study period with around 35% of modern trucks with GPS transponders and the study months being presumed to be persistent with the other months in the study, yielding a consumption of 10.6 million gallons of diesel from 2011-2014 in Mississippi alone, a daily average of 1,218 gallons a day (Table 23).

Data Analysis

We can look at each hour of truck traffic to help identify where performance measures are not optimized. As an example, we will use just trucks from 7 to 8 AM from 2011-2014. By separating out the 1.35 million unique trucks by speed, we can see a picture of the state develop. Starting with the stopped trucks, about 35% of trucks are stopped or about 375,000 unique trucks (Figure 64). The urban centers and major road network can be interpolated. The next speed is rolling under 25 mph, or crawling, usually meaning shifting is constant and fuel consumption is greatly increased. About 657,000 unique trucks, the largest number, can be seen in urban centers and along highways (Figure 65). Trucks rolling between 25 and 45 are city and rural cruising speeds and have only about 87,000 unique trucks concentrated around the timber region of the mid-south, urban centers, and highways (Figure 66). The second highest number of trucks fall in the 45-70 mph category or cruising speed, which required less shifting and maximizing fuel efficiency. This category shows the highway network very well and the urban centers are not as apparent, beyond multiple highways converging (Figure 67). The last category are trucks exceeding 70 mph or speeding trucks that can go as fast as they deem they can without getting ticketed (Figure 68). This category shows only the highways and interstates that trucks are free to move and the main thoroughfares are evident. By combining all of these truck speeds, the fast moving trucks (red) and cruising trucks (dark green) reveal the truck network of Mississippi (Figure 69). The remaining colors of light green and orange show urban centers and distribution areas. Comparing a road atlas, the non-interstate roads are revealed for overuse (Figure 69). The non-Interstate heavy traffic is evident on Highways 25, 45 and ALT 45, 61, 72, 78, 82, 84, and 98. Of note, Highway 25 has a significant truck slowdown near Louisville, Mississippi. Running the FPMs for Jackson, Meridian, and Hattiesburg along the highway system, showed no major hotspots for consistent congestion.

The port of Gulfport can also be evaluated by number of trucks rolling (Figure 70). At the bottom of the Figure 70 is the port that extends into the Mississippi Sound with highway 90 running along the coast line. Here traffic has to move away from the port but beyond rail, has to travel through local streets with lights and stop signs. Trucks take highway 49 north mainly, but also take 30th avenue north to 28th street west to Canal road then north. Trucks also head up 49 to Pass Road east to Cowen/ Lorraine then north to the interstate. Figure 70 shows highway 49 in the middle heading north with the western branch south of the interstate being 28th and Canal and the eastern branch being Pass Road and Cowen/ Lorraine. The evidence of three alternate paths by trucks to the interstate from the port of Gulfport shows a severe need for transportation improvement. Further, the rate of speed the trucks are moving are below 45 and 25 during the transit. Highway 90 also has evidence of heavy truck traffic that will typically need to head to I-10 as well, but through different routes. Evidence of the truck flow in the area can be seen by the numerous stopped vehicles at the Flying J Truck Stop at the intersection of Canal and I-10 in Harrison County, where many trucks from the port are joined by through traffic on I-10 (Figure 71). FPMs for Gulfport show that Highway 49 is a point of congestion between I-10 and the port.

Meridian, Mississippi has a good location for intermodal with truck and rail converging with connections to Gulfport and Mobile, Alabama with its heaviest roads being I 20 and I 59 followed by heavy traffic on Highway 45 with slower connections on Highways 19 and 39 (Figure 72). Hattiesburg also could see intermodal opportunities with connections to Gulfport and New Orleans with its heaviest roads being I 59 and Highways 49 and 98 (heading east) with slower connections on Highways 11, 42, 98 (heading west), and 589 (Figure 73).

Conclusions and Recommendations

The port of Gulfport needs better north-south corridor to allow for truck freight to be moved. Potential roads for upgrading would be Highways 25, 45 and ALT 45, 61, 72, 78, 82, 84, and 98. Minimally, it would be recommended to reduce traffic impediments, lights, etc. Much of the traffic flow in the State of Mississippi is through traffic using Interstates and highways to bisect the state. The urban centers continue to show expected slowdown of traffic.

Table 21 Unique Trucks by Year

Truck Transponder Signals in March June October	
2011	4,185,400
2012	6,138,500
2013	7,239,500
2014	9,006,900

Table 22 Unique trucks by Day of the week

Signals by Day of Week in March June October	
Sunday	1,760,000
Monday	4,046,200
Tuesday	4,660,400
Wednesday	4,697,100
Thursday	4,604,200
Friday	4,258,000
Saturday	2,544,400

Table 23 Costs and Consumption of Diesel

	Average Fuel Cost ($) at $2.43	Total Gallons Consumed at 6.5 mpg
Total Cost	$12,975,268	4,571,228
Annual Costs	$4,325,089	1,523,743
Monthly Costs	$1,441,696	507,914
Daily Costs	$47,398	16,699
4 Year Costs	$17,300,357	6,094,971
Total Inferred Costs	$30,275,625	10,666,199
Average Inferred Daily Cost	$82,947	29,222
Average inferred Hourly Cost	$3,456.12	1,218

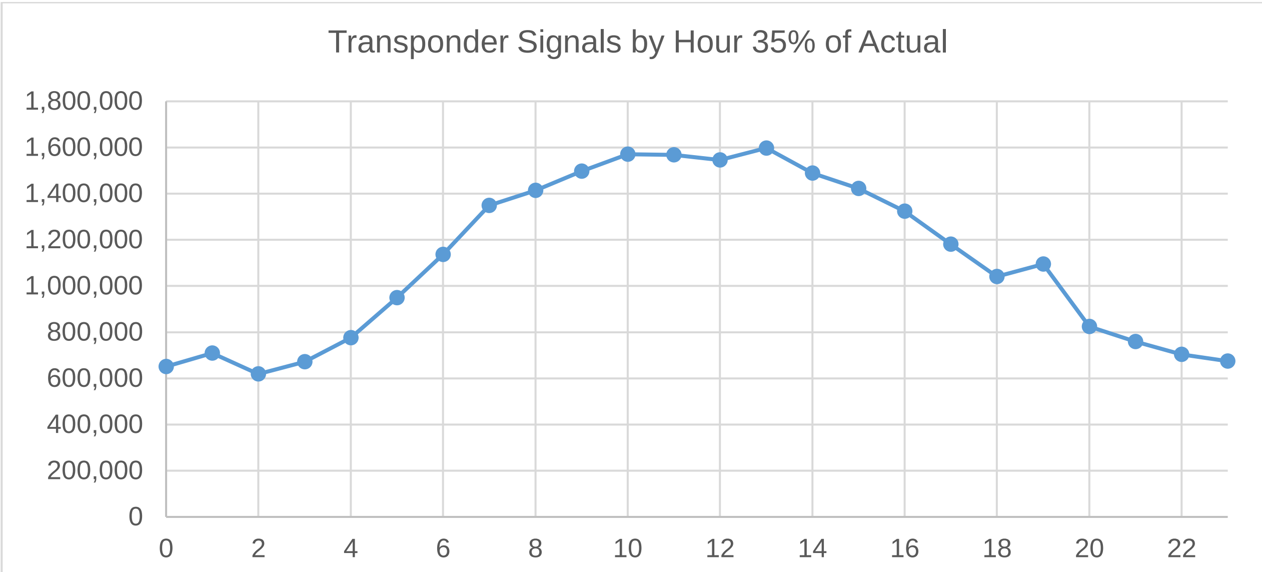


Figure 59 Unique trucks by hour with GPS transponders volume total over 2011-2014

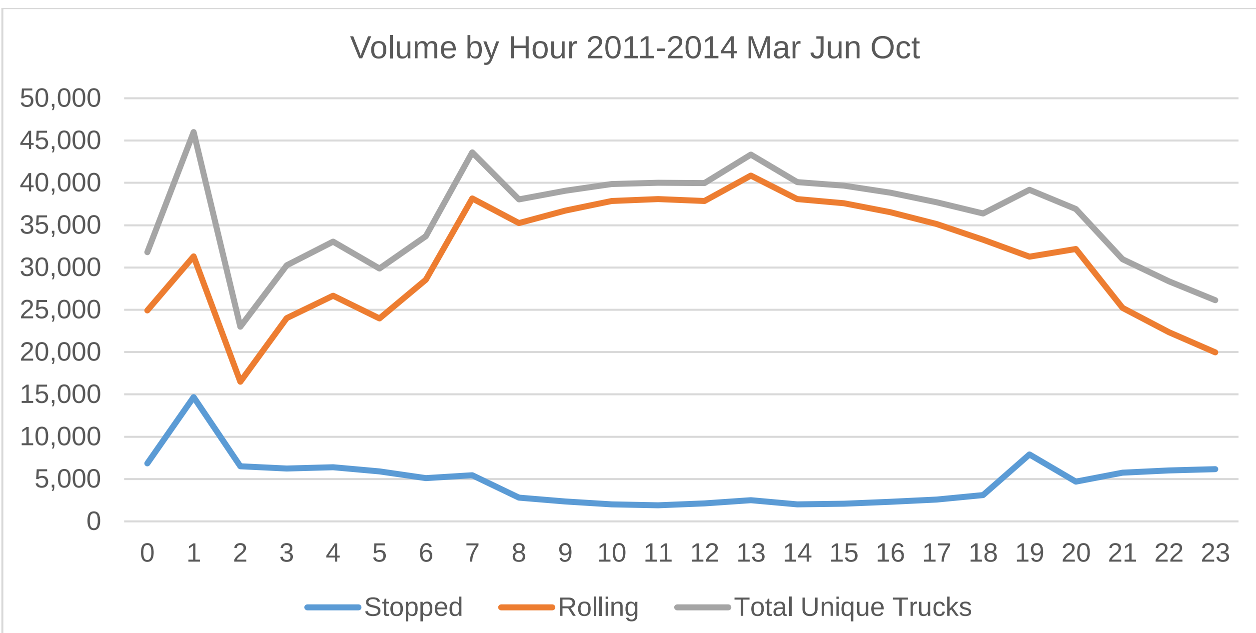


Figure 60 Volume of individual trucks with trucks registering zero mph shown as stopped trucks, which are idling

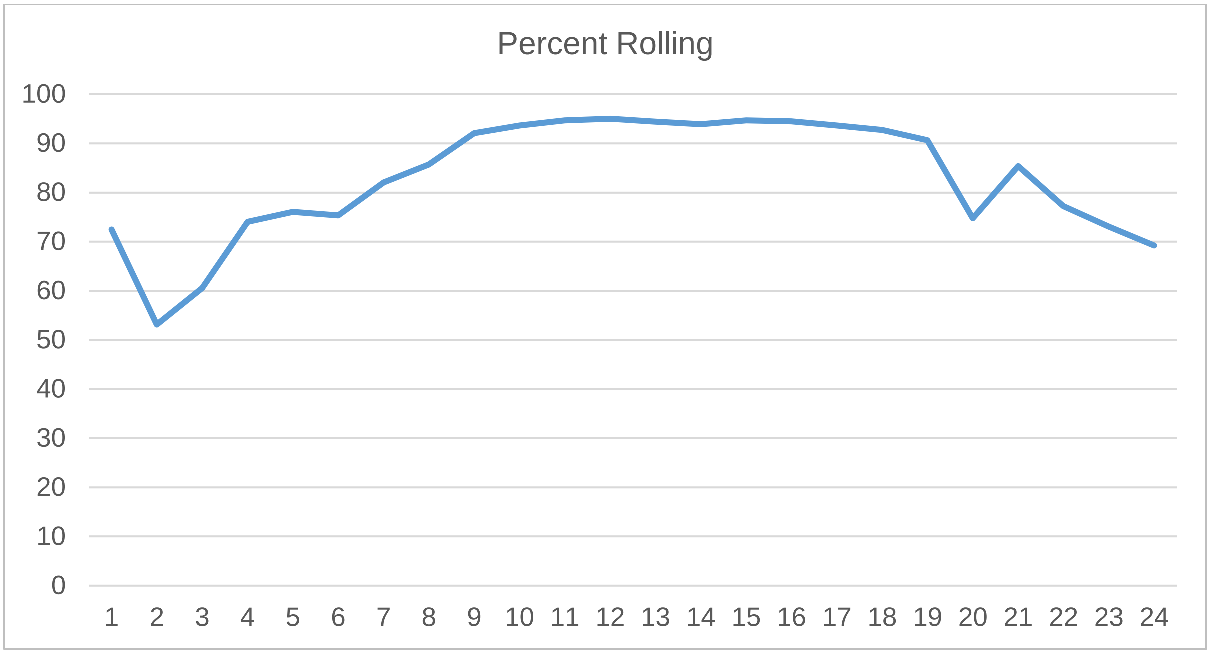


Figure 61 Percentage of trucks rolling to those stopped over the whole dataset by hour

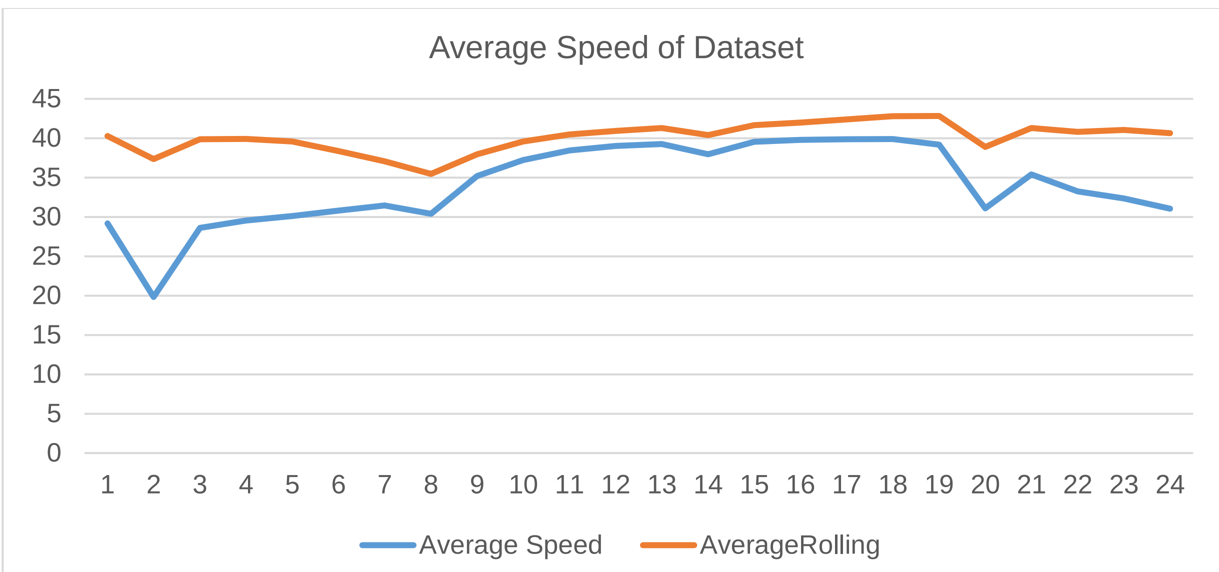


Figure 62 Average speed of trucks with and without stopped trucks

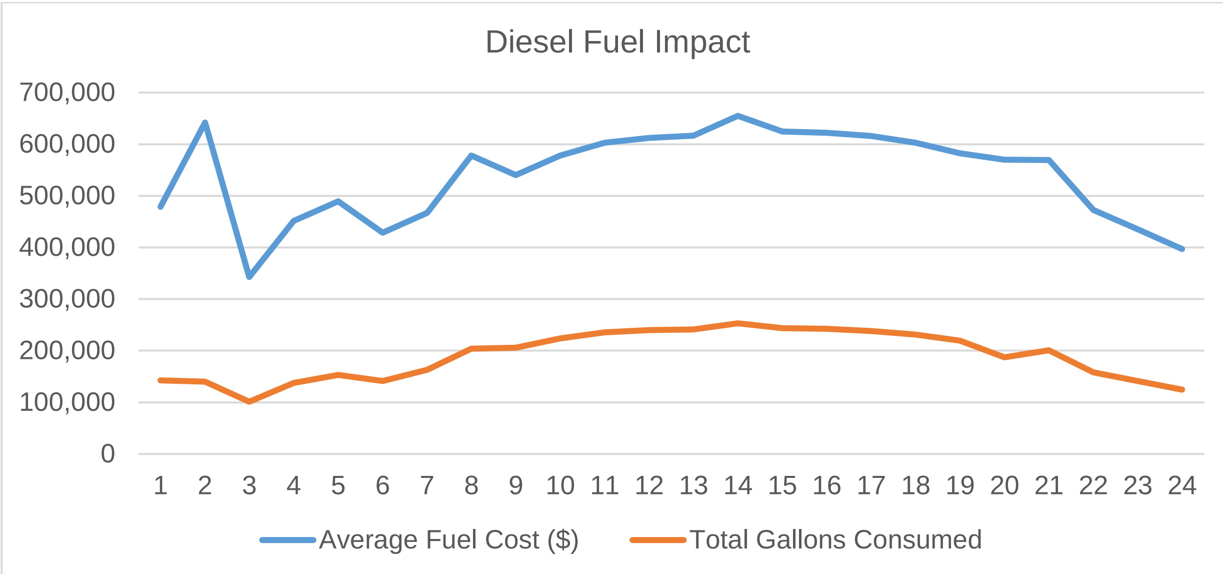


Figure 63 Diesel cost for the trucks in Mississippi by hour

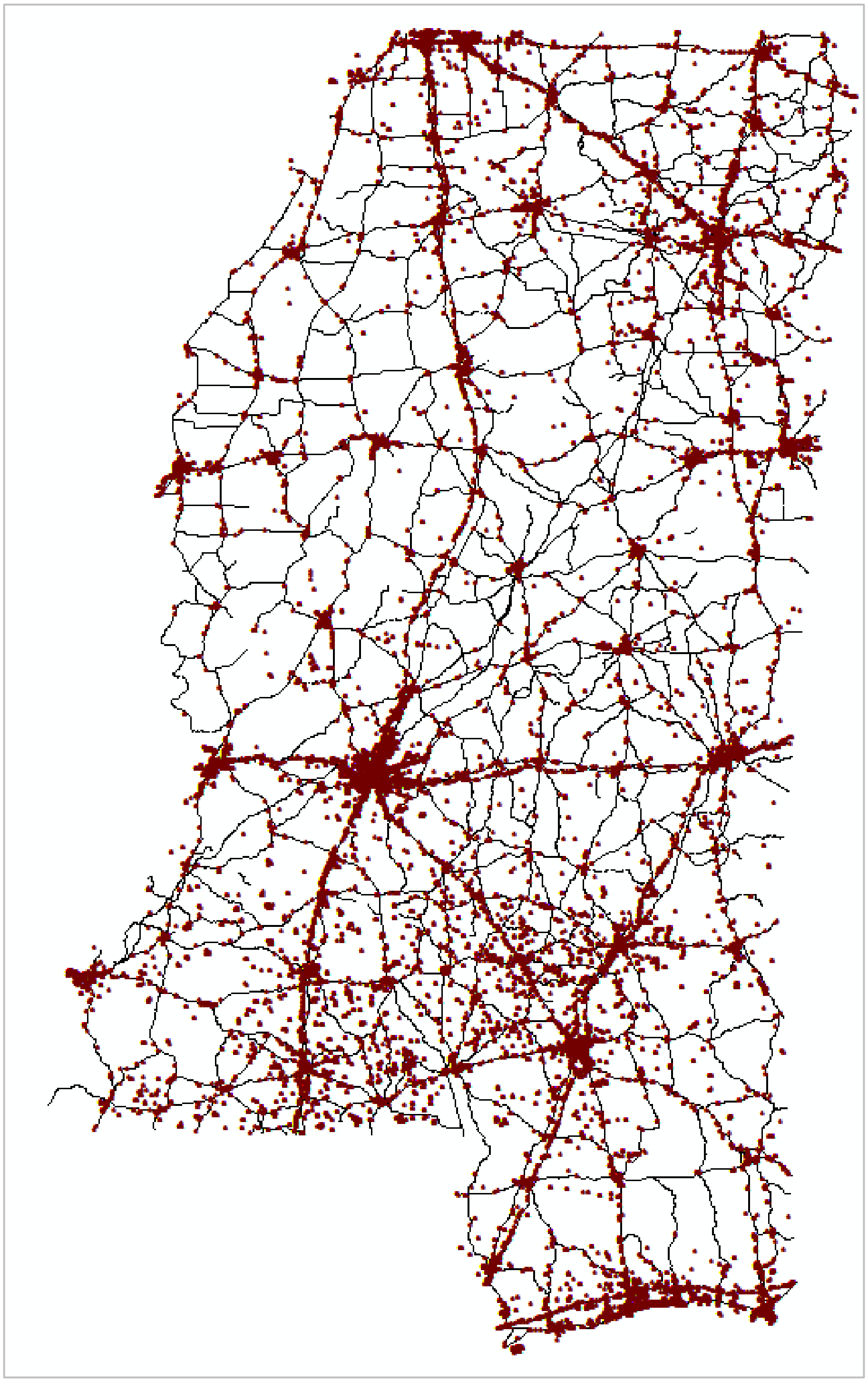


Figure 64 Trucks stopped between 7 and 8 AM (357,000 trucks)

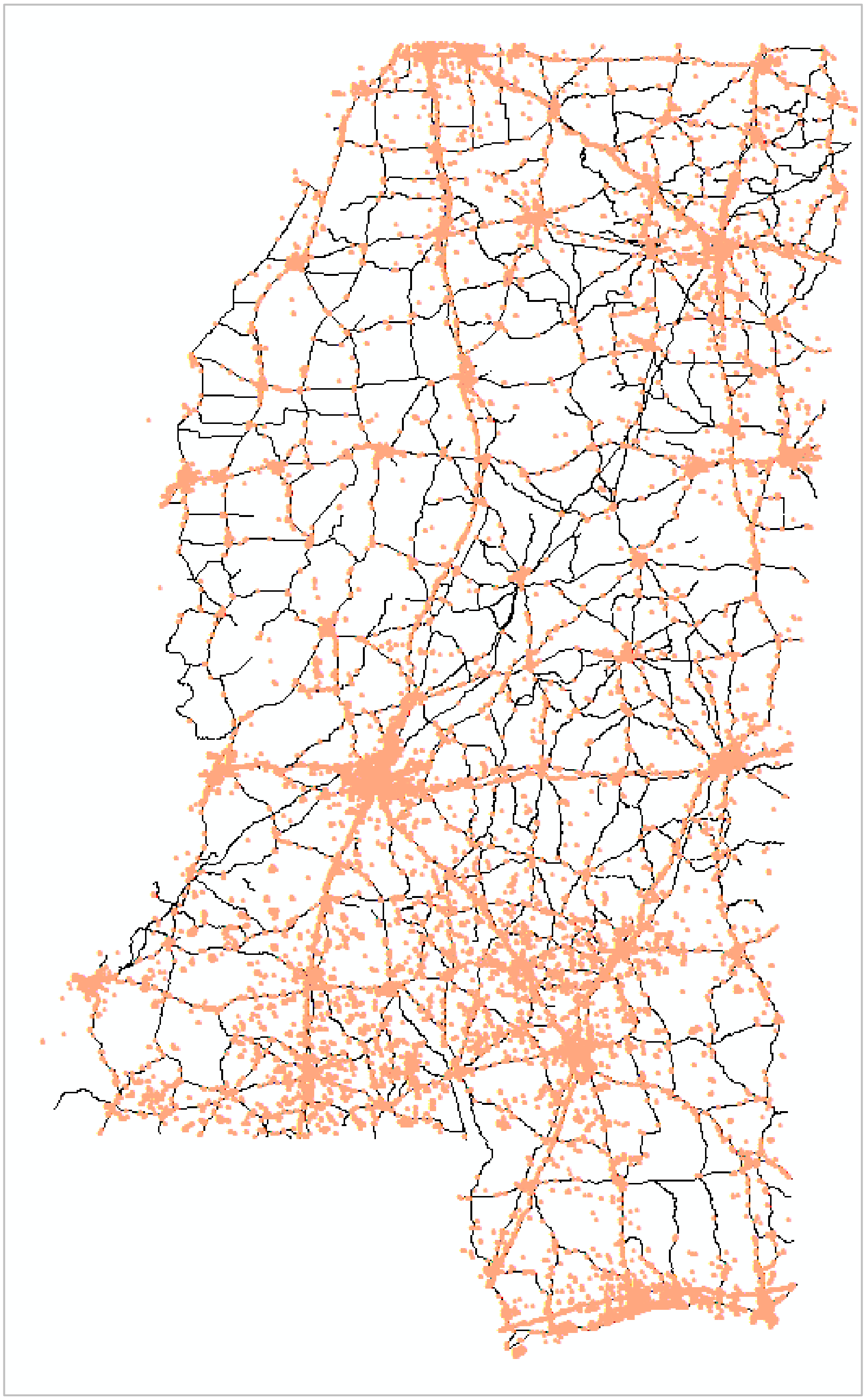


Figure 65 Trucks crawling from 0-25 mph (675,000 trucks)

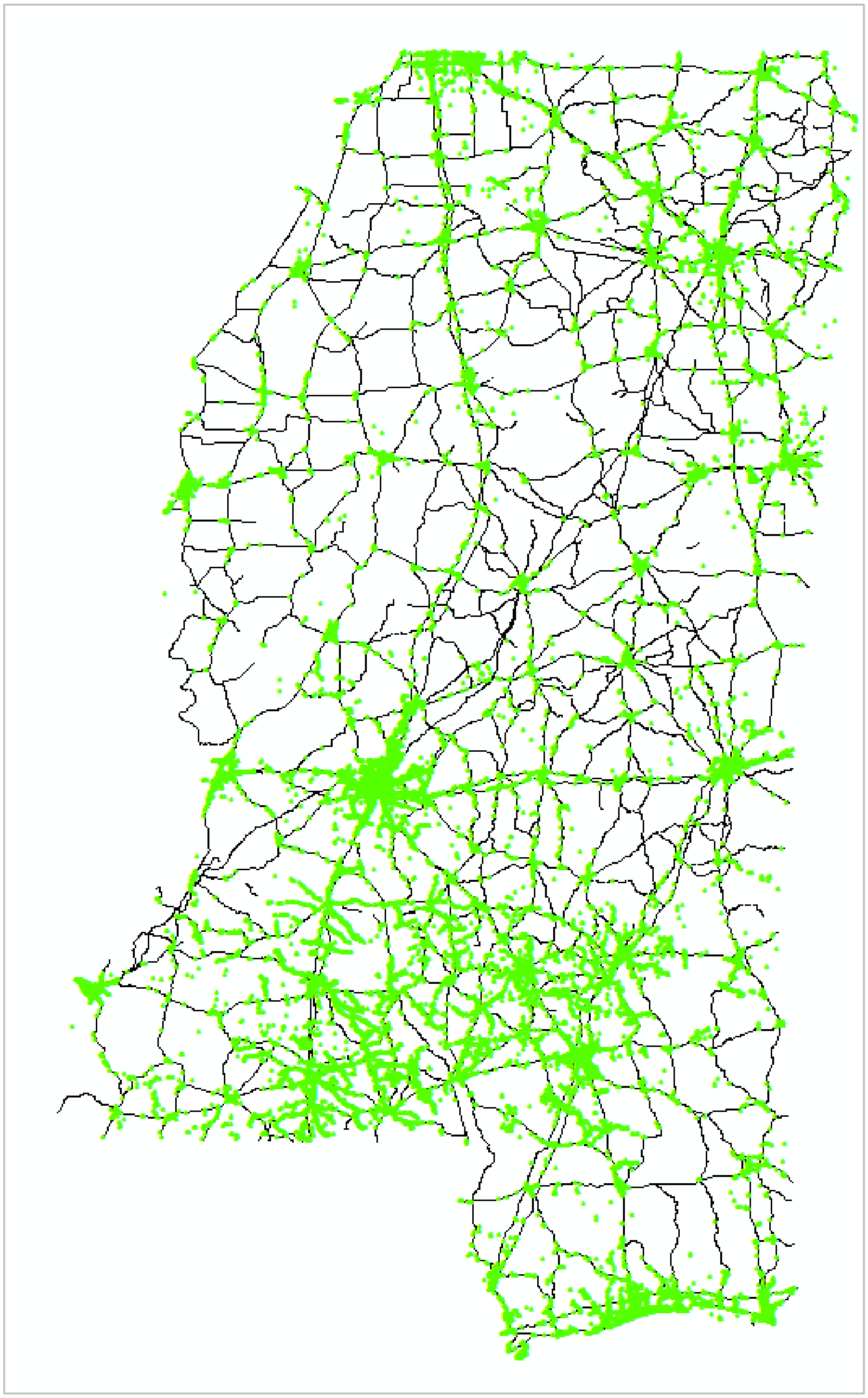


Figure 66 Trucks rolling at 25-45 mph (87,000 trucks)

Figure 67 Trucks cruising at city and rural speeds from 45-70 mph (550,000 trucks)

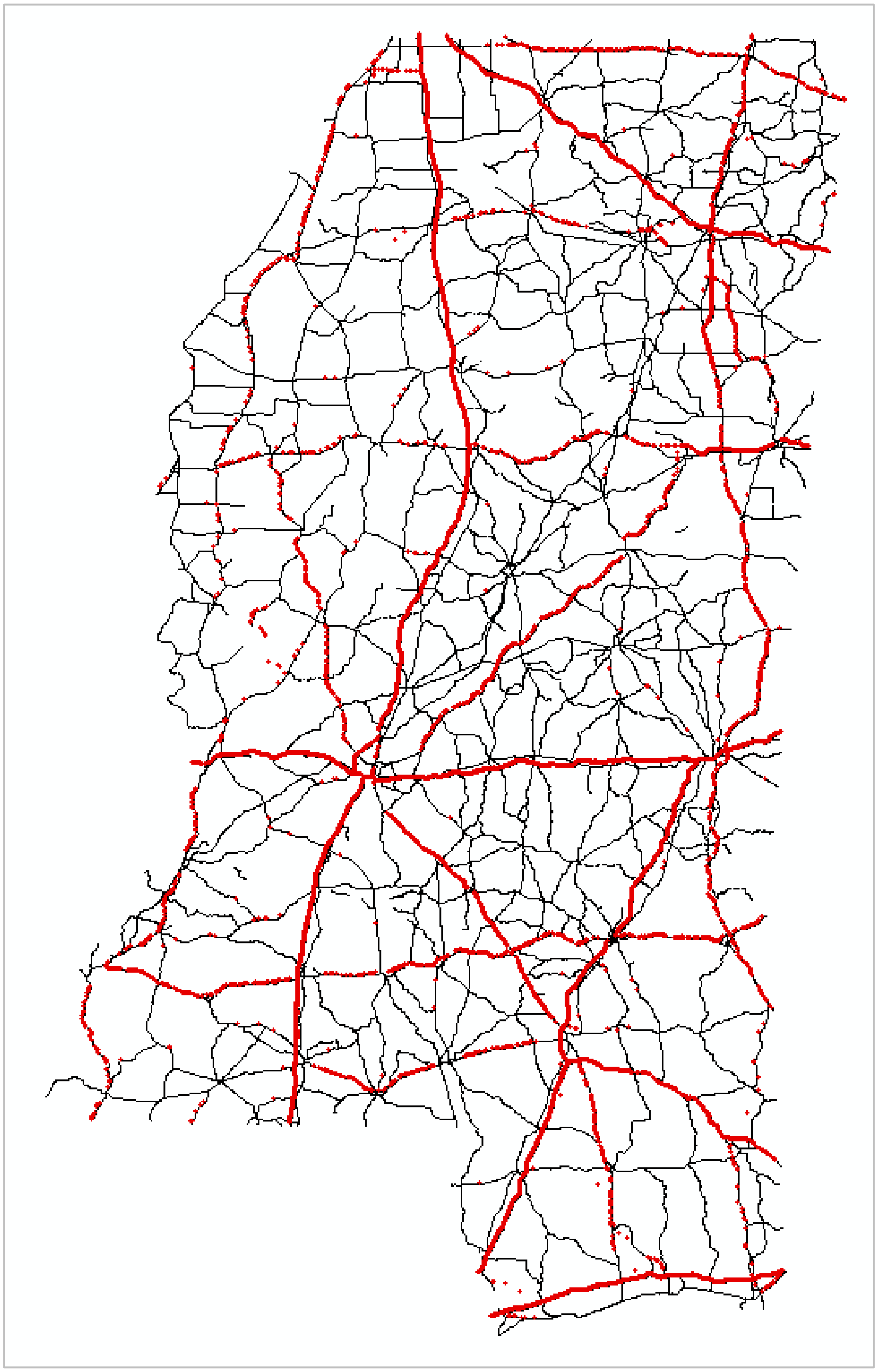


Figure 68 All trucks travelling in excess of 70 mph (37,000 trucks)

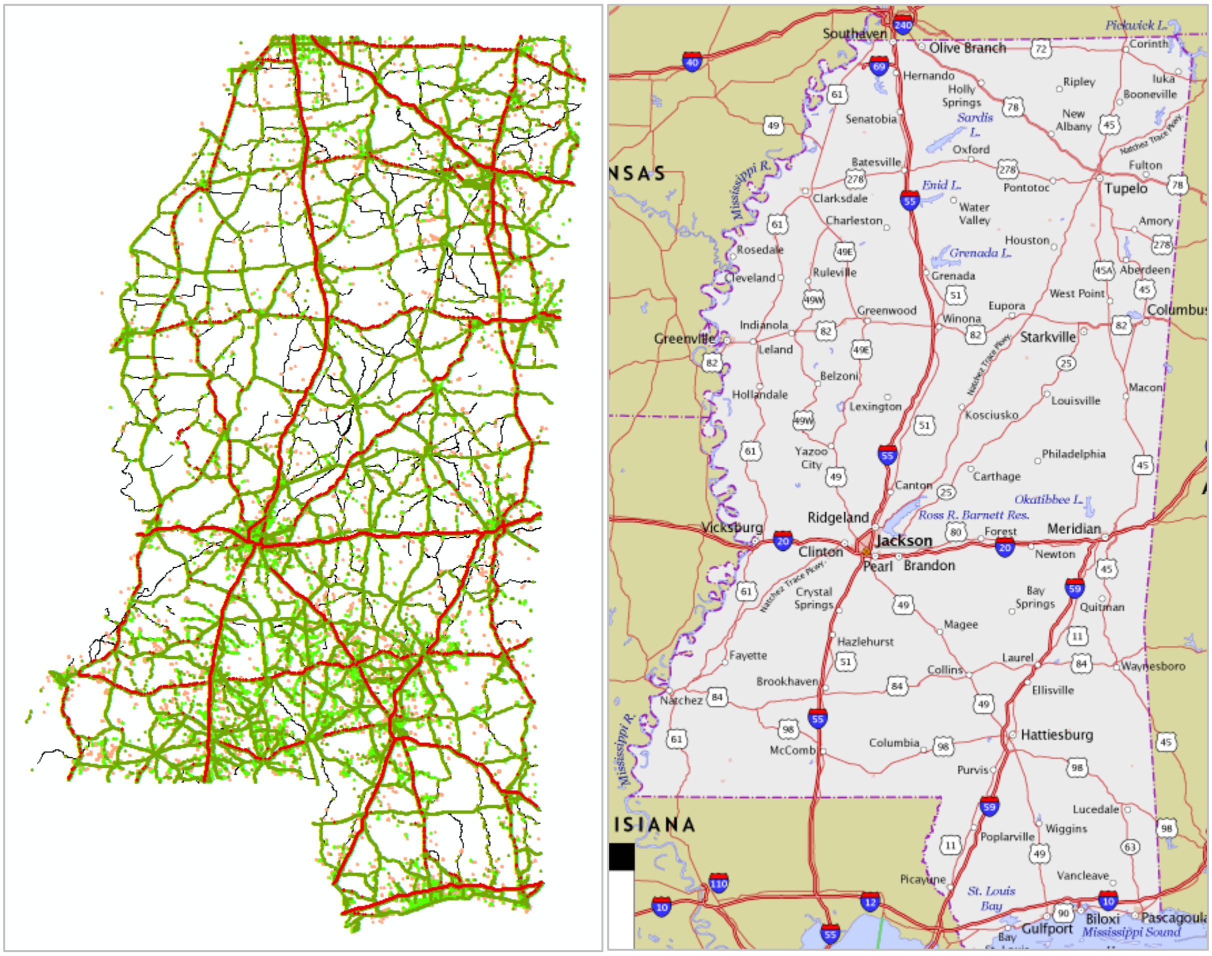


Figure 69 All trucks by speed (1.35 million trucks)

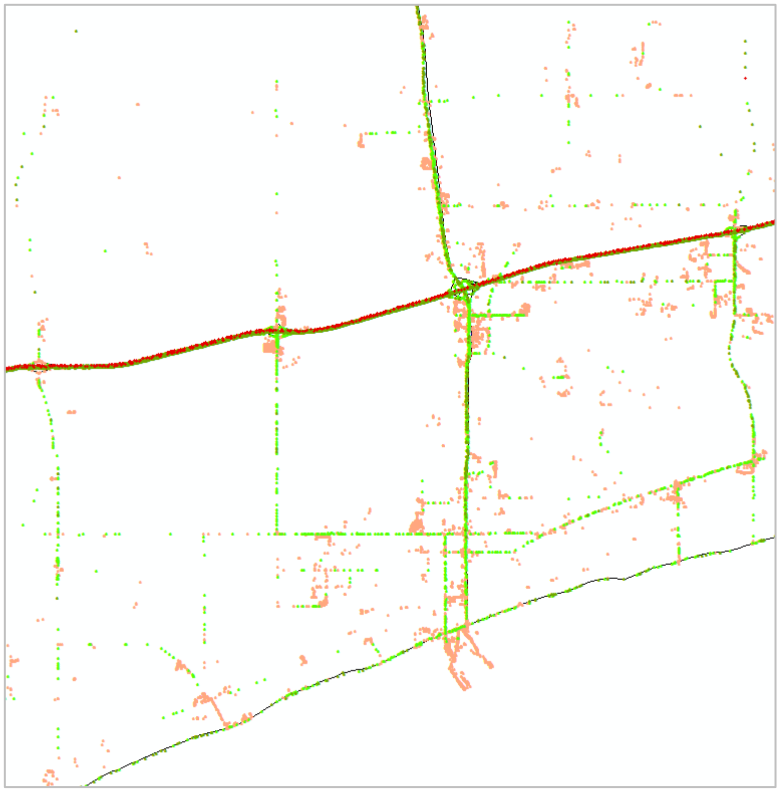


Figure 70 Port of Gulfport (bottom center) and the intersection of I-10 and Highway 49

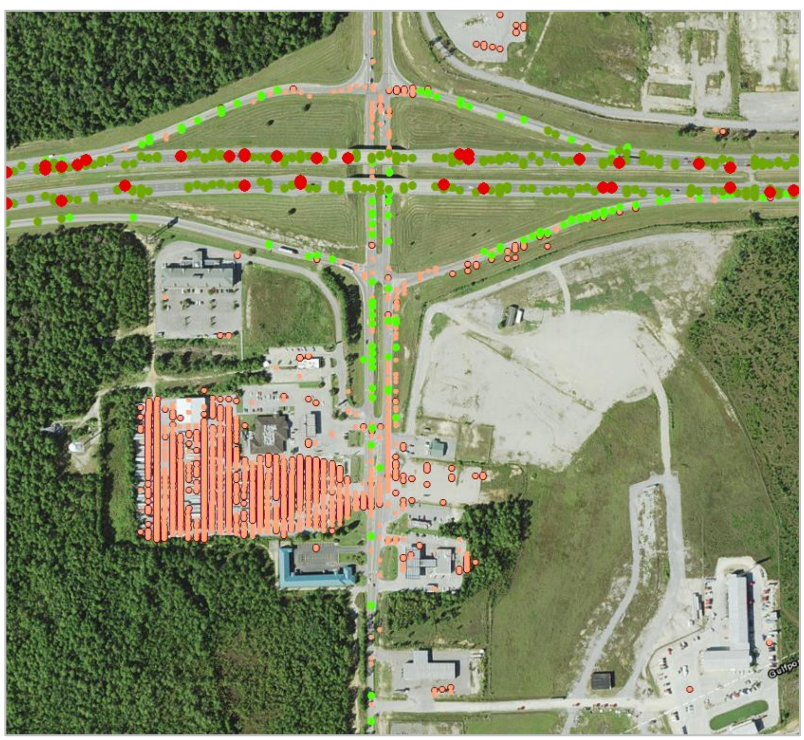


Figure 71 Flying J truck stop at I-10 and Canal Road in Harrison County, Mississippi

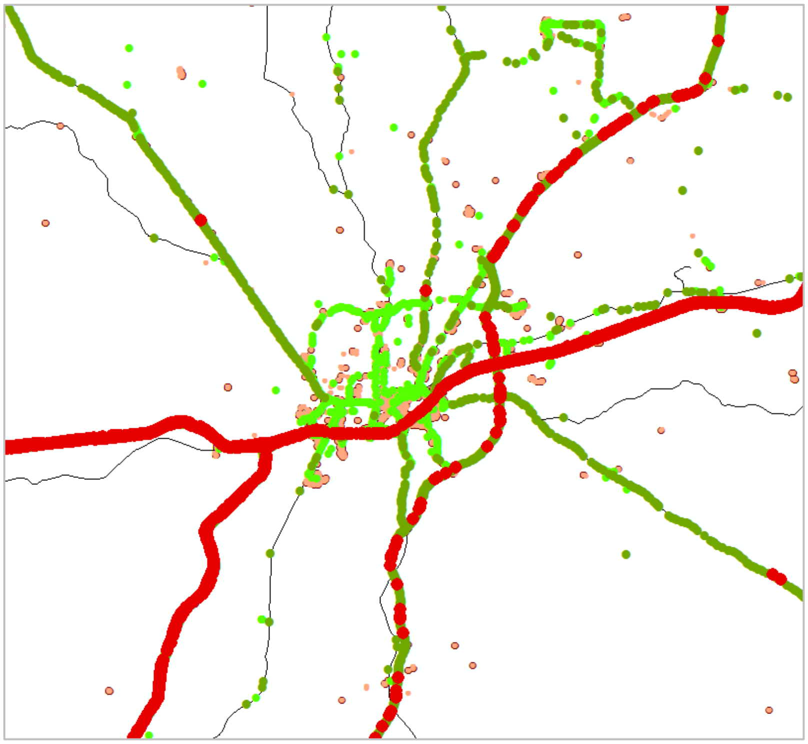


Figure 72 Meridian, Mississippi and the road network surrounding it

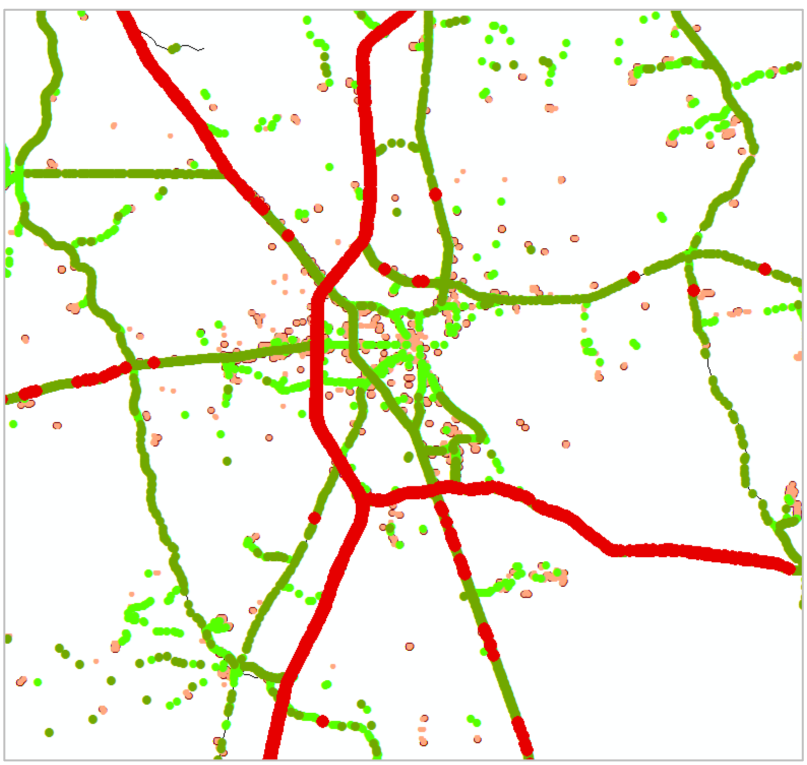


Figure 73 Hattiesburg, Mississippi and the road network surrounding it

APPENDIX- B. ALABAMA CASE STUDY

Data Overview of Alabama

The analysis performed in Alabama focused on the using the data to determine the daily average truck speeds in Alabama along roadway facilities. As to the convention, TSa was used for trucks traveling north or east (as the mileposts counted up in these directions; and TSb was used for trucks traveling south and west. The daily average speed is calculated based on the GPS data of October 2014, where the Monday to Friday consist of two different days and Saturday and Friday consist of three different days.

The average speeds were calculated using the following convention:

1. Join the LinkFPMs results on the Alabama initial network.
2. There are so many zero values of link speed data attributes.
3. Count the number of zero values than the attribute cell is excluded for the calculation
4. $\frac{TSa1+TSa2}{2-number\ of\ zero\ cell}$ or $\frac{TSa1+TSa2+TSa3}{3-number\ of\ zero\ cell}$

Statewide Truck Speeds

Figure 74-75 represent daily truck speeds for the roadways in Alabama based on the tool.

One item of particular importance was the number of zero values. For the TSa, there were 191 zero values out of 2838 total links; and for the TSb, there were 193 zero values out of 2838 total links. Although it is important to note that there were only 3 links with zero values that were identified as interstate links.

Conduct speed analysis of I-65

To refine the analysis, Interstate 65 was selected as a roadway of interest. This facility run the length of the state (north-south) and connects the Port of Mobile to locations within Alabama as well as cities within the Ohio Valley including Nashville, Louisville, Indianapolis and Chicago. The analysis was performed for each day of the week using the tool and the Figure 76 thru Figure 89 show difference in truck speed for the different days.

Figure 74 Result of Travel Speed (TSa) of all Alabama Network on Wednesday

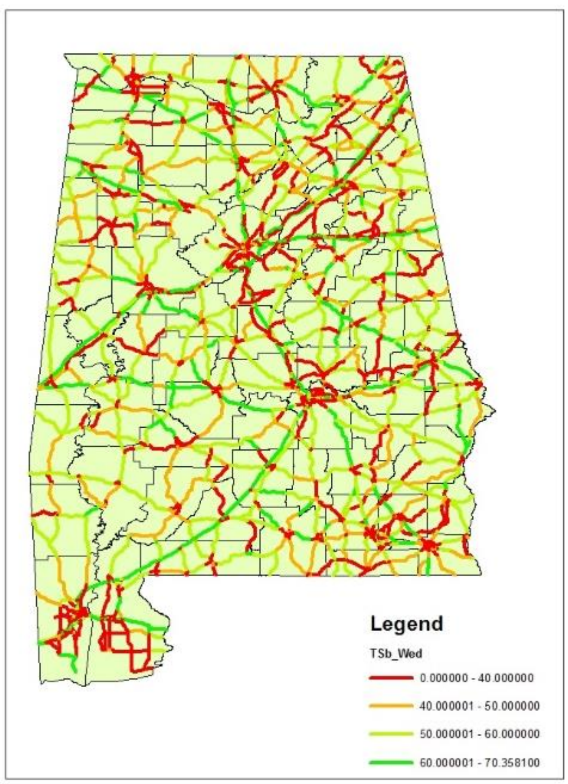


Figure 75 Result of Travel Speed (TSb) of all Alabama Network on Wednesday

Figure 76 I-65 speed on Monday (TSa)

Figure 77 I-65 speed on Monday (TSb)

Figure 78 I-65 speed on Tuesday (TSa)

Figure 79 I-65 speed on Tuesday (TSb)

Figure 80 I-65 speed on Wednesday (TSa)

Figure 81 I-65 speed on Wednesday (TSb)

Figure 82 I-65 speed on Thursday (TSa)

Figure 83 I-65 speed on Thursday (TSb)

Figure 84 I-65 speed on Friday (TSa)

Figure 85 I-65 speed on Friday (TSb)

Figure 86 I-65 speed on Saturday (TSa)

Figure 87 I-65 speed on Saturday (TSa)

Figure 88 I-65 speed on Sunday (TSa)

Figure 89 I-65 speed on Sunday (TSb)

As can be seen from the figures, the majority of truck speeds were in the 60 miles per hour or greater range, indicating that is limited congestion or slowing of the travel speed within the state and along the roadway. In addition, it can be seen that the majority of congestion along Interstate 65 is located in the Birmingham Alabama area.

Birmingham Area Speed Analysis

To further examine the data, a focused view of Birmingham Alabama was performed to view changes in truck travel speeds for the community. There were four days in October from four separate years selected.

- Y 2011 results from data of 2011 Oct. 19
- Y 2012 results from data of 2012 Oct. 17
- Y 2013 results from data of 2013 Oct. 16
- Y 2014 results from data of 2014 Oct. 15

Below Figure 90 thru Figure 97 are showing the truck average travel speed for the Birmingham area.

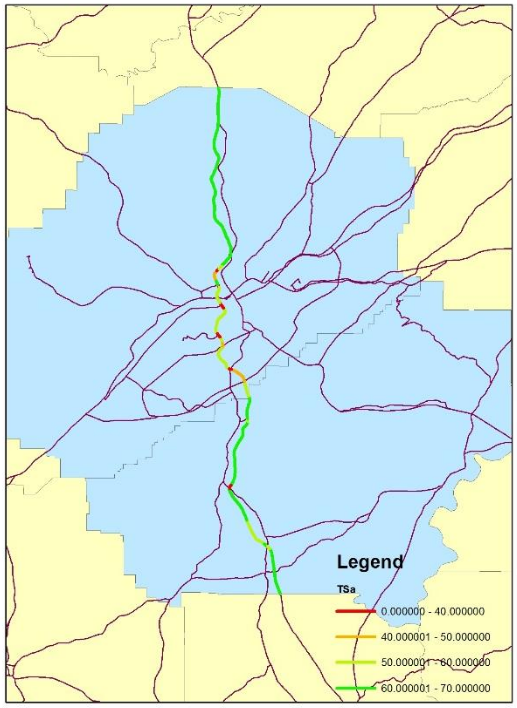


Figure 90 2011 I-65 PM Peak hour Travel Speed at Birmingham Area (TSa)

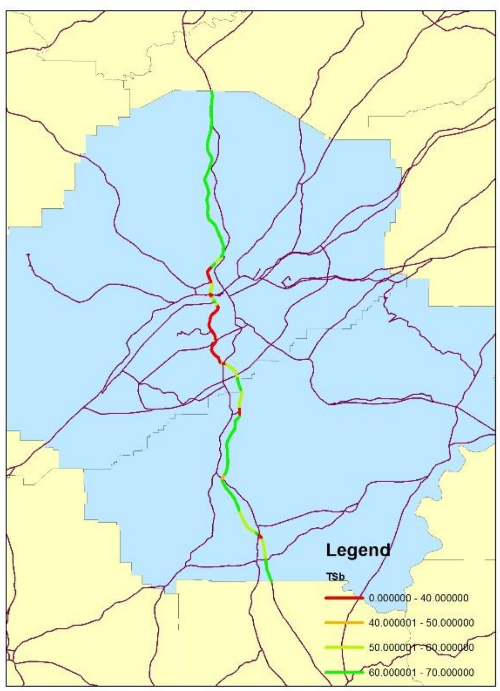


Figure 91 2011 I-65 PM Peak hour Travel Speed at Birmingham Area (TSb)

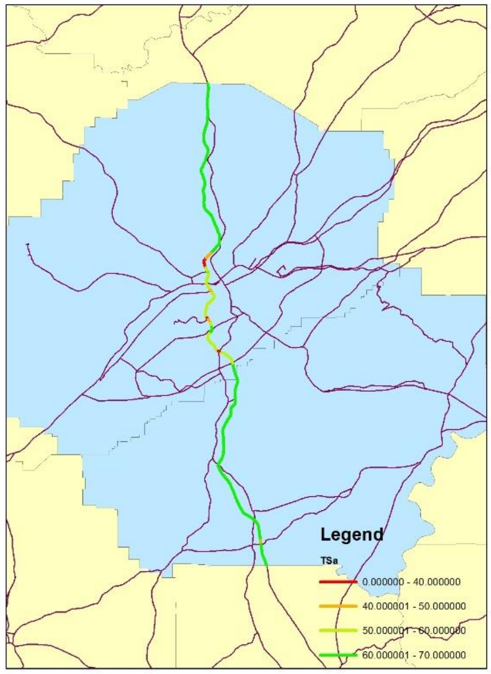


Figure 92 2012 I-65 PM Peak hour Travel Speed at Birmingham Area (TSa)

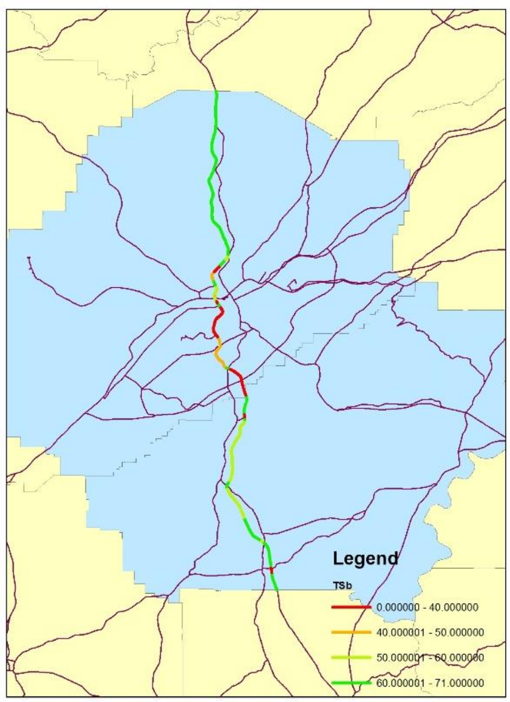


Figure 93 2012 I-65 PM Peak hour Travel Speed at Birmingham Area (TSb)

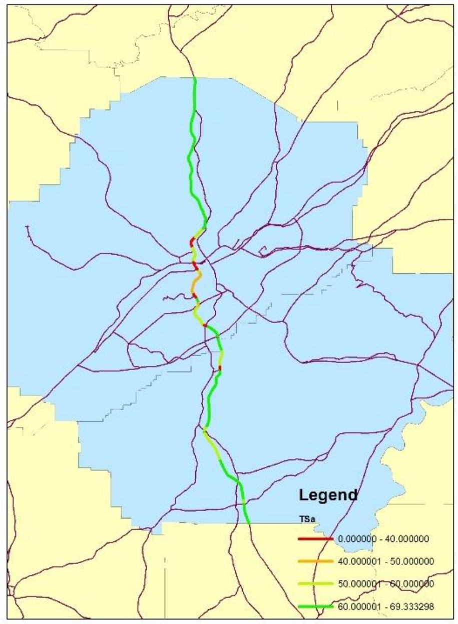


Figure 94 2013 I-65 PM Peak hour Travel Speed at Birmingham Area (TSa)

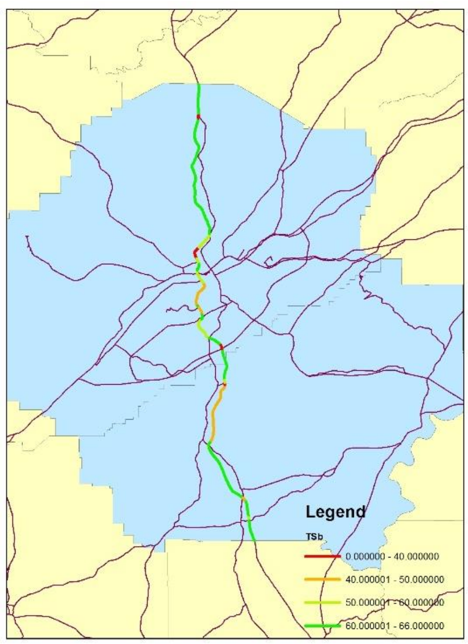


Figure 95 2013 I-65 PM Peak hour Travel Speed at Birmingham Area (TSb)

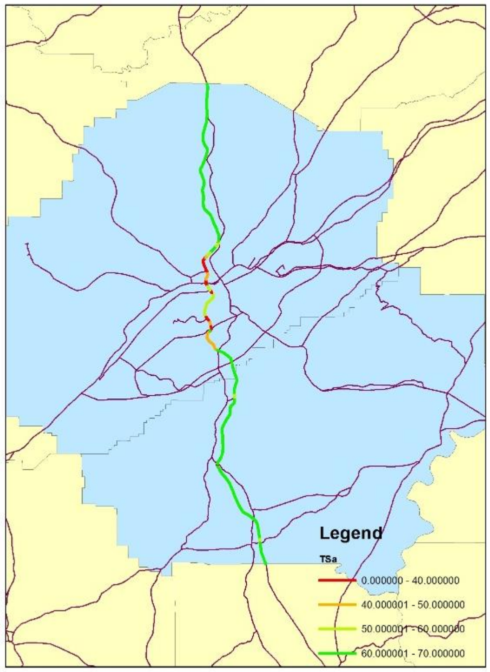


Figure 96 2014 I-65 PM Peak hour Travel Speed at Birmingham Area (TSa)

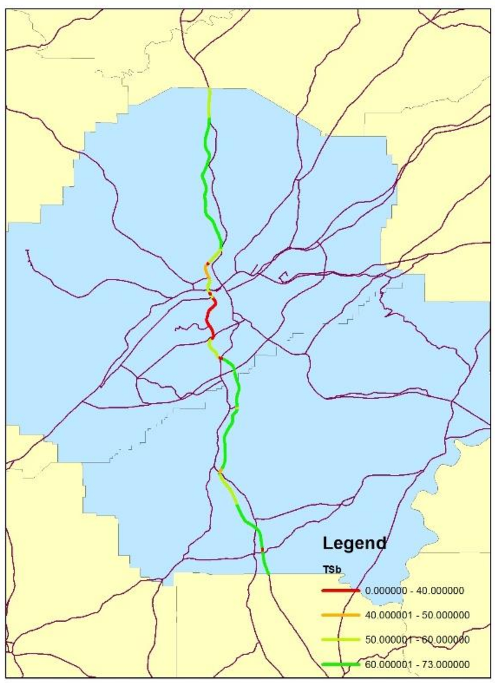


Figure 97 2014 I-65 PM Peak hour Travel Speed at Birmingham Area (TSb)