



FINAL REPORT

**Evaluation of Roadway Reallocation Projects:
Analysis of Before-and-After Travel Speeds and
Congestion Utilizing High-Resolution Bus Transit Data**

NITC-RR-887 ■ November 2017

*NITC is a U.S. Department of Transportation
national university transportation center.*



**EVALUATION OF ROADWAY REALLOCATION
PROJECTS
ANALYSIS OF BEFORE-AND-AFTER TRAVEL SPEEDS AND
CONGESTION UTILIZING HIGH-RESOLUTION BUS
TRANSIT DATA**

Final Report

NITC-RR-887

by

Miguel Figliozzi
Travis Glick
Portland State University

for

National Institute for Transportation and Communities (NITC)
P.O. Box 751
Portland, OR 97207



November 2017

Technical Report Documentation Page			
1. Report No. NITC-RR-887	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Evaluation of Roadway Reallocation Projects Analysis of Before-and-After Travel Speeds and Congestion Utilizing High-Resolution Bus Transit Data		5. Report Date November 2017	
		6. Performing Organization Code	
7. Author(s) Miguel Figliozzi Travis Glick		8. Performing Organization Report No.	
9. Performing Organization Name and Address Portland State University		10. Work Unit No. (TRAVIS)	
		11. Contract or Grant No. 887	
12. Sponsoring Agency Name and Address National Institute for Transportation and Communities (NITC) P.O. Box 751 Portland, OR 97207		13. Type of Report and Period Covered	
		14. Sponsoring Agency Code	
15. Supplementary Notes			
16. Abstract The traditional process of identifying corridors for road diet improvements involves selecting potential corridors (mostly based on identifying four-lane roads) and conducting a traffic impact analysis of proposed changes on a selected roadway before implementing changes. The evaluation of roadway reallocation projects should include the analysis of traffic volumes, level of service, speeds, queue lengths and bus operations. There are tools and equipment to evaluate effectively traffic volumes and level of service changes in before-and-after studies. However, the detailed evaluation of speed and queue length distributions along a segment are significantly more cumbersome. In addition, the exhaustive evaluation of bus operations requires detailed data and specific tools and cannot be readily accomplished with field measurements or conventional level of service studies. This research presents a general methodology for the detailed evaluation of transit operations and speed and queue length distributions along roadway reallocation projects. The strategy and methodology is based on the utilization of high-resolution transit datasets. The focus of this research is on the development of a practical, general and theoretically sound methodology that can be applied to future roadway reallocation projects and applicable in a wide range of traffic conditions and locations. The methodology proposed in this research was applied successfully to two separate road diets in Portland, OR.			
17. Key Words Roadway reallocation, high-resolution data, before-and-after study, speed profiles, speed variability		18. Distribution Statement No restrictions. Copies available from NITC: www.nitc-utc.net	
19. Security Classification (of this report) Unclassified	20. Security Classification (of this page) Unclassified	21. No. of Pages 56	22. Price

ACKNOWLEDGEMENTS

The authors would like to acknowledge Steve Callas and Miles Crumley from TriMet for providing transit data and for their useful feedback and data analysis suggestions. The authors of this report are responsible for any errors, mistakes or omissions.

DISCLAIMER

The contents of this report reflect the views of the authors, who are solely responsible for the facts and the accuracy of the material and information presented herein. This document is disseminated under the sponsorship of the U.S. 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 U.S. Government. This report does not constitute a standard, specification, or regulation.

RECOMMENDED CITATION

Figliozi, Miguel and Travis Glick. *Evaluation of Roadway Reallocation Projects Analysis of Before-And-After Travel Speeds and Congestion Utilizing High-Resolution Bus Transit Data*. NITC-RR-887. Portland, OR: Transportation Research and Education Center (TREC), 2017.

TABLE OF CONTENTS

EXECUTIVE SUMMARY	1
1.0 INTRODUCTION.....	2
1.1 MOTIVATION.....	2
1.2 RESEARCH OBJECTIVE	2
2.0 LITERATURE REVIEW	3
2.1 ROAD DIETS.....	3
2.2 ROAD DIET BEFORE-AND-AFTER STUDIES	3
2.3 TRANSIT BEFORE-AND-AFTER STUDIES.....	4
2.4 BUS SPEEDS AND TRAVEL TIMES.....	4
2.5 TRANSIT ROUTE-LEVEL ANALYSIS.....	5
3.0 DATA SOURCES	7
3.1 TRAFFIC VOLUMES.....	7
3.2 TRANSIT DATA.....	7
3.2.1 Stop-to-Stop Data.....	8
3.2.2 Stop-Disturbance Data	8
3.2.3 High-Resolution Data	8
4.0 METHODOLOGY	10
4.1 TRIMET DATA INTEGRATION	10
4.1.1 Spatial Data Cleaning	11
4.1.2 GIS Tool Definitions.....	11
4.1.3 Final Combined Dataset.....	12
4.1.4 Data Irregularities	13
4.2 DATA ANALYSIS.....	13
4.2.1 Unknown Distributions.....	13
4.2.2 Peak-Hour vs. Whole-Day Performance.....	17
4.3 PROPAGATION OF UNCERTAINTY.....	17
4.3.1 Assumption of Normality	17
4.3.2 Propagation of Uncertainty Addition and Subtraction.....	18
4.3.3 Propagation of Uncertainty Product and Quotient.....	19
4.3.4 Confidence Intervals	19
4.4 TRAVEL TIMES.....	20
4.4.1 Space-Time-Speed Diagrams.....	20
4.4.2 Before-and-After Comparisons.....	20
4.4.3 Travel Speeds, Travel Times and Peak Performance	21
5.0 CASE STUDIES.....	22
5.1 CASE STUDY I: LOMBARD STREET	22
5.2 CASE STUDY II: NE 16 TH STREET.....	25
6.0 RESULTS	27
6.1 CASE STUDY I: LOMBARD STREET	27
6.1.1 Travel Times	27
6.1.2 Travel Speeds.....	28
6.2 CASE STUDY II: NE 16 TH STREET.....	29
6.2.1 Travel Times	29

6.2.2	Travel Speeds.....	29
7.0	DISCUSSION.....	32
8.0	REFERENCES.....	33
9.0	APPENDIX A.....	36
9.1	LOMBARD STREET HEATMAPS	37
9.2	LOMBARD STREET INDICES	39
10.0	APPENDIX B.....	41
10.1	NE 16 TH STREET HEATMAPS	41
10.2	NE 16 TH STREET INDICES	43

LIST OF TABLES

Table 4-1	– GIS tools definitions with visuals.....	12
-----------	---	----

LIST OF FIGURES

Figure 3-1	– Bus trajectories using combined SE and SD data versus high-resolution data.....	9
Figure 4-1	– Bus trajectory using the combined dataset.....	10
Figure 4-2	– GIS flowchart for processing non-overlapping route data.....	11
Figure 4-3	– Hypothetical overlapping transit routes.....	12
Figure 4-4	– Point distribution density for real data (top) and normal approximation (bottom)..	14
Figure 4-5	– Point distribution density for real data compared with normal approximation.....	14
Figure 4-6	– Speed histogram at $x = 2,000$ ft. (left) and $x = 200$ ft. (right) from Figure 4-5.....	15
Figure 4-7	– Speed histogram at $x = 125$ ft. (left) and $x = 1,925$ ft. (right) from Figure 4-5.....	15
Figure 4-8	– Percentile estimation process using random non-normal data.....	16
Figure 4-9	– Speed histogram at $x = 700$ ft. from Figure 4-5 and associated distributions.....	18
Figure 5-1	– Map and aerial views of study area for Lombard Street.....	23
Figure 5-2	– Streetmix and Google Streetview of Lombard Street cross section before (top) and after (bottom).....	24
Figure 5-3	– Map and aerial views of study area for NE 16 th Street.....	25
Figure 5-4	– Streetmix and Google Streetview of 16 th Street cross section before (top) and after (bottom).....	26
Figure 6-1	– Change in travel time by percentile on Lombard Street.....	27
Figure 6-2	– Difference in travel speeds between before-and-after cases along Lombard Street	28
Figure 6-3	– Travel time by percentile for 16 th Street.....	29
Figure 6-4	– Difference in travel time by percentile for 16 th Street.....	29
Figure 6-5	– Difference in travel speeds between before-and-after cases along 16 th Street.....	30
Figure 6-6	– Northbound speed variability and speed variability index on 16 th Street.....	31
Figure 9-1	– Eastbound travel speed by percentile on Lombard Street. Top: before road diet. Bottom: after road diet. Used for Figure 6-2 (top).....	37

Figure 9-2 – Westbound travel speed by percentile on Lombard Street. Top: before road diet.
Bottom: after road diet. Used for Figure 6-2 (bottom). 38

Figure 9-3 – Eastbound speed variability and speed variability index on Lombard Street. 39

Figure 9-4 – Westbound speed variability and speed variability index on Lombard Street. 40

Figure 10-1 – Northbound travel speed by percentile on 16th Street. Top: before road diet.
Bottom: after road diet. Used for Figure 6-5 (top)..... 41

Figure 10-2 – Southbound travel speed by percentile on 16th Street. Top: before road diet.
Bottom: after road diet. Used for Figure 6-5 (bottom). 42

Figure 10-3 – Northbound speed variability and speed variability index on 16th Street. 43

Figure 10-4 – Southbound speed variability and speed variability index on 16th Street. 44

EXECUTIVE SUMMARY

In the last decade there has been a national trend toward projects that involve roadway space reallocation across modes. Many of these projects may include road diets and are typically very controversial when automobile travel lanes and/or curbside parking are proposed for removal.

The traditional process of identifying corridors for road diet improvements involves selecting potential corridors (mostly based on identifying four-lane roads) and conducting a traffic impact analysis of proposed changes on a selected roadway before implementing changes. The evaluation of roadway reallocation projects should include the analysis of traffic volumes, level of service (LOS), speeds, queue lengths and bus operations. There are tools and equipment to evaluate effectively traffic volumes and level of service changes in before-and-after studies. However, the detailed evaluation of speed and queue length distributions along a segment are significantly more cumbersome. In addition, the exhaustive evaluation of bus operations requires detailed data and specific tools and cannot be readily accomplished with field measurements or conventional LOS studies.

The data collection that is necessary to perform a detailed evaluation of speed and queue length distributions along a segment may be too expensive if conducted using traditional speed measurement tools. Furthermore, a complex methodology to evaluate of speed and queue length distributions may also be too burdensome for most transportation agencies evaluating potential road reallocation projects. To avoid these potential barriers, this research utilizes high-resolution transit data. This data is freely available, though it requires a lot of data processing. However, as data analytics tools and packages (many of them open source and freely available online) become more mainstream, the latter potential barrier is likely to decrease rapidly.

This research presented a general methodology for the detailed evaluation of transit operations and speed and queue length distributions along roadway reallocation projects. The proposed strategy and methodology is based on the utilization of high-resolution transit datasets. The focus of this research is on the development of a practical, general and theoretically sound methodology that can be applied to future roadway reallocation projects and applicable in a wide range of traffic conditions and locations.

This research provides a strategy and formulas to quantify changes in transit speeds and travel times and use confidence intervals (without the need to assume a distribution) to determine if that change was significant. The integration of transit high-resolution, stop-event and stop-disturbance data provides more information than any one dataset can provide, which improves resolution of the results. Finally, the methodology is applicable across a range of locations, traffic volumes, roadway types and roadway modification projects; as such, it can be applied broadly to any segment or network. The methodology proposed in this research was applied successfully to two separate road diets in Portland, OR.

1.0 INTRODUCTION

1.1 MOTIVATION

The national trend toward projects that involve roadway space reallocation across modes includes different types of treatments or projects. Sometimes roadway space reallocations are related to the concept of Complete Streets that are designed and operated for all users (not just motorized vehicles). The Federal Highway Administration (FHWA) defines road diets as “removing travel lanes from a roadway and utilizing the space for other uses and travel modes” (Knapp et al., 2014). Other types of projects may narrow motorized travel lanes to provide medians or other treatments that tend to reduce automobile travel speeds. In this research, the term *roadway reallocation* is utilized whenever the space allocated for motorized travel or parking is reduced.

Roadway reallocation projects are typically very controversial when automobile travel lanes and/or curbside parking are proposed for removal. When the removed travel lane or parking is reallocated for pedestrian, bicycle or transit utilization policymakers are often bombarded with complaints that emphasize the cost of the reallocation, potential travel time delays for passenger vehicles, and potential customer access problems for businesses. Some complaints also argue about “fairness” and cite the fact that, in many cases, roadway reallocation projects are paid by fuel taxes.

The evaluation of roadway reallocation projects should include the analysis of traffic volumes, level of service (LOS), speeds, queue lengths, and bus operations (Knapp et al., 2014) . There are tools and equipment to evaluate effectively traffic volumes and LOS changes in before-and-after studies. However, the detailed evaluation of speed and queue length distributions along a segment are significantly more cumbersome. The exhaustive evaluation of bus operations requires detailed data and specific tools, and cannot be readily accomplished with field measurements or conventional LOS studies.

1.2 RESEARCH OBJECTIVE

The objective of this research is to formulate a strategy and methodology for the detailed evaluation of transit operations and speed and queue length distributions along roadway reallocation projects. The proposed strategy and methodology is based on the utilization of high-resolution transit datasets. High-resolution GSP data integrated with stop-event (SE) and stop-disturbance (SD) data can be successfully used to examine speed and travel time changes on a roadway before and after a modification. Two case studies are used to show the range of applicability of this methodology and to highlight how it may be applied in the future.

The focus of this research is on the development of a practical methodology that can be applied to future roadway reallocation projects and applicable in a wide range of traffic conditions and locations. This research does not addresses general questions about road diets or their effectiveness.

2.0 LITERATURE REVIEW

2.1 ROAD DIETS

A road diet is a technique in transportation engineering and planning where the number of vehicle travel lanes or lane widths are reduced by altering lines without changing physical road structure. The additional space is often used to add sidewalks, bicycle travel lanes, center left-turn lanes, transit lanes, or other non-travel features such as planters. Once implemented, their effectiveness can be examined through a variety of before-and-after studies. According to the Federal Highway Administration (FHWA), road diets offer low-cost and high-return improvements when applied to typical four-lane highways and can result in crash reductions from 19-47% (Federal Highway Administration, 2016). Furthermore, the addition of bike lanes has been shown to increase the number of cyclists during peak commuting periods by more than 200% without negatively affecting automobiles, despite vehicle lane-count or lane-width reductions (Gudz, Fang & Handy, 2017).

Annual Average Daily Traffic (AADT) and peak-hour traffic volumes provide metrics by which road diet feasibility is determined. Road diets have been implemented on roads with AADTs from 8,500 to 24,000 vehicles per day (vpd) (Knapp et al., 2014). Many case studies were reported to have improved or unchanged operations. The FHWA recommends that roadways with an AADT less than 20,000 vpd are good candidates for a road diet (Federal Highway Administration, 2004). In terms of peak-hour traffic volumes, the FHWA recommends that road diets should be considered for roadways with less than 750 vehicles per hour per direction.

2.2 ROAD DIET BEFORE-AND-AFTER STUDIES

It is widely accepted that road diets present a beneficial tradeoff in terms of safety, cost and improvements for automobiles, cyclists and pedestrians. Several sources, including the FHWA, state that adding transit lanes is often a part of a road diet and it is generally understood that transit-only lanes improve transit service; however, there is little to no research about the effect of a road diet on transit when the addition of a transit lane is not part of the project.

A study funded by Smart Growth America (Anderson & Searfoss, 2015) argues that Complete Street projects tend to improve safety and increase levels of bicycle and pedestrian activity. After reviewing 37 projects this study claims that Complete Streets projects, when compared to conventional automobile-focused projects, are an inexpensive way to achieve transportation goals. Safety improvements and potential savings associated with reduced crashes, fatalities and injuries are largely responsible for the direct economic benefits that outweigh construction costs. This study also claims that Complete Streets projects are linked to economic gains like increased employment, higher property values and, in some cases, higher private investment. The Complete Streets study also recognizes that more data are needed to conclusively connect Complete Streets with economic success (Anderson & Searfoss, 2015). Out of the 37 projects, the authors were able to collect a moderate amount of data in 11 projects; in many communities, positive economic outcomes are backed mostly by anecdotal evidence.

2.3 TRANSIT BEFORE-AND-AFTER STUDIES

Studies that quantify the changes to routes before and after a roadway design change are not new. These studies exist for personal vehicles, roadway geometry, airplanes, buses, bikes, safety, passengers, and most other features of transportation. Measuring the impact of a change is a primary way that new systems are validated and added to the accepted practice.

For transit, before-and-after studies have used a variety of methods to study the effect of changes on perception, ridership, property values and transit performance. Watkins et al. (2011) used survey and observational data to examine how riders' perceptions of wait times changed with the introduction of real-time bus tracking. Dell'Olio et al. (2010) used focus groups to understand what types of information most improve riders' opinions of transit quality. Brown & Werner (2008) also used surveys but added ridership reports to determine, quantitatively, changes resulting from a new light rail stop. Rodríguez & Targa (2004) examined property values within walking distance of new bus rapid transit stops, and Kimpel et al. (2005) used archived transit data to quantify changes in travel times, on-time performance, and passenger wait times following implementation of transit signal priority (TSP). However, the research by Kimpel et al. was limited to TriMet's stop-level data, as other automated data collection methods had not yet been implemented. While high-resolution archived transit data has been used to quantify bus performance metrics, it has yet to be incorporated into before-and-after studies on public transportation.

Descriptive before-and-after studies that quantify performance or explain behaviors are also used for traffic assessment. For example, Farmer et al. (1997) looked at fatal crashes before and after the introduction of anti-lock brakes, and compared results using the difference in an estimated risk ratio and the associated 95% confidence interval calculated from the collected data. Huang & Cynecki (2000) looked at the impact of traffic calming systems across the United States for pedestrian wait times using t-tests to measure differences in mean values. These types of statistical comparisons are common with large sample sizes and appear in many fields. For transportation, before-and-after studies are less common than studies that aim to quantify performance along known segments.

2.4 BUS SPEEDS AND TRAVEL TIMES

The *Transit Capacity and Quality of Service Manual* indicates that a combination of internal factors (e.g., vehicle quality, age, availability, driver experience, route length, scheduling, and control strategies) and external factors (e.g., weather, traffic signals, congestion, demand variable, construction, wheelchair use, and passenger movements) contribute directly to service reliability and variability (Kittelsohn & Associates et al., 2013). For years, research efforts have focused on quantifying the impact of these factors.

To estimate travel times and trajectories, researchers needed proxies. Early research efforts revealed that buses experience the same long delays as automobiles, but that vehicles do not experience the same delays as buses, as is the case when buses dwell at stops because they are ahead of schedule (Hall & Vyas, 2000; Cathey & Dailey, 2002). In particular, TriMet buses have been used to evaluate arterial performance for automobiles and transit (Bertini & Tantiyanugulchai, 2004; Berkow et al., 2008).

Travel times have been estimated using SE data combined with aggregated data from signal loop detectors, green times, cycle lengths, and offsets for the signals in a corridor. Researchers have also used this data to help study factors that may affect bus travel time and service reliability at the point-segment level, the stop-to-stop segment level, and the route level (Hall & Vyas, 2000; Bertini & El-Geneidy, 2003; Chakroborty & Kikuchi, 2004). In particular, recent research focused on the performance of the adaptive traffic signal system (SCATS (Slavin, Feng & Figliozzi, 2013) and the impact of TSP on transit performance (Albright & Figliozzi, 2013). Other research focused on the impact of air quality at bus stops (Moore et al., 2012), sidewalks at intersections (Slavin & Figliozzi, 2013), and sidewalks at mid-block locations (Moore et al., 2012). Additionally, Feng et al. (2015) successfully integrated detailed signal timing SE data to estimate the impact of traffic volumes and intersections simultaneously on bus travel times. Each of these methodologies have added useful information to the stop-level data, but the representation remains an average between stops due to the nature of stop-to-stop data.

2.5 TRANSIT ROUTE-LEVEL ANALYSIS

Recently, some researchers began using the high-resolution data provided by TriMet. High-resolution data can come from other sources, such as cellphone Bluetooth data (Zhu, Holden & Gonder, 2017); however, only a handful of agencies have this data collection technology onboard their buses. The prediction of dwell times improved with the introduction of variables generated from the high-resolution data. These models included stop events that many previous studies dropped, such as stop location directly preceding or following intersections. Additional variables indicating whether a specific bus stopped at a red light prior to crossing an intersection as well as variables for the traffic speed immediately surrounding a stop improved the R^2 of the models. For stops near intersections, the new variables improved the adjusted R^2 by 200-300% (Glick & Figliozzi, 2017a)

High-resolution data can be used to create higher-resolution bus trajectories between bus stops, categorized speed breakdowns, and identified signal/queuing delays without including additional data sources (Glick et al., 2014). High-resolution data reduced much of the guesswork and the need for non-bus proxies in understanding bus performance between stops, which improved the applicability of buses as probe vehicles. These same researchers expanded their use of high-resolution data to multistop segments by producing space-time-speed diagrams that highlighted the locations of slow speed, but only provided an average speed by time or location (Stoll, Glick & Figliozzi, 2016). The first step in this process used heat maps to show high-density clusters of GPS data. Because the GPS data records at a steady rate, these clusters indicate areas of slow speeds. While this study showed the location and usage of bus stops, intersections and crosswalks, it did not provide a means to quantify the stopping behavior or identify what was causing the delays.

Sidhu et al. have used the same data to make improvements to existing travel time models using a statistically significant inter-stop trip time model to determine the number of signalized intersections encountered on a given route (Sidhu, Bertini & Pande, 2017).

Improving on previous results, high-resolution data were used to create performance metrics for larger segments by aggregating the data by time and location. This aggregation allowed for

percentiles and confidence intervals to be calculated for specific times and locations. Perhaps more importantly, this study created the basis of the methodology used in this research for removing the influence of bus stops from the data. Segments immediately surrounding stop events are removed from the dataset, leaving only pass-by (Glick & Figliozzi, 2017b). This type of analysis creates performance metrics that overcome the traditional problem of using buses as probes: Buses stop to serve passengers at locations that cars do not; thus, buses have not been able to accurately represent vehicle travel before now.

The ability to quantify behaviors between stops using the buses themselves is a rapidly evolving field. However, none of the current studies were able to fully remove the effects of the stops themselves to use buses as probes. The research presented here introduces disturbance data into the data, which allows for the influence of timepoint and pseudo-timepoint stops that remained in previous studies. Furthermore, this study outlines an approach to examine overlapping and diverging routes, which none of the previous analysis used.

3.0 DATA SOURCES

3.1 TRAFFIC VOLUMES

AADT and peak-hour traffic volumes provide key metrics by which road diet feasibility is determined. Traffic data were obtained from the Portland Bureau of Transportation (PBOT, 2017).

3.2 TRANSIT DATA

Archived transit data are widespread across transit agencies and their uses are integrated into the transit system. Tried-and-true data collection systems make up the core of transit data collection; operators, planners and app makers apply well-established methodologies daily. However, when archived data are used, stop-event (SE) data still dominate the discourse and practice. While newer high-resolution systems exist, they are not widespread and few practice-ready methodologies exist.

This study relied on two types of data: automatic vehicle location (AVL) and geographic information systems (GIS) shapefiles. The Tri-County Metropolitan Transportation District of Oregon (TriMet), Portland's public transportation provider, has archived AVL data for all trips since 1997; in 2013, it updated its bus dispatch system (BDS) with a high-resolution, GPS-based data collection system with five-second, time-based, resolution. The SE, stop-disturbance (SD) and high-resolution data serve as the basis of TriMet's data collection systems. TriMet also maintains onboard video recordings of all trips; however, TriMet, as with many agencies, erases video on a weekly cycle unless an incident occurs or if requested for a specific date.

Upon request, TriMet provided three sets of AVL data: SE, SD, and high resolution. Each of the AVL datasets represent the same buses, routes and times, which allows for comparisons and integration. Each type of data includes information absent in the others; as such, the visuals and comparisons obtained using a combined dataset provide a more detailed account of individual and aggregated bus behaviors.

To augment these datasets, this research also uses GIS files provided on TriMet's public website. The GIS data provide a base network to compare route locations; this process divides networks into individual segments with a unique set of routes. Each unique segment can analyze all routes at once to correct for bias created when only one route is examined on a segment with multiple lines. It also provides a base line to account and correct for segments in which buses have deviated from their assigned paths.

While TriMet focuses on data collection specific to its vehicles, past researchers have incorporated other data collection systems in their work, including propriety collection and survey methods of surface streets such as stationary Bluetooth or roadside radar sensor. This research relies on some

of the assumptions of those past studies; for example, when buses are between stops, they maintain speeds akin to the rest of traffic.

3.2.1 Stop-to-Stop Data

Stop-to-stop data, also called stop event (SE) data for this study, collect information at bus stops whether or not a bus actually serves passengers at a given stop. These data consist of bus operational data information including, but not limited to, arrive time, leave time, dwell time, average speed between stops, and passenger movements. SE data are widespread across transit agencies and usually records the number of passenger boardings (ons), alightings (offs), lift usage, door usage, and estimated passenger load; this study, which focuses on performance between stops, does not consider passenger movements.

Except when SD or high-resolution data collection system are available through transit agencies, SE data provide the primary means for researchers to determine route-level performance metrics for that agency. Unfortunately, the use of SE data only allow for averages between bus stops. As such, performance metrics near signalized intersections, on congested segments or with spaced bus stops lacks spatial accuracy. While it may be possible to determine that a problem is occurring between two stops with a high degree of accuracy, the specific location of the problem remains uncertain without additional data sources.

3.2.2 Stop-Disturbance Data

SD data expand on the information collected in the SE dataset by also including points between stops where the wheels of the bus stop moving. At each of these locations, the dataset records time, door activity and stop type. Timepoints, a stop type that denotes locations where buses have a specific arrival time, are locations where drivers correct for discrepancies in their arrival time versus TriMet's posted schedule. When late, they attempt to depart quickly; when early, they stay until they are back on schedule. Other stop types, such as unscheduled stops and pass-throughs, are also included and can provide additional insights into bus behavior. Unlike SE data, no passenger movement information is included in SD data.

Despite this lack, SD data can provide a more accurate view of transit behavior between stops than traditional SE data. Estimates are still required between points of zero motion, but periods of no motion that occur between stops provide a broader picture. However, this would not be helpful in determining the difference between an individual bus that traveled at 41 mph (66.0 kph) for two minutes and then 5 mph (8.0 kph) for one minute from another bus that traveled at 29 mph (46.7 kph) for three minutes; for that, additional information is required.

3.2.3 High-Resolution Data

High-resolution data, which collects data in up to five-second intervals, augments TriMet's previously implemented SE and SD datasets and provides a means to overcome some of the limitations of the other data sets, revisit factors influencing bus and route performance metrics, and examine inter-stop behavior of buses.

High-resolution data collects no passenger movement information or bus operational data except time and position information; while the resolution can be up to five seconds, data are not recorded if the bus is not in motion at the time of the next scheduled recording. This creates a situation where the high-resolution data shows low speeds in segments where a bus stopped rather than no speed. In Figure 3-1, the horizontal lines of the stop-level data, shown in red, are bus stops; the high-resolution data always shows a positive slope even at locations where it is known the bus stops. Integrating the datasets can correct for these problems.

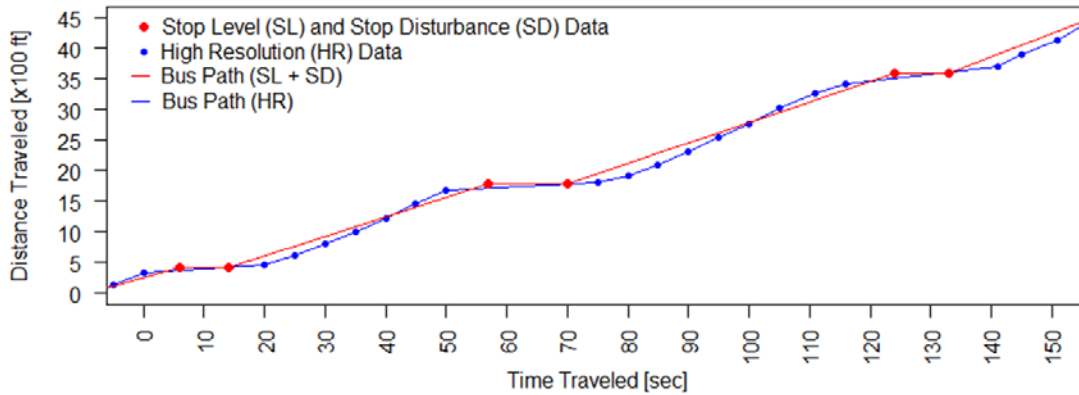


Figure 3-1 – Bus trajectories using combined SE and SD data versus high-resolution data.

4.0 METHODOLOGY

This research relies heavily on the R programming language in R-Studio interface and ArcMap software to clean and process the data. After merging the datasets, the three routes in the analysis had a combined 37.1 million rows and 37 columns. Due to the amount of data, all codes incorporate open source memory management, multithreading and multicore packages. A substantial amount of effort was dedicated to data cleaning and integration. The novel methodological contributions of this research are detailed in the remainder of this section.

4.1 TRIMET DATA INTEGRATION

Three data sets exist (SE, SD and high resolution) that each cover the same time periods, routes and buses with some exceptions, often the result of differing collection parameters. These datasets must be cleaned to provide a uniform set of trips, which requires a unique identification number across all sets. Since all datasets include a bus number, time and a date, these serve as a starting point.

Data integration, using a unique identification number (UID), begins with SD and SE data. Individual trips are separated based on UID, route number and direction along that route. High-resolution data does not include route and direction and therefore must be compared to the other sets to separate out individual trips based on time.

When a bus dwells at a bus stop, these locations are also recorded as SD so the events are duplicated in the combined set; pass-by stops would not be recorded as SD and are therefore not duplicated. Once a single record exists for each event, individual events with a different arrive time and leave time (i.e., events where the bus stops for any length of time) are duplicated so each row represents a single point in time. This duplication allows for direct integration with high-resolution data. Following this step, the high-resolution data is interwoven by timestamps with the SE and SD data to provide a complete picture of the bus's trajectory, an example of which is shown in Figure 4-1.

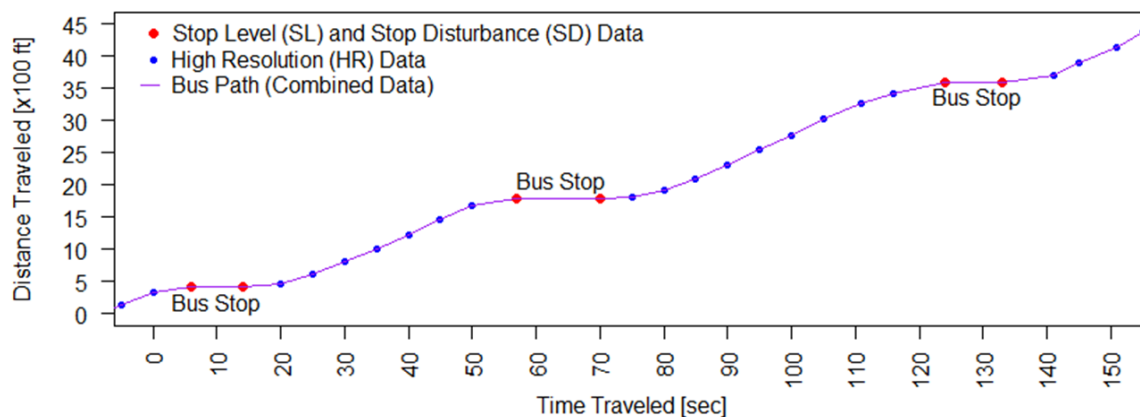


Figure 4-1 – Bus trajectory using the combined dataset.

The integrated data also provide a means to quantify roadway behaviors that account for or remove bus-stopping behavior, which allows for estimates of general traffic behavior. However, the data still contain multiple routes that sometimes, but not always, overlap. As such, GPS data must be assigned to an underlying map that shows which routes are present on each segment and to separate out different directions of travel for the same route.

4.1.1 Spatial Data Cleaning

The spatial data provided by TriMet shows the location of all stops and routes as GIS shapefiles (a file format commonly used in mapping programs). However, the GIS files lacked the spatial accuracy to allow for simultaneous analysis for multiple routes. The problem arises from routes that travel on the same road but are encoded in such a way that do not exactly overlap. These files are loaded into ArcMap and processed to clean up discrepancies between routes traveling on the same streets.

The process of correcting for problems in the spatial data begins with grouping routes on the same road. In their raw form, routes do not precisely overlap. Following the process outlined in Figure 4-2, a final map of routes with a single line representing each segment is produced. Importantly, this map can be exported to provide GPS coordinates for additional processing within R.

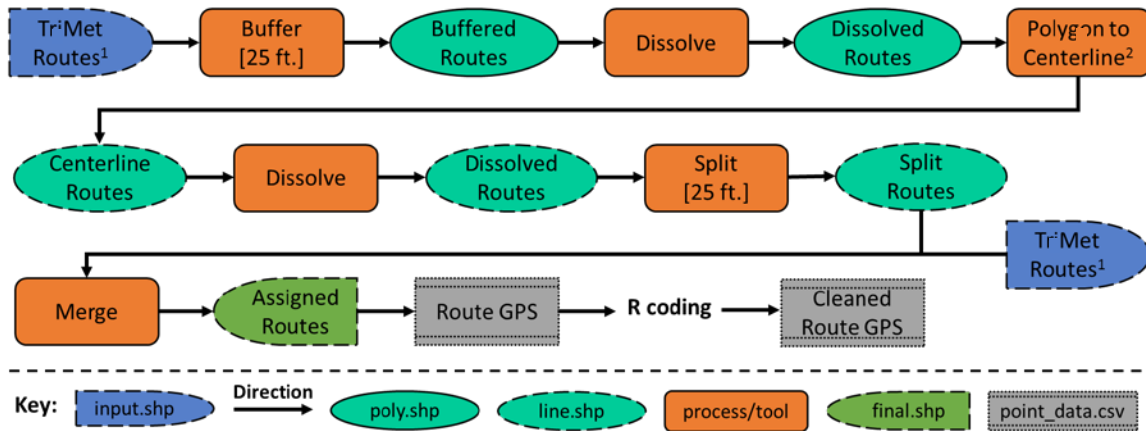


Figure 4-2 – GIS flowchart for processing non-overlapping route data.

4.1.2 GIS Tool Definitions

The tools used in Figure 4-2 are defined in Table 4-1. Most are part of ArcMap’s basic license. However, Polygon to Centerline is an added tool that requires an advanced license.

Table 4-1 – GIS tools definitions with visuals

Tool	Definition	Visual Representation
Buffer	Creates a polygon around input features at a specified distance.	
Dissolve	Combines like features based on a specific attribute or combination of attributes.	
Polygon to Centerline	Creates a centerline profile of a polygon input.	
Split	Divides an input based on a specified distance into a set of inputs that sum to the size of the original input.	
Trim Line	Removes a portion of a line that extends a specific distance past the intersection of two lines.	
Merge	Combines multiple input datasets of the same type into a single, new output dataset with combined attributes.	

4.1.3 Final Combined Dataset

The final combined dataset is not necessarily one continuous route. Instead, multiple routes overlap to create a system of routes. This system is divided into a set of unique road segments where the same routes run from the beginning of each to the end. In Figure 4-3, a hypothetical system of routes has been divided into a set of unique road segments represented by numbers 1-9. Routes shown next to each other shall exactly overlap in the data. While segments 4 and 9 have the same routes, they are numbered separately, as they are not continuous. Combining segments 1, 4, 6 and 9 would give a complete picture of the red route after analysis.

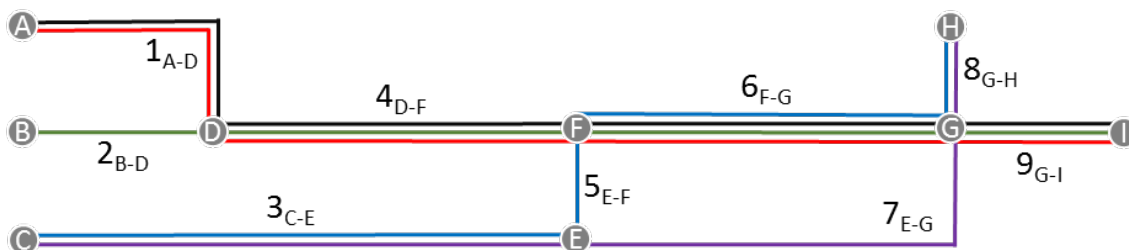


Figure 4-3 – Hypothetical overlapping transit routes.

4.1.4 Data Irregularities

Processing the high-resolution data revealed irregularities in the data collection that deserve special attention and further study. A small but notable percent of trips reported in the data were not limited to the route designated by the data (i.e., the buses deviated from their assigned routes). These buses included information from deadheads, which are trips made while the bus was not in service, bus parking locations and detours. Cumulative distance calculations using the data itself without corrections for location using coordinates of a known route may result in errors of incorrectly assigned locations if not properly filtered. It appears that this problem is limited to the high-resolution data; as such, integrating the SL and SD data can eliminate many of the incorrect reports. The remainder of the errors can be corrected by checking every point against the GIS spatial route data provided by TriMet; however, this process is computationally intensive.

4.2 DATA ANALYSIS

Because each route segment is composed of a combination of pieces of several distinct overall bus routes, the data in each segment must be examined separately. For example, while analyzing the red route from Figure 4-3, segment 1 consists of two bus routes, 4 consists of three, 6 consists of four, and 9 consists of three. Once each route segment has been processed, the results can be strung together to form a complete picture of the red route.

Each unique route segment, I , is divided into a set of n_I equal-length and non-overlapping sub-segments, i . A centerpoint, c_i marks the midpoint of each of these sub-segments. The combined data includes a set of n_{j_i} bus trips that pass through each sub-segment i , with each individual bus designated by the index j . Due to the data-cleaning process, n_{j_i} may fluctuate between adjacent sub-segments.

$$\begin{aligned} i \in I &= \{1, 2, 3, \dots, n_I\} \\ j \in J_i &= \{1, 2, 3, \dots, n_{j_i}\} \end{aligned}$$

4.2.1 Unknown Distributions

For each sub-segment, i , the velocity of each bus, v_{ij} , that passes a center point, c_i , is extracted. V_i , the set of velocities in sub-segment, i , do not always follow a known distribution. To illustrate, bus speeds were extracted for a 2,500-foot segment along Powell Boulevard at 25-foot increments. Figure 4-4 shows the point density distribution (as they would appear in a histogram) at each location (x) using real data (top) and random normal data based on the mean and standard deviation of the real data (bottom).

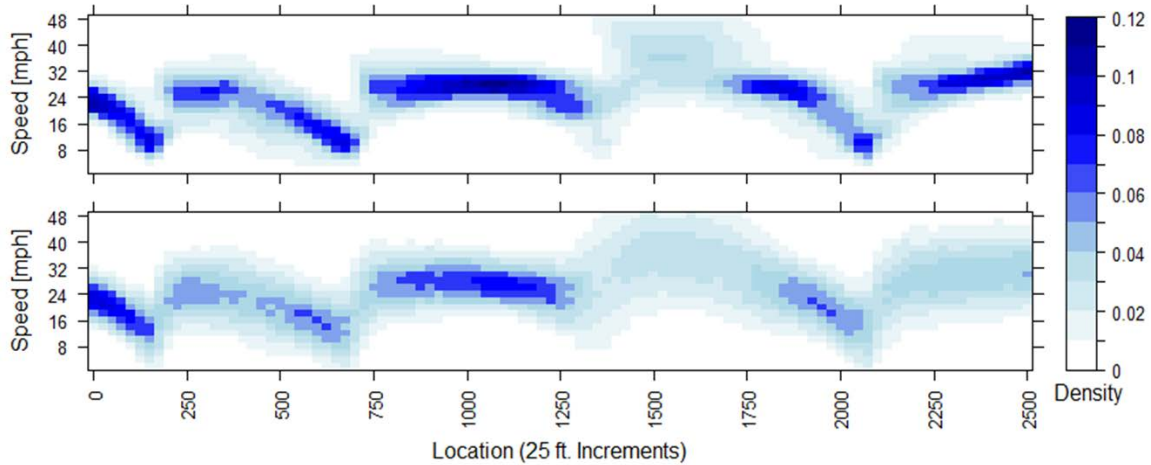


Figure 4-4 – Point distribution density for real data (top) and normal approximation (bottom).

For each location and speed-bracket combination, a percent difference in the histogram densities is calculated by subtracting normal densities from real densities then dividing by the normal densities. Figure 4-5 shows this percent difference between Figure 4-4 (top) and Figure 4-4 (bottom). To reduce visual clutter, Figure 4-5 shows only the percent difference for values within two standard deviations of the estimated normal mean. If the real density is higher than the normal, it is blue; if lower, it is red.

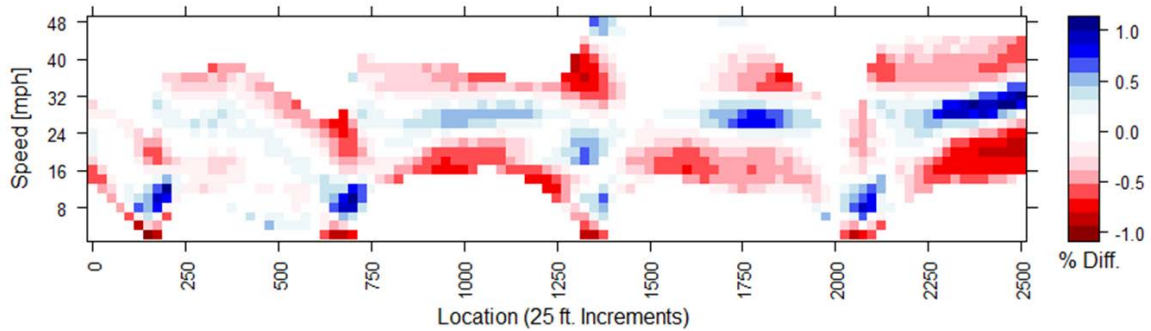


Figure 4-5 – Point distribution density for real data compared with normal approximation.

Some locations appear to follow a normal distribution, such as $x = 2000$ (Figure 4-6: left); however, most locations, like $x = 200$ (Figure 4-6: right), do not, nor do they follow any typical distribution.

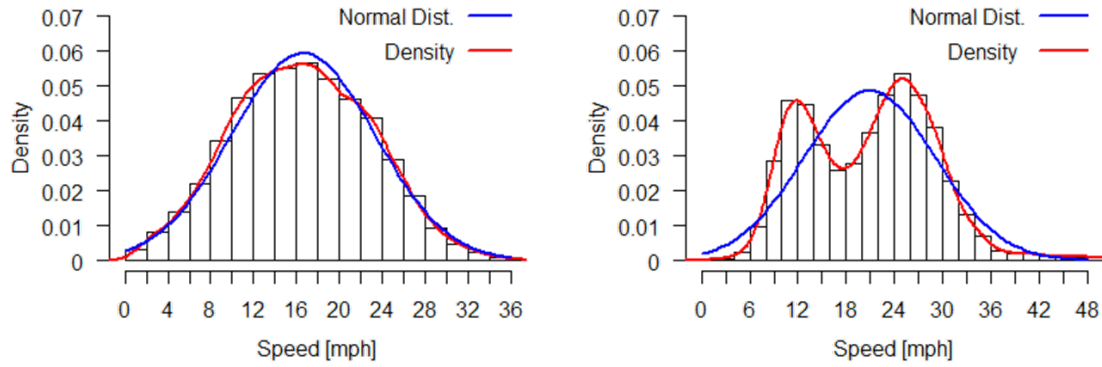


Figure 4-6 – Speed histogram at $x = 2,000$ ft. (left) and $x = 200$ ft. (right) from Figure 4-5.

Furthermore, when data do not follow a normal distribution, the mean and harmonic mean have the potential to give misleading or less than useful information; this is the case when data are skewed. Skewness is a measure of asymmetry of the probability distribution about a mean. Figure 4-7 shows a positive skew distribution from Figure 4-5 at $x = 125$ ft. (Fig. 4-7, left) and negative skew distribution from $x = 1,925$ ft. (Fig. 4-7, right). Non-skewed normal data typically have near-equal means, medians and mode. This is not the case for skewed data.

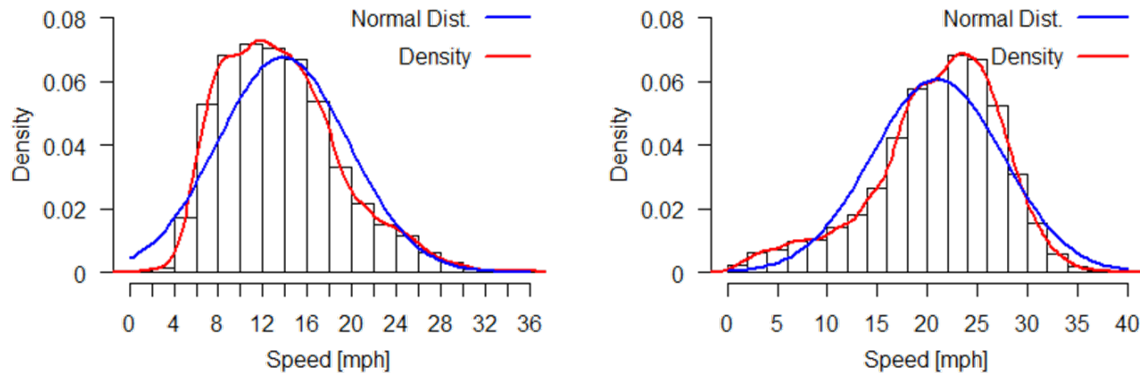


Figure 4-7 – Speed histogram at $x = 125$ ft. (left) and $x = 1,925$ ft. (right) from Figure 4-5.

Due to the range of possible speed distributions, an approach that does not require the distribution to be known is used to calculate statistics about the set of speeds, V_i . The methodology for calculating percentile speeds stems from a previously published journal article using high-resolution data (Glick & Figliozzi, 2017b), but with added methodologies for percentiles and confidence intervals.

p = percentile

$\hat{v}_{i,p}$ = estimated speed in segment s_i at percentile p

Any s_i has associated percentile travel speeds found by ordering the data and finding an observation of a specified percent up or down a list. The estimate of the variance for any percentiles of univariate data can be estimated through a cumulative distribution function (CDF) and its derivative, the probability distribution function (PDF). This is true of any given set of data of known or unknown distribution. Figure 4-8 shows a set of randomly generated non-normal data to provide an example of the percentile estimation process along with the progression used to

estimate confidence. A histogram of the data is shown (upper left). After ordering the data points and normalizing to form a CDF (upper right), spline smoothing is applied to create a continuous function that approximates the CDF (lower left). From this generated spline-smoothed function, the probability of each point can be calculated by taking the derivative of the CDF to produce an estimate of the PDF (lower right) for the set of data.

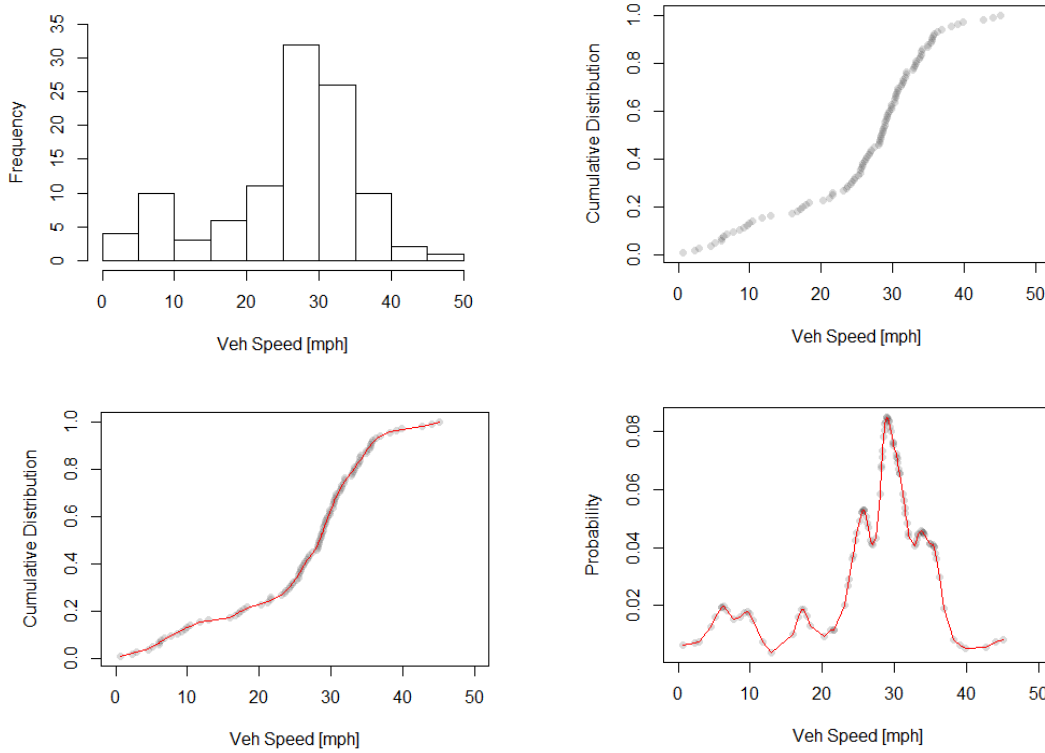


Figure 4-8 – Percentile estimation process using random non-normal data.

The estimate for the variance of speed for a given percentile, p , in segment s_i is denoted as $\hat{\sigma}_{v_{i,p}}^2$. To estimate this value, the following equation is utilized (Brown & Wolfe, 1983):

$$\hat{\sigma}_{v_{i,p}}^2 = \frac{p(1-p)}{f(\hat{v}_{i,p})^2 \cdot n_{J_i}} \quad (1)$$

Here, $f(\hat{v}_{i,p})$ is the probability of the PDF given the input velocity, $\hat{v}_{i,p}$, and n_{J_i} is the number of observations in each segment i . Assuming the number of observations is large (>160) (Brown & Wolfe, 1983), this estimate of variance may be used to estimate the confidence intervals for each $\hat{v}_{i,p}$, assuming a normal distribution. For a confidence level α and its associated z-score, $z(\alpha)$, the range of percentile values that may represent an estimated percentile is found:

$$[\hat{v}_{i,p} - \hat{\sigma}_{v_{i,p}} \cdot z(\alpha), \hat{v}_{i,p} + \hat{\sigma}_{v_{i,p}} \cdot z(\alpha)] \quad (2)$$

This interval provides extremes of the α confidence interval around $v_{i,p}$. Variances and standard deviations are calculated with above methodology (equation 1) (Brown & Wolfe, 1983) and a methodology that assumes normality is used in a previous publication (Glick & Figliozzi, 2017b).

4.2.2 Peak-Hour vs. Whole-Day Performance

A speed variability $\Delta\hat{v}_i$, is used to identify segments that are more heavily congested during the peak hour (Glick & Figliozzi, 2017b). It is calculated by subtracting the 15th percentile travel speed from the 85th percentile travel speed. When divided by the median travel time, a speed variability index ($\hat{\mu}_i$) or SVI is obtained for each segment (Glick & Figliozzi, 2017b).

$$\Delta\hat{v}_i = \hat{v}_{i,85} - \hat{v}_{i,15} \quad (3)$$

$$\hat{\mu}_i = \frac{\Delta\hat{v}_i}{\hat{v}_{i,50}} = \frac{\hat{v}_{i,85} - \hat{v}_{i,15}}{\hat{v}_{i,50}} \quad (4)$$

If a SVI is greater than 1, it indicates a location where median travel speed is more similar to the 15th percentile travel speed than the 85th.

4.3 PROPAGATION OF UNCERTAINTY

The estimates of uncertainty, $\hat{\sigma}$, that this has been previously calculated for each percentile as a standard deviation, $\hat{\sigma}_{v_{i,p}}$, should be carried through in the calculations for speed variability and speed variability index. Each of the estimated percentile travel speeds ($\hat{v}_{i,15}$, $\hat{v}_{i,50}$, and $\hat{v}_{i,85}$) has an associated and normally distributed error (i.e., the standard deviations, $\hat{\sigma}_{v_{i,15}}$, $\hat{\sigma}_{v_{i,50}}$, and $\hat{\sigma}_{v_{i,85}}$); thus, the speed variability and speed variability index are uncertain themselves. The resulting error for correlated and uncorrelated data can be estimated through the properties of normal distributions for $\Delta\hat{v}_i$ and appropriate formulae for the propagation of uncertainties for ratios, such as $\hat{\mu}_i$, (Taylor, 1997).

4.3.1 Assumption of Normality

The top left of Figure 4-9 shows the density and normal approximation of the distribution of travel speeds at $x = 700$ of Figure 4-5. This distribution does not appear to be normal (upper right). Using 100 samples of 150 non-replaced data points. The 15th, 85th and 50th percentiles were estimated 100 times. The distribution of the 15th (upper right), 50th (middle left) and 85th (middle right) percentiles appears to follow somewhat normal distributions.

Additionally, the distribution of the speed variability (equation 3) (bottom left) and speed variability index (equation 4) (bottom right) also appear to follow normal distributions. This indicates that the equations for the propagation of error will result in standard deviations where normality can be assumed.

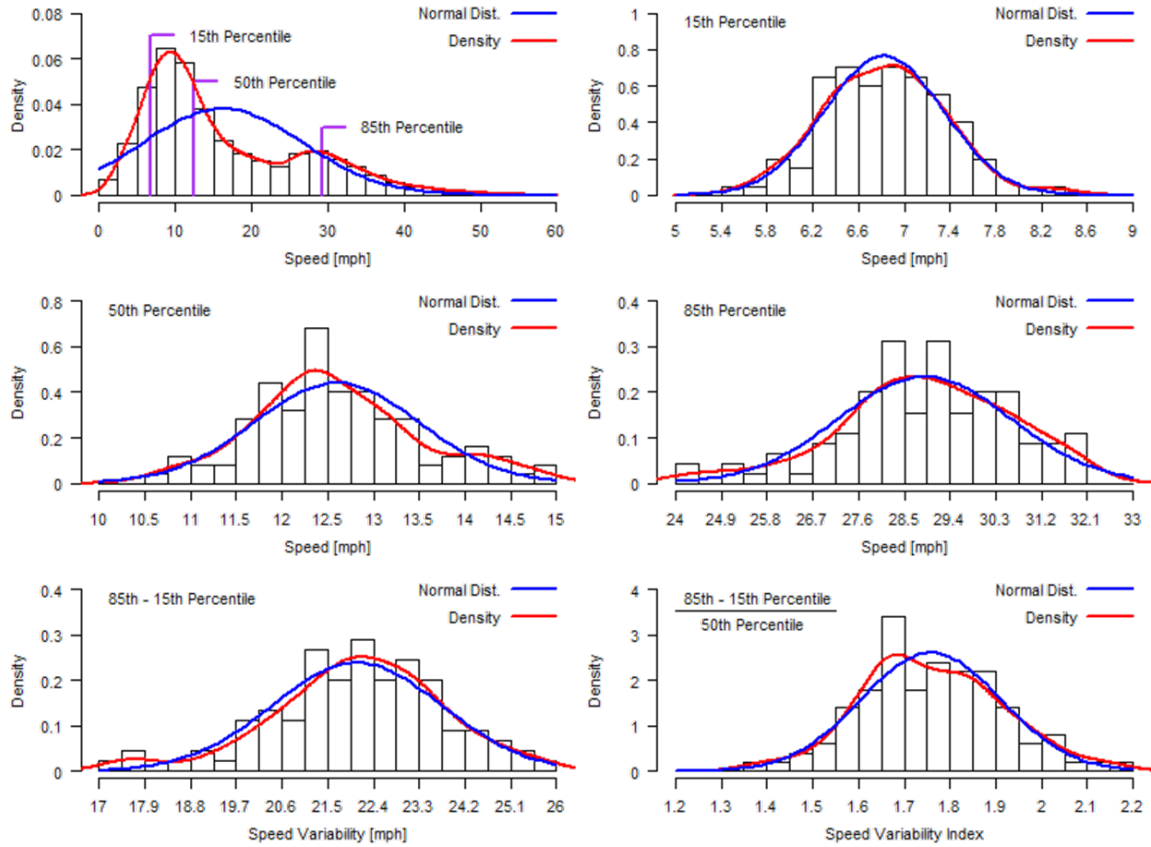


Figure 4-9 – Speed histogram at $x = 700$ ft. from Figure 4-5 and associated distributions.

4.3.2 Propagation of Uncertainty Addition and Subtraction

The addition and subtraction of normal distributions is also a normal distribution. The addition and subtraction of N values with independent uncertainties is defined generally as:

$$q = x_1 + \dots + x_N$$

where q is the result and the values x_1, \dots, x_N are independent and carry standard deviations $\sigma_{x_1}, \dots, \sigma_{x_N}$, respectively, then

$$\sigma_q = \sqrt{\sigma_{x_1}^2 + \dots + \sigma_{x_N}^2}$$

For $\hat{v}_{i,85}$ and $\hat{v}_{i,15}$, which are correlated values, the error for the speed variability can be estimated as:

$$\sigma_q = \sqrt{\left(\frac{\partial q}{\partial x_1}\right)^2 \sigma_{x_1}^2 + \left(\frac{\partial q}{\partial x_2}\right)^2 \sigma_{x_2}^2 + 2 \left(\frac{\partial q}{\partial x_1}\right) \left(\frac{\partial q}{\partial x_2}\right) \text{cov}(x_1, x_2)}$$

The partial derivatives for subtraction (e.g. $q = x_1 - x_2$), are $\frac{\partial q}{\partial x_1} = 1$ and $\frac{\partial q}{\partial x_2} = -1$. As such, the propagation of uncertainty for the difference, $\Delta \hat{v}_i$, can be written:

$$\hat{\sigma}_{\Delta v_i} = \sqrt{\hat{\sigma}_{v_{i,85}}^2 + \hat{\sigma}_{v_{i,15}}^2 - 2 \cdot \text{cov}(\hat{v}_{i,85}, \hat{v}_{i,15})} \quad (5)$$

4.3.3 Propagation of Uncertainty Product and Quotient

For the propagation of uncertainty for the product or quotient, normality (in the result) cannot be assumed. Estimates can be obtained through the Taylor series expansion. For a case of a general function q consisting of N projects and M quotients:

$$q = \frac{x_1 \cdot \dots \cdot x_N}{y_1 \cdot \dots \cdot y_M}$$

with independent uncertainties $\sigma_{x_1}, \dots, \sigma_{x_N}$ and $\sigma_{y_1}, \dots, \sigma_{y_M}$, the uncertainty of the quotient σ_q is defined by:

$$\frac{\sigma_q}{|q|} = \sqrt{\left(\frac{\sigma_{x_1}}{x_1}\right)^2 + \dots + \left(\frac{\sigma_{x_N}}{x_N}\right)^2 + \left(\frac{\sigma_{y_1}}{y_1}\right)^2 + \dots + \left(\frac{\sigma_{y_M}}{y_M}\right)^2}$$

However, for data involving correlated variables (such as the SVI), the covariance must be taken into account. For a simple ratio of two correlated values x_1 and x_2 with standard deviations σ_{x_1} and σ_{x_2} , respectively, q and its standard deviation, σ_q , are estimated as (Lee & Forthofer, 2006):

$$q = \frac{x_1}{x_2}$$

$$\frac{\sigma_q}{|q|} = \sqrt{\left(\frac{\sigma_{x_1}}{x_1}\right)^2 + \left(\frac{\sigma_{x_2}}{x_2}\right)^2 - 2 \cdot \frac{\text{cov}(x_1, x_2)}{x_1 \cdot x_2}}$$

As such, the uncertainty for SVI is estimated as:

$$\frac{\hat{\sigma}_{\mu_i}}{|\hat{\mu}_i|} = \sqrt{\left(\frac{\hat{\sigma}_{\Delta v_i}}{\Delta \hat{v}_i}\right)^2 + \left(\frac{\hat{\sigma}_{v_{i,50}}}{\hat{v}_{i,50}}\right)^2 - 2 \cdot \frac{\text{cov}(\Delta \hat{v}_i, \hat{v}_{i,50})}{\Delta \hat{v}_i \cdot \hat{v}_{i,50}}} \quad (6)$$

Finally, by propagating the uncertainty from the speed variability into equation 6, the final estimate for the SVI is:

$$\frac{\hat{\sigma}_{\mu_i}}{|\hat{\mu}_i|} = \sqrt{\frac{\hat{\sigma}_{v_{i,85}}^2 + \hat{\sigma}_{v_{i,15}}^2 - 2 \cdot \text{cov}(\hat{v}_{i,85}, \hat{v}_{i,15})}{\hat{v}_{i,85} - \hat{v}_{i,15}^2} + \left(\frac{\hat{\sigma}_{v_{i,50}}}{\hat{v}_{i,50}}\right)^2 - 2 \cdot \frac{\text{cov}(\hat{v}_{i,85} - \hat{v}_{i,15}, \hat{v}_{i,50})}{(\hat{v}_{i,85} - \hat{v}_{i,15}) \cdot \hat{v}_{i,50}}} \quad (7)$$

4.3.4 Confidence Intervals

The standard deviations, $\hat{\sigma}_{\Delta v_i}$ and $\hat{\sigma}_{\mu_i}$ can be used to estimate confidence intervals (CI) for peak hour performance when combined with a z-score and specified alpha:

$$CI_{\Delta v_i} = [\Delta \hat{v}_i - \sigma_{\Delta \hat{v}_i} \cdot z(\alpha), \Delta \hat{v}_i + \sigma_{\Delta \hat{v}_i} \cdot z(\alpha)] \quad (8)$$

$$CI_{\mu_i} = [\hat{\mu}_i - \sigma_{\hat{\mu}_i} \cdot z(\alpha), \hat{\mu}_i + \sigma_{\hat{\mu}_i} \cdot z(\alpha)] \quad (9)$$

If 0 falls within the upper and lower bound of the confidence interval, no speed variability can be said to exist between the 15th and 85th percentile (i.e., the null hypothesis cannot be rejected). An SVI greater than 1 indicates severe speed variability in a given segment, i .

4.4 TRAVEL TIMES

Travel times, t_j , between any two points are extracted from the data where j is a single bus. Percentiles and confidence intervals are calculated using the same methodology as before where n_j is the total number of buses and p is the percentile. \hat{t}_p is the percentile travel time with an estimated standard deviation of $\hat{\sigma}_{t_p}$.

$$\hat{\sigma}_{t_p}^2 = \frac{p(1-p)}{f(\hat{t}_p)^2 \cdot n_j} \quad (10)$$

The estimated average daily travel time for all buses, \bar{t} , is found by summing each percentile travel time as if it were an individual bus, then correcting for the number of buses n_j . On average each percentile travel time should be seen an equal number of times.

$$\bar{t} = \frac{1}{99} \left(\sum_{p=1}^{99} \hat{t}_p \cdot n_j \right) \quad (11)$$

The standard deviation of the average daily travel time, $\sigma_{\bar{t}}$, is found by summing the squares of the percentile travel times.

$$\hat{\sigma}_{\bar{t}} = \sqrt{\sum_{p=1}^{99} \hat{\sigma}_{t_p}^2} \quad (12)$$

Average cost per day is found by multiplying the average daily travel time, converted into hours, by the operational cost of TriMet for 2015, \$93.27 (TriMet, 2016).

4.4.1 Space-Time-Speed Diagrams

Speed data can also be viewed after aggregating the data by time of day through the use of a moving average within a range of times. These moving averages are calculated using the harmonic mean within each segment i for vehicles that fall within the time window, w . The set of velocities within the time window is denoted as W_i , which is a subset of V_i . Percentiles are not used for this visual, as this methodology already highlights areas of high performance and low performance.

$$\bar{v}_{ij} = \frac{n_{W_i}}{\sum_{W_i} \left(\frac{1}{v_{iw}} \right)} \quad \forall v_{iw} \in W_i \subseteq V_i \quad (13)$$

There are no variances calculated with this moving harmonic mean due to non-normality and low number of points within the moving window.

4.4.2 Before-and-After Comparisons

All methodologies previously discussed provide the means to determine the initial and current conditions of the route, which will be compared for travel speeds, travel times, speed variability and speed variability indexes. For all data two additional indexes, β_0 and β_1 , represent the initial conditions and current conditions, respectively. A δ added before a variable will indicate a value calculated from β_0 and β_1 .

4.4.3 Travel Speeds, Travel Times and Peak Performance

The differences in the percentile speeds, travel times and peak-hour performance metrics of β_0 and β_1 are compared using simple subtraction.

$$\delta \hat{v}_{i,p} = \hat{v}_{i,p,\beta_1} - \hat{v}_{i,p,\beta_0} \quad (14)$$

$$\delta \bar{t}_p = \bar{t}_{p,\beta_1} - \bar{t}_{p,\beta_0} \quad (15)$$

$$\delta \Delta \hat{v}_i = \Delta \hat{v}_{i,\beta_1} - \Delta \hat{v}_{i,\beta_0} \quad (16)$$

$$\delta \hat{\mu}_i = \hat{\mu}_{i,\beta_1} - \hat{\mu}_{i,\beta_0} \quad (17)$$

Since these equations all follow the same format, X represents the specific variables (i.e. $\hat{v}_{i,p}$, \bar{t}_p , $\Delta \hat{v}_i$, and $\hat{\mu}_i$)

$$\delta X = X_{\beta_1} - X_{\beta_0} \quad (18)$$

The estimate for standard deviation is the same for all speed and travel time variables denoted $\hat{\sigma}_{X_{\beta_0}}$ and $\hat{\sigma}_{X_{\beta_1}}$ for the before-and-after case, respectively. Since the before-and-after values are often correlated, the covariance is included in the estimate for their standard deviation:

$$\delta \hat{\sigma}_X = \sqrt{\hat{\sigma}_{X_{\beta_1}}^2 + \hat{\sigma}_{X_{\beta_0}}^2 - 2 \cdot \text{cov}(X_{\beta_1}, X_{\beta_0})} \quad (19)$$

As before, these estimated standard deviations can be used to estimate confidence intervals (CI) for all δX by including a z-score and specified alpha.

$$CI_{\delta X} = [\delta X - \delta \hat{\sigma}_X \cdot z(\alpha), \delta X + \delta \hat{\sigma}_X \cdot z(\alpha)] \quad (20)$$

If 0 falls within the CI, no statistically significant change can be said to have occurred (i.e., cannot reject null). If $\delta X - \delta \hat{\sigma}_X \cdot z(\alpha) \geq 0$, then the metric of interest (e.g., speeds, travel times, speed differences, speed variability, etc.) can be said to have increased. If $\delta X + \delta \hat{\sigma}_X \cdot z(\alpha) \leq 0$, then the metric of interest can be said to have decreased.

5.0 CASE STUDIES

The two case studies are located in Portland, OR: Lombard Street and NE 16th. Each location underwent a notable roadway change. The effect of those changes was measured using data collected before and after implementation of each change. The segment performance before the change is compared to performance after the change to determine overall effect. The purpose of these case studies is not to make broad claims about road diets. Rather, the purpose is to provide examples of how the proposed methodology can be applied in the future to other road diets or general before-and-after case studies.

5.1 CASE STUDY I: LOMBARD STREET

The first analysis is of a road diet consisting of a lane-width reduction on Lombard Street that extended for 1,100 feet. Counts available from the Portland Bureau of Transportation (PBOT, 2017) show an ADT of 6,800 and p.m. peak-hour volume of 650 vehicles for westbound travel, and an ADT of 6,400 with 500 vehicles during the p.m. peak hour for eastbound travel. These volumes easily meet the recommended criteria by FHWA for road diet implementation. Both the ADT and peak-hour traffic volumes fall well below the thresholds where the FHWA would recommend the consideration of other factors and, based on volumes, should not experience significant change to operations (Knapp et al., 2014).

Figure 5-1 shows the location of the road diet. The measurements on the map correspond to the x-axis of results. Road diet begins at 1,175 feet and ends at 2,275 feet. The middle and bottom images are the aerial view of the before-and-after conditions, respectively.

Figure 5-2 shows a Streetmix cross section with lane dimensions marked and a Google Streetview of the lane configuration before and after the road diet. The total width of roadway is 50 feet. The data collected before the roadway change (before data) includes approximately 2,300 buses from all weekdays between June 15 and July 31, 2015. The data collected after the change (after data) includes approximately 1,900 buses for all weekdays between July 20 and August 26, 2016.



Figure 5-1 – Map and aerial views of study area for Lombard Street.

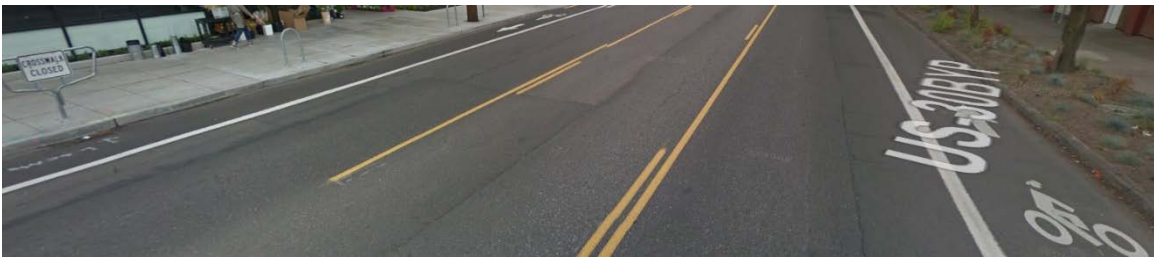
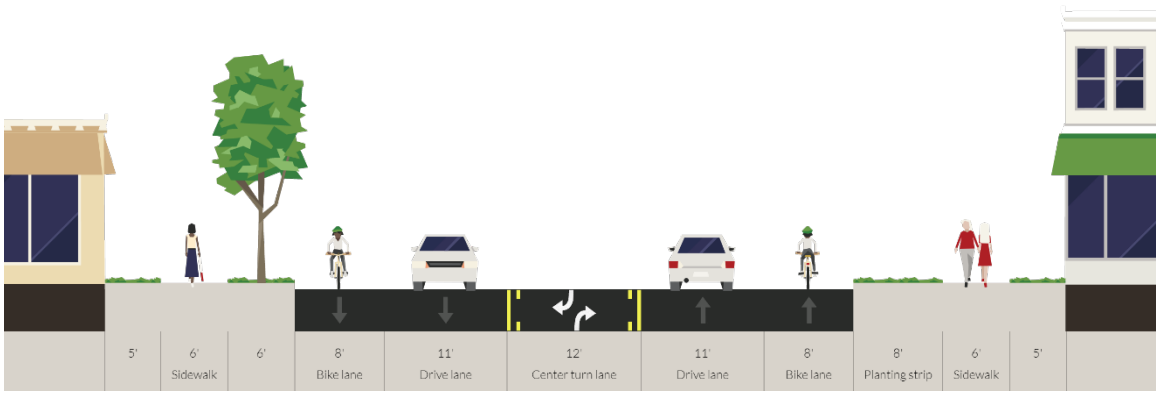
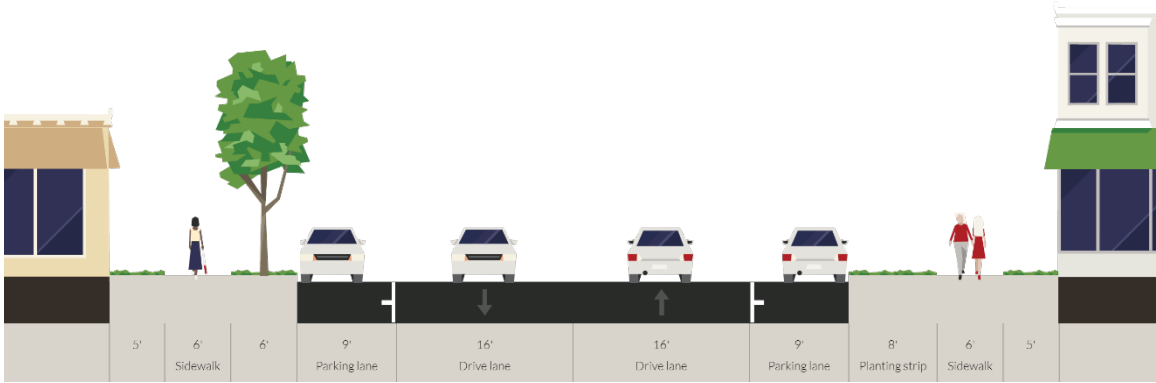


Figure 5-2 – Streetmix and Google Streetview of Lombard Street cross section before (top) and after (bottom).

5.2 CASE STUDY II: NE 16TH STREET

The second analysis is of a road diet consisting of a travel lane reduction at NE 16th Street. The initial cross section included a bike lane and two travel lanes in each direction. These lane markings were replaced with an 8-foot bike lane, 5-foot buffer and an 11-foot vehicle travel lane. Figure 5-3 shows a map (top) that includes measurements that correspond to an x-axis of results. The road diet begins at 125 feet and ends at 1,100 feet. The two aerial views show the road before (middle) and after (bottom) the road diet.

Using PBOT (2017) traffic counts, the average daily traffic counts are approximately 3,730 and 4,160 southbound and northbound, respectively, with p.m. peak-hour counts of 324 and 473. These daily volumes are lower than those of Lombard; as such, this segment is also not expected to have significantly altered operations. The before data include 1,100 buses from all weekdays between June 12 and July 3, 2015. The after data include 1,700 buses all weekdays between July 14 and August 12, 2016.

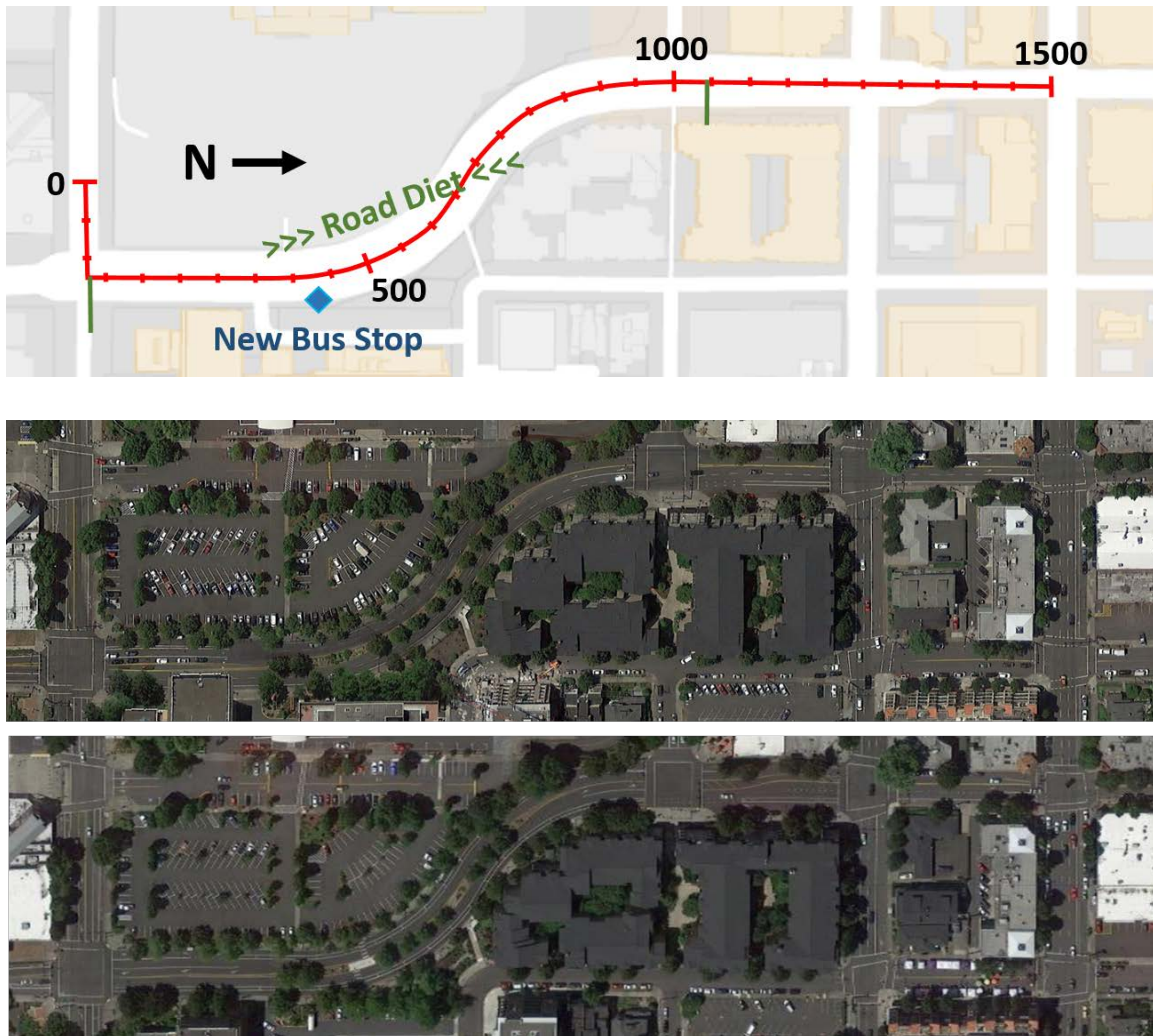


Figure 5-3 – Map and aerial views of study area for NE 16th Street.

The Streetmix and Google Streetview cross sections of 16th Street before and after the road diet are shown in Figure 5-4. The dimensions of the lanes given and the total width on either side of the median is 25.5 feet.

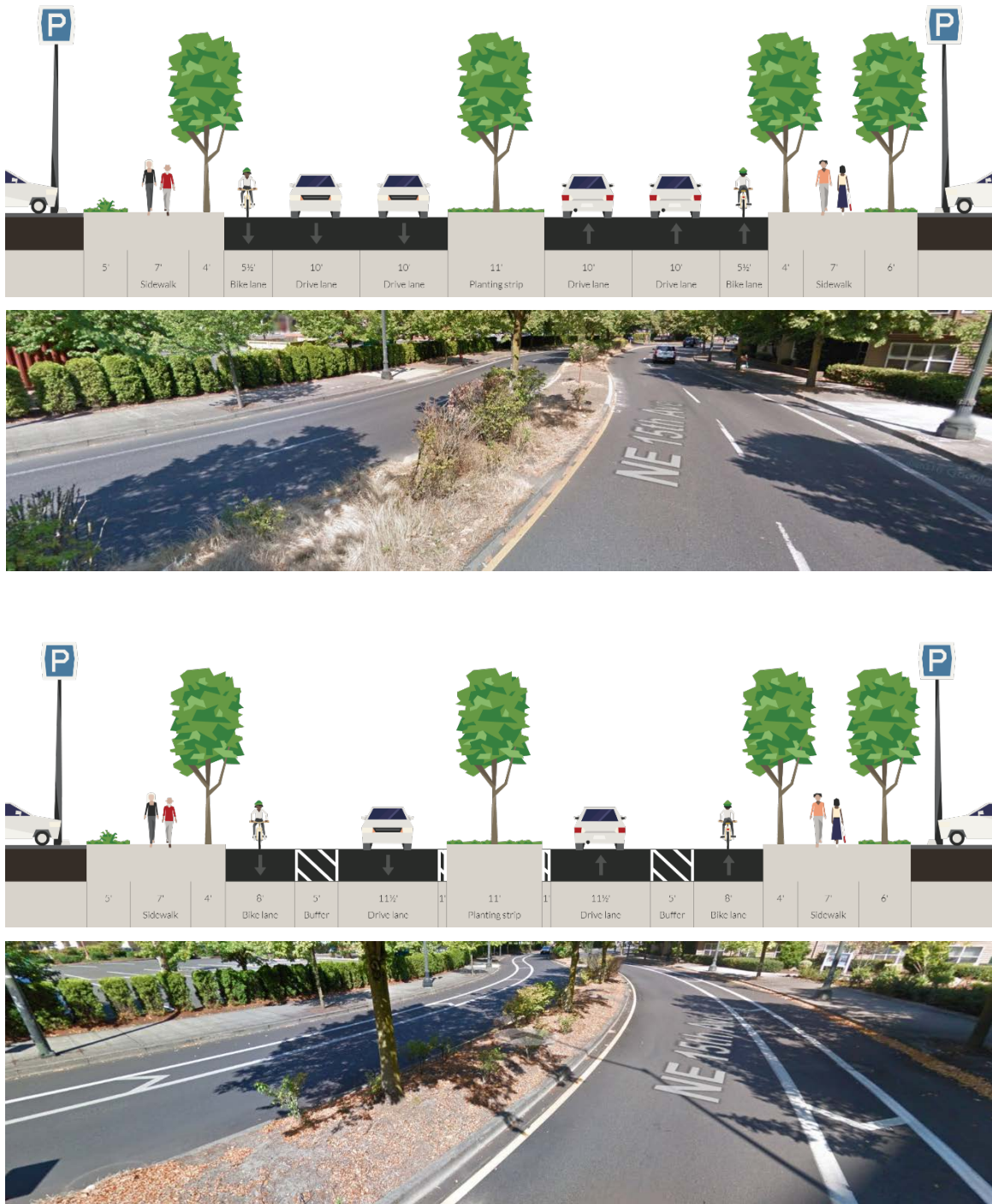


Figure 5-4 – Streetmix and Google Streetview of 16th Street cross section before (top) and after (bottom).

6.0 RESULTS

The methodology outlined in this thesis was applied successfully to three study areas, and the results of travel time and speed changes are presented for each location with commentary about possible causes and effects. However, these results should not be used as a motivation or deterrent for future road diets or transit route changes, since each change must be considered on a case-by-case basis. Instead, these results provide evidence for when and where the methodology can be applied and what types of information it can provide.

6.1 CASE STUDY I: LOMBARD STREET

6.1.1 Travel Times

Figure 6-1 shows the change in travel times over the Lombard study area. The top two plots summarize all trips taken throughout the day and the bottom two plots are for the morning commute only (between 7 a.m. and 10 a.m.). The width of each line represents the 95th percentile confidence interval. Travel times did not see a statistically significant change following the implementation of the road diet for the majority of trips. Some of the slower trips (higher percentile travel times) saw a decrease in travel times of less than 30 seconds over the segment. The decrease in travel time is still small, but more pronounced in the morning commute for westbound travel. None of the trips in the evening commute saw a statistically significant decrease in travel times.

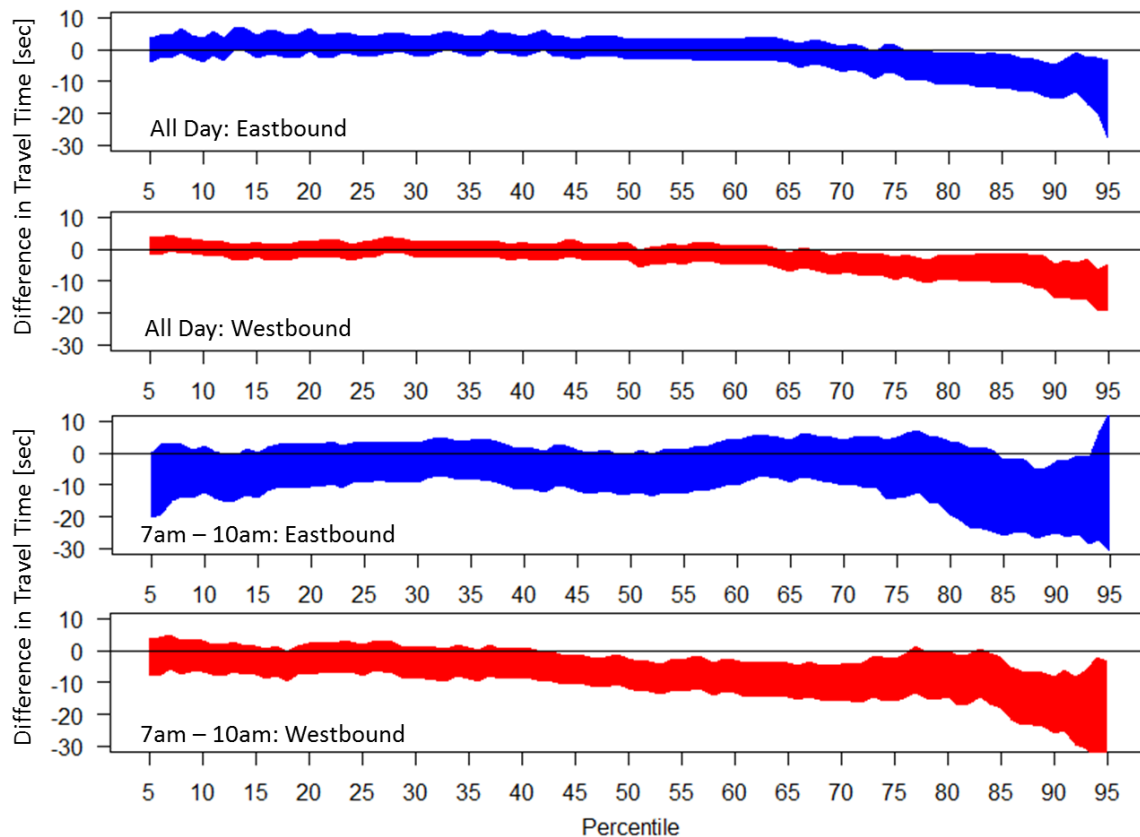


Figure 6-1 – Change in travel time by percentile on Lombard Street.

6.1.2 Travel Speeds

Figure 6-2 shows change in travel speeds by location (x-axis) and percentile (y-axis) for eastbound (top) and westbound (bottom) travel. The area of the road diet is marked. The change in speeds remained within five miles per hour of their original speed when analyzed for an entire day. This indicates little to no known operational change for transit or general traffic as a result of the lane narrowing.

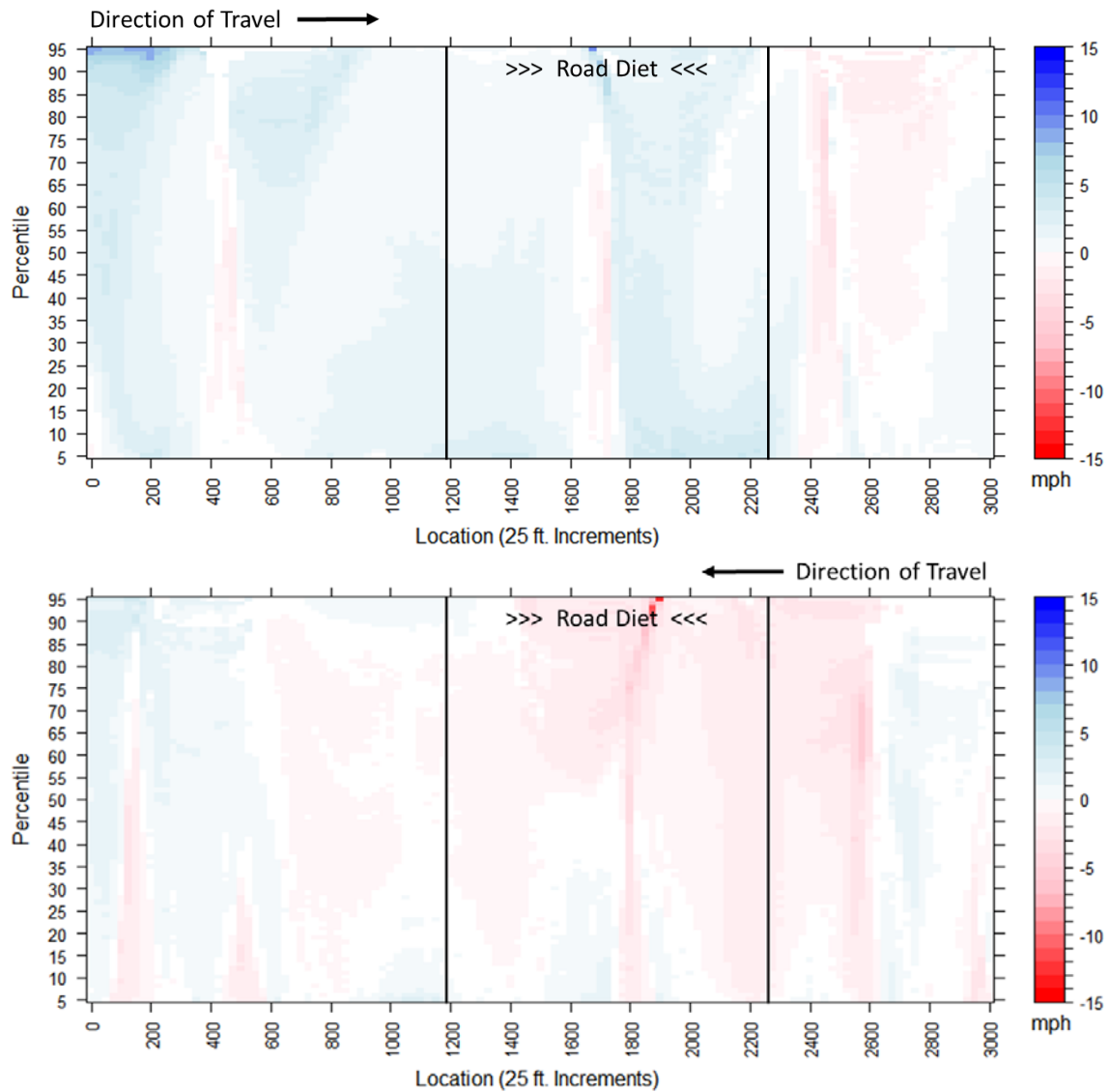


Figure 6-2 – Difference in travel speeds between before-and-after cases along Lombard Street

6.2 CASE STUDY II: NE 16TH STREET

6.2.1 Travel Times

Figure 6-3 shows travel times by percentile over the study area of NE 16th Street for both directions of travel. The width represents the 95th percentile confidence interval.

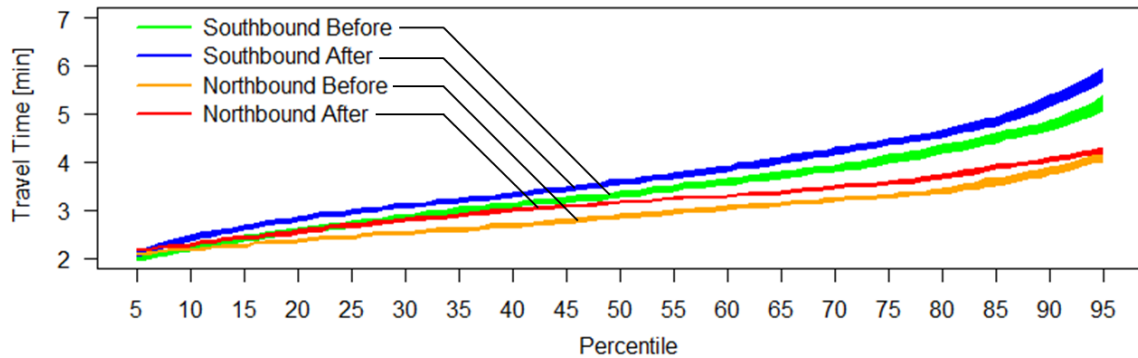


Figure 6-3 – Travel time by percentile for 16th Street.

After subtracting the before-and-after cases (Figure 6-4), a statistically significant increase in travel times is observed for both directions of travel; but, for all percentiles, the increase was less than one minute. Assuming that all travel increases were attributed to the 1,100-foot road diet, the travel time increase would remain below five minutes per mile in the worst case and less than one minute per mile on average.

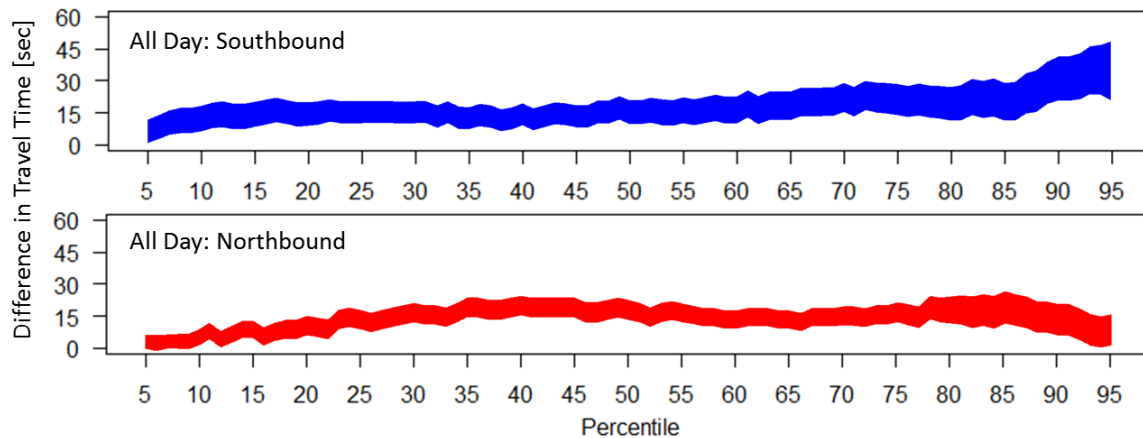


Figure 6-4 – Difference in travel time by percentile for 16th Street.

6.2.2 Travel Speeds

The changes in speed observed (Figure 6-5) are within 5 miles per hour of the original speed for most locations and percentiles. The speed decrease above this range in northbound direction at $x = 400$ feet is the result of a bus stop being added to the route following the completion of the road diet.

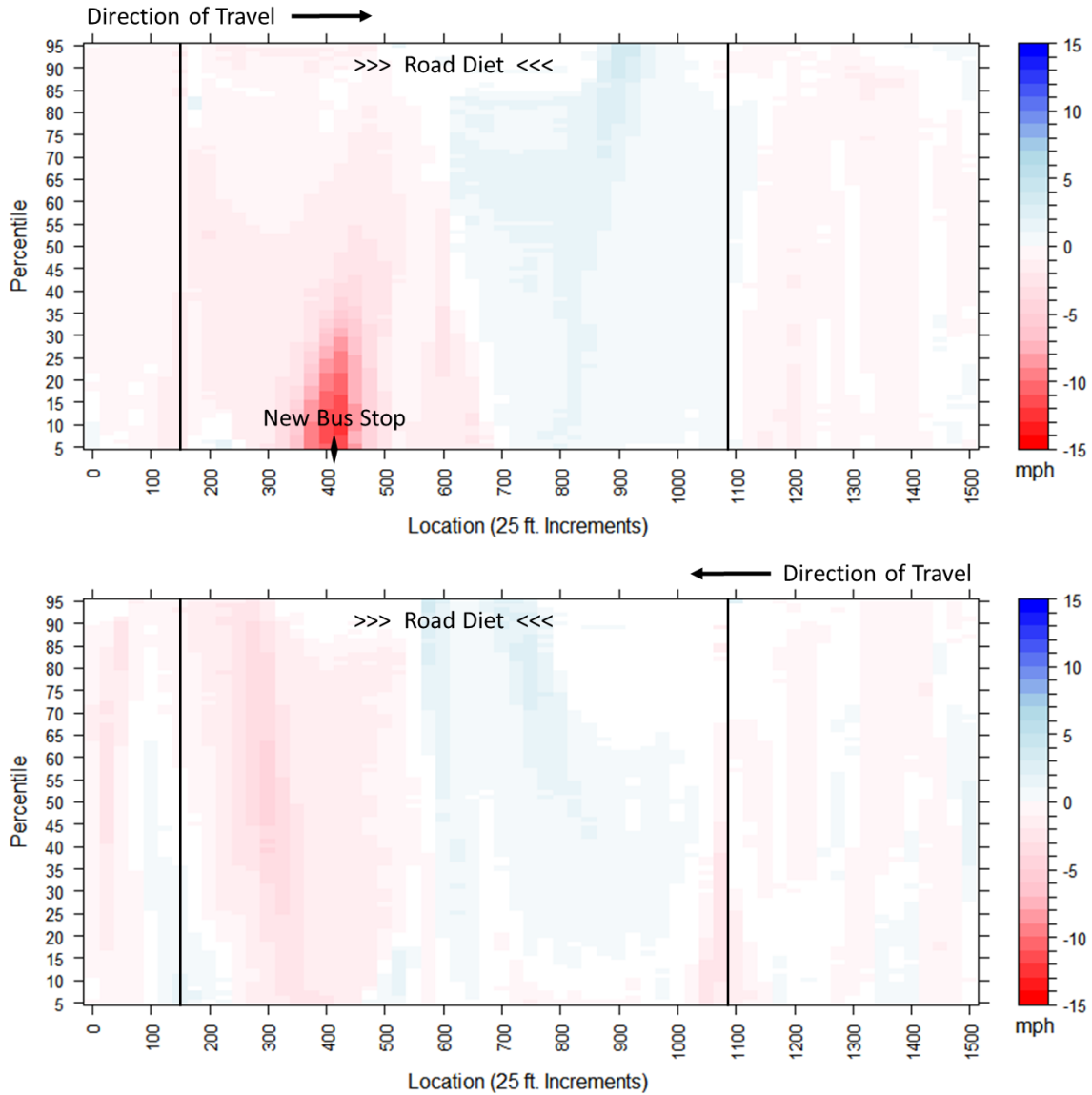


Figure 6-5 – Difference in travel speeds between before-and-after cases along 16th Street.

The added bus stop is unlikely to impact regular traffic because when buses stop to serve passengers they pull to the right, using the bike and buffer lane as a loading zone. This has been observed to leave enough space for other vehicles to pass. Additionally, the bus stops located at $x = 1,350$ feet in both direction did not see any marked change in performance, nor did the left turn for northbound travel and the right turn for southbound travel at $x = 100$ feet.

The minor differences in speeds can be seen in Figure 6-6, which shows speed variability (equation 3) and speed variability index (equation 4). The top two plots show the before-and-after conditions for these metrics separately, while the bottom two plots show the difference in speed variability (lower middle) and difference in speed variability index (bottom). The width of each line corresponds to the 95th percentile confidence interval.

The difference in the before-and-after cases of both speed variability and speed variability index remained near zero. The exception occurs at $x = 400$ feet, where the new stop was added. Overall, these results indicate little total impact on the operational speeds of transit caused by a lane reduction.

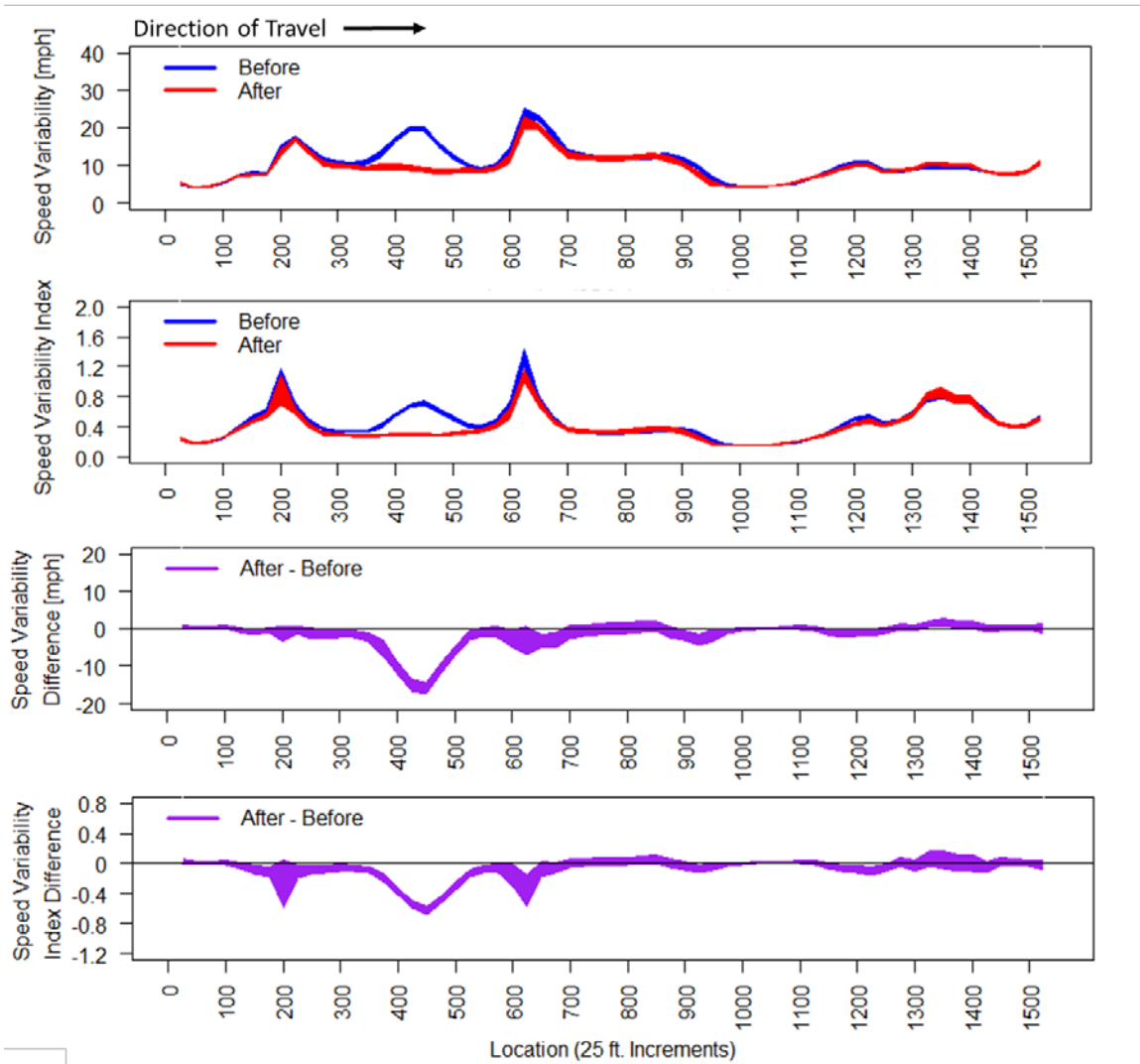


Figure 6-6 – Northbound speed variability and speed variability index on 16th Street.

7.0 DISCUSSION

The traditional process of identifying corridors for road diet improvements involves selecting potential corridors (mostly based on identifying four-lane roads), and conducting a traffic impact analysis of proposed changes on a selected roadway before implementing changes. The evaluation of roadway reallocation projects should include the analysis of traffic volumes, level of service (LOS), speeds, queue lengths and bus operations (Knapp et al., 2014). There are tools and equipment to evaluate effectively traffic volumes and LOS changes in before-and-after studies. However, the detailed evaluation of speed and queue length distributions along a segment are significantly more cumbersome. In addition, the exhaustive evaluation of bus operations requires detailed data and specific tools, and cannot be readily accomplished with field measurements or conventional LOS studies.

This research presented a general methodology for the detailed evaluation of transit operations and speed and queue length distributions along roadway reallocation projects. The proposed strategy and methodology is based on the utilization of high-resolution transit datasets. The focus of this research is on the development of a practical, general and theoretically sound methodology that can be applied to future roadway reallocation projects and applicable in a wide range of traffic conditions and locations.

This research provides a strategy and formulas to quantify changes in transit speeds and travel times and use confidence intervals (without the need to assume a distribution) to determine if that change was significant. The integration of transit high-resolution, stop event and stop disturbance data provides more information than any one dataset can provide, which improves resolution of the results. Finally, the methodology is applicable across a range of locations, traffic volumes, roadway types, and roadway modification projects; as such, it can be applied broadly to any segment or network. The methodology proposed in this research was applied successfully to two separate road diets.

8.0 REFERENCES

- Albright, E. & Figliozzi, M., 2013. Factors Influencing Effectiveness of Transit Signal Priority and Late Bus Recovery at Signalized-Intersection Level. *Transportation Research Record: Journal of the Transportation Research Board*, Volume 2311.
- Anderson, G., Searfoss, L., 2015, Safer Streets, Stronger Economies Complete Streets project outcomes from across the country, Smart Growth America.
- Berkow, M., Wolfe, M., Monsere, C. & Bertini, R., 2008. Using Signal System Data and Buses as Probe Vehicles to Define the Congested Regime on Arterials. *Transportation Research Record: Journal of the Transportation Research Board*, pp. 35-45.
- Bertini, R. & El-Geneidy, A. M., 2003. Generating Transit Performance Measures with Archived Data. *Transportation Research Record: Journal of the Transportation Research Board*, Volume 1841, pp. 109-119.
- Bertini, R. & Tantiyanugulchai, S., 2004. Transit Buses as Traffic Probes: Use of Geolocation Data for Empirical Evaluation. *Transportation Research Record: Journal of the Transportation Research Board*, Volume 1870, pp. 35-45.
- Brown, B. B. & Werner, C. M., 2008. Before and After a New Light Rail Stop: Resident Attitudes, Travel Behavior, and Obesity. *Journal of the American Planning Association*, 75(1), pp. 5-12.
- Brown, M. B. & Wolfe, R. A., 1983. Estimation of the Variance of Percentile Estimates. *Computational Statistics & Data analysis*, Volume 1, pp. 167-174.
- Cathey, F. & Dailey, D., 2002. Transit Vehicles as Traffic Probe Sensors. *Transportation Research Record: Journal of the Transportation Research Board*, Volume 1804, pp. 23-30.
- Chakroborty, P. & Kikuchi, S., 2004. Using Bus Travel Time Data to Estimate Travel Times on Urban Corridors. *Transportation Research Record: Journal of the Transportation Research Board*, Volume 1870, pp. 18-25.
- Dell'Olio, L., Ibeas, A. & Cecín, P., 2010. Modelling User Perception of Bus Transit Quality. *Transport Policy*, 17(6), pp. 388-397.
- Farmer, C. M., Lund, A. K., Trepel, R. E. & Braver, E. R., 1997. Fatal Crashes of Passenger Vehicles Before and After Adding Antilock Braking Systems. *Accident Analysis & Prevention*, 29(6).
- Federal Highway Administration, 2004. Summary Report: Evaluation of Lane Reduction "Road Diet" Measures and Their Effects on Crashes and Injuries, s.l.: Department of Transportation.
- Federal Highway Administration, 2016. Road Diets (Roadway Reconfiguration). [Online] Available at: https://safety.fhwa.dot.gov/road_diets/
- Feng, W., Figliozzi, M. & Bertini, R., 2015. Quantifying the joint impacts of stop locations, signalized intersections, and traffic conditions on bus travel time. *Public Transport*, 7(3), pp. 391-408.
- Glick, T. B., Feng, W., Bertini, R. & Figliozzi, M., 2014. Exploring Applications of Second Generation Archived Transit Data for Estimating Performance Measures and Arterial Travel Speeds. *Transportation Research Record: Journal of the Transportation Research Board*, Volume 2538.

- Glick, T. B. & Figliozzi, M., 2017a. Measuring the Determinants of Bus Dwell Time: New Insights and Potential Data Biases. Forthcoming Transportation Research Record: Journal of the Transportation Research Board.
- Glick, T. B. & Figliozzi, M., 2017b. Traffic and Transit Travel Time Reliability Indexes and Confidence Intervals. Forthcoming Transportation Research Record: Journal of the Transportation Research Board.
- Gudz, E., Fang, K. & Handy, S., 2017. When a Diet Prompts a Gain: Impact of Road Diet on Bicycling in Davis, California. Forthcoming Transportation Research Record: Journal of the Transportation Research Board.
- Hall, R. & Vyas, N., 2000. Buses as a Traffic Probe: Demonstration Project. Transportation Research Record: Journal of the Transportation Research Board, 1731(1), pp. 96-103.
- Huang, H. F. & Cynecki, M. J., 2000. Effects of Traffic Calming Measures on Pedestrian and Motorist Behavior. Transportation Research Record: Journal of the Transportation Research Board, Volume 1705.
- Kimpel, T., Strathman, J., Bertini, R. & Callas, S., 2005. Analysis of Transit Signal Priority using Archived TriMet Bus Dispatch System Data. Transportation Research Record: Journal of the Transportation Research Board, Volume 1925, pp. 156-166.
- Kittelson & Associates, et. al. , 2013. Transit Capacity and Quality of Service Manual, Third Edition. Washington D.C.: The National Academies Press.
- Knapp, K., Chandler, B., Atkinson, J., Welch, T., Rigdon, H., Retting, R., Meekins S, Widstrand E, & Porter, R. J. (2014). Road Diet Informational Guide (No. FHWA-SA-14-028).
- Levinson, H., 1983. Analyzing Transit Travel Time Performance. Transportation Research Record: Journal of the Transportation Research Board, Volume 915, pp. 1-6.
- Moore, A., Figliozzi, M. & Bigazzi, A., 2014. Modeling the Impact of Traffic Conditions on the Variability in Midblock PM_{2.5} Urban Arterial Concentrations. Transportation Research Record: Journal of the Transportation Research Board, Volume 2428, pp. 35-43.
- Moore, A., Figliozzi, M. & Monsere, C., 2012. Bus Stop Air Quality: An Empirical Analysis of Exposure to Particulate Matter at Bus Stop Shelters. Transportation Research Record: Journal of the Transportation Research Board , Volume 2270, pp. 78-86.
- PBOT, 2017. PBOT Traffic Counts. [Online]
Available at: pdx.maps.arcgis.com/
- Persuad, B. & Lyon, C., 2007. Empirical Bayes Before–After Safety Studies: Lessons Learned from Two Decades of Experience and Future Directions. Accident Analysis & Prevention, 39(3), pp. 546-555.
- Rodríguez, D. A. & Targa, F., 2004. Value of Accessibility to Bogotá's Bus Rapid Transit System. Transport Reviews, 24(5), pp. 587-610.
- Sidhu, B., Bertini, R. & Pande, A., 2017. Exploiting New High Resolution Transit Data Sources: Developing Improved Data Driven Models of Transit Travel Time and Delay. Forthcoming Transportation Research Record: Journal of the Transportation Research Board.
- Slavin, C., Feng, W., Figliozzi, M., & Koonce, P. (2013). Statistical study of the impact of adaptive traffic signal control on traffic and transit performance. Transportation Research Record: Journal of the Transportation Research Board, (2356), 117-126.
- Slavin, C., & Figliozzi, M. (2013). Impact of traffic signal timing on sidewalk-level particulate matter concentrations. Transportation Research Record: Journal of the Transportation Research Board, (2340), 29-37.

- Stoll, N., Glick, T. B. & Figliozzi, M., 2016. Utilizing High-Resolution Bus GPS Data to Visualize and Identify Congestion Hot Spots in Urban Arterials. *Transportation Research Record: Journal of the Transportation Research Board* 2539, 20-29.
- Taylor, J. R., 1997. *An Introduction to Error Analysis The Study of Uncertainties in Physical Measurements*. Sausalito, CA: University Science Books.
- Watkins, K. E. et al., 2011. Where Is My Bus? Impact of Mobile Real-Time Information on the Perceived and Actual Wait Time of Transit Riders. *Transportation Research Part A: Policy and Practice*, 45(8), pp. 839-848.
- Zhu, L., Holden, J. & Gonder, J., 2017. A Trajectory Segmentation Map-Matching Approach for Large-Scale, High-Resolution GPS Data. *Forthcoming Transportation Research Record: Journal of the Transportation Research Board*

9.0 APPENDIX A

LOMBARD STREET

The figures presented in this appendix include visuals of the roadway conditions before and after the completion of the case-study project as well as additional comparative results that did not show significant changes. Speed heat maps show travel speed profiles for a single direction of travel at one location. Each figure includes two plots where the upper heatmap shows before conditions while the lower shows after conditions. The x-axis of each indicates location and the y-axis will show either the 5th through 95th percentile or time of day. Speeds are displayed in miles per hour and direction of travel is shown. Using the 15th, 50th and 85th percentile travel speeds, other figures show the speed variability (top) and speed variability index (upper middle) for the before-and-after conditions as well as the difference in speed variability (lower middle) and difference in speed variability index (bottom). The width of each line is the 95th percentile confidence interval.

9.1 LOMBARD STREET HEATMAPS

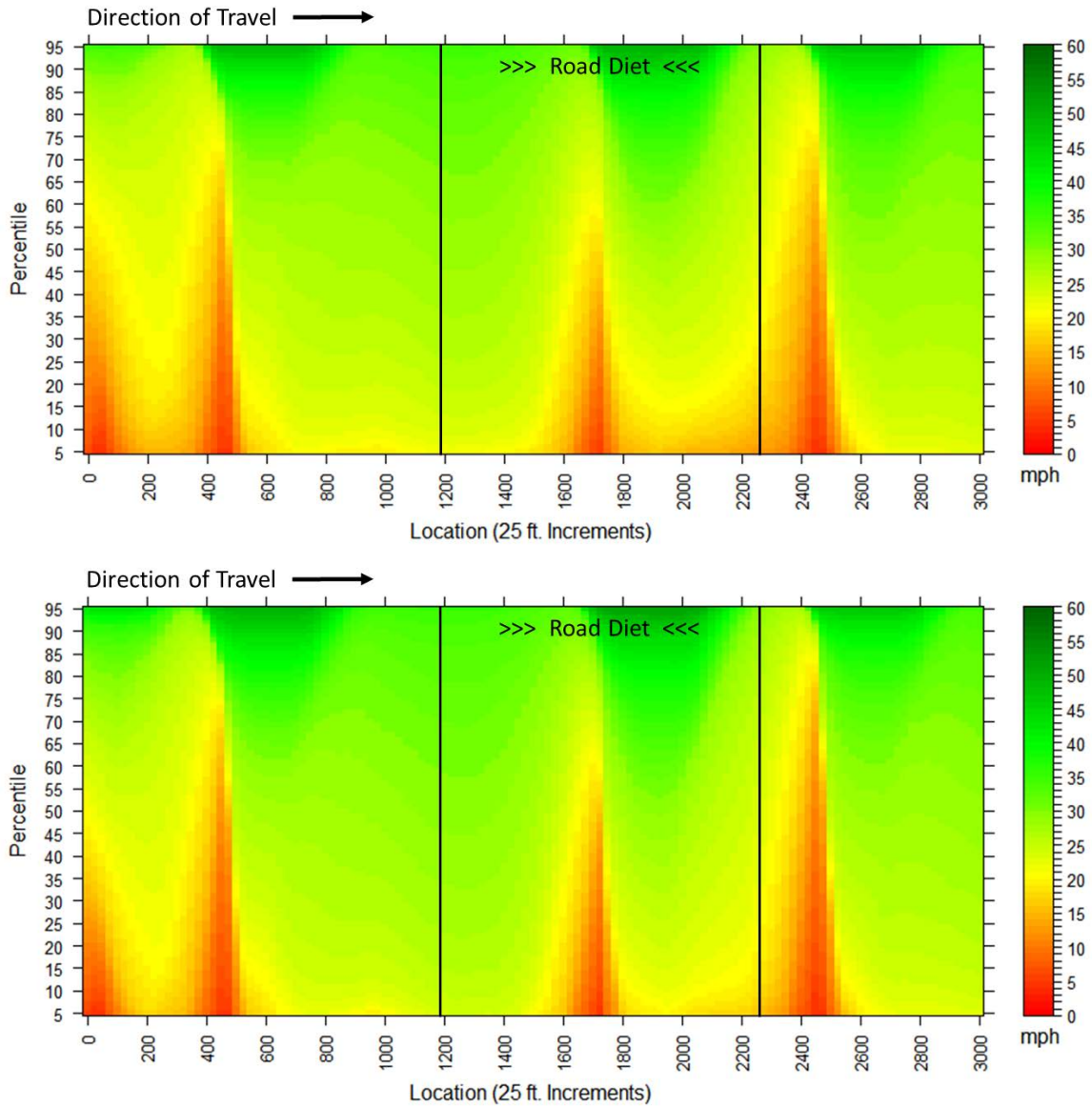


Figure 9-1 – Eastbound travel speed by percentile on Lombard Street. Top: before road diet. Bottom: after road diet. Used for Figure 6-2 (top).

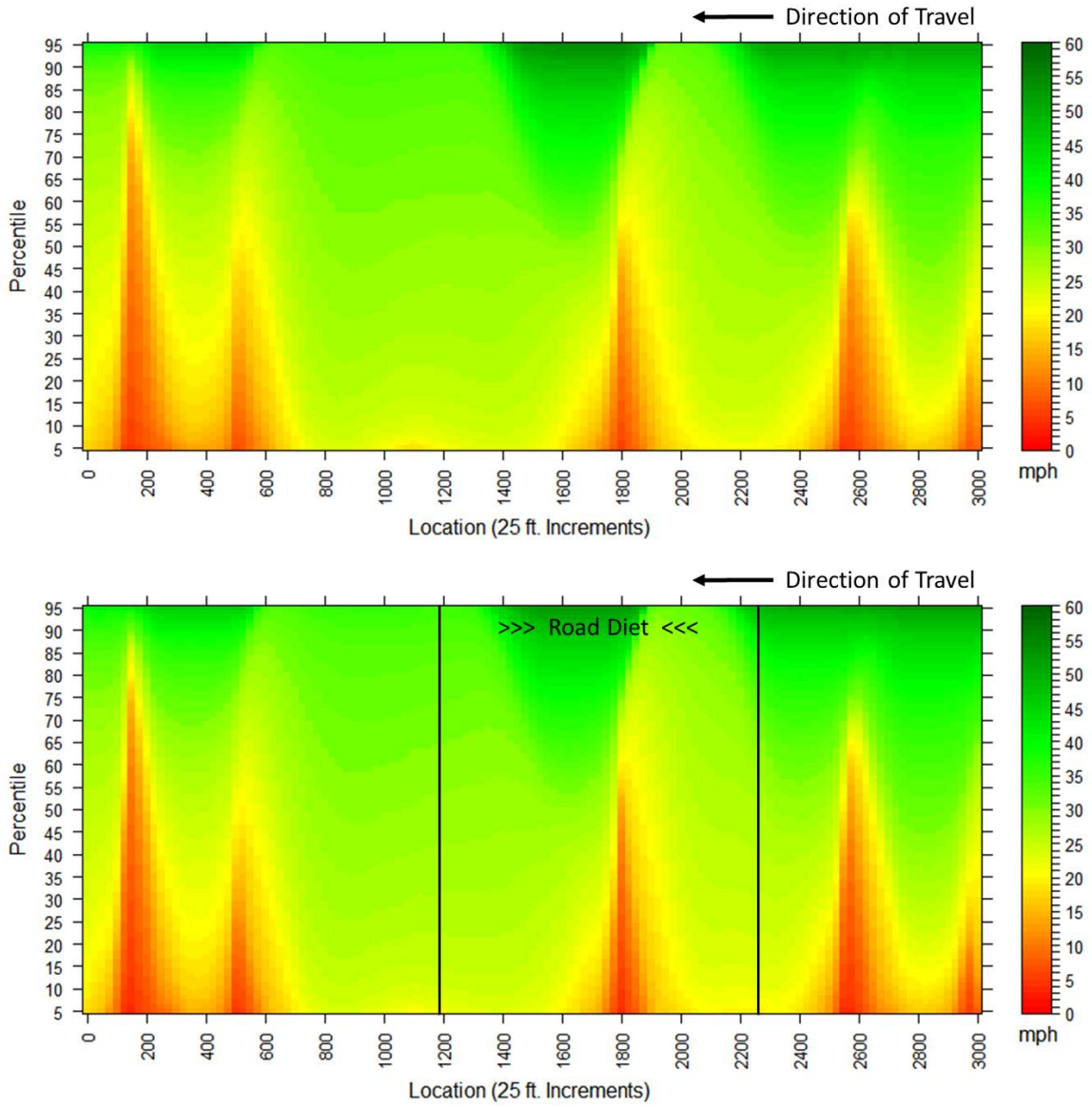


Figure 9-2 – Westbound travel speed by percentile on Lombard Street. Top: before road diet. Bottom: after road diet. Used for Figure 6-2 (bottom).

9.2 LOMBARD STREET INDICES

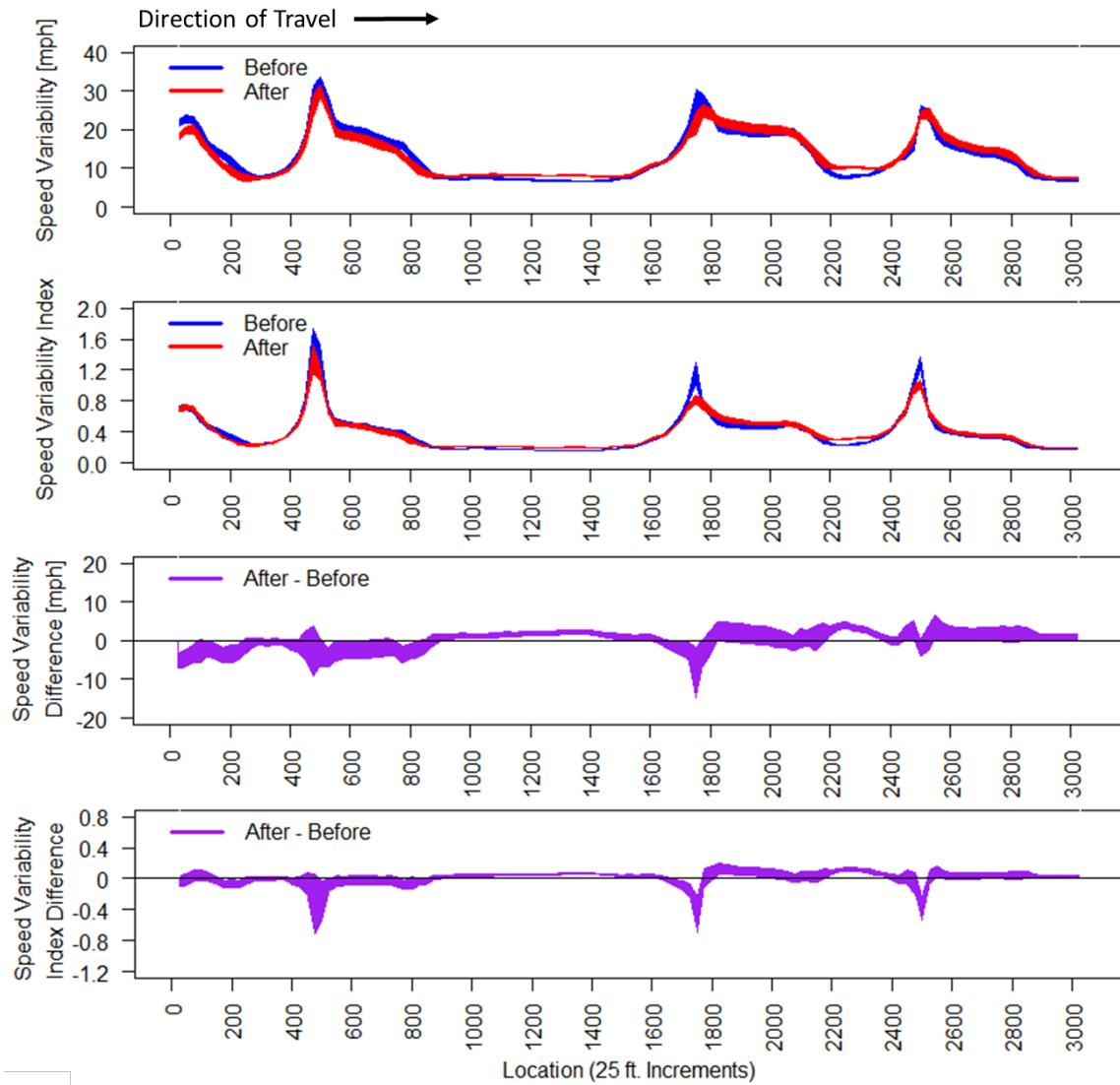


Figure 9-3 – Eastbound speed variability and speed variability index on Lombard Street.

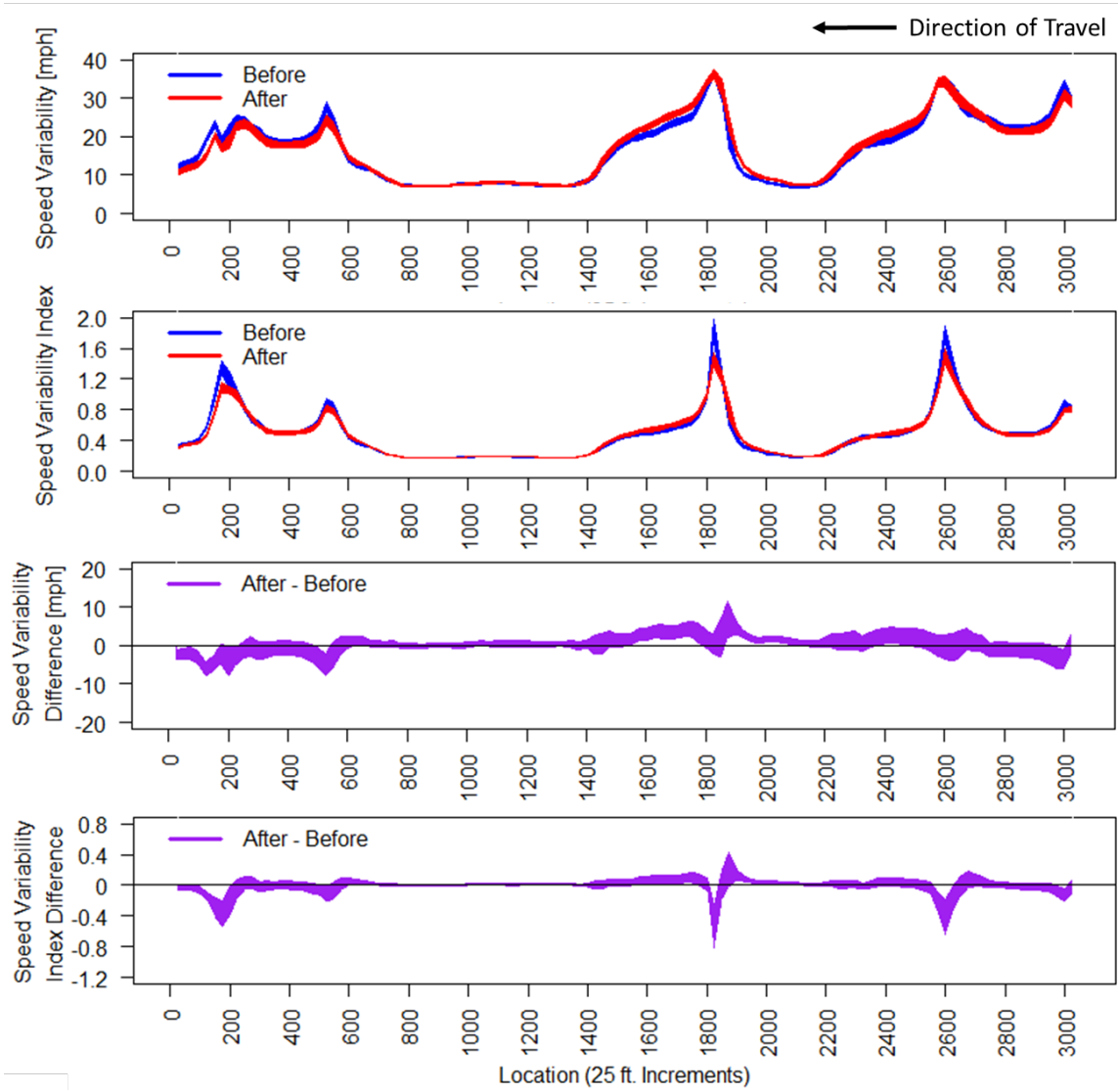


Figure 9-4 – Westbound speed variability and speed variability index on Lombard Street.

10.0 APPENDIX B

NE 16TH STREET

10.1 NE 16TH STREET HEATMAPS

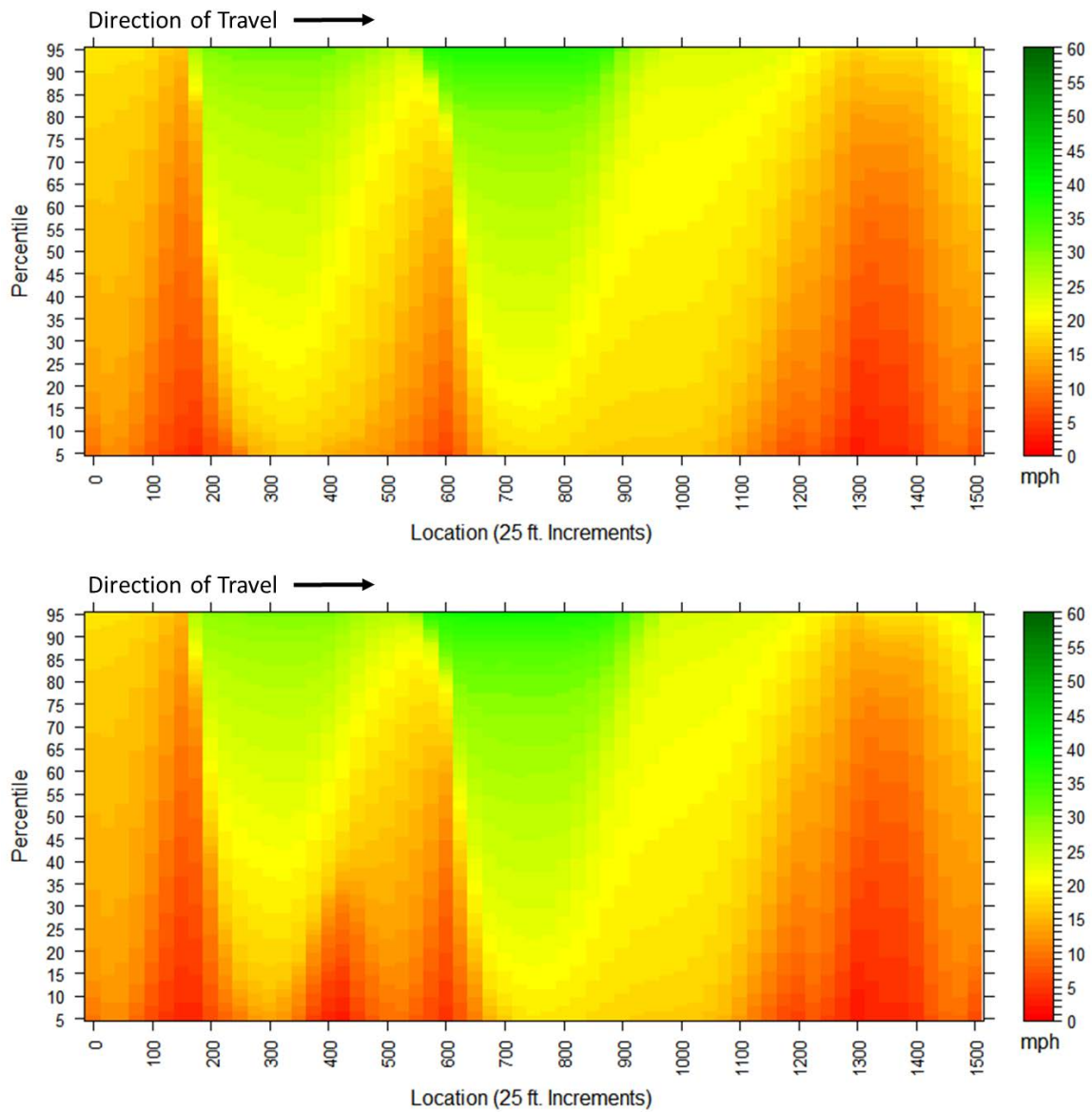


Figure 10-1 – Northbound travel speed by percentile on 16th Street. Top: before road diet. Bottom: after road diet. Used for Figure 6-5 (top).

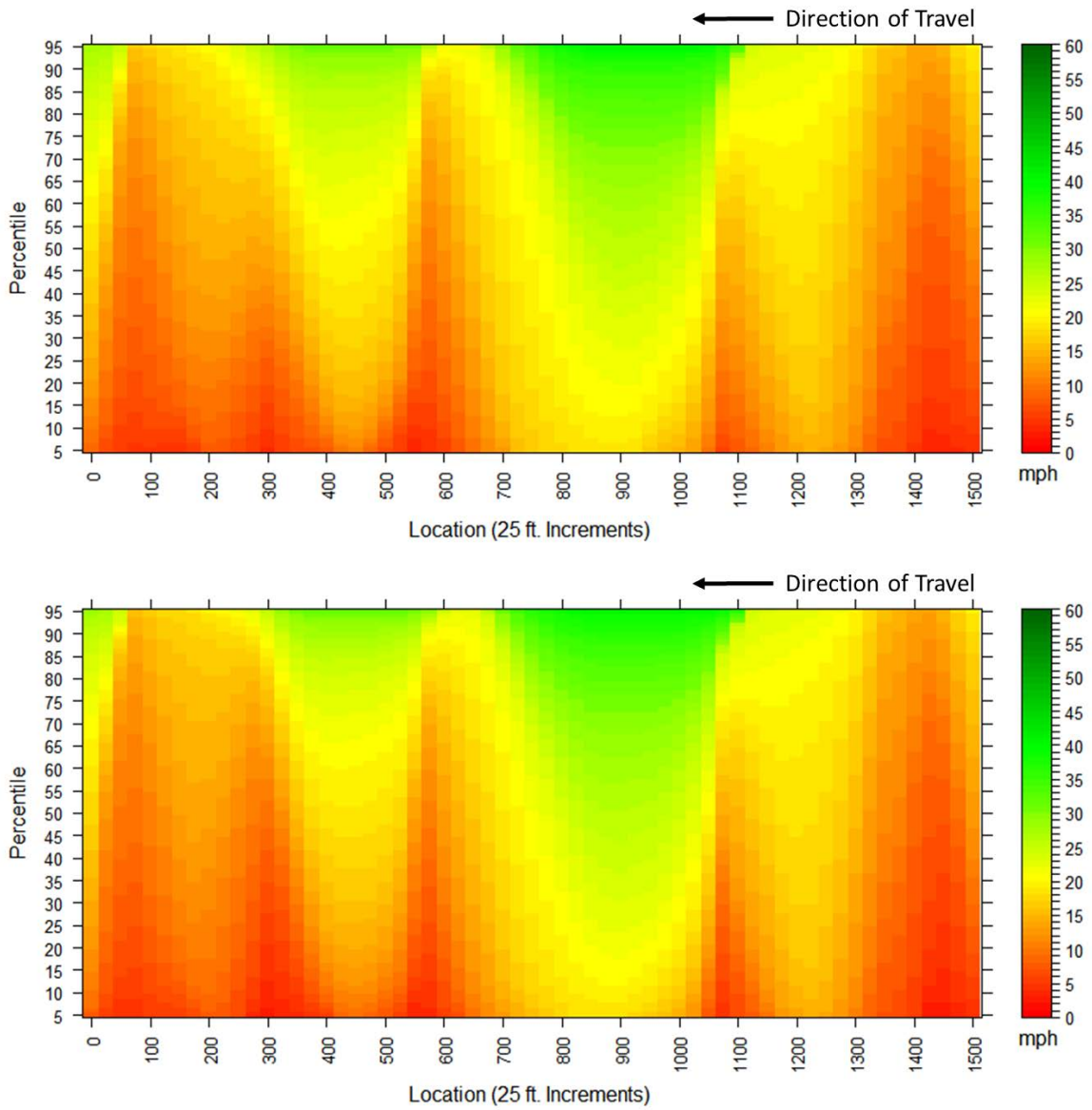


Figure 10-2 – Southbound travel speed by percentile on 16th Street. Top: before road diet. Bottom: after road diet. Used for Figure 6-5 (bottom).

10.2 NE 16TH STREET INDICES

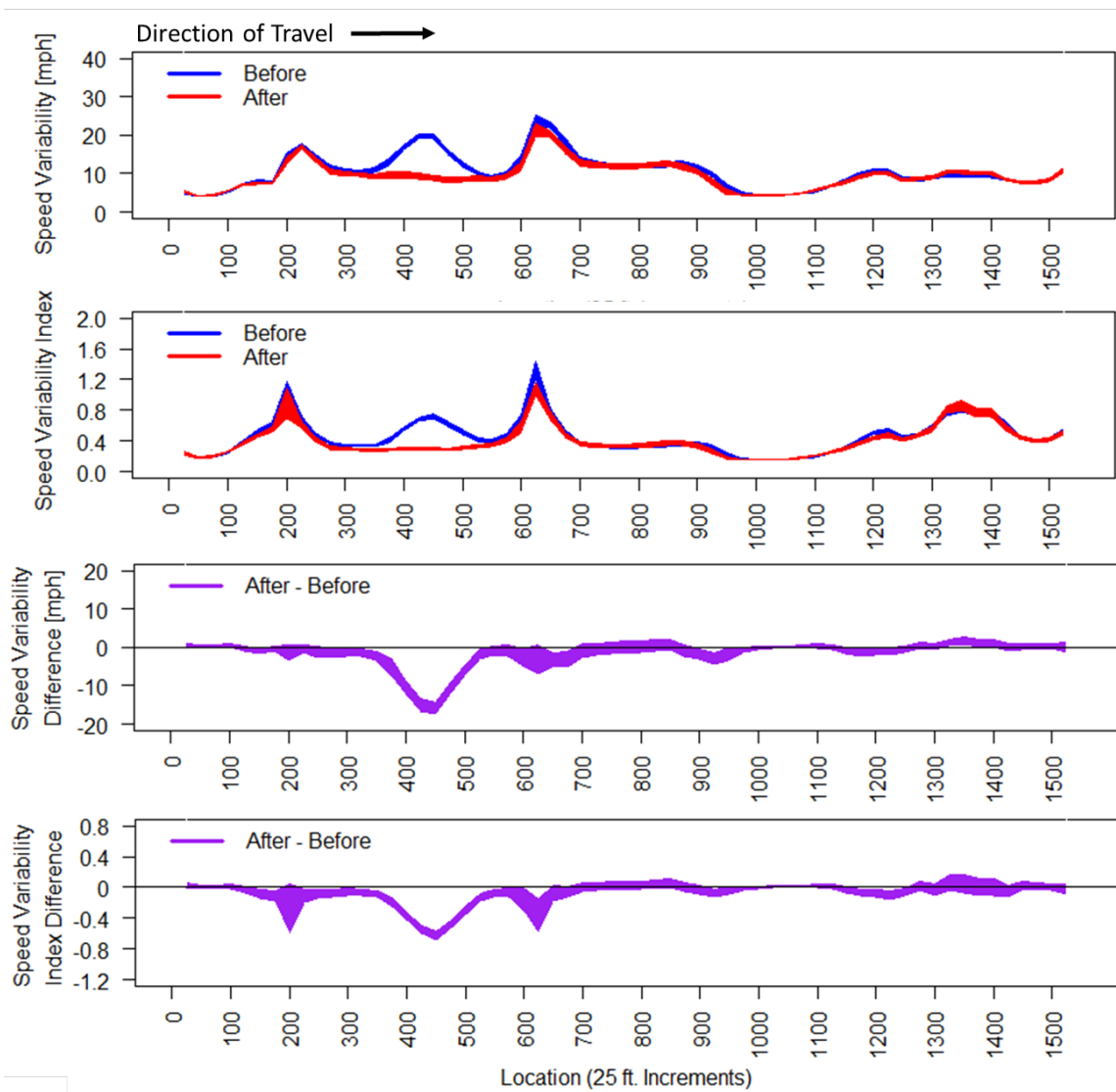


Figure 10-3 – Northbound speed variability and speed variability index on 16th Street.

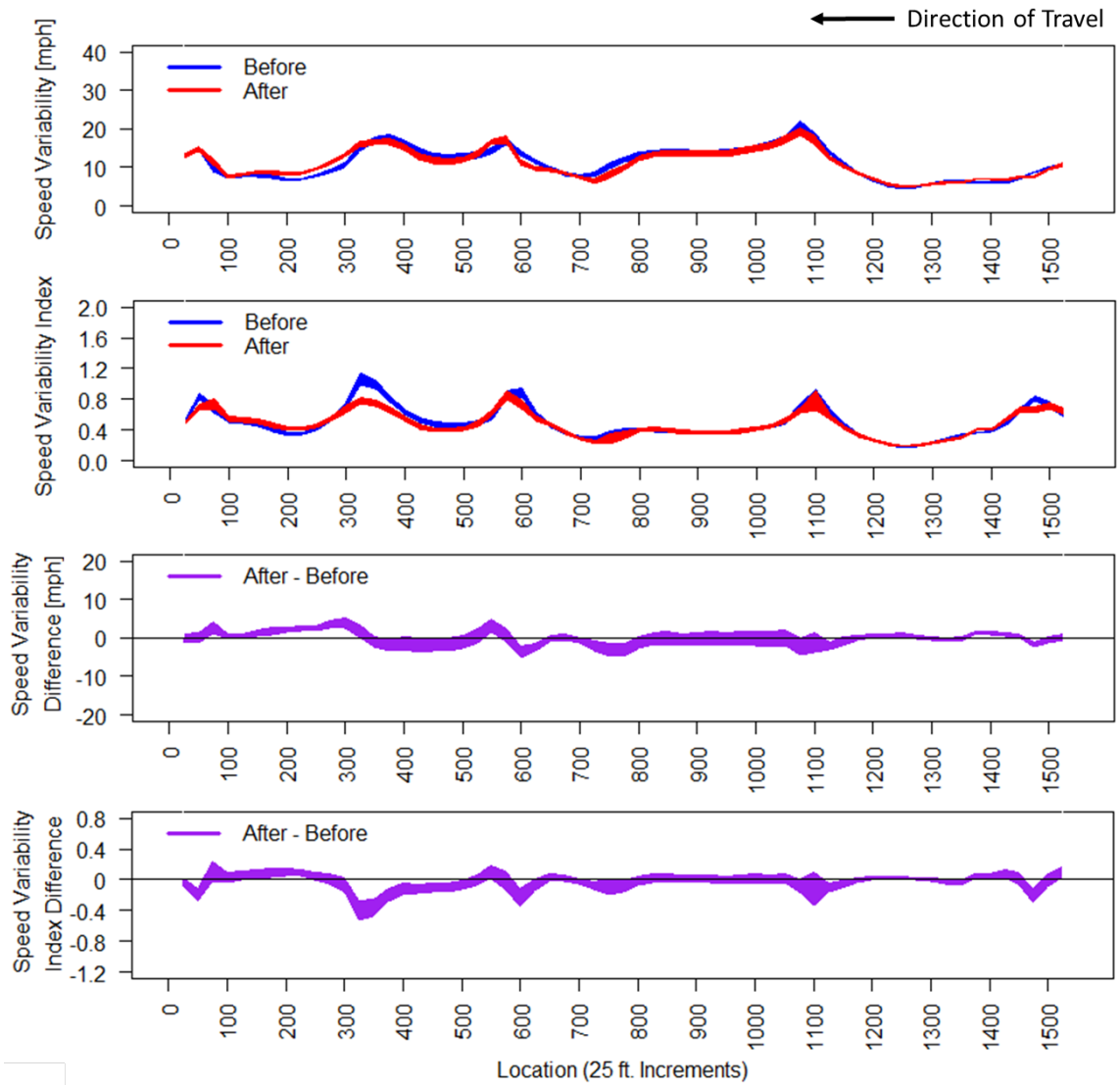
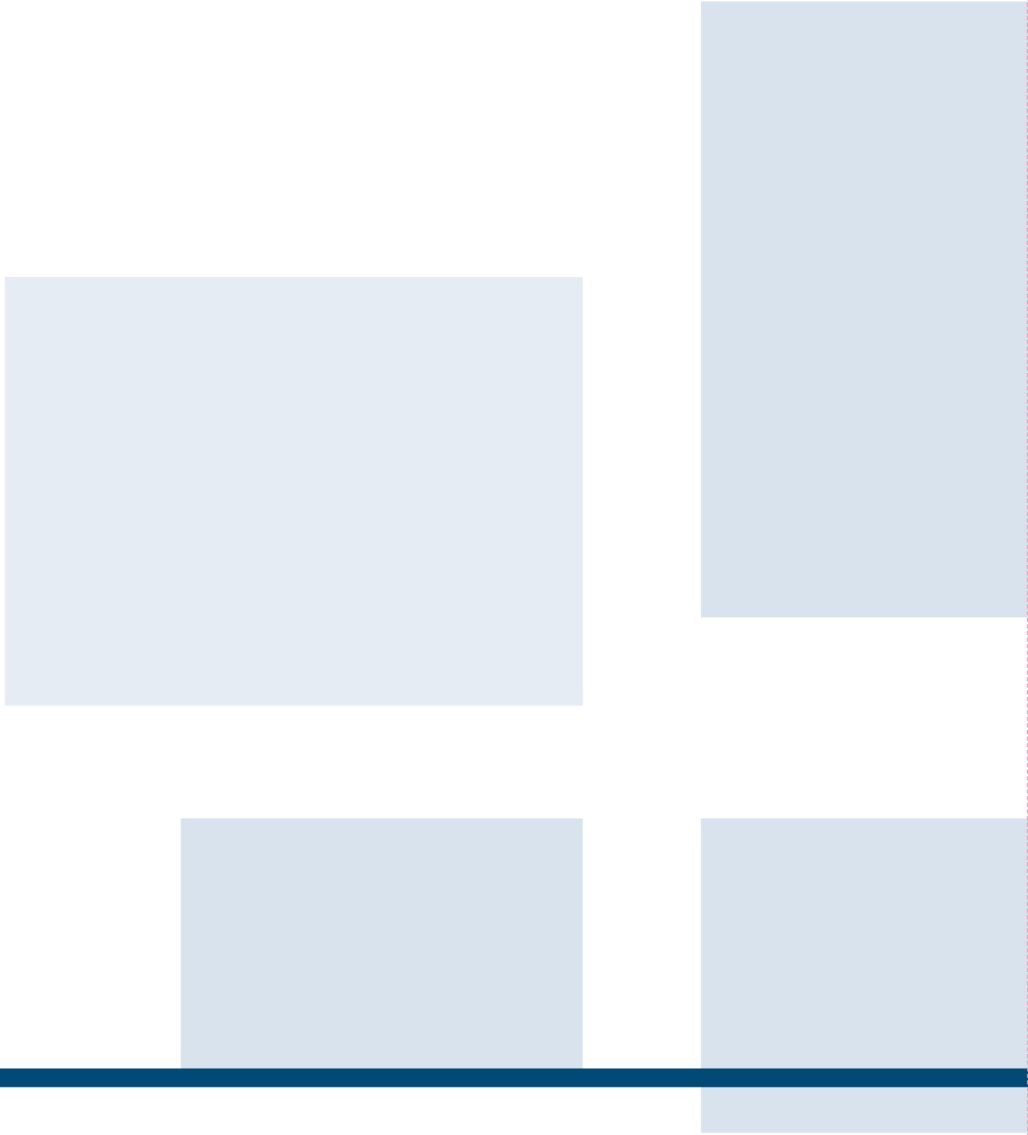


Figure 10-4 – Southbound speed variability and speed variability index on 16th Street.



Transportation Research and Education Center
Portland State University
1900 S.W. Fourth Ave., Suite 175
Portland, OR 97201