Utilizing High-Resolution Archived Transit Data to Study Before-and-After Travel-Speed and Travel-Time Conditions

Travis Bradley Glick
Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds

Part of the Civil and Environmental Engineering Commons, and the Transportation Commons

Let us know how access to this document benefits you.

Recommended Citation
Glick, Travis Bradley, "Utilizing High-Resolution Archived Transit Data to Study Before-and-After Travel-Speed and Travel-Time Conditions" (2017). Dissertations and Theses. Paper 4065.
https://doi.org/10.15760/etd.5949

This Thesis is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. For more information, please contact pdxscholar@pdx.edu.
Utilizing High-Resolution Archived Transit Data to Study
Before-and-After Travel-Speed and Travel-Time Conditions

by

Travis Bradley Glick

A thesis submitted in partial fulfillment of the
requirements for the degree of

Master of Science
in
Civil and Environmental Engineering

Thesis Committee:
Miguel Figliozzi, Chair
Christopher Monsere
Avinash Unnikrishnan

Portland State University
2017
ABSTRACT

Travel times, operating speeds, and service reliability influence costs and service attractiveness. This paper outlines an approach to quantify how these metrics change after a modification of roadway design or transit routes using archived transit data. The Tri-County Metropolitan Transportation District of Oregon (TriMet), Portland’s public transportation provider, archives automatic vehicle location (AVL) data for all buses as part of their bus dispatch system (BDS). This research combines three types of AVL data (stop event, stop disturbance, and high-resolution) to create a detailed account of transit behavior; this probe data gives insights into the behavior of transit as well as general traffic. The methodology also includes an updated approach for confidence intervals estimates that more accurately represent range of speed and travel time percentile estimates. This methodology is applied to three test cases using a month of AVL data collected before and after the implementation of each roadway change. The results of the test cases highlight the broad applicability for this approach to before-and-after studies.
DEDICATION

To my wife, friend, and main critic, Lucille Glick: Thank you for your encouragement, patience, and hours of proofreading.
ACKNOWLEDGEMENTS

Miguel Figliozzi, my advisor and committee member: your guidance and feedback have made this thesis possible.

Casey, Melanie and Steve Glick: thank you for your emotional and financial support through this process.

Nick Stoll, thank you for your assistance with R and coding.

TriMet staff members, Steve Callas and Miles Crumbly: thank you for graciously providing the data and definitions used in the analysis.

Chris Monsere and Avinash Unnikrishnan, professors and committee members: thank you for your comments and instruction.

Finally, a thank you to the members of Transportation Research and Innovation Lab (TRAIL). (It was a fun and friendly laboratory to work in these past years.)
# TABLE OF CONTENTS

Abstract ........................................................................................................................................... i

Dedication ........................................................................................................................................... ii

Acknowledgements ........................................................................................................................... iii

List of Tables ..................................................................................................................................... vi

List of Figures ................................................................................................................................. vii

1. Introduction and contribution ........................................................................................................ 1
   1.1. Structure of Thesis ................................................................................................................... 2

2. Literature Review ............................................................................................................................ 4
   2.1. Bus Speeds and Travel Times ................................................................................................. 4
       2.1.1. Buses as Roadway Probes ............................................................................................... 4
       2.1.2. Point Level to Route Level Analyses ............................................................................... 5
   2.2. Before-and-After Studies ......................................................................................................... 7
       2.2.1. Using Bayesian Statistics .................................................................................................. 8
       2.2.2. Descriptive Analysis ......................................................................................................... 9
   2.3. Road Diets ............................................................................................................................... 9

3. Data Sources .................................................................................................................................... 11
   3.1. Stop-to-Stop Data .................................................................................................................. 12
   3.2. Stop Disturbance Data ........................................................................................................... 13
   3.3. High-Resolution Data ............................................................................................................ 13

4. Methodology .................................................................................................................................... 15
   4.1. TriMet Data Integration ......................................................................................................... 15
   4.2. Spatial Data ............................................................................................................................ 16
       4.2.1. Spatial Data Cleaning ....................................................................................................... 17
       4.2.2. GIS Tool Definitions ........................................................................................................ 17
       4.2.3. Final Combined Data Set ................................................................................................ 18
   4.3. Data Analysis .......................................................................................................................... 19
   4.4. Unknown Distributions ........................................................................................................... 20
   4.5. Peak Hour Vs Whole Day Performance ................................................................................. 24
   4.6. Propagation of Uncertainty ..................................................................................................... 25
       4.6.1. Assumption of Normality .................................................................................................. 25
       4.6.2. Propagation of Uncertainty Addition and Subtraction ................................................... 26
       4.6.3. Propagation of Uncertainty Product and Quotient .......................................................... 27
       4.6.4. Confidence Intervals ....................................................................................................... 28
   4.7. Travel Time .............................................................................................................................. 29
   4.8. Space-Time-Speed Diagrams .................................................................................................. 30
   4.9. Before and After Comparisons ............................................................................................... 30
       4.9.1. Travel Speeds, Travel Times, and Peak Performance .................................................... 30
5. Study Areas .................................................................................................................. 32
   5.1. Case Study I: Lombard Street .............................................................................. 32
   5.2. Case Study II: 16th street .................................................................................. 35
   5.3. Case Study III: Powell Blvd. and Tilikum Crossing ........................................ 38

6. Results .......................................................................................................................... 40
   6.1. Case Study I: Lombard Street ............................................................................ 40
       6.1.1. Travel Times .......................................................................................... 40
       6.1.2. Travel Speeds ....................................................................................... 41
   6.2. Case Study II: 16th Street ................................................................................ 43
       6.2.1. Travel Times ......................................................................................... 43
       6.2.2. Travel Speeds ....................................................................................... 44
   6.3. Case Study III: Powell Boulevard ..................................................................... 46
       6.3.1. Westbound Travel Times (Bridge) .......................................................... 46
       6.3.2. Eastbound Travel Times (Bridge) ............................................................ 47
       6.3.3. Hourly Travel Times (Bridge) ................................................................ 49
       6.3.4. Costs ....................................................................................................... 49
       6.3.5. Eastbound and Westbound Travel Times (Half Route) .......................... 50
       6.3.6. Travel Speeds ........................................................................................ 51

7. Discussion .................................................................................................................... 54
   7.1. Road Diets ......................................................................................................... 54
   7.2. Powell Blvd. and the Tilikum Crossing ............................................................... 54
   7.3. Recommendations ............................................................................................ 55

8. Conclusion ...................................................................................................................... 56

References ......................................................................................................................... 57

Appendix A: Supplemental Figures ................................................................................ 61
   A.1. Lombard Street ................................................................................................. 62
   A.2. 16th Street ....................................................................................................... 66
   A.3. Powell Blvd....................................................................................................... 70

Appendix B: Code for Standard Deviation Example ..................................................... 77

Appendix C: Travel Time Confidence Estimates ............................................................. 80
LIST OF TABLES

Table 1 – GIS tools definitions with visuals................................................................. 18
Table 2 – Route 9 hourly travel times over river on Powell Blvd. ............................. 49
Table 3 – Average daily TriMet operational costs before and after the roadway change. 50
Table 4 – Route 9 hourly travel times over 7.2 miles on Powell Blvd. ....................... 51
LIST OF FIGURES

Figure 1 – Bus trajectories using combined SE and SD data versus HR data. .................. 14
Figure 2 – Bus Trajectory from Figure 1 using the combined data set.......................... 16
Figure 3 – GIS flowchart for processing non-overlapping route data. ............................ 17
Figure 4 – Hypothetical overlapping transit routes...................................................... 19
Figure 5 – Point distribution density for real data (top) and normal approximation (bottom) ......................................................................................................................... 20
Figure 6 – Point distribution density for real data compared with normal approximation 21
Figure 7 – Speed histogram at $x = 2000$ ft. (left) and $x = 200$ ft. (right) from Figure 6. 21
Figure 8 – Speed histogram at $x = 125$ ft. (left) and $x = 1925$ ft. (right) from Figure 6. 22
Figure 9 – Percentile estimation process using random non-normal data ....................... 23
Figure 10 – Speed histogram at $x = 700$ ft. from Figure 6 and distributions of $15^{th}$, $50^{th}$, and $85^{th}$ percentile and speed variability and speed variability index. ................. 26
Figure 11 – Map and aerial views of study area for Lombard Street. ............................. 33
Figure 12 – Streetmix and Google Streetview of Lombard Street cross-section before (top) and after (bottom) road diet................................................................. 34
Figure 13 – Map and aerial views of study area for $16^{th}$ Street. ................................ 36
Figure 14 – Streetmix and Google Streetview of $16^{th}$ Street cross-section before (top) and after (bottom) road diet ................................................................. 37
Figure 15 – Map of Study Area for Powell Blvd......................................................... 38
Figure 16 – Map of route 9 before and after change.................................................... 39
Figure 17 – Change in travel time by percentile on Lombard Street............................ 41
Figure 18 – Difference in travel speeds between before and after cases along Lombard Street ......................................................................................................................... 42
Figure 19 – Travel time by percentile for $16^{th}$ Street................................................ 43
Figure 20 – Difference in travel time by percentile for $16^{th}$ Street............................. 43
Figure 21 – Difference in travel speeds between before and after cases along $16^{th}$ Street 44
Figure 22 – Northbound speed variability and speed variability index on 16th Street...... 46

Figure 23 – Travel times and travel time differences over river in westbound direction for all buses................................................................................................................. 47

Figure 24 – Travel times and travel time differences over river in westbound direction for PM-peak travel................................................................................................................. 47

Figure 25 – Travel times and travel time differences over river in eastbound direction for all buses...................................................................................................................................... 48

Figure 26 – Travel times and travel time differences over river in westbound direction for PM-peak travel................................................................................................................. 48

Figure 27 – Travel times over 7.2 miles of Powell Blvd for all trips. Top: westbound.
Bottom: eastbound ...................................................................................................................... 50

Figure 28 – Difference in eastbound travel speeds along Powell Blvd. ...................... 52

Figure 29 – Difference in westbound travel speeds along Powell Blvd. ...................... 53
1. INTRODUCTION AND CONTRIBUTION

Transit performance metrics influence the transportation planning process and future transportation decisions, which in turn affect operating speeds, travel times, ridership, costs, and efficiency (Levinson, 1983). Research on transit operations and performance metrics evolve as agencies integrate and change their data collection systems. The installation of global position systems (GPS) technology onboard buses has opened new avenues of study and provided a means to quantify transit behaviors hidden by the structure of previous data collection system. While high-resolution (HR) GPS data collection expands analysis options, these data collection systems are new and not yet widespread across transit agencies. As such, many applications remain untested and understudied.

Archived transit data is widespread across transit agencies and its 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 is used, SE data still dominates the discourse and practice. While newer high-resolution systems exist, they are not widespread and few practice-ready methodologies exist.

Using high-resolution data integrated with more ubiquitous stop event (SE) and stop disturbance (SD) data (discussed below), a methodology is presented to examine speed and travel time changes on a roadway before and after a modification. Three test cases are then used to show the range of applicability of this methodology and to highlight how it may be applied in the future.
This thesis expands on the existing systems of transit analysis to create applicable methodologies using higher-resolution data that quantify transit performance changes following a roadway modification. It then applies it to three case studies including road diets, where transit has been traditionally excluded. Three archived data sets (SE, SD, and HR) are merged to provide a more detailed account of segment activities than has been used in previous studies with the addition of confidence intervals to give a more accurate account of transit performance.

The results of the application of this methodology, as applied to the case studies, are not intended to definitively answer general questions about road diets or transit bridge effectiveness. Rather, they are intended to highlight how the methodology is applicable in a wide range of ADTs, locations, and project and has the potential to be expanded to more in-depth and broader analyses, such as road diets at the upper limits of recommended AADT.

1.1. Structure of Thesis

This thesis can be divided into three broad parts: (1) background and literature review, (2) methodology, and (3) case studies.

(1) First, Section 2 examines the existing literature on travel speeds and travel times as a means to quantify roadway changes. Section 3 discusses the three types of data sources provided by TriMet that are used in this analysis.

(2) Next, Section 4 provides an in-depth look into this thesis’ methodology, dataset cleaning and processing, statistical distributions and uncertainty propagation, and definitions for summary statistics and diagrams.
Finally, the thesis presents three before-and-after case studies for changes in Portland. Section 5 summarizes the three study areas, Lombard Street, 16th Street, and Powell Blvd and the new Tilikum Crossing. Section 6 presents the results of the quantitative analysis methods introduced in Section 4. Section 7 concludes with a qualitative discussion of the changes in the three study areas.
2. LITERATURE REVIEW

This literature review presents the background material for this thesis’ examination of travel speeds and travel times as a means to quantify roadway changes. It includes previous approaches to performance metric analysis of transit routes, types of before-and-after studies, and an introduction to road diets.

2.1. 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 (Kittelson & Associates, et al., 2013). For years, research efforts have focused on quantifying the impact of these factors.

2.1.1. Buses as Roadway Probes

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 the
 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, et al., 2012) and the impact of transit signal priority (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, 2011), and sidewalks at mid-block locations (Moore, et al., 2014). Additionally, Feng, et al. (2014) (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.

### 2.1.2. Point Level to Route Level Analyses

Recently, Zhu et al. (Forthcoming 2017) began using the HR data provided by TriMet. High-resolution data can come from other sources, such cellphone Bluetooth data; 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, Forthcoming 2017).

HR 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). HR data reduced the need much of the guesswork and the need for non-bus proxies in understanding bus performance between bus stops, which improved the applicability of buses as probe vehicles. These same researchers expanded their use of HR data to multi-stop 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, et al., Forthcoming 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. (Forthcoming 2017) has 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.

Improving on previous results, HR data was 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 data set leaving only pass-by buses (Glick
& Figlioizzi, Forthcoming 2017). 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 time point 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.

2.2. 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 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.

2.2.1. Using Bayesian Statistics

One of the primary comparison methodologies is the empirical Bayesian approach, which has been employed for more than thirty years to perform observational before-and-after analysis. Persuad & Lyon (2007) summarized the literature of Bayesian methodologies up to 2006, claiming that many of the problems of these types of studies stem from insufficient data (e.g. traffic or collision counts). Since there exists variation in traffic counts and crash rates between months or between years, results using just a few counts may be insufficient. Even with defined experimental approaches that include control groups and randomly assigned treatments, most methodologies do not account for how changes in traffic volume, weather, or driving trend affect results.

The debate over frequentist versus Bayesian statistics continues; each is useful in some situations and less useful in others. The choice of which approach to use is highly dependent on the type of data and whether that data will follow a known distribution.
Despite its wide use in before-and-after studies, the methodology proposed will not use Bayesian statistics as there is little-to-no need to correct for insufficient counts and a known distribution is unnecessary to calculate variances and standard deviations of percentiles. For this research, which is descriptive, an estimate of variance can be directly calculated through asymptotic variance and probability mass function (Ringner, 2009), which are described in the methodology section.

2.2.2. **Descriptive Analysis**

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.

2.3. **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 have 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, et al., Forthcoming 2017).

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 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.

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., 2003). All seven 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.
3. DATA SOURCES

This study relied on two types of data: automatic vehicle location (AVL) and geographic information systems (GIS) shape-files. 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, they updated their bus dispatch system (BDS) with high-resolution (HR), GPS based data collection system with five-second, time-based, resolution. The SE, SD, and HR 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 (Crumley, 2017).

Upon request, TriMet provided three sets of AVL data: SE, SD, and HR. Each of the AVL data sets 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 data set provide a more detailed account of individual and aggregated bus behaviors.

To augment these data sets, this research also uses GIS files provided on TriMet’s public website (TriMet, 2016). The GIS data provides 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 their 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 (Stoll, 2016).

3.1. Stop-to-Stop Data

Stop-to-stop data, also called stop event (SE) data for this study, collects information at bus stops whether or not a bus actually serves passengers at a given stop. This data includes bus operational data information including, but not limited to, arrive time, leave time, dwell time, average speed between stops, and passenger movements. SE data is 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 HR data collection system are available through transit agencies, SE data provides the primary means for researchers to determine route-level performance metrics for that agency. Unfortunately, the use of SE data only allows 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. Stop Disturbance Data

Stop disturbance data expands on the information collected in the SE data set by also including points between stops where the wheel of the bus stop moving. At each of these locations, the data set records time, door activity, and stop type. Time-points, 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-thrus, 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 2 minutes then 5 mph (8.0 kph) for one minute from another bus that traveled at 29 mph (46.7 kph) for 3 minutes; for that, additional information is required.

3.3. High-Resolution Data

HR data, which collects data in up to five-second intervals, augments TriMet’s previously implemented SE and SD data sets 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.
HR 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 is not recorded if the bus is not in motion at the time of the next scheduled recording. This creates a situation where the HR data shows low speeds in segments where a bus stopped rather than no speed. In Figure 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 it is known the bus stops. Integrating the data sets can correct for these problems.

Figure 1 – Bus trajectories using combined SE and SD data versus HR data.
4. METHODOLOGY

This research relies heavily on the R programming language in R-Studio interface (RStudio Team, 2016) and ArcMap (ESRI, 2016) to clean and process the data. After merging the data sets, 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.

All equations that are directly used for this research are numbered.

4.1. TriMet Data Integration

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

Data integration, using a created unique identification number (UID), begins with SD and SE data. Individual trips are separated based on UID, route number, and direction along that route. HR 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 stop disturbance, so the events are duplicated in the combined set; pass-by stops would not be recorded as disturbance 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 HR 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 2.

![Bus Trajectory](image)

Figure 2 – Bus Trajectory from Figure 1 using the combined data set.

The integrated data also provides 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 contains multiple routes that sometimes, but not always, overlap. As such, GPS data must be assigned to an underlying map that show which routes are present on each segment and to separate out different directions of travel for the same route.

4.2. Spatial Data

The spatial data provided by TriMet shows the location of all stops and routes as GIS shape files (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.

4.2.1. Spatial Data Cleaning

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 3, 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.

![GIS flowchart for processing non-overlapping route data.](image)

4.2.2. GIS Tool Definitions

The tools used in Figure 3 are defined in Table 1 (ESRI, 2016). Most are part of ArcMap’s basic license. However, Polygon to Centerline is an added tool that requires an advanced license (Dilts, 2015).
Table 1 – GIS tools definitions with visuals

<table>
<thead>
<tr>
<th>Tool</th>
<th>Definition</th>
<th>Visual Representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buffer</td>
<td>Creates a polygon around input features at a specified distance.</td>
<td><img src="image" alt="Buffer Visual" /></td>
</tr>
<tr>
<td>Dissolve</td>
<td>Combines like features based on specific attribute or combination of attributes.</td>
<td><img src="image" alt="Dissolve Visual" /></td>
</tr>
<tr>
<td>Polygon to Centerline</td>
<td>Create a centerline profile of a polygon input.</td>
<td><img src="image" alt="Polygon to Centerline Visual" /></td>
</tr>
<tr>
<td>Split</td>
<td>Divides an input based on a specified distance into a set of inputs that sum to the size of the original input.</td>
<td><img src="image" alt="Split Visual" /></td>
</tr>
<tr>
<td>Trim Line</td>
<td>Removes portion of a line that extend a specific distance past the intersection of two lines.</td>
<td><img src="image" alt="Trim Line Visual" /></td>
</tr>
<tr>
<td>Merge</td>
<td>Combines multiple input datasets of the same type into a single, new output dataset with combined attributes.</td>
<td><img src="image" alt="Merge Visual" /></td>
</tr>
</tbody>
</table>

4.2.3. **Final Combined Data Set**

The final combined data set 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, 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.
4.3. 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, 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 $n_{ji}$ bus trips that pass through each sub-segment $i$, with each individual bus designated by the index $j$. Due to the data-cleaning process, $n_{ji}$ may fluctuate between adjacent sub-segments.

$$i \in I = \{1, 2, 3, ..., n_I\}$$

$$j \in J_i = \{1, 2, 3, ..., n_{ji}\}$$
4.4. Unknown Distributions

For each sub-segment, $i$, the velocity of each bus, $v_{ij}$, that passes a centerpoint, $c_i$, 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 2500 ft. segment along Powell Blvd. at 25 ft. increments. Figure 5 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 5](image-url)

Figure 5 – 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 6 shows this percent difference between Figure 5 (top) and Figure 5 (bottom). To reduce visual clutter, Figure 6 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 6 – Point distribution density for real data compared with normal approximation

Some locations appear to follow a normal distribution, such as $x = 2000$ (Figure 7: left); however, most locations, like $x = 200$ (Figure 7: right), do not, nor do they follow any typical distribution.

Figure 7 – Speed histogram at $x = 2000$ ft. (left) and $x = 200$ ft. (right) from Figure 6.

Furthermore, when data does 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 is skewed. Skewness is a measure of asymmetry of the probability distribution about a mean. Figure 8 shows a positive skew distribution from Figure 6 at $x = 125$ ft. (left) and negative skew distribution from $x = 1,925$ ft. (right). Non-skewed normal data typically has near-equal means, medians, and mode. This is not the case for skewed data.
Due range of possible speed distributions, an approach that does not required 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 HR data (Glick & Figliozzi, Forthcoming 2017), but with added methodologies for percentiles and confidence intervals.

$$p = \text{percentile}$$

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

Any $s_i$ has associated percentile travel speeds found by ordering the data and finding an observation 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 9 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. See

![Histogram and CDF](image1)

![Spline smoothed function](image2)

![PDF estimate](image3)

Figure 9 – 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):

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

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

23
(>160) (Brown & Wolfe, 1983), this estimate of variance may be used to estimate the confidence intervals for each \( \hat{\nu}_{i,p} \), assuming a normal distribution. For a confidence level \( \alpha \) and its associated \( z \)-score, \( z(\alpha) \), the range of percentile values that may represent an estimated percentile is found:

\[
[\hat{\nu}_{i,p} - \hat{\sigma}_{\nu_{i,p}} \cdot z(\alpha), \hat{\nu}_{i,p} + \hat{\sigma}_{\nu_{i,p}} \cdot z(\alpha)]
\]  \hspace{1cm} (2)

This interval provides extremes of the \( \alpha \) confidence interval around \( \nu_{i,p} \).

See Appendix (page 77) for example of percentile standard deviation estimates with randomly generated and normally distributed data. Variances and standard deviations are calculated with above methodology (equation 1) (Brown & Wolfe, 1983) and methodology that assumes normality used in a previous publication by Glick et al. (Glick & Figliozzi, Forthcoming 2017). Also, see Appendix (page 80) for difference in confidence using real data.

### 4.5. Peak Hour Vs Whole Day Performance

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

\[
\Delta \hat{\nu}_i = \hat{\nu}_{i,85} - \hat{\nu}_{i,15}
\]  \hspace{1cm} (3)

\[
\hat{\mu}_i = \frac{\Delta \hat{\nu}_i}{\hat{\nu}_{i,50}} = \frac{\hat{\nu}_{i,85} - \hat{\nu}_{i,15}}{\hat{\nu}_{i,50}}
\]  \hspace{1cm} (4)
If a SVI is greater than 1, it indicates a location where median travel speed is more similar to the 15\textsuperscript{th} percentile travel speed than the 85\textsuperscript{th}.

4.6. Propagation of Uncertainty

The estimates of uncertainty, $\hat{\sigma}$, that been previously calculated for each percentiles, as a standard deviation, $\hat{\sigma}_{v_{lp}}$, should be carried through in the calculations for speed variability and speed variability index. Each of the estimated percentile travel speeds ($\hat{v}_{l,15}$, $\hat{v}_{l,50}$, and $\hat{v}_{l,85}$) has an associated and normally distributed error (i.e., the standard deviations, $\hat{\sigma}_{v_{l,15}}$, $\hat{\sigma}_{v_{l,50}}$, and $\hat{\sigma}_{v_{l,85}}$); 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 $\Delta\hat{v}_l$ and appropriate formulae for the propagation of uncertainties for ratios, such as $\hat{\mu}_l$, (Taylor, 1997).

4.6.1. Assumption of Normality

The top left of Figure 10 shows the density and normal approximation of the distribution of travel speeds at $x = 700$ of Figure 6. This distribution does not appear to be normal (upper right). Using 100 samples of 150 non-replaced data points. The 15\textsuperscript{th}, 85\textsuperscript{th} and 50\textsuperscript{th} percentiles were estimated 100 times. The distribution of the 15\textsuperscript{th} (upper right), 50\textsuperscript{th} (middle left), and 85\textsuperscript{th} (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 follows normal
distributions. This indicates that the equations for the propagation of error will results in standard deviations where normality can be assumed.

![Figure 10](image)

Figure 10 – Speed histogram at \( x = 700 \) ft. from Figure 6 and distributions of 15\(^{th}\), 50\(^{th}\), and 85\(^{th}\) percentile and speed variability and speed variability index.

### 4.6.2. Propagation of Uncertainty Addition and Subtraction

The addition and subtraction of normally distributions data results in normally distributed data. The addition and subtraction of \( N \) values with independent uncertainties is defined generally as:

\[
q = x_1 + \cdots + x_N
\]
where \( q \) is the result and the values \( x_1, \cdots, x_N \) are independent and carry standard deviations \( \sigma_{x_1}, \cdots, \sigma_{x_N} \), respectively, then

\[
\sigma_q = \sqrt{\sigma_{x_1}^2 + \cdots + \sigma_{x_N}^2}
\]

For \( \hat{v}_{l,85} \) and \( \hat{v}_{l,15} \), 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:

\[
\hat{\sigma}_{\Delta v_i} = \sqrt{\hat{\sigma}_{v_{l,85}}^2 + \hat{\sigma}_{v_{l,15}}^2 - 2 \cdot \text{cov}(\hat{v}_{l,85}, \hat{v}_{l,15})}
\]  

(5)

### 4.6.3. 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. (Taylor, 1997).

For a case of a general function \( q \) consisting of \( N \) projects and \( M \) quotients:

\[
q = \frac{x_1 \cdot \cdots x_N}{y_1 \cdot \cdots y_M}
\]

with independent uncertainties \( \sigma_{x_1}, \cdots, \sigma_{x_N} \) and \( \sigma_{y_1}, \cdots, \sigma_{y_M} \), the uncertainty of the quotient \( \sigma_q \) is defined by:

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

27
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 $\sigma_{x_1}$ and $\sigma_{x_2}$, respectively, $q$ and its standard deviation, $\sigma_q$, 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:

$$\frac{\widehat{\sigma}_{\mu_i}}{|\widehat{\mu}_i|} = \sqrt{\left(\frac{\widehat{\sigma}_{\Delta v_i}}{\Delta v_i}\right)^2 + \left(\frac{\widehat{\sigma}_{\mu_{L50}}}{\mu_{L50}}\right)^2 - 2 \cdot \frac{\text{cov}(\Delta v_i, \mu_{L50})}{\Delta v_i \cdot \mu_{L50}}} \quad (6)$$

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

$$\frac{\widehat{\sigma}_{\mu_i}}{|\widehat{\mu}_i|} = \sqrt{\left(\frac{\widehat{\sigma}_{\mu_{L85}}}{\mu_{L85}}\right)^2 + \left(\frac{\widehat{\sigma}_{\mu_{L15}}}{\mu_{L15}}\right)^2 - \frac{2 \cdot \text{cov}(\mu_{L85}, \mu_{L15})}{\mu_{L85} - \mu_{L15}^2}} + \left(\frac{\widehat{\sigma}_{\mu_{L50}}}{\mu_{L50}}\right)^2 - 2 \cdot \frac{\text{cov}(\mu_{L85} - \mu_{L15}, \mu_{L50})}{(\mu_{L85} - \mu_{L15}) \cdot \mu_{L50}}}$$

(7)

4.6.4. Confidence Intervals

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

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

$$CI_{\mu_i} = [\widehat{\mu}_i - \sigma_{\mu_i} \cdot z(\alpha), \widehat{\mu}_i + \sigma_{\mu_i} \cdot z(\alpha)] \quad (9)$$
If 0 falls within the upper and lower bound of the confidence interval, no speed variability can be said to exist between the $15^{th}$ and $85^{th}$ percentile (i.e. the null hypothesis cannot be rejected). An SVI greater than 1 indicates severe speed variability in a given segment, $i$.

4.7. Travel Time

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 $n_f$ is the total number of buses, $p$ is the percentile. $\hat{t}_p$ is the percentile travel time with an estimated standard deviation of $\hat{\sigma}_{t_p}$.

$$\hat{\sigma}_{t_p}^2 = \frac{p(1-p)}{f(i_p)^2 \cdot n_f}$$

(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 $n_f$. On average each percentile travel time should seen an equal number of times.

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

(11)

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

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

(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).
4.8. **Space-Time-Speed Diagrams**

Speed data can also be viewed after aggregating the data by time of day through the use of 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.

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

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

4.9. **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: $\beta_0$ and $\beta_1$, represent the initial conditions and current conditions, respectively. A $\delta$ added before a variable will indicate a value calculated from $\beta_0$ and $\beta_1$.

4.9.1. **Travel Speeds, Travel Times, and Peak Performance**

The differences in the percentile speeds, travel times and peak hour performance metrics of $\beta_0$ and $\beta_1$ are compared using simple subtraction.
\[
\delta \hat{v}_{i,p} = \hat{v}_{i,p,\beta_1} - \hat{v}_{i,p,\beta_0} \\
\delta \bar{t}_p = \bar{t}_{p,\beta_1} - \bar{t}_{p,\beta_0} \\
\delta \Delta \hat{v}_i = \Delta \hat{v}_{i,\beta_1} - \Delta \hat{v}_{i,\beta_0} \\
\delta \hat{\mu}_i = \hat{\mu}_{i,\beta_1} - \hat{\mu}_{i,\beta_0}
\]

(14) \hspace{1cm} (15) \hspace{1cm} (16) \hspace{1cm} (17)

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

\[
\delta X = X_{\beta_1} - X_{\beta_0}
\]

(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:

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

(19)

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

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

(20)

If 0 falls within the confidence interval, 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. STUDY AREAS

The three areas used for analysis (i.e. Lombard Street, 16th Street, and Powell Blvd.) are all in Portland, OR. Each location underwent a notable change that had the potential to improve or disrupt transit operations. The effect of those changes were measured using data collected before and after implementation of each change by the TriMet. The segment performance before the change is compared to performance after the change to determine overall effect.

The purpose of the three study areas included is not to make broad claims about road diets, bridges, or route changes. Rather, the purpose is to provide examples of how the proposed methodology can be applied to different situations and how it can be applied in the future to other road diets or general case before-and-after 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 in Portland, OR that extended for 1,100 ft. Counts available for the Portland Bureau of Transportation (PBOT) show an ADT of 6800 and pm peak-hour volume 650 vehicles for westbound travel and an ADT of 6400 with 500 vehicles during the pm peak-hour for eastbound travel (PBOT, 2017). 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., 2003).
Figure 11 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 ft. and ends at 2,275 ft. The middle and bottom images are the aerial view of the before and after conditions, respectively.

Figure 12 shows a Streetmix (Code for America, 2017) cross section with lane dimensions marked and a Google Streetview (Google, 2017) of the lane configuration before and after road diet. The total width of roadway is 50 ft.

The data collected before the roadway change (before data) includes approximately 2,300 buses from all weekdays between 15 June and 31 July in 2015. The data collected after the change (after data) includes approximately 1,900 buses for all weekdays between 20 July and 26 August 2016.
Figure 12 – Streetmix and Google Streetview of Lombard Street cross-section before (top) and after (bottom) road diet.
5.2. Case Study II: 16th street

The second analysis is of a road diet consisting of a travel lane reduction at 16th Street in Portland, OR. The initial cross-section of 16th street included a bike lane and two travel lanes in each direction. These lane markings were replaced with an 8 ft. bike lane, 5 ft. buffer, and an 11 ft. vehicle travel lane.

Figure 13 shows a map (top) that includes measurements that correspond to x-axis of results. The road diet begins at 125 ft. and ends at 1,100 ft. The two aerial views show the road before (middle) and after (bottom) the road diet.

Using PBOT traffic counts (PBOT, 2017), the average daily traffic counts are approximately 3730 and 4160, southbound and northbound, respectively with pm 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 includes 1,100 buses from all weekdays between 12 June and 3 July in 2015. The after data includes 1,700 buses all weekdays between 14 July and 12 August 2016.
The Streetmix and Google Streetview cross sections of 16th Street before and after the road diet are shown in Figure 14. The dimensions of the lanes given and the total width on either side of the median is 25.5 ft.
Figure 14 – Streetmix and Google Streetview of 16th Street cross-section before (top) and after (bottom) road diet.
5.3. Case Study III: Powell Blvd. and Tilikum Crossing

The third analysis covers changes to bus operation on Powell Boulevard resulting from the completion of the Tilikum Crossing, a new bridge over the Willamette River. Route 9, which historically crossed the Ross Island Bridge, was diverted to the new crossing, which is the largest bridge in the United States that does not allow private passenger vehicles or freight (PBOT, 2017). The bridge is designed for light rail, streetcar, bikes, buses, and pedestrians. TriMet, which built and operates the bridge, expected the new bridge would reduce travel times and improve efficiency for Route 9 and Route 17 (TriMet, 2017). This case study examines those claims.

The changes to bus operation are first examined for the section of Powell Blvd. where the diverted Route 9 converge on the east side of the river. Traditionally, this area is highly congested during peak hours as Powell Blvd carries upwards of 40,000 vehicles daily. In addition to the travel speed percentile comparisons that are conducted on the road diet segments, Powell was also examined statistically for changes by time of day to isolate when the roadway experiences increases in speed. Figure 15 shows the location of the analysis area. This segment saw no physical changes to road structure but is examined for performance changes that may be the result of the diverted route.

![Figure 15 – Map of Study Area for Powell Blvd.](image-url)
Figure 16 shows the change in route 9. The red line follows route 9 before the opening of the bridge; the green line follows route 9 after. Eastbound buses travel the same distance while westbound buses must travel an additional 0.2 miles.

Figure 16 – Map of route 9 before and after change.
6. 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 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 17 shows the change in travel times over the Lombard study area. Top two plots summarize all trips taken throughout the day and the bottom two plots are for the morning commute only (between 7:00am and 10:00 am). 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 thirty 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 17 – Change in travel time by percentile on Lombard Street.

### 6.1.2. Travel Speeds

Figure 18 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 results indicate little no know operational change for transit or general traffic as a results of the lane narrowing.
Figure 18 – Difference in travel speeds between before and after cases along Lombard Street
6.2. Case Study II: 16th Street

6.2.1. Travel Times

Figure 19 shows travel times by percentile over the study area of 16th Street for both directions of travel. The width represents the 95th percentile confidence interval.

After subtracting the before and after cases (Figure 20), a statistically significant increase in travel times is observed for both directions of travel; but, for all percentiles, the increase was less than 1 minute. Assuming that all travel increases were attributed to the 1,100 ft. road diet, the travel time increase would remain below five minutes per miles in the worst case and less than one minute per mile on average.
6.2.2. Travel Speeds

The changes in speed observed (Figure 21) are within 5 miles per hour of the original speed for most locations and percentiles. The speed decrease above this range in northbound direction Figure 21, at \( x = 400 \) ft., is the result of a bus stop being added to the route following the completion of the road diet.

Figure 21 – 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 $x = 1350$ ft. 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 $x = 100$ ft.

The minor differences in speeds can be seen in Figure 22, which shows speed variability (eq. 3) and speed variability index (eq. 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 $x = 400$ ft., where the new stop was added. Overall, these results indicates little total impact on the operational speeds of transit caused by a lane reduction.
6.3. Case Study III: Powell Boulevard

6.3.1. Westbound Travel Times (Bridge)

Figure 23 shows the travel time across the two miles segment of changed route shown in Figure 16 using data from buses at all times of day. Using data only from the AM-peak (7:00 am to 10:00 am) revealed a near identical pattern to all-day performance (Figure 43).
Data from the PM-peak (4:00 pm to 7:00 pm) (Figure 24) shows a similar travel time increases as is seen in Figure 23, except for the slowest ten percent of trips, which saw a reduction in travel times.

6.3.2. **Eastbound Travel Times (Bridge)**

Eastbound travel over the area shown in Figure 16 experienced a larger decrease in travel time for a higher percentage of buses than westbound travel. By time-of-day (Figure
the majority of trips are shown to experience a travel time increase of approximately four minutes. While the majority (70%) of trips in the PM commute saw a travel time decrease (Figure 26), none of the trips during the AM-peak saw a decrease in travel time (Figure 44).

Figure 25 – Travel times and travel time differences over river in eastbound direction for all buses.

Figure 26 – Travel times and travel time differences over river in westbound direction for PM-peak travel.
6.3.3. *Hourly Travel Times (Bridge)*

Additionally, average travel time became much more consistent by time of day after (A) the change then before (B) for both eastbound (EB) and westbound (WB) travel (Table 2). For example, the sharp increase in travel times in eastbound travel at 17:00 is notably reduced after the route change. Travel times are for the same distance covered in Figure 23 through Figure 26.

Table 2 – Route 9 hourly travel times over river on Powell Blvd.

<table>
<thead>
<tr>
<th>Time</th>
<th>EB</th>
<th>WB</th>
<th>EB</th>
<th>WB</th>
<th>EB</th>
<th>WB</th>
<th>EB</th>
<th>WB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>6:00</td>
<td>9.8</td>
<td>13.7</td>
<td>8.2</td>
<td>12.2</td>
<td>2.1</td>
<td>1.1</td>
<td>0.9</td>
<td>1.3</td>
</tr>
<tr>
<td>7:00</td>
<td>10.4</td>
<td>14.2</td>
<td>10.6</td>
<td>13.9</td>
<td>1.1</td>
<td>1.5</td>
<td>2.5</td>
<td>1.3</td>
</tr>
<tr>
<td>8:00</td>
<td>10.5</td>
<td>14.2</td>
<td>11.0</td>
<td>13.9</td>
<td>1.1</td>
<td>2.0</td>
<td>2.0</td>
<td>1.5</td>
</tr>
<tr>
<td>9:00</td>
<td>9.6</td>
<td>14.1</td>
<td>9.7</td>
<td>13.4</td>
<td>1.3</td>
<td>1.0</td>
<td>1.5</td>
<td>1.4</td>
</tr>
<tr>
<td>10:00</td>
<td>9.8</td>
<td>14.4</td>
<td>8.6</td>
<td>12.9</td>
<td>1.2</td>
<td>1.3</td>
<td>1.8</td>
<td>1.4</td>
</tr>
<tr>
<td>11:00</td>
<td>9.7</td>
<td>14.6</td>
<td>8.8</td>
<td>13.3</td>
<td>0.9</td>
<td>1.4</td>
<td>1.7</td>
<td>1.6</td>
</tr>
<tr>
<td>12:00</td>
<td>10.4</td>
<td>14.5</td>
<td>8.8</td>
<td>13.3</td>
<td>1.6</td>
<td>1.4</td>
<td>2.0</td>
<td>1.3</td>
</tr>
<tr>
<td>13:00</td>
<td>10.6</td>
<td>14.7</td>
<td>8.9</td>
<td>13.5</td>
<td>1.1</td>
<td>1.3</td>
<td>2.1</td>
<td>1.3</td>
</tr>
<tr>
<td>14:00</td>
<td>11.1</td>
<td>14.6</td>
<td>9.3</td>
<td>14.1</td>
<td>1.7</td>
<td>1.7</td>
<td>2.0</td>
<td>2.3</td>
</tr>
<tr>
<td>15:00</td>
<td>15.5</td>
<td>15.0</td>
<td>10.8</td>
<td>13.6</td>
<td>6.1</td>
<td>2.0</td>
<td>3.2</td>
<td>1.5</td>
</tr>
<tr>
<td>16:00</td>
<td>19.9</td>
<td>15.8</td>
<td>10.4</td>
<td>13.5</td>
<td>5.3</td>
<td>2.0</td>
<td>2.6</td>
<td>1.1</td>
</tr>
<tr>
<td>17:00</td>
<td>22.7</td>
<td>16.3</td>
<td>12.8</td>
<td>13.8</td>
<td>6.7</td>
<td>2.4</td>
<td>4.5</td>
<td>1.2</td>
</tr>
<tr>
<td>18:00</td>
<td>14.8</td>
<td>15.0</td>
<td>10.2</td>
<td>13.5</td>
<td>5.9</td>
<td>1.8</td>
<td>3.8</td>
<td>1.2</td>
</tr>
<tr>
<td>19:00</td>
<td>10.3</td>
<td>14.6</td>
<td>8.1</td>
<td>13.2</td>
<td>1.7</td>
<td>2.1</td>
<td>1.6</td>
<td>1.6</td>
</tr>
<tr>
<td>20:00</td>
<td>9.6</td>
<td>14.3</td>
<td>6.9</td>
<td>12.6</td>
<td>1.2</td>
<td>1.7</td>
<td>0.8</td>
<td>1.1</td>
</tr>
</tbody>
</table>

### 6.3.4. Costs

The increased travel time for the majority of trips increased TriMet operational costs in both directions over the altered portion of the route. Table 3 shows these costs including the range of the 95th percentile confidence interval. However, the decreased variability has the potential to reduce overall cost as it reduces TriMet’s need to add unscheduled buses when routes are delayed.
Table 3 – Average daily TriMet operational costs before and after the roadway change.

<table>
<thead>
<tr>
<th></th>
<th>Westbound</th>
<th>Eastbound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost [$/hour]</td>
<td>Before 92.79</td>
<td></td>
</tr>
<tr>
<td></td>
<td>81</td>
<td>92.79</td>
</tr>
<tr>
<td>Trips per Day</td>
<td>Before 81</td>
<td>After 83</td>
</tr>
<tr>
<td>Mean Cost per Day [$/day]</td>
<td>1.165 ± 4.2</td>
<td>1.650 ± 1.6</td>
</tr>
<tr>
<td>Difference (after-before) [$/day]</td>
<td>485 ± 4.5</td>
<td>223 ± 8.9</td>
</tr>
<tr>
<td></td>
<td>1.630 ± 8.5</td>
<td>1.853 ± 2.5</td>
</tr>
</tbody>
</table>

6.3.5. Eastbound and Westbound Travel Times (Half Route)

Expanding the travel time estimates to include an additional 5 miles of route 9 on Powell Blvd., which represents the most heavily congested part of route 9 showed that travel times still increased for approximately 85% of all trips when examined for a whole day (Figure 27). Figure 45 in Appendix shows travel times for the same segment in the AM-peak (top) and PM-peak (bottom) travel periods.

![Travel Time Graphs](image)

Figure 27 – Travel times over 7.2 miles of Powell Blvd for all trips. Top: westbound. Bottom: eastbound

For the 7.2-mile stretch along Powell Blvd.,
Table 4 also shows that travel time consistency improved by time of day when more of the route is examined while still showing an increase in travel times for most times of day.

Table 4 – Route 9 hourly travel times over 7.2 miles on Powell Blvd.

<table>
<thead>
<tr>
<th>Time</th>
<th>Mean Travel Time [min]</th>
<th>Travel Time Standard Deviation [min]</th>
<th>Average Load</th>
<th>EB Trips Per Day</th>
<th>WB Trips Per Day</th>
</tr>
</thead>
<tbody>
<tr>
<td>6:00</td>
<td>31.4 36.2 34.2 38.9</td>
<td>2.5 1.4 3.3 2.9</td>
<td></td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>7:00</td>
<td>34.1 40.1 47.8 48.5</td>
<td>2.8 2.7 10.2 5.7</td>
<td>4</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>8:00</td>
<td>35.1 40.2 50.5 51.2</td>
<td>2.4 2.4 10.7 7.7</td>
<td>4</td>
<td>13</td>
<td>6</td>
</tr>
<tr>
<td>9:00</td>
<td>34.9 40.2 40.9 44.0</td>
<td>3.0 2.5 6.7 3.5</td>
<td>5</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>10:00</td>
<td>33.0 40.9 36.4 42.8</td>
<td>19.9 1.8 2.8 3.3</td>
<td>5</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>11:00</td>
<td>35.6 42.3 38.2 44.4</td>
<td>2.5 2.3 2.9 4.0</td>
<td>6</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>12:00</td>
<td>37.8 44.6 38.6 44.1</td>
<td>2.3 3.6 3.2 2.7</td>
<td>7</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>13:00</td>
<td>39.3 44.3 39.0 43.8</td>
<td>2.5 2.3 3.7 3.0</td>
<td>8</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>14:00</td>
<td>39.1 45.8 41.0 43.9</td>
<td>2.3 2.8 5.5 3.7</td>
<td>10</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>15:00</td>
<td>46.4 49.2 41.8 45.1</td>
<td>7.0 4.4 4.6 4.2</td>
<td>12</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>16:00</td>
<td>50.4 51.4 42.0 44.8</td>
<td>5.7 4.5 6.6 2.2</td>
<td>13</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>17:00</td>
<td>52.9 50.4 47.4 44.8</td>
<td>6.0 4.2 11.8 3.4</td>
<td>12</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>18:00</td>
<td>45.4 45.5 38.3 42.2</td>
<td>7.7 3.9 9.0 3.1</td>
<td>12</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>19:00</td>
<td>34.3 41.9 33.4 40.9</td>
<td>14.5 2.9 2.4 2.6</td>
<td>10</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>20:00</td>
<td>33.3 40.8 30.8 38.2</td>
<td>2.2 3.8 2.3 2.4</td>
<td>8</td>
<td>7</td>
<td>4</td>
</tr>
</tbody>
</table>

6.3.6. Travel Speeds

The travel speed differences in the eastbound direction, following the changes to Powell Blvd indicate substantial changes to bus performance. However, the study area (excluding where route changed) was not significantly altered physically. The road was repainted without altering the location of lines. The decrease in speed in Figure 28 at \( x = 250 \) ft. is accounted for by the addition of a new bus stop. Prior to the change in route, that bus stop was a nearside stop approximately fifty feet before the start of the analysis area. The before and after routes do not overlap. As such, the previous bus stop and the new left turn on the route are excluded.
By percentile, much of eastbound travel appears to have decreased in speed. By
time of day (calculated through a moving average), the majority of the decreased speed is
seen to be in the PM-peak period.

In the westbound direction, travel speeds increased significantly for the 10th
through 25th percentiles. This decreased travel time is concentrated during the evening
commute. Typically, traffic patterns suggest that vehicles are attempting to leave the city
center at that time-of-day. The time of day plot in Figure 29 shows that there is less congestion heading into the city after the route change then there was before.

Figure 29 – Difference in westbound travel speeds along Powell Blvd.
7. DISCUSSION

7.1. Road Diets

The methodology proposed in this research was applied successfully to two separate road diets. The results of that application indicate that road diets are unlikely to have a notable impact on the operations of transit when lanes have been narrowed or removed in cases where ADTs and peak-hour traffic volumes remain below recommended thresholds. The research by Knapp, et al. (2003) indicates that operations should be unaffected for road diets on low ADT streets. 16th Street and Lombard Street confirm these results as both showed little-to-no operational changes.

Road diets have been well studied and these results do not add significant information to the body of research already available. However, these test cases are evidence that the methodology is generally applicable and may be applied to examine other and more complicated road diet projects.

7.2. Powell Blvd. and the Tilikum Crossing

When applied to route 9, the methodology provides the means to quantify changes before and after the implantation of the route 9 diversion over the Tilikum crossing. The results showed that for a majority trips, travel times actually increased for route 9. One possible cause for the decreased travel times is the 25 miles per hour speed limit for buses and trains (TriMet, 2017). This limit is a product of line-of-sight requirements, which are limited by the grade of the bridge, which to provide enough clearance for ships is just under five percent for the majority of its length (Anderston, 2015).
Applied for just the peak travel periods, travel times decreased for a majority of trips. Additionally, the range of possible travel times during the peak periods was reduced. This increased predictability in trip duration may reduce the need for added and unscheduled buses during peak periods, which may reduce long-term costs.

The results of this case study are not comprehensive, as they do not analyze the full route 9 or the other bus routes that cross the bridge. As such, generalizations about the effectiveness of transit-only bridges or the Tilikum Crossing specifically cannot be made. Rather, this research indicates that before-and-after studies can be applied when routes move as well as to test the effect of a physical change to a roadway.

7.3. Recommendations

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 his problem is limited to the HR 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.
8. CONCLUSION

This thesis expands on the existing systems of transit analysis to create applicable methodologies using higher-resolution data that quantify transit performance changes following a roadway modification. When applied to determine the congestion impact caused by the implementation of road diets with low ADTs, this approach indicates that overall congestion remains near constant before and after the change. Additionally, this approach highlights the impact of a route change over a segment of a congested transit line. Results indicated the change reduced congestion in the transit line leading up to the modified route, increased travel times for most trips while reducing travel time during peak periods, and significantly lower travel time variability across all trips.

The successful application of the proposed methodology for the two road diets and Powel Blvd. indicate applicability over a wide range of ADTs and roadway configurations. These test cases shows the potential to apply the same methodologies in future research of longer road diets and road diets with higher ADTs, as well as being generally applicable to short and long corridors.

More generally, this research provides a means 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 high-resolution, stop event, and stop disturbance data provides more information than any one data set 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 transit network with the appropriate data collection systems.
REFERENCES


Glick, T. B. & Figliozzi, M., Forthcoming 2017. Novel Methodology to Estimate Traffic and Transit Travel Time Reliability Indexes and Confidence Intervals at Corridor and Segment Levels. *Transportation Research Record: Journal of the Transportation Research Board*.


Gudz, E., Fang, K. & Handy, S., Forthcoming 2017. When a Diet Prompts a Gain: Impact of Road Diet on Bicycling in Davis, California. *Transportation Research Record: Journal of the Transportation Research Board*.


60
APPENDIX A: SUPPLEMENTAL FIGURES

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, like Figure 30, 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 condition. 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 percentiles travel speeds, plots like Figure 31 shows 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.
A.1. Lombard Street

Figure 30 – Eastbound travel speed by percentile on Lombard Street. Top: before road diet. Bottom: after road diet. Used for Figure 18 (top).
Figure 31 – Eastbound speed variability and speed variability index on Lombard Street.
Figure 32 – Westbound travel speed by percentile on Lombard Street. Top: before road diet. Bottom: after road diet. Used for Figure 18 (bottom).
Figure 33 – Westbound speed variability and speed variability index on Lombard Street.
A.2. 16th Street

Figure 34 – Northbound travel speed by percentile on 16th Street. Top: before road diet. Bottom: after road diet. Used for Figure 21 (top).
Figure 35 – Northbound speed variability and speed variability index on 16th Street.
Figure 36 – Southbound travel speed by percentile on 16th Street. Top: before road diet. Bottom: after road diet. Used for Figure 21 (bottom).
Figure 37 – Southbound speed variability and speed variability index on 16th Street.
A.3. Powell Blvd

Figure 38 – Eastbound travel speed by percentile on Powell Blvd. Top: before road diet. Bottom: after road diet. Used for Figure 28 (top).
Figure 39 – Eastbound travel speed by time-of day on Powell Blvd. Top: before road diet. Bottom: after road diet. Used for Figure 28 (bottom).
Figure 40 – Westbound travel speed by percentile on Powell Blvd. Top: before road diet. Bottom: after road diet. Used for Figure 29 (top).
Figure 41 – Westbound travel speed by time-of-day on Powell Blvd. Top: before road diet. Bottom: after road diet. Used for Figure 29 (bottom).
Figure 42 – Westbound speed variability and speed variability index on Powell Blvd.
Figure 43 – AM-peak travel times and travel time differences in westbound direction on Powell Blvd.

Figure 44 – PM-peak travel times and travel time differences in eastbound direction on Powell Blvd.
Figure 45 – Travel time on Powell Blvd. Top: 7:00 am to 10:00 am. Bottom: 4:00pm to 7:00pm.
APPENDIX B: CODE FOR STANDARD DEVIATION EXAMPLE

#The following is for the estimation of:
#cumulative distribution function
#probability distribution function
# standard deviations calculated through probability and assumption of normality

#Options 1: non-normal data
x <- c(rnorm(85,30,5.5),rnorm(25,6,5,5))

#Options 2: normal data (used in this example)
x <- rnorm(1000,30,5.5)

#subset between 0 and 50 to simulate real speeds
x <- subset(x,x>0)
x <- subset(x,x<=50)

#plot the histogram of data
hist(x,xlim=c(0,50),xlab="Veh Speed [mph]", ylab = "Frequency")

#sort data by size and apply an index
x <- sort(x)
y <- seq(1:length(x))
z <- seq(1:length(x))/length(x)

#apply a spline smoothing with a smoothing parameter of .618
x2 <- smooth.spline(x, z, spar=.618)

#plots the resulting CDF and spline smoothed graph
plot(x,z,xlim=c(0,50),xlab="Veh Speed [mph]",
     ylab = "Cumulative Distribution",
     pch=16, col=rgb(0,0,0,maxColorValue = 1, a = .15))
lines(x2,col=2)
#calculate the derivative at the CDF at the location of data.
x3 <- predict(x2, deriv=1)

#plot the resulting PDF
plot(x3, xlim=c(0, 50), xlab="Veh Speed [mph]",
     ylab = "Probability",
     pch=16, col=rgb(0,0,0, maxColorValue = 1, a = .15))
lines(x3, col=2)

#Create a matrix to hold the difference between the normal approximation and probability version of the standard deviation
abc <- as.data.frame(matrix(nrow=3, ncol=3))

#rename rows and columns
colnames(abc) <- c("speed", "sigma_prob", "sigma_norm")
rownames(abc) <- c(.15, .5, .85)

#set the percentiles to be used
px1 <- c(.15, .5, .85)

for (i in 1:3) {
  px <- px1[i]

  #calculate the the value of the given percentile
  p <- as.numeric(quantile(x, px))
# calculate the slope of the CDF line at the percentile
temp <- predict(x2,p,deriv=1)
p_d1 <- temp$y

# calculate the standard deviation based on 1983 equations using probability
sigma2 <- (1/((p_d1)^2))*((px)*(1-px))/(length(x))
sigma <- sqrt(sigma2)

# calculate the standard deviation index based on assumption of normality
sigma2idx <- length(x)*px*(1-px)
sigmaidx <- round(sqrt(sigma2idx),0)

# use index to find upper and lower bound of one standard deviation for percentile
sigmaidx_upper <- x[round(length(x)*px,0)+sigmaidx]
sigmaidx_lower <- x[round(length(x)*px,0)-sigmaidx]

# subtract values and divide by two to find standard deviation for percentile
sigmaidx_temp <- (sigmaidx_upper - sigmaidx_lower)/2

# store the results
abc[i,1] <- p
abc[i,2] <- sigma
abc[i,3] <- sigmaidx_temp
}

# display results
abc

## p   speed   sigma_prob   sigma_norm
## 0.15 24.34177   0.2753419    0.3214624
## 0.5  30.16335   0.2141526    0.2022942
## 0.85 35.82190   0.2225545    0.2152049
APPENDIX C: TRAVEL TIME CONFIDENCE ESTIMATES

Instead of using probabilities, as this research did, confidence intervals may be calculated assuming a normal distribution of the data (Glick & Figliootti, Forthcoming 2017). Figure 46 and Figure 47 show differences in confidence intervals calculated based on real probabilities verses when normality is assumed.

Figure 46 – Difference in confidence interval for travel time depending on if probabilities are used or normality is assumed.

Figure 47 – Difference in confidence interval for change in travel time depending on if probabilities are used or normality is assumed.