EXPLORING TRANSIT PERFORMANCE AND TRAFFIC CONGESTION IN DOWNTOWN
TORONTO USING BIG DATA

by

Christopher Chun Kong Yuen
Bachelor of Commerce, University of British Columbia, 2012

A Major Research Paper
Presented to Ryerson University

in partial fulfillment of the requirements for the degree of

Master of Planning
in
Urban Development

Toronto, Ontario, Canada, 2017

Author's Declaration for Electronic Submission of MRP

I hereby declare that I am the sole author of this MRP. This is a true copy of the MRP, including any required final revisions.

I authorize Ryerson University to lend this MRP to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this MRP by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my MRP may be made electronically available to the public.

EXPLORING TRANSIT PERFORMANCE AND TRAFFIC CONGESTION IN DOWNTOWN
TORONTO USING BIG DATA
© Christopher Chun Kong Yuen, 2017

Master of Planning

in

Urban Development

Ryerson University

Abstract

This exploratory research evaluates the linkages between roadway operations and mixed-traffic
transit performance on three arterial corridors in Toronto- King Street, Queen Street, and Dundas
Street. Using Inrix traffic speed probe data as well as GPS location data from the Toronto Transit
Commission's vehicles between January 2014 and June 2016, this research visualizes spatial and
temporal trends in traffic congestion and transit headway regularity. Three regression models were
estimated that indicate both traffic congestion and terminus departure times are statistically
significant, but weak predictors of mixed-traffic transit reliability. These models leave most of the
variability unexplained. The findings highlight opportunities and limitations for congestion
management and transit scheduling as tools for improving headway reliability. They also illustrate
the complexity of the relationships between transportation modes in downtown Toronto.

Acknowledgements

I would like to thank Dr. Matthias Sweet, for offering me the opportunity to carry out this study despite my obvious lack of technical experience with Big Data analysis at the beginning of this project. His patience and guidance throughout the year has been instrumental to the success of this work and I owe him a great debt of gratitude. Dr. Sweet's dedication to fostering impactful research led to his co-founding of the TransForm Transportation and Land Use Laboratory with Dr. Raktim Mitra in 2016. The Ryerson School of Urban and Regional Planning is fortunate to have Dr. Matthias Sweet as a part of its faculty and I have been lucky to have him as a supervisor.

I would also like to thank the Transit Planning and Transit Data teams at the City of Calgary. During my summer employment with Calgary Transit in 2016, they provided constant guidance and inspiration for me to learn more about data-driven decision making in transportation planning.

Lastly, I would like to thank my colleagues and friends at the School of Urban and Regional Planning. Working with quantitative data within a specialized transportation project, it is easy to lose sight of the big picture in city building. Spending time with friends from many backgrounds and with diverse interests has helped me maintain the broad perspective that urban planners, as generalists, need.

I especially thank my good friends Andrew Sgro and Kailey Laidlaw, with whom I have shared experiences in three provinces in Canada and three countries overseas, for spending countless hours chatting with me about nearly every topic under the sun. From organizational behaviour to fiscal policy, and from global geopolitics to culinary arts, our conversations over the past two years have probably made a better person and a better urban planner. I'm not really sure how, but I'm quite confident nonetheless. Their professionalism in project settings, and friendship and camaraderie outside of the classroom embodies the spirit of the tight knit culture that makes this graduate planning program such a success.

Dedicated to my father

Contents

Table of Figures

1.0 Introduction

Faced with an forecasted annual growth of over 130,000 residents by year 2031, the immense pressures placed on the City of Toronto transportation system has been met with policy directives that list public transit as an integral tool for mitigating the negative impacts of traffic congestion. Significant transit investment in the Greater Toronto Area is currently underway with projects such as the Eglinton Crosstown Light Rail Transit line as well as the Yonge-Spadina subway extension expected to have a significant impact on travel patterns. Despite the continued implementation of transit expansion on grade-separated (ROW A) or exclusive right-of-way corridors (ROW B), much of the surface transit network in the City of Toronto continues to operate in mixed-traffic (ROW C), with potential implications on schedule and headway reliability. Given that the performance of mixed-traffic transit may be directly impacted by traffic congestion itself, are additional policy levers needed to enable transit as a path to mitigate the impacts of congestion?

Across the City of Toronto arterial roadway network, rush-hour congestion levels have been steadily increasing over the past years (Sweet, Harrison, & Kanaroglou, 2015). While policies for development intensification within built-up areas such as the Ontario Places to Grow Act and the City of Toronto's Official Plan aim to increase in the share of trips made by walking, cycling, and transit, it is unlikely that the trend of increasing roadway congestion will reverse in the foreseeable future. At the same time, concerns regarding transit service unreliability in Toronto have recently been brought to light and specifically targeted by new operating procedures by the TTC (Toronto Transit Commission, 2015).

Traditionally, roadway performance has been measured based on impacts on general vehicular traffic, but as the City of Toronto adopts a broader, multi-modal approach that recognizes the diverse role of streets in the public realm, the implicit impacts of policy actions on transit

operations needs to be better understood. As such, estimating the extent to which streetcar performance is a function of traffic congestion, external factors, or operating characteristics is important for effectively improving transit services. If reduced transit performance is largely a function of traffic congestion, this would suggest policy recommendations such as dedicated lanes, increased transportation demand management strategies, or signal timing changes may be effective. However, if traffic plays no role and transit reliability is fundamentally determined by initial operating conditions, this suggests that a focus on improving transit operating procedures would yield the best results.

This exploratory research exploits newly available Big Data sources on transit operations and roadway traffic in Toronto to increase understanding of reliability trends across three arterial corridors- King Street, Queen Street, and Dundas Street. Furthermore, it attempts to establish several models for roadway congestion and transit operating conditions as predictors of transit reliability in Toronto.

2.0 Background & Previous Studies

2.1 Congestion trends in Toronto

Roadway congestion in the City of Toronto has grown significantly between 2011 and 2014, with speeds across city-wide and downtown arterials decreasing by 7kph during the peak period, (Sweet, Harrison, & Kanaroglou, 2015). While travel speeds within the downtown core are generally lower than on the arterial network elsewhere in the city, travel time reliability is higher (Sweet, Harrison, & Kanaroglou, 2015). This is fairly consistent with overall trends across other North American cities where hours of delay per automobile commuter has been increasing over the past decades in cities of all sizes (Schrank D. , Eisele, Lomax, & Bak, 2015).

2.2 Factors influencing transit reliability

The Transit Capacity and Quality of Service Manual, 3rd edition, provides a reference guide and standardized operational metrics to be used by transit agencies. It describes several factors that influence transit reliability and finds that they can be generally characterized into two categories- "internal" factors under a transit agency's control, and "external" factors outside of a transit agency's control (Transportation Research Board, 2013). Figure 1- Factors Influencing Transit Reliability, lists these items in detail.
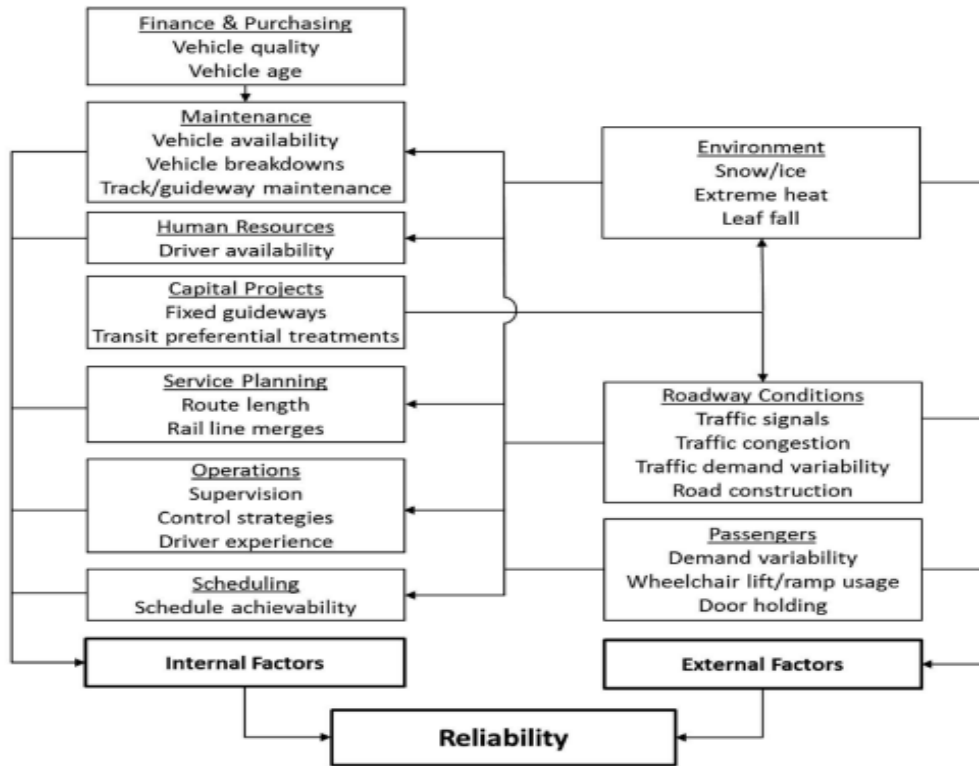
Figure 1: Factors Influencing Transit Reliability (Transportation Research Board, 2013)

The causes of day-to-day transit travel time variability for on-road public transport has been studied through a Melbourne case study in 2008. In this research, a model was developed to predict travel time variability through explanatory variables such as route length, stop spacing, time of the day, land use, and timing point departure delay (Mazloumi, Currie, & Rose, 2008). This study found departure delay from timing points to be significantly associated with travel time variability, with every minute of early running per unit section length correlated to an increase in travel time variability of 4% (Mazloumi, Currie, & Rose, 2008).

In 2015, the Toronto Transit commission piloted a streetcar service quality improvement initiative on the 512 St Clair and 504 King routes whereby operational procedures such as scheduling, additional supervising staff, and the location of spare vehicles were adjusted to improve service reliability (Leary, 2015). Prior to the pilot, short turning- the procedure of prematurely

taking a transit vehicle out of service along a route to turn it around and fill a gap in service in the opposite direction caused by headway unreliability, was regularly used, to the detriment of customer satisfaction. By the end of the year, short turns on the 504 King route decreased from approximately 310 per week to 40 per week (Toronto Transit Commission, 2015). This suggests that transit agencies' own operating procedures have major impacts on schedule reliability.

2.3 Bunching phenomenon

On frequent transit lines with short headways between vehicles, variability in headway is exacerbated by the "bunching" phenomenon. This behaviour, also known as "platooning", is caused by two properties that work in tandem with a positive feedback effect. Firstly, dwell time at transit stops are positively correlated with the number of passengers boarding and alighting (Pilachowski, 2009). At the same time, the number of people waiting to board at each stop generally increases with the time between vehicle arrivals. When the first vehicle begins to fall behind schedule, the next vehicle faces shorter dwell times as little time has elapsed since the preceding vehicle's departures from each stop, causing the second vehicle to catch up.

Transit bunching is associated with "longer waiting times for some riders, uneven passenger distribution, overcrowding in late [vehicles], and an overall decrease on level of service and capacity" (Feng & Figliozzi, 2010).

2.4 Roadway congestion's link to transit operations

Based on a review of literature, I found no previous study directly linking roadway congestion and transit reliability. Instead, studies have focused on different elements of congestion or transit operations and transit performance. For example, the New Jersey Department of Transportation and US Federal Highway Administration explored the links between roadway

congestion and transit operating speeds in 2003. This study used a regression model to estimate

transit run times under free flow conditions, comparing that with actual observed run times

(McKnight, Levinson et al., 2003).

One study used a linear regression approach to model variations between scheduled

headways and actual headways, finding that a 15 percent increase in scheduled travel time

correlates with a 160 percent increase in deviation from scheduled headway" (Cooper & Gould,

1994). Scheduled travel times were used as a proxy for congestion as it was expected that the

transit agencies would schedule in extra travel time during periods of the day when congestion is

common.

2.5 Measurements of headway reliability

Transit agencies use several reliability metrics, the most common of which is schedule

adherence. The Toronto Transit Commission uses an on-time performance standard of +- 3 minutes

at timing points (Toronto Transit Commission, 2005). Schedule adherence is considered to be

important to the customer experience particularly when headways exceed 15 minutes. When transit

customers make their travel decisions based on a transit schedule, schedule adherence determines

whether the customer experiences excessive wait time at the stop or misses the vehicle altogether.

Conversely, for transit services operating at short headways like Toronto's extensive "10

minute or better" network, "headway adherence becomes more important from the perspective of a

passenger" (Saberi & Zockaie, 2013). For example, if every transit vehicle on a route scheduled at

10-minute headways is 10 minutes late, a schedule adherence metric would determine this to be a

failure even though customers' perceptions may be very favorable.

On the other hand, if vehicles along a 10-minute headway route alternated between being

exactly 3 minutes early and 3 minutes late, the schedule adherence metric may indicate that these

services are very successful, while in actuality, passengers may experience up to a 16-minute wait when the average is 10 minutes. As such, headway regularity can be a more meaningful metric for customer experience on routes with frequent service.

3.0 Knowledge Gap

Transportation planning in urban areas involves trade-offs between many competing interests. Especially within the spatially constrained context of urban centres, some cities are attempting to optimize the use of public streets not just for mobility objectives, but also as components of the public realm that enhances quality of life for city dwellers. For example, the City of Toronto's Transportation Services Division is currently implementing measures in the 2016-2020 Congestion Management Plan, not to fight traffic congestion outright, but to limit the economic impacts of it, and to manage the effects of it on strategically important travel modes (City of Toronto, 2016). However, for cities to improve their understanding of the linkages between different transportation modes in terms of congestion and reliability, planning practitioners need high quality data on speed and travel time. For transit, complete datasets on schedule adherence and headways are essential. This kind of data has historically been difficult to capture.

3.1 Traditional data sources

Traditionally, travel time and delay studies which measure real-world traffic speeds, have been dependent on data collected through manual sampling methods. The most common techniques have required the use of test vehicles, also referred to as "floating cars", where drivers are employed to drive in the traffic stream for the sole purpose of data collection (Turner & Eisle, 1998). These methods were labour intensive and costly, and therefore impractical for the collection of large samples. Schedule adherence and headway reliability were similarly difficult to capture for transit vehicles. The resulting small sample sizes, while useful for estimating averages, were generally insufficient for estimating extreme values, such as the 85th percentile (Furth & Hemily, 2006). Yet, for many transportation outcomes, such as travel time and transit reliability, it is these

extreme values that are most frustrating for commuters as they must plan for them during their day to day lives.

3.2 Big Data Promises

The recent emergence of "Big Data" has brought about new opportunities for transportation planning analysis. With anonymous, aggregated data being generated from mobile devices and vehicle fleets, transportation studies can now cover a much larger geographic and temporal spread than was possible with the manual sampling approaches of the past. Not only that, Big Data allows for the study of extreme events and the statistical distribution of transportation outcomes. By combining datasets of performance measures of different transportation modes, it is now also possible to model the relationships between potentially competing planning outcomes.

4.0 Research Design

This project is exploratory in nature and as such, focuses on the descriptive analysis of trends and patterns in both transit performance and streetcar reliability. It also explores the linkages between traffic congestion and transit performance through a regression framework. This predictive model takes the following functional form:

$$y = \beta x + \varepsilon$$

where y represents transit headway reliability, β represents a matrix of coefficients representing estimated links between exogenous variables (x) and the dependent variable of interest (y), and ε represents a vector of error terms.

4.1 Data Sources

Data to complete this analysis was procured from two sources. Traffic data was purchased from Inrix as a part of the City of Toronto Big Data Innovation project. Inrix traffic data consists of billions of anonymous, crowdsourced "GPS probe" speed data points collected from private and commercial vehicles, as well as from GPS-enabled smartphones such as iPhone and Android devices (Eagle & Greene, 2014). These speed and travel time readings are recorded at the link level, and are coded by directionally-specific "Traffic Message Centre" (TMC) identifiers. These links are illustrated below in figure 2.
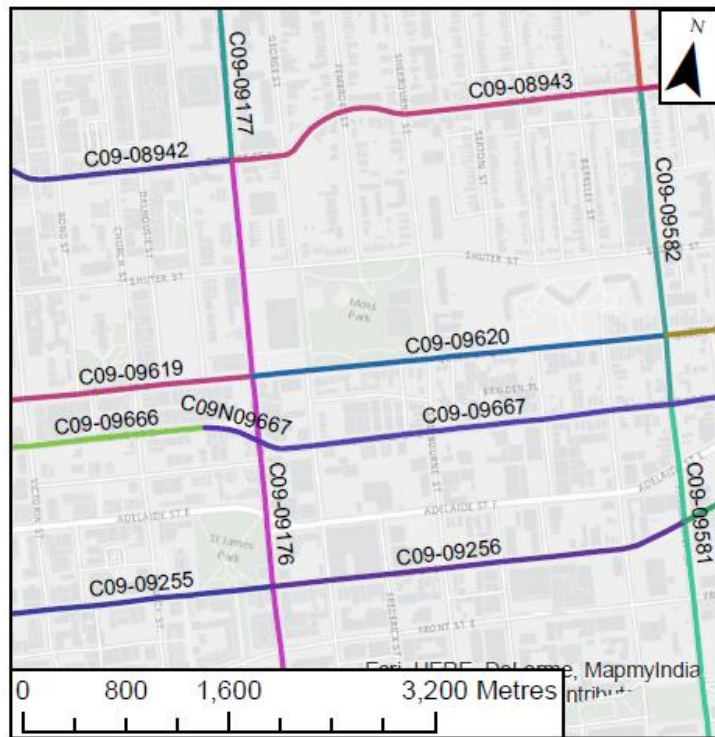
Figure 2: Illustration of roadway links and Inrix TMCs. Note that TMCs are directional and this illustration only includes westbound and southbound TMCs for clarity

Stop locations were drawn from the TTC's General Transit Feed Specification (GTFS) dataset published on the City of Toronto's open data portal (Toronto Transit Commission, 2016). In addition to the location coordinates of each stop, the GTFS data also provided a unique Stop ID code and a name for each stop.

Archived transit vehicle location data was provided by the Toronto Transit Commission through an agreement with the City of Toronto. While the Toronto Transit Commission is in the process of procuring a modern Computer Aided Dispatch/Automatic Vehicle Location (CAD/AVL) system which is expected to collect high quality location data, CAD/AVL data was not available at the time of this study. As such, vehicle location data was retrieved from TTC's legacy Communications and Information system (CIS). This system collects location-at-time data every 20 seconds, creating a complete record of every bus and streetcar's physical movements throughout

the day. Every record includes a timestamp, a route identifier, a run number, a vehicle identifier, and geographic coordinates.

4.2 Transit Routes and Study Corridors

This research focuses on three major transit routes in Toronto- 501 (Queen St), 504 (King St), and 505 (Dundas St). These routes were chosen for their unique characteristics as high-frequency surface routes operating in mixed traffic. The 504 (King St.) route is the busiest surface route in Toronto, with over 65,000 boardings per weekday (Toronto Transit Commission, 2016). These routes were also chosen because they travel through Downtown Toronto, an area targeted for additional development intensification in the Official Plan (City of Toronto, 2015).   Downtown Toronto is a part of the TOcore planning initiative, which will potentially introduce changes to the use of public right-of-way to increase multi-modal transportation options and to allow for public realm improvements (City of Toronto, 2017).

While all three routes are primarily served by streetcars, during the peak hours, route 504 and 505 are often supplemented by buses. This is evident in the CIS data, which differentiates between vehicle number and type. With the ability to maneuver around obstacles, buses by themselves may be less susceptible to headway unreliability. However, as buses operate on the same street and serve the same customers as streetcars, it is difficult to isolate each from the other for analysis. As such, this study assumes that buses and streetcars running on these routes do not exhibit drastically different operating characteristics.

It should be noted that for the purposes of this project, the analysis of Route 501 was truncated to eliminate the western portion of the route, some of which operates on dedicated right of way (ROW B), and where Inrix traffic data is unavailable. Maps of the analyzed stop locations of these routes are below:

Route 501 Stops Analyzed
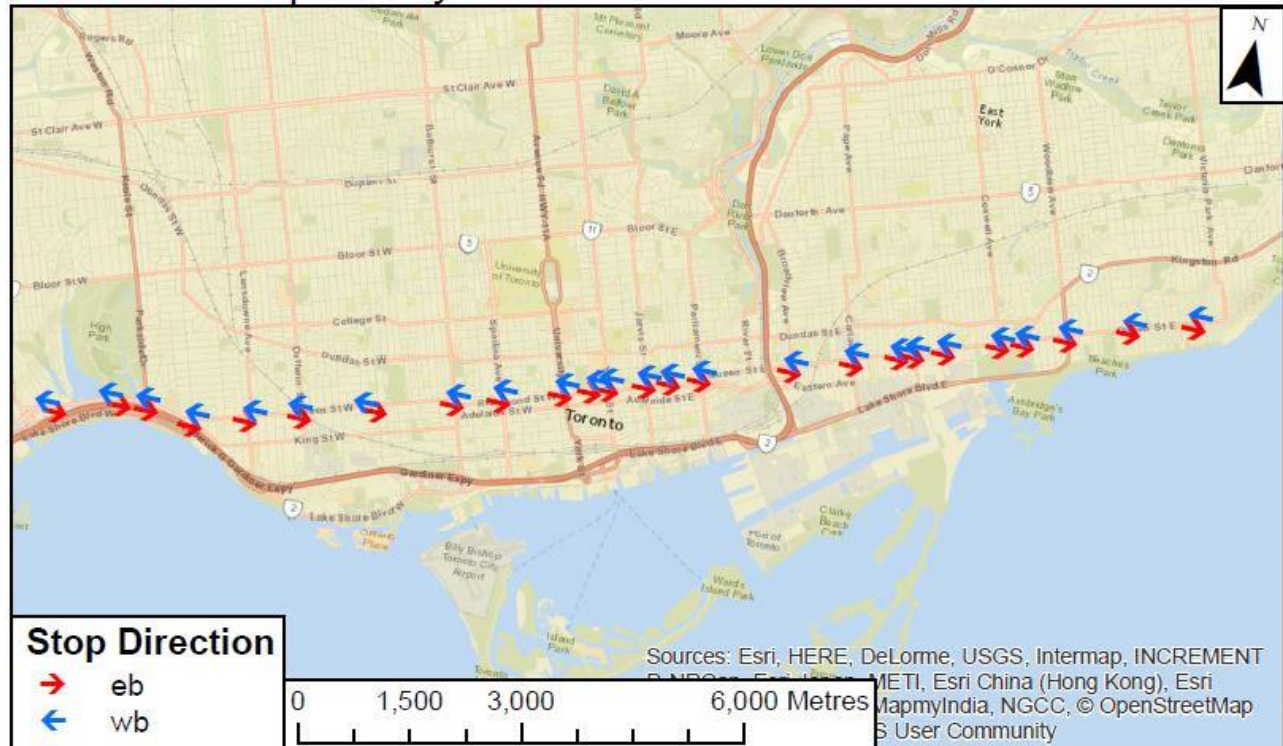
Figure 3: Stop locations included in analysis of Route 501 (Queen St)

Route 504 Stops Analyzed
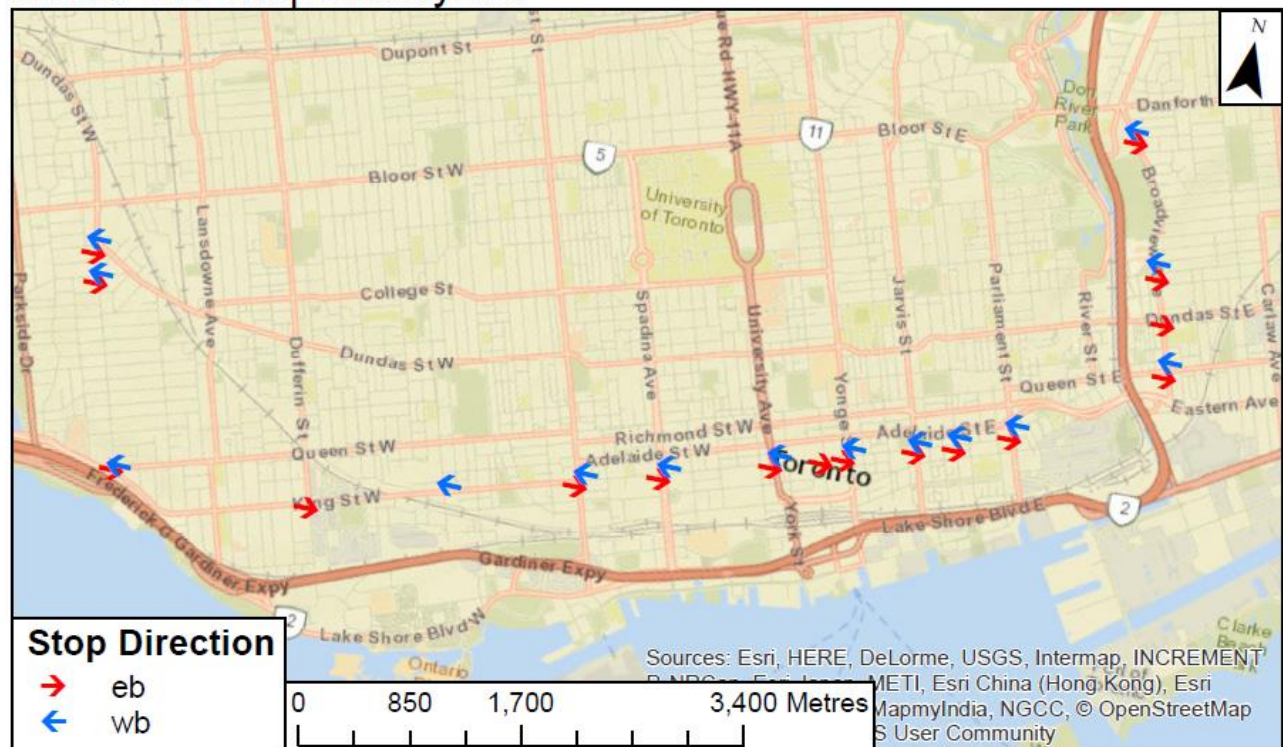
Figure 4: Stop locations included in analysis of Route 504 (King St)

Route 505 Stops Analyzed
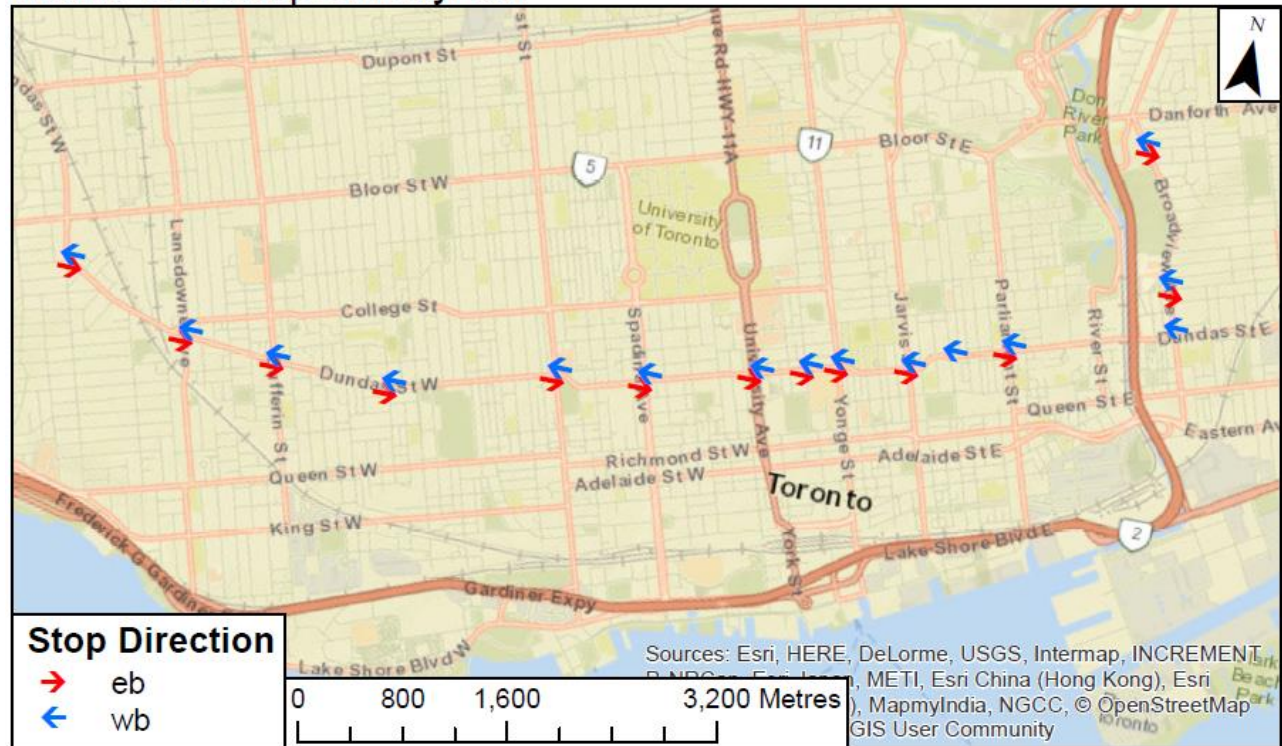
Figure 5: Stop locations included in analysis of Route 505 (Dundas St)

4.3 Methodology

Vehicle location data can be presented in two ways. The first is "location-at-time", where a location record is generated at a regular, fixed time interval. The second is "time-at-location", where a record is generated only when a vehicle passes by a pre-determined location of interest, such as a transit stop or major intersection. Headway analysis requires the use of "time-at-location" data because headway is defined by the time between successive arrivals at a fixed location. Since the Toronto Transit Commission's CIS data is in "location-at-time" format, this study processes the data to estimate "time-and-location" at the transit stops of interest.

This project used the R programming language for Transit CIS and Inrix traffic data processing. The next section provides an overview of the main steps of analysis.

14

4.3.1 Mapping Stop Locations

Using the GTFS to Shapefile conversion tool (Morang, 2017), all TTC transit stops were mapped in ArcGIS. To reduce the number of stops analyzed, the timepoints of route 501, 504, and 505 were isolated from the full dataset based on information on the TTC's website (Toronto Transit Commission, 2016). While they are considered timepoints by the TTC, the terminus station of all the routes were removed from the analysis and replaced with the second stop and the penultimate stop on each route. This was done to mitigate challenges with estimating travel direction and departure times from the terminus stations where transit vehicles loop around and layover in preparation for their next trip. Because stops used in this study correspond with TTC timepoints, whereby observed headways can be expected to best match their scheduled intent, deviations from scheduled headway may be higher at other stations. As such, the estimates of headway variability in this study may best be interpreted as lower bounds.

4.3.2 CIS Data Processing

The sheer size of the CIS dataset presented computational challenges for headway analysis. With over 13 gigabytes of "location-at-time" records for each route, it was necessary to subset and reduce the data to be processed for each stop level analysis. Furthermore, the CIS system occasionally produced rogue data points, with GPS coordinates that were nowhere near the transit routes, some of which were in the middle of Lake Ontario.
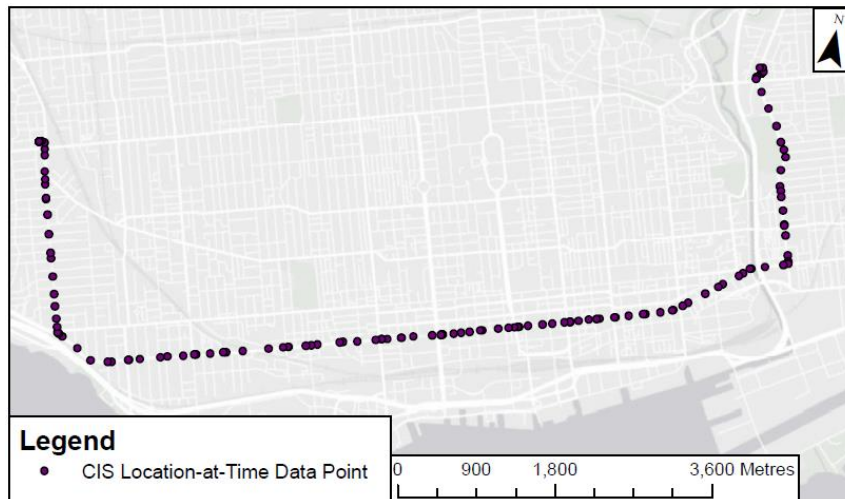
Figure 6: CIS Data for one trip in one direction

To subset the CIS data by the geographic location of each station, CIS data points within 125 metres of each stop of interest were isolated. The timestamp of the CIS data point closest to each stop was then used to approximate arrival times at the stop.
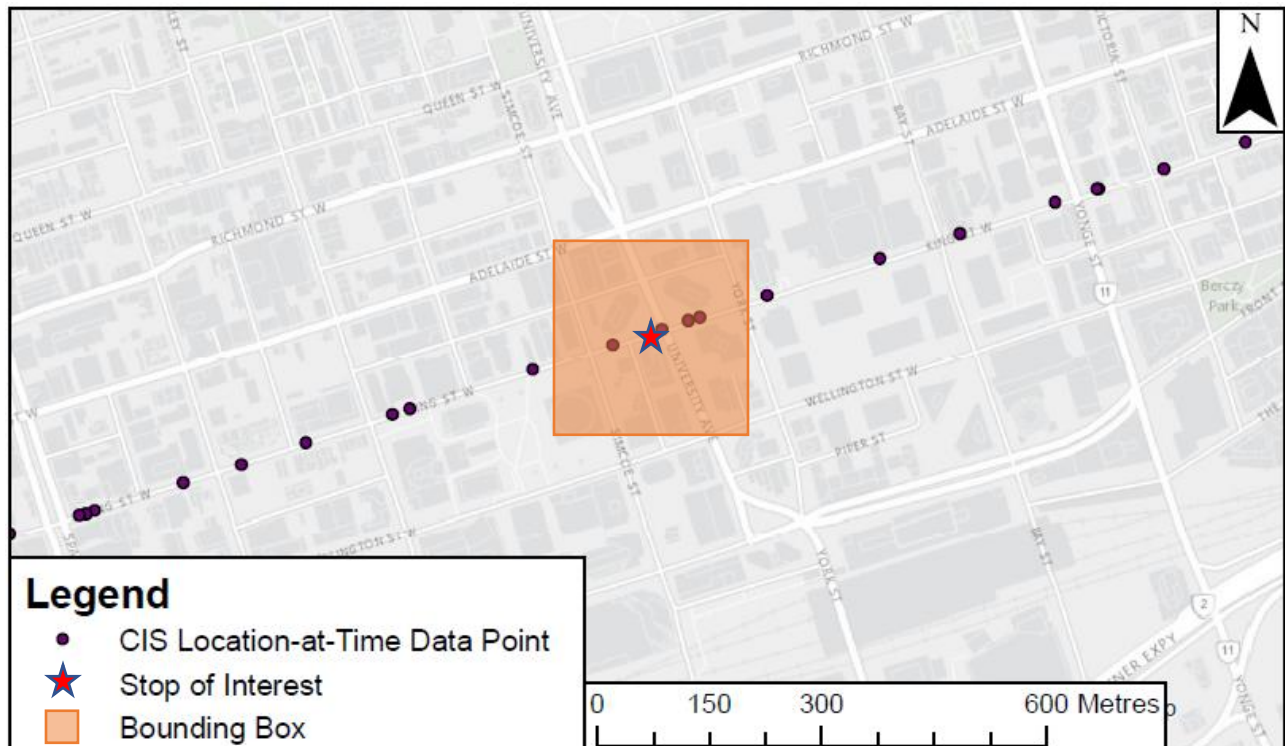
Figure 7: Illustration of bounding box applied around one stop of interest

Within the weekday PM peak period, this analysis was limited to a two-hour window during the weekday PM peak between 4:00 pm and 6:00 pm. This study assumes that any linkages between traffic congestion and transit reliability sufficiently propagate within a 2-hour window to be reflected in the results. While it is likely that the onset of peak hour congestion can have immediate impacts on transit running speeds, its effects on headway reliability is not immediate as it takes time for bunching and gapping to develop between transit vehicles. Likewise, transit base operating conditions at the first stop may not have an immediate impact on headway reliability down the line as it takes every transit vehicle approximately one hour to travel from one terminus to the other.

Since the CIS data did not differentiate by travel direction, this analysis determined directionality using a lag function for each trip based on time and location. This was done through a comparison of the geographic coordinate of the subject data point and the point seven observations prior.

4.3.3 Estimating Base Operating Conditions

To enable analysis of the impact of transit base operating conditions, headway statistics from the first stop of every route in each direction was appended to each record. This allowed for further investigation into whether headway regularity at the start of each route was linked with headway regularity throughout the remainder of each route.

4.4 Aggregation by Month

All data in this study was aggregated by month, and by year. The study assumes that monthly aggregation does not occlude important day to day effects. For example, this study assumes that if a given month has five days of extreme traffic congestion that has severe impacts on transit, this congestion is reflected in the buffer-time index of that month. However, depending on

the actual distribution of periods of abnormal traffic congestion, and whether there is sufficient month to month variation in the aggregated values, it is possible for monthly aggregation to hide important trends. This is a limitation of the research design.

Summary statistics of the main outputs are available in Appendix A: Table of Summary Statistics by Route.

4.5 Dependent Variable

Headway Buffer Time Index (BTI) 85 was chosen as the dependent variable that represents transit reliability. This measure is the difference between 85th percentile headways and 50th percentile headways, divided by the mean headway. For example, based on a headway BTI85 of 2.0, a route with a scheduled headway of every 5 minutes, would see actual headways of shorter than 10 minutes, 85 percent of the time.

$$Headway\ BTI85 = \frac{85th\ percentile\ headway - 50th\ percentile\ headway}{mean\ headway}$$

Headway BTI85 differs slightly from the industry standard metric for headway reliability-the coefficient of variation of headways (Cvh) as defined below.

$$Cvh = \frac{SD\ of\ headway\ deviations}{scheduled\ headway}$$

Headway BTI85 and Cvh are similar in that they both include a measure of deviation from typical headway in the numerator, and a measure of typical headway in the denominator. For both measures, a higher number corresponds to a worse experience for transit customers.

4.6 Independent Variables

Five predictors of streetcar headway reliability were explored in this analysis. These variables are explained in detailed below:

4.6.1 Scheduled Headway

Mean headway, the average amount of time between successive transit vehicle arrivals, measured in minutes, was included to show the actual number of trips that serve each stop. It serves as a control for the possibility of route diversions that lead to no service being offered at some stops during construction or other disruptions. It also serves as a proxy for the scheduled headway at each stop.

4.6.2 Distance from Start of Route

The Stop Number variable is an ordinal estimate of the distance from the starting terminus of each route in each direction. As the distances between selected stops in the analysis were not consistent, this variable should not be treated as a perfectly linear estimate of distance from the starting terminus. Stop order was expected to be positively correlated with transit unreliability based on a previous study that showed a general upward trend in transit unreliability as transit routes progressed from the starting terminus to ending terminus (Feng & Figliozzi, 2010).

4.6.3 Transit Base Operating Conditions

Transit Headway Buffer Time Index (BTI) 85 was selected from the first stop along the route to determine the regularity of departures from the starting terminus. This measure was included as an estimate of the base operating conditions allowed by the TTC's schedules of terminus recovery times and dispatching practices. It was expected that unreliability at the first stop would be positively correlated with unreliability further down the line.

4.6.4 Traffic Unreliability

Derived from Inrix travel time data, Buffer Time Index (BTI) represents the ratio between the 95th percentile slowest travel times to the average travel time during the same time of the day. It reflects the amount of extra "buffer" time a person must budget to be on-time 95 percent of the

time. Since BTI captures extreme values, it is indicative of some months having more days where PM peak traffic is abnormally slow. Especially in mixed traffic environments like King Street, Queen Street, and Dundas Street, it was expected that BTI would be positively linked with transit unreliability.

4.6.5 Traffic Congestion

Travel Time Index (TTI) is a comparison between the average travel time during the peak period and travel time under free flow conditions like what a driver may experience at 4am in the morning. It represents the typical rush hour congestion during every month of study instead of the extreme days that BTI covers.

4.7 Expected Signs

Independent Variable	Expected Sign
Transit Headway Mean (ln)	?
Stop Order	+
Transit Headway BTI 85 (ln) at beginning of route	+
Traffic Buffer Time Index (BTI) (ln)	+
Traffic Travel Time Index (TTI) (ln)	+

5.0 Results

Using the outlined methodology, this study explores the potential linkages between road operations and transit performance in Toronto's city centre. Specifically, it estimates traffic congestion and transit base operating conditions as predictors of transit headway reliability. Results are presented first by focusing on general descriptive statistics and then by highlighting model results.

5.1 Descriptive Statistics

As this study is based on a complete dataset of transit movements along three routes, as well as data on traffic speed along corresponding roadway links, some trends and patterns are visible through descriptive statistics. Specifically, two topics are explained below- seasonality and location.

5.1.1 Seasonality

Seasonal trends are apparent in both transit reliability and roadway traffic congestion. The next few graphs explore these seasonal trends using several different metrics.
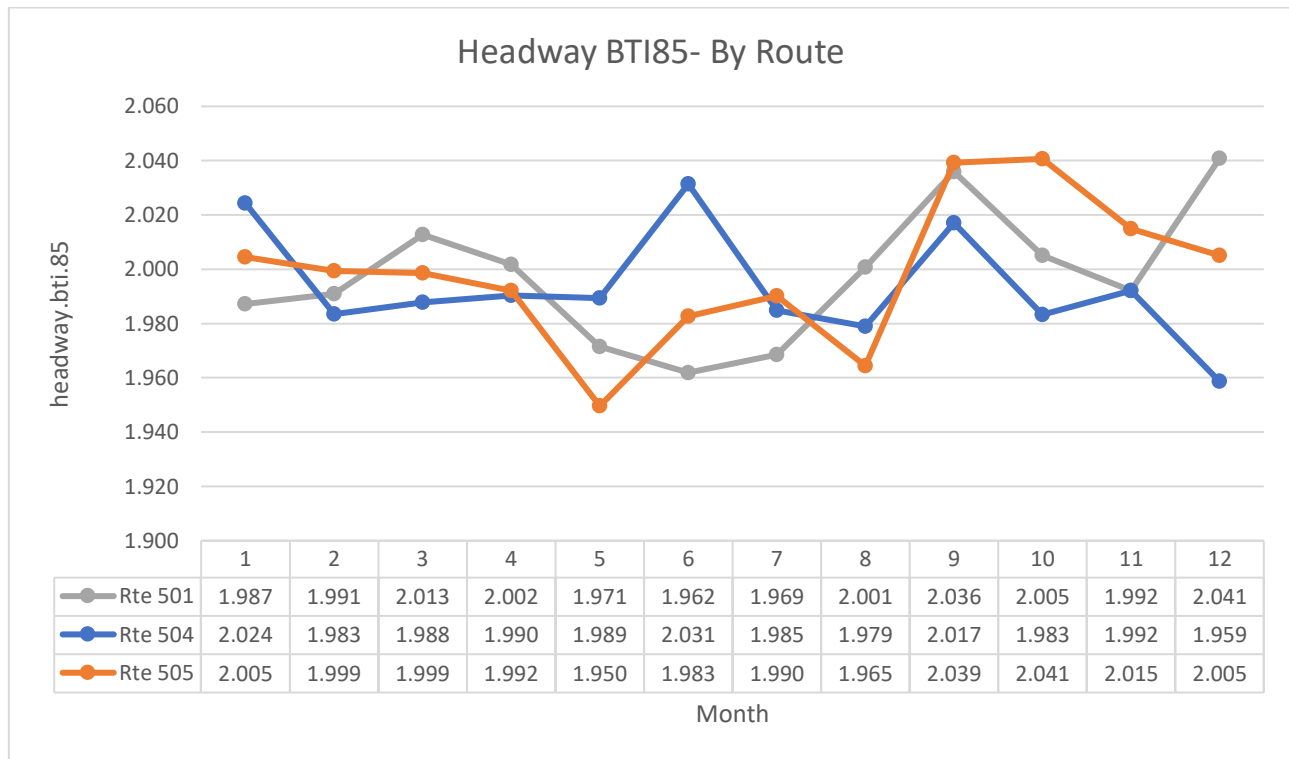
Figure 8: Headway BTI85- By Route

The mean of the Headway BTI85 measure is approximately 1.99 for all three routes. This shows that 85 percent of the time, wait times between two arrivals are lower than twice the expected average headway.

By using mean headway as a denominator, this measure controls for times during which transit services may have been diverted from the normal route, resulting in a lower observed mean headway than reality. The plots suggest that seasonal variation is generally consistent between all three routes. The summer months tend to fare better than the winter months, with an apparent peak in headway unreliability during the month of September. Many factors may be behind this trend, but it is conceivable that weather and ridership may both be contributors to these seasonal trends. Based on the figures from 2009 to 2015, the TTC typically sees the highest systemwide ridership in the month of September (Toronto Transit Commission, 2016).

Distribution of Headways- Route 501

	1	2	3	4	5	6	7	8	9	10	11	12
85Percentile	9.1	9.0	9.5	10.3	10.4	10.0	9.6	10.1	10.3	10.8	9.6	10.2
Mean	4.7	4.6	4.8	5.2	5.5	5.2	5.0	5.2	5.2	5.6	4.9	5.1
15Percentile	1.2	1.2	1.1	1.4	1.4	1.4	1.4	1.5	1.2	1.6	1.2	1.2
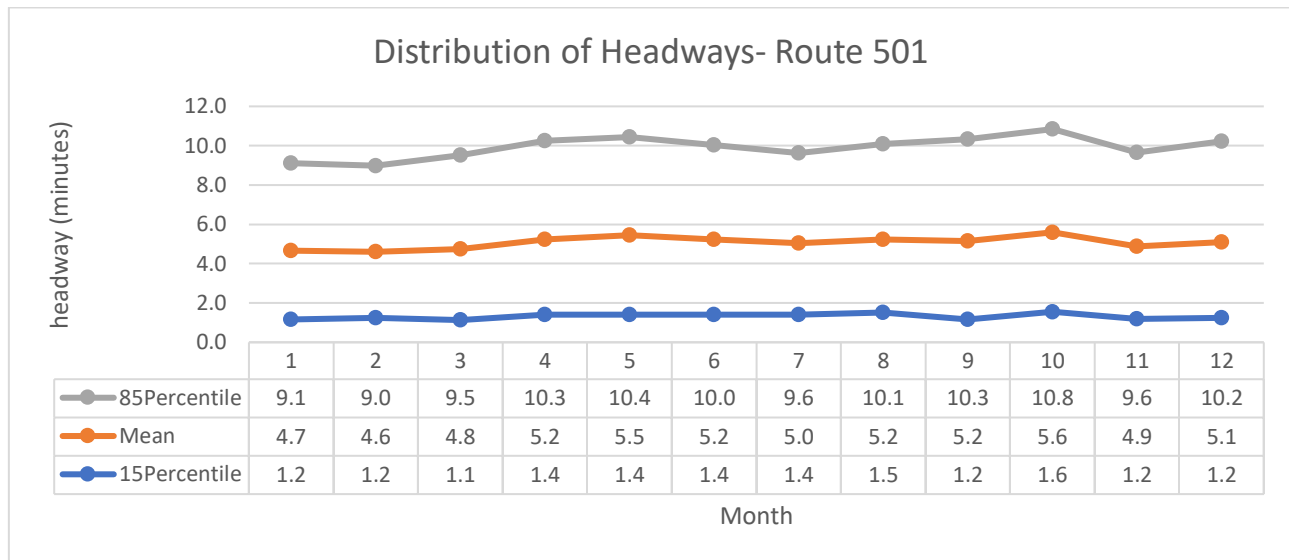
Figure 9: Distribution of Headways- Route 501 (Queen St).

Distribution of Headways- Route 504

	1	2	3	4	5	6	7	8	9	10	11	12
85Percentile	6.2	5.6	6.0	5.8	5.6	6.0	6.4	6.6	5.7	5.6	5.5	5.5
Mean	3.1	2.8	3.0	3.0	2.8	3.0	3.3	3.4	2.9	2.8	2.8	2.8
15Percentile	0.6	0.5	0.5	0.6	0.5	0.6	0.6	0.8	0.6	0.4	0.5	0.4
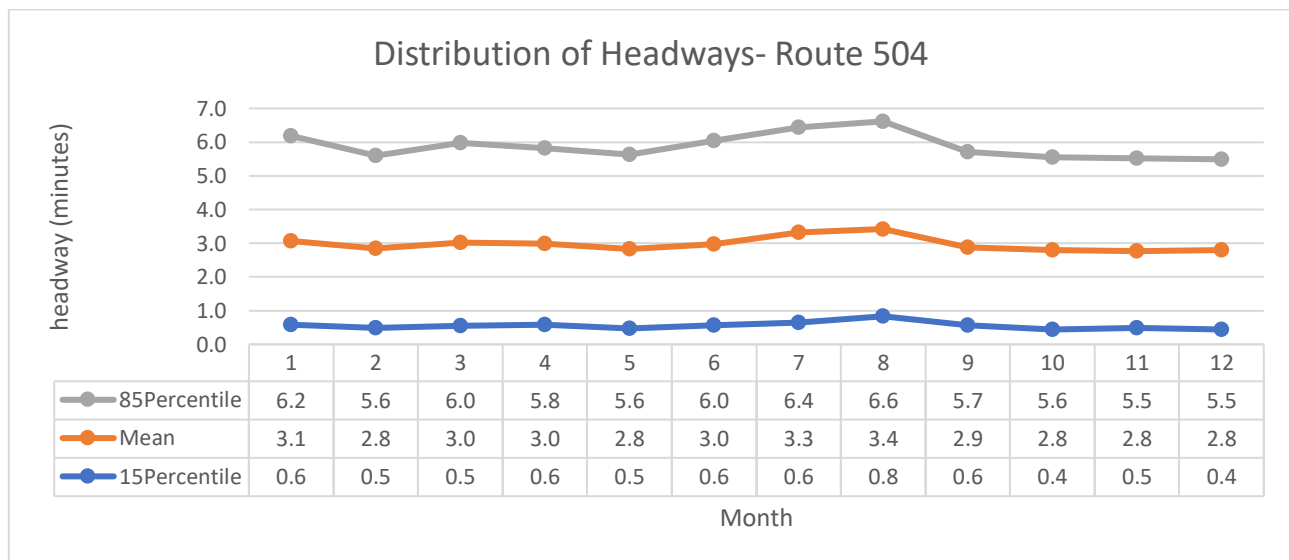
Figure 10: Distribution of Headways- Route 504 (King St.)

23

Figure 11: Distribution of Headways- Route 505 (Dundas St.)

85Percentile	12.2	11.6	10.8	11.5	11.4	11.2	11.8	11.8	13.5	13.4	12.8	12.2
Mean	6.2	5.9	5.5	5.8	6.0	5.7	6.1	6.1	6.8	6.7	6.5	6.2
15Percentile	1.4	1.5	1.4	1.4	1.6	1.4	1.6	1.6	1.4	1.5	1.5	1.4
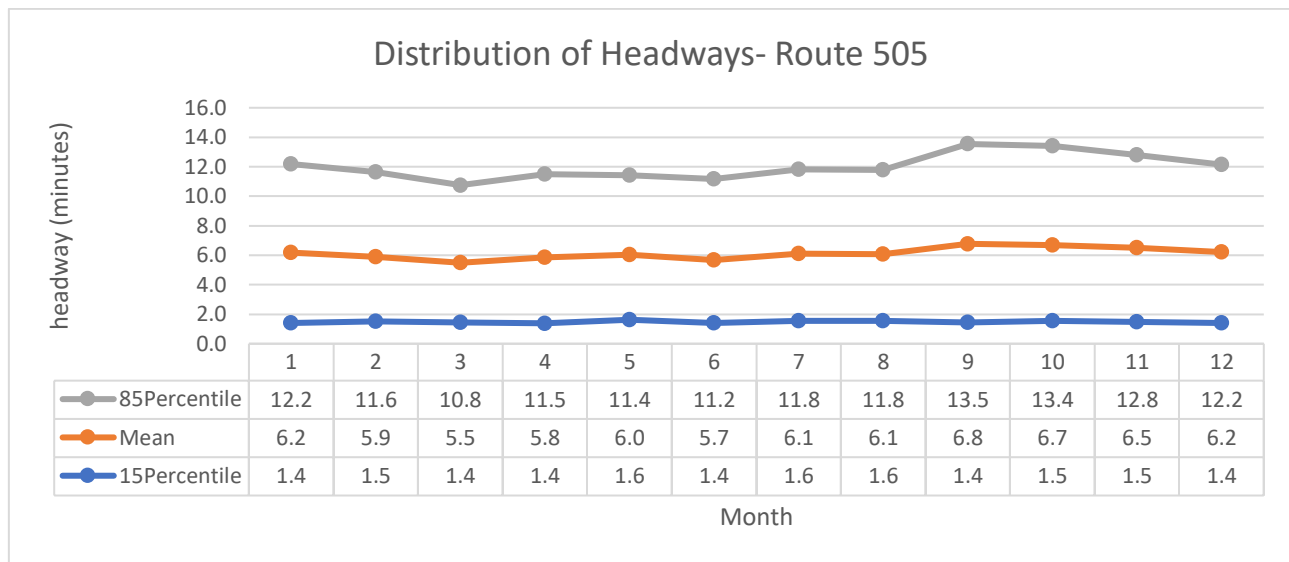
Figure 11: Distribution of Headways- Route 505 (Dundas St.)

Figures 2, 3, and 4 illustrate the distribution of headways for the three routes based on the TTC CIS data analysis. As expected, the data shows that Route 501 has a mean headway of approximately 5 minutes during the 4 pm to 6 pm hours, just as indicated in the TTC's schedule. Route 504, as the TTC's surface route with the highest ridership, shows a mean headway of the 3 minutes, while Route 505 shows a mean headway of 6 minutes. For all routes, the 85th percentile headways appear to have more variation than the mean, potentially suggesting some months are skewed by days with very poor reliability.

5.1.1.3 Transit Headway Reliability- by Year

Headway Reliablity- By Year- All Routes

headway.bti.85	1	2	3	4	5	6	7	8	9	10	11	12
2014 BTI85	2.002	1.975	2.011	1.989	1.950	1.974	1.966	1.992	2.031	2.006	1.987	1.995
2015 BTI85	1.992	2.018	2.014	2.028	2.008	1.989	1.992	1.977	2.033	2.013	2.010	2.020
2016 BTI85	2.014	1.982	1.980	1.970	1.951	1.998						

Month

Figure 12: Headway Reliability- By Year

Aggregating the headway reliability, as measured by "Headway Buffer Time Index 85" reveals differences in performance between the years, but no clear pattern as to which year had better performance.

5.1.1.4 Traffic Congestion

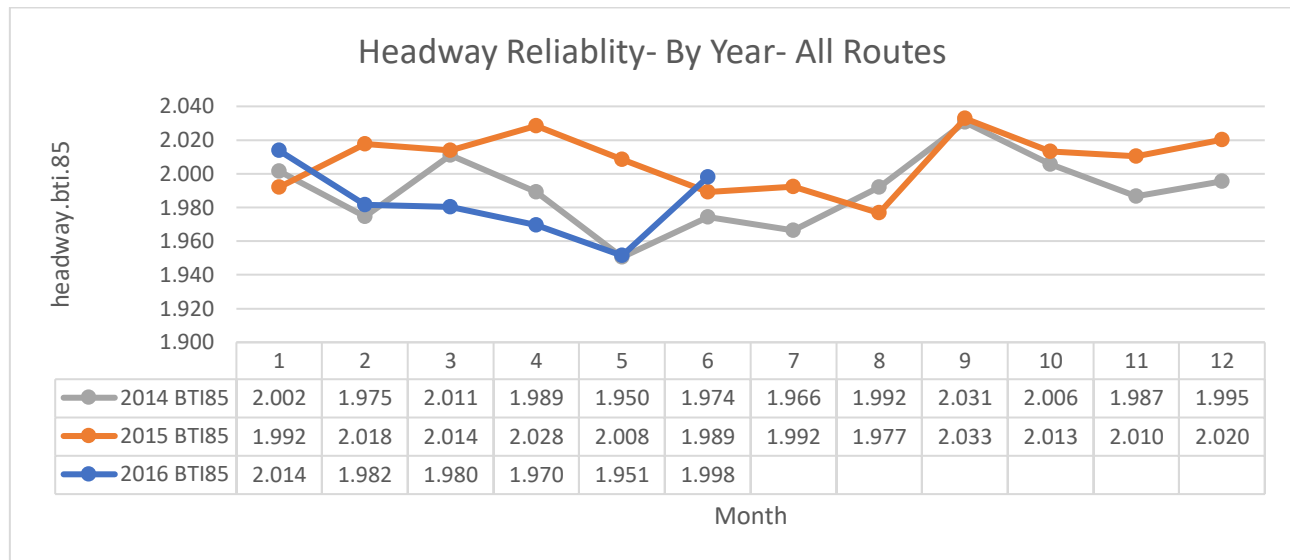
Traffic Reliability- By Route

	1	2	3	4	5	6	7	8	9	10	11	12
Rte 501	1.28	1.28	1.26	1.31	1.29	1.31	1.28	1.27	1.37	1.40	1.36	1.37
Rte 504	1.28	1.26	1.25	1.28	1.29	1.31	1.27	1.26	1.35	1.38	1.33	1.36
Rte 505	1.25	1.22	1.23	1.26	1.25	1.29	1.22	1.22	1.32	1.36	1.29	1.32
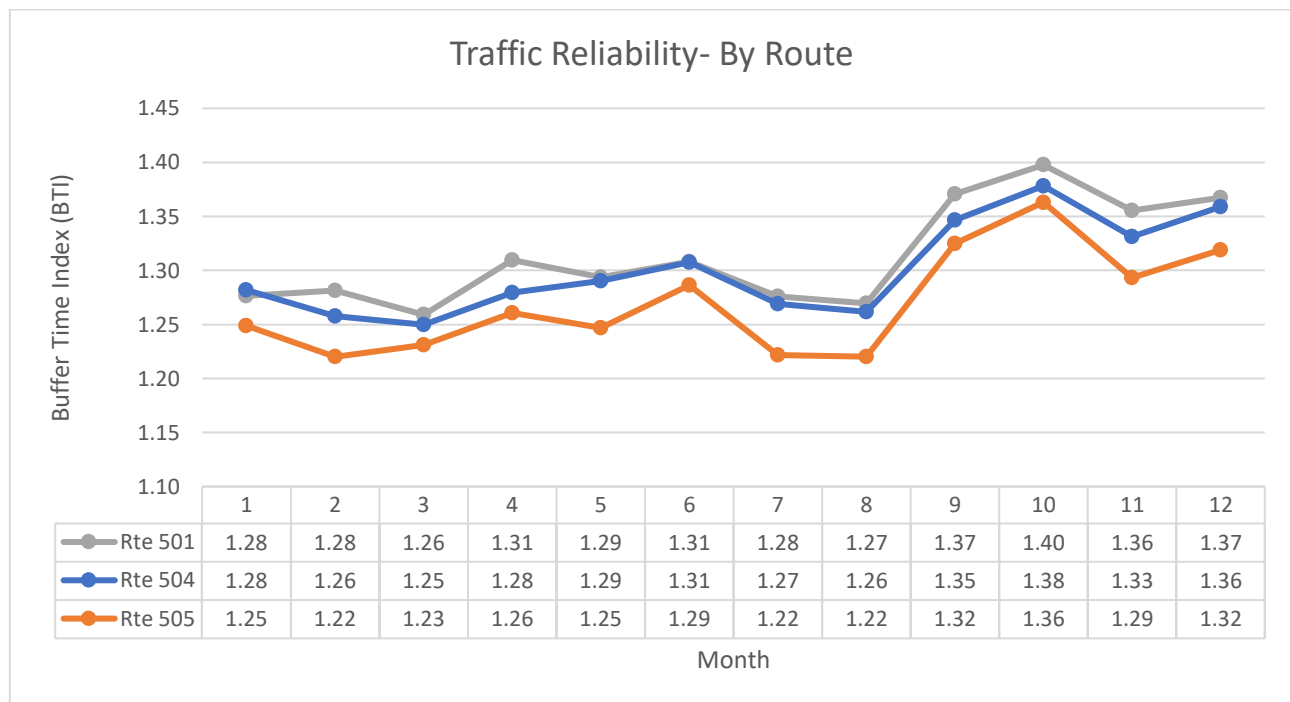
Figure 13: Traffic congestion along transit routes between 16:00 and 18:00 in year 2014, 2015, and 2016.

As explained previously, Buffer Time Index (BTI) is a measure of extreme events in traffic congestion. By isolating the road segments that each transit route runs on, it is apparent that traffic congestion along each route follows a similar seasonal pattern along the King, Queen, and Dundas street transit corridors. Congestion is shown to be lower during the summer months, and appear to peak during the fall.
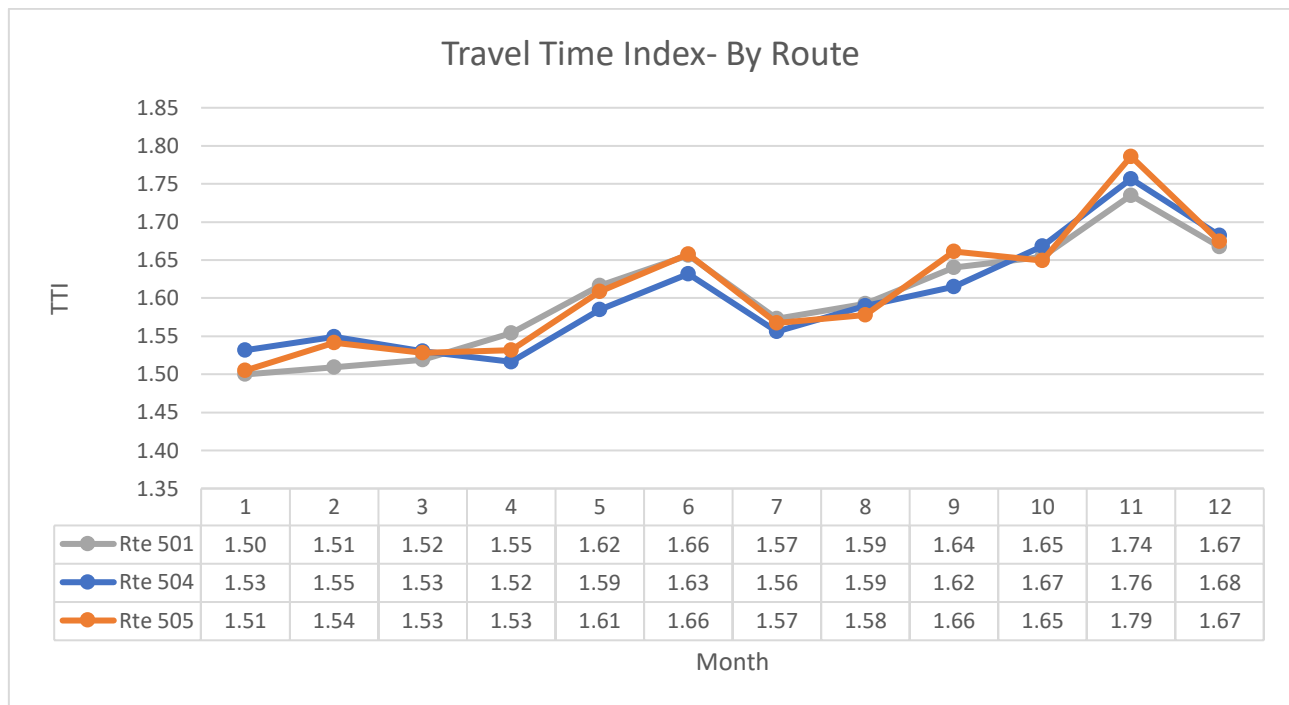
Figure 14: Travel Time Index for mixed traffic along transit corridors between 16:00 and 18:00 in year 2014, 2015, and 2016

Based on the Travel Time Index (TTI) measure of typical congestion, Dundas St, Queen St, and King St all appear to follow a similar seasonal pattern, with PM peak travel times between 1.5 and 1.8 times as long as during free flow conditions.

5.1.2 Spatial Distribution for Transit Unreliability

Each route has a unique spatial profile for reliability, with some sections of every route exhibiting more variation than others. Although headway reliability appears show a slight positive relationship with the number of stops from the terminus, the trend is not clear, especially towards the end of each route. This pattern may explain why "Stop Order" is a poor predictor variable for headway reliability in the regression models in Section 5.2.
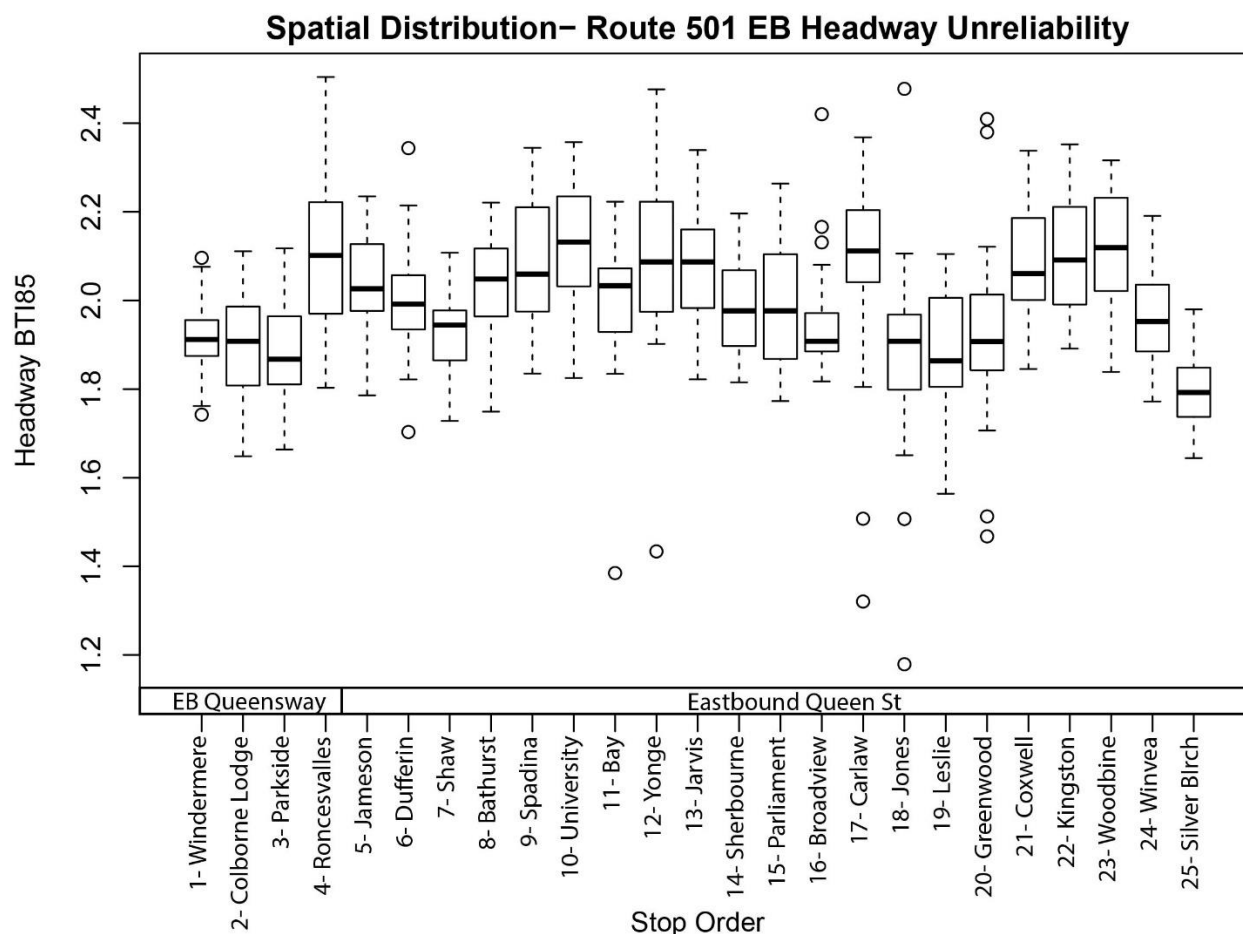
Figure 15: Headway Unreliability by Stop- Route 501 (Queen St.) Eastbound

Based on the Headway BTI85 reliability measure, it appears that the Route 501 Queen Street service has varying levels of reliability in the eastbound direction. A sudden increase in unreliability is apparent at Stop 4 (Eastbound Queensway at Roncesvalles Ave). This is the final stop before the route goes through a major intersection and transitions from dedicated right-of-way to mixed traffic. Another potential contributor to this sudden spike in measured unreliability may be that the TTC's Sunnyside operations and maintenance facility is located just before this stop. It is possible that additional transit vehicles are dispatched from this facility, or that some vehicles are short-turned at this stop. Headway regularity appears to recover between stop 4 and 7, before deteriorating again as the route enters the central business district. Conditions generally recover as

the route leaves the central business district until the route reaches Coxwell Ave, when it deteriorates once again. An isolated peak at Carlaw Ave exists with no obvious explanation.

Note that while they are included in the descriptive analysis, stops 1, 2, and 3 were removed from the regression analysis explained in Section 5.2 since they were situated in a section of the route with dedicated right of way (ROW B) and no associated Inrix traffic data.
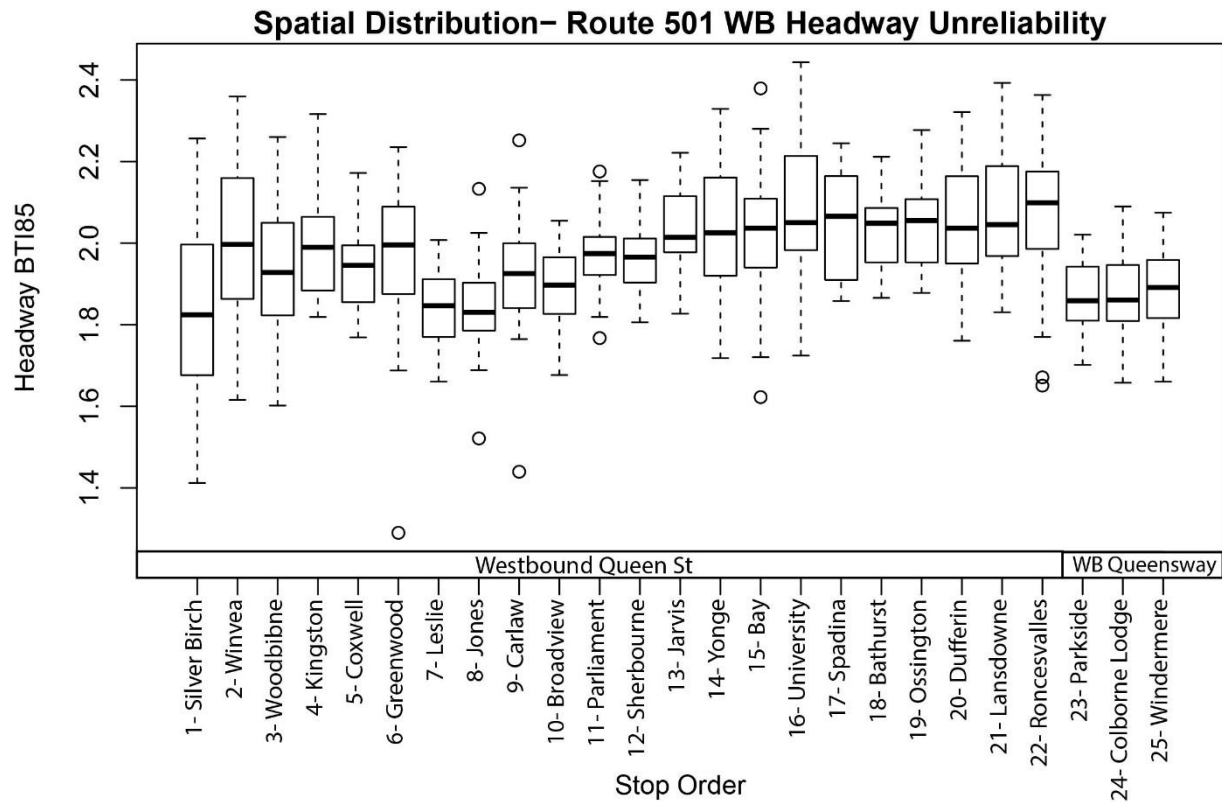
Figure 16: Headway Unreliability by Stop- Route 501 (Queen St.) Westbound

In the Westbound direction, the route appears to exhibit predictable gradual changes in unreliability. A "re-adjustment" in headway regularity is apparent at Leslie Street before gradually deteriorating as vehicles progress through downtown Toronto. A second "re-adjustment" is apparent at stop 23. At this location, the route changes from mixed-traffic to dedicated right-of-way and gores through a major intersection. As well, the Sunnyside operations and maintenance facility may be a location where some vehicles are turned around.
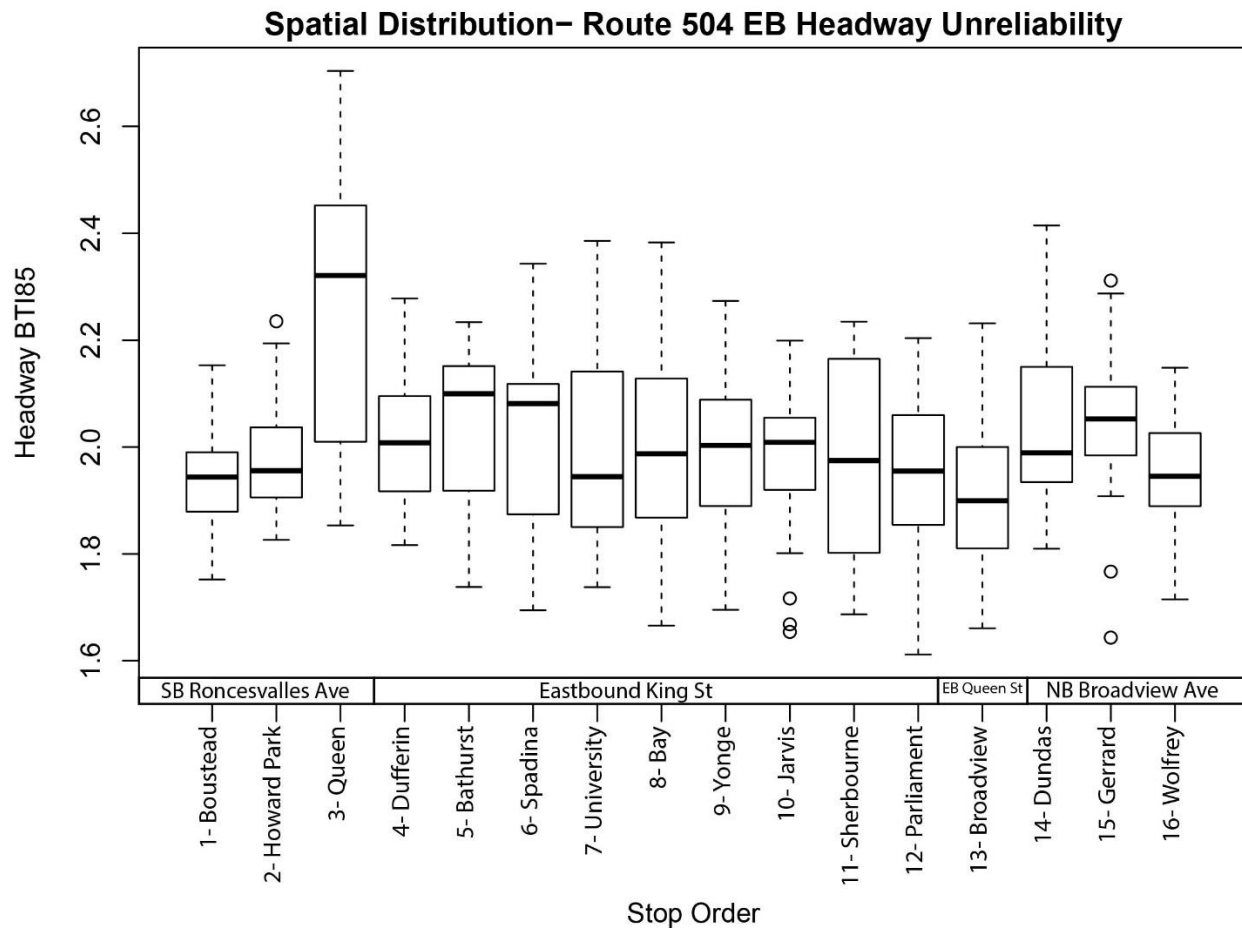
Figure 17: Headway Unreliability by Stop- Route 504 (King St) Eastbound

The route 504 King Street service has varying reliability in the eastbound direction. A spike in unreliability is visible at stop 3. The cause of this spike is not apparent as headway regularity appears to recover at stop 4. Potential causes of this abnormality include the change in travel direction and that stop 3 just prior to a major intersection.
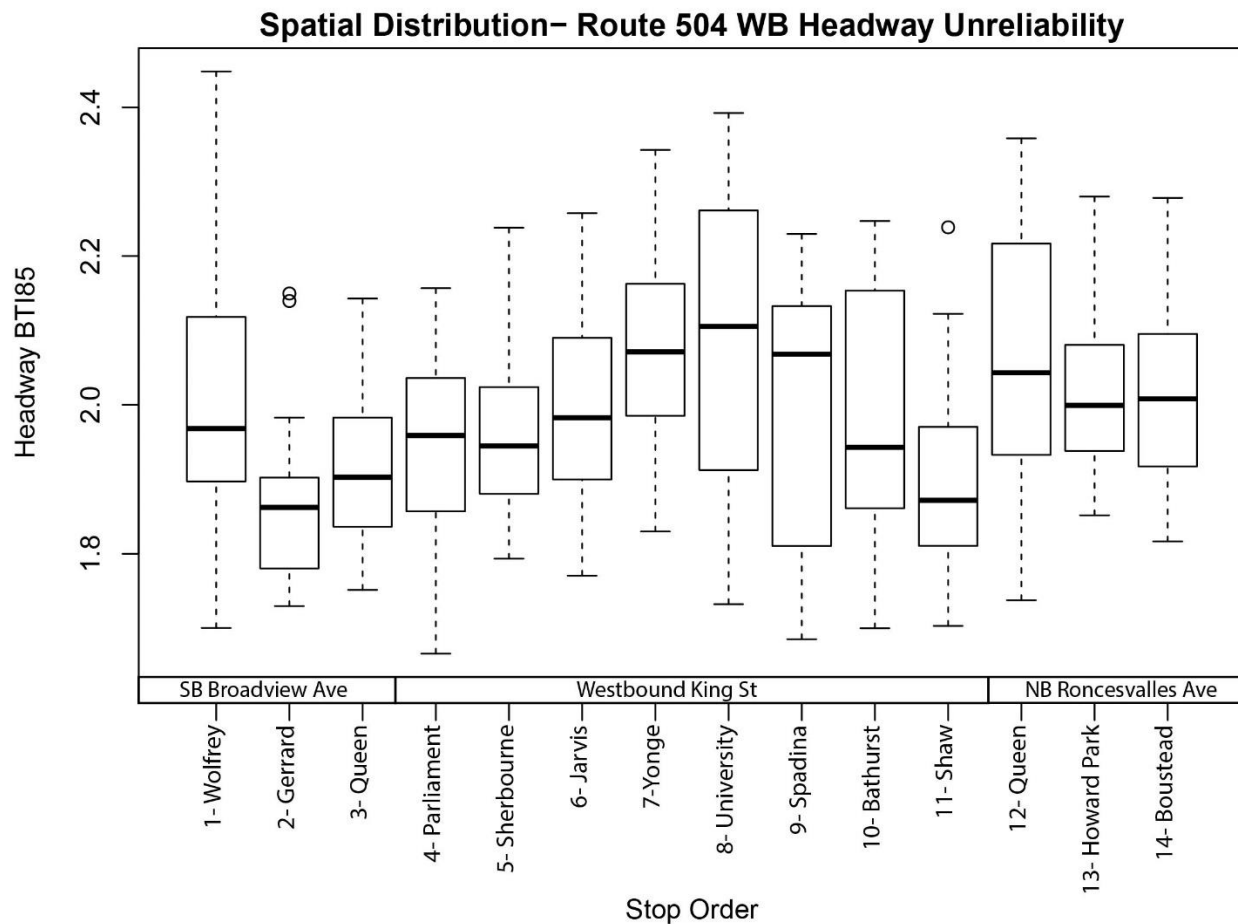
Figure 18: Headway Unreliability by Stop- Route 504 (King St) Westbound

In the Westbound direction, the plots show that unreliability generally increases as the route nears

the central business district. It peaks at stop 8- University Avenue, before appearing to recover.

Unreliability "re-adjusts" to a higher level at stop 12, when the route changes from a westbound to a

northbound direction.

Spatial Distribution– Route 505 EB Headway Unreliability
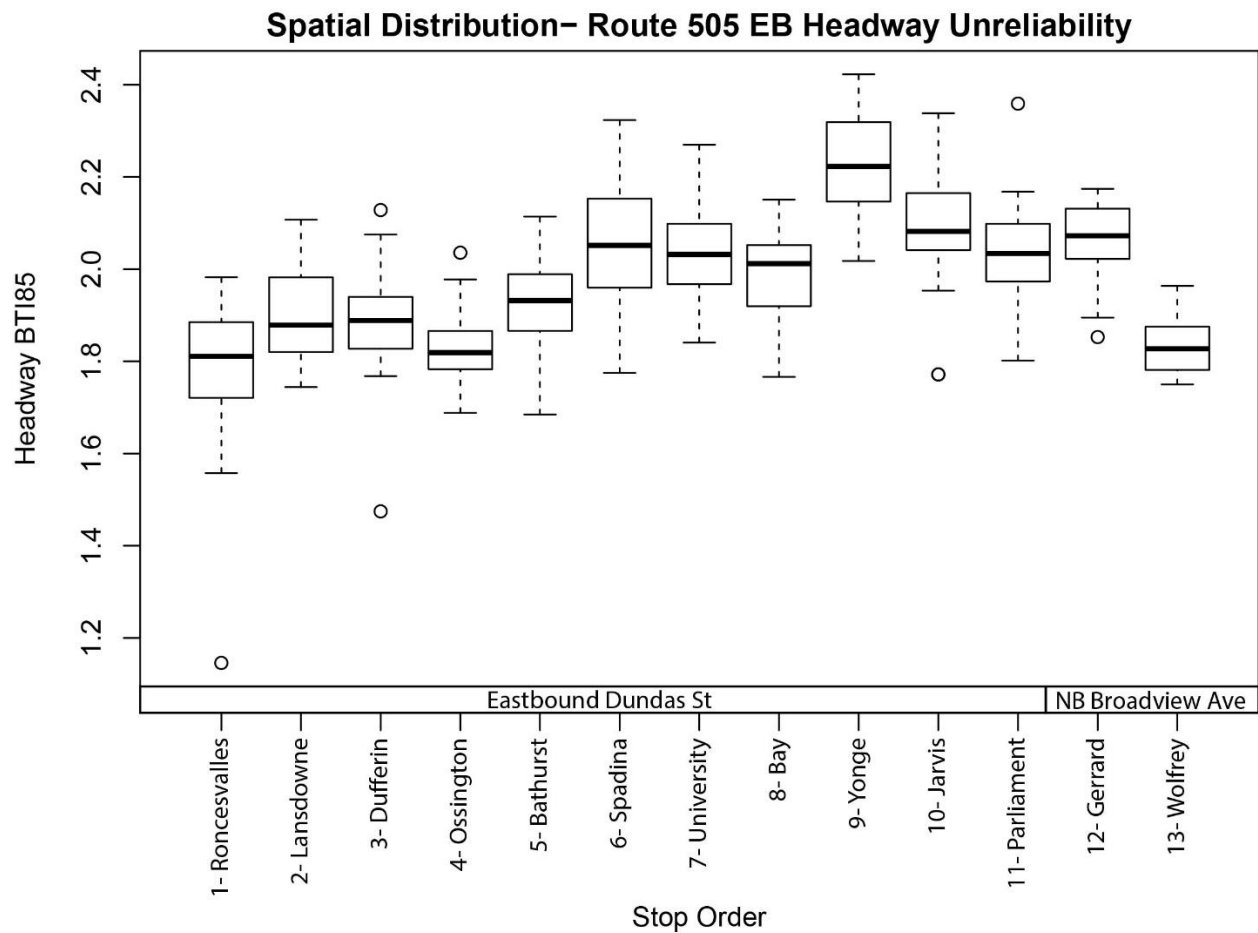
Figure 19: Headway Unreliability by Stop: Route 505 (Dundas St.) Eastbound

The route 505 Dundas Street service appears to become more unreliable as vehicles enter downtown. Headways recover after Yonge street.
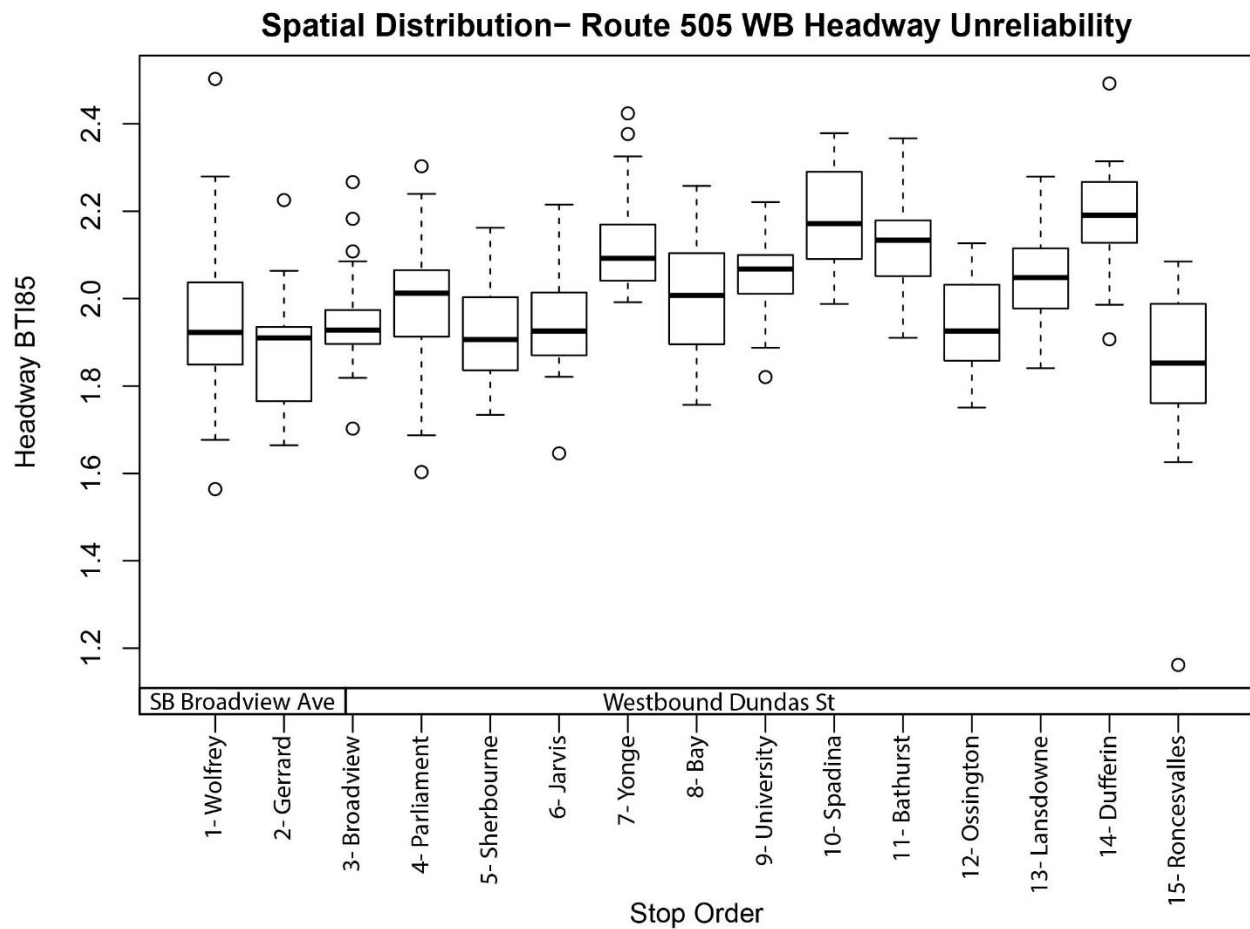
Figure 20: Headway Unreliability by Stop: Route 505 (Dundas St.) Eastbound

In the westbound direction, Route 505 seems to fluctuate in unreliability throughout the

route, with peaks at stop 10 (Spadina) and stop 14 (Dufferin).

5.2 Models of Headway Reliability

This study produced three models using Ordinary Least Squares regression, all of which predict

Transit Headway Buffer Time Index (BTI) 85 using variables related to transit operations and

traffic congestion variables. All three models control for year, and for route and direction. All

continuous variables except stop order, are natural logged. These models are presented below:

Dependent variable = Transit Headway Buffer Time Index (BTI) 85 (ln)

Variables	Model 1 coef.est	sig.	Model 2 coef.est	sig.	Model 3 coef.est	sig.
(Intercept)	0.781	***	0.709	***	0.702	***
Headway Mean (ln)	-0.055	***	-0.051	***	-0.050	***
Transit Operations						
Stop Order*	-0.001	**				
Headway BTI 85 (ln) at beginning of route			0.091	***	0.095	***
Traffic Congestion						
Buffer Time Index (ln)	0.072	***	0.065	***		
Travel Time Index (ln)					0.051	***
Dummy Variables						
Year 2014	*control*		*control*		*control*	
Year 2015	0.004		0.003		-0.006	
Year 2016	-0.010	*	-0.007	.	-0.015	**
Route 501 Eastbound	-0.030	***	-0.040	***	-0.038	***
Route 501 Westbound	-0.021	***	-0.018	***	-0.017	***
Route 504 Eastbound	-0.055	***	-0.053	***	-0.051	***
Route 504 Westbound	-0.052	***	-0.051	***	-0.050	***
Route 505 Eastbound	-0.021	***	-0.011	.	-0.012	*
Route 505 Westbound	*control*		*control*		*control*	
Adjusted R-squared:	0.118		0.125		0.127	
Degrees of freedom	3015		2990		2990	

Notes: * Statistical significance at the p = 0.10 level; ** statistical significance at the p = 0.05 level; *** Statistical significance at the p = 0.01 level.

5.2.1 Model 1

Model 1 predicts transit headway reliability using Stop Order and Traffic Buffer Time Index. While both variables are statistically significant factors, the impact of Stop Order is extremely small and opposite of what was expected. Using Traffic Buffer Time Index as a representation of the extreme values within each aggregated month, this model suggests a 7.2% decrease in headway unreliability can be expected for every unit decrease in traffic congestion. Despite this, the Adjusted R-squared for this model is only 0.118, indicating that the vast majority of variation in headway unreliability is not explained with these variables.

5.2.2 Model 2

Model 2 removes the stop order variable due to its low explanatory power and instead, introduces Transit Headway BTI85 from at the start of every route as a new variable. Since the dependent variable would be the same as the dependent variable at stop 1 for each route in each direction, all stop 1 records were removed from this analysis, resulting in a reduced sample size. In this model, both variables are statistically significant predictors of headway reliability. This model suggests a 9.1% increase in headway unreliability throughout each route can be expected by doubling unreliability at the start of the route. A 6.5% increase in headway unreliability can be expected for every doubling in Traffic Buffer Time Index. The Adjusted R-squared for this model is 0.125- higher than Model 1, but still leaving most of variation in headway unreliability unexplained.

5.2.3 Model 3

Model 3 replaces Traffic Buffer Time Index with Travel Time Index and therefore depends on average peak hour levels of traffic congestion instead of extreme values that may be experienced only several days a month. This model suggests that a 9.5% increase in headway unreliability throughout each route can be expected with every unit increase in unreliability at the start of the route. A 5.1% increase in headway unreliability can be expected for every unit increase in traffic congestion as measured by Travel Time Index. The Adjusted R-squared for this model is 0.127- the highest of all three models, but still leaving most of variation in headway unreliability unexplained.

6.0 Discussion

This study's descriptive analysis shows that some patterns in mixed-traffic transit reliability do exist in downtown Toronto. Seasonal patterns are apparent, with the spring months being the most reliable, and the month of September being the least reliable. Similarly, traffic congestion along these corridors also follows a seasonal trend peaking in the fall months. Reliability along each route fluctuates, showing localized spatial patterns. Changes in reliability along each route generally progress gradually, although exceptions exist.

Models show that despite estimated links between several of the model variables, approximately 88% of variation in streetcar headway unreliability remains unexplained. Given the low explanatory power of these models, it may be prudent to further explored the results using alternative datasets and other contexts in the future. While this study suggests that congestion influences transit service reliability, without knowledge of the extent to which the balance of variation in unreliability can be explained by other factors, these results should be interpreted cautiously.

7.0 Conclusion

This study illustrates some ways in which big data can be leveraged to quantify the linkages between related transportation performance outcomes. Analyzing roadway congestion and transit performance data, this study produced both descriptive statistics, and regression models. Conclusions from this study address two issues: 1) the complexity of downtown transportation, and 2) how transportation big data may improve data-driven decision making in the future.

7.1 Substantive Findings and Implications

Based on two-hour PM peak data aggregated by month, traffic congestion is positively associated with transit headway unreliability, with elasticities between 0.051 to 0.072 depending on the specific model and measure used. The irregularity of transit departures from the originating terminus is also positively linked to transit headway unreliability, with elasticities between 0.091 and 0.095. Under the specific spatial and temporal limitations of this study, these two factors only explain a small portion of the total variation in transit reliability.

Results indicate that there is no panacea for improving transit headway reliability and that the contributors to headway reliability are complex in urban environments. As such, the results of this study serve as a caution for planning practitioners to temper their expectations for mixed-traffic transit. The lack of predictability for headway reliability also suggests that cities should avoid overengineering their mixed-traffic corridors with the expectation of both reduced traffic congestion and improved transit reliability.

However, this study shows that there is an outlook, albeit limited, for improving transit through transit operations and through managing traffic congestion. Adopting scheduling practices-such as incorporating more recovery time at terminus stations, that allow for more regular departures at the start of each route, can be expected to lead to reduced unreliability throughout the

entire route. Insofar that the cities can influence traffic congestion through expedited incidence response, or measures such as traffic signal optimization and on-street parking policies, reducing congestion can also be expected reduce transit unreliability.

Extrapolating further, under the assumption that dedicated right-of-way transit functions like transit on a non-congested arterial road, this study also suggests that projects such as the King Street dedicated transit right-of-way pilot are likely to see headway reliability benefits from shielding transit vehicles from traffic congestion. It should be noted that beyond headway unreliability, existing studies demonstrate poor peak hour transit performance in other ways. With the current mixed-traffic design, typical streetcar speeds in the downtown portion of the 504 King route often fall below the brisk walking speed of 8km/h during the PM peak (City of Toronto, 2017).

7.2 Big Data Opportunities for Transportation Planning

As the quality and quantity of transportation data available to cities continues to grow, it is likely that cities will seek to take a more data driven approach in decision making. As this research demonstrates, the linkages between different transportation outcomes may not always be clear cut and easy to interpret. Moreover, a dependence solely on big data may bring the risk of de-prioritizing qualitative city planning outcomes such as those related to pedestrian experience and quality of life.

However, despite the inherent complexity of urban transportation that this study illustrates, it is clear that cities will have take a proactive approach in managing change. Given the 16.2% population growth rate over five years in downtown Toronto, four times the rate of the rest of the city (City of Toronto, 2017), targeted policies should continue to be explored to manage demand pressures and allow Toronto to meet its liveability, environmental and mobility goals set out it Official Plan. Big Data offers promise for cities as a part of the decision-making process to help

link policy to outcome. For example, pilot projects such as the King Street transit priority corridor may be expeditiously implemented to demonstrate effectiveness and to gauge public support. Big Data can provide the possibility to evaluate certain quantitative transportation outcomes across the transportation network that the traditional four-step travel forecasting models may not capture.

A final observation from the process of conducting this study is that institutional barriers present limitations to the use of Big Data in planning policy. While the Toronto Transit Commission is a wholly owned arm of the City of Toronto, and while this research was funded by the City's Transportation Services Division, acquiring the TTC's transit location data proved to be more challenging than expected. In the end, for this data to be provided by TTC, senior leadership involvement was required. Furthermore, although the TTC has developed tools to further process the CIS transit location data into a more useable format which this project would have greatly benefited from, this post-processed data was unavailable for this research due to policies that restricted sharing of data developed for internal TTC use. Institutional barriers like these are likely not limited to the City of Toronto and the TTC. To realize the full potential of Big Data for urban planning, increased clarity in data sharing agreements ought to be established between government institutions, particularly within organizations which are part of the same public entity.

7.4 Recommendations

The conclusions of this study suggest several actions that the City of Toronto and the Toronto Transit Commission should consider. The general unpredictability of headway deviation observed in this study suggests that drastic measures may be necessary to have make a significant impact. To improve the speed and reliability of transit services, the city should continue to pursue measures to protect critical transit corridors from the impacts of congestion. This can be achieved through implementing a dedicated transit right-of-way on these streets, or through managing

upstream traffic volumes through signalization or bottlenecks. City-wide demand-management measures such as road pricing may also be effective for lessening the negative impacts of congestion on surface-running transit.

Given the measurable relationship between terminus departure regularity and reliability throughout the line, the Toronto Transit Commission should continue to explore operations management practices for improved reliability at the start of each route. Measures may include scheduling more recovery time at the terminus or staging spare vehicles at strategic locations. As this is an general operational recommendation with financial impacts for the transit operator, the TTC should make decisions regarding this recommendation based on cost-benefit analysis.

To address the institutional barriers that hinder the potential of big-data in municipal decision making, the City of Toronto should establish a data-sharing agreement between all its departments and subsidiaries such as the TTC. Such an agreement would enhance the ability for departments such as Transportation Services to build off of the TTC's operational expertise and experience with CIS data.

A final recommendation is for this study to be repeated with new data once the TTC's new Computer Assisted Dispatch / Automatic Vehicle Location (CAD/AVL) system is in place and producing more accurate time-at-location data. This, combined with more fine-grain traffic speed and congestion data may help pinpoint specific roadway links that are problematic and require special attention.

7.3 Funding

Appendix A: Table of Summary Statistics by Route

Route 501 Summary of Descriptive Statistics

Statistic	N	Mean	St. Dev	Min	Max
transit headway mean	1,320	5.1	5.5	1.2	50.3
transit headway- 85th percentile	1,320	9.9	10.3	2.0	92.5
transit headway- 50th percentile	1,320	3.4	4.7	0.0	48.0
transit headway bti85	1,320	2.0	0.2	1.2	2.5
traffic mean speed	1,320	13.2	3.0	6.4	19.8
traffic speed- 5th percentile	1,320	10.3	2.9	3.9	18.0
traffic speed- 50th percentile	1,320	13.0	3.1	6.0	20.0
traffic buffer time index	1,320	1.3	0.1	1.1	2.0
traffic free flow speed	1,320	20.2	2.4	16.5	24.0
traffic travel time index	1,320	1.6	0.3	1.1	2.9

Route 504 Summary of Descriptive Statistics

Statistic	N	Mean	St. Dev	Min	Max
transit headway mean	838	3	2.6	0.9	22.5
transit headway- 85th percentile	838	5.9	5	2	48
transit headway- 50th percentile	838	1.8	1.7	0	18
transit headway bti85	838	2	0.2	1.6	2.7
traffic mean speed	838	12.9	2.9	7.7	20.8
traffic speed- 5th percentile	838	10.1	2.7	5.4	18.7
traffic speed- 50th percentile	838	12.7	2.9	7	20.5
traffic buffer time index	838	1.3	0.1	1.1	1.8
traffic free flow speed	838	19.9	2.5	16	23
traffic travel time index	838	1.6	0.3	1.1	2.2

Route 505 (Dundas) Summary of Descriptive Statistics

Statistic	N	Mean	St. Dev	Min	Max
transit headway mean	868	6.1	3.8	2.8	31.8
transit headway- 85th percentile	868	11.9	7.1	6	53.6
transit headway- 50th percentile	868	4.1	2.9	1	32
transit headway bti85	868	2	0.2	1.1	2.5
traffic mean speed	868	12.5	2.9	6.3	20.3
traffic speed- 5th percentile	868	10	2.7	4.4	16.5
traffic speed- 50th percentile	868	12.4	3	6.2	20
traffic buffer time index	868	1.3	0.1	1.1	1.7
traffic free flow speed	868	19.2	2.2	16.4	24
traffic travel time index	868	1.6	0.3	1.1	2.6

References

Button, K. J. (2008). *Handbook of transport modelling (2nd ed.).* Elsvier.

City of Toronto. (2013). *Guidelines for the Preparation of Transportation Impact Studies.* Toronto: City of Toronto.

City of Toronto. (2015). *Official Plan.* Toronto: City of Toronto.

City of Toronto. (2016). *Congestion Management Plan (CMP) 2016-2020.* Toronto: City of Toronto.

City of Toronto. (2017, 02 13). *King Street Pilot Study- Projects & Studies | City of Toronto*. Retrieved from City of Toronto: http://www1.toronto.ca/wps/portal/contentonly?vgnextoid=4f149ca0cc3f8510VgnVCM100 00071d60f89RCRD

City of Toronto. (2017, 02 10). *TOcore - Planning Toronto's Downtown- City Planning | City of Toronto*. Retrieved from City of Toronto: http://www1.toronto.ca/wps/portal/contentonly?vgnextoid=222101f2e9745410VgnVCM100 00071d60f89RCRD

Cooper, S., & Gould, L. (1994). *Faster than Walking? Street Congestion.* New York: New York City Transit Authority, Operation Planning.

Eagle, N., & Greene, K. (2014). *Reality Mining: Using Big Data to Engineer a Better World.* Boston: MIT Press.

Feng, W., & Figliozzi, M. (2010). *Using Archived AVL/APC Bus Data to Identify Spatial-Temporal.* Portland, OR: Portland State University.

Furth, P., & Hemily, B. (2006). *Using Archived AVL-APC Data to Improve Transit Performance and Management.* Washington: Transportation Research Board.

Leary, R. (2015). *Performance Based Service.* Toronto: Toronto Transit Commission.

Mazloumi, E., Currie, G., & Rose, G. (2008). *Causes of travel time unreliability - a Melbourne case study.* Melbourne: Institute of Transport Studies, Monash University.

McKnight, C. E., Levinson, H. S., Ozbay, K., Kamga, C., & Paaswell, R. E. (2003). *Impact of Congestion on Bus Operations and Costs.* New York: University Transportation Research Centre.

Morang, M. (2017, 03 2017). *Yay, transit!* Retrieved from Using GTFS Data in ArcGIS Network Analyst: http://transit.melindamorang.com/

Pilachowski, J. M. (2009). *An Approach to Reuducing Bus Bunching.* Berkley, CA: University of California, Berkley.

Saberi, M., & Zockaie, A. K. (2013). Definition and Properties of Alternative Bus Service Reliability Measures at the Stop Level. *Journal of Public Transportation, 16*(1), 97-122.

Schrank, D., Eisele, B., Lomax, T., & Bak, J. (2015). *2015 Urban Mobility Scorecard.* Texas: Texas A&M Transportation Institute and INRIX.

Sweet, M., Harrison, C., & Kanaroglou, P. (2015). *Congestion Trends in the City of Toronto: 2011-2014.* Hamilton, ON: McMaster Institute for Transportation and Logistics.

Toronto Transit Commission. (2005). *Surface Transportation Route Management And Key Performance Measures.* Toronto, ON: Toronto Transit Commission.

Toronto Transit Commission. (2015). *Proposed Service Improvements for the 2016 Operating Budget.* Toronto, ON: Toronto Transit Commission.

Toronto Transit Commission. (2016, January 15). *Ridership numbers and Revenues Summary- Data Catalogue- Open Data | City of Toronto*. Retrieved from City of Toronto: http://www1.toronto.ca/wps/portal/contentonly?vgnextoid=a3b7c87477438310VgnVCM1000003dd60f89RCRD

Toronto Transit Commission. (2016, 12 20). *Schedules and Maps*. Retrieved from Toronto Transit Commission: https://www.ttc.ca/Routes/index.jsp

Toronto Transit Commission. (2016, December 31). *TTC Ridership- All Day Weekday for Surface Routes- Data Catalogue- Open Data | City of Toronto.* Retrieved from City of Toronto: http://www1.toronto.ca/wps/portal/contentonly?vgnextoid=34b6c316f16e8410VgnVCM10000071d60f89RCRD

Toronto Transit Commission. (2016, 11 15). *TTC Routes and Schedules- Data Catalogue- Open Data | City of Toronto.* Retrieved from City of Toronto: http://www1.toronto.ca/wps/portal/contentonly?vgnextoid=96f236899e02b210VgnVCM1000003dd60f89RCRD

Transportation Research Board. (1998). *Transit Scheduling: Basic and Advanced Manuals-Runcutting.* Washington: Transportation Research Board.

Transportation Research Board. (2013). *Transit Capacity and Quality of Service Manual, 3rd Edition.* Washington DC: Transportation Research Board.

Turner, S., & Eisle, W. (1998). *Travel Time Data Collection Handbook.* Texas: Texas Tranportation Institute.