2023-07

Level of Service Measures for an Urban Bus Route

Wimalasiri Devasurendra, Kaushan

UNIVERSITY OF CALGARY

Level of Service Measures for an Urban Bus Route

by

Kaushan Wimalasiri Devasurendra

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN CIVIL ENGINEERING

CALGARY, ALBERTA

JULY, 2023

Abstract

The ability to measure the level of the quality of transit service provided is of utmost importance for customers to assess the level of service they receive and for the transit agency to assess the effectiveness of the service improvements made. Despite its importance, the transportation industry lacks an efficient, widely accepted, and widely applicable overall level of service (LOS) measure. Specifically, one that can assess and compare the overall quality of service (QOS) of transit lines or systems or one that can compare different operational performances of the same transit line or system is needed.

The content of the thesis consists of four major parts. The first part critically reviews major domains of transit level of service (TLOS) measures in industry and academic literature. It focuses on the success in achieving anticipated goals as opposed to the requirement of such a measure. Existing measures fall short in incorporating a combined view of both the passenger and operator and in assessing the overall TLOS by a single measure. A new approach to evaluate TLOS is proposed that has the potential to address these drawbacks.

The second part of the thesis proposes a novel approach to measure the LOS with respect to the value of time (VoT) distribution of the passengers. An implied VoT representing the LOS of a particular attribute, a combination of attributes, or overall service is derived and is compared with the respective VoT distribution of the passengers to obtain the LOS. An approach to distinguish LOS grades depending on the standard deviation (SD) of the VoT distribution is proposed.

The third part of the thesis engages in developing three LOS measures representing five attributes of concern in the thesis. Accordingly, a measure to represent headway and crowding attributes, a measure to represent access and travel time attributes, and a measure to represent the reliability attribute are developed. Each measure represents an implied VoT figure obtained by simulating an existing operation using an analytical model of optimum operation related to the service attributes of concern. The analytical model of optimum operation is developed from the basics for reliability LOS measure, while for other measures, existing models in the literature are modified and used. Finally, the three measures developed are combined using a novel approach to represent the overall LOS of a bus route. The development of each LOS measure is accompanied by a numerical example explaining the calculation of the LOS of a bus route.

The fourth and final part of the thesis applies the developed measures to a bus route operation in Calgary. The data for the bus route is obtained from Calgary Transit for the year 2021. While each chapter discusses the derived LOS measure and draws conclusions, the final chapter provides insights into potential improvements to the suggested approaches and potential future research related to the developed work.

Acknowledgments

With a heart full of gratitude and a spirit of joy, I express my sincerest thanks to those who have helped me reach the pinnacle of my academic journey. First and foremost, I would like to extend my sincere gratitude to my supervisor, Dr. Chan Wirasinghe, for being a steadfast presence in my life for almost six years, supporting me through the ups and downs of my thesis and life. His exceptional guidance has been the beacon of light that has helped me navigate the murky waters of academia, and I am forever grateful for the trust he has placed in me.

To Dr. Lina Kattan, my co-supervisor, I owe a debt of gratitude for her unwavering support and guidance, which has been an invaluable asset to my research. I extend my heartfelt thanks to Dr. Saman Bandara for his exceptional support in securing an opportunity to study for a Ph.D. at the University of Calgary. I would also like to thank my supervisory committee members, Dr. Nigel Michael Waters and Dr. Saeid Saidi, for their dedication and unwavering commitment to improving my work.

To Calgary Transit, I am indebted for their support in providing me with the necessary data at short notice. My journey would not have been possible without the financial support and resources provided to me by the University of Calgary, Provincial Government of Alberta in Canada, Natural Sciences and Engineering Research Council (NSERC) of Canada, and the Queen Elizabeth Scholarship program.

I also wish to express my deepest gratitude to my wife, Upeksha Rathnavibushana, whose love, care, encouragement, and understanding have been my constant companions through the long and stressful six years of my Ph.D.

Lastly, I thank my parents, who have been my lifelong supporters and encouragers, and my sons Keon and Cayden, whose laughter and joy have been a constant reminder of what truly matters in life. Together, they have made this journey of mine a joyous one; for that, I am eternally grateful.

Dedication

To my dear wife, children, parents, family, and friends.

This thesis is dedicated for your unwavering love and loyalty.

Table of Content

List of Tables

List of Figures

Table of Acronyms

Acronym	Description
A&T LOS	Accessibility and Travel Time Level of Service
A&TT LOS	Accessibility and Travel Time Level of Service
ATTC	Total Expected Cost due to Additional (delayed) Travel Time
AWTC	Total Expected Cost due to Additional (delayed) Wait Time
CPF	Crowding Penalty Factor
DPC	Total Expected Delay Penalty Cost
DRPF	Delayed Riding Penalty Factor
DWPF	Delayed Waiting Penalty Factor
GHG	Green House Gas
H&C LOS	Headway and Crowding Level of Service
HCM	Highway Capacity Manual
HLOS	Headway Level of Service
ITS	Intelligent Transportation Systems
ITT	In-vehicle Travel Time
IVT	In Vehicle Time
LoC	Level of Crowding
LOS	Level of Service (Also the level of the quality of service)
LTA	Local Transit Accessibility
MAD	Mean Absolute Deviation
OC	Operator Cost
OTP	On Time Performance
PT	Public Transportation
QoS/QOS	Quality of Service
RP	Revealed Preference
RTTIS	Real Time Traveler Information Systems
SD	Standard Deviation
SP	Stated Preference
SQI	Service Quality Index

TCQMS	Transit Capacity and Quality of Service Manual
TCRP	Transit Cooperate Research Program
TLOS	Transit Level of Service
TRB	Transportation Research Board
TTTC	Total Expected Cost of Passenger's Budgeted Travel Time
VoAT	Value of Access Time
VoC	Value of Crowding
VoRT	Value of Ride Time
VoT	Value of Time
VoWT	Value of Wait Time
WHO	World Health Organization

1. Chapter 1: Background

1.1. Introduction

Access to basic human needs is a human right. Different needs being spatially apart, makes mobility and access an essential part of human lives. Public transportation is a way a community fulfills its own mobility needs for both choice and captive riders. Therefore, access to a public transport service can be identified as a right for choice riders and captive riders. The extent to which their needs and preferences are met reflects the 'quality' of life of a community (Flanagan 1982). How efficiently and effectively the need for mobility and accessibility is met through the public transportation service reflects its 'Quality of Service' (QOS). Therefore, an efficient transit system will satisfy the needs and preferences of a community to a greater extent promising a better quality of life. The role of a transit service provider is to efficiently meet the passenger's needs for mobility and accessibility to provide a higher quality of life for the community in part from public funds. As a result, improving the QOS for transit is one of the primary concerns of transit agencies.

Quality improvement in transit service also improves the quality of life of non-transit riders. Transit modal share increment as a result of transit service quality improvement (FitzRoy & Smith, 1998; Hensher, 1998; Manville & Cummins, 2015; Morpace International Inc & Cambridge Systematics Inc, 1999), reduces greenhouse gas (GHG) emission, road traffic crashes, fuel costs as well as road traffic congestion (Eboli & Mazzulla, 2007). While reducing GHG improves environmental quality, reducing traffic congestion also improves mobility and accessibility of non-transit riders - especially automobile users - through reduced travel times in the short-term future. Therefore, service quality improvement in transit benefits both transit and non-transit riders.

Hensher (2014) states that improvement of QOS results in reduced average cost per kilometer of bus service provision after controlling for fleet size, loading and speed. The primary reason is the induced modal shift towards transit as a result of higher QOS, which leads to a higher income per unit of service provision. This supports the finding of Newell (1979) that a higher QOS can be provided in a particular route when passenger demand for that service is higher. Therefore, public transit improvement imposes a positive economy of scale. This implies, the higher the QOS of transit, the lesser the generalized cost per trip made by transit, which leads to a higher efficiency in fulfilling needs of a community and hence a better quality of life. Therefore, improving public transit QOS benefits the whole community, including both the service provider and passenger.

To this end, this study formulates a measure to assess the overall QOS of an urban bus route, which is necessary for assessing and improving QOS. Unique QOS measures are derived for each or combinations of attributes affecting the overall QOS that includes both passenger and operator concerns. Individual QOS measures are then combined into one measure that represents the overall LOS of the route. What follows in this chapter is the context of the research problems and justification of the approach taken in this study to address the problems.

1.2. Problems of Concern

Incorporating the rider perception in a QOS measure is critical to ensure an improved QOS is sustained by increased ridership. On the contrary, the current practice in the transit industry, as described in transit manuals (Canadian Urban Transit Association., 1993; TCRP Project A-15C, 2013), does not seem to meet the requirement in two main ways. First, the measures described in manuals are based on expert opinion as opposed to passenger perceptions of service quality. In the same way, measures have also been developed to assess customer satisfaction where operator concerns have been disregarded. None of the measures therefore represent a combination of the views of both passengers and operators.

Second, the measures in current practice assess specific attributes of QOS separately rather than the service package that riders experience. What riders perceive as the QOS is a holistic view of the entire service they experience. An improvement in a single attribute of the transit service does not necessarily represent an equivalent improvement in the quality of the entire service provided. Therefore, an effort to evaluate the QOS by its constituent attributes separately compromises the necessity of such a measure. A nearly accurate estimate of the system-wide perspective of the QOS, with due attention to the characteristics of the unique ridership market, of a transit route or a system will satisfy this requirement.

Many researchers (Aronstein, 1976; Fu, 2007; Güner, 2018; Hensher et al., 2003; Prioni & Hensher, 2001) have attempted to tackle these problems completely or partly by developing QOS measures that incorporate several attributes of a transit service through customer satisfaction surveys which are time and money consuming. A major drawback of this approach is that such surveys need to be done pre- and post-interventions to determine the improvements made. This will consume a significant amount of agency funds. Although such initiatives have been able to tackle the drawbacks in current practice to a certain extent, there are still no measures that also incorporate transit agency concerns – mainly costs.

The term 'Level of Service' (LOS) is the most popular term in the field of transit in terms of measuring service quality (Higgins & Ryan, 1983). The level of service would ideally mean the level of amount of service that has been provided. For example, the level of frequency of service can be denoted as 4 to 6 buses per hour, or the level of service based on headway can be denoted as 10-15 minutes for the same LOS. The term LOS has initially been used as a measure of performance on quantifiable attributes of service by managers of transit to decide on disbursement of funds (Higgins & Ryan, 1983). As transit transitioned to more of a welfare service with time, passenger perception was also taken into account. Thus, qualitative attributes like comfort and convenience related factors were incorporated into the measures of service. Now, agencies are measuring the 'quality' of service (which includes direct performance as well) rather than performance and yet keep using the word LOS rather than QOS. Measures of QOS are continuous and are categorized into different levels for easy representation and use. Certain ranges of QOS of a single attribute of service are assigned to a certain level. This is evident from the way the 2nd edition of the Transit Capacity and Quality of Service Manual (TCQSM) has assigned letter grades from 'A' to 'F' to various ranges of services in different attributes. The term QOS can also be used to express an abstract idea of a certain aspect of the service or an idea of the whole service. An example is how the passengers perceive the quality of the service of a bus route rather than the headway of that bus route. In this study, we will refer to the level of the quality of entire service of a bus route or a system by the term 'LOS'. For the level of QOS of a certain attribute of service, for example headway, we will refer to it as 'Headway LOS'.

1.3. Literature review

1.3.1. History

From the 1950s onward the LOS concept has been widely utilized in the context of highway planning and design. Yet, by the end of the 1970s, there have not been well standardized measures for evaluating public transport services (Taylor & Brogan, 1978). As early as 1965, the LOS for highways had been defined in terms of the effect of several factors including speed, traffic interruptions, safety, comfort, freedom to maneuver, travel time, convenience and cost (Spring, 2000). The LOS methodologies for transit have been developed to help answer transit efficiency questions by transit authorities (Taylor & Brogan, 1978).

As public transportation changed from a private, profit oriented operation to a public service, profitability alone was no longer a viable measure for the evaluation of such a service and as a result, the LOS concepts were developed with due attention to the non-monetary values for the community that has compelled the inclusion of measures related to the service receiver (Taylor & Brogan, 1978). Earliest reported work attributes a transit LOS concept has been performed by the National Committee on Urban Transportation which recommends standards for routing, loading, frequency of service or headway, stop frequency, speed and regularity of service (National Committee on Urban Transportation, 1958).

Higgins & Ryan (1983), in their work, review the research attempts related to transit LOS and measuring QOS by that time. They show how it was stressed in the USA in 1976 and in Australia in 1980 where transit agencies were compelled to concentrate on more mundane aspects of public transport supply than being concerned with a LOS approach. Because in an LOS approach, user attitudes on scheduling, comfort and convenience are given comparatively significant attention in the evaluation process that determines the disbursement of funds. This has been a result of rising operating deficits where the transport authorities have been pushed to utilize more accounting of performance measures. It was further highlighted that central issues pertaining to the debate were: (1) LOS being a very complex matter in the broader sense where mutually opposing views of user and operator have to be considered together; (2) having no simple and useful measure on urban transit LOS as simpler measures have tendencies of being biased in their applications; (3) popular measures by that time (cost per passenger kilometer and deficit per passenger) not being able to measure benefits to the community that may be obtained from higher LOS standards. The study of Higgins & Ryan (1983) further argues that development of a composite index for transit LOS is worthless on the grounds that such an index could be subject to mutations (some components having a confounding effect on other components) and can be constrained by assumptions by the time it becomes of usable form (Higgins & Ryan, 1983). They therefore presented a comprehensive framework consisting of five LOS measures for three market segments.

Heighington and Brogan presented a composite set of guidelines that defines minimum, desirable and maximum levels for transit system characteristics based on several earlier individual attempts by transit operations/authorities (Heathington & Brogan, 1975). Their work is only comprised of two levels that state whether a transit system is adequate or not, thus depriving the authorities of assessing minor system improvements. Botzow developed a methodology capable of assessing the ability of a transit system to provide a comfortable ride and reasonable travel times and made it analogous to a highway LOS concept by introducing six distinct levels (A to F) of services (Botzow, 1974). Botzow's concept was simple to comprehend and had applications in both evaluation and planning where a LOS B has been used for planning of transit systems (Taylor & Brogan, 1978). Addressing the issue with the assumption of Botzow's concept where all users assign equal weights to each attribute, Aronstein develops a system standard accounting to performance measures as perceived by travelers that assigns different weights for different characteristics depending on their importance as determined by Aronstein (1976). Taylor and Brogan further examine the different factors constituting the LOS of Public Transportation (Taylor & Brogan, 1978).

Due to the fact that measuring QOS is of utmost importance for service providers, different transit authorities have developed their own measures over time (Higgins & Ryan, 1983; Taylor & Brogan, 1978). In the absence of a commonly accepted standard defining QOS in transit, the Transit Cooperative Research Program (TCRP) of the Transportation Research Board (TRB) developed the first edition of the TCQSM in 1999 which specified six designated levels of the QOS denoted by letters "A" (Best) to "F" (Worst) as it is easier to refer QOS with several levels rather than a continuous scale. TCQSM suggests a framework for assessing QOS that includes availability and quality aspects of transit. Here, TCQSM specifies different ranges of service quality for each of six sub aspects but does not specify an overall service quality for a transit system or a transit route (TCRP Project A-15C, 1999).

There are several studies that develop LOS measures for individual elements of transit using passenger perception acquired through stated preference or revealed preference surveys (dell'Olio et al., 2011; Prioni & Hensher, 2001). Some studies have also developed overall QOS measures using passenger survey data but output a continuous measure (Fu, 2007; Hensher et al., 2003). These measures are data and time intensive and hence expensive. Therefore, they are not viable to be estimated frequently for maintenance and quality adherence requirements. Having individual LOS grading for different service elements of a transit line or a system, however, compromises the requirement of having a LOS measure which is to assess the overall quality of the service. Such methods were mostly the result of transit agency efforts as it is easier to maintain and assess the service quality in individual service elements and help administrative decisions (Taylor & Brogan, 1978).

On the other hand, selective focus on individual elements of a service can result in incomplete evaluation conclusions where it is possible that better alternatives are overlooked (Hodge & Orell, 1995). Accordingly, the traditional transit LOS measures on service frequency, route density, access distance stress the same outcomes everywhere without regard to the actual market's (riders of a certain transit line or the system) specific service requirements (e.g., implementing a headway of 20 minutes assuming a LOS of 'C' as per the standards whereas the LOS perceived by the riders of that transit ride can be 'E' where 'E' is two grades inferior in quality than 'C'). Irrespective of the fact that many studies highlight the importance of passenger perception in evaluating transit LOS (Das & Pandit, 2013), almost all the LOS measures currently in practice, use expert opinion in evaluating transit LOS (Highway Research Board, 1950; TCRP Project A-15C, 2013). Universal application of measures developed based on expert opinion can impose perceptive differences from what passengers really feel which depends on geographical location, culture and income of passengers.

The TCQSM 3rd edition in 2013, introduced a multimodal LOS concept. It provides a single LOS letter grade for all the transport modes operating on a certain segment of the road which is subjective to the conditions of that particular road segment. Accordingly, this procedure provides a single LOS grade for transit only limited to a particular road segment. This LOS depends on transit operational characteristics, transit amenity conditions and pedestrian environment. The LOS of transit derived for one road segment can be different from that on the next road segment.

1.3.2. Factors affecting Transit LOS

1.3.2.1. Frequency/Headway and Wait Time

Frequency of the service is one of the main factors affecting transit LOS. Having a more frequent service implies potential transit users have access to transit more often and have a lower potential wait time at a bus stop. Therefore, service frequency induces ridership and hence improves QOS. From the transit agency point of view, service frequency is one of the main factors affecting operating costs and will incur a capital cost beyond a certain limit where they must buy new vehicles and expand existing fleet size to operate at new higher frequencies. Doubling the frequency can make the operating cost almost double with the exception of automated transit where drivers will no more be needed (TCRP Project A-15C, 2013). As the term headway is more tangible than frequency of service, most of the researchers that studied frequency have given attention mainly to 'average headway' for minimizing passenger wait time at station with the intention of increasing QOS. However, Hensher et al., (2010) moved beyond the average headway and studied service headways in a more disaggregated level by considering headways in different time periods such as evening, day and weekends. Studying frequency at different time periods will offer the opportunity to better evaluate service performance (Hensher et al., 2010). Pushkarev et al., (1977) suggests a different approach for headways which can be used in the land use design to achieve a higher QOS on public transit by recommending ranges of headway levels for different dwelling unit densities (e.g., number of dwelling units/acre) (Pushkarev & Zupan, 1977).

1.3.2.2. Service Span

Hours of service (Service Span) is another factor affecting Transit LOS. Frequency or headway alone does not provide the possibility for a potential transit user to make a trip unless the service is provided at the time of day where the user needs to make the trip. A longer service span than usually needed provides more flexibility for a potential user and therefore improves the confidence of users in choosing transit. This helps expand the potential market group that transit can serve. The main negative aspect of utilizing a longer service span is obviously the higher operating costs. If all other conditions are constant, increasing service hours by 20% is estimated to increase operating cost by 20% (TCRP Project A-15C, 2013) and increasing the service hours beyond the normal working hours makes the situation even worse for obvious reasons: transit demand declines with time and operator's pay is higher for overtime. Therefore, a critical trade-off between passenger and operator perspectives needs to be achieved in improving LOS.

1.3.2.3. Access

Apart from transit headway and service span, access to transit also determines the choice of a potential user to use transit for their trip or not. Although there are several modes possible to be used in accessing transit, walking is the dominant way in the urban context. Service coverage (area where transit service is provided), therefore, has been identified as the area located within an acceptable walking distance from a transit service. Route miles per square mile (Route density), percentage system area (population) served (Geographic coverage) and transit supportive area served (transit market coverage) are some of the measures employed and described in TCQSM (3rd edition) as potential measures of service coverage. A methodology to calculate the transit system coverage area is suggested in detail which utilizes a concept called “Transit stop service radius”. The ideal transit stop service distance is identified as 400 m factored for effects of street connectivity, grade of the street, population and pedestrian crossings. It is also important to note that weather, temperature, individual characteristics, and station characteristics can affect the walking distance to a bus stop (El-Geneidy et al., 2014).

1.3.2.4. Passenger Load

Bunker (2015) investigates the relationship between the passenger loading factor and trip travel time. He found that there is a strong relationship between passenger load factor and average travel time across the entire span of service. In other words, passengers make longer commutes with significantly less comfortable conditions towards peak direction than in the off-peak direction. This correlation further implies that higher passenger load factors are associated with longer commutes and hence reduced transit LOSs (Bunker, 2015). Shen et al., (2016) propose a bus comfort model using survey data to evaluate passenger comfort level which takes both passenger load and in-vehicle time into account (Shen et al., 2016). It has been found that, the passenger loading factor can be related to perceived in-vehicle time rather than the real in-vehicle time (Wardman, 2004).

1.3.2.5. Reliability

Reliability is the ability for a transit user to have confidence in the expected service performance of the transit system. This is one of the most important measures that forms an attitude towards a transit service and affects the perceived LOS of passengers. This is evident from the findings of Golob et al., (1972) which support the fact that passengers perceive ‘arriving when planned’ as one of the most important service attributes of a transit system (Golob et al., 1972). TCQSM presents ‘on time performance’ (percent of schedule deviations that fall within a defined range), ‘headway adherence’ (coefficient of variation of headways) and ‘excess wait time’ (average of the non-negative schedule deviations) as the three main measures of transit reliability (TCRP Project A-15C, 2013). It is further reported that on-time performance needs to be measured at locations that make sense. As an example, if it is measured at a terminal point where normally almost all the passengers have alighted (gotten off) before, it makes no improvement to the passenger perception of QOS. Next-to-last timepoints and timed-transfer centers are recognized as the most efficient locations to measure on time performance due to their potential impacts on perceived QOS. 'On-time performance' was found to be the most widely utilized measure on reliability across the North American transit industry (Benn, 1995). Not having a widely accepted standard to identify the limits (acceptable threshold on deviations from schedule) of being 'on-time' is recognized as one of the negative aspects of this measure. It was found that in the mid-1990s, about 42% of the transit agencies considered vehicle arrivals that are more than 5 minutes late as 'on-time' (Benn, 1995) whereas it was found that in the Canadian context in 2000, 11 out of 17 agencies only considered vehicles as 'on-time' if they are no more than 3 to 4 minutes late (Canadian Urban Transit Association, 2001). In such a situation, TCQSM defines being 'on-time' as arriving 1 min early to 5 min late as the recommended value.

For frequent transit services where the headway is approximately 10 minutes or less (maximum average wait time at a stop is half the headway), passenger arrival at a stop is normally assumed to follows a random distribution whereas the average headway is greater than or equal to 15 minutes, passengers tend to time their arrival at the stop based on the information available on the schedule. Passengers tend to arrive at their station a few minutes earlier to account for the not-perfectly reliable service in case they miss their bus due to early departure and to account for possible slight uncertainties of accessing the stop. Depending on this situation, TCRP Report 113 (2006) defines four measures of wait time as Excess wait time, Excess platform wait time, Potential waiting time and Budgeted waiting time (Furth et al., 2006). The lower the reliability of the transit service, the higher the values of the above-mentioned measures and hence the lower the perceived transit LOS.

1.3.2.6. In-vehicle Travel Time

It is obvious that travel time is one of the major factors affecting transit QOS where most often, the reason behind using a personal vehicle is for saving travel time. An average person will compare the transit travel time with that of using their next possible alternative which usually is their private vehicle. Therefore, one of the most successful measures to account for this factor is the ratio between transit and auto (private vehicle) travel times that is in-vehicle transit travel time divided by the in-vehicle auto travel time (TCRP Project A-15C, 2013). Lower transit travel time will obviously induce an increase in transit ridership while a lower auto travel time will not always decrease the transit ridership as there are other benefits of using public transit as compared to private vehicles such as lower costs, less stress in traffic congestions, ability to utilize the transit travel time to undergo other activities (e.g. reading, phone calls, personal works etc.) which will outperform the potential travel time saving from private vehicles up to a certain extent depending on the circumstances. TCQSM provides passenger and operator perspectives for several levels of transit-auto travel time ratios.

1.3.2.7. Information Availability and Intelligent Transportation Systems (ITS)

Information Availability (IA) is another major factor found to have a significant impact on passengers' perception of the service quality (NRG Research Group, 2017). This has been made possible due to the technological advances in the field of ITS. ITS applies advanced data communication technologies to integrate users, vehicles and infrastructure which has made better management of transit operations and effective dissemination of information possible. Real Time Traveler Information Systems (RTTIS) is one of such outcomes that has become popular among both transit users and operators for its benefits over costs (Hatcher et al., 2017). Real time IA has impacts on several other major attributes of transit quality. RTTIS for example has reduced the passenger waiting time specially in low frequency routes and perceived value of wait time (Lu et al., 2018; Watkins et al., 2011). The perceived wait time has been reduced as a result of reducing passengers' anxiety in waiting for the bus through bus arrival information provision at bus stops (Fries et al., 2011). A survey found that without any changes to actual on time arrival performance of buses, 64% of the passengers has perceived that on time performance (reliability of the service) has improved, after the implementation of a bus arrival information system in London (Transit Cooperative Research Program, 2003). Apart from RTTIS, ITS has many other forms that have significant impacts on other attributes of transit quality. Automatic vehicle location systems have been proven to improve reliability (Camus et al., 2005). Transit signal priority (TSP) and other transit priority measures (e.g., dedicated bus lanes and queue jumping) reducing transit travel times and bus delays significantly (Hatcher et al., 2017) are only a few of the impacts of ITS on transit quality.

1.3.2.8. Other Factors

TCRP Report 88 (2003) presents several other factors important for passengers in assessing comfort and convenience of transit service which are not as easy to assess as the factors discussed so far. They are identified as Passenger Safety and Security, Customer Service, Compliment and Complaint Tracking (Kittelson & Associates et al., 2003). The most recent development in the area of LOS can be recognized as the formulation of a composite LOS index – known as the multimodal LOS – that can be derived separately for each mode of transport operating on the same street section (not grade separated) developed by the TCRP Report 47 (Morpace International Inc & Cambridge Systematics Inc, 1999). This measure provides an opportunity to compare relative service quality experienced by users of every mode operating on the same street. The interesting aspect of this method is that it incorporates several components of a transit trip (Walking to the stop, Waiting for transit at station and on-board satisfaction in terms of crowding level and speed) into one single letter of LOS. It uses data on transit operation, transit amenity, and pedestrian environment to calculate a transit wait-ride score and pedestrian environment score to calculate the Transit LOS score that has been divided to six ranges, each denoting a letter from A through F. This measure does not represent the LOS of the transit service but the transit LOS of a certain road section that takes all the transit services operating in that road section into account.

1.3.3. Efforts towards an Overall LOS Measure

To the best of our knowledge, there are no studies that provide a single LOS measure assessing a transit line or a system that takes the concerns of both operator and users into account nor such measures in practice. The reasons behind such attempts being missing from literature are the many challenges such a measure would pose. Some of them can be described as: (1) there are different interested parties such as transit agency, passengers, community; (2) there is no specified, widely accepted framework to assess LOS; (3) there are several governing attributes both qualitative and quantitative; (4) some attributes are competing – e.g., Access time and Travel time; and (5) passenger concerns influencing LOS depend on the spatial and temporal ridership market. In fact, Higgins and Ryan (1983) claim that it is "extremely difficult" to formulate a single measure that incorporates most of the important attributes of a LOS framework owing to the competing nature (trade-offs between attributes) of attributes. However, there exist a few studies that have taken this risk and been able to address the above challenges in part, but not without their limitations. Table 1 provides a summary of such studies with an emphasis on their limitations and the framework used.

Botzow in 1974 proposes six different levels for each of the transit characteristics that he identified as important (e.g., adjusted speed, delay, space, acceleration, temperature etc.), and a point scale of his own that assigns different amounts of points for achieving each LOS in each characteristic. Different amounts of points are assigned for the same service level (for example LOS 'B') of different attributes depending on the perceived relative importance of the considered attributes by the author. The points for these characteristics are then summed up to come up with the total amount of points. Then he introduces the ranges of total amount of points for six service levels (A through F) that represent the LOS of that transit system. This methodology can be used to compare different modes of transit or the variations in the same transit mode. The levels of service suggested in this methodology are solely based on the judgement of the author and there is no user perception involved.

Table 1 - Research on formulating an overall LOS

Name	Study	Constituent Attributes	Limitations
Level of Transit Service (LOTS) - A point scale and letter grades	Botzow, 1974	7 Variables (Adjusted Speed, Delay, Space for passengers, Temperature, Ventilation, Noise)	Solely depends on expert/author opinion, Subjective, Not the passenger perception, weights given to attributes and the ranges of points assigned for each letter grade (A through F) were determined by author
Standard Rating for Transit - A point scale	Aronstein, 1976	14 variables both Quantitative and Qualitative (Velocity, Travel time, Wait time, Cost, Safety etc.)	Depends on expert opinion and CS Surveys (SP). Only a point scale - No impression on whether the quality is good or bad, can be used only to compare transit lines. Data intensive. Need to be done before and after each improvement

Service Quality Index (SQI)	Hensher, 2003	13 Variables (Reliability, Travel time, Bus fare, frequency, having a seat, cleanliness, info. Etc.)	Solely depends on SP CS Surveys. The weighting factors and points for each attribute derived from the extensive survey itself. Data intensive. Calculated for segments of a route. Need to be done before and after each improvement
Transit Service Indicator (TSI)	Liping Fu, 2007	Service frequency, Service span, access, travel time, demand	The measure is the ratio of perceived total travel time by auto and transit. This indicator highly depends on the characteristics of the region (Congestion, road conditions, traffic prioritizing techniques etc.) Can be used to compare. No levels or quality implication. Needs to be done before and after each improvement

Aronstein (1976) proposes a standard rating (i.e., benchmark rating) for a particular transportation system that incorporates traveler experience-based rating for attributes of a transit service and assigning a weighting factor depending on the importance of such factors for a given journey. For quantitative attributes of service, lowest and highest possible values have been identified depending on common sense or previous work. The range of service values between lower and upper bounds are then scaled from 0 to 10 where not all the scales are linear. For qualitative attributes, the percentage of riders satisfied with a particular attribute will represent the rating of that particular attribute where 25% of the riders in the sample are satisfied, the rating will be 2.5 out of 10. A total of 14 attributes have been introduced and 3 weighting factors are assigned to 3 clusters of attributes that are identified according to the author's opinion. The author has then introduced values for each of the attributes of a nominal transit system. The final calculated value is used to compare the quality of other transit systems/lines depending on its total score. The weighting factors assigned, however, have no logical basis and the ratings given for a nominal transport system solely reflect the judgment of the author and hence are subjective. Therefore, this methodology, while it shows progress in formulating a single indicator to represent service quality, fails to represent an unbiased overall passenger perception as some of the attributes (quantitative) are still dependent on expert opinion.

One of the main drawbacks of these measures, as described in Table 1, is requiring customer satisfaction surveys to be conducted as an integral part of the assessment. Conducting passenger surveys is significantly time and money intensive. One of the objectives of having an LOS measure is to assess the impacts of interventions (e.g., operation changes, investments, extensions etc.) in service improvements, LOS assessments need to be carried out before and after each intervention to assess the efficiency of such and number of improvements made. Doing pre and post surveys to assist this method can sometimes demand more money than the interventions themselves, especially in a context where labor costs are comparatively high. Using expert opinion about passenger perception on service attributes to counter this issue results in a difference of assessed and experienced service quality depending on the characteristics of the ridership market.

The most recent development in the area of LOS can be recognized as the formulation of a composite LOS index that can be derived separately for each mode of transport operating on the same street (not grade separated), for a particular section of the street, developed by the TCRP Report 47 (Morpace International Inc & Cambridge Systematics Inc, 1999) and incorporated in the third edition of the TCQSM (the full version is presented in the Highway Capacity Manual). This measure provides an opportunity to compare relative service quality experienced by users of every mode operating on a street segment. The uniqueness of this method is that it incorporates all the components of a transit trip (walking to the stop, waiting for transit at station and on-board travel time) into one single letter of LOS which also depends on the characteristics of other transport modes/facilities in the road segment.

1.4. Measuring QOS based on perspectives

1.4.1. Importance of both the perspectives of passengers and operators in a measure

The role of transit provider being to maximize QOS improvement under limited funds, it is necessary to identify where the transit QOS is lowest or where there can be the maximum possible QOS improvement for a unit of investment. For this purpose, it is necessary to be able to measure the QOS of public transit. QOS from a passenger perspective can be defined as ‘to what extent does the transit service meet my expectations?’. The same can be described from the operator perspective as ‘to what extent does the service meet passenger expectations?”. To be able to quantify, the expectations should be measurable and achievable. Having a method to quantify QOS helps the service provider identify potential increments in QOS for a unit of investment. It is therefore possible to decide on the ways that funds should be utilized to achieve a maximum return on investment. An important point to be noted from a service provider's point of view is the operator cost. It is also a component of the generalized total cost that needs to be minimized, as the operator is also funded by the community, which is a cost for the community. A significantly high operator cost can lower generalized passenger costs significantly and hence make the perceived passenger QOS higher which is not in turn sustainable from a community perspective. The third edition of the TCQSM points out that "transit operators must strike a balance between the quality of service that passengers would ideally like and the quality of service that a transit agency (a) can afford to provide or (b) would reasonably provide, given the demand for transit service" (TCRP, 2013). Yet, the TCQSM or any other transit guidelines have not suggested at least a conceptual approach, let alone a comprehensive methodology, through which these concerns can be addressed. Therefore, in an attempt to suggest such an approach, this study looks at the operator cost as similar to the user cost where the summation of both these costs (user and operator) needs to be minimized. This will help QOS to be assessed effectively in an overall sense.

1.4.2. Evolution of Transit Level of Service (TLOS) measures

From the 1950s the LOS concept has been widely utilized in the context of highway planning and design. As early as 1965, the level of service for highways had been defined in terms of the effect of several factors including speed, traffic interruptions, safety, comfort, freedom to maneuver, travel time, convenience and cost (Spring, 2000). Yet, until the end of 1970s, there have not been standardized measures for evaluating public transport services (Taylor & Brogan, 1978). The Level of Service methodologies for transit have first been developed to help answer transit efficiency questions posed by transit authorities (Taylor & Brogan, 1978).

As public transportation changed from a private, profit oriented operation to a public service, profitability alone was no longer a viable measure for the evaluation of such a service and as a result, the LOS concepts were developed with due attention to the non-monetary values for the community that has caused the inclusion of measures on service receiver (Taylor & Brogan, 1978). Earliest reported work attributed to a TLOS concept has been performed by the US National Committee on Urban Transportation which recommends standards for routing, loading, frequency of service or headway, stop frequency, speed and regularity of service (National Committee on Urban Transportation, 1958).

Due to the fact that measuring QOS is of utmost importance for service providers, different transit authorities have developed their own measures over time (Higgins & Ryan, 1983; Taylor & Brogan, 1978). In the absence of a commonly accepted standard defining QOS in transit, TCRP of the TRB, in their first edition of the TCQSM in 1999, specified six designated levels of the QOS denoted by letters "A" (Best) to "F" (Worst) as it is easier to refer QOS with several levels rather than a continuous scale.

1.4.3. Passenger Perspective

As better QOS is possible with a higher passenger demand, passenger perception is important for the service provider. Incorporating passenger perception into a measure assessing QOS can improve transit ridership and QOS in return. There is a large body of research pointing out the importance of user perception in the assessment of transit LOS (Das & Pandit, 2015; Eboli et al., 2014; Noor et al., 2014; Tyrinopoulos & Antoniou, 2008). Irrespective of the fact that many studies stress the importance of passenger perception in evaluating transit LOS, almost all the LOS measures currently in practice, use expert opinion (Highway Research Board, 1950; TCRP Project A-15C, 2013). Universal application of measures developed based on expert opinion can be different from what passengers actually perceive. Overall passenger perspective of the service of a particular transit line or a system is subject to the characteristics of the ridership market of that transit line/system such as average income, age distribution, trip characteristics, demand profile, etc. For example, the same headway of two bus routes along two ridership markets with different average incomes and trip purposes will pose different levels of QOS as riders in these two routes tend to perceive the value of their wait times differently.

What is most important in assessing TLOS based on passenger perspective is the identification of the set of service attributes that most influence the perceived TLOS. As transit ridership is the major concern of the transit agencies and government as higher transit ridership would imply higher QOS and hence higher urban living standards, it is critical for the agencies to find out the factors that can induce an increase in ridership. This helps transit authorities prioritize their future funding on improvement of such factors. Assessing the service attributes affecting transit user satisfaction therefore becomes of greater importance for transit agencies. A comprehensive set of potential service attributes (48 attributes) that a transit agency can utilize to identify the factors that are important for their potential market is presented in the Transport Corporative Research

Program (TCRP) Report 47 (Morpace International Inc & Cambridge Systematics Inc, 1999) along with a detailed description on how to perform a customer satisfaction survey. There are many other researchers who also contributed towards the literature on public transit quality of service parameters (Andaleeb et al., 2007; dell'Olio et al., 2010, 2011; Eboli & Mazzulla, 2007; Hensher et al., 2010; Mamun et al., 2013).

There are several studies that develop LOS measures for individual attributes of transit service using passenger perception acquired through stated preference or revealed preference surveys (dell'Olio et al., 2011; Prioni & Hensher, 2001). Some studies have also developed overall QOS measures using passenger survey data that output a continuous measure (Fu, 2007; Hensher et al., 2003). These measures are data and time intensive and hence expensive. Also, they need customer satisfaction surveys to be done pre- and post-intervention to estimate the improvements or change in TLOS. Therefore, they are not viable to be estimated frequently for maintenance and quality adherence requirements.

1.4.4. Operator Perspective

Earliest reported work on TLOS measures was purely based on operator perspective and consisted of direct quantitative measures like loading, frequency of service, stop frequency, speed etc. and were used mainly to help transit agencies with management. Although, later, owing to the significant number of studies pointing out the importance of the passenger perspective towards TLOS, most of the studies that have been carried out are based on passenger perception (service receiver's point of view). Due to this reason, there are fewer studies carried out to assess the transit LOS that incorporate both the concerns of service receiver and provider. There have been attempts to incorporate the operator's point of view as well into the quality-of-service measures so that the assessed level of service will imply the concerns of the operator (budgetary and facility constraints) as well. Significant research attention has been paid to the perspective of the operator within the LOS framework recently as the rising budget deficit between the higher operating costs of a higher LOS and low farebox revenue where transit agencies are pushed to seek subsidies from the government (Ibarra-Rojas et al., 2014; Sheth et al., 2007). This is evident from the fact that third edition of the TCQSM has included the operator perspective in each and every quality of service measure presented for assessing fixed route transit (TCRP Project A-15C, 2013).

Irrespective of the demands in academia through literature for TLOS measures to include passenger perspectives, the majority of industry practices utilize expert opinion on how passengers perceive a given service parameter rather than a way to include real passenger perspective. The newest (3rd) edition of the TCQSM bears witness for this claim (TCRP Project A-15C, 2013). Therefore, the quality of a transit service assessed on such grounds can most of the time be biased towards the operator who would like to reduce costs as much as possible. Operator bias in the assessment will provide the operator with the ability to trade operator costs to passenger costs at a higher value leaving passengers undervalued. As a result, passenger perception of the service rendered will not be accurately represented by the TLOS obtained. Therefore, a balanced view that combines both the perspectives of passengers and operator will provide a fair representation of the quality of the service provided.

1.5. A unified measure of TLOS

1.5.1. The abstract idea of TLOS

As discussed in the 'Introduction', what passengers perceive as TLOS is an overall experience of the total service package provided by a particular transit line or a system. This is also what operators are concerned about, i.e., sustaining and improving ridership, with respect to the amount of funds invested. For the ease of analysis, this abstract idea needs to be estimated as accurately as possible. For this purpose, a framework for QOS is required. The greater the number of constituent attributes used in the framework, the more accurate or closer the answer would be to the actual average abstract idea of passengers. As indicated in section 1.4.2 of this chapter - "Evolution of TLOS measures", one study had identified about 48 potential attributes that can be considered as constituents in such a framework. However, studying such a large number of attributes to assist in assessing TLOS is not practical. The best approach would be to identify the few, most significant constituent attributes to be used in the framework that can yet provide an accurate enough estimate about the average passenger perception with least effort (using least number of attributes). There have been individual attempts (Garrido & Ortuzar, 1994) and industrial attempts (TCRP Project A-15C, 2013) to find the most significant constituent attributes. TCQSM suggests a framework for assessing TLOS for the north American context that includes availability and quality aspects of transit as constituent attributes. Here, TCQSM specifies different ranges of service quality for each of six attributes of the framework separately but does not specify a unified measure that represent TLOS of a transit system or a transit route (TCRP Project A-15C, 1999).

1.5.2. A Framework for Transit LOS

In the earliest research, a LOS framework was mainly utilized as a tool to decide on transit agency funding where not much emphasis was given to community requirements and benefits; and therefore, the measures used to assess LOS were mainly performance measures assessing the operation (Higgins and Ryan 1983). As public transport was accepted more as a welfare service and community concerns got the better of financial aims of transit agencies, more emphasis on user perspective was given in the LOS framework utilized by transit authorities and this is evident from the fact that LOS framework of TCQSM 1st edition and 2nd edition are mainly developed on user perception. There is a significant amount of research pointing out the importance of incorporating user perception in the assessment of transit LOS (Das & Pandit, 2015; Eboli et al., 2014; Noor et al., 2014; Tyrinopoulos & Antoniou, 2008). However, significant research attention has also been paid to the concerns of the operator within the LOS framework recently due to the rising budget deficit between the higher operating costs of a higher LOS and low fare box revenue where transit agencies are pushed to seek subsidies from the government (Ibarra-Rojas et al., 2014; Sheth et al., 2007). This is evident from the fact that the third edition of the TCQSM has included the operator perspective in each and every quality of service measure presented for assessing fixed route transit (TCRP Project A-15C, 2013).

Florida Department of Transport developed a Transit LOS indicator (TLOS) which is a composite measure that partially represents the transit LOS using a single value. TLOS indicator only represents the availability area of QOS framework suggested in the TCQSM, yet it accounts for all the measures included under availability and more. Accordingly, TLOS indicator provides a single measure of transit availability that constitutes service frequency, service coverage, hours of service, population and job densities (Ryus et al., 2000).

With transit ridership being the major concern of the transit agencies and government, a higher transit ridership would imply higher QOS and hence higher urban living standards, it is critical for agencies to identify the factors that can induce an increase in ridership. This helps transit authorities prioritize their future funding on improvement of such factors. An improved customer satisfaction, which can be achieved through an improved service quality in the eyes of users can induce a rise in ridership (Eboli & Mazzulla, 2007). Assessing the service attributes affecting transit user satisfaction therefore becomes of greater importance for transit agencies. A comprehensive set of potential service attributes (48 attributes) that a transit agency can utilize to identify the factors that are important for their captive market is presented in the TCRP Report 47 (Morpace International Inc & Cambridge Systematics Inc, 1999) along with a detailed description on how to perform a customer satisfaction survey. Garrido & Ortuzar (1994) used the Delphi method to methodologically identify 12 factors affecting public transport QOS in the context of a developing city (Garrido and Ortuzar 1994). There are many other researchers who also contributed towards literature on public transit QOS parameters. In a study carried out in South Africa, Roebuck showed that the mean journey speed, ratio of flow to capacity and load factor are some of the parameters that indicate the transit LOS (Roebuck, 1979). Through a revealed preference survey of commuters in the city of Accra, Birago et al., (2017) identified that factors such as over-crowding of buses, non-adherence to time schedule, risk of seats being unavailable, unavailability of buses at trip origin and destination areas, long waiting times and long in-vehicle times have negatively affected the transit LOS.

The TCRP of the Transport Research Board has conducted a significant amount of the pioneering research work that integrated other independent research findings as well in establishing a widely accepted QOS framework for transit and has published in first, second and third editions of TCQSM. Going through their QOS framework and service measures suggested provides a broader yet comprehensive insight of potential factors affecting Transit LOS. Table 2 shows the transit QOS framework together with the core service attributes.

Table 2 - Fixed Route Transit QOS framework of TCQSM 3rd Edition

Availability	**Comfort and Convenience**
Frequency	Passenger Load
Access	Reliability
Service Span	Travel Time

Source – TCQSM 3rd edition (2013)

QOS measures presented in this framework are categorized under two areas namely "Availability" and "Comfort and Convenience" as each area affects the mode choice behavior of potential users at two stages. Looking for transit 'availability' is a day-to-day quick decision that people take to decide on using transit or not whereas 'comfort and convenience' is an experience people get gradually by using the transit service over time. From a different perspective, the measures falling into the availability category like service frequency, service span and transit access are tangible performance measures that can be measured directly using mostly the transit operational data whereas measures under comfort and convenience like passenger loading, reliability and travel time can be identified as intangible measures that cannot be directly measured and need to undergo data collection and analysis. One thing to note here is that the service span can only be categorized as a measure of transit availability but not the quality of transit availability while the other two measures – headway and access – change the passenger perception of transit depending on the amount of service provided. Service span only helps them decide whether or not to take transit and does not improve or reduce the quality.

1.6. Challenges and Issues

To the best of our knowledge, there are no studies that output a single LOS measure assessing the overall QOS of a transit line or a system that takes both the concerns of operator and user into account. Attempts like that have been rare due to the many challenges such a measure poses. Some of them can be described as: (1) there are different interested parties like transit agency, passengers, community; (2) there is no specified, widely accepted framework to assess LOS; (3) there are several governing attributes both qualitative and quantitative; (4) some attributes are competing – e.g., Access time and Travel time; (5) passenger concerns influencing LOS depend on the spatial and temporal ridership market. In fact, Higgins & Ryan (1983) claim that it is "extremely difficult" to formulate a single measure that incorporates most of the important attributes of a LOS framework owing to the competing nature (trade-offs) of attributes.

As an example, one would like to walk less to the nearest bus stop – or reduce the spacing of the bus stops – and reduce the total travel time at the same time. But both these requirements have tradeoffs with one another. To reduce walk distance or the access time of the commuters, the operator needs to establish bus stops closer on a bus route which will increase the number of bus stops. This will increase the total travel time of the route as now the bus has to stop at a higher number of bus stops to board and alight passengers (load and unload passengers) although under the same operating speed. There are no studies addressing such issues in coming up with an overall

LOS measure. Instead, the normal practice has been to utilize separate TLOS gradings for individual constituent attributes of the framework (TCRP Project A-15C, 1999).

Having individual LOS gradings for different service attributes of a transit service, however, compromises the requirement of having a LOS measure which is to assess the overall quality of the service. Such methods were mostly the results of transit agency efforts as it is easier to maintain and assess the service quality in individual service aspects and help administrative decisions (Taylor & Brogan, 1978). On the other hand, selective focus on individual attributes of a service can result in incomplete evaluation conclusions where it is possible that better alternatives could be overlooked (Hodge & Orell, 1995). Accordingly, the traditional transit LOS measures on service frequency, route density, access distance etc., stress universal outcomes, without regard to the actual service requirements – e.g., a given level of headway, say 10 to 15 minutes, can be a preferable headway for a suburban bus route, but the same level of headway will not be preferable for an urban bus route, especially in the peak period.

1.7. Contributions and Novelty

Although there is a significant amount of work carried out in the field of LOS of urban bus transit, to the best of our knowledge, there are no studies proposing a method to assess the overall LOS of a transit route, with due attention to both passenger and operator concerns. This study, for the first time, formulates a LOS measure to assess the overall LOS of an urban bus route which is the most important contribution. Furthermore, unique LOS measures are derived for each or combinations of attributes of the LOS framework which includes both passenger and operator concerns. Accordingly, the study develops LOS measures for the attribute reliability, for the combination of attributes headway and crowding, and for the combination of attributes, travel time and access time. The study then combines these three LOS measures into one measure that represents the overall LOS of the route.

The LOS criteria developed for an individual or a combination of attributes of the service will provide transit agencies and the public with more useful and efficient insights depending on the unique passenger market of a transit service. Current methods usually provide universal measures in assessing these attributes which may not be representative of the unique passenger market of that transit service.

2. Chapter 2: Proposed Approach

2.1. Introduction

This study utilizes a similar QOS framework to that presented in TCQSM (described in Chapter 1) to develop individual QOS measures and combine them. Accordingly, the proposed LOS framework that will be considered in this study is shown in the Figure 1.

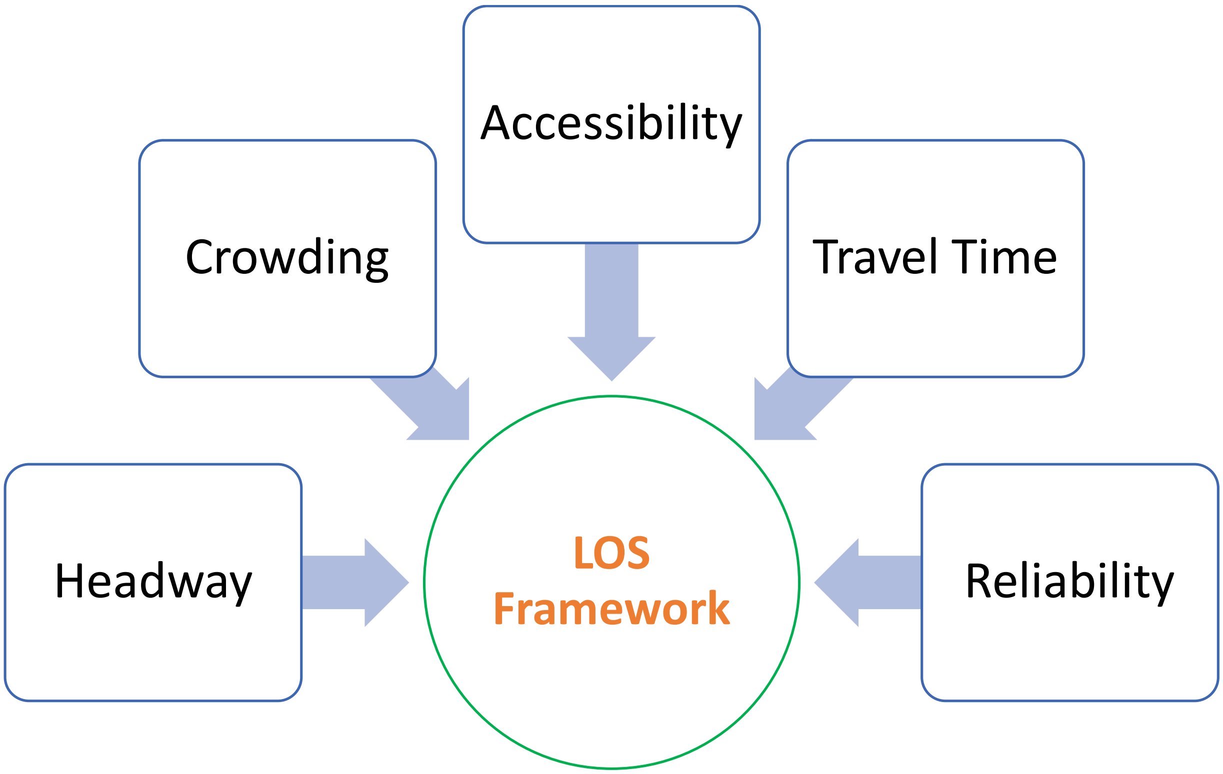


Figure 1 - Proposed LOS framework

One of the attributes considered in the TCQSM framework is the ‘service span’. As also described in Chapter 1: Background, improving service span is not going to affect the passenger perception of transit service quality. For instance, Calgary Transit customer satisfaction survey done in 2016 reveals that although the service span has become a significant factor for the non-transit users, it has not been considered a significant factor affecting transit users – not among the ten most important factors affecting the transit quality as identified by the passengers (NRG Research Group, 2017). Therefore, service span is not a factor significantly affecting the people taking the transit. If the proposed approach concerns the people who are currently not taking the transit but possible to take transit in future, service span could have been a significant factor and hence required to be included in the framework. Given this background, the attribute ‘service span’ has been excluded from the QOS framework considered in this study while all the other attributes in the TCQSM framework have been considered. It is important to note that this study does not intend to dive into the problem of determining the appropriate framework but to use one that is available in the literature with minor adjustments.

Owing to the dependence of some of the attributes in the framework – travel time competes with access distance while headway affects the crowding – some of the attributes have been combined to a single LOS measure. Accordingly, the headway and crowding attributes are represented through the Headway and Crowding LOS, while access and travel time attributes are represented through Access and Travel Time LOS. Only reliability is represented by a single measure. The combined LOS measure together with sub-LOS measures will be developed in a way that they can be used for any urban bus route where key operational parameters are known or can be derived using available data, and where the value of time (VoT) distribution of passengers (including mean and standard deviation) is available. The approach for deriving the combined LOS measure for the overall LOS of a bus route is illustrated in Figure 2.

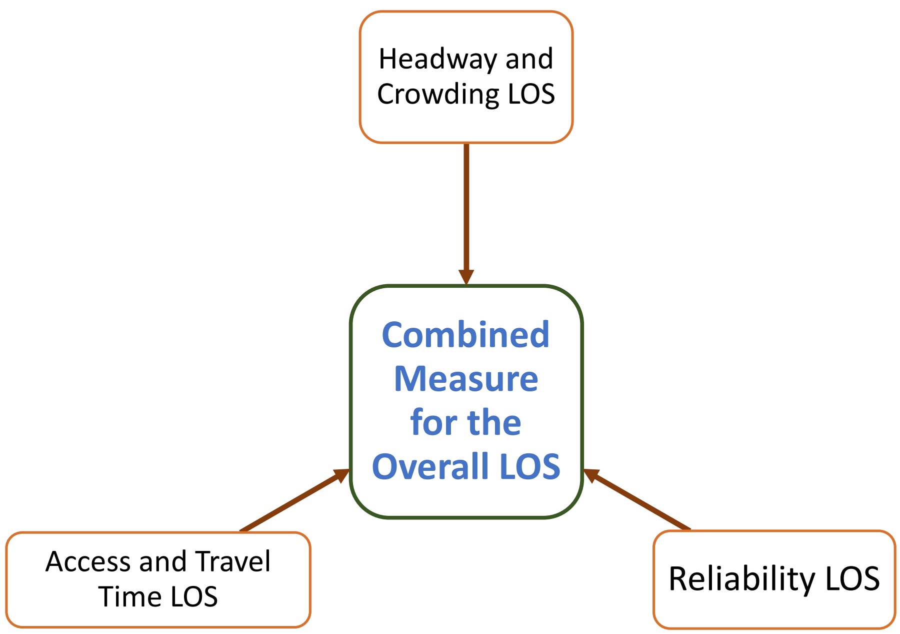


Figure 2 - Composition of the overall LOS measure

2.1.1. Optimal operation of transit routes and square root policy

The concept of optimum service rates was first propounded in a study attempting to price public utilities based on marginalized cost pricing (Vickrey, 1955). However, the pioneering work in the field of optimal transport services was carried out by G. F. Newell and Herbert Mohring (Mohring, 1972; Newell, 1971) with the introduction of the “square root dispatching policy” that states the optimal rate at which buses are dispatched is proportional to the square root of the demand rate. Although there exists no evidence to date on transit agencies adapting to such strategies, transit manuals can be found elaborating the benefits of the use of such methods (Watson et al., 2002). However, there exists a significant number of studies based on the pioneering work of Newell on square root dispatching policy (Wirasinghe, 1990) where Newell’s dispatching policy has been applied to various types of problems ranging from, models of bus service (Clarens & Hurdle, 1975) to design of airport limousine services (Banks et al., 1982). The main objective of these attempts was to optimize the total cost of operation (user cost plus operator cost) to find out the optimum operational parameters. Such outcomes are derived using analytical models of related costs that provide valuable insights and functional relationships between variables. S. C. Wirasinghe has expanded on Newell’s work on the square root rule through several studies (G. Liu & Wirasinghe, 1991; Wirasinghe, 1990; Wirasinghe & Liu, 1995b).

Models of optimal transit operation can be used to estimate perceived user costs using several parameters of the service (Watson et al., 2002). Perceived user cost is the multiplication of the time spent by a passenger in a certain part of the trip, and the value that a passenger would assign – perceived value – for a unit of that time spent. In an optimal operational behavior, a transit agency economizes on the cost of a user’s time until it does not cost the operator more than what the user’s time is worth (Watson et al., 2002).

Despite this recommendation, transit agencies most of the time do not operate under optimal conditions. It is possible to obtain service parameters required to model an optimal operation with respect to a service attribute of interest. For example, the service parameters required to model bus frequency (i.e., a service attribute) according to the square root dispatching rule (i.e., optimal behavior) developed by Newell (Watson et al., 2002; Wirasinghe, 1990) are: existing frequency (headway), bus dispatch cost (available with operator), demand, and average value of wait time of passengers. Through these parameters, using an assumed optimal operational behavior, one can estimate the average value that operator assumes (or what the operation indirectly assumes) for the value of the wait time of passengers. This can also be described as the value that the operator is willing to trade off per unit time of a passenger to operate the service. As the difference between this value and the values those passengers would associate themselves with, implies a perceptional difference on the quality of the service experienced, this poses an opportunity to identify the quality of the service provided depending on the characteristics of the ridership market. This approach is used in the proposed methodology to derive the LOS measures on individual or a combination of attributes.

Likewise, a user's VoT with respect to each aspect or a combination of aspects can be derived using already existing models (e.g., square root rule) or models that this study will develop of optimal operational behavior. These time values will be compared to the available mean VoT distributions of passengers to derive LOS of corresponding attributes. The higher the derived (implied) value of time, the higher the LOS. Where only the mean and standard deviation is available instead of the type of distribution – mostly the case when time value data are obtained through a demand model – a normal distribution is assumed. Therefore, in the approach proposed, the VoT of the passengers plays an important role.

2.1.2. Role of the value of time (VoT)

The extent to which one's needs and wants are satisfied implies the quality of life (Flanagan, 1982). One of the obstacles in achieving those needs and wants is the time taken through transportation. The ability to meet more needs and wants in a shorter period of time contributes towards a higher quality of time. A transportation system with a higher QOS will help the community achieve a higher amount of 'needs and wants' at the cost of less time. Due to this importance of time in transportation, the amount of time incurred in transportation has been studied widely in attempts to improve the QOS. It could be found that the time spent at different components of a trip had different perceived values and hence different costs. Therefore, the time spent on a trip has basically been studied in two aspects.

One aspect is to convert the time spent at each component of a trip – access, wait, ride and transfer – to an equivalent perceived ride time i.e., in-vehicle time (Pratt & Evans, 2004; Wardman, 2004). The other aspect is to convert the value of a unit time spent at each component of a trip to a single equivalent value of a unit time mostly as a percentage of income or wage rate (Concas & Kolpakov, 2009; Watson et al., 2002). Further research in this subject denotes that non-time quality attributes of a trip also influence the perceived time and hence the perceived total time or perceived total cost of a trip. Such findings were evident in studies where load factor of a bus is related to the perceived in-vehicle time, and bus stop amenities – shelter, lighting, type of seats – are related to equivalent in-vehicle time (Balcombe et al., 2004). Both approaches, however, lead to the same outcome to quantify the perceived equivalent total cost of the trip which is, in other words, to obtain a generalized total cost.

Wardman presented a detailed review of the value of wait time studies (Wardman, 2004). Initially, the studies were only concerned with the value of in-vehicle time for private transport modes. In addition, studies have been conducted on wait and walking times compared to in vehicle times of private modes where it was found that walking and waiting times are worth 2.5 times more on average than in vehicle times (Quarmby D. A., 1967). However, on average, wait time has been valued 2 to 2.5 times more than the in-vehicle time and it is valued more than any time component of a transit trip (Mohring et al., 1987; Wardman, 2004). Popular methods of finding out the values of time are revealed preference (RP) and stated preference (SP) surveys, which were found to have effects on the results derived (Wardman, 2004).

2.1.3. Normally distributed VoT

Time values of passengers (value of wait time, value of access time, value of ride time etc.) are dependent on many factors like income, age, trip purpose, weather, thus they can be considered random (Ben-Akiva et al., 1993) and better represented through a VoT distribution (Mohring et al., 1987). In a general population, it is possible to assume that VoT of passengers is normally distributed with a unique mean (μ) and a variance (σ^2) (J. H. Banks, 1991; Rahman, Wirasinghe, et al., 2018). VoT is a perception that involves a psychological evaluation that varies by chance. From the early beginning of the field of psychology, such a variable is assumed to follow a normal distribution according to the 'Quetelet's Normal Law of Error' (Boring, 1961). However, when the ridership population deviates from being apparently homogeneous, researchers have also assumed VoT be log-normally distributed (Mohring et al., 1987; Yang et al., 2001).

As also discussed in section 3.2.3 of the Chapter 3, the most widely available statistics for transit agencies from the willingness to pay studies and stated preference surveys are the mean and standard deviation of the probability distribution of passenger VoT. The shape of the distribution is mostly not available. In such circumstances, the shape of the distribution that carries the maximum information entropy (therefore minimizes the risk of errors in the prediction) is the normal distribution (Jordaan, 2005). Therefore, this study utilizes a normally distributed VoT of passengers with a mean 'μ' and a standard deviation 'σ' to relate to the LOS. Therefore, depending on where the implied VoT falls in the distribution with respect to the mean and the standard deviation, the LOS can be determined. Simply comparing this derived value with the mean value of the distribution alone is not able to provide useful insights as the spread of the population also plays a significant role. Therefore, a methodology that involves the mean and a measure of the spread of data will provide a better relative assessment.

2.2. Standard deviation to distinguish the levels of QOS

The applicability of the 'Normal Law of Error' for psychological variables such as perceived VoT has facilitated the conversion of ordinal scales in psychological testing into equal intervals using standard deviation (Boring, 1961). Müller & Gosling (1991), Correia & Wirasinghe (2013) and Ndoh & Ashford (1993) utilize this approach to convert the results of passenger satisfaction survey data on an ordinal scale to an equal interval continuous scale. In Chapter 3 – Headway LOS, an approach is developed to measure the QOS of the attribute of the headway of a bus route using the implied mean Value of Wait Time (VoWT) and the mean VoWT distribution of the passenger population. Here, the VoWT distribution acts as the continuous scale and is divided into equal intervals to denote the LOS using categories. Accordingly, the standard deviation (SD) of the VoT distribution becomes a viable candidate for dividing the distribution into equal intervals those denote the LOS categories. There are several reasons supporting this decision.

Inequality in transport has been identified as one of the key elements of the broader issue of economic inequality, that contributes to social exclusion and wellbeing and ultimately to the economic development. Banister (2018) further points out that the normal practice of comparing the performance only with the mean value to meet transport justice leaving out the SD - spread of the data, has enforced inequality. The response to this issue most of the time has been to provide a minimum required service (Banister, 2018). A solution for this issue can be reached using SD along with the mean value in assessing transit quality, to reduce transit inequality.

The suggested approach, being based on the VoT distribution of the passengers, takes into account the income inequality of the passenger population. The suggested approach also considers the heterogeneity of the passengers – standard deviation of the VoT parameter accommodates the unobservable preference heterogeneity of the population (Hossain et al., 2015). One example is the age difference – middle aged people have higher incomes leading to higher VoTs while young and older aged have low incomes leading to having lower VoTs. If the levels of service cover the whole VoT distribution, the LOS spectrum represents the whole population where lowest and highest income earners have been accounted for. Two potential approaches, as shown in Figure 3 and Figure 4, are considered. In both the approaches, the majority of the distribution population – the peak – is divided using the SD. The two tails, however, have been extended beyond the width of a SD to include the entire population. This will only add 0.13% and 0.63% to both the LOS grades at the two tails of the distribution in Figure 3 and Figure 4, respectively, compared to the LOS grade with a width of one SD. Therefore, both approaches represent the entire population. We use a spectrum of letter grades starting from "A" – best –to "F" worst, where a range of one SD of values of time represents one LOS grade.

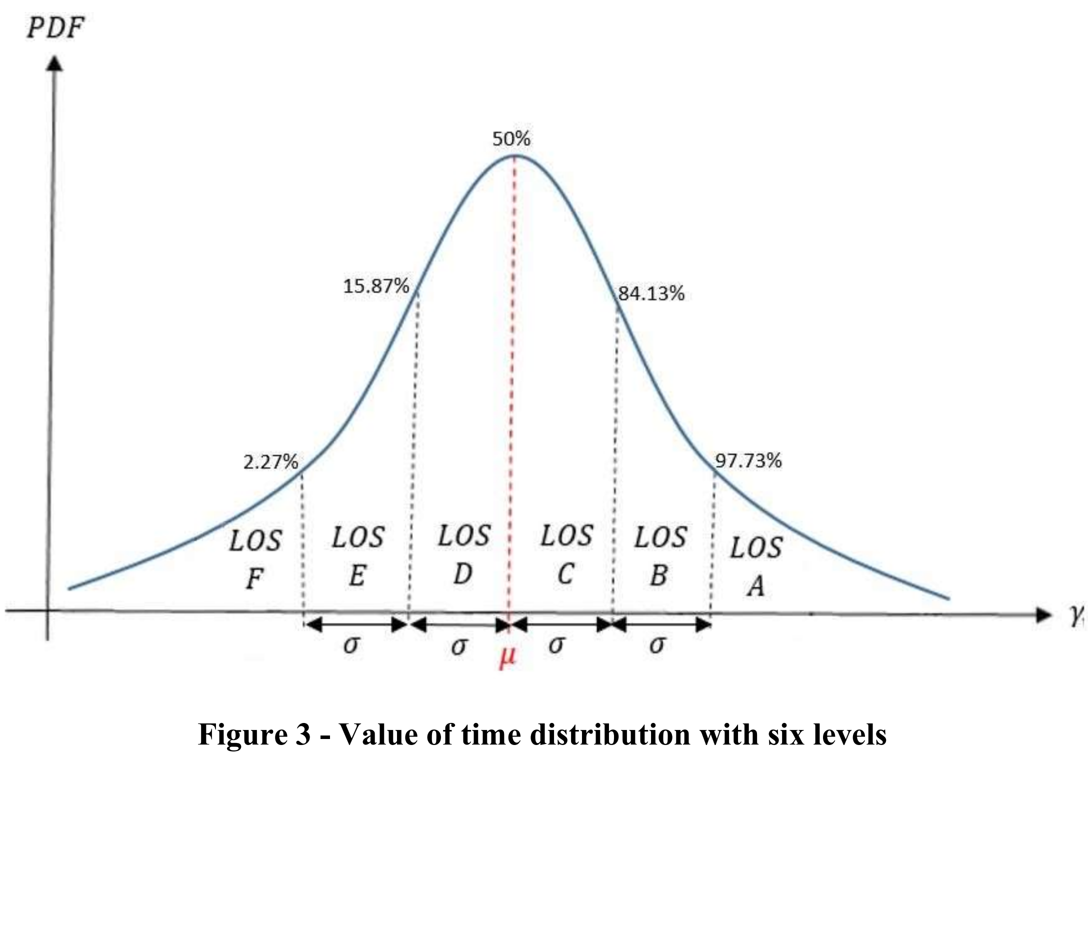


Figure 3 - Value of time distribution with six levels

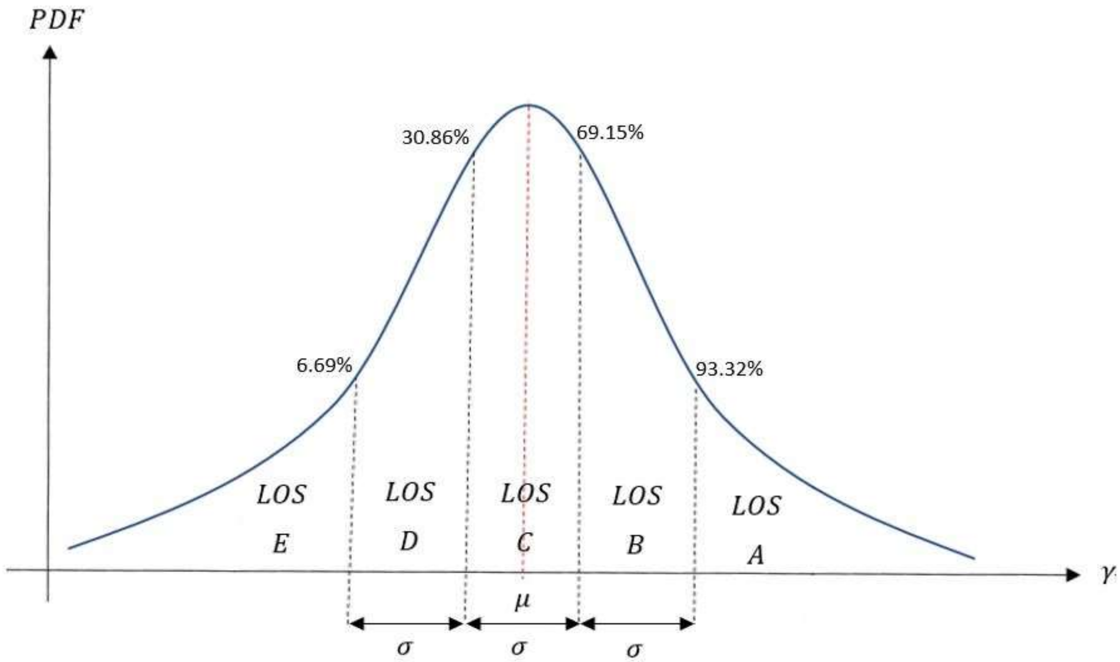


Figure 4 - Value of time distribution with five levels

$\gamma = Mean\ Value\ of\ Time\ of\ passengers$

$\mu = Mean\ value\ of\ the\ distribution$

$\sigma = Standard\ Deviation\ of\ the\ distribution$

For the approach suggested in Figure 3 and Figure 4, the LOS can be represented by obtaining the standardized value – also known as the Z score - of the implied VoT. Comparing this value with the standardized values of the category boundaries, the corresponding LOS can be obtained. Here, the Z scores for the lower boundary of the lowest quality category – LOS E, and the upper boundary of the highest quality category – LOS A, cannot be calculated as they are minus infinity and positive infinity respectively.

Given the implied VoT is γ', the corresponding Z-score value can be calculated as,

$$Z(\gamma') = \frac{\gamma' - \mu}{\sigma} \qquad 1$$

Accordingly, depending on the derived Z-score value the corresponding LOS grade can be assigned as shown below.

$$-\infty < Z(\gamma') \leq -1.5 \rightarrow LOS\ E$$

$$-1.5 < Z(\gamma') \leq -0.5 \rightarrow LOS\ D$$

$$-0.5 < Z(\gamma') \leq 0.5 \rightarrow LOS\ C$$

$$0.5 < Z(\gamma') \leq 1.5 \rightarrow LOS\ B$$

$$1.5 < Z(\gamma') \leq \infty \rightarrow LOS\ A$$

For a normally distributed random variable, the SD (σ) can be expressed as follows.

$$\sigma = \left[\int (\gamma - \mu)^2 f(\gamma) d\gamma\right]^{1/2} \qquad 2$$

$f(\gamma)$ = *probability density function of the VoTs*

According to Equation (2), the deviation of each value (VoT of each passenger) from the mean VoT – the distance from each value (γ) to the mean – has a direct effect on the SD i.e., a squared effect on the variance. Therefore, SD is more sensitive to the values further away from the mean as squaring the difference provides more weight. In contrast, the difference between each value and the mean has a direct effect on the mean absolute deviation (MAD) as MAD treats all deviations equally. MAD can be denoted by $\int |\gamma - \mu| f(\gamma) d\gamma$.Accordingly, the further the points from the mean, the higher the weightage given to those values in the calculation of SD. MAD on the other hand does not provide a special consideration to the data positioned further apart from the mean. Each LOS grade – each SD representing each LOS grade – represents a unique portion of the population in terms of their values of time. If MAD is utilized to distinguish LOS grades, the portion of the population represented by a particular LOS grade is not unique in terms of their VoTs. When the range of the distribution changes - increases or decreases, the portion of the population represented by a LOS grade as measured by MAD changes, compared to the case with SD. This is evident from the fact that the percentage population covered by each SD on the sides of the mean is always a constant while the percentage covered by a MAD, changes when the distribution changes – when the range changes as an example.

The passengers in the lower most percentiles are the ones who are significantly affected by the service quality depreciation as they are mostly the captive riders. Therefore, it is important that those representing the tails of the distribution are given due recognition in a measure assessing the quality of the service. Utilizing this method guarantees that the same LOS is represented by each LOS grade irrespective of the VoT distribution of the population - some cities have more adult percentages leading to a higher mean VoT, some cities have a higher mean income – cities in developed countries, while some have a lower mean income – cities in less developed and emerging economy countries. Utilizing a LOS measure that takes into account the difference in the income inequality of the population of a city/community ensures a sense of transport justice and better represents the collective passenger perception on service quality. For example, the same range of a given operational parameter – 10 to 20 minutes range of headway – will not represent the same LOS grade in two unique passenger populations like routes, cities– this aspect is discussed in the Headway LOS chapter in detail.

2.3. Entropy maximizing arrangement of the levels of service

Requirement of a letter grade system to represent the LOS of a transit system is more of an administrative requirement than a scientific one. For this reason, determining a preferable arrangement of levels or grades of service lacks a sound theoretical base therefore depriving anyone of building a proper argument, making it difficult to prefer one LOS grade arrangement over another. Both the approaches suggested in Figure 3 and Figure 4 satisfy the intended requirements to a great extent. Yet, one can argue one approach is better or biased than the other. Choosing the better arrangement of the LOS grades as denoted by the VoT intervals measured by the SD is challenging. The approach with six grades has a slight superficial advantage as it closely resembles the approach utilized in the Highway Capacity Manual (HCM) in representing the LOS of the road traffic condition (National Research Council (U.S.), 2010). This helps transport professionals to easily refer to the letter grade system. However, the HCM does not provide any proof that the six-grade system is the best one. This section intends to propose a method that provides at least some logical sense to select a LOS grade arrangement.

Each LOS grade in each approach accounts for a certain amount of probability. This leads each arrangement to a discrete probability distribution where each LOS grade has a probability. This can be further explained using Figure 5 and Figure 6 below.

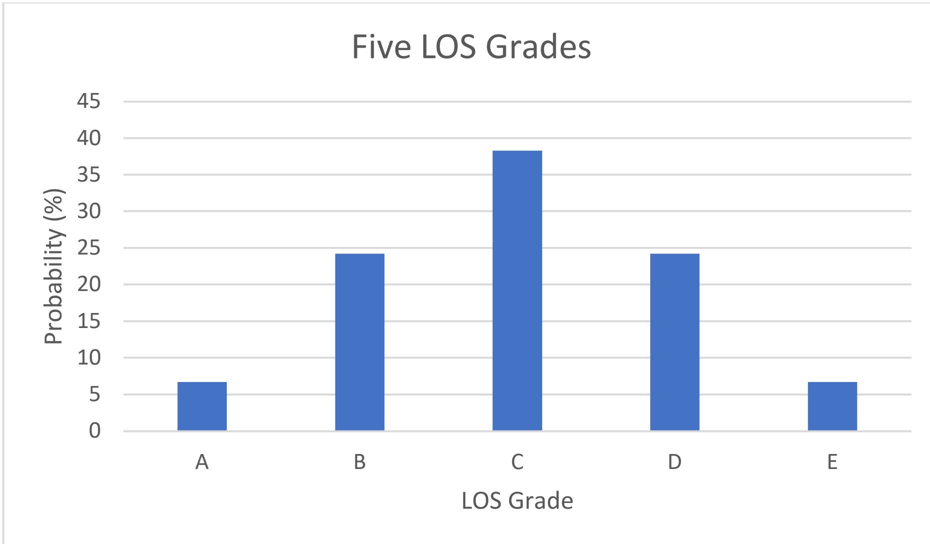


Figure 5 - Discrete probability distribution of the five LOS grades approach

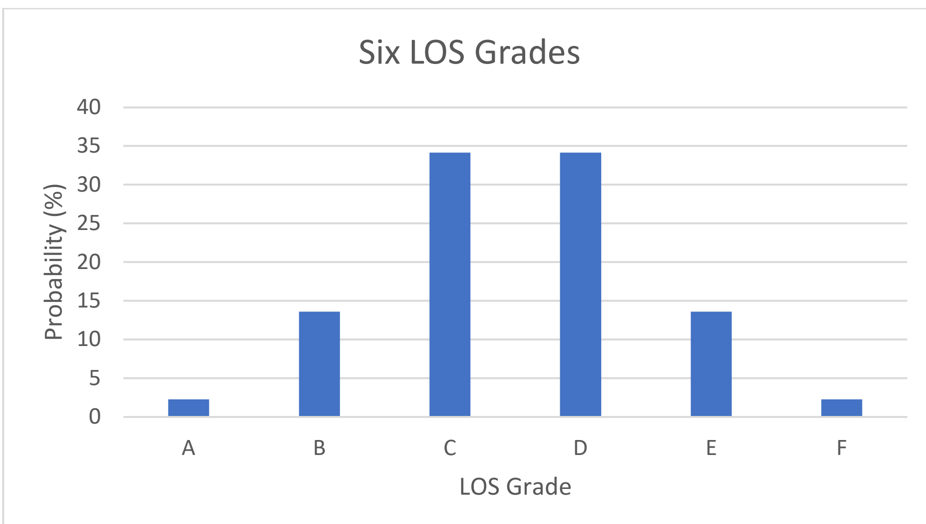


Figure 6 - Discrete probability distribution of the six LOS grades approach

Now we have two discrete probability distributions equivalent to the two approaches suggested. As mentioned earlier in section 1.3 of this chapter, the distribution that maximizes the information entropy, given the information; mean and SD of the distribution, is the normal distribution ((Jordaan, 2005) – p.294). This is the distribution that minimizes the risk of prediction by maximizing the entropy. The concept of information entropy, however, is described as "a jewel with many facets" by Jordaan (2005) for its usefulness in many aspects (Sun et al., 2007). Wilson (1970) further describes these uses in detail with examples. While SD is a subjective measure of uncertainty – a measure that depends on the numerical values of the VoT distribution, entropy is an absolute measure – does not depend on the numerical VoTs. This is also evident from the fact that the calculation of the entropy – Equation (3) – does not involve the real values of the variable but only the probabilities. For example, in a bi-modal distribution, while changing the distance between two peaks affects the SD, it does not affect the entropy.

Shannon's information entropy has gained further recognition recently, especially in the field of finance, for its explanatory power over the SD-related measures as a risk measure (Bentes & Menezes, 2012; Ormos & Zibriczky, 2014). Entropy can be used to determine the distribution with the least bias toward the information given about the distribution – mean and SD. The distribution with the maximum entropy, makes it the least presumptive distribution that can be used for inference (Jaynes, 1957). The higher the entropy of the discrete probabilities, the less the bias towards given information. Kapur (1993) also mentions that given partial information about a random variable, the distribution with the maximum uncertainty should be used to represent the random variable. Therefore, we calculate the entropy of both the discrete probability distributions to determine the least biased arrangement of LOS grades. This is the approach as taken in determining the least presumptive distribution given the information mean and SD.

The entropy of a discrete distribution can be denoted by $H(x)$, where x is the discrete variable where in our case, the variables are the LOS grades.

$$H(x) = -\sum_{i=1}^{n} P(x_i)\ln(P(x_i)) \qquad 3$$

where x_i are the variables and n is the number of variables – see Jordaan, (2005).

<u>Approach with six LOS grades</u>

$$H(x) = -\sum_{i=A}^{F} P(x_i)\ln\big(P(x_i)\big) \qquad 4$$

where $P_X(\gamma)$ is the probability density function of γ – the VoT.

For the LOS grade F,

$$P(x_F) = \int_{-\infty}^{\mu-2\sigma} P_X(\gamma)d\gamma \qquad (5)$$

where μ and σ are the mean and SD of the VoT distribution. Using the cumulative probability function for a normally distributed random variable, we can rewrite the Equation (5) with limits as follows.

$$\int_{-\infty}^{\mu-2\sigma} P_X(\gamma)d\gamma = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{\gamma-\mu}{\sqrt{2}\sigma}\right)\right]_{-\infty}^{\mu-2\sigma} \qquad (6)$$

Where the erf (.) represents the error function given as,

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}}\int_0^z e^{-t^2}dt \qquad (7)$$

Therefore,

$$P(x_F) = \frac{1}{2}\left[\operatorname{erf}\left(\frac{-2}{\sqrt{2}}\right) - \operatorname{erf}(-\infty)\right] = 0.02275$$

and thus

$$-P(x_F)\ln\big(P(x_F)\big) = -0.02275 \times \ln(0.02275) = 0.0861$$

In the same way,

$$P(x_E) = \int_{\mu-2\sigma}^{\mu-\sigma} P_X(\gamma)d\gamma$$

$$P(x_E) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{\gamma-\mu}{\sqrt{2}\sigma}\right)\right]_{\mu-2\sigma}^{\mu-\sigma} = \frac{1}{2}\left[\operatorname{erf}\left(\frac{-1}{\sqrt{2}}\right) - \operatorname{erf}\left(\frac{-2}{\sqrt{2}}\right)\right]$$

$$P(x_E) = 0.1359$$

$$-P(x_E)\ln\big(P(x_E)\big) = -0.1359 \times \ln(0.1359) = 0.2712$$

Accordingly,

$$-P(x_D)\ln\big(P(x_D)\big) = 0.3413$$

$$-P(x_C)\ln\big(P(x_C)\big) = 0.3413$$

$$-P(x_B)\ln\big(P(x_B)\big) = 0.2712$$

$$-P(x_A)\ln\big(P(x_A)\big) = 0.0861$$

Therefore, the entropy for the approach with six LOS grades is,

$$H(x) = -\sum_{i=A}^{F} P(x_i)\ln(P(x_i))$$

$$H(x) = 1.3972$$

<u>Approach with five LOS grades</u>

$$H(x) = -\sum_{i=A}^{E} P(x_i)\ln\big(P(x_i)\big) \qquad 8$$

For the LOS grade F,

$$P(x_E) = \int_{-\infty}^{\mu-\frac{3\sigma}{2}} P_X(\gamma)d\gamma \qquad 9$$

Therefore,

$$P(x_E) = \frac{1}{2}\left[\text{erf}\left(\frac{-3}{2\sqrt{2}}\right) - \text{erf}\,(-\infty)\right] = 0.06681$$

$$-P(x_E)\ln\big(P(x_E)\big) = -0.06681 \times \ln(0.06681) = 0.1808$$

Accordingly,

$$-P(x_D)\ln\big(P(x_D)\big) = 0.3432$$

$$-P(x_C)\ln\big(P(x_C)\big) = 0.3676$$

$$-P(x_B)\ln\big(P(x_B)\big) = 0.3432$$

$$-P(x_A)\ln\big(P(x_A)\big) = 0.1808$$

Therefore, the entropy for the approach with five LOS grades – from Equation (8) – is,

$$H(x) = 1.4156$$

Information entropy of the approach with five LOS grades is higher by 7% than that of six LOS grades. This means the approach with five LOS grades is less biased toward given information – statistics of the population of riders in the bus route.

Therefore, it can be recommended to adopt the approach with the five LOS grades – Figure 5 – for its lower bias in denoting the LOS.

It can be one of the goals for transit agencies to provide and maintain the LOS C as more than half the population might find it exceeding expectations. Having a LOS, A would mean that the provided service quality is higher than that of expected by 93.32% of the population. That is the operator is making the transit service preferable even to the topmost 6.69%. That is, the provided quality is much higher than the average expected quality. These passengers are mostly the choice or affluent riders and most of the time have other modes to take a trip. This would mean an over-utilization of the resources where the marginal return is at a minimum. The LOS E represents the opposite of this story. That is, the provided (or has been planned to provide) transit service is preferable to only the lowermost 6.69%. This portion of the population is mostly the set of people who are adversely affected by negative impacts of the service as they often have no other means of mobility than transit.

Combining the derived LOS for different attributes of the service and formulating a single LOS measure that represents the overall LOS of a bus route becomes challenging as different components are competing with each other. For example, higher LOS in access time can result in a lower LOS in trip speed as low stop spacing to improve access also increases the trip time as a higher number of stops delay the trip and vice versa. The outline of the overall process is shown in the Figure 7. Therefore, a mechanism will be proposed to combine the LOS of different aspects and come up with a unique LOS for the bus route.

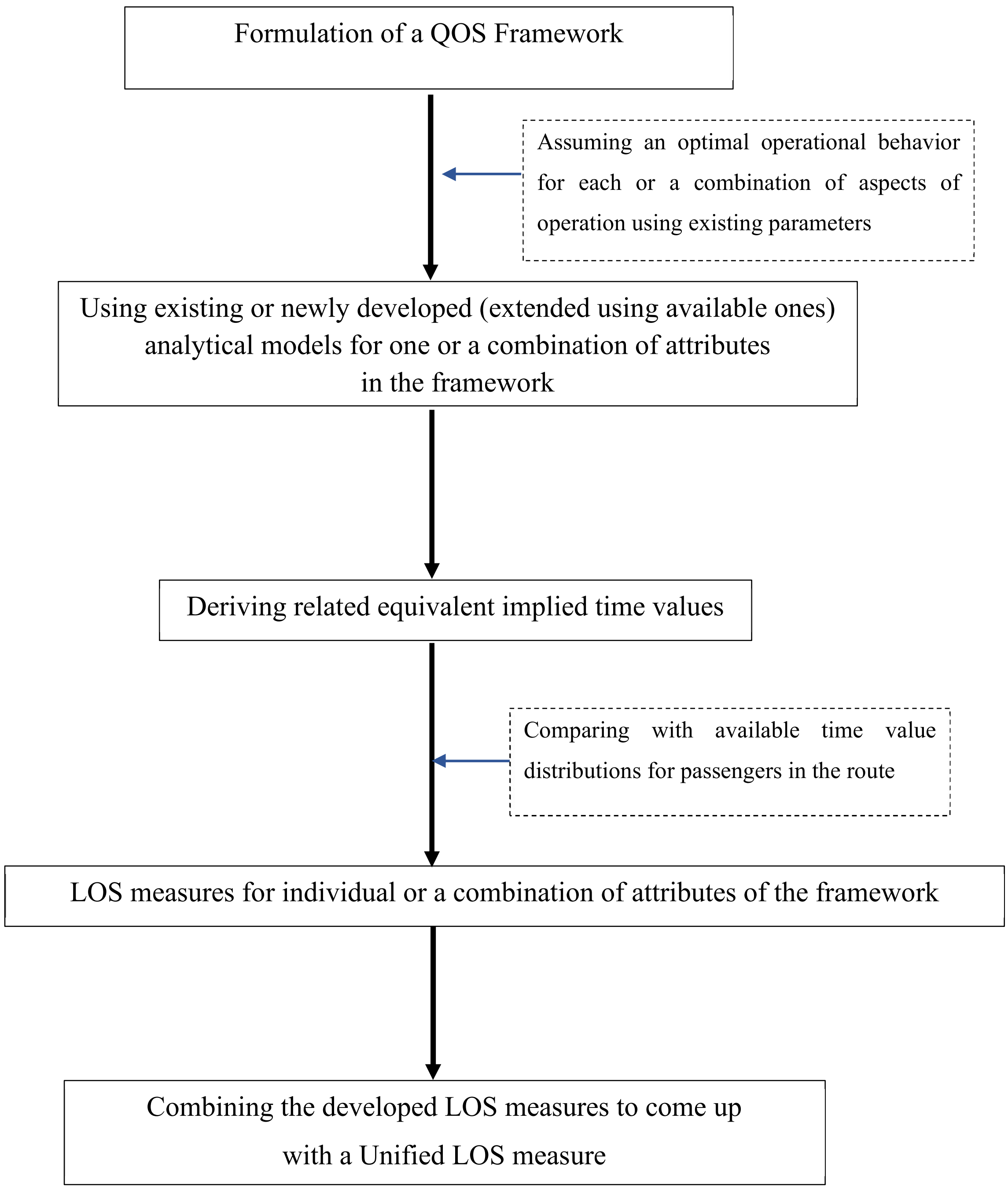


Figure 7 - A general framework of the proposed approach for the combined LOS measure for a bus route

2.4. Contributions and Novelty

Although there is a significant amount of work carried out in the field of LOS of urban bus transit, to the best of our knowledge, there are no studies proposing a method to assess the overall LOS of a transit route, with due attention to both passenger and operator concerns. This study, for the first time, formulates a LOS criterion to assess the overall LOS of an urban bus route which is its most important contribution. Furthermore, unique LOS criterions are derived for each or combinations of attributes of the LOS framework which includes both passenger and operator concerns.

The LOS criteria developed for an individual, or a combination of attributes of the service will provide transit agencies and the public with more useful and efficient insights depending on the unique passenger population of a transit service. Current methods usually provide universal measures in assessing these attributes which may not be representative of the unique passenger population of that transit service.

3. Chapter 3 - Headway Level of Service of an Urban Bus Route

3.1. Introduction

Due to the fact that measuring QOS is of utmost importance for a service provider, different transit agencies have developed their own measures over time. In the absence of a commonly accepted standard defining QOS in transit, the TCRP of the TRB developed the first edition of the TCQSM (TCRP Project A-15C, 2003b) in 1999 which specifies six designated levels of the QOS denoted by letters "A" (Best) through "F" (Worst). This addressed the convenience and availability aspects of transit stops, systems and route segments. In this widely accepted transit LOS framework suggested by the TCQSM, transit availability measures have shown a significant importance in both the perspectives of passengers and operators.

The three main measures for assessing transit availability in time and space are the service frequency, service span (the time period of a day that the service is available), and access. Attention in this chapter is mainly pointed towards the service frequency - the inverse of headway. Increasing headway of a service results in increasing average wait time for passengers making the service less preferable. Reducing wait times requires headway to be reduced and hence service frequency to be higher. This results in higher operational costs as more buses need to be dispatched per unit time and in some cases requiring a larger fleet size. It is assumed that a bus route operates under the condition where the tradeoff between the operations cost and average passenger wait time cost is minimized - optimum operation. A new methodology is presented in this chapter to relate LOS with service headway with and without the constraints of bus capacity using the distribution of the mean value of wait times (given in $/hr./passenger) of passengers. For this study, we assume the optimal operation with respect to the headway attribute (optimum service frequency), which is also well known as the 'square root dispatching policy' by G.F. Newell, (1971), as given. Nevertheless, the square root headway model is still derived in this chapter to support the rest of the work.

3.2. Background

3.2.1. Headway as a LOS Measure

As wait time has been found to be an important parameter in assessing the quality of service from a passenger perspective (dell'Olio et al., 2011; Turnquist, 1978), many studies have been conducted on obtaining a tradeoff between passenger wait time cost and operating cost through achieving an optimal frequency/headway (J. H. Banks, 1990; Jansson, 1980; G. Liu & Wirasinghe, 1991; Newell, 1971; Wirasinghe, 1990). Hensher relates bus frequency (reciprocal of headway) to a QOS measure called Service Quality Index (SQI) (Hensher, 2014; Hensher et al., 2003; Prioni & Hensher, 2001). His studies only describe the behavior of SQI depending on different values of bus frequency using passenger stated preference surveys.

TCQSM specifies six designated levels for QOS within a framework concerning availability and comfortability aspects of transit. Accordingly, TCQSM specifies six headway ranges for different levels of QOS (TCRP Project A-15C, 2003b) as shown in the Figure 8, for example – only two ranges are shown.

Average Headway	Passenger Perspective	Operator Perspective
11–15 min	• Relatively frequent service, but passengers will usually check scheduled arrival times to minimize their waiting time at the stop or station • Maximum desirable wait time for the next service if a bus or train is missed	• Often branded as "frequent service" in conjunction with long service hours, including weekends • Feasible in higher-density corridors (e.g., 15 dwelling units/net acre for bus service [*3*]), routes with strong anchors on both ends, and park-and-ride–based peak-period commuter bus service • Typically the longest feasible off-peak headway that would justify light rail or BRT service
16–30 min	• Passengers will check scheduled arrival times to minimize their waiting time • Passengers must adapt their travel to the transit schedule, often resulting in less-than-optimal arrival or departure times for them	• Typically provided as 20- or 30-min headways (e.g., 3 or 2 buses per hour) • Other headways can also be seen when traffic congestion increases bus running time, but budget not available to add service • Feasible in moderate-density corridors (e.g., 7 dwelling units/net acre for bus service [*3*]) • Typical commuter rail headway; longest commuter bus headway

Figure 8 - Headway ranges as specified by TCQSM (Source - TCQSM 3rd edition 2013)

3.2.2. User Costs through Square Root Policy

The concept of optimum service rates was first propounded in a study attempting to price public utilities based on marginal cost pricing (Vickrey, 1955). However, the pioneering work in the field of optimum transport service operations was carried out by G. F. Newell and Herbert Mohring (Mohring, 1972; Newell, 1971) with the introduction of the concept of the "square root dispatching policy". Although evidence is scarce on transit agencies adopting or not adopting such strategies, transit manuals can be found elaborating on the benefits of the use of such methods (Watson et al., 2002). Additionally, there exists a significant number of studies based on the pioneering work of Newell on square root dispatching policy (Wirasinghe, 1990) where Newell's dispatching policy has been applied to various types of problems ranging from, models of bus service to design of airport limousine services (J. Banks et al., 1982; Clarens & Hurdle, 1975; Pathak et al., 2020; Tirachini et al., 2010). Many authors have expanded on Newell's work on the square root rule (Ansari et al., 2017; Hawkins et al., 2020; Klumpenhouwer & Wirasinghe, 2016; G. Liu & Wirasinghe, 1991; Wirasinghe, 1990). The main objective of these attempts was to optimize the total cost of operation - user cost plus operator cost - to obtain the optimum operational parameters. Such outcomes are derived using closed form analytical models of related costs that provide valuable insights and functional relationships between variables. (Ansari et al., 2017; Hawkins et al., 2020; Klumpenhouwer & Wirasinghe, 2016; G. Liu & Wirasinghe, 1991; Wirasinghe, 1990). The main objective of these attempts was to optimize the total cost of operation - user cost plus operator cost - to obtain the optimum operational parameters. Such outcomes are derived using closed form analytical models of related costs that provide valuable insights and functional relationships between variables.

Models of optimal transit operation can be used to estimate 'perceived' user costs using several parameters of the service (Watson et al., 2002). Perceived user cost is the multiplication of the time spent by passengers with respect to a certain attribute of service and the value that a passenger would assign – perceived value – for a unit of that time spent. In the case of Headway LOS (HLOS), it is the average wait time of passengers, multiplied by the mean value of wait time per passenger. In an optimal operation, a transit agency seeks to minimize the cost of a user's time up to an amount that the operator assumes as the average value of time of passengers (Watson et al., 2002). In other words, the mean value of a unit of time of passengers derived from a model on an optimal operation can be depicted as the value that the operator is willing to trade off for a unit of time of an average passenger (Wirasinghe, 1990). Although transit agencies usually do not operate under optimal conditions, it is however possible to obtain service parameters – average demand, cost per bus dispatch, average headway – required to model an equivalent optimal operation with respect to the service attribute Headway. This way, it is possible to obtain the representative value of a unit of wait time that the operator would trade-off – the mean VoWT that an equivalent optimal operation will have – on behalf of passengers. The difference between the derived value and the values those passengers would associate themselves with implies a perceptional difference in the quality of the service experienced. This perceptional difference poses an opportunity to identify the quality of service provided depending on the characteristics of the ridership market. This approach is used in the proposed methodology to derive the HLOS measure and is described in detail in the following section.

3.2.3. Value of Wait Time (VoWT) Distribution

Every trip made by transit has a wait time component greater than or equal to zero incurred either at the beginning, at a transfer point or at the end of the trip depending on the trip purpose. The higher the amount of this unwilling component, the higher the disutility to make the trip. As this component directly affects the choice behavior of potential transit riders, it can be identified as a main element of the LOS provided by the transit service to its riders. Due to this reason, the wait time in transit has been studied widely to identify its effects on service characteristics such as frequency and reliability and hence on LOS (Ansari Esfeh et al., 2022; J. H. Banks, 1991; Hawkins et al., 2020; Shires & de Jong, 2009; Turnquist, 1978; Wardman et al., 2016).

What really affects the choice behavior of potential transit riders is the 'value' (cost) of total wait time rather than the 'amount' of wait time. Value of total wait time is a function of both the amount of wait time and the value of a unit of wait time. Therefore, wait time in transit has been explored in two major ways which are the amount of wait times and value of a unit of wait time as perceived by passengers. As referenced in the section 3.2.2 of this chapter, several researchers starting from Newell (1971) have attempted to find the effects of wait time on service attributes such as frequency/headway of service and reliability and to find the optimal frequency/headway that minimizes the total combined cost – user cost plus operator cost – in a route/network assuming different base conditions such as variable demand, fixed demand, different passenger arrival distributions to stops etc. Deviating from the traditional trend, (Ansari Esfeh et al., 2022) proposed wait time as different portions of headway depending on the proportions of transit user types and types of service (Ansari Esfeh et al., 2022). The value of a unit of wait time has also been widely studied as there were differences regarding the value of wait time as perceived by the passengers not being the same as perceived by the operators (Jansson, 1980).

Wardman presented a detailed review of the VoWT studies that has been conducted (Wardman, 2004). Initially, the studies have only concerned the value of in vehicle time for private transport modes. In addition, studies have been conducted on wait and walking times compared to in vehicle times of private modes where it was found that walking and waiting times are worth 2.5 times, on an average, more than in-vehicle times (Quarmby D. A., 1967). However, on average, wait time has been valued 2 to 2.5 times the in-vehicle time and it is valued more than any other time component of a transit trip (Mohring et al., 1987; Wardman, 2004). Methods of determining the values of time are RP and SP surveys, which were found to have effects on the results (Wardman, 2004).

The VoWT of a passenger varies depending on several factors such as income, age, and trip purpose. VoWT is more a personal taste unique to an individual that depends on many factors. Mohring et al., (1987) further pointed out that the value assigned for a unit of wait time in transit has a close relationship with income of the passengers and hence the average value of wait time differs at peak and off-peak times as there are less wage earners at off peak hours (Mohring et al., 1987). Mohring has done pioneering work deriving distributions for the values of wait time for transit riders under different conditions and for peak and off peak. A number of studies assume the value of time of passengers to be either normally distributed (Banks 1991; M. M. Rahman, Wirasinghe, and Kattan 2018) or log-normally distributed (Leurent, 1994; Mohring et al., 1987; Yang et al., 2001).

Determining the distribution of the mean VoWT of passengers can be difficult. Yet, it is possible to obtain key statistics of the distribution through widely conducted studies such as willingness to pay and stated preference surveys. For example, see (Hossain et al., 2015; Richardson, 2002). Most available key statistics from such studies are the mean and standard deviation. The maximum entropy distribution – minimizing the estimation bias – that one can estimate depending on these data (fixed data points) is the normal distribution (Jordaan, 2005). Therefore, for this study, we will assume the VoWT distribution of the passengers to be normally distributed.

3.2.4. Bus Capacity and Operating Cost

For operations under capacity, the size (capacity) of the bus can have an influence on the headway/frequency and hence on wait time. For this reason, it is required to identify what type of relationship exists between bus capacity (size) and operating costs to minimize tradeoffs between wait times and operating costs. A comprehensive analysis of bus operating costs is presented by

Jansson, (1980). He showed that 80% of the bus company costs are operating costs and stated the relationship between the bus capacity and operating cost is linear, and in the form of;

$$\mathrm{OC} = k + \alpha C \qquad 10$$

where OC represents the operation cost of a bus per unit time, C denotes the capacity (passenger carrying capacity) of the bus and k *and* α are constants. When the round-trip time including layover time of a particular bus route is given by 'T', the cost for a single dispatch of a bus in that route is given by;

$$\lambda_D = \mathrm{T}(k + \alpha C) \qquad 11$$

where λ_D represents the cost per bus dispatch.

3.3. Approach

Transit trips are produced by the transit agency contributing vehicles, labor, resources and passengers contributing their time and transit fare. Just as the operation of public transit has a cost, passengers also associate a cost for their time spent making a trip which depends on their value of time for each component of the trip – time spent making a part of the trip multiplied by the mean value of a unit of time of a passenger for that part of the trip. E.g., Accessing transit, riding, waiting, and transferring have different values for a unit of time spent in each of these activities. The conditions – operational characteristics such as the values of frequency, passenger demand, speed, stop distance, etc. – under which the demand for transit trips is met, can be used to approximately estimate the costs incurred by users and operators in a given operation. Optimum operations that minimize the tradeoffs between total user cost and operator cost also minimize the total combined cost of operator and users. The transit fare cancels out here as it is a cost to passengers and a revenue for the operator so that the total cost is unaffected. Therefore, the return on investment of public funds – i.e., the operator costs and user costs – is highest when the total cost is the lowest. Therefore, a significant amount of transit literature is based on optimized operations. Despite the academic interests of optimized operations, it is not clear that the transit industry has incorporated such concepts into their practice. However, we are using an equivalent optimum operation for the same operational parameters of an existing service to estimate the implied value of wait time of passengers to relate it to the HLOS. With this intention we first explore analytical aspects of an optimized bus route under different, close to real, base conditions that will be used later in denoting the HLOS of a bus route.

3.3.1. Case I – A bus route with a demand that is constant with time

Consider a bus route with a many-to-many demand of P passengers per hour for the whole route where buses are never filled – i.e., demand is less than the capacity of the route.

Let the passenger demand to board a bus (number of passengers that needs to get into the bus) at the stop i at time t given in passengers per hour be $P_i(t)$. If the time taken from the origin to station i is given by t_i, the total passenger demand to board the entire route can be denoted as a demand at the origin. Let the total passenger demand to board – (hereafter referred to as the "demand") the bus route at time t be given by

$$P(t) = \sum_{i=0}^{n} P_i(t + t_i) \tag{12}$$

where n is the number of stops in the transit line.

When the demand at the bus stops is fixed over time, Equation (12) becomes,

$$P = \sum_{i=0}^{n} P_i \tag{13}$$

The frequency of the bus route is known, passenger arrivals at the bus stops are random. We assume that the bus route has a fixed headway over the period considered. The headway of the bus route is denoted by H hours, cost of dispatching a bus in this route by λ_D dollars per dispatch, and the mean value of wait time of passengers in the route by γ_w dollars per hour per passenger.

We assume the bus route operation is frequency based, and not schedule based. In frequency-based bus routes – normally bus routes with a shorter headway than ten minutes – the passengers are not checking the schedule – sometimes the schedule is not even published (Ansari Esfeh et al., 2022). Therefore, the passenger arrivals to the bus stop are random. The average wait time for a passenger depends also on the type of passenger – whether they are planning or non-planning/having a fixed arrival time or flexible arrival time. As described by Ansari et. al (2022), calculating the average wait time with due attention to these factors yields a better approximation and is data intensive. To keep the methodology simple, the type of passenger has not been considered leading to the assumption that all the passengers on the bus route are non-planning and hence flexible on their arrival time. The methodology presented can however be extended to include the passenger type and type of service – whether frequency based, or schedule based – using the suggested modifications as described in the Appendix A.

Average wait time of a passenger = $H/2$

When the passenger arrivals to the stop are random, the maximum wait time a passenger can experience is H, and the minimum wait time a passenger can experience is zero. Therefore, the average wait time is half the headway.

Average hourly wait time cost for the passengers on the route;

= (Avg wait time per passenger) × (Value of passenger wait time) × (Passenger demand)

$= \left(\frac{H}{2}\right)\gamma_w \mathrm{P}$

Hourly operating cost of the route;

(Cost per bus dispatch) × (number of bus dispatches per hour) = $\lambda_D \left({}^{1}\!/_{H}\right)$

$= {}^{\lambda_D}\!/_{H}$

Therefore, the total cost, Z – passenger cost plus operating cost

$$Z = \left(\frac{H}{2}\right)\gamma_w P + {}^{\lambda_D}\!/_{H} \tag{14}$$

The minimum total cost operations in terms of headway take place when $\frac{\partial Z}{\partial H} = 0$;

Therefore, the optimum headway that minimizes total cost is given by H' as follows by taking the derivative of Z by H (λ_D, P, γ_w *assumed constant*)

$$H' = \sqrt{\frac{2\lambda_D}{P\gamma_w}} \tag{15}$$

The second derivative of the Equation (14) can be shown as;

$$\frac{\partial^2 Z}{\partial H^2} = \frac{2\lambda_D}{H^3} \tag{16}$$

As λ_D and H are positive values, the second derivative in Equation 16 is positive. Therefore, the total cost function is strictly convex and the value of headway at the global minimum gives the optimum headway.

For an existing operation of a bus route, it is possible to observe the value of headway, and obtain the value of λ_D and P from the transit agency. Using these values, it is possible to derive the implied VoWT of passengers for an equivalent optimal operation from the Equation (15) as follows.

$$\gamma_w = \frac{2\lambda_D}{PH'^2} \tag{17}$$

Implied VoWT is not necessarily the actual mean value of wait time of passengers. It can be interpreted to be used as an indicator of the quality of service provided (Wirasinghe, 1990), and this can be measured based on the properties (e.g.- mean VoWT distribution of riders) of the ridership population. It is not essentially the QOS experienced by the passengers nor the QOS provided as perceived by the operator. It is a measure of the combined QOS in comparison to an optimized bus operation with the same parameters, i.e., demand, cost of dispatching, and headway. The higher the implied VoWT than the mean of the VoWT distribution of passengers, the higher the QOS and hence the LOS provided.

3.3.2. Case II – A bus route with a demand varying with time

Consider an urban bus route where the buses are never full – demand of the route is always lower than the capacity of the route and varies with time - $P(t)$. The demand, as derived in Equation (12), is used here. Therefore, the optimum headway, which is a function of demand, also becomes a function of time – varies with time. However, it is common for transit agencies to use a constant headway during a low demand period where line capacity is not reached (normally in off-peak operations). Therefore, we assume a constant headway for the round-trip time 'T'. The total cost of the bus route – given round trip time 'T' – can then be expressed as;

$$\int_T Z(t)\,dt = \int_T \frac{1}{2}\gamma_w HP(t)dt + \int_T \frac{\lambda_D}{H}dt \tag{18}$$

Equation 18 can be further simplified as follows.

$$= \frac{1}{2}\gamma_w H \int_T P(t)dt + \frac{\lambda_D}{H}T$$

$$= \frac{1}{2}\gamma_w H\bar{P}T + \frac{\lambda_D}{H}T$$

Taking the derivative in terms of H to obtain the optimum headway,

$$\frac{1}{T}\frac{\partial}{\partial H} = \frac{1}{2}\gamma_w\bar{P} - \frac{\lambda_D}{H^2} = 0$$

$$H^*(Optimum\ Headway) = \left[\frac{2\lambda_D}{\gamma_w\bar{P}}\right]^{\frac{1}{2}}$$

$$\gamma'_w(implied\ VoWT) = \left[\frac{2\lambda_D}{\bar{P}H^{*2}}\right] \qquad 19$$

where $\bar{P}$ is the mean passenger demand during the period T.

Appendix A derives the implied VoWT for two routes with equal headways and time varying demands and discusses the extension to multiple routes. The derivation of the implied VoWT to denote the HLOS for the case of stochastic passenger demand is discussed in the Appendix A.

3.3.3. Effects of demand, cost of dispatch, and VoWT on HLOS of a bus route

(1) Demand

Consider a bus route operating with P_1 and P_2 average demands given in passengers per hour for the line (P_1>P_2) for the time periods T_1 and T_2 respectively (T_1 and T_2 are larger than round trip time) and the bus route does not reach the demand capacity during any of these time periods. Note that here the dispatch cost (λ_D) of the bus route is the same for both peak and off-peak operations and the operator utilizes the same headway for both the time periods T_1 and T_2. Therefore, according to Equation (19), the implied VoWT for the operation during the time period T_1 – having a demand of P_1 – is smaller than for the operation during the time period T_2 – having a demand of P_2. That means, in a particular bus route, if the operator decides to keep the headway the same artificially for time periods with different demands where the demand varies during off-peak yet not reaching the line capacity, the implied VoWT and hence the Headway LOS varies with demand. The higher the demand, the lower the γ'_w and hence Headway LOS. A passenger will not notice any difference in the Headway LOS that is experienced. Therefore, the Headway LOS measured by the implied VoWT here is not entirely about the passenger perception.

In an overall operational sense, increasing the demand will increase the contribution (cost) of passengers – $\int_T \frac{1}{2}\gamma_w HP(t)dt$. If the bus route is keeping the same operational parameters, irrespective of the increased demand, the contribution of the operator - $\int_T \frac{\lambda_D}{H} dt$ stays the same as headway and dispatch costs are constant. Generalizing this unequal state of contributions to assume an optimum condition where user cost is made equal to operator cost, underestimates the contribution from passengers - underestimate the average VoWT while demand increases as the time component of the trip does not change – in a way that total passenger cost is made equal to the operator cost. Therefore, the implied HLOS can be categorized as low from an overall passenger perspective because now when the demand is higher the operator undervalues the VoWT of passengers in an equivalent optimized operation. This measure can therefore be identified as a way of measuring HLOS in an overall perspective where both passenger and operator concerns are taken into account.

(2) Cost of dispatch

Impact of the dispatch cost on the implied VoWT and hence on the Headway LOS can be described in a similar manner to the impact of demand as described above. Assume in a certain bus route, the operator decides to utilize a larger bus size with a higher capacity while all the other parameters stay the same (i.e.: Demand, Mean VoWT and Headway). As it is obvious that with larger buses now the fuel costs and discounted capital costs are higher, the corresponding dispatch cost becomes high. Therefore, the hourly operating cost in the bus route - $\int_T \frac{\lambda_D}{H} dt$ is now higher than it was before. To assume an optimum operation, the equivalent total passenger cost - $\int_T \frac{1}{2}\gamma_w HP(t)dt$ now must be high and equal to that of hourly operating cost – two terms in the right side of Equation (18) should be equal. As demand is the same, the only way it can be done is to assume a higher average VoWT for passengers. The operator functioning with the same headway with a larger bus size hence a higher hourly operating cost means that the implied VoWT of passengers in an equivalent optimum operation is now higher. The same change can be explained using Equation (17) as well. This is also an operational change that a single passenger alone will not experience about the service being offered in terms of headway.

(3) Headway;

As opposed to the previous two parameters, Headway has a squared effect on implied VoWT (Equation (17)) and hence on Headway LOS. In a bus route where all the parameters except headway stay the same, increasing headway will increase the total passenger contribution (cost) - $\int_T \frac{1}{2}\gamma_w HP(t)dt$, while also reducing the hourly operating cost - $\int_T \frac{\lambda_D}{H} dt$. Now the difference between total passenger cost and operator cost has increased - Difference between the two terms of the right side of Equation (18). In an assumed equivalent optimum operation where operator and passenger costs are equal, the equivalent pseudo VoWT has to be significantly reduced to bridge the gap between user costs ($\int_T \frac{1}{2}\gamma_w HP(t)dt$) and operator costs ($\int_T \frac{\lambda_D}{H} dt$) resulting in a severe impact on Headway LOS. Effect of headway changes affects the individual passenger perception while also affecting the overall perspective on headway LOS of the bus route. The derivation of the implied VoWT to denote the headway LOS considering the effects of headway variation is presented in the Appendix A.

3.3.4. Operations under capacity

Consider a bus route with fixed one-to-many or many-to-one demand with a fixed headway over the period of calculation. There is a limit for the optimum headway imposed due to the limited capacity (person capacity) of the buses. TCQSM (third edition) outlines that the transit agencies adopt a design person capacity of a bus ranging from 1.25 – 1.5 times the seated capacity – number of seats in a bus, depending on the quality that the transit agency needs to provide (TCRP Project A-15C, 2013).

During higher-than-average demand, e.g., peak periods, dispatching buses according to the optimal headway might not be enough to meet demand, which leads to some passengers not being able to board the bus. Given the demand for the route is P passengers per hour for the entire route, and the capacity of a bus is 'C' number of passengers per bus, the minimum rate of bus dispatches required to meet the passenger demand can be derived as below;

Minimum rate of bus dispatches required to meet the demand = P/C per hour

Utilizing this rate of bus dispatches will transport all the passengers in the stations without leaving anyone behind. Otherwise, left behind passengers need to wait for another headway causing the total cost of the system to rise significantly – the average wait time of the passengers would be greater than half the headway. Therefore, the rate P/C has to be adopted whenever the rate corresponding to the optimum headway is less than P/C.

The operator can increase the number of buses in operation per hour. We assume the operator has buses only with the capacity of C. Therefore, the maximum time interval - headway - that the operator can have between consecutive bus dispatches to at least serve the demand is given by;

Maximum headway = $\frac{Hour}{Minimum\ number\ of\ busses\ required} = \frac{1}{P/C} = \frac{C}{P}$

This headway can be also called the Capacity Headway - H_c.

$$H_c = \frac{C}{P} \qquad 20$$

Therefore, any headway that can be utilized by the operator needs to be less than H_c.

The dispatching policy introduced by Newell (1971) and Wirasinghe (1990) does not allow passengers to be left behind. According to Newell (1971), the time varying headway under capacity operations for a many-to-one time-varying demand can be presented as

$$H_c(t) = \frac{C}{P(t)} \qquad 21$$

where C is the design capacity of the bus. In a many-to-one or one-to-many bus route, each passenger-space is used by one passenger only or not used at all. But, in a many-to-many bus route, a passenger-space can be used by several passengers in series. Therefore, to be used in a many-to-many bus route, Equation (21) can be modified by replacing P(t) with the demand in passenger-spaces per hour Q(t) for the bus route (Wirasinghe, 1990) as given by the number of passenger spaces per hour.

$$H_c(t) = \frac{C}{Q(t)} \qquad 22$$

Wirasinghe (1990) shows how Q(t) can be derived for a bus route using boarding and alighting data. It is important to note here that Q(t) is different from the load of a bus at time t. He further shows that, in the common case of a bus route having a fixed maximum load point, Q(t) is equal to the rate at which the passengers pass the maximum load point and can be measured by placing an observer at the maximum load point.

To obtain a uniform headway for a time period, e.g., peak period, we can use the average value of the demand in passenger-spaces per unit time at the maximum load point Q. Hence, we have

$$H_c = \frac{C}{Q} \tag{23}$$

where H_c is the uniform headway under capacity operations during peak-period.

Although Q is defined as a mean value, demand varies with time as $Q(t)$, and capacity C must be able to deal with demand at the height of the peak. Thus, C must be chosen carefully as the sum of the seated capacity plus only a portion of the standing capacity. At the height of the peak, the bus is then still able to maintain the same uniform headway, but with a higher load, and does not leave passengers behind.

The normal industry practice is to use a "policy-headway," h, which is subjective to a transit agency, as an upper limit of the allowable headways (Wirasinghe, 1990). The intention is to limit the maximum waiting time of passengers in the transit system below a given threshold. That is, even though there is minimal demand, which makes the optimal headway significantly high, i.e., $H' > h$, the transit agency will run buses using h. Having a policy headway ensures a minimum quality of service provided by transit agencies because passengers are aware that the waiting time will not be longer than a given value under any circumstances, which creates trust.

Dispatching policy is to utilize the minimum of optimum headway - Equation (15), capacity headway - Equation (23), and policy headway. Accordingly, the dispatching policy H_p can be presented as

$$H_p = Min\begin{cases} H' \\ H_c \\ h \end{cases} \tag{24}$$

This formulation is carried out for a single line of an urban public bus route with many-to-many demand where fleet size (number of buses available to be dispatched on the line) is not a constraint, passengers arrive randomly at bus stops, and passengers are served by the first arriving vehicle (i.e., passengers are not left behind).

As far as the headway being used in the bus route is less than the capacity headway, the HLOS of the bus service can be related to the implied VoWT from the methods described in sections 3.1 and 3.2. Under this section, a new method is sought to determine the LOS of a bus service when it is operating under the capacity headway.

As the aim of the bus operator is to minimize social costs and the operator is also looking to minimize the operator's own costs, it can be assumed reasonably that the operator would use a bus size (capacity) that minimizes the total cost (user + operator).

For a many-to-many bus route operating under the capacity headway, the total cost function can be presented as follows; (by substituting $H = C/Q$, and $\lambda_D = \mathrm{T}(k + \alpha C)$ in Equation (14))

$$Z = \frac{P\gamma_w C}{2\mathrm{Q}} + \frac{QT(k + \alpha C)}{C} \tag{25}$$

Total cost would become a minimum when $\frac{\partial Z}{\partial C} = 0$ and if $\frac{\partial^2 Z}{\partial C^2} > 0$;

As $\frac{\partial^2 Z}{\partial C^2} = \frac{2kQT}{C^3}$; $\frac{\partial^2 Z}{\partial C^2} > 0$ ($k, Q\ and\ C$ are greater than zero). Therefore, for the minimum total cost,

$$\frac{\partial Z}{\partial C} = \frac{P\gamma_w}{2\mathrm{Q}} - \frac{QTk}{C^2} = 0 \tag{26}$$

Provides;

$$C = \sqrt{\frac{2kQ^2T}{P\gamma_w}} \tag{27}$$

Therefore, capacity of a bus operating under optimal capacity headway conditions can be presented as a function of the mean value of wait time of passengers. $C = f(\gamma_w)$.

Accordingly, the implied VoWT for any operation under capacity headway of a bus route can be expressed as;

$$\gamma'_w = \frac{2kQ^2T}{PC^2} \tag{28}$$

3.3.5. Peak period headway level of service

During the peak period, a bus route can operate under; (a) normal conditions – a predetermined headway (b) capacity headway (c) a combination of both the pre-determined and capacity headway. In a bus route where the demand is less than the line capacity all the time, it can run according to a pre-determined headway throughout the peak period. In this situation, the HLOS can be referred to using the implied VoWT that can be derived using ${2\lambda_D}/{P(t)H^2}$. If the demand of the line is always greater than the capacity of the line, the buses have to be dispatched using the capacity headway – this varies and is always lower than the optimum headway. In such a situation, the HLOS varies with the demand of passenger spaces per hour and the HLOS can be referred to using the implied VoWT that can be derived using Equation (28). The demand of a bus route can change significantly during the peak period, especially in a bus route mostly running through a downtown area. This can lead to a situation where the operator has to operate the bus route under both a predetermined headway and a capacity headway. In this section, we seek to develop a methodology to derive HLOS for the peak period.

It can be fairly assumed that the operator runs the line with a pre-determined policy headway (not variable) until the demand reaches the line capacity. A transit agency does not have the ability to adjust the bus size (capacity) with the varying demand during peak time to maintain a given HLOS when the demand is higher than the capacity. Continuing with the same bus size by just adjusting the headway when the existing demand cannot be catered with the policy headway, will lead to a bus size (C') running under capacity condition which is not equal to the optimum bus size C. Allocation of a non-optimal bus size C' changes the implied VoWT for the bus route. A lower implied VoWT according to the Equation (28) will imply a lower LOS in the transit route and vice versa.

We assume that the demand curves in terms of boardings per hour and seats per hour for the bus route are available. It is important to determine the critical demand for a given bus size C when the bus route starts to run under capacity condition.

$$Q' = \frac{C}{H} \tag{29}$$

Q' denotes the demand of passenger spaces at which the bus route starts to run under capacity headway. C denotes the bus size. H denotes the pre-determined headway the operator has been operating the bus route with. When the passenger space demand of the bus route $Q(t)$ is greater than Q', the bus route runs under the capacity headway.

Let's assume, for a certain bus route, the peak period starts at time t_S and finishes at t_F. The bus route starts to run under capacity headway at time t_{CS} and finishes at time t_{CF}. The bus route is assumed to be operating under a predetermined headway H. This headway does not vary with time or the variable demand until the demand is large enough so that the bus route has to be operated under capacity conditions to avoid passengers being left behind. Beyond this demand, the bus route operates with capacity headway - C/Q(t) - which varies with time. The demand beyond which the bus route operates under capacity headway is denoted by Q'. The bus route is assumed to have a demand profile during the peak period T' as follows.

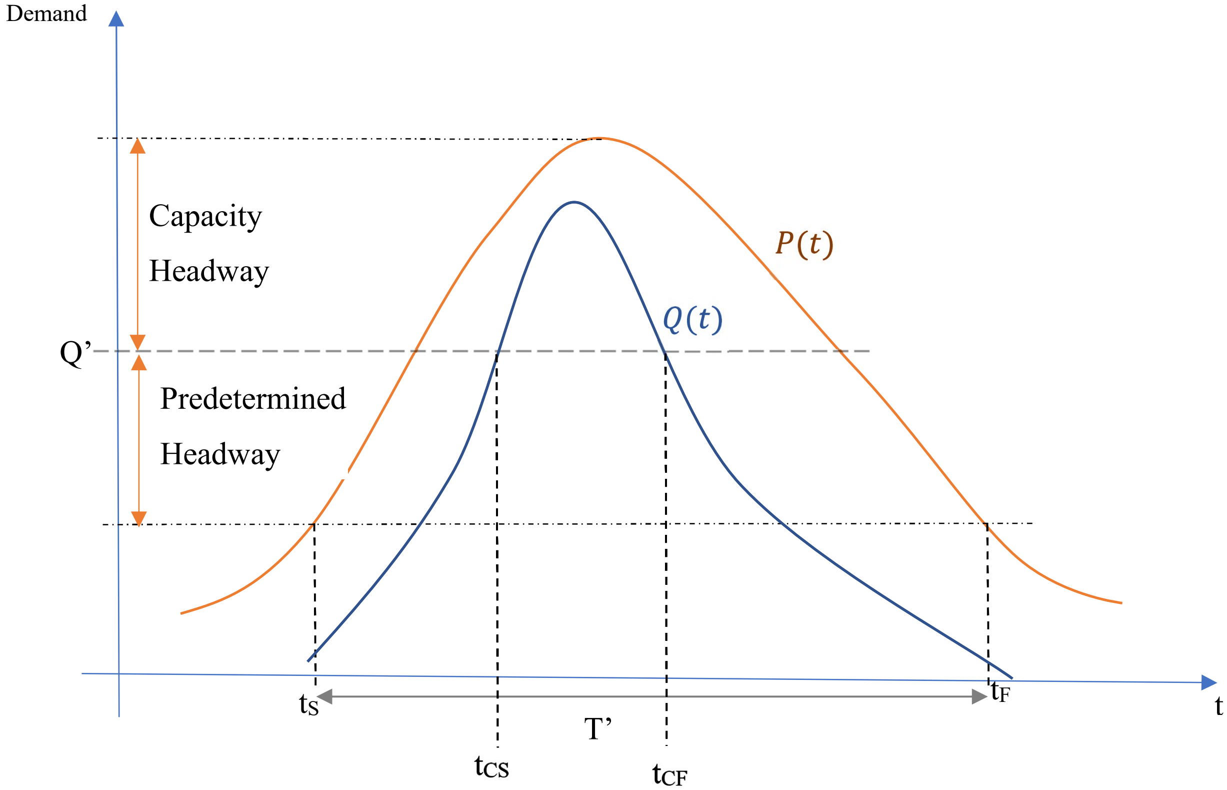


Figure 9 - Peak period demand profile of the bus route

During normal operation, the implied VoWT can be expressed using Equation (17), and during the operations under capacity headway the implied VoWT can be expressed using Equation (28). Therefore, we can derive the average implied VoWT for the peak period operation by multiplying the implied VoWT of the bus route operation for a given moment by the number of passengers affected by that implied VoWT at that particular moment for the peak period and dividing that amount by the total number of passengers affected by the varying HLOS throughout the peak period. It is important to note here that the number of passengers affected by the implied VoWT during an infinitely small time period 'dt' - we are assuming the implied VoWT stays constant during time period 'dt' as the change in demand is small - is the number of passengers boarding the bus route during that particular time period 'dt'. This is because the HLOS only affects the wait time of passengers who are boarding. Accordingly, the average implied VoWT of the bus route during the peak period can be obtained as follows.

$$\gamma'_{w,avg} = \frac{\int_{t_S}^{t_{CS}} \frac{2\lambda_D}{H^2 P(t)} P(t)dt + \int_{t_{CS}}^{t_{CF}} \frac{2kQ(t)^2 T}{C^2 P(t)} P(t)dt + \int_{t_{CF}}^{t_F} \frac{2\lambda_D}{H^2 P(t)} P(t)dt}{\int_{t_S}^{t_F} P(t)dt} \qquad 30$$

The above expression can be simplified as below.

$$\gamma'_{w,avg} = \frac{\frac{2\lambda_D}{H^2}[T' - (t_{CF} - t_{CS})] + \frac{2kT}{C^2} \int_{t_{CS}}^{t_{CF}} Q(t)^2 dt}{\int_{t_S}^{t_F} P(t)dt} \qquad 31$$

Where $T' = t_F - t_S$, which is the duration of the peak period. When $\overline{P_P}$ is the average demand for boarding during the peak period T', the term $\int_{t_S}^{t_F} P(t)dt$ can be replaced by the term $\overline{P_P}T'$.

We can relate the corresponding $\gamma'_{w,avg}$ to the HLOS using the methods outlined in the section 3.4 of this chapter.

3.3.6. Example

A numerical example has been designed to demonstrate the methodology developed to obtain the implied VoWT for a peak hour bus operation. Consider a bus route operation during peak period with the following characteristics.

Table 3 - Operational parameters for the bus route

Parameter	Value
λ_D	100 ($/dispatch)
H	0.333 (hrs.) = 20 min
C	50 passenger spaces
K	100 ($/hr./dispatch)
T	1.1 (hrs.)

The transit agency plans to operate this bus route with a pre-determined headway (H) of 0.333 hrs. (20 minutes) during the peak period – 16:00-20:00 hours. The demand for the bus route, however, varies significantly during the peak period with a minimum demand of 100 passenger spaces per hour and a maximum demand of 200 passenger spaces per hour. As the passenger capacity of the buses used on the bus route is 50 passenger spaces, the operational capacity that this bus route can achieve with a policy headway of 0.333 hrs. is 50/0.333 = 150 passengers per hour. The variation of the type of headway used with the demand is described below.

Table 4 - Variation of predetermined headway and capacity headway with demand

Q(t)	Predetermined Headway (H)	Capacity Headway (H_c)
100	**0.333**	0.500
110	**0.333**	0.455
120	**0.333**	0.417
130	**0.333**	0.385
140	**0.333**	0.357
150	**0.333**	**0.333**
160	0.333	**0.313**
170	0.333	**0.294**
180	0.333	**0.278**
190	0.333	**0.263**
200	0.333	**0.250**

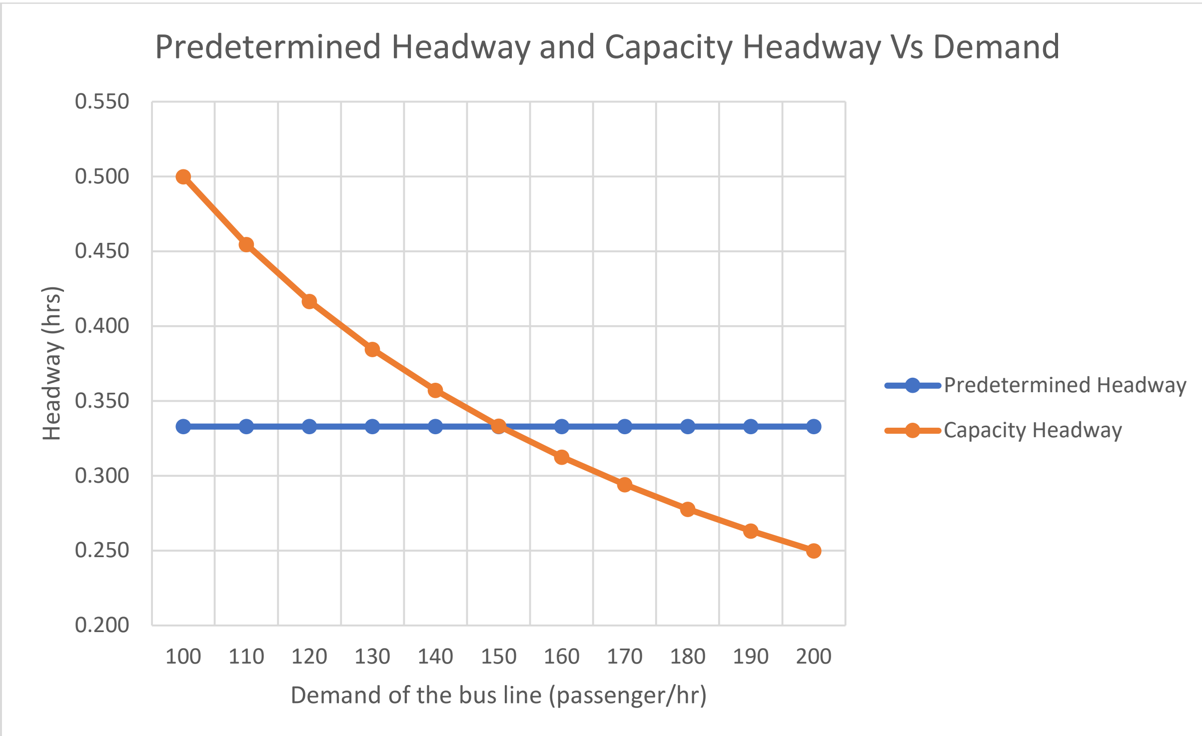


Figure 10 - Variation of predetermined and capacity headways with the demand of passenger spaces

The normal practice is to use the minimum headway out of these two headways (Newell, 1971; Wirasinghe, 1990). According to Figure 10, the bus route can be operated with the predetermined headway up to a demand of 150 pass. spaces/hr. When the demand is higher than 150 pass. spaces/hr., the bus route has to be operated according to the capacity headway (H_c). As it is evident here, the implied VoWT now has to be obtained using two methods separately for capacity operation and normal operation. The implied VoWTs for corresponding demand values can be obtained and described below.

As an example, for the normal operation, the implied VoWT can be obtained using Equation (17) with respect to an optimized operation. We consider a boarding demand P(t) when Q(t) is less than 150 passengers/hr. as an example. The boarding demand P(t) will always be higher than or equal to the demand for passenger spaces.

When P(t) = 140 pass. /hr.,

$$\gamma'_w = {}^{2\lambda_D}/_{PH^2} = \frac{2 * 100}{140 * 0.333} = 12.88\ (\$/hr - pass.)$$

Now let's assume an example of capacity operation when the demand for passenger spaces Q(t) is higher than 150 pass./hr.. In this instance, the implied VoWT can be obtained using the Equation (28) with respect to an optimized capacity operation.

For this example, we assume the boarding demand P(t) and demand for passenger spaces Q(t) to vary as shown in Figure 11 and Equations (32) and (33) that conforms with the minimum and highest demands stated previously in the text.

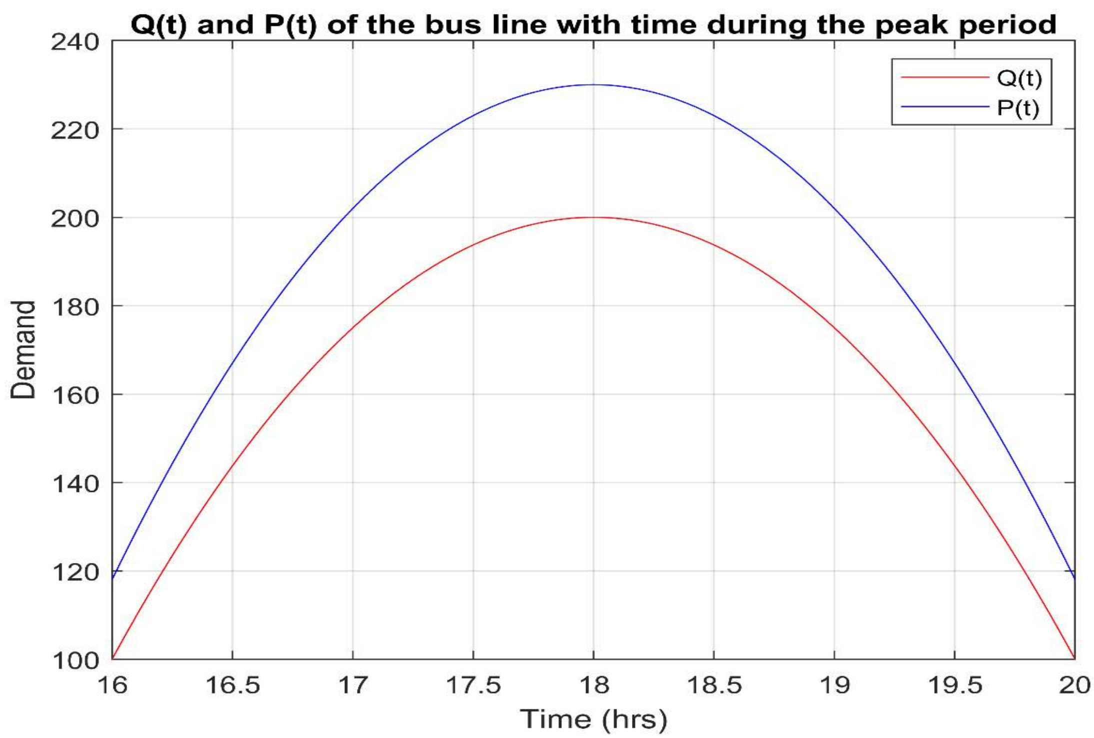


Figure 11 - Variation of Q(t) and P(t) during peak period

$$Q(t) = 200 - 25(t - 18)^2 \qquad 32$$

$$P(t) = 230 - 28(t - 18)^2 \qquad 33$$

When Q(t) = 160 pass. spaces/hr., P(t) = 185 pass. /hr.

$$\gamma'_w = \frac{2kQ^2T}{PC^2} = \frac{2 * 100 * 160^2 * 1.1}{185 * 50^2} = 12.18\ (\$/hr - pass.)$$

Therefore, it is evident that different passengers getting onto the transit system during different time intervals having different demand volumes will experience different γ'_w (implied VoWT) values and hence different HLOS. And the variation of this value with the demand is different with respect to the type of operation of the bus route – normal operation or capacity operation. In order to represent the HLOS of the transit line with one value for the peak period, we need to find a representative value for the average implied VoWT for the passengers during peak period. We pursue an approach where an average value for the implied VoWT of the passengers riding the transit line during the peak period is obtained by multiplying the number of boarding passengers by their corresponding implied VoWT using the Equation (31).

According to Equation (32) and Figure 10, the time at which the bus route starts to run under the capacity headway t_{CS} – the time when the demand for passenger spaces is higher than the capacity (150 pass. spaces/hrs.) is 16.58 hours. Also, the time when the bus route finishes running under the capacity headway t_{CF} – the time when the demand for passenger spaces becomes lower than the capacity (150 pass. spaces/hr.) is 19.41 hours. Therefore, the bus route will operate with the capacity headway from 16.58 hours to 19.41 hours. The peak period starts at 16:00 hours, will end at 20:00 hours, and therefore the length of the peak period T' is 4 hours.

Accordingly, the representative average implied VoWT can be obtained as follows;

$$\gamma'_{w,avg} = \frac{\frac{2\lambda_D}{H^2}[T' - (t_{CF} - t_{CS})] + \frac{2kT}{C^2}\int_{t_{CS}}^{t_{CF}} Q(t)^2 dt}{\int_{t_S}^{t_F} P(t)dt}$$

$$\int_{t_{CS}}^{t_{CF}} Q(t)^2 dt = \left[125t^5 - 11250t^4 + \frac{1205000}{3}t^3 - 7110000t^2 + 62410000t\right]_{16.58}^{19.41}$$

$$\int_{16.58}^{19.41} Q(t)^2 dt = 95729.96$$

Here, $\int_{t_S}^{t_F} P(t)dt$ is also equal to $\overline{P_P}T'$ where $\overline{P_P}$ is the mean demand for boarding during the peak period. For the example that we are considering, we will adopt the expression $\int_{t_S}^{t_F} P(t)dt$ directly as we know the functional form $P(t)$ and the integral limits. Accordingly,

$$\int_{t_S}^{t_F} P(t)dt = \left[-\frac{28}{3}t^3 + 504t^2 - 8842t\right]_{16}^{20} \cong 771$$

Therefore,

$$\gamma'_{w,avg} = 13.5\ (\$/hr - pass.)$$

We assume the operator has the mean and the standard deviation of the distribution of the mean value of wait time of passengers. Therefore, it is possible to obtain the corresponding HLOS for the transit route during the peak period using the obtained $\gamma'_{w,avg}$ value according to the methods described in the section 3.4 in this chapter.

3.4. Headway LOS and the Probability Distribution of Mean Value of Wait Time

3.4.1. A Letter Grade Scale to Indicate LOS

The VoWT of passengers varies depending on several reasons such as income, age, trip purpose etc. VoWT is more a personal taste unique to an individual that depends on many factors. Therefore, in an apparently homogeneous market, it is possible to assume that VoWT of passengers are normally distributed with a unique mean (μ) and a variance (σ^2) (J. H. Banks, 1991; Rahman, Kattan, et al., 2018). However, when the ridership market deviates from being apparently homogeneous, researchers have also assumed VoWT be log-normally distributed (Mohring et al., 1987; Yang et al., 2001). As discussed in section 3.2.3 in this chapter, the most widely available statistics for transit agencies from willingness to pay studies and stated preference and revealed preference studies, conducted as part of the demand estimation processes, are the mean and standard deviation of the probability distribution of the mean VoWT of the passengers. The shape of the distribution is mostly not available. In such circumstances, the shape of the distribution that carries the maximum information entropy - therefore minimizes the risk of errors in the prediction (minimum bias) is the normal distribution (Jordaan, 2005). As it is fair enough to assume a normal distribution for the mean VoWTs of the passengers, this study utilizes a normally distributed VoWTs of passengers to relate to the HLOS. Therefore, depending on where the found implied VoWT falls in the distribution with respect to the mean 'μ' and the standard deviation 'σ', the HLOS can be determined. In this case, the implied VoWT can be referred through characteristics of the actual VoWT distribution with the use of a constant 'k' (k>0) as follows;

$$\gamma'_w = \mu + k * \sigma \qquad 34$$

$5/2 \leq k < 3/2 \quad \rightarrow \quad LOS\ A$

$1/2 \leq k < 3/2 \quad \rightarrow \quad LOS\ B$

$-1/2 \leq k < 1/2 \quad \rightarrow \quad LOS\ C$

$-\frac{3}{2} \leq k < -1/2 \quad \rightarrow \quad LOS\ D$

$k < -3/2 \quad \rightarrow \quad LOS\ E$

A similar approach that was briefly suggested by Wirasinghe and Vandebona, where a value with one standard deviation higher than the mean value of time, denotes a one LOS higher (Wirasinghe & Vandebona, 2011), has been utilized in referring to a certain headway LOS. Also, the importance and reasoning behind differentiating LOS letter grades depending on the standard deviation (σ) of the distribution and the arrangement of the LOS grades across the VoT distribution is further described under the Chapter 2 - 'Proposed Approach' using Shannon's information entropy theory.

The HLOS can be defined using the method shown above as shown below in Figure 12.

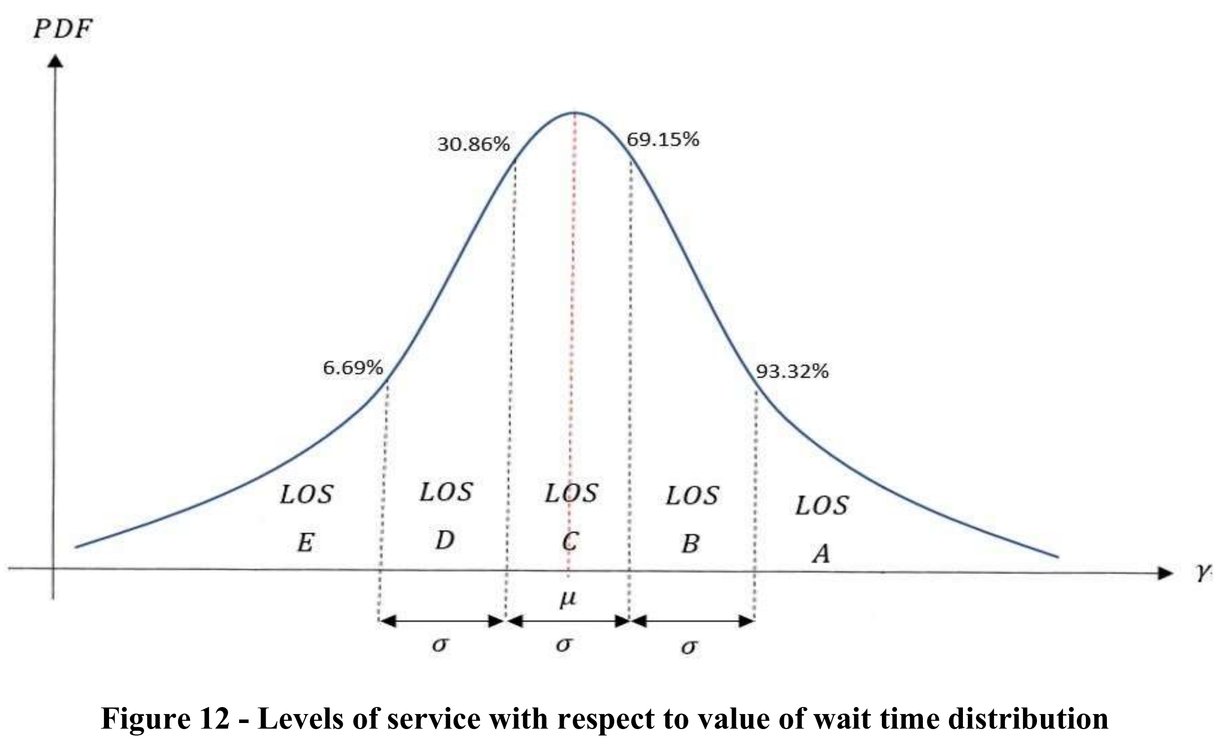


Figure 12 - Levels of service with respect to value of wait time distribution

3.4.2. A Continuous Scale (Percentage) to Indicate LOS

There are complaints (TCRP Project A-15C, 2013) from some transit agencies/practitioners that a letter grade to indicate the LOS compromises one of the needs of having an LOS measure, i.e., being able to quantify and indicate the quality improvements upon interventions (investments, operational changes, network restructuring etc.). This is important for transit agencies as LOS measures have first been used to help disbursement of limited funds to the best possible alternatives in terms of available interventions and report whatever the improvements made upon these interventions. According to a letter grade system, where a certain range of quality improvements falls into the same level of QOS (one letter grade), the minor quality improvements and even significantly large quality improvements (within $'\sigma/2'$ to $'\sigma'$ of γ_w values in the above case) might not indicate any improvements in terms of letter grades. Transit agencies claim that most of the interventions result in minor improvements which cannot be depicted in terms of a letter grade system.

In such a case, the LOS can be indicated as a percentage as follows;

For a normally distributed VoWT population, a certain value γ_{wx} in that population can be expressed as follows,

$$\gamma_{wx} = E(\gamma_w) + Z_k\sqrt{Var(\gamma_w)} \qquad 35$$

$E(\gamma_w)$ = Expected value (mean) of the VoWT distribution

$Var(\gamma_w)$ = Variance of the VoWT distribution

Z_k = Value of standard normal distribution corresponding to the percentage k (confidence level in the normal case)

Implied VoWT (γ'_w) can be expressed in this form instead of γ_{wx}. Therefore,

$$Z_k = \frac{\gamma'_w - E(\gamma_w)}{\sqrt{Var(\gamma_w)}} \qquad 36$$

There is a corresponding percentage for the Z_k value derived above. Therefore;

Headway LOS = $\lfloor k \rfloor$%

In the same way, for a log normally distributed VoWT population, γ'_w (despite γ_{wk}) can be expressed as follows;

$$\gamma'_w = \exp(\mu + Z_k.\sigma) \qquad 37$$

From these methods, it is possible to express LOS as a percentage which has a simple cognitive implication of the how good/bad the LOS is and address the concerns expressed by transit agencies in a letter grade system where it is now possible to quantify and express even the smallest improvements made.

3.4.3. Example I – Normal operation with bus and fleet size not binding

Assume, for example, an urban bus route with the following service parameters.

$$\lambda_D = 200\$\ per\ bus\ dispatch,$$

$$P = 200\frac{pass}{hr},$$

$$H = 0.33\ hrs\ (20\ min)$$

According to the currently available popular transit guidelines, the HLOS can be chosen as follows;

Table 5 - Frequency LOS for urban transit by TCQSM 1st edition

LOS	Headway (min)	Veh/h	Comments
A	<10	>6	Passengers don't need schedules
B	10-14	5-6	Frequent service, passengers consult schedules
C	15-20	3-4	Maximum desirable time to wait if bus/train missed
D	21-30	2	Service unattractive to choice riders
E	31-60	1	Service available during hour
F	>60	<1	Service unattractive to all riders

According to TCQSM (1st edition), the HLOS can be referred to as '**C**'. It is important to note that this measurement does not depend on any other parameters but only upon the headway itself. And this does not change with where and in what conditions the bus route operates.

An equivalent optimal operation, based on the available service parameters (allowing only the VoWT to vary, as it is what we are about to find out to refer to the LOS), allows us to estimate user costs based on operation costs. Hence, we can find out the implied VoWT through the existing operation.

Assuming a large fleet and bus sizes with a constant demand, from Equation (17),

$\gamma'_w = \frac{2\lambda_D}{PH'^2} = \frac{2*200}{200*0.33^2} = 18$ $/hr/pass.

Let's now assume the actual VoWT of passengers are normally distributed with a mean (**μ**) of 15 $/hr./pass. And standard deviation (σ) of 4 $/hr./pass. We also propose the 5 levels of quality be distinguished by σ as follows justified in the chapter 'Introduction'.

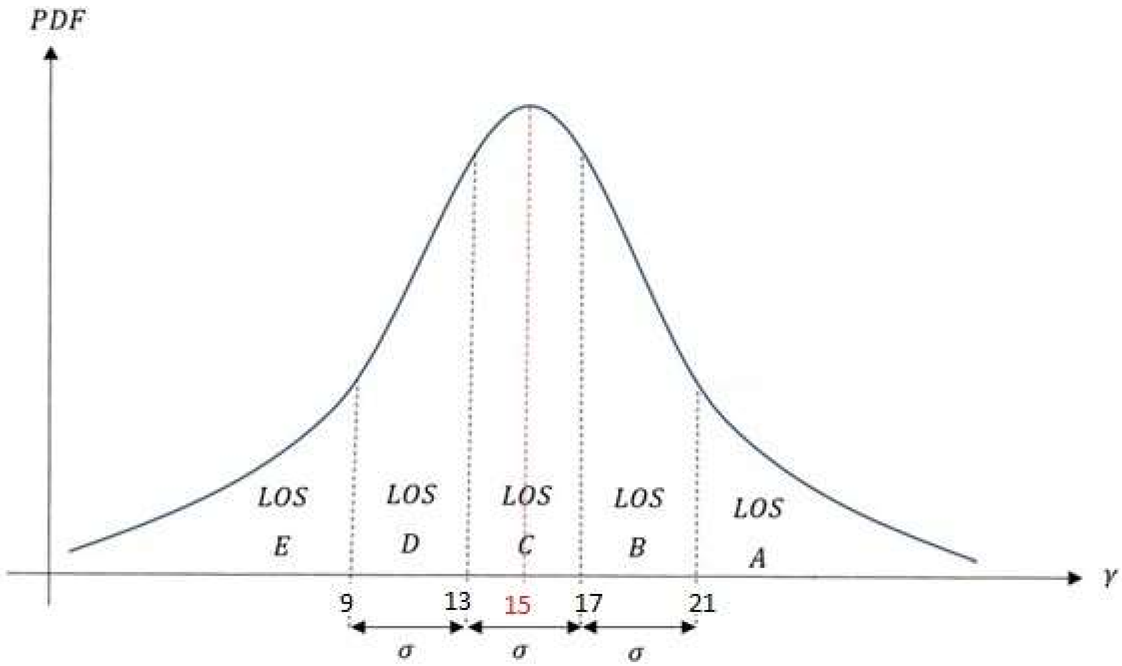


Figure 13 - VoWT distribution for the example problem with LOS letter grades

According to this proposed methodology, the previously calculated implied VoWT (γ'_w = 18 $/hr./pass) falls into the LOS interval B (between 17 and 21).

Or we can calculate the HLOS as a percentage from the Equation (35) as follows;

$$\gamma_{wx} = E(\gamma_w) + Z_k\sqrt{Var(\gamma_w)}$$

$$18 = 15 + Z_k \times 4$$

$$Z_k = \frac{3}{4} = 0.75$$

$$\text{K} = 0.77 = 77\% = LOS$$

Therefore, the LOS is 77%

3.4.4. Example II – A high demand bus route in the peak time with bus size as a constraint

Assume a proposed bus route having the following service parameters;

$\lambda_D = 100\$$

$Q = 180 \frac{pass.spaces}{hr}$

Bus capacity (size) = 50 passengers per bus

We assume that a bus cannot take any more passengers than 50 and hence there is no crowding effect.

The operator would normally dispatch buses using square root headway policy with an assumed mean VoWT of 10$/hr/pass for the passengers in the bus route.

First, we need to determine whether this bus route can operate under the optimum (square root) headway or not. We assume that the boarding demand to be 200 passes. /hr. at the time when the passenger space demand is 180 passenger spaces/hr.

Sqrt headway = $H' = \sqrt{\frac{2\lambda_D}{P\gamma'_w}} = \sqrt{\frac{2*100}{50*10}} = 0.31\ hrs$

Capacity headway H_c = C/Q = 50/180 = 0.27 hrs.

It can be noted that the capacity headway is lower. If the operator adopts the square root headway, passengers will be left behind en route creating significant negative impact on long-term choice behavior of passengers leading to a lower passenger demand in future.

Expected number of passengers per bus if the operator is to adopt the square root headway,

= $H' \times P$ =0.31*200 = 62

62 > 50 (bus size) ➔ Passengers will be left behind en-route

As the operator doesn't want this to happen, the bus route has to run under capacity headway when increasing the bus size is not an option.

Under capacity operations, we can find out the implied VoWT through an equivalent optimal operation using Equation (28) if we have the corresponding values for 'T', 'P' and 'k'. We assume T=1.1, k=100 as also given in Table 3 in this chapter. Accordingly,

$$\gamma'_w = \frac{2kQ^2T}{PC^2} = \frac{2 * 100 * 180^2 * 1.1}{200 * 50^2} = 14.25\ (\$/hr - pass.)$$

It can be noted here that the resulting VoWT is higher than the value that the operator has assumed first - 10$/hr./pass. If we assume the characteristics of the passengers in this bus route are the same as for the bus route in the Example I, we can obtain the HLOS by referring to the Figure 13.

Accordingly, the resulting LOS is '**C**'

$$(\mu - \sigma/2) < 14.25 < (\mu + \sigma/2) \ \Rightarrow LOS\,'C'$$

3.5. Discussion and Concluding Remarks

The value of a unit of wait time can differ with the purpose of the trip. A person waiting for a bus to go to work can have a higher VoWT than a person waiting for a bus to go shopping. The value of a unit of wait time for the same passenger for the same trip purpose can also differ with other factors like weather conditions (very low temperatures), security concerns at stops, facilities given at the stop. From an operator perspective, it is possible to take measures to reduce the perceived value of waiting time of passengers. For example, having heated stops, facilities like Wi-Fi at stops, and improving security at stops can lower the perceived value of the waiting time of passengers and hence shift the value of wait time distribution a little towards zero which will improve the HLOS in turn. However, research concerning this aspect is scarce. The value of a unit of wait time can be studied further accounting for the above-mentioned factors. Literature has also discussed the aspect of the VoWT being a function of the elapsed wait time at the stop (Wirasinghe, 1990) – this aspect is further investigated in the Appendix A as it relates to the HLOS.

Another important aspect that needs more attention is the amount of waiting time of a passenger. The methodology presented has assumed the average wait time of a passenger to be half the headway assuming the passengers arrive uniformly, and buses operate with a high frequency without a posted schedule. Although most of the related literature has followed a similar method, the amount of waiting time of a passenger depends on whether that passenger is planning their arrival or not – purpose of the trip makes the passenger plan or not plan their arrival at the destination. E.g., passengers traveling to work plan their arrivals and passengers traveling for shopping do not usually plan their arrivals. Ansari Esfeh et al., (2022) has conducted an in-depth exploration of the effects of these aspects on average passenger wait times. Therefore, the average amount of waiting time of the passengers in a route depends on the percentages of planning and non-planning passengers. Also, the type of service offered – whether it is frequency based (high frequency) or headway based (low frequency – affects the amount of wait time of the passengers. Effects of these aspects on the derivation of an implied VoWT is discussed in the Appendix A.

The mean value of wait time distribution can be developed through different methods such as revealed and stated preference surveys of passengers about their trip choice behaviors under different conditions. Mohring et al., (1987) and Wardman (2004) in their studies conduct an in-depth review of some of the methods that had been utilized to develop value of time distributions and effects of such methods on the results obtained.

If the VoWT implied by the operation is less than the actual mean VoWT of a given passenger, the service offered to the passenger implies a lower level of service and vice versa. In the same way, if the value of a unit wait time assigned by the operator is higher than the values of wait time of the majority of passengers, the service provided by operator is preferable for the majority of the passengers and hence can be categorized to have a better LOS.

With the increase of bus size, the weight and engine capacity of the bus will also be increased. Therefore, fuel consumption of the bus increases and the cost per bus kilometer increases resulting in an increase in operational costs. At the same time there are other costs that do not vary with the bus size like the administration cost, maintenance facilities, and driver-related costs (wage, insurance, pension). Therefore, the relationship between bus capacities and operating cost has one linearly varying component and a fixed component. The values of the components can vary depending on the policies of the operator, type of service offered and where the service is offered etc. During the application of this method, the bus operator can come up with their own model for the relationship between bus capacity and cost per bus dispatched.

4. Chapter 4 – Headway and Crowding Level of Service[1]

4.1. Introduction

As described under section 1.5.2 of the chapter 1 - 'Background', passenger crowding – also known as passenger loading – is a key attribute of the LOS framework considered in this study. It can be seen through the recent literature that crowding in transit has become a pressing issue that has drawn significant attention. This can be attributed at least partly to the notion that crowding is detrimental to the health of both the transit industry – in terms of delays and low efficiency – and passengers – in terms of safety and stress (Cox et al., 2006). Accordingly, there are many studies attempting to quantify the effects of crowding in different stages of a transit trip – in-vehicle, waiting and walking – as means to assess its role in modal usage and modal choice in terms of service quality. Most of such studies have taken an approach of estimating the value of crowding (VoC) empirically through choice experiments for passengers' willingness to pay for reduced crowding. (Li & Hensher, 2011) reviews several crowding valuation studies that present different crowding valuation estimates throughout different parts of the world where the common finding is the positive correlation of value of passengers' time and level of crowding.

Although one of the main intentions of crowding valuation is to monitor the level of service – to provide incentives for operators for maintaining a certain level of service in transit operations, the relationship between the quality of service and level of crowding is not well established. Çelebi & İmre (2020) in such an attempt related the level of crowding in transit to the perceived quality of service of passengers through a customer satisfaction survey. Another approach practiced in the industry is to define the quality of service depending on the crowding levels as measured by the percentage of passengers above the seat capacity (TCRP Project A-15C, 2013). These approaches suffer from the drawbacks discussed in section 1.6 of Chapter 1 - 'Background'.

Another particular drawback of the existing crowding valuation studies is the assumption that all passengers will value a given crowding level equally. However, passengers can exhibit observed heterogeneity in their perception toward crowding depending on the trip purpose, income and distance (Whelan & Crockett, 2009). Willingness to pay for reduced crowding can also be affected by the unobserved heterogeneity among the passengers in a transit line (Alexander, 1993).

1. [1] This chapter has been published in the journal 'Transportation Research Interdisciplinary Perspectives' under the title "Integrating COVID-19 health risks into crowding costs for transit schedule planning" in March 2022

Therefore, the passenger valuation of a given level of crowding can take a range of values which is better represented through a distribution. The approach taken in this study to represent a given performance level of an attribute that depends on the mean value of time distribution of the passengers, resolves this issue. Because the observed and unobserved heterogeneity of the passenger perception is now considered through the mean value of time distribution of the passengers in the transit line.

Our approach for developing a LOS measure, as described in detail in the chapter 'Background', is based on incorporating the concerns of both the operator and passengers by minimizing the tradeoffs between their generalized costs. Therefore, in this chapter, we first develop a function to represent the average cost of crowding for an hour in a bus route. Then we revisit the optimum headway problem in conjunction with the crowding costs. During the time of this study, the COVID-19 pandemic took the world by storm creating the worst impact ever on transit worldwide leaving transit agencies and operators with many challenges. Recent studies identifying the cost of crowding, that is a part of the riding cost in transit, have been amplified owing to significant changes in the passenger perception toward crowding affected by the COVID-19 related health risks. Therefore, the last part of this chapter investigates this disruption as it relates to crowding and headway of the bus route and makes suggestions to alleviate the negative impacts on transit.

4.2. Background

This section describes the basic principles used in the methodology to calculate crowding costs and measure crowding using the related literature.

4.2.1. Cost of crowding on an urban bus route

There are three main approaches that capture the cost due to crowding: the time multiplier, the monetary value per unit time, and the monetary value per trip (Li & Hensher, 2011). The first approach is to quantify the cost of disutility due to crowding as a markup on the average value of riding time. For instance, in (Qin, 2014) and (Whelan & Crockett, 2009), the average value of a unit of riding time spent by a passenger is multiplied by a crowding penalty factor to represent the disutility experienced by a passenger in a crowded situation inside a transit vehicle. The second approach calculates the cost of discomfort due to crowding as a portion of user costs. For example, in (Klumpenhouwer & Wirasinghe, 2016) and in (Lu et al., 2008), a value of discomfort (measured in dollars per unit time per person) is multiplied by a discomfort factor (a function of crowding level/loading) to represent different levels of crowding. The third approach calculates the cost of crowding discomfort as a "per trip" value for a given level of crowding as in (Polydoropoulou & Ben-Akiva, 2001) and (Hensher et al., 2011). This study expands the first approach because the approach has the advantage of transferring to different contexts.

4.2.2. Measuring crowding

Different individuals can view a transit vehicle that has a certain number of passengers as having different levels of crowding discomfort. In other words, one person may consider a transit vehicle crowded and another person may view the same vehicle as not crowded. Therefore, the level of crowding depends on passengers' perceptions (Whelan & Crockett, 2009). We are interested in calculating the user cost associated with crowding discomfort as a markup of the average value of riding time as in (Whelan and Crockett 2009); the value of riding time (VoRT) is multiplied by the crowding penalty factor (CPF), which is also called the VoRT multiplier. The CPF is a function of the level of crowding (LoC). The LoC is measured mainly using the loading factor (L), which is the ratio of the number of passengers on board a vehicle to the number of seats in that vehicle (Qin, 2014). For a transport mode mostly designed for seated passengers, it is suitable to use "L" as the measure of crowding, while for modes designed mostly for standees, the best measure is the number of passengers per square meter (TCRP Project A-15C, 1999).

4.2.3. Crowding Penalty Factor (VoRT multiplier) with loading factor

CPF is the factor by which the VoRT of a passenger is multiplied to account for the discomfort caused by the level of crowding. VoRT can be obtained using stated preference and revealed preference surveys (Basu & Hunt, 2012; Lam et al., 1999; Ojeda-Cabral et al., 2016; Rizzi et al., 2012). Studies have proposed methods to obtain the CPF by modelling the marginal utility of travel time as a function of the level of crowding in a transit vehicle (Batarce et al., 2016; Haywood & Koning, 2015).

4.3. Methodology

Figure 14 shows the outline of the methodology followed in this chapter.

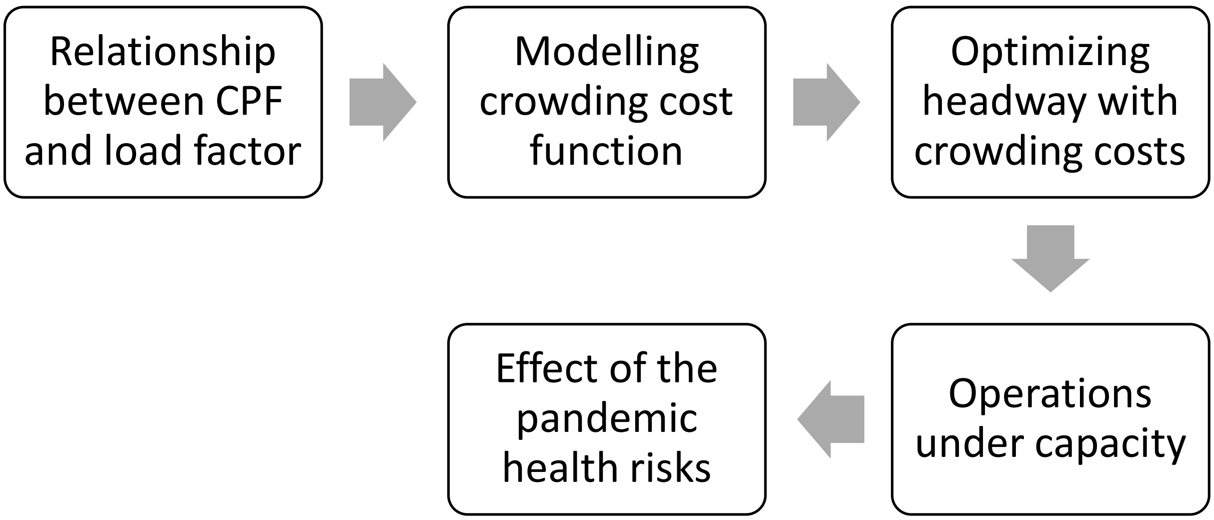


Figure 14 - Outline of the methodology

This section first presents the relationship between the VoRT multiplier (CPF) and the loading factor (L) to account for crowding discomfort. An analytical model is developed in the next step which represents the cost of crowding of a bus route in the objective function followed by a methodology to optimize the headway in the presence of crowding. The above steps are developed assuming a linear relationship between CPF and L for computational simplicity; if the relationship between CPF and L is exponential, limitations exist, and these are shown in the Appendix C.

4.3.1. Modelling the crowding cost function

The average cost due to crowding of a given bus route is first calculated. We assume an average demand rate and a uniform headway for the time period considered.

$$\text{Number of buses operating (running) on the route at any given moment} = \frac{T'}{H} \qquad 38$$

where T' is the round-trip time of a bus from start to end excluding the layover time, and H is the uniform headway of the bus route.

Average load of passengers per bus has been shown to be $PH(\bar{l}/D)$ by Tirachini, Hensher, and Jara-Díaz (2010) and Qin (2014) where $\bar{l}$ is the average trip distance of passengers in the bus route, and D is the round-trip length of the bus route. The expression for the average load of passengers per bus by Qin (2014) is extended and proved true for a many-to-many demand bus route in the Appendix B.

Therefore, the number of passenger hours spent riding in the bus route per hour

$$\frac{PH\bar{l}}{D}\frac{T'}{H} = \frac{PT'\bar{l}}{D} \qquad 39$$

The average loading factor L is defined as the average number of passengers on a bus divided by the seat capacity of the bus:

$$L = \frac{PH(\bar{l}/D)}{S} \qquad 40$$

where S is the number of available seats on the bus.

Let the average CPF (the VoRT multiplier) be

$$\beta = 1 + \emptyset L \qquad 41$$

where $\emptyset$ denotes the rate of change in average CPF with the average loading factor.

In the literature, CPF has been modelled using different functional forms, namely stepwise linear (de Palma et al., 2015), exponential (Qin, 2014), stepwise exponential (Qin, 2014), and quadratic (Tirachini et al., 2010). Depending on the situation, it can be more appropriate to use an exponential form or a quadratic form to model the variation of CPF with the loading factor. We use a linear function of CPF and explore the issue of using an exponential form in the Appendix.

The average perceived VoRT per passenger under crowding, as given by γ_r, can be obtained by multiplying the basic average VoRT per passenger (without the effect of crowding), as given by γ_r^o, by CPF:

$$\gamma_r = \gamma_r^o \beta = (1 + \emptyset L)\gamma_r^o \qquad 42$$

CPF is always larger than one. That is, the VoRT with discomfort of crowding can take a minimum value of VoRT without crowding. Therefore, the ratio of the VoRT with crowding to VoRT without crowding is always greater than or equal to one. Hence, the intercept of Equation (41) is 1.

Therefore, using Equations (40) and (42), we have

$$\gamma_r = (1 + \emptyset \frac{PH\bar{l}}{DS})\gamma_r^o \tag{43}$$

It is important to note that in Equation (43), the possible different values of seated and standing passengers' perceived ride time have been considered through the average value.

Using Equations (39) and (43), we can calculate the user costs due to crowding by multiplying the number of passenger hours spent on the bus route per hour by the average value of perceived ride time of passengers in a bus given by dollars per passenger per hour.

The cost of riding in crowded conditions per hour is as follows:

$$TC_r = \frac{PT'\bar{l}}{D}(1 + \emptyset \frac{PH\bar{l}}{DS})\gamma_r^o \tag{44}$$

Equation (44) denotes the total cost of both riding and crowding bus route passengers per hour. However, headway only affects the level of crowding but not the value of riding time of a passenger. Therefore, to account only for the additional cost incurred by passengers due to crowding discomfort, Equation (44) can be modified. Instead of the perceived average value of riding time, we use only the markup applied to the average value of riding time due to crowding, which is

$$\gamma_r - \gamma_r^o = (1 + \emptyset \frac{PH\bar{l}}{DS})\gamma_r^o - \gamma_r^o = \left(\emptyset \frac{PH\bar{l}}{DS}\right)\gamma_r^o \tag{45}$$

Therefore, TC_{cr}, the cost of only crowding discomfort, can be obtained as follows:

$$TC_{cr} = \frac{PT'\bar{l}}{D}(\gamma_r - \gamma_r^o) = \frac{PT'\bar{l}}{D}\left(\emptyset \frac{PH\bar{l}}{DS}\right)\gamma_r^o \tag{46}$$

Let the average running speed of a bus on a bus route be

$$\bar{V} = \frac{D}{T'} \tag{47}$$

Using Equations (46) and (47),

$$TC_{cr} = \frac{P\bar{l}}{\bar{V}}\emptyset \frac{PH\bar{l}}{DS}\gamma_r^o \tag{48}$$

The ratio of the average trip distance ($\bar{l}$) to the average running speed ($\bar{V}$) is the average trip time ($\bar{t}$) of passengers in the transit line and can therefore be denoted as

$$\frac{\bar{l}}{\bar{V}} = \bar{t} \tag{49}$$

Therefore, Equation (48) can be modified as follows:

$$TC_{cr} = \frac{P^2 \bar{l}\bar{t}\emptyset\gamma_r^o H}{DS} \tag{50}$$

Equation (50) represents the average cost of crowding per hour given in dollars per hour.

As it can be seen in the Equation (50), the average cost of crowding in the bus route varies linearly with the headway for a given bus route with a given average demand in a given period of time. The average cost of passenger waiting in the bus route, as given by $\left(\frac{H}{2}\right)\gamma_w P$ in the Chapter 3 'Headway Level of Service of an Urban Bus Route', is also varying linearly with headway. Therefore, the optimum headway should minimize the tradeoffs of waiting and crowding costs of passengers and operating costs.

4.3.2. Optimum headway of operating with consideration of crowding costs

Using Equation (14) in Chapter 3 and Equation (50), the cost of users, i.e., costs due to both crowding discomfort and waiting, and the cost of operators due to the headway of operating bus routes can be stated as

$$Z = \left(\frac{H}{2}\right)\gamma_w P + \frac{P^2 \bar{l}\bar{t}\emptyset\gamma_r^o H}{DS} + {\lambda_D}/{H} \tag{51}$$

The total cost function when the CPF varies exponentially with the loading factor is derived in the Appendix C.

The optimum headway (H*) that minimizes the total cost of the bus route associated with headway can be obtained by taking the derivative of Equation (51) with respect to H and setting the derivative equal to zero.

$$H^* = \left[\frac{2\lambda_D DS}{\gamma_w PDS + 2P^2 \bar{l}\bar{t}\gamma_r^o \emptyset}\right]^{1/2} \tag{52}$$

The second derivative of Equation (51) with respect to H is positive.

Hence, H^* provides the optimum headway that minimizes the trade-offs between user costs and operator costs.

Utilizing Equation (44) instead of (50) in the objective function does not affect the result as the first term within the brackets of Equation (44) disappears when taking the derivative. While the variation in headway affects the level of crowding and the cost due to crowding discomfort, it does not affect the cost of riding a bus without crowding. Therefore, the derivative that represents the rate of variation of different cost components with respect to headway variation is independent of the cost of riding even though it is included in the objective function. However, Equation (50) must be used because it conveys the concept accurately.

Equation (52) can be further rearranged as follows.

$$H^* = \left[\frac{2\lambda_D/\gamma_w P}{1 + (2P\bar{l}\bar{t}\gamma_r^o \emptyset/\gamma_w DS)}\right]^{1/2} \qquad 53$$

Since the denominator of the H^* is greater than 1 and the numerator of the Equation (53) is the optimum headway without crowding costs (H'), the optimum headway with crowding cost (H^*) is always smaller than H'. This finding is expected because user costs are higher when crowding is considered, and the operator has to balance this extra cost utilizing a higher frequency in the bus route, which increases operator costs.

The analytical model represented by Equation (52) shows the interrelation between a bus route's parameters and the optimum headway in the presence of crowding costs. As shown, H^* varies with passenger demand on a bus route, while H' only varies with the square root of passenger demand because H^* is affected by crowding cost and passenger demand has a squared effect on crowding cost. It should also be noted that as values of average trip distance, the basic VoRT, VoWT of passengers, and $\emptyset$ increase, the optimum headway with crowding becomes smaller. Also, for longer bus routes and larger bus sizes, the optimum headway with crowding is greater. However, all these parameters, except a line's passenger demand, have a square root effect on the optimum headway/frequency.

4.3.3. Operations under capacity

The minimum required headway to meet an available maximum demand of passenger-spaces Q of a bus route with buses of passenger carrying capacity C, to avoid passengers being left behind, is given by the capacity headway H_c from Equation (23) in the Chapter 3 - 'Headway LOS'.

Substituting the expression for H_c instead of H in the Equation (50) provides the average passenger cost of crowding in the bus route per hour under capacity operations as follows.

$$TC_c = \frac{P^2\bar{l}\bar{t}\emptyset\gamma_r^o C}{QDS} \tag{54}$$

Extending the Equation (25) in Chapter 3 for the total cost function under capacity operations to include the crowding cost provides.

$$Z = \frac{P\gamma_w C}{2Q} + \frac{P^2\bar{l}\bar{t}\emptyset\gamma_r^o C}{QDS} + \frac{QT(k+\alpha C)}{C} \tag{55}$$

Taking the partial derivative of the Equation (55) assumes that the other parameters are constant. That is, the bus size is also constant. Allowing the passenger carrying capacity (C) to vary while the bus size is a constant is another way of allowing the load factor to vary for a given bus size. The context simulated by this assumption is analogous to a scenario where the maximum allowable load factor, C/S, in the bus route is allowed to vary with which the cost of crowding will be variable.

Total cost would become a minimum when $\frac{\partial Z}{\partial C} = 0$ and if $\frac{\partial^2 Z}{\partial C^2} > 0$;

As $\frac{\partial^2 Z}{\partial C^2} = \frac{2kQT}{C^3}$; $\frac{\partial^2 Z}{\partial C^2} > 0$ (As k, Q, T and C are greater than zero). Therefore, for the minimum total cost,

$$\frac{\partial Z}{\partial C} = \frac{P\gamma_w}{2Q} + \frac{P^2\bar{l}\bar{t}\emptyset\gamma_r^o}{QDS} - \frac{QTk}{C^2} = 0 \tag{56}$$

Provides the optimum design capacity (C^*);

$$C^* = \sqrt{\frac{2kQ^2T}{P\gamma_w + (2P^2\bar{l}\bar{t}\emptyset\gamma_r^o/DS)}} \tag{57}$$

Therefore, the design capacity of a bus operating under optimal capacity headway conditions can be presented as a function of the mean VoWT and mean VoRT of passengers.

$$C^* = f(\gamma_w, \gamma_r^o) \tag{58}$$

In the context of planning a bus route, operators might need to find a suitable bus size that can provide a specified/desirable LOS during capacity operations. The best way to approach such a situation is to assume a maximum allowable load factor which then becomes a constant in the Equation (55). Transit agency can decide their maximum allowable load factor constrained by a maximum of crush capacity. If the maximum allowable load factor is L_m,

$$C = L_m S \tag{59}$$

Then the problem becomes about finding the bus size that can deliver a predetermined LOS. The approach to this is by modeling the total cost function in terms of the bus size and finding the optimum bus size as follows.

Substituting for C in the Equation (55) from the Equation (59),

$$Z = \frac{P\gamma_w L_m S}{2\text{Q}} + \frac{P^2 \bar{l}\bar{t}\emptyset\gamma_r^o L_m}{QD} + \frac{QT(k + \alpha L_m S)}{L_m S} \tag{60}$$

Total cost would become a minimum when $\frac{\partial Z}{\partial S} = 0$ and if $\frac{\partial^2 Z}{\partial S^2} > 0$;

As $\frac{\partial^2 Z}{\partial S^2} = \frac{2kQT}{(L_m S)^3}$; $\frac{\partial^2 Z}{\partial S^2} > 0$. Therefore, for the minimum total cost,

$$\frac{\partial Z}{\partial C} = \frac{P\gamma_w L_m}{2\text{Q}} - \frac{QTk}{(L_m S)^2} = 0 \tag{61}$$

Provides the optimum bus size (S^*);

$$S^* = \sqrt{\frac{2kQ^2T}{P\gamma_w L_m^3}} \tag{62}$$

S^* is a function of the VoWT - γ_w.

It can be seen from Equation (61) that the crowding cost component disappears while taking the partial derivative with respect to the bus size. This is because now the load factor is fixed, the cost of crowding is independent of the bus size. Therefore, the optimum bus size no longer depends on the cost of crowding – as the bus size, given in number of seats per bus, does not concern the level of crowding in the bus (different bus sizes can have the same level of crowding/loading).

For a given bus size S and maximum allowable load factor L_m, the HLOS can be estimated from finding the implied VoWT γ_w'.

$$\gamma_w{}' = \frac{2kQ^2T}{PL_m^3S^2} \tag{63}$$

The HLOS can be obtained following the general procedure outlined in chapter 2 - 'Proposed Approach'. According to Equation (63), the HLOS depends on the bus size and the maximum allowable load factor – the load factor at the design capacity - L_m. This is because, utilizing a larger design capacity - L_mS – increases the wait time of the passengers. Also, L_m has a more significant influence on the headway LOS than the bus size.

4.3.4. Headway and Crowding Level of Service

4.3.4.1. Normal operations

According to Equation (50), the cost of crowding is a function of the headway of the bus route. As headway increases, the operator cost decreases, and the cost of crowding and waiting for the passengers increases. Therefore, the tradeoff is in between the sum of crowding and waiting costs of the users and the operator cost. As described in Chapter 1 – 'Background', the transit trips happen when users contribute their time – generalized cost of time – and operators contribute their resources – cost of operations. The concept behind deriving the LOS measures is to compare the existing operation with an optimum operation where the tradeoff between user and operator costs is at a minimum. Accordingly, in a situation where the summation of both the crowding and waiting cost of passengers impacts the user cost that is in tradeoff with the operator costs, it is more appropriate to use a single LOS measure for the attributes headway and crowding combined.

Comparing Equation (53) with Equation (15) in Chapter 3, 'Headway LOS', where the optimum headway is a function of the mean VoWT of the passengers, the optimum headway in the presence of the crowding costs is also a function of both the mean VoWT and the mean value of ride time (VoRT) of passengers. Time valuation studies have mainly explored and reported the VoRT of the passengers – also known as in-vehicle-time (IVT) – as opposed to other time valuations like VoWT, value of access (walk) time (VoAT), etc., (Wardman, 2004; Wardman et al., 2016). Where the VoWT and VoAT are reported, they were typically secondary to VoRT (Wardman, 2004). Wardman (2004) in his study – one of the most comprehensive review studies to date on public transport time valuations – provides emphasis on this and presents formulae for deriving the public transport VoRT, VoWT and VoAT depending on the VoRT for car users which, according to Wardman, is the most available and reliable source for time valuations. (Wardman et al., 2016) in their study take a similar approach by developing a meta-model for deriving VoRT, and the VoWT and VoAT in terms of VoRT, depending on the gross domestic product of a country for the European context. This evidence bears witness to the versatility of the VoRT in representing the other time related attributes in a trip made using public transportation. Accordingly, Wardman (2004) suggests that it is more appropriate to value the VoWT and VoAT at 2.5 and 2 times the VoRT as opposed to the conventional wisdom at that time in transportation planning to value both the VoWT and VoAT at twice the VoRT. Taking a similar approach, the mean VoWT in Equation (52) can be expressed by multiplying the mean VoRT by a factor (ρ).

$$\gamma_w = \rho \gamma_r^o \qquad 64$$

Chapter 4 - Quality of Service Concepts - of the TCQSM third edition reports the nominal values and ranges for ρ in the context of United States (US) - Exhibit 4-5 (Quality of Service Concepts, 2013). . Balcombe et al., (2004) presents the theoretical framework to derive the factor ρ using the utility specification used in time valuation studies.

Accordingly, the Equation (52) can be modified using Equation (64) as a function of only the VoRT.

$$H^* = \left[\frac{2\lambda_D DS}{(\rho PDS + 2P^2 \bar{l}\bar{t}\emptyset)\gamma_r^o}\right]^{1/2} \qquad 65$$

For an existing bus operation, using the practiced headway, the implied base VoRT ($\gamma_r^{o\prime}$) can be derived as follows using the Equation (65).

$$\gamma_r^{o\prime} = \frac{2\lambda_D}{(\rho P + 2P^2 \bar{l}\bar{t}\emptyset / DS)H^2} \qquad 66$$

The implied VoRT derived from Equation (66) can be compared against the mean VoRT distribution of the passengers to derive the combined headway and crowding LOS (H&C LOS) of the bus route as described in the Chapter 2 - ‘Proposed Approach’.

In a typical situation of a city having origin destination survey done for planning purposes, the values of $\bar{l}$, $\bar{t}$ and P can be easily derived. Section 4.5.1 of this chapter describes an approach to obtain generic values for $\emptyset$ with loading and section 4.4.3 of this chapter presents an approach for developing values for $\emptyset$ for a specific context using a stated preference study if more accurate estimations are required. Similarly, values reported in the TCQSM – chapter 4 or in the literature can be used as generic values for ρ or a methodology presented in the cited studies in this section can be utilized for deriving more accurate estimations if necessary.

It is important to note here that, as compared to the implied VoWT to derive headway LOS in Equation (17) in the Chapter 3, the implied VoRT in the case of combined headway and crowding LOS depends on several other parameters specific to the bus route and operation such as D, S, $\bar{l}$ and $\bar{t}$. It also depends on population specific values such as $\emptyset$ and ρ. The value of $\emptyset$ relates the implied VoRT to the crowding cost while the value of ρ relates it to the waiting cost through the VoWT. The higher the sensitivity of CPF for the loading factor ($\emptyset$), the lower the $\gamma_r^{o\prime}$ representing a lower H&C LOS as perceived by the passengers. A similar relationship exists between the value of ρ and H&C LOS. The value of ρ that depends on how passengers value their wait time versus ride time normally remains unchanged over time unless there are significant behavioral changes such as technological disruptions for passenger travel e.g., having the ability to engage in other relevant work en-route through internet. This is not the case for different geographical locations as US and United Kingdom (UK) report values with slight variations ("Quality of Service Concepts," 2013). Increasing demand and headway can be seen to decrease $\gamma_r^{o\prime}$ in a similar scale implying that decreasing H&C LOS from increasing demand can be compensated by increasing the frequency of the bus route. D on the other hand can be seen to have confounding effects on both the denominator and enumerator of the Equation (25). In a typical situation of a bus route, there can be a relationship between the average trip distance ($\bar{l}$) and the length of the bus route (D) where the $\bar{l}$ can be represented as a fraction of D, then the value of $\gamma_r^{o\prime}$ becomes independent of the value of D. If this happens, the H&C LOS tends to increase with the ratio $\bar{l}/D$, i.e., the higher the average trip distance of the passengers in the route compared to the route length, the lower the H&C LOS. This can be attributed to the increased average crowding cost of the passengers for traveling a comparatively longer time in a given loading condition.

Substituting for $\bar{l}, D$, and S in the Equation (66) from Equation (40) provides,

$$\gamma_r^{o\prime} = \frac{2\lambda_D}{(\rho H + 2L\bar{t}\emptyset)PH} \qquad 67$$

that shows increasing the average load factor (L) of the bus route decreases the H&C LOS.

4.3.4.2. Capacity Operations

Under capacity operations when a bus route operates with a given bus size, the optimum design becomes a function of both the mean VoRT and mean VoWT.

Using Equation (64), Equation (57) can be modified as follows to make the optimum design capacity a function of the mean VoRT.

$$C^{*} = \sqrt{\frac{2kQ^{2}T}{\rho P\gamma_{r}^{o} + (2P^{2}\bar{l}\bar{t}\emptyset\gamma_{r}^{o}/DS)}} \qquad 68$$

For an existing capacity operation, using the practiced design capacity C, the implied base VoRT ($\gamma_{r}^{o\prime}$) can be derived as follows using the Equation (68).

$$\gamma_{r}^{o\prime} = \frac{2kQ^{2}T}{C^{2}\rho P + (2C^{2}P^{2}\bar{l}\bar{t}\emptyset/DS)} \qquad 69$$

Equation (69) can be further simplified using the Equation 20 in the Chapter 3 for capacity headway H_c as follows.

$$\gamma_{r}^{o\prime} = \frac{2kT}{\left[\rho P + (2P^{2}\bar{l}\bar{t}\emptyset/DS)\right]{H_{c}}^{2}} \qquad 70$$

Comparing the implied VoRT for capacity operation with the implied VoRT for normal operations – Equations (70) and (66) – shows that the term in the denominator is the same except for the corresponding headway used depending on the type of operation. The term in the numerator of Equation (66) can be expanded using the Equation (11) from Chapter 3 to show that the numerator of the Equation (70) is a part of it as follows.

$$2\lambda_{D} = 2\mathrm{T}(k + \alpha C) = 2kT + 2T\alpha C \qquad 71$$

Therefore, while $\gamma_{r}^{o\prime}$ for normal operations increases with the increasing design capacity, $\gamma_{r}^{o\prime}$ for capacity operations decreases with the increasing design capacity and has a squared effect as ${H_c}^{2} = C^{2}/Q^{2}$. This is because at normal operations, increasing the design capacity – bus size – will create more space inside the bus for the passengers and hence reduce the cost component of crowding leading to an increased implied VoRT. This will increase the H&C LOS. At capacity operations, increasing the design capacity will not change the level of crowding as buses are dispatched fully through the maximum load section. Although this increases the number of people experiencing a given level of crowding which in turn increases the cost of crowding leading to a reduced implied VoRT. This will decrease the H&C LOS.

4.4. Integrating COVID-19 health risks into crowding costs

4.4.1. Introduction

This study was carried out in a time when the COVID-19 pandemic had taken the world by storm, impacting public transportation systems worldwide severely. The pandemic has created unprecedented challenges needing transit authorities to reevaluate the existing practices of transit operations. One of the areas affected is transit quality of service. Existing approaches assessing transit quality had to be changed as now passenger perception of the transit service had significantly changed. As a result, this study explores the issue of transit passenger perception being affected by the pandemic health risks and suggests an approach that can address this issue.

The rapid spread of COVID-19 had profound effects on the lives of people worldwide. The World Health Organization (WHO) declared COVID-19 a pandemic on 11th March 2020 (WHO, 2020). Among the steps taken to curb the spread of COVID-19, social distancing measures and travel restrictions had a detrimental effect on the public transportation (PT) sector. Fear of contamination from exposure to other people, along with government restrictions, resulted in a plunge in PT demand of 70% - 90% (UITP, 2020). The loss of PT demand can mostly be attributed to reduced travel demand resulting from a combination of public health restrictions, working from home, online schooling, shift to online shopping, transit service cuts, and loss of employment. PT as a mode absorbed some of the most adverse effects (Tirachini & Cats, 2020). The coronavirus pandemic may prove to be a catalyst for long-term changes in trip making behavior. Many people might continue to work remotely, at least part time, which means travelling to work fewer days per week. In addition, because of passengers' concerns about PT hygiene (Beck & Hensher, 2020) and their reluctance to use transit (Labonté-Lemoyne et al., 2020), there is a risk of significantly lower PT ridership levels even in the aftermath of the pandemic.

The negative impacts of reduced PT demand are widespread. While it directly impacts transit agencies, it will have a ripple effect on several layers of society. Revenue loss due to lower ridership reduces a transit agency's capability to provide the same service levels. This revenue loss therefore results in reduced frequency and time span of service, which are evident during and after lockdowns due to COVID-19 (Gkiotsalitis & Cats, 2020). Cuts to service adversely affect the quality of service offered, leading to an even lower demand. This downward spiral trend will crucially affect social equity and deprive parts of some communities of their accessibility and mobility needs that could not be otherwise met for seniors, school children, essential workers, and other segments of the population. The resulting shift away from transit also leads to increased passenger car use, which creates other problems, such as increased greenhouse gas (GHG) emissions, traffic accidents, fuel costs, and traffic congestion (Eboli & Mazzulla, 2007). The importance of preventing transit demand and mode share degradation during the phased re-opening of the economy and the aftermath of the pandemic is evident.

Transit agencies face many challenges in adapting their services while considering COVID-19 health risks and the changes in transit demand levels. Although the economy will eventually reopen from lockdown, it will go through many ups and downs because reopening will be affected by the penetration rate of COVID-19 vaccines and the variants of the virus that might cause further waves of spread. Therefore, for a given community, the state of the pandemic has been dynamic and uncertain. Because of the lack of useful decision support tools to help redesign services in times of dynamic uncertainty, transit agencies have resorted to ad-hoc measures (Gkiotsalitis & Cats, 2020). To accommodate reduced demand due to COVID-19, most transit agencies in North America have responded by cutting services and operating on reduced schedules and capacities. Such cuts in service cause a further downward spiral in transit demand (Tirachini & Cats, 2020). While fare revenue is reduced, unit operating cost has increased because of the need to address varying health safety measures and social distancing guidelines.

This chapter revisits the optimum headway problem and considers crowding in conjunction with the issues of public health, quality of service, capital, and operating costs of the system. In this approach, COVID-19 related issues are considered as part of passengers' perceived health risks associated with crowding using a crowding penalty factor, i.e., the value of ride time multiplier under crowding. At a particular crowding level, the perceived risk of crowding is incorporated differently for various pandemic levels. The developed approach is numerically tested using a simple case of a single bus route with and without the constraint of bus size.

4.4.2. Background

This section is divided into five parts. First, it reviews the recent literature on the impact of the COVID-19 pandemic on transit operations. It then discusses the recent attempts to revise transit operations to cope with the challenges posed by the pandemic. The following section briefly explores the new generalized user and operator costs due to the pandemic, which is integral to minimizing the tradeoffs between user and operator costs. The last two parts of this section describe the basic principles used in the methodology to calculate crowding costs and measure crowding using the related literature.

4.4.2.1. COVID-19 impacts on transit

During different pandemic stages, the PT demand was significantly lower due to travel and activity restrictions and fear of infection (Honey-Rosés et al., 2020). Even in the aftermath of the pandemic, researchers anticipate a potential increase in teleworking and teleshopping (Orro et al., 2020; Shamshiripour et al., 2020). While these new trends in trip making behavior point towards a sustained reduction in PT demand, they do not necessarily mean a reduction in transit modal share. In contrast, a reduction in transit modal share during and post-COVID-19 expresses a preference for other modes. Studies show that there is a significant reduction in the modal share of transit in response to COVID-19 (Apple, 2021; Orro et al., 2020). This shift can be attributed, at least in part, to negative passenger perceptions of the potential health risks of riding transit (Przybylowski et al., 2021) t. If these negative perceptions are not properly addressed during the transition stage, the reduction in the modal share of transit may be sustained for an extended period in the aftermath of the pandemic (Przybylowski et al., 2021; Tirachini & Cats, 2020).

4.4.2.2. Attempts to alleviate pandemic induced issues in transit operations

Studies on addressing pandemic related issues in transit scheduling attempt to allocate transit agency resources optimally while introducing capacity reductions brought on by social distancing measures as constraints (Gkiotsalitis & Cats, 2020; Tirachini & Cats, 2020). This approach can be further explored through models that set frequencies for a single line (Furth and Wilson 1981) or for a network with limited resources as constraints (Yu et al., 2011); models that maximize resource allocation can also be used (Verbas & Mahmassani, 2013). Gkiotsalitis and Cats (2021) extend the approach of Furth & Wilson (1981) to a network-wide problem of optimal frequency setting. Their approach considers the lost revenue from passengers left behind because of transit vehicles adhering to social distancing measures (restricting adjacent seats, etc.) with the objective of minimizing the tradeoffs between user and operator costs. These approaches, however, do not consider the new perceived discomfort cost, induced by COVID-19, in riding in a crowded vehicle.

Perceived risk is the primary factor that influences human decision-making behavior (Bavel et al., 2020). The perceived health risk of being in an enclosed space with a crowd has reduced the use of public transport during the pandemic (Dandapat et al., 2020) and will continue to do so post-pandemic. In fact, crowding was found to be the most important factor in mode choice during the pandemic (Shin et al., 2021). Research has revealed that passengers' willingness to return to transit significantly depends on perceived safety and comfort of using transit during the pandemic (Kopsidas et al., 2021; Przybylowski et al., 2021). Tan & Ma (2020) found that people who perceived a higher risk of contracting the virus by taking transit had a lower probability of taking transit. Further, passenger perception of public transit being safe was found to have increased overall satisfaction with transit (Dong et al., 2021). Supporting these claims, studies report that negative perception of crowding was magnified by an amount of 1.04 – 1.23 due to the pandemic (Cho & Park, 2021). Shin et al., (2021) reported that metro passengers in the city of New York showed ride time crowding multiplier values of 2.13 and 2.65 for sitting and standing passengers, respectively. Pollock et al., (2021) reported an implied cost increment of about $37 CAD due to the disutility of high pandemic severity compared to low pandemic severity as measured by the number of daily infections. Therefore, it is evident that transit mode choice during the re-opening stages and the aftermath of COVID-19 will continue to be significantly influenced by the perceived health risk/safety.

4.4.2.3. Pandemic induced generalized costs of transit

Although the evidence that reveals the significance of crowding induced health risks for transit riders in determining transit attractiveness post-COVID-19, there are no studies that investigate the effects of crowding induced health risks in transit scheduling. Consequently, some transit agencies have adopted the physical distancing measures recommended by health professionals to curb the spread; these measures mean that there is a specific maximum level of crowding allowed in transit vehicles (Kamga & Eickemeyer, 2021). Nevertheless, different people can view this same level of crowding as having different levels of risk, which will affect transit demand. This situation can even propagate to the aftermath of a pandemic when transit systems may not require social distancing measures, but passengers may still have residual anxiety from COVID-19. To develop an analytical expression for the optimal headway of a transit line, we first briefly explore the generalized costs of users and operators.

The pandemic has had a multifaceted impact on generalized operator and user costs. Table 6 highlights some cost increments due to the pandemic. Of the operator and user-related cost increments, all except crowding discomfort due to COVID-19 health risk can be directly quantified. For example, the increased operator costs are reflected in the cost of dispatching a transit vehicle on a particular route that is discussed later in the paper. Increased passenger costs are directly reflected in the increased values of ride and wait time of the passengers. The discomfort due to crowding depends on passenger's perceptions of COVID-19 health risks while riding transit and therefore becomes a challenging issue that requires considerable theoretical support.

Table 6 - Sudden effects of the covid-19 pandemic on operator and user costs

Increases in Operator Costs	Increases in User Costs
Cost of mandated safety measures (cleaning vehicles and infrastructure)	Increased ride time cost due to adopted health precautions (passenger metering at stations and using specific doors to board and alight)
Cost due to increased operating time (higher round trip time due to cleaning activities and safety measures for passengers boarding alighting and in-vehicle circulation)	Increased wait time costs due to cuts in service (longer headways cause longer average wait times)
Cost due to reduced capacity (more vehicles need to be dispatched per unit time)	Increased cost of crowding discomfort due to COVID-19 health risks

This research is mainly to find a way to quantify the cost of crowding discomfort due to COVID-19 and to incorporate the cost into an analytical model that can support scheduling transit services. We will first investigate crowding discomfort without the effects of pandemic health risks on an urban bus route. This provides a reference to then explain the cost of crowding discomfort associated with COVID-19 health risks.

4.4.2.4. Effect of COVID-19 on crowding discomfort

It is evident that the extent of discomfort corresponding to a given level of crowding is partly influenced by COVID-19 related health risks (Cho & Park, 2021; Shin et al., 2021). Before the pandemic, discomfort due to crowding was influenced by the sense of a lack of privacy (Fried & Defazio, 2016) because of the standing passengers and the difficulty of boarding and alighting. Discomfort is further amplified by COVID-19 related health risks. Due to its integral role in terms of crowding, we consider the pandemic health risks as one of the two components that influence discomfort due to crowding. In other words, discomfort and the associated costs corresponding to a given level of crowding have increased because of the pandemic. If we assume that the resulting crowding discomfort varies linearly with the level of crowding (Jara-Diaz & Gschwender, 2003), the effects of the pandemic health risks can be expressed along with crowding discomfort in a form similar to Figure 15.

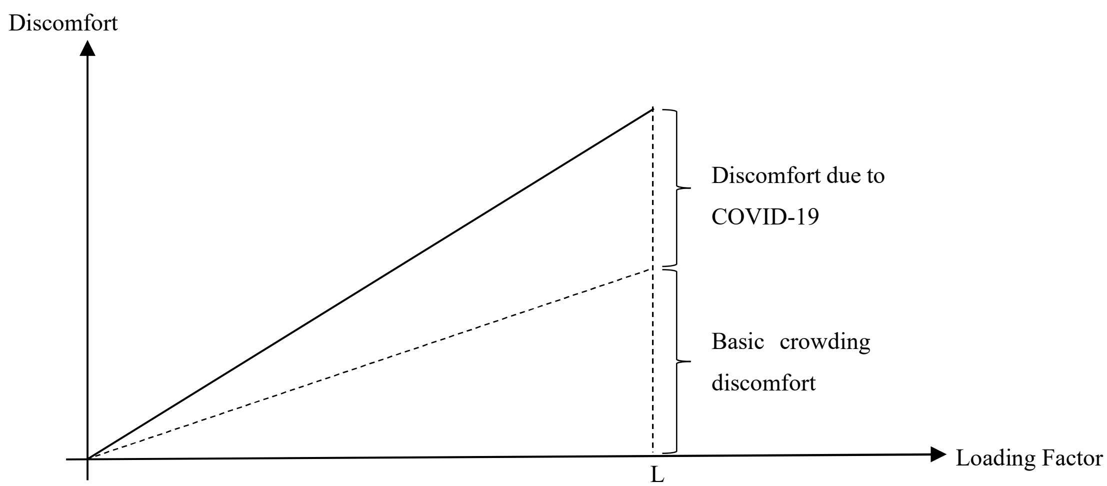


Figure 15 - Basic and COVID-19 health risks related crowding discomfort with the loading factor

The crowding discomfort of a passenger due to COVID-19 depends on the perception of the prevailing health risks associated with riding transit (Cho and Park, 2021). Perception of the overall health risks depend on cleaning strategies undertaken by the operator; safety measures onboard and at stations, such as mask policy and social distancing; rate of compliance to the safety measures (Beck & Hensher, 2020; Elias & Zatmeh-Kanj, 2021); and the stage of the pandemic expressed in terms of measures such as daily infection rates and number of infected persons in the city/province per 10,000 persons (Shelat et al., 2020). The collective effect of the above factors can be represented as the perceived chance of getting infected while taking transit. Most of these factors can be controlled or remain unchanged over long periods, but the state of the pandemic cannot be easily predicted. Therefore, the factor “state of the pandemic” is selected to represent the perceived health risk of taking transit. Selecting the factor “state of the pandemic” helps the study achieve one of its primary intentions: to help estimate passenger perception towards crowding during different pandemics.

4.4.3. Incorporating pandemic health risks into the calculation of crowding costs

As the level of crowding (or loading factor) increases, passengers are willing to pay more to travel using a mode that is less crowded (Yap et al., 2020). Therefore, a higher load factor indicates a higher disutility per unit time. At the onset of the COVID-19 pandemic, a particular crowding level posed different levels of disutility depending on the severity of the pandemic (Pollock et al., 2021; Shelat et al., 2020). For example, a person may be comfortable travelling on a given transit line at a given loading factor when a city's daily infection rate is around a typical intermediate level of 300 (NAIT, 2021). The same person might not be comfortable travelling on the same transit line at the same level of crowding when the daily infection rate in the city is around 800, which is near a lockdown state (Przybylowski et al., 2021). Therefore, the disutility of travelling time (utils per unit time), and hence, the VoRT, depends on both the loading factor and the health risks posed by the pandemic at that time. Accordingly, the resulting change to CPF can be expressed by adding another term to Equation (41) as follows:

$$\beta = 1 + \emptyset L + \sigma RL = 1 + (\emptyset + \sigma R)L \qquad 72$$

where R represents the level of pandemic severity (e.g., daily infection rate), and σ is a constant that indicates the rate of change in CPF for a unit change in the RL. The term “σRL” is the average amount by which passengers might increase their CPF to trade off the health risks caused by a given level of pandemic severity and a given level of loading. Here, σL can be interpreted as the rate of change in β (CPF) with the pandemic severity (R), and ‘$\emptyset + \sigma R$’can be interpreted as the rate of change in β with loading factor (L).

Studies have reported that the COVID-19 virus can live on different types of surfaces for different durations of time up to a few days (Moriyama et al., 2020; Nishiura et al., 2020). Thus, despite regular cleaning and safety measures, passengers could still perceive a health risk in travelling on a transit vehicle even when the bus has no other passengers, i.e., zero loading. This health risk can increase as the severity of the pandemic increases; more active cases in a community increase the chance of getting infected. This component of the health risk also causes the value of β to increase in a fashion that only depends on pandemic severity and not on the level of crowding. Accordingly, Equation (72) can be modified by adding the term ηR, where η is a constant that represents the rate of change in the portion of CPF that depends only on R, the pandemic severity.

Tirachini & Cats (2020) emphasize the possibility that there might be a residual health risk even in the aftermath of the pandemic. In other words, there is a portion of the health risk, independent of both the level of crowding and the severity of the pandemic, that will increase CPF and remain even when the pandemic is over. This portion of the health risk can be expressed by adding the term ω to Equation (72), where ω is a constant that represents the residual health risk unique to the population considered.

Equation (72) can be modified using the expression for L from Equation (40):

$$\beta = 1 + (\emptyset + \sigma R)\frac{PH\bar{l}}{DS} \qquad 73$$

Figure 16 illustrates Equation (73) and the terms ηR and ω

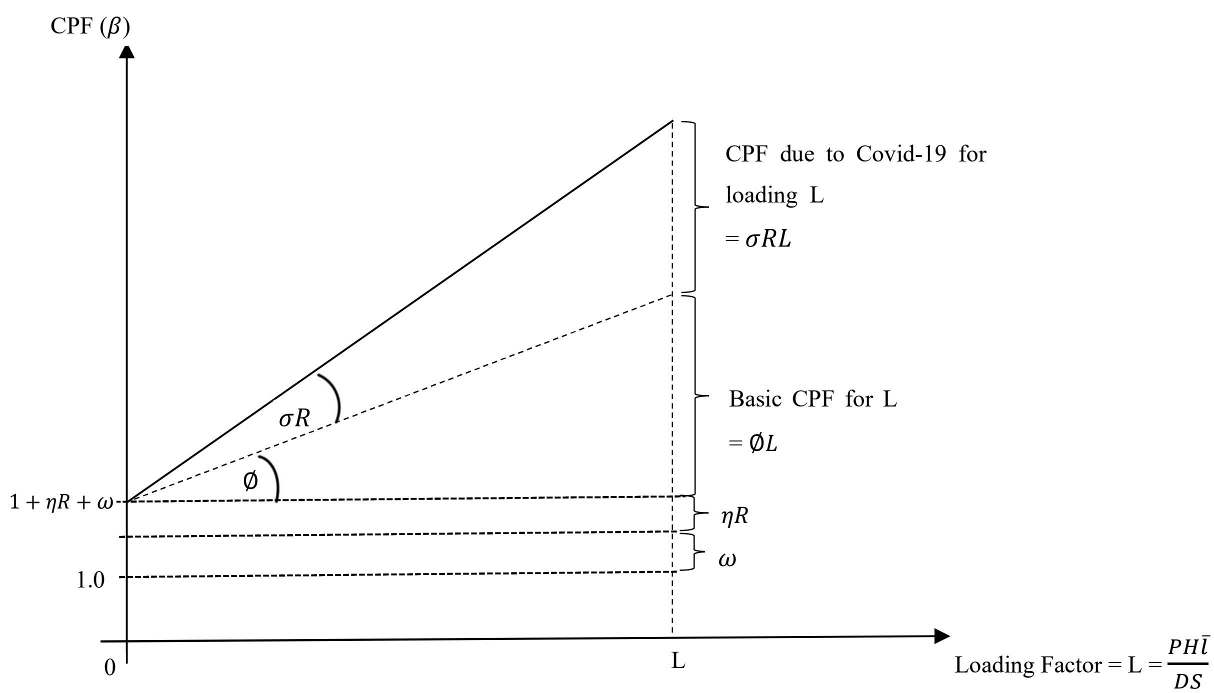


Figure 16 - Variation of CPF with loading considering COVID-19 health risks

Figure 16 shows the ways in which pandemic health risks can affect a passenger's perceptions of riding transit. However, ηR and ω are eliminated in the process of obtaining the optimum headway, i.e., taking the derivative in terms of H to obtain the optimum headway eliminates these terms as they are not dependent on the headway. On the other hand, we expect ηR and ω to be negligible because having no other users inside a transit vehicle does not pose a significant health risk to most users most of the time. The residual perceived health risk (ω) also does not affect most users and dissipates over time (and into the aftermath).

Substituting from Equation (72) in Equation (43) to develop the objective function as in Equation (51) and taking the derivative in terms of H, the optimum headway, considering both crowding and pandemic health risk related costs (H**), can be obtained as follows:

$$H^{**} = \left[\frac{2\lambda_D/\gamma_w P}{1 + 2P\bar{l}\bar{t}\gamma_r^o(\emptyset + \sigma R)/\gamma_w DS}\right]^{1/2} \qquad 74$$

As shown in Equation (74), in the presence of pandemic health risks, the optimum headway is further reduced as expected by the addition of σR to the value of $\emptyset$. As the severity of the pandemic increases, the cost of perceived health risks increases, which lowers the optimum headway.

The estimations for $\emptyset$ and σ can be obtained using a stated preference survey of transit passengers. For example, Batarce et al., (2016) has shown that the marginal utility of transit travel time can be modelled as a linear function of the loading:

$$\delta_t(L) = \delta_0 + \delta_1 L \qquad 75$$

where δ_0 (utils per hour) can be expressed as the basic marginal utility of travel time – without the effect of crowding and pandemic health risks, and δ_1 (utils per hour per unit loading) can be expressed as the marginal utility of travel time with crowding. During a pandemic, the value of δ_1 is further affected by the pandemic severity level (R). Therefore, taking a similar approach to Batarce et al., (2016) in Equation (75), we propose that the value of δ_1 can be modelled as a linear function of pandemic severity:

$$\delta_1 = \theta_0 + \theta_1 R \qquad 76$$

where θ_0 (utils per hour per unit of loading) can be expressed as the basic marginal utility of travel time with crowding – without the effect of pandemic health risks, and θ_1 (utils per hour per unit of loading and pandemic severity) can be expressed as the marginal utility of travel time with crowding under pandemic health risks.

Using this modelling approach, CPF can be obtained for different levels of loading and pandemic severity through stated preference and revealed preference surveys. The obtained values reflect the average passenger perception of pandemic health risks under different crowding and pandemic severity conditions. Using these values, the function for CPF in the presence of crowding and pandemic severity, Equation (73), can be developed, which provides the values for $\emptyset$ and σ. A practical example of deriving a function for CPF is presented in section 4.5.1 of this chapter using available CPF values from the literature for different loading conditions. Similarly, the method described herein supports the derivation of the CPF as a function of loading factor and health risk measure.

Substituting for γ_w in Equation (74) from Equation (64), the implied base value of riding time can be obtained as follows for normal bus operation under pandemic conditions.

$$\gamma_r^{o\prime} = \frac{2\lambda_D}{\left(\rho P + 2P^2 \bar{l}\bar{t}(\emptyset + \sigma R)/DS\right)H^2} \tag{77}$$

The implied VoRT derived from Equation (77) can be compared against the mean VoRT distribution of the passengers to derive the H&C LOS of the bus route as described in the chapter 2 - ‘Proposed Approach’. Comparing Equation (77) with Equation (66), it can be seen that the addition of the term $\emptyset + \sigma R$ instead of $\emptyset$ in the denominator has reduced the corresponding value of the implied base value of ride time to account for the increased cost of crowding due to pandemic health risks. This reduction in $\gamma_r^{o\prime}$ represents a reduction in the H&C LOS in the bus route.

4.4.4. Operations under capacity during the pandemic

The minimum required headway to meet an available maximum demand of passenger-spaces Q on buses of passenger carrying capacity C, is given by the capacity headway $H_c = C/Q$. The regular policy is to operate a given bus route with the minimum out of the optimum headway and the capacity headway to avoid passengers being left behind.

The norm in capacity operations is to dispatch buses when full. In a many-to-many demand bus route, the concept of “dispatching buses when full” entails meeting the demand of the rate of passengers passing through the maximum load point using the bus capacity (a bus full of passengers). Yet, the term "full" can be subjective, i.e., the person operating the transit vehicle and the passengers inside the transit vehicle may have different opinions on what "full" means; it can be even more subjective under the effect of COVID-19 health risks (T. Dai & Taylor, 2020). There is a maximum number of passengers that can physically fit in a transit vehicle of a given size, commonly known as the crush capacity, which is sometimes referred to as 1.5 times the number of seats available (TCRP Project A-15C, 2003a). The regular practice is to use a number that represents the number of passengers allowed in a bus that is lower than the crush capacity C'. Therefore, C is normally in the range of 1 to 1.5 times the seat capacity.

The load factor resulting from operation under optimum headway can rise above the crush capacity under extreme values of some parameters (e.g., λ_D), which significantly increases operator costs. Therefore, it is necessary to have an upper limit to the loading factor. This upper limit is normally the crush capacity (1.5 times the seat capacity). Under pandemic conditions, such a value indicates a significant health risk. Therefore, during a pandemic, transit agencies need to consult the health authorities to agree on an upper limit to the level of loading/crowding allowed to manage the health risks of the pandemic. If the health authorities request a minimum distance between passengers, this required distance can dictate a particular design capacity depending on the available space inside the bus. Likewise, if the upper limit for the design capacity is C^u, the capacity headway H_c becomes the ratio of C^u to Q:

$$H_c = \frac{C^u}{Q} \qquad 78$$

Accordingly, the resulting dispatching policy is to use the minimum H_c and H^{**}.

As discussed in Section 3.3.4 of Chapter 3, the policy headway is related to the quality of service and trust. Significant increases in operator costs and reduced demand due to a pandemic can make adhering to the same policy headway challenging for transit agencies. In a pandemic, essential workers, health care and front-line workers, depend on transit services (METROLINX, 2020). Reducing the policy headway will significantly affect essential workers' travel.

Therefore, the dispatching policy is

$$H_p = Min \begin{cases} H^{**} \\ H_c^u \\ h \end{cases} \qquad 79$$

Transit agencies can adopt the dispatching policy presented in Equation (79) during a pandemic and in the aftermath to help schedule transit.

4.5. Applications and discussion

4.5.1. A numerical example for a function of crowding penalty factor

An example function for the CPF is derived first. A 12 m long bus that has 44 seats (S=44) is used for this example. According to the Transit Capacity and Quality of Service Manual (TCRP Project A-15C, 2013), the design capacity is $C = 1.25S = 55$, and the crush capacity is $C' = 1.5S = 66$.

Table 7 shows the factors by which passengers increase their riding time values under different passenger load conditions. The variable "n" represents the number of passengers on board the vehicle, and "S" represents the number of seats in the bus.

Table 7 - Riding Time Multiplier Values under Different Loading Conditions

Load condition	Type of passenger	VoRT multiplier
n<(S/2)	For a seated passenger - one passenger per dual seat (no one in the next seat)	1
n>(S/2)	For a seated passenger - two passengers per dual seat	1.1
S<n<(1.25*S)	For a seated passenger	1.25
S<n<(1.25*S)	For a standing passenger	2.1
(1.25*S) <n<(1.5*S)	For a seated passenger	1.4
(1.25*S) <n<(1.5*S)	For a standing passenger	2.25

(Source – Information obtained from Transit Capacity and Quality of Service Manual, Third Edition 2013)

The average VoRT multiplier for passengers on the bus (S=44, C=55, C*=66) can be calculated by adding passengers one-by-one to the bus and calculating the average VoRT multiplier. Average VoRT multipliers with corresponding loading factors – number of passengers on the bus divided by the number of seats in the bus – are shown in the Figure 17.

.

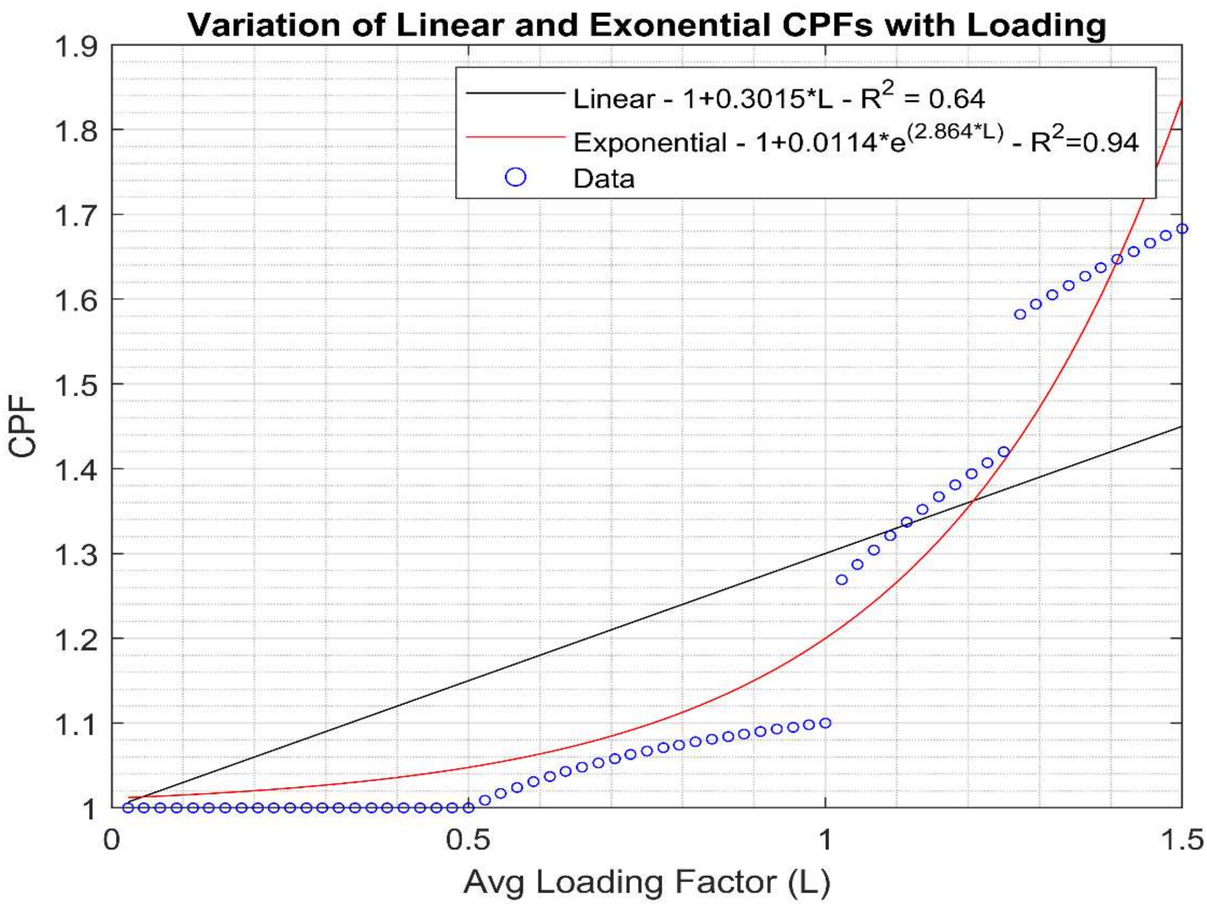


Figure 17 - Variation of the average CPF (VoRT multiplier) with L

Curves that have different functional forms can be fitted to this dataset:

Exponential:

$$Avg\ CPF = 1 + ae^{bL} = 1 + 0.0114 * e^{2.864*L} \qquad (\mathrm{R2}\ =\ 0.94) \qquad 80$$

Linear:

$$Avg\ CPF = 1 + aL = 1 + 0.3015 * L \qquad (\mathrm{R2}\ =\ 0.64) \qquad 81$$

CPFs (VoRT multipliers) are normally acquired from stated preference surveys. Results from such surveys determine CPFs for the corresponding levels of crowding. The number of data points that can be obtained depends on the number of levels utilized in the survey on crowding. If more levels are used, more data points can be obtained, and hence, the derived functional form is more accurate. However, more levels require a larger number of survey responses to maintain a certain accuracy level for the data points.

4.5.2. Numerical example

4.5.2.1. Optimum headway with the effect of crowding

In this section, we present a numerical example designed to determine the optimum headway with and without crowding costs. The parameters in Table 8 are assumed for an existing transit line.

Table 8 - Operational parameters of a transit line

Parameter	Value	Units	Reference
P	150	pass. /hr.	Assumed
Q	130	pass. Spaces/hr	Assumed
γ_w	15	\$/hr/pass.	(Hess et al., 2004) adjusted for inflation; (Hossain, 2019)
$\bar{t}$	0.5	hrs	Assumed
ϑ	0.3	-	Figure 17
S	44	seats	(TCRP Project A-15C 2003)
γ_r^o	10	\$/hr/pass.	(Ettema and Verschuren 2007) adjusted for inflation
λ_D	100	\$/dispatch	Assumed
$\bar{l}$	10	km	Assumed
D	20	km	Assumed

For the parameters shown in Table 8, the optimum headway without crowding from Equation (15) in the Headway LOS chapter, $H' = \mathbf{18}\, minutes$. Capacity headway, $H_c = 20\, minutes$ from Equation (23) in the Headway LOS chapter. In this example, passenger carrying capacity is assumed to be equal to bus size S to obtain a benchmark value, i.e., buses are dispatched when all seats are filled. Since $H' < H_c$, according to Equation (24) in the Headway LOS chapter, the bus route runs under optimum headway. Optimum headway with crowding is $H^* \cong \mathbf{15}\, minutes$ from Equation (53). Consequently, the effect of crowding costs reduces the optimum headway by nearly 3 minutes.

4.5.2.2. Sensitivity analysis of the optimum headway with crowding and pandemic health risks

To simplify the variation between β and L in Equation (72), let us replace $\emptyset + \sigma R$ by ϑ for a given level of pandemic severity:

$$\beta = 1 + \vartheta L \qquad 82$$

When there is no pandemic, R is zero and the value of ϑ is equal to $\emptyset$. As shown in Figure 16, the value of ϑ (rate of change in CPF with L and R - gradient) is higher in the presence of perceived health risks in taking transit. This value increases or decreases during different stages of a pandemic. For example, during a near lockdown (the stage when the infection rate is close to highest but not yet locked down and, hence, the health risks are highest), the value of ϑ is at the highest point, and during the aftermath of a pandemic, it is at the lowest point.

The sensitivity of H^* (optimum headway with crowding costs) and H' (optimum headway without crowding costs) are assessed against some key operational parameters. The effects of perceived health risks during different imaginary stages of a pandemic (expressed using different ϑ values) are explored. It should be noted that the effects of the perceived health risks during a near lockdown stage can increase the value of ϑ to more than two times the value of $\emptyset$ (Shin et al., 2021). The value of ϑ can also be nearly as low as $\emptyset$ (the effect of the pandemic health risks is negligible), and this situation occurs in the aftermath when passengers are no longer affected by COVID-19. However, the average variation of ϑ with the different stages of the pandemic, i.e., the variation of CPF with respect to the severity of the pandemic, is unique for a given community.

Accordingly, the values of optimum headway with crowding (H*) and without crowding (H') are allowed to change with changes in critical parameters, such as demand, bus size, and passengers' average trip time - Figure 18, Figure 19, and Figure 20. Average trip distance has a similar effect to that of average trip time, and the length of the bus route has a similar effect to that of bus size owing to the nature and positioning of these variables in Equation (52). These sensitivity analyses are also extended to represent different potential severity levels of a pandemic by increasing the value ϑ by 50% (H* @ 50%ϑ) and by 100% (H* @ 100%ϑ). The value of ϑ is assumed to represent no pandemic severity (R=0) and is equal to 0.3, which is the value of $\emptyset$ in the example presented in Figure 17. These results are compared with the optimum headway without crowding costs to show the significant impact of crowding costs on optimum headway.

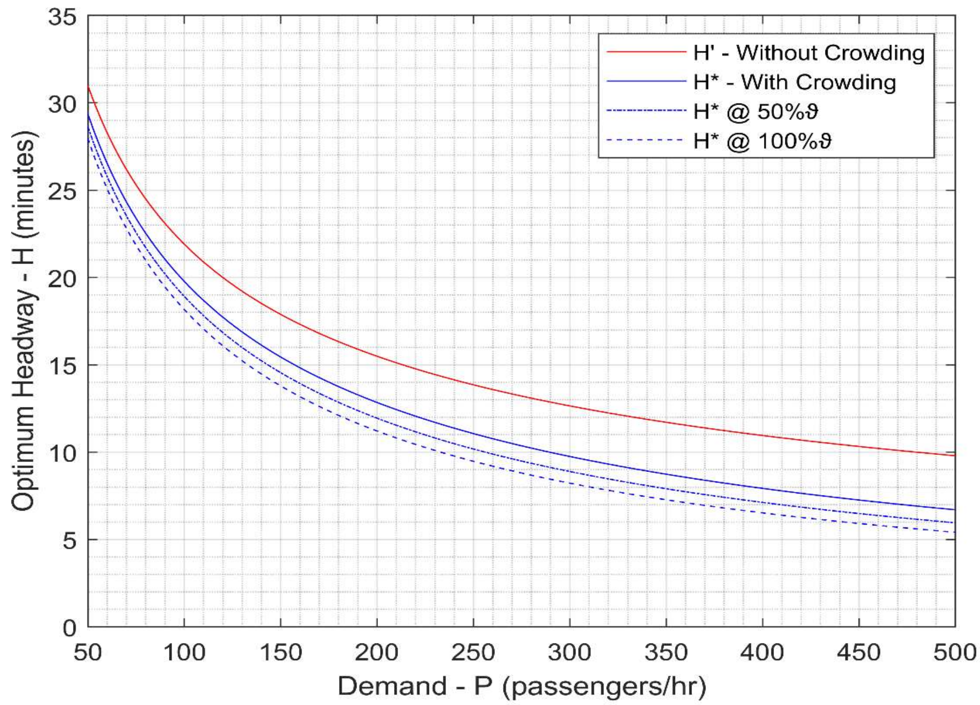


Figure 18 - Sensitivity of optimum headway to demand variations with and without consideration of crowding

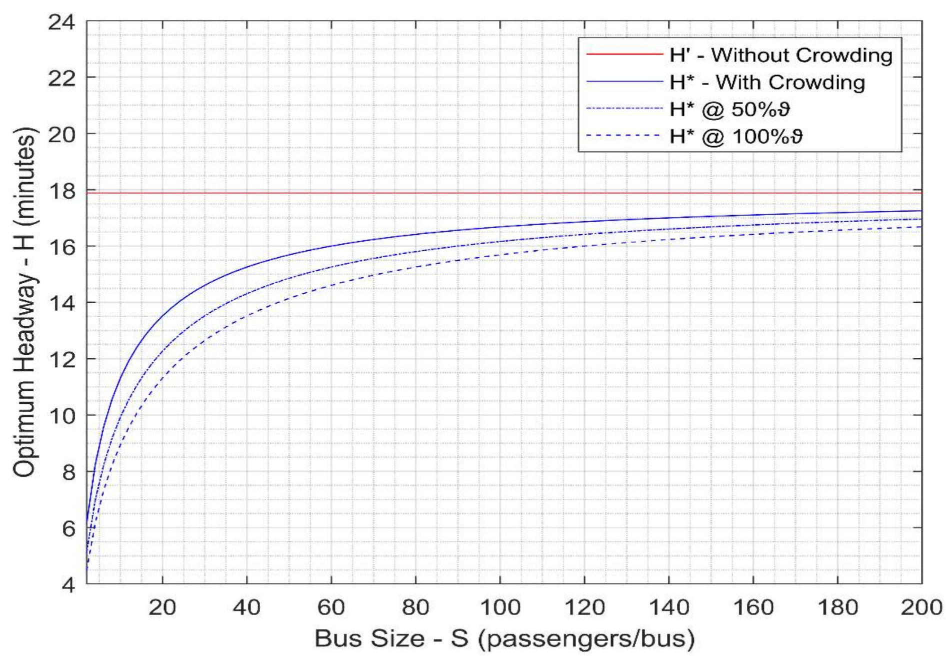


Figure 19 - Sensitivity of optimum headway to variations in bus size with and without consideration of crowding

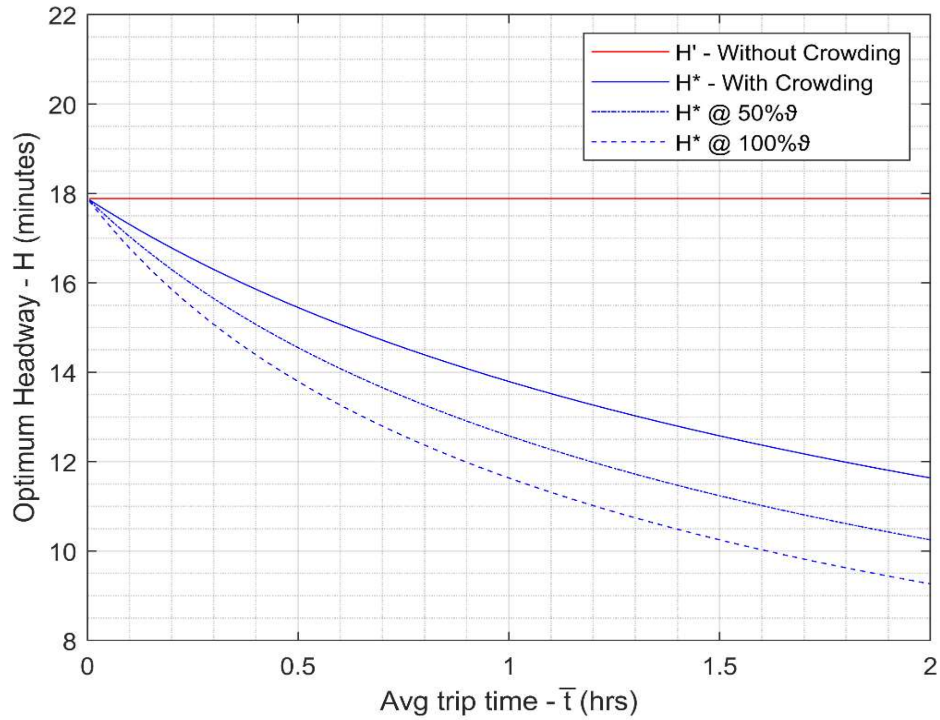


Figure 20 - Sensitivity of optimum headway to variations in average trip time of passengers with and without consideration of crowding

Figure 18 indicates that as demand increases, the values of both the optimal headways (H* and H') decrease. Higher demand implies higher passenger costs for the same headway value, which also means the same operator cost. An optimized headway balances this difference between operator and user costs by reducing headway. Figure 18 shows that as demand increases, the difference between H* and H' increases. This difference is more significant for higher values of demand, and at such values, there is a higher chance of the bus route running under capacity headway and not at optimal headways. The increasing difference between H* and H' indicates that as demand increases, crowding cost of passengers also increases, which decreases the value of optimum headway with crowding. Nevertheless, for all demand levels, the effect of crowding and health risks on optimum headway are distinct. This difference is sufficiently significant such that transit agencies should consider using a similar methodology to set frequencies for their transit lines because it will help retain and improve transit modal share while optimally allocating resources. However, according to the methodology suggested in this study, which balances the tradeoffs between increased operator and user costs, transit agencies will have to increase their resource contribution because H* will have to be reduced to match the increased user costs, which will in turn increase operator costs. Increased operator costs also mean that the transit agency needs increased funding/subsidy for their operations. During COVID-19, allocations of emergency funding have been made to promote and maintain sustainable transportation modes such as public transit (Mallett, 2020).

Figure 19 shows that smaller bus sizes are related to lower optimum headway with crowding. The optimum headway without crowding (H') stays constant as it does not depend on bus size. This result also demonstrates that smaller bus sizes are related to higher effects of crowding on H*. If the same dispatching policy is adopted for small and big buses for a particular demand, passengers on small buses will find the bus more crowded. Perceived health risks have been shown to amplify this effect: the difference between H*@100%ϑ' and 'H* increases as bus size decreases (Figure 19), which makes sense because as spaces become more confined, discomfort due to crowding and health risks grows.

As average trip time on a given transit line increases, the difference between H* and H' increases as shown in Figure 20. If the same frequency setting policy is adopted for different transit lines that have different average trip times, service will not be equitable/favorable, especially for the bus routes that have higher average trip times. This effect is amplified in the presence of COVID-19 health risks; as average trip time increases, the effect of health risks on passenger costs also increases. Unfavorable transit service may cause a decrease in transit ridership. As such, according to the developed model, optimum headway is sensitive to the average trip time. The degree of success of the methodology depends on accurately estimating the perceived health risks of different pandemic stages. The methodology presented in Section 4.4.3 in this chapter helps accurately estimate $\emptyset$ values that consider discomfort due to both crowding and health risks during different stages of a pandemic.

The proposed methodology supports managing the significantly reduced demand during a pandemic, and it also applies to normal demand levels. It is presented in a way such that it can be easily translated to potential future pandemic situations by deriving CPF functions for different pandemic severity levels. It is important to note that this method is developed for public bus routes and not for private bus routes.

4.5.3. H&C LOS with pandemic health risks

H&C LOS with and without the effect of the pandemic health risks can be evaluated by calculating the implied base VoRT for an existing bus operation. We will calculate the implied base VoRT to evaluate the H&C LOS first and extend the same example calculation to include the effects of the pandemic health risks. As the initial step, a figure to represent the base VoRT distribution of the passenger population is produced using the VoWT distribution presented in the Figure 13 of the Chapter 3, and assuming a value of 2.5 for the parameter ρ as suggested in literature, as shown in Figure 21.

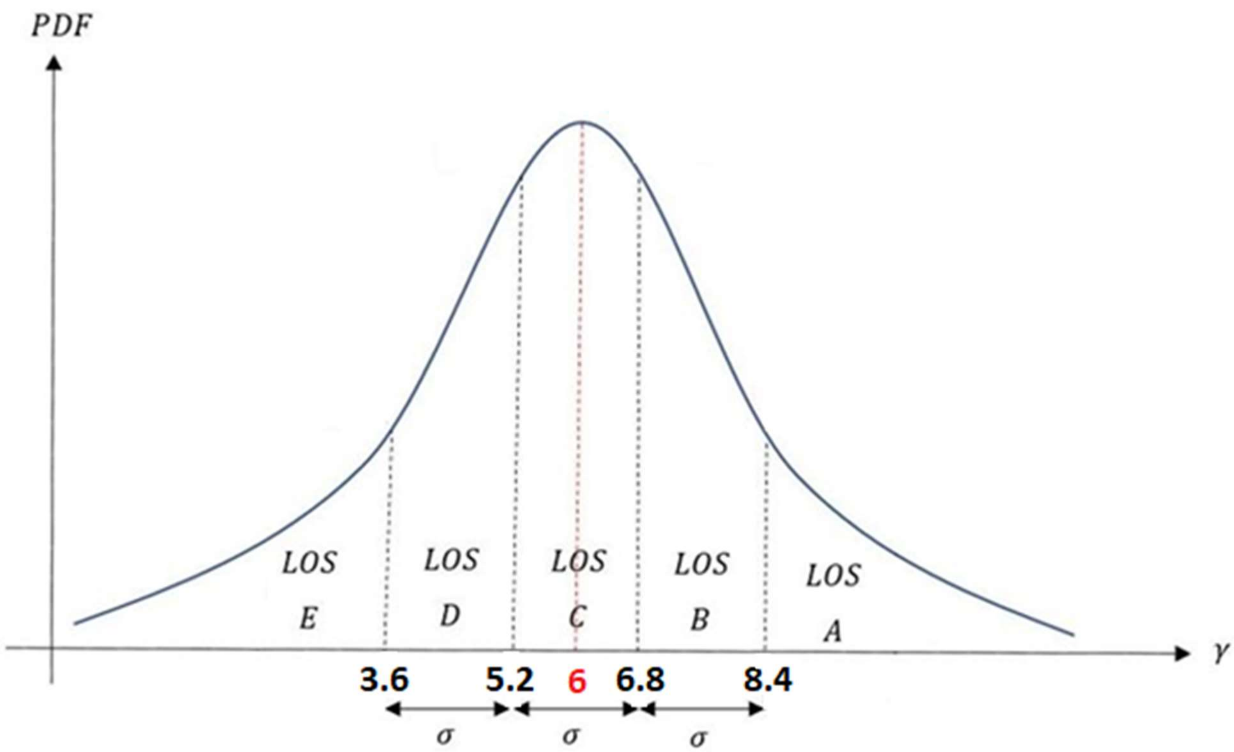


Figure 21 - Mean VoRT distribution

A bus route with the parameter values in Table 9 is considered.

Table 9 - Operational parameters of a bus route operating under a pre-selected headway

Parameter	**Value**	**Units**
P	150	pass. /Hr
ρ	2.5	-
$\bar{t}$	0.5	hrs
$\emptyset$	0.3	-
S	44	seats
H	0.333 (20)	hrs (mints)
λ_D	150	\$/dispatch
$\bar{l}$	10	km
D	20	km

For the bus route operation with the values presented in Table 9, the implied base VoRT can be calculated using the Equation (66) as 5.98 \$/hr./pass. This value falls under the LOS C for the VoRT distribution presented in Figure 21. Therefore, for the bus route operation consisting of the parameter values in Table 9, the H&C LOS is ‘C’.

If the same bus route is to operate with the same parameters under a condition with pandemic health risks. An average CPF of 2.5, closer to those reported in the literature – cited in section 4.4.2.2 – is assumed to represent the pandemic situation. Accordingly, the value of $\emptyset$ corresponding to a CPF of 2.5 can be calculated to be 2.64 using the Equation (43). Using this value instead of the value of $\emptyset$ in Table 9 will yield an implied base VoRT of 2.57 \$/hr./pass. according to the Equation (66). This value falls under the LOS E for the H&C LOS of the bus route under a pandemic condition in reference to the VoRT distribution represented in Figure 21. It should be noticed that the H&C LOS, under the influence of the pandemic health risks, has fallen significantly to the LOS E from LOS C which stresses the requirement of new interventions to maintain the transit attractiveness.

4.5.4. Operations under capacity

Consider a bus route that has the following parameters:

Table 10 - Operational parameters of a bus route

Parameter	Value	Units	References
P	200	pass. /Hr	Assumed
Q	180	pass. Spaces/hr	Assumed
γ_w	12	\$/hr/pass.	(Hossain 2019)
$\bar{t}$	0.5	hrs	Assumed
$\emptyset$	0.3	-	Figure 17
S	44	seats	(TCRP Project A-15C 2003)
γ_r^o	10	\$/hr/pass.	(Ettema and Verschuren 2007), adjusted for inflation
λ_D	150	\$/dispatch	Assumed
$\bar{l}$	10	km	Assumed
D	20	km	Assumed

Optimum headway under crowding for this bus route can be obtained using Equation (53) for the parameters presented in Table 10 as 17 minutes ($H^* = 0.3\ hrs = 17\ minutes$). In this example, it is assumed that there are no pandemic health risks involved (no additional cost due to health risks). If the design capacity of the bus in Table 10 is 50, the capacity headway can be calculated as 16.67 minutes using the Equation (52) in the chapter 'Headway LOS'. According to the dispatching policy – Equation (53) in the chapter 'Headway LOS' – the bus in this line now has to be dispatched using the capacity headway.

At near lockdown stages of the pandemic, vehicle capacities were restricted sometimes to half the seating capacity - blocking adjacent seats - by some transit agencies to adhere to social distancing regulations, for example in Calgary Transit, Alberta, Canada. In this situation, the load factor is 0.5, and the corresponding vehicle capacity C^u becomes bus size times the load factor, i.e., 44*0.5 = 22. Therefore, according to Equation (77), the capacity headway H_c is (22/150) *60 ≅ 9 minutes. According to the dispatching policy, Equation (78), buses need to be dispatched with a headway of 9 minutes on this line. This process can be used to meet the guidelines of health authorities that limit vehicular capacity.

Consider a situation where the bus route is running under the capacity headway – for the bus service having operational characteristics as in Table 10, and it is required to find out the H&C LOS of the bus service. Equation (70) can be used for this purpose. Assuming an average cruise speed of 30 km/hr. for the bus route – including stopping, boarding, alighting, and accelerating delays, provides a route travel time of 40 minutes for the route length of 20 km. Adding a layover time of 5 minutes provides an approximation of 45 minutes (0.75 hrs.) for the value of 'T'. The value of 'k' is assumed to be 120 $/hr./dispatch. Accordingly, the implied base VoRT - $\gamma_r^{o'}$ - can be calculated as 3.66 $/hr./pass. which falls under H&C LOS 'D' with reference to the Figure 21. As described in the section 4.5.3, assuming a value of 2.64 for the parameter $\emptyset$ to represent the pandemic health risks in a pandemic condition will further decrease the resulting value of $\gamma_r^{o'}$ reducing the H&C LOS experienced by the passengers.

4.6. Discussion and Concluding Remarks

This chapter integrates one of the key aspects of transit passenger comfort – crowding – into a measure assessing transit service level of service. Due to the nature of crowding discomfort which is mainly influenced by the practiced headway of a bus route, the quality of crowding cannot be assessed in isolation of the quality of headway. Owing to this background, this chapter proposes a methodology that assesses the headway and crowding LOS using a single measure. This suggested methodology is then extended to include a condition of a health pandemic that can affect passenger perception on taking transit due to the prevalent COVID-19 pandemic at the time of writing this thesis. The extended methodology can be adopted to any similar pandemic situation.

This chapter makes two important methodological contributions: it presents a new methodology to estimate the cost of crowding on a bus route, and it presents an approach that incorporates transit riders' perceptions of COVID-19 health risks as part of headway optimization. The perceived health risks of crowding are modelled using a crowding penalty factor (CPF), which is the value of ride time multiplier under crowding. It is postulated that discomfort, and hence cost, for a given level of crowding is also affected by the perceived health risks of the severity of a pandemic. An analytical model is then developed to obtain the optimum headway that considers average passenger cost of crowding. This model is then extended to be used in a context where the bus route is operating under the capacity headway.

The cost of crowding for users, which is also influenced by pandemic health risks, can be reduced by decreasing the headway of a bus route that operates under normal conditions. However, if the headway is decreased, operation costs increase. A trade-off is achieved by finding optimum headway. This revised optimal headway provides favorable conditions for both users and operators.

One of the main challenges transit agencies face during a pandemic is to schedule transit while managing crowding and increased operating costs. The presented methodology addresses this issue by deriving optimum headway that balances increased passenger costs due to crowding and operator costs in meeting the existing demand. As reported in the literature, passengers' perceptions of safety/health risks while riding transit and adapting to the changing perceptions during different pandemic stages must be addressed. Modelling the CPF with loading factor and pandemic stages helps transit agencies address this issue. A sensitivity analysis is presented to show the effect of different parameters and the pandemic on the optimum headway with and without crowding. Examples show the calculation and assessment of the H&C LOS of bus routes operating under policy headway and capacity headway.

5. Chapter 5 - Accessibility and Travel Time LOS

5.1. Introduction

The word accessibility has many facets even in the field of transportation. Therefore, the literature talks about both the importance and difficulty of measuring and quantifying accessibility (Polzin et al., 2002). One definition of accessibility is the ease of accessing locations (Luo and Wang 2003). Another definition is the ease of entering the transit system (Langford et al., 2012). As a result, the studies evaluating the concept of transit accessibility can be divided into two parts known as spatial and temporal accessibility. Temporal accessibility deals with mainly the availability of the service in time of day, wait time and origin-destination travel time. Spatial accessibility on the other hand deals with the proximity to transit stops/stations and origin-destination connectivity (Polzin et al.., 2002). There is another body of research that explores the spatial transit accessibility in two aspects known as *system accessibility* and *system-facilitated accessibility* (Langford et al., 2012; Malekzadeh & Chung, 2020). While system-accessibility – also known as local transit accessibility (LTA) – is concerned with the access 'to' transit system, system-facilitated accessibility is concerned with the access 'by' transit system. LTA is considered an important and a critical aspect of transit performance as system-facilitated access becomes meaningless if users have difficulties entering the system (Langford et al.., 2012). Boisjoly & El-Geneidy (2017) point out the methods used for evaluating LTA in the literature as part of assessing transit service quality. Their study found that of the 343 surveyed transit practitioners around the world the majority were using the count of bus stops as an accessibility measure (Bree et al., 2020). However, methods employed for transit accessibility assessment range from gravity-based and utility-based models to simple stop counts (Malekzadeh and& Chung, 2020). Bree, Fuller, and Diab et al. (2020) assessed several LTA measures by validating them against ridership data and found that the gravity-based model – "filtered frequency accessibility measure calculated using a 400m network buffer and a Butterworth filter with a distance decay bandpass value of 250m" – performed best.

This chapter deals with the service operations aspects of an urban bus route rather than the planning aspects under which system-facilitated accessibility comes. LTA has been explored in two aspects – density of bus routes that is concerned with the urban form and density, and bus stop spacing concerned with the access distance along the transit line. Density of bus routes is more of a planning aspect of the bus system while bus stop spacing is more an operational aspect that can be changed for a given bus route to attend to passenger requirements. The spacing between the bus routes/routes is chosen to minimize the access distance to the routes. Therefore, the distance travelled orthogonal to the bus routes affects the spacing of the bus routes whereas distance travelled along the bus route affects only the stop spacing problem (Wirasinghe & Ghoneim, 1981). While major bus stops are fixed – such as transfer points, major trip attractors/generators, time points, etc., minor bus stops can be changed with minimal intervention/cost, and therefore come under the purview of operational planning in most cases. In the special case of LTA, as it pertains to bus transit, most of the studies in the literature have focused on distance to bus stops (Langford et al.., 2012). It is reasonable, to a certain extent, as bus stops are mostly accessed through the mode of walking. Biba et al., (2010) found that the distance to a bus stop to be a primary factor affecting transit usage. O'Sullivan and Morrall (1996) also found that walking distance to a transit stop as one of the main factors affecting the choice of using transit among many other factors like age, income, car ownership and travel distance. The empirical studies in the literature suggest that on average, pedestrians will walk 400m to catch a bus factored for effects of street connectivity, grade of the street, population and pedestrian crossings (TCRP Project A-15C, 2013).

An interesting, yet cumbersome fact, is that transit accessibility in terms of transit stop spacing (along the bus route) and the in-vehicle travel time (ITT) are intertwined. Increasing the number of bus stops along the route – reducing the average walking distance of the passengers along the route – with the intention of increasing the level of accessibility and hence the accessibility LOS also increases the travel time along the route. More bus stops require stopping and alighting/boarding passengers which takes time that otherwise would have been saved for the passengers already riding the bus. Reducing the average stop spacing also reduces the average bus travel speed, therefore restricting the distance of the route reducing the system-facilitated accessibility. This is evident from the fact that many measures assessing system-facilitated accessibility take into account the ITT (Hillman & Pool, 1997; Leake & Huzayyin, 1979; Tribby & Zandbergen, 2012).

It is obvious that travel time is one of the major factors affecting transit QOS while most of the time the reason behind using a personal vehicle is for saving travel time. An average person will compare the transit travel time with that of using their next possible alternative which most often is their private vehicle. Therefore, one of the most used measures to account for this factor is the ratio between transit and auto (private vehicle) travel times that is in-vehicle transit travel time divided by the in-vehicle auto travel time (TCRP Project A-15C, 2013). Lower transit travel time will induce an increase in transit ridership. A lower transit travel time will not always decrease the transit ridership as there are other benefits of using public transit as compared to private vehicles. Some of these benefits are lower costs, hassle-free traffic congestion, ability to utilize the transit travel time to undergo other activities - e.g., reading, phone calls, personal work. These benefits will outperform the potential travel time saving from private vehicles up to a certain extent. TCQSM provides passenger and operator perspectives for several levels of transit-auto travel time ratio. The ratio of transit to auto travel time comes down to an aspect known as the directness of the routes which is entirely a planning concern and not an aspect that affects the service. The only component that affects the travel time is the local stop spacing that can be changed during the service with little effort while not changing the routes. From a passenger perspective, the negative correlation between bus stop spacing and the ITT means that two main attributes of the transit quality compete with each other stressing an optimal solution to serve the passenger interests. As a result, this chapter looks into finding a solution for this issue.

The methodology presented is built on the optimum spacing model developed by Wirasinghe, (2021a) and Wirasinghe & Ghoneim, (1981). A methodology is presented to find an implied VoT for the service that represents the Accessibility and Travel-Time LOS (A&T LOS). An example calculation is designed to demonstrate the method to be followed.

5.2. Methodology

As the methodology to denote A&T LOS depends on the optimum spacing model developed by Wirasinghe (2021) and Wirasinghe and Ghoneim (1981), the related key assumptions and their model are briefly described below.

The passengers whose origins are on the route will always walk to the nearest bus stop irrespective of the stop being upstream or downstream. This is a reasonable assumption since the average bus cruising speeds are around ten times higher than the average walking speeds. Passengers whose origins are not on the route are assumed to walk to the nearest route and then also walk along the bus route to the nearest stop. It is further stated that, if the time taken for boarding and alighting a passenger is constant, then the increase in travel time for any passenger due to boarding and alighting passengers is independent of the stop spacing problem. Buses on the route stop at all the stops along the bus route. The stop spacing problem is investigated as it pertains to the local stops in between major stops – generators. The goal of this chapter being to demonstrate the new approach to evaluating LOS, only a short route consisting of local stops in between two generators as terminal points is being considered. Depending on these assumptions, a continuum approximation has been derived for the sum of passenger access costs, additional travel time cost of passengers on board lost due to stopping, and operator cost associated with additional time lost stopping at a bus stop for a stopping distance centered at an x distance from the origin of the bus route. Hence, the total cost per day per unit distance at x – Z(x), can be shown below.

$$Z(x) = \frac{\gamma_k P(x)S(x)}{4} + \frac{\gamma_r C(x)\emptyset(x)}{S(x)} + \frac{n\lambda_B \phi(x) + \lambda_s}{S(x)} \qquad 83$$

The first, second and third terms in the right-hand side of Equation (83) represent the access cost of passengers, additional ride time cost of passengers on board and additional operations cost of the operator respectively.

$P(x)$ = number of passengers boarding and alighting the bus service in a day per unit distance at the distance x.

$S(x)$ = stop spacing centered at x

$\bar{S}$ = average stop spacing of the bus route

$C(x)$ = number of passengers on board passing a point located from an x distance from the origin of the route in a day

$\emptyset(x)$ = amount of additional time taken due to the stopping of a bus

N = number of bus dispatches per day

λ_B = cost of owning and operating a bus measured in $ per unit time

λ_s = cost of a bus stop per day

γ_k = mean value of a unit of access distance of the passengers

γ_r = mean value of a unit of riding time of passengers

$\emptyset(x)$ has been shown as a function of the average cruising speed and the acceleration and deceleration rates of the buses by Wirasinghe and Ghoneim (1981) as follows.

$$\emptyset(x) = \frac{u(x)\left[\frac{1}{a} + \frac{1}{d}\right]}{2} + \Omega \qquad 84$$

$u(x)$ is the average cruising speed of the buses at the distance x over the day. a and d are the acceleration and deceleration rates of the buses respectively. Ω is the opening/closing time of the doors and the lag times between the door close/open and passengers start boarding/alighting. It should be noted that the time taken to board/alight passengers has not been included in the total cost function within the user cost of time lost and operator cost of operating due to stopping a bus. This is because it has been shown that the time taken to board/alight passengers depends linearly only on the number of passengers and is therefore independent of the number or the distance between the bus stops along the route. It is assumed that the level of crowding does not affect the time taken for a passenger to board or alight the bus.

We assume that the demand variations along the route do not have significant differences between the stop distances along the route. Stop spacing within a route section tend to be similar in distance (Wirasinghe 2021). Therefore, for the rest of the methodology development, we assume a uniform stop spacing – s – for the bus route. The optimum stop spacing for the route that minimizes the tradeoffs in between access cost, travel-time cost and operating cost is achieved by minimizing the total cost Z over the route. This can be achieved by integrating Equation (83) over the route and replacing s(x) by s as shown in Equation (85).

$$\int Z(x) = \frac{\gamma_k S \int P(x)}{4} + \frac{\gamma_r \int C(x)\emptyset(x)}{S} + \frac{N\lambda_B \int \phi(x) + \lambda_s}{S} \qquad 85$$

To further simplify Equation (85), it is assumed that the route only has bus stop signs to demarcate a bus stop along the route which makes the cost of a bus stop per day – λ_s – negligible. This cost is significantly higher if the bus stops placing needs; pavement strengthening, setting up shelters and seats, regular maintenance over vandalism and litter cleanup, and ITS equipment to show real-time bus schedule and cannot be disregarded in the equation. The stopping cost – $\phi(x)$ – depends on the cruising speed of the bus which is normally not significantly different along the bus route and therefore an average value – ϕ – can be used without changing the objective function drastically. After these simplifications, taking the partial derivative of the objective function – Equation (85) –in terms of the stop distance – S – provides the optimum spacing for the route – S'.

$$S' = 2\left[\frac{\gamma_r \phi}{\gamma_k}\frac{\int C(x)}{\int P(x)} + \frac{N\lambda_B \phi D}{\gamma_k \int P(x)}\right]^{1/2} \quad 86$$

D is the distance of the route in usual notation as in the previous chapters. Wirasinghe (2021) further states that for a route with uniform stop spacing, $\frac{\int C(x)}{\int P(x)}$ can be interpreted as half the mean trip length of the passengers - $\bar{l}/2$. $\int P(x)$ is the number of passengers that have boarded and alighted in the bus route in a day which is equal to twice the number of passengers those have boarded in a day. If the total demand to board the bus route in a day is $\hat{P}$, then $\int P(x)$ is equal to $2\hat{P}$. Accordingly, the Equation (86) can further be modified as,

$$S' = 2\left[\frac{1}{2\gamma_k}\left[\gamma_r \phi \bar{l} + \frac{N\lambda_B \phi D}{\hat{P}}\right]\right]^{1/2} \quad 87$$

Unlike the expression for the optimum headway without crowding effects in Chapter 3, the expression for the optimum spacing has two VoT parameters. According to the suggested methodology in the chapter 'Proposed Approach' to denote the LOS for the combination of attributes access and travel-time, it is required to have a single implied VoT parameter that can be compared with the distribution of the values of that parameter. Therefore, it is now required to convert the expression in Equation (87) for the optimum spacing into an expression which contains only one VoT parameter. This can be achieved through the two methods suggested below.

5.2.1. Method 1

Let us consider the following expression inside the Equation (87).

$$\gamma_* = \gamma_r \phi \bar{l} + \frac{N\lambda_B \phi D}{\hat{P}} \quad 88$$

For the parameters considered, Equation (87) represents a linear function of a random variable which can be considered a separate random variable. Let us denote this new random variable as γ_* where the mean – μ_* – and the SD – σ_* – can be obtained using the mean and SD of γ_r as follows.

$$\mu_* = \phi \bar{l} \mu_r + \frac{N\lambda_B \phi D}{\hat{P}} \quad 89$$

$$\sigma_* = \phi \bar{l} \sigma_r \quad 90$$

Where μ_r and σ_r are the mean and SD of the distribution of γ_r.

Accordingly, the optimum spacing as denoted by the Equation (87) becomes a function of a ratio of random variables γ_* and γ_{k}. The ratio $\gamma_*/\gamma_{\mathrm{k}}$ can also be represented through a unique random variable γ_{rk} where the mean – μ_{rk}– and the SD – σ_{rk} – can be approximated using the mean and SD of γ_k and γ_* as follows (Kendall, 1994).

$$\mu_{rk} \cong \frac{\mu_*}{\mu_k} - \frac{Cov(*,k)}{{\mu_k}^2} + \frac{{\sigma_k}^2\mu_*}{{\mu_k}^3} \qquad 91$$

$$\sigma_{rk} \cong \frac{\mu_*}{\mu_k}\left[\frac{{\sigma_*}^2}{{\mu_*}^2} - 2\frac{Cov(*,k)}{\mu_*\mu_k} + \frac{{\sigma_k}^2}{{\mu_k}^2}\right]^{1/2} \qquad 92$$

Where μ_* and σ_* are the mean and SD of the distribution of γ_r, and μ_{k} and σ_{k} are the mean and SD of the distribution of γ_k. Equations (91) and (92) materialize only when the values of γ_k are greater than zero - $\gamma_k > 0$.

Therefore, using the derivations obtained in Equations (88) – (92), Equation (87) can be modified as,

$$S' = 2\left(\frac{\gamma_{rk}}{2}\right)^{1/2} \qquad 93$$

Accordingly, using the existing average stop spacing in the bus route – $\bar{S}$ – in the Equation (93) in place of S', the implied VoT that represents the combined LOS of the attributes access and travel-time – γ'_{rk} – can be obtained as below.

$$\gamma'_{rk} = \frac{\bar{S}^2}{2} \qquad 94$$

Let us call γ'_{rk} the implied pseudo VoT (IPVoT). Now using the method suggested in the chapter 'Proposed Approach' using μ_{rk} and σ_{rk} of the γ_{rk} distribution, the A&TT LOS grade can be obtained.

5.2.2. Method 2

Equation (87) has been developed assuming mean value for the VoT parameters of the passengers in the route. But, for the evaluation process of the LOS, we are deriving an implied value that denotes the LOS of the route using an equivalent optimum operation with existing operational parameters of the service. While passenger VoT parameters are random variables, the implied values that are derived in the process are not random variables but values that we compare with the distributions of the corresponding random variables.

In order to convert the two VoTs into a single VoT, we use the method suggested under the 'H&C LOS' chapter where VoAT - γ_k - and VoWT - γ_w - can be represented as a multiplication of the VoRT - γ_r. Hence,

$$\gamma_k = \omega\gamma_r \qquad 95$$

Where ω is the value of the ratio γ_k/γ_r. Value of ω has been identified in literature as 2 for the North American context (Wardman 2004).

The Equation (87) can be modified using the implied VoRT - γ_r', the implied VoAT - γ_k' and $\bar{S}$ to model an equivalent optimal operation for an existing operation of a bus route as follows.

$$\bar{S} = 2\left[\frac{1}{2\gamma_k'}\left[\gamma_r'\phi\bar{l} + \frac{N\lambda_B\phi D}{\hat{P}}\right]\right]^{1/2} \qquad 96$$

Unlike the random variables, the implied values can be modified using Equation (95). Accordingly,

$$\bar{S} = 2\left[\frac{\phi}{2\omega}\left[\bar{l} + \frac{N\lambda_B D}{\hat{P}\gamma_r'}\right]\right]^{1/2} \qquad 97$$

Therefore, the implied VoRT that represents the A&TT LOS is obtained as,

$$\gamma_r' = \frac{2N\lambda_B D}{\hat{P}(\omega\bar{S}^2/\phi - 2\bar{l})} \qquad 98$$

The implied VoRT obtained from Equation (98) can be compared against the VoRT distribution of the passengers to obtain the A&TT LOS. Here, the value of ω takes into account the effect of the VoAT.

5.3. Numerical Example

A numerical example is presented here that is designed to exhibit the proposed methodology in method 2 for calculating the implied VoRT that can be used to obtain the A&TT LOS of an existing bus operation. Method 1 is not explored due to the unavailability of the covariance values between the VoRT and VoAT. The parameters in Table 11 are assumed for a bus route.

Table 11 - Operational parameters of a bus route

Parameter	Value	Units
$\hat{P}$	1000	pass./hr.
N	10	
ω	2	
$\bar{S}$	0.3	km
ϕ	0.416/60	Hrs (25 Seconds)
λ_B	150	$/hr.
$\bar{l}$	7	km
D	20	km

According to Equation (98), the implied VoRT - γ_r' can be calculated to be 5.02 $/hr./pass. According to the VoRT distribution presented in Figure 21 of the Chapter 4, the implied VoRT of 5.02 $/hr./pass. falls into the LOS D. Therefore, the A&TT LOS of the bus route is LOS D.

5.4. Discussion and Concluding Remarks

The value of implied VoRT representing the A&TT LOS is sensitive to the values of average stop distance and average time lost due to a stop. For example, increasing the average stop distance of the bus route in the numerical example to 400 meters from 300 meters, decreases the implied VoRT to 1.86 $/hr./pass. In the same way, decreasing the value of time lost due to a stop from 25 seconds to 15 seconds, reduces the implied VoRT of the bus route to 2 $/hr./pass. This example has been calculated assuming that the bus route does not incur a significant cost for the bus stops per day – most bus stops are just a sign pole. If this is an express bus route in an urban area, the cost of a bus stop per day as represented by λ_s can incur a higher cost. In such an instance, the implied VoRT will increase.

Method 1 presented in section 5.2 is also a promising way of obtaining the implied VoRT. This however requires the covariance value between the VoRT distribution and VoAT distribution. To the best of the author's knowledge, there are no values currently presented in the related literature regarding the covariance between different VoTs. Formulation is easier for the case where the covariance values are zero. But this is unlikely because the VoRT and VoAT are two random variables derived for the same source, hence dependent.

6. Chapter 6 - Reliability LOS

6.1. Introduction

6.1.1. Importance of reliability

Reliability is one of the most important factors affecting the transit mode and departure time choice (Abkowitz et al., 1979). Abkowitz (1979) defines reliability as the "invariability of service attributes which influence the decision of travelers and transportation providers". Improving transit reliability increases passenger satisfaction and loyalty by increasing the efficiency of the system (Diab et al., 2015). Transit passengers in Calgary, Canada found reliability to be more important than ride comfort in choosing transit (Habib et al., 2011). Improving reliability is also beneficial for transit operators as it will increase efficiency, reduce operating costs and increase farebox revenue by attracting and retaining users (Diab et al., 2015). Therefore, improving reliability benefits both users and the transit operator.

6.1.2. Improving Reliability

Reliability is understood mostly in terms of the variability in the bus service – variability in travel time and variability in arrival time at stops hence the variability in passenger wait time. For frequency-based services, reliability is also understood in terms of headway variation (Mohammadi, 2023). In reliability assessment schemes found in the literature, the stop-to-stop travel time includes the running time and passenger handling time at the stop - described in detail in Section 6.2 in this chapter. The variations in the time stop-to-time stop travel time – hereafter referred to as link travel time – can therefore be attributed to the variations in the traffic conditions, en-route incidents, variations in the number of passengers boarding and alighting as well as on board. The variation in the travel time can be transformed to the variation in the average waiting time of the passengers depending on the variation of the arrival time around a scheduled departure time at a stop. Therefore, unreliable wait times occur due to the variation of travel times (Chang, 2010).

In a perfect case where the demand rate for boarding and alighting is uniform at each stop of a scheduled bus route, getting delayed at a certain stop is translated to an increased dwell time along the route. This is because now there are more passengers to board at that particular stop and more passengers to alight in subsequent stops which consume additional time. This causes the stop-to-stop travel time to further increase along the route in a cycling effect sometimes creating a bus bunching effect. Therefore, it is important to improve and maintain reliability at acceptable levels on a bus route although it is not practical to completely eliminate unreliability. While improved reliability will reduce the average travel time – by reducing the slack time – the time saved on the route through the reduced travel time can be converted to more service runs or saved operating time and hence cost. Measures aiming at improving reliability must therefore reduce the variation in the travel time or the dwell time.

Accordingly, transit agencies have utilized different measures to improve reliability. Implementing transit signal priority, reserved bus lanes, limited stop services (express services), bus rapid transit (BRT) and BRT like systems have reduced the travel time variations while utilizing new low floor or/and articulated buses and smart card systems have reduced the dwell time at a stop (Diab et al.., 2015). Intelligent transportation systems together with automatic vehicle location and automatic passenger count systems have contributed to improving reliability through providing real-time information.

It is found that providing real-time information on the crowding levels inside transit vehicles can increase the reliability of the service. Real-time information on crowding helps some passengers to choose a departure time with less crowding which distributes the peak of crowding reducing the crowding levels inside buses. This will decrease the number of passengers boarding, reducing the time required for boarding. This will also reduce the time required for alighting for a given number of passengers as now the existing level of crowding inside the bus is less, reducing the time taken for a passenger to alight. Eliminating both these delays will improve the regularity of headway and travel time, and reduce bus bunching (Drabicki et al., 2022).

6.1.3. Role of the scheduled travel/arrival/departure time on reliability

Despite implementing different strategies to improve reliability, there can still be a significant amount of variability to the transit service in a bus route which can induce a negative passenger perception. Thus, most transit agencies use holding control measures in further reducing unreliability. This is done by selecting a few bus stops on the route and making them timed-stops where transit vehicles arriving at these stops have a scheduled departure time. Vehicles arriving early will wait until the scheduled departure time and the vehicles arriving late will depart as soon as the passenger handling finishes at time points. Wirasinghe (1993) states that the scheduled departure time at a particular time point is derived by adding a slack time to the mean travel time

– including the passenger handling time – between a previous adjacent timepoint and adding it to the scheduled departure time of the upstream adjacent time point. The slack time is intended to absorb a significant portion of the late arrivals at a particular bus stop so that the bus route will have a satisfactory on-time performance (OTP).

Deriving scheduled departure times by selecting suitable slack time for each time-point and selecting the number and location of the time points are two important steps in this process. While too many time points will slow down the operation increasing wait time, ride time of passengers and operating cost, too few time points are not successful in improving the reliability which will increase passenger costs – riding, waiting and a delay penalty, and operating costs as the risk of bus bunching requires additional buses in the operation. Also, while larger slack times retard the operation, too small slack times lack control over the travel time variations (Wirasinghe, 1993). Wirasinghe and Liu (1995) provides a comprehensive background on selecting the number and location of the time points. Liu and Wirasinghe (2001) show how the slack time can be calculated for some special cases. Klumpenhouwer and Wirasinghe (2018) provide a solution for the general case using a numerical approach. A transit agency may select the scheduled arrival/departure/travel time depending on the percentage of OTP they want to maintain in the bus operations (Diab et al., 2015) – the longer the scheduled travel time, the larger the number of busses that are on time, depending on the actual travel time distribution. These measures are widely utilized in the transit industry in increasing the OTP or the schedule adherence of transit routes which is a measure for assessing the reliability performance of transit services.

6.1.4. Measuring reliability – passenger and operator-oriented measures

Accurate measuring of transit reliability facilitates the ability of transit operators to understand and improve reliability. Accordingly, there are many measures assessing reliability that can be found in the literature (Abkowitz et al., 1979; Chang, 2010; Huo et al., 2014). Such measures can be divided mainly into two categories: (1) travel-time oriented measures, (2) operator-oriented measures. Some of the measures in the first category are: the compactness of the travel time distribution – as measured by the coefficient of variation, likelihood of extreme delays – as measured by the % of bus arrivals that are taking N minutes (decided by the transit agency) longer than the mean/scheduled travel time, excess waiting and riding time. Some of the measures in the second category are: OTP or the schedule adherence – as measured by the percentage of bus arrivals/departures withing a range of time around the scheduled arrival/departure time (e.g., 1 minute early to 5 minutes late) – used mostly for scheduled services, headway regularity – as measured by the coefficient of variation of the headways – used mostly for operations with a headway less than 10 minutes (frequency based services) (Camus et al., 2005; Mohammadi, 2023), standard deviation around the mean/scheduled arrival/departure time – depending on the service being scheduled or frequency based, average difference between mean and scheduled arrival/departure time. Out of these measures, the most popular operator-oriented measure is the OTP (Camus et al., 2005; Diab et al., 2015). The TCQSM presents the reliability LOS as different levels distinguished depending on different OTP ranges (TCRP Project A-15C, 2013).

OTP tends to disregard some of the factors affecting the passenger's perception. OTP introduces only the number or the percentage of late arrivals and not the amount of time that a transit vehicle is late. According to the definition of on-time buses from different transit organizations (Camus et al.., 2005), the on-time buses are also later in different amounts than the scheduled/mean arrival/departure time. Improving OTP will not essentially improve the reliability for the passengers. OTP as a reliability measure fails to recognize the amounts of travel-time and wait-time delay of the passengers. Most of the studies on reliability pay attention mainly to the travel time variations but rarely emphasize its impact on the wait time variations of the passengers which is a significant factor affecting passengers mode choice in the long term. Average passenger wait time is a function of the headway distribution (Furth & Muller, 2006) and therefore can be related to the travel time distribution.

6.1.5. Proposed approach through an optimum scheduled travel time

Selecting a scheduled travel time in between time points or selecting a scheduled arrival/departure time has different implications for passengers and operators. For a given travel time distribution, increasing the scheduled travel time will increase the total travel time cost, increase operating cost – more operating time and decrease the additional wait/ride time costs of the passengers due to late bus arrivals/departures. Therefore, the scheduled travel time should be selected in a way such that it minimizes the trade-offs between passenger travel time plus operating cost and the summation of the additional passenger wait time and ride time costs. Reasoning behind selecting these specific cost components as they relate to the scheduled travel time and the reliability is discussed in detail in section 6.2 – Methodology.

In line with the procedure presented in Chapter 2 – Proposed Approach, the reliability LOS is evaluated with reference to an implied value of the time component. The implied value of time referring to the reliability LOS is derived using the mathematical model simulating the optimum scheduled travel time. Accordingly, the study formulates the different passenger and operator cost components as a function of the scheduled travel time. The summation of these cost components – the objective function – is used to obtain the optimum scheduled travel time between time points. The implied value of time component is derived using the function for the optimum scheduled travel time to represent the reliability LOS. Section 6.3 presents example calculations designed to show the practical uses of the developed method.

6.2. Methodology

Formulating a methodology for the reliability assessment of realistic situations is complex (Newell, 1977; Wirasinghe, 1993). Improving reliability is a different problem than the one investigated here. The present problem involves figuring out the number and location of time-points and finding the optimum slack times at those time-points (W. Chen & Chen, 2009; Huo et al., 2014; Klumpenhouwer & Wirasinghe, 2018; Wirasinghe, 1993; Wirasinghe & Liu, 1995a). This study investigates the ability to assess the Reliability LOS that considers both the passenger and operator concerns for a bus route that is currently implementing a bus holding control strategy. As the goal of this study is to demonstrate the concept of incorporating user perception and operator concerns to a LOS measure and then combining them to determine a combined measure, an ideal case of a schedule-based bus route where all major stops are time points is considered. Stops with low demand where buses stop only occasionally are not considered. Applicability of the suggested methodology for a real-world case would require further fine tuning of the model. In a real-world situation, not all the stops are time-points. A generalization of some aspects of a real-world bus route has been discussed in section 6.4 of this chapter. The methodology proposed here intends to approach an estimate for the reliability LOS using average parameter values for the bus route.

This approach is proposed considering a time period of twenty-four hours – a day – but can be used for a different time period depending on sufficient data availability. The inter-time-point travel time is scheduled – this includes the passenger handling time as well. It makes sense to assume that passengers boarding at a certain bus stop do not feel a disutility for waiting during the passenger handling time after the bus arrival as it takes the anxiety of the waiting passengers away.

Passenger handling time includes the time taken to open/close doors, the delay between door opening/closing and passengers start/finish boarding/alighting, and to board/alight passengers. This will be included as part of the riding time. Accordingly, each time point has a scheduled arrival time – except, in some cases the starting stop has a scheduled departure time. In such circumstances we assume the scheduled departure time is equal to the scheduled arrival time at the starting stop of the route. Therefore, all the stops are time points; that is, they have a scheduled arrival time.

Figure 22 - A bus route with n timed-points as stops

The bus route under consideration – as shown in Figure 22 – has n stops where all the stops are timed points and the inter-timepoint link travel time from timepoint $i-1$ to timepoint i is S_i, which helps form the scheduled departure times at each stop. We define the scheduled link travel time between two consecutive timepoints as the time between the scheduled arrival times. Accordingly, the average inter-timepoint travel time for the route - S - can be calculated as the average of S_i for all the n number of links in the route. Theoretically, all stops can be used as timepoints (Laskaris, G. et. al. 2020). Fu and Yan (2002) showed that an all stop control strategy performed better in reducing the average passenger wait time at the cost of higher bus travel times. The inter-timepoint link travel time – hereafter referred to as the link travel time – can be represented through a random variable with a unimodal travel time distribution due to the randomness associated with the link travel time. The literature has also followed a similar approach in representing the link travel time (Z. Liu et al., 2013; Newell, 1977; Wirasinghe, 1993; Wirasinghe & Liu, 1995b). It is required to obtain a representative link travel time distribution for the entire route as represented through the random variable t. One of the factors affecting link travel time of a given link – hence the reliability of the finishing stop - is the randomness of the previous link travel time. Wirasinghe, (2021b) has shown how the reliability measured by the OTP of a stop is associated with the reliability of the previous stop. For the ease of problem formulation, it is best if this dependable randomness can be avoided. Accordingly, the travel time of the link i, i.e., t_i, is defined as the time between the 'scheduled departure time' of the stop $i-1$ and the actual time for arriving at stop i. This isolates the randomness associated with the link travel time from that of the previous link so that the randomness of a link can be considered independent.

There are two main parts to the randomness of t_i – the variation in running time between stops that depends on the en-route traffic condition and the variation of passenger handling time that depends on the demand to board/alight at each stop. As the factors affecting these two parts are similar and independent from one link to another, t_i for all i can be assumed to be independent random variables having identical distributions as also suggested in the literature (Wirasinghe, 1993, 2021b; Wirasinghe & Liu, 1995a) with different standard deviations and mean values. Given the nature of link travel times that contain a minimum possible value for the travel time, the shape of these distributions is suggested to be similar to a gamma or a log-normal distribution (B. Y. Chen et al., 2017; Z. Dai et al., 2019; Rahman, Wirasinghe, et al., 2018; Uno et al., 2009; Wirasinghe & Liu, 1995b). Accordingly, a representative random variable for the link travel time t can be defined as,

$$t = \sum_{i=1}^{n} \frac{1}{n} t_i \qquad 99$$

It should be noted that t is a linear combination of the random variables t_i which is again a random variable having a unimodal distribution with the probability density function $f(t)$ where $t \geq t_m$. t_m is the minimum possible link travel time. The values for t_i can be obtained using the schedule for the bus route and the automatic vehicle location (AVL) data. The mean and the standard deviation of the random variable t can be obtained through statistical theory as described in Kendall (1994) using means and standard deviations of t_i. Hence the probability distribution of t – for all links – can be shown as in Figure 23 below with respect to the scheduled travel time S.

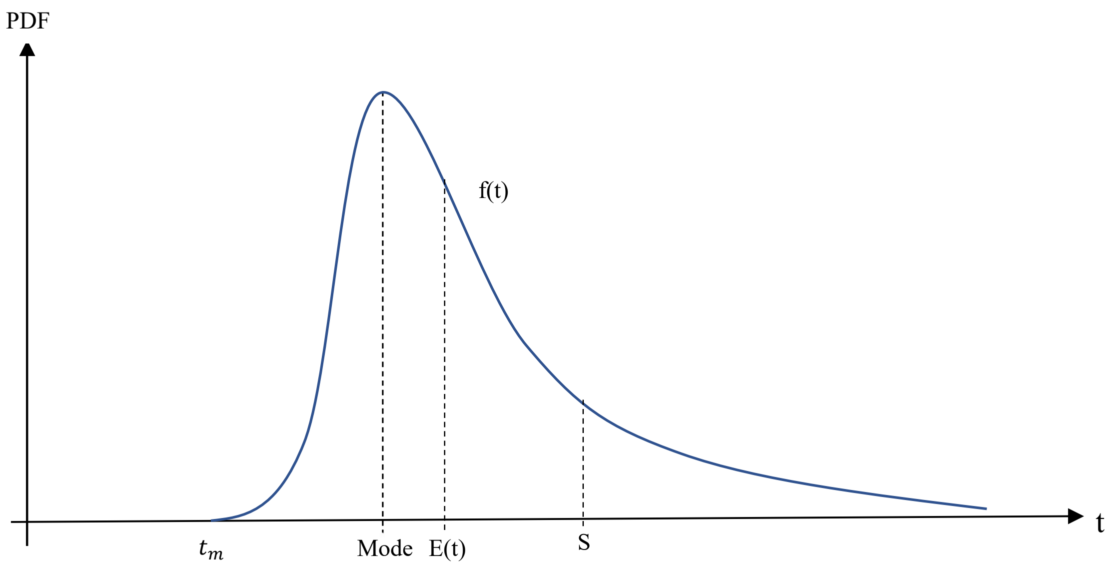


Figure 23 - PDF of the link travel time with scheduled travel time

An existing bus route with n inter-timepoint links with different link travel times and scheduled travel times can be modelled as a bus route with n number of inter-timepoint links with t link travel time (which is a random variable) and with S scheduled travel time for the purpose of reliability assessment using the suggested methodology. Here, the passengers are assumed to board/alight only at the timed stops and not in the intermediate stops. The impact of this assumption on the intended results is discussed in section 6.4 with suggested counter measures.

As there is a scheduled travel time for all the links and this helps form the potential trip time for a passenger, the scheduled travel time in the links is assumed to be budgeted by the passengers. That is, the budgeted travel time is valued at the rate of normal VoRT. The additional travel time caused by the delay due to the unreliability is valued at a higher rate than the normal VoRT – multiplied by a factor to account for the inconvenience caused by the delay.

6.2.1. Costs associated with unreliability

Wirasinghe (2021) shows that the expected delay in the link travel time can be reduced by increasing the scheduled travel time in the link. It is intended to minimize the expected total generalized cost associated with the reliability – the scheduled travel time and the average delay – in the route.

The expected delay in the link travel time, E(d) can be denoted as,

$$E(d) = \int_{s}^{\infty} (t - s) f(t) dt \qquad 100$$

Therefore, an optimum scheduled travel time in the link is achieved by minimizing the trade-off between the disutility of the expected delayed travel time and the budgeted trip travel time. Accordingly, this study formulates an analytical model on the expected total generalized cost of the bus route operation to obtain the optimum scheduled link travel time in a route.

6.2.1.1. User Costs

Delays in the scheduled travel time cause inconvenience to the passengers causing them to value their delayed traveling time at a higher rate than the value of travel time under normal conditions. This cost is taken into consideration through an additional travel time cost where passengers' value of travel time is weighted by a factor 'a' to account for the inconvenience. To minimize the delays, the scheduled travel time must be increased leading to an increase in the budgeted travel time of passengers.

As discussed in the introduction, the reliability of the bus route operation is expected to change with the choice of the scheduled travel time between the time points subject to the inter-time-point travel time distribution. This is because the travel time variation, the delay in travel time, the delay in waiting time and the operating cost, are going to be affected by the scheduled travel time and the travel time distribution. Depending on the scheduled travel time, the passenger's budgeted travel time will change. That is, if a passenger must travel a few time points along the bus route, increasing the scheduled travel time will increase the budgeted travel time of the passengers. Therefore, the objective includes the generalized total expected cost of the passengers budgeted travel time – BTTC - and can be calculated as follows.

Average trip distance of a passenger $\bar{l}$ (km) – can be normally obtained using origin-destination data. The average running speed of a bus in the route can be calculated as D/nS – note that this speed excludes the time associated with any delays when arriving at a time point. D is the distance of the bus route in kms, S is the scheduled travel time between bus stops measured in hours and n is the number of links (i.e., between consecutive transit stops) in the bus route. The average travel time of a passenger can be calculated to be $\bar{l}nS/D$ by dividing the average trip distance from the average speed on the route. If the passenger demand for the entire route (demand to board) is P

passengers per day, the average total budgeted travel time cost of the passengers in a route in a day – BTTC – can be calculated as,

$$BTTC = \frac{P\bar{l}nS\gamma_r}{D} \qquad 101$$

where γ_r is the value of ride time of passengers in \$/hr./passenger. Equation (101) represents the budgeted travel time cost of all the passengers – including the ones who had a delayed arrival at their destination stop.

All the passengers who board the bus must alight at some point on the bus route where there is a probability of them being late for their destination. Only the passengers alighting at a specific station will be affected by the delayed arrival of the bus at that particular station. The passengers riding through a timed stop which has a delayed arrival are assumed to experience no disutility as long as he/she arrives on time for their respective destination stops – a notion supported by Seneviratne (1990). Therefore, only the portion of passengers who are late at their respective destination stops are having an additional cost for being delayed. According to the representative travel time distribution t that can be derived using the bus travel time data for a day in the route, the probability of a bus being late at a stop - $p(t > S)$ - can be calculated as,

$$p(t > S) = \int_{S}^{\infty} f(t)dt \qquad 102$$

It can be fairly assumed that the portion of the passengers having a delayed arrival at their destination stop is approximately equal to the portion of buses that are delayed in a day. It is assumed that the delays in bus arrivals are purely associated with the reliability of link travel time. As described in section 6.1.2, one of the most significant factors affecting the link travel time variation is the number of passengers boarding and alighting. Therefore, in the peak hour, higher passenger demands will result in higher late arrivals at stops. This leads to the probability of a late arrival – the link travel time being greater than the scheduled travel time – to have a linearly increasing relationship with the passenger demand. Owing to this background, it is safe to associate the portion of passengers having a late arrival is approximately equal to the portion of late arrivals in the buses in the route in a given day. Applying this methodology to a given hour or a certain smaller time period will increase the accuracy of the relationship between the portion of late passengers and portion of late buses while decrease the accuracy of the link travel time distribution due to limited amount of data.

And the average delay in the bus route for a passenger is expressed in Equation (100). The demand to board the bus route in a given day – P passengers per day– is equal to the demand to alight the bus route in the same day. The passengers experiencing a delayed travel time are assumed to value their delayed portion of the travel time at a rate of $a\gamma_{\mathrm{r}}$ where a is the delayed riding penalty factor (DRPF) that accounts for the additional inconvenience of the delayed travel time. Therefore, the expected total additional cost of the delayed travel time – ATTC – can be calculated as the multiplication of the passenger demand – P, percentage of trips having a delayed travel time – Equation (102), average delay in the travel time – Equation (100), and the average value of a unit of delayed travel time - $a\gamma_r$.

$$ATTC = aP\gamma_r \int_S^{\infty} f(t)dt \int_S^{\infty} (t-S)f(t)dt \tag{103}$$

There are two components for the wait time of the passengers. The average wait time of the passengers due to the size of the headway and the amount of additional wait time of the passengers due to the delay of the bus in arrival at the bus stop. The average wait time of the passengers, that depends on the size of headway, also depends on the trip purpose of passengers, and the type of service. Ansari Esfeh et al., (2020) provide a comprehensive background on evaluating the average passenger wait time for bus routes with different service types and passengers with different trip purposes. Changing the scheduled travel time does not affect the average wait time of the passengers as it does not affect the headway – increase in scheduled travel time increases the round-trip time requiring more buses in the route operation to maintain a given headway. Therefore, the average wait time cost of the passengers has no trade-off with the scheduled travel time, and hence is not included in the objective function. This enables the monetization of only the waiting time after the scheduled departure time – additional wait time.

Only the delayed buses will cause an additional waiting cost for passengers. Only the portion of passengers experiencing a delayed bus arrival at their origin stop, out of the total passenger demand to board in a day, are having an additional wait time cost which is valued at a higher rate than their normal value of wait time γ_{w}. To account for this increased disutility the value of a unit of delayed wait time of the passengers is obtained by multiplying γ_{w} by a factor 'b' – the delayed waiting penalty factor (DWPF). For the passengers who are experiencing an additional wait time, their expected delayed wait time is equal to the average expected delay at a stop as expressed by Equation (100). Therefore, the additional expected wait time cost of the passengers in a day – AWTC – can be calculated as the multiplication of the passenger demand – P, percentage of passengers having a delayed travel time – Equation (102), average delay in the arrival time – Equation (100), and the average value of a unit of delayed waiting time - $b\gamma_w$.

$$AWTC = bP\gamma_w \int_S^{\infty} f(t)dt \int_S^{\infty} (t - S)f(t)dt \tag{104}$$

A bus exceeding the scheduled travel time can have another fixed cost for some passengers that is independent of the amount of the delayed time. The passengers having a planned arrival time can miss the purpose of their trip causing them inconvenience. This can be considered through a delay penalty. An accurate estimation of the delay penalty includes the number of passengers with planned arrivals, types of their purposes at the destination, what percentage of passengers with planned arrivals miss their purpose due to the delay – some passengers can miss their appointments when delayed more than a certain time. For the purpose of this study, the delay penalty – γ_{p}, measured in dollars per passenger, is considered a fixed amount for all the passengers when a bus is delayed. As a delay penalty is affecting only the trip cost of the passengers experiencing a delayed arrival at their destination, the delay penalty cost for the bus route in a day – DPC – can be calculated as the multiplication of the passenger demand – P, the percentage of passengers having a delayed arrival time – Equation (102), and the delay penalty per passenger - γ_p.

$$DPC = P\gamma_p \int_S^{\infty} f(t)dt \tag{105}$$

6.2.1.2. Operating Cost

Increasing the scheduled travel time of a link increases the total travel time of the route that the transit operator has to budget. To maintain a predetermined headway, the operator has to introduce more buses to the route when the total route travel time is increased. This cost is considered through the operating cost of a bus per hour on the route. The delay occurs when the actual travel time exceeds the scheduled travel time and causes additional operating time for the operator. In a route with multiple inter time point links, it is possible for the bus to catch up for the delay caused at one time point at the next time point. The operator can increase the speed a little or there can be normal stops where there are no passengers to board or alight saving the stopping delay at that stop reducing the total inter timepoint travel time. Although the probability of this scenario is comparatively less, here the operator does not incur any additional operating cost. What can really cost the operator more is the additional travel time that occurs at the end of the route. This is countered when the layover time at the end of the route is larger than the delay at the end. Normally the layover time is significantly higher than any potential delay at the end of the route if not for any en-route incident. We assume that no operator ends their shift at the scheduled trip end time of the route so that there is no additional crew cost associated with the delay at the end of the route – no overtime pay. The only additional cost the operator experiences when this happens is the fuel cost from the additional running time on the route. But we consider this cost to be negligible.

The budgeted route travel time is equal to $nS + T'$, where n is the number of links in the bus route and T' is the layover time of the route in hours. If the operating cost of a bus per hour – including the capital, maintenance, fuel, and crew costs – is λ in \$/hr, the operating cost of a round trip in the route is given by $(nS + T')\lambda$. In addition, if the number of dispatches per day in the route is N, the operating cost of the bus route per day – OC – can be calculated as,

$$OC = (nS + T')\lambda N \qquad 106$$

6.2.1.3. Expected total cost of the bus route

The objective function - $Z(S)$ – is the summation of the quantities BTTC, ATTC, AWTC, DPC and the OC as defined above. Although the BTTC and OC have no variation with the travel time variation, other components depend on the travel time variation and hence there is an expected cost. Therefore, the objective function is also an expected cost of the bus route in a day that needs to be minimized.

$$\begin{aligned} Z(S) = \frac{P\bar{l}nS\gamma_r}{D} &+ aP\gamma_r \int_S^\infty f(t)dt \int_S^\infty (t-S)f(t)dt \\ &+ bP\gamma_w \int_S^\infty f(t)dt \int_S^\infty (t-S)f(t)dt \\ &+ P\gamma_p \int_S^\infty f(t)dt + (nS + T')\lambda N \end{aligned} \qquad 107$$

Equation (107) can be further simplified to be,

$$Z(S) = \frac{P\bar{l}nS\gamma_r}{D} + P(a\gamma_r + b\gamma_w)\int_s^{\infty} f(t)dt \int_s^{\infty} (t - S)f(t)dt + P\gamma_p \int_s^{\infty} f(t)dt + (nS + T')\lambda N \qquad 108$$

6.2.2. The optimum scheduled travel time

An optimum scheduled travel time 'S^*' for the bus route can be obtained by taking the partial derivative of the objective function in the Equation (108) in terms of S and making it equal to 0.

$$\frac{\partial(Z(S))}{\partial S} = \frac{P\bar{l}n\gamma_r}{D} - \left[f(s)\left(\int_s^{\infty} tf(t)dt - S\big(1 - F(s)\big)\right) - \big(1 - F(s)\big)^2\right] P(a\gamma_r + b\gamma_w) + n\lambda N - P\gamma_p f(s) \qquad 109$$

$F(s)$ is the cumulative probability of the function $f(t)$ up to S. $f(s)$ is the probability of the travel time being equal to S. $\int_s^{\infty} tf(t)dt$ is the average travel time in a link for the delayed buses. Simplification of Equation (108) to Equation (109) is presented in the Appendix E

Let,

$$Q(S) = f(S)\left(\int_S^{\infty} tf(t)dt - S\big(1 - F(S)\big)\right) - \big(1 - F(S)\big)^2 \qquad 110$$

In the chapters "Headway and Crowding LOS" and "Access and Travel Time LOS", it has been shown how the VoWT can be represented as a multiplier of the VoRT using the values that can be found in the literature. Therefore, let γ_w be represented as,

$$\gamma_w = \theta\gamma_r \qquad 111$$

where θ is the ratio of VoWT to VoRT.

Using Equations (110) and (111), Equation (109) can be further simplified as,

$$\frac{P\bar{l}n\gamma_r}{D} - Q(S^*)P\gamma_r(a + b\theta) + n\lambda N - P\gamma_p f(S^*) = 0 \tag{112}$$

Equation (112) holds true for the optimum value of S for the bus route and for the mean value of ride time of the passengers. Applying the Equation (112) to an existing bus operation using the existing parameter values, an implied value of ride time – $\gamma_r{}'$ – representing the reliability of the bus route can be obtained as,

$$\gamma_r' = \frac{\gamma_p f(S) - \left(\frac{n\lambda N}{P}\right)}{\left(\frac{\bar{l}n}{D}\right) - Q(s)(a + b\theta)} \tag{113}$$

It is necessary to ensure that the value of the scheduled travel time at which the expression in the Equation (112) becomes zero, is a minimum – not a maximum or an inflection point. For this to be realized, the second derivative of the objective function – Equation (108) – at S^* needs to be greater than zero.

The second derivative of Equation (108) can be obtained as,

$$\begin{aligned}\frac{\partial^2(Z(S))}{\partial S^2} = (P(a\gamma_r + b\gamma_w))\Bigg[&3f(S^*)\big(1 - F(S^*)\big) \\ &+ S^* f'(S^*)\big(1 - F(S^*)\big) - E(t)f'(S^*) \\ &+ f'(S^*)\int_0^{S^*} tf(t)dt\Bigg] - P\gamma_p f'(S^*)\end{aligned} \tag{114}$$

Equation (114) can be further simplified to be,

$$\begin{aligned}\frac{\partial^2(Z(S))}{\partial S^2} = (P(a\gamma_r + b\gamma_w))\Bigg[&3f(S^*)\big(1 - F(S^*)\big) \\ &- f'(S^*)\left(\int_{S^*}^{\infty}(t - S^*)f(t)dt\right)\Bigg] - P\gamma_p f'(S^*)\end{aligned} \tag{115}$$

The operators consider bus travel time on a link as an on-time arrival if the bus travel time in the link is less than or equal to the scheduled travel time. A typical reliability measure from the operator's perspective is the percentage on-time arrivals in a given time period (TCRP Project A-

15C, 2013). Owing to the issues like bus bunching and negative passenger perception associated with lower reliability, operators try to maintain reliability as measured by the percentage on-time arrivals as high as possible. TCQSM states that passengers assume a transit service is highly unreliable when the percentage on-time performance falls below 70% (TCRP Project A-15C, 2013). The highest level is where the on-time performance is higher than 95%. Therefore, it can be fairly assumed that operators try to maintain the on-time performance at a percentage higher than 70%. That is, in an inter-time-point travel time distribution, 70% is higher than the median that accounts for 50%. As discussed earlier, the literature suggests that such a travel time distribution is positively skewed (Rahman, Wirasinghe, et al., 2018). While the 50th percentile marks the mean, median and mode of a normal distribution, the 50th percentile for a positively skewed distribution is located further away from the mode/peak in the tail of the lognormal distribution. The higher the skewness, the further away the median – 50th percentile is from the peak of the distribution. Therefore, an operator wanting to choose a scheduled travel time that keeps the on-time performance at a rate higher than the 70th percentile must choose a value that falls within the right-hand-side tail of the distribution. For any value in the right-hand-side tale of the distribution, the corresponding derivative of the probability density function is a negative value as also shown in the Figure 24.

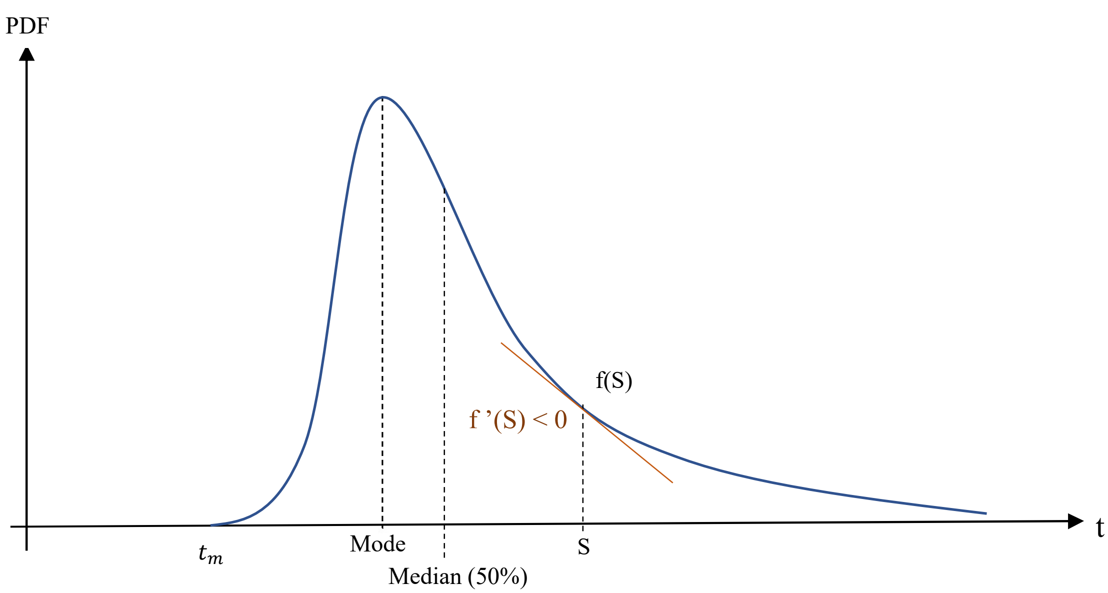


Figure 24 - Derivative of the travel time distribution at the value of scheduled travel time

Owing to the background described here, it is possible to rule that the derivative of the probability density function (PDF) at the scheduled travel time is negative. With this knowledge, it can be shown that the second derivative of the objective function – Equation (115) – is positive.

Since this study is using an existing transit operation, the scheduled travel time used by the operator – S – is used instead of the optimum scheduled travel time - S^* - to come up with the implied VoRT corresponding to the reliability LOS of the bus route. Assuming that transit operators use an S value that makes the on-time performance of the bus route higher than 70% - which is the case almost all the time – it is shown that $f'(S^*)$ is negative. Therefore, the term $P\gamma_p f'(S^*)$ is negative as the values of P and γ_p are positive. According to the definitions of the parameters in the term $P\gamma_r(a + b\theta)$, it is positive. Since $f(S^*)$ and $\left(1 - F(S^*)\right)$ are probabilities – therefore are positive values in the range of 0 to 1 – the first term within the parentheses – $3f(S^*)\left(1 - F(S^*)\right)$ - is positive. The term – $\int_{S^*}^{\infty}(\mathrm{t} - \mathrm{S}^*)f(t)dt$ – is positive as the value of t ranges from S^* to infinity. $f'(S^*)$ is negative. Accordingly, the term $-f'(S^*)\left(\int_{S^*}^{\infty}(\mathrm{t} - \mathrm{S}^*)f(t)dt\right)$ is positive for the range of S^* that a transit agency would normally use. Therefore, the second derivative of the objective function – Equation (115) – is positive and the Equation (112) represents a minimum, allowing the Use of Equation (113) for obtaining an implied VoRT representing the reliability LOS.

6.3. Example calculation

An example calculation is designed to show the applicability of the methodology developed in this chapter for determining the reliability of LOS. We assume the following parameters of an existing bus route.

Table 12 - Values of operational parameters of an example bus route

Parameter	Name	Value	Units
P	Demand for boarding	1000	passengers/day
n	Number of inter-time-point links	6	No.
N	Number of bus dispatches per day	30	dispatches/day
D	Length of the route	20	km
$\bar{l}$	Average trip distance of the passengers	10	km
λ	Operating cost of a bus per hour	120	\$/hr./bus

γ_p	Delay penalty of a passenger	20	$/passenger
a	Delayed riding penalty factor	1.25	
b	Delayed waiting penalty factor	1.25	
θ	Ratio of VoWT to VoRT	2.5	
μ	Mean link travel time	15/60	hours
σ	standard deviation of the link travel time	5/60	hours
S	Average scheduled link travel time	20/60	hours

The link travel times – t – are distributed log-normally.

Using the Equation (113), the implied VoRT representing the reliability of a bus route can be obtained as 7.45 $/hr./pass. The value of S – 20/60 hrs, in this example corresponds to an on-time performance of 85% and can be calculated as follows.

A function for the link travel time can be denoted as,

$$f(t) = \frac{1}{t\sigma\sqrt[2]{2\pi}} exp\left[-\frac{(\ln t - \mu)^2}{2\sigma^2}\right] \qquad 116$$

$$P(t > \bar{S}) = \int_{20/60}^{\infty} f(t)dt = 85\% \qquad 117$$

According to the methodology suggested in Chapter 2 - 'Proposed Approach', the reliability LOS can be obtained by comparing the derived implied VoRT with the VoRT distribution of the passengers. It is important to note the difference between reliability and the reliability LOS. Reliability is normally measured using OTP in the industry and the focus of this chapter is on the measuring the reliability LOS and not the reliability. According to the VoRT distribution in Figure 21 of Chapter 4 – Headway and Crowding LOS, the derived implied VoRT represents a LOS grade 'B'. Figure 25 and Figure 26 show the behavior of $Q(S)$ and γ_r' with S respectively modeled using Matlab version 2020b (MathWorks, 2020) – the code used in this exercise is attached in the Appendix D. The range of S is selected in a way that it covers the travel times greater than 60th percentile assuming the transit agencies are aiming to maintain an average on-time performance greater than 60%.

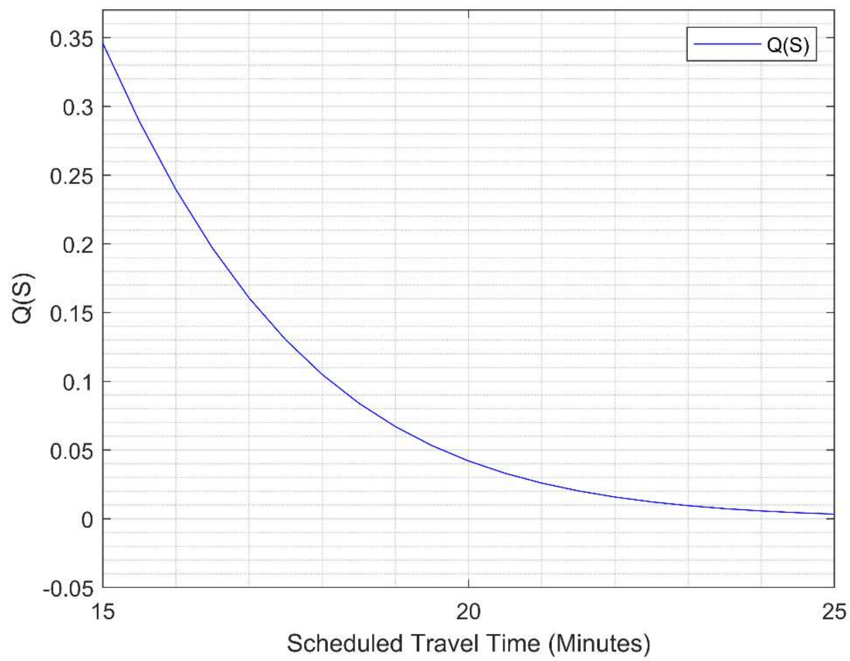


Figure 25 - Behavior of Q(s) with Scheduled Travel Time

According to Figure 25, the value of $Q(S)$ reduces with the increasing scheduled travel time and converges to zero.

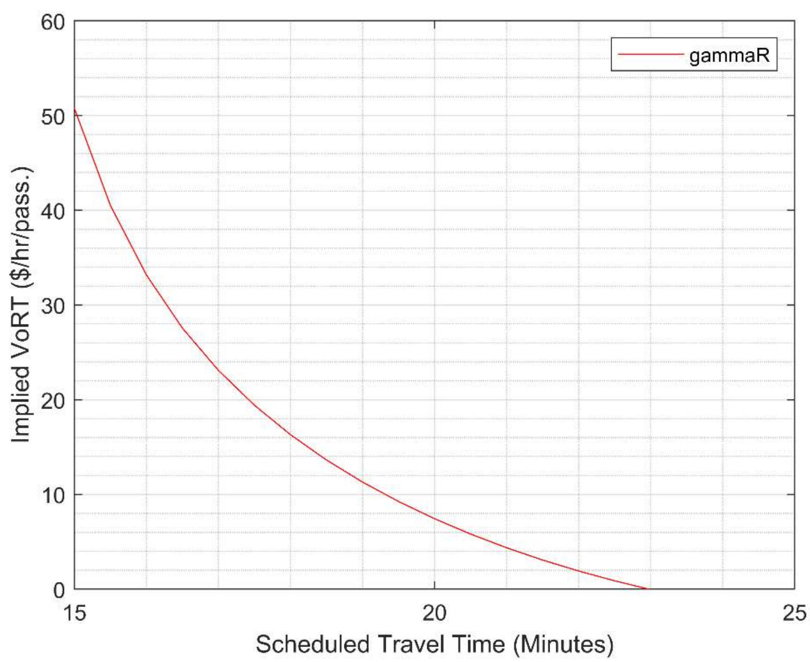


Figure 26 - Behavior of the implied VoRT with the Scheduled Travel Time

According to Figure 26, the implied VoRT for the operation in the example decreases with the scheduled travel time for the range of S chosen. Although increasing S increases the reliability – as measured by OTP – and decreases the additional travel/wait time cost and the delay penalty of the passengers, it increases the budgeted travel time cost of the passengers and the transit service operating cost at a higher rate within the range of S considered.

6.4. Discussion and concluding remarks

This study has been carried out for a bus route with a schedule – scheduled inter-time-point travel times or scheduled departure/arrival times at time points. According to TCQSM, a scheduled bus operation is best suited for a bus route with headways longer than 10 minutes (TCRP Project A-15C, 2013). For a high frequency bus service, the more appropriate, hence widely used measure of reliability is a measure of the headway variation than the travel time variation (Mohammadi, 2023). But the proposed method of assessing the reliability LOS may be extended to a high frequency bus route.

The proposed methodology utilizes a travel time distribution that is the average of the inter-time-point travel time distributions in a route that consists only of the timed points. The average delay derived from the travel time distribution depends on the variation of the data, i.e., the standard deviation. The cost of the additional travel time and the additional wait time of the passengers depends on the average delay calculated in the proposed methodology. This makes sense as the passengers will only be boarding/alighting at the timed points in the proposed approach. This is different when there are intermediate stops in between timed points where people board/alight. In most practical situations, the average delay time for the stops between timed points (normal stops) can be smaller at the beginning and reach the calculated value at the end. Therefore, applying the same average delay for the calculation of additional travel time and wait time is an overestimation of these costs. An approximate and simple correction for this issue can be achieved by using half of the average delay time at the end of the inter-time point link for the calculation of additional travel time and wait time costs.

The developed methodology does not require the probability distribution to take any particular shape. As Equation (113) contains only the values of the probability density function and the cumulative probability function at S, this methodology can be applied to any type of travel time distribution whether it is log-normal, normal or gamma.

The suggested approach in this chapter tries to address a few of the gaps in the reliability LOS literature. Studies that consider the wait time of the passengers in measures of the quality of the reliability are scarce, irrespective of the fact that several studies have emphasized the importance of the wait time on reliability. Studies that take the wait time aspect into account are either concerned with only the effects of reliability on the wait time or have taken a heuristic approach in formulating a solution. The suggested approach taking the wait time aspect into a measure assessing reliability is one of the important contributions of this chapter. Similarly, studies finding an optimum scheduled travel time through a closed form solution of an analytical model by also taking wait time into account are rare. The suggested approach addresses this gap as well. Methodologies found in the literature deriving an optimum scheduled travel time or slack time through an analytical model assume near perfect situations such as a route with one inter-time-point link, and do not consider routes with many such links at least with equal distances/scheduled travel times. This is partly due to the inter dependability of link travel times. The suggested approach here addresses this issue to a certain extent by assuming a route with many inter-time-point links those having equal scheduled travel times and by suggesting a method to isolate the inter dependent randomness between adjacent link travel times. The approach is further extended to make it applicable to a real situation by suggesting a methodology to change/remodel the data collected from a bus route operation to be used in the suggested methodology that approximates the assumed situation. In conclusion, the proposed methodology in this chapter makes the current literature on the reliability level of service and optimum scheduled travel time, one step closer to the real-world situation.

The approach presented in this chapter provides the tools for transit agencies and other interested parties to assess the reliability LOS of a bus route that considers both the perspectives of passengers and the operator. Deciding on how much reliability – based on OTP – to maintain in the route by changing the scheduled travel time, affects passenger's budgeted travel time cost, passenger's additional travel time and additional wait time costs, and operations cost in the bus route. The methodology suggested provides the ability for the transit agency to choose a level of reliability through choosing an inter timepoint scheduled travel time (S) to provide a desired reliability LOS depending on what the passengers' values of time are worth in the bus route. It provides insights into how the scheduled travel time may be changed to achieve the required reliability LOS. The methodology presented has potential future work into extending the suggested approach that considers the different inter timepoint travel time distributions of inter timepoint links rather than an average value which will improve the accuracy of estimation.

7. Chapter 7 - A Combined Measure for the Overall LOS and Conclusion

7.1. Introduction

This chapter is aimed at addressing one of the critical issues in the transit LOS literature – developing an overall LOS measure for a bus route that incorporates both passenger and operator concerns. A comprehensive review of this issue has been carried out in section 1.2 of Chapter 1 while section 1.3.3 of the same chapter investigates the efforts towards an overall LOS measure in industry and academia with their associated limitations. The approach suggested here to come up with a combined measure for the overall LOS of a bus route depends on the LOS framework proposed in Figure 1. Accordingly, the methodology presented here combines the three sub-LOS measures derived in the previous chapters that cover all the attributes presented in the LOS framework.

7.2. Background

This study has developed three LOS measures concerning several aspects of a bus transit operation while taking the passenger and operator concerns into account. What follows in this section is a snapshot of the sub-LOS measures developed in chapters 4, 5, and 6. The variable notations for different LOS measures are different from one chapter to another. Therefore, for each LOS measure, notation descriptions of the variables used in each corresponding chapter are stated under the respective LOS measure.

7.2.1. Headway and Crowding LOS – H&C LOS

H&C LOS considers the aspects of waiting time and riding time of the passengers. Crowding affects the quality of riding and therefore the effect of crowding is considered by penalizing the base value of riding time. An implied VoRT representing the quality of headway and crowding attributes is obtained through modeling the existing bus operation using the equation for the optimum VoRT. The closed form solution for the optimum VoRT is developed using an objective function that minimizes the trade-offs between the passenger costs associated with waiting and riding in a crowded condition and the costs of operation.

The implied value of ride time is obtained using the following equation from Chapter 4.

$$\gamma_r^{o\prime} = \frac{2\lambda_D}{\left(\rho P + 2P^2\bar{l}\bar{t}\emptyset/DS\right)H^2} \qquad 118$$

The objective function used in the derivation of the Equation (118) is as follows.

$$Z = \left(\frac{H}{2}\right)\gamma_w P + \frac{P^2 \bar{l}\bar{t}\emptyset\gamma_r^o H}{DS} + {\lambda_D}/{H} \qquad 119$$

The first term of Equation (119) considers the passengers' average cost of waiting for the bus while the second term considers the cost of additional riding time due to crowding discomfort. The third term represents the cost of operation of the bus route.

The notations used in the Equations (118) and (119) are shown below in Table 13.

Table 13 - Notations used in the Equations 118 and 119

Abbreviation	Description
D	Length of a bus route (km)
H	Uniform headway of a bus route (hr./bus)
$\bar{l}$	Average trip distance of passengers riding a bus route (km)
P	Average demand to board a bus route (pass./hr.)
S	Number of available seats on a bus (pass./bus)
γ_r^o	Basic value of γ_r without the effect of crowding nor pandemic health risks
$\emptyset$	Rate of change in average CPF with the average loading factor
ρ	Ratio of the value of wait time to value of ride time
$\bar{t}$	Average travel time of the passengers (hrs)
λ_D	Operations cost of a bus per dispatch ($/dispatch)

In Equation 1, the VoWT - γ_w – is referred to as a function of the VoRT - γ_r^o.

$$\gamma_w = \rho\gamma_r^o \qquad 120$$

According to the literature, the value of ρ is 2.5 (Wardman, 2004).

7.2.2. Access and Travel Time LOS – A&TT LOS

A&TT LOS considers the aspects of accessibility and travel time of a bus route. Similar to H&C LOS, an implied VoRT represents the quality of access and travel time attributes of a trip taken by transit is obtained through modeling the existing bus operation using the equation for the optimum stop distance for a bus route. The equation for the optimum stop distance is developed using an objective function that minimizes the trade-offs between the passenger costs of access distance along the bus route, the average travel time in the bus route and the cost of bus operation.

The implied VoRT representing the A&TT LOS is obtained using the Equation (121) below.

$$\gamma_r' = \frac{2n\lambda_B D\phi}{\hat{P}(\omega\bar{S}^2 - 2\bar{l}\phi)} \quad (121)$$

The objective function used in the derivation of the Equation (121) is as follows.

$$Z(x) = \frac{\gamma_k P(x)S(x)}{4} + \frac{\gamma_r C(x)\emptyset(x)}{S(x)} + \frac{n\lambda_B \phi(x) + \lambda_s}{S(x)} \quad (122)$$

The notations used in the Equations (121) and (122) are shown below in Table 14.

Table 14 - Notations used in the Equations 121 and 122

Abbreviation	Description
P(x)	Number of passengers boarding and alighting the bus service in a day per unit distance at the distance x (pass./km)
S(x)	Stop spacing centered at x (km)
C(x)	Number of passengers passing a point located from an x distance from the origin of the route in a day (pass./day)
$\emptyset(x)$	Amount of additional time taken due to stopping of a bus (hrs)
n	Number of bus dispatches per day
γ_r	Mean value of a unit of riding time of passengers ($/hr./pass.)
λ_B	Cost of operating a bus ($/hr)
λ_s	Cost of a bus stop per day ($/stop/day)
γ_k	Average travel time of the passengers (hrs)
D	Route distance (km)

$\bar{l}$	Mean trip length of passengers (km)
$\emptyset$	Average amount of additional time taken due to stopping of a bus (hrs)
$\hat{P}$	Total demand to board the bus route in a day (pass./day)
$\bar{S}$	Average stop spacing in the bus route (km)
ω	Value of the ratio γ_k/γ_r
γ_r'	Implied VoRT representing the A&TT LOS

In Equation (121), the VoWT - γ_k – has been referred to as a function of the VoRT - γ_r.

$$\gamma_k = \omega\gamma_r \qquad 123$$

According to literature, the value of ω is 2 (Wardman, 2004).

7.2.3. Reliability LOS – R LOS

Reliability LOS considers the aspect of unreliability of the service as it relates to the passengers' average wait time and travel time in the bus route. Similar to H&C LOS, an implied VoRT is representing the quality associated with the reliability of service that relates to the wait time and travel time attributes of a trip taken by transit. This is obtained through modeling the existing bus operation using the equation for the optimum scheduled travel time between consecutive time points for the bus route – hereafter referred to as the scheduled travel time. The equation for the optimum scheduled travel time is developed using an objective function that minimizes the trade-offs between the passenger costs of budgeted travel time, additional travel time due to unreliability of the service, additional wait time due to unreliability of the service, and the operations cost.

The objective function used in this process is as follows.

$$Z(S) = \frac{P\bar{l}nS\gamma_r}{D} + P(a\gamma_r + b\gamma_w)\int_s^\infty f(t)dt \int_s^\infty (t-S)f(t)dt + P\gamma_p \int_s^\infty f(t)dt + (nS + T')\lambda N \qquad 124$$

This objective function is used to derive an implied VoRT as follows, that represents the reliability LOS in comparison with the mean VoRT distribution of the passengers.

$$\gamma_r' = \frac{\gamma_p f(S) - \left(\frac{n\lambda N}{P}\right)}{\left(\frac{\bar{l}n}{D}\right) - Q(s)(a + b\theta)} \qquad 125$$

Where,

$$Q(S) = f(S)\left(\int_S^\infty tf(t)dt - S\big(1 - F(S)\big)\right) - \big(1 - F(S)\big)^2 \qquad 126$$

Table 15 - Notations used in the Equations 127, 128 and 129

Abbreviation	Description
P	Passenger demand to board the entire route in a day (passengers/day)
S	Scheduled travel time in between consecutive time points (hrs)
t	Link / inter-time-point travel time in the route (hrs)
$f(t)$	Probability density function of t
n	Number of inter-time-point links
γ_r	mean value of a unit of riding time of passengers ($/hr./pass.)
γ_w	mean value of a unit of waiting time of passengers ($/hr./pass.)
γ_p	Delay Penalty
λ	cost of operating a bus ($/hr)
D	Route distance (km)
$\bar{l}$	Mean trip length of passengers (km)
a	Delayed riding penalty factor
b	Delayed waiting penalty factor
T'	Layover time (hrs)
N	Number of dispatches per day in the route
θ	Ratio of VoWT to VoRT
γ_r'	Implied VoRT representing the A&TT LOS
$F(s)$	Cumulative probability of t being less than or equal to S

7.3. Methodology

Importance of the service attributes, their performance and passenger satisfaction are related (Börjesson & Rubensson, 2019). The performance of each attribute affects the overall quality of the service on different scales depending on how important each attribute is for the overall LOS. Therefore, figuring out a weighting factor for each of the attributes, or a combination of the attributes in the case of this study, is important. There are two main approaches considered in the literature to obtain the attribute level importance: (1) explicitly asking passengers for their level of importance for each attribute using a common scale; and (2) estimating the importance by modeling the explanatory power of each quality attribute on the trip utility by means of, for example, regression analysis, structural equations, Pearson correlation or path analysis (Börjesson & Rubensson, 2019). The proposed methodology in this study takes a novel approach of using the relationship between different VoT figures of the passengers to measure the relative importance of the sub-LOS measures for the overall LOS measure.

The literature provides a significant amount of evidence on the relationship between the VoRT of passengers to their VoWT, VoAT as a ratio (Wardman 2004). Therefore, the VoWT and VoRT can be denoted as a function of the VoRT. Unlike the VoRT of the passengers, the ratios of VoWT to VoRT and the VoAT to VoRT do not change from one population to the other and over time as values of time do (Wardman, 2013). Also, the vast majority of the studies on the values of time of passengers provide only the VoRT and therefore, the VoRT figures for a certain geographical region (for a province or a city) are already available from several sources and this makes it easy to carry out the calculations. Also, the value of a unit of time spent waiting being a times the value of a unit of time spent riding is a good indication that a person values their aspect of waiting at a times the aspect of riding. The same analogy exists within the relationship between the attributes of access time and riding time. This background supports the formulation of a weighted average approach in combining the LOS measures.

In formulating all the LOS measures introduced before, each LOS measure has been denoted by an implied VoRT. For example, the H&C LOS is denoted by an implied VoRT and then this value is compared with the VoRT distribution of the passengers to assess the LOS grade as introduced in the chapter 'Proposed Approach'. As described above, the H&C LOS represents both the aspects of riding and waiting.

Let us assume,

$$VoWT = a * VoRT \qquad 130$$

$$VoAT = b * VoRT \qquad 131$$

Since the H&C LOS measure represents the waiting time aspect – valued at a times the riding time aspect – apart from the riding time aspect, the H&C LOS measure can be weighted by a factor of $(1 + a)$. Similarly, the A&TT LOS measure can be weighted at a factor of $(1 + b)$ as this measure represents the aspects of riding and access. The reliability LOS measure, as it represents the aspects of riding time and waiting time, can be weighted at a factor of $(1 + a)$. We will name the implied VoRT representing the H&C LOS as $\gamma_r'(H\&C)$. The same way, the implied values of time representing the A&TT LOS and the Reliability LOS will be named as $\gamma_r'(A\&TT)$ and $\gamma_r'(Rel)$ respectively. Therefore, the overall LOS that is the combination of the LOS measures H&C, A&TT and Reliability, can be obtained by taking the weighted average as follows.

$$\gamma_r^*(Overall) = \frac{(1+a)\gamma_r'(H\&C) + (1+b)\gamma_r'(A\&TT) + (1+a)\gamma_r'(Rel)}{(1+a) + (1+b) + (1+a)} \qquad 132$$

Where $\gamma_r^*(Overall)$ represents the implied VoRT representing the overall LOS.

Table 16 provides an overview of the process of combining LOS measures.

Table 16 - Weights of sub-LOS measures to derive the overall LOS

LOS Measure	Aspects represented	Weight assigned
H&C LOS	Waiting Time - $\gamma_w = a\gamma_r$	(1+a)
	Riding Time - γ_r	
A&TT LOS	Access Time - $\gamma_a = b\gamma_r$	(1+b)
	Riding Time - γ_r	
Reliability	Waiting Time - $\gamma_w = a\gamma_r$	(1+a)
	Riding Time - γ_r	

The value obtained for $\gamma_r^*(Overall)$ in Equation (132) is similar to those of $\gamma_r'(H\&C)$, $\gamma_r'(A\&TT)$, and $\gamma_r'(Rel)$. Therefore, the same procedure used to obtain the LOS grades in each of these attributes – as described in the chapter 'Proposed Approach' – is used in obtaining a LOS grade for the overall LOS. For example, in the same way $\gamma_r'(H\&C)$ represents the implied VoRT representing the H&C LOS, the $\gamma_r^*(Overall)$ represents the overall/combined LOS. As the $\gamma_r'(H\&C)$ is compared with the mean VoRT distribution of the passengers to obtain the H&C LOS grade, $\gamma_r^*(Overall)$ is used in the same way to obtain the overall LOS by comparing it with the mean VoRT distribution of the passengers. Here the Overall LOS represents the combined LOS of the three separate LOS measures through a weighted average approach.

The next section presents the derivation of sub-LOS grades and the overall LOS grade for a real bus route in Calgary.

7.4. Example

The data for the bus route 7 – Marda Loop and City Centre as shown in **Figure 27**, obtained from Calgary Transit for a day in fall 2021, are shown in Table 17 as follows. It is important to note that the bus route has now changed as of June 2023.

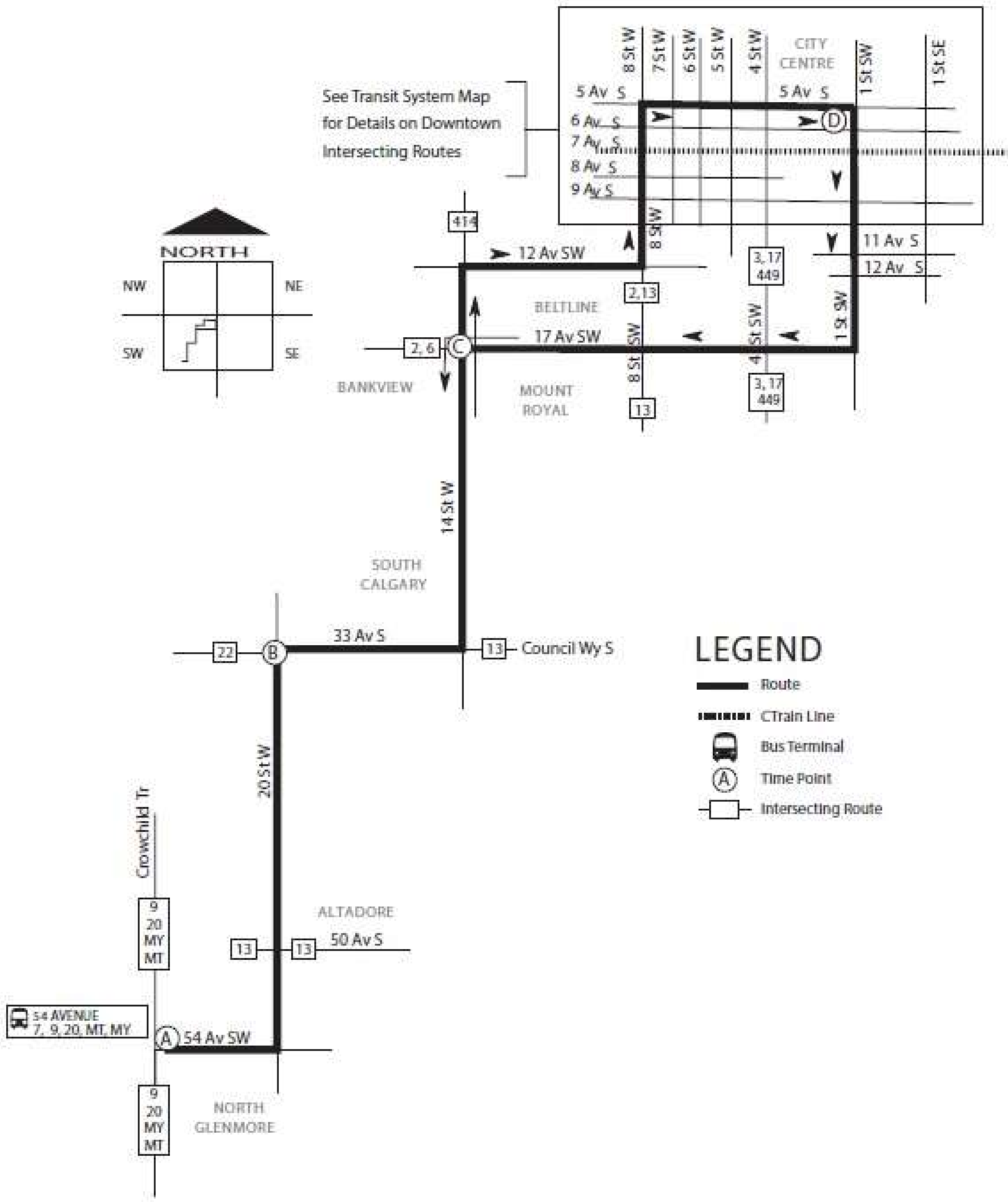


Figure 27 - Marda Loop Bus Route Map (Source - Calgary Transit Web Site)

Table 17 - Operational parameters of Route 7

Parameter	Description	Value	Units
P	Average passenger demand to board the entire bus route	35	Pass. /hr.
Q	Number of passengers passing through the maximum load point of the bus route during the peak period	48	Pass. /hr.
$\bar{t}$	Average trip duration of the passengers in the route	0.14	hrs
S	Number of available seats on a bus	35	Seats/bus
λ_D	Operating cost per dispatch	130	$/dispatch
$\bar{l}$	Average trip distance of the passengers in the route	3.61	km
D	Length of the bus route	17.1	km
λ_B	Cost of operating a bus	110	$/hr.
$\hat{P}$	Passenger demand to board the bus route in a day	1302	Passengers/day
$\bar{S}$	Average stop spacing in the bus route	0.23	km
n	Number of inter-time-point links	6	
N	Number of bus dispatches per day	50	

A stated preference survey carried out in the City of Calgary, AB, Canada studying the effect of the COVID-19 pandemic on the transit passenger perception in 2021 reports the mean and standard deviation of the VoRT of the passengers to be 24 $/hr/passenger and 10 $/hr/passenger respectively – 757 respondents and 4542 choice scenarios (Pollock, 2022). These values provide a VoRT distribution as shown in Figure 28.

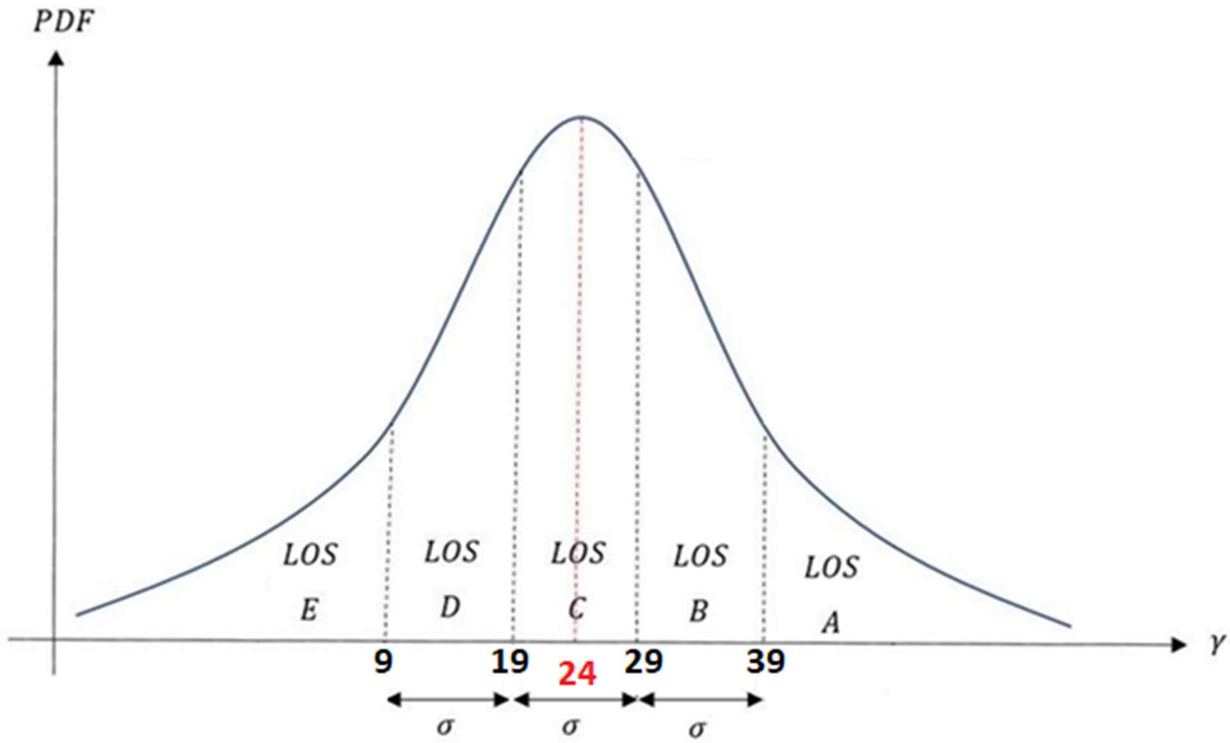


Figure 28 - VoRT Distribution with LOS grades

7.4.1. Headway & Crowding LOS of the Bus Route 7

The value of ρ – the ratio of the VoWT to VoRT – is assumed as 2.5 according to the literature.

The headway of the route changes from 20 to 30 minutes during the day. As other parameters assumed are average values, an average value of 25 minutes (0.416 hrs) of headway is assumed for the bus route to represent an average LOS. H&C LOS can be calculated for different time durations depending on the data availability. The value of ∅ is taken as 0.3 with reference to the calculation of ∅ carried out in the chapter 4 – 'Headway and Crowding LOS'. All the parameters in Equation (118) except for ∅ and ρ, are same as in Table 17.

Therefore, according to the Equation (118), the implied value of ride time representing the H&C LOS can be calculated to be 17.11 $/hr./pass. According to the VoRT distribution shown in the Figure 28, the H&C LOS falls in the LOS grade D.

7.4.2. Access and Travel Time LOS of the Bus Route 7

Value of ϕ – average additional time lost due to stopping of a bus was not available and is assumed to be 20 seconds (0.0056 hrs) for the route 7. According to Equation (121), the implied VoRT representing the A&TT LOS - $\gamma_r'(A\&TT)$ can be calculated to be 12.37 \$/hr/pass. that falls into the LOS grade of D according to the mean VoRT distribution of the passengers shown in Figure 28.

7.4.3. Reliability LOS

Depending on the time point log data obtained from Calgary Transit for route 7 for a Tuesday in September 2021, 6 inter-timepoint links were identified, i.e., n = 6. The average of the link travel time is calculated to be 0.2094 hrs – 12.57 minutes and the standard deviation of the link travel time is calculated to be 0.0269 hrs – 1.62 minutes. Mean and standard deviation of each link is shown in the Table 18. This mean and standard deviation helps us form the probability density function of the link travel time that is lognormally distributed.

Table 18 - Mean and standard deviation of link travel times

Link No.	Mean Travel Time (Hrs)	Standard Deviation (Hrs)
1	0.13603	0.02753
2	0.34999	0.03152
3	0.13416	0.02534
4	0.36770	0.13957
5	0.12301	0.03984
6	0.14578	0.05197

Accordingly, the shape parameter and scale parameter of the log normal function is calculated using the mean and standard deviation of the link travel times where the outcome values are used in the calculation. The delay penalty can be assumed to be the cost of passenger's travel time plus the bus fare (Wirasinghe, 1993). Accordingly, using the average trip time of passengers in the route – 0.14 hrs – with a value of ride time 24 \$/hr/pass. the delay penalty per passenger can be calculated as 7.01 \$/pass – Calgary had a flat rate bus fare of 3.65 CAD per passenger in October 2022. The values of a and b - the delayed waiting and riding time penalty factor – are assumed to be equal to 1.25 in line with the suggestion by Wirasinghe (1993). The value of θ – the ratio of VoWT to VoRT is assumed to be 2.5 inline with the assumptions in previous calculations.

Therefore, according to Equation (125), the implied VoRT representing the Reliability LOS - $\gamma_r'(Rel)$ is calculated to be 36.75 \$/hr/pass. representing a Reliability LOS grade of B with reference to the mean VoRT distribution of the passengers in the route as represented by the Figure 28.

7.4.4. The Overall LOS of route 7

The values obtained for $\gamma_r'(H\&C)$, $\gamma_r'(A\&TT)$, and $\gamma_r'(Rel)$ are substituted in the Equation (132). In the Equation (132), the value of a represents the ratio of VoWT to VoRT which is 2.5, and b represents the ratio of VoAT to VoRT which is 2 as used in the previous calculations under section 7.4.

Therefore, the implied VoRT - $\gamma_r^*(Overall)$ - for the route 7 can be calculated to be 22.56 \$/hr./pass. which falls into the LOS grade of C with reference to the mean VoRT distribution of the passengers in the route as represented by the Figure 28. Therefore, according to the methodology suggested in this study, the overall LOS of the bus route 7 in Calgary can be identified as LOS C.

7.5. Discussion

The accuracy of the approach presented herein for combining the sub-LOS measures depends on the accuracy of the weighting factors used. Weighting factors depend on the ratio of VoT components associated with sub-LOS measures to VoRT. Therefore, the accuracy of the mean and standard deviation of the VoRT is important. Other VoT components such as VoWT and VoAT used in the study are estimated using the VoRT and the ratios of VoWT and VoAT to VoRT that can be found in the literature. As discussed before, literature provides values for the ratios of VoAT and VoWT to VoRT. If such values have been derived in a similar geographical and economic environment to Calgary – where the LOS is being assessed – accuracy is higher. The ratios of VoTs do not depend on time as much as the VoTs depend on time due to inflation and economic development. Availability of ratios of VoTs for the context of the area where the LOS is being assessed – Calgary in this study – will increase accuracy.

8. Chapter 8 - Conclusion

8.1. Limitations and Recommended Future Work

Headway LOS has been derived for a bus route with frequency operations – headway is less than 10 minutes in normal terms. But the application of the overall LOS has been calculated for a bus route with scheduled operations. This issue can be eliminated by developing the model for implied VoRT of Headway LOS using the mean waiting time suggested by Ansari Esfeh et al., (2020). This study has not sought to address this issue since the scope was to demonstrate the applicability of the concept being developed.

Access and Travel Time LOS on the other hand consider only the access distance along the bus route whereas it can be more accurate if a better estimate for the access time can be achieved that includes the passengers living away from the bus route. A good approximation for the access time/distance of the passengers is proposed in the TCQSM that considers many other factors such as topography of the bus route and barriers such as pedestrian crossings in the approach path (TCRP Project A-15C, 2013). This LOS measure also considers only the travel time in transit using passengers value of travel time. Another factor that impacts the quality of travel time – as also identified in the chapter – 'Background' – is the directness of the route. Route directness compares the travel times using transit and a private vehicle. This can be accounted for within the proposed model by multiplying the average travel distance of the passengers by a factor that depends on the directness of the route. A comprehensive approach to calculating the directness of the route is presented in the TCQSM (TCRP Project A-15C, 2013).

The proposed method for reliability assessment depends heavily on data availability. Although technological advancement with transit vehicles has made this a possibility in many parts of the world, there is still a significant portion of transit agencies without access to such technologies. This method is developed for a route with only timepoints but can be extended to a real-world case with normal stops in between two time stops. As also discussed in the reliability chapter, making the average additional wait time half of the one used in the model is a good first step towards improving the model. Another limitation as it pertains to reliability LOS is that the methodology presented in the thesis depends on travel time variation rather than the headway variation. Depending the reliability assessment on headway variation is the suitable approach for a headway based (high frequency) service – which can be accounted through the coefficient of variation of the headway.

One of the key issues is that obtaining LOS values through optimization models can provide unrealistic values – and, for example, increasing the delay penalty factor from 5.75 $/pass. to 10 $/pass. will increase the implied VoRT of the reliability LOS to 63 $/hr/pass. Given that this represents a very high value of LOS for reliability, using this value in the Equation (132) will increase the overall LOS also to a very high value. This implies that increasing the LOS of one sub-measure will increase the overall LOS as well on the same scale. This might not represent the passenger expectations correctly as passengers tend not to perceive an importance in service performance improvements beyond a certain level. This is discussed comprehensively by Börjesson & Rubensson (2019). Improving the method suggested for obtaining the overall LOS to accommodate these concerns – such as introducing a stepwise weights function for the sub-LOS measures depending on their performance where the weight reduces with performance increases beyond a certain level – is a potential future extension of this study.

One of the key problems with the way the operator's concerns have been incorporated into the LOS service measures is that when operator's cost increases due to inefficiencies or non-optimal service provisions, it reflects in an equivalent improvement in the LOS measures. Because now the operator's contribution for making a transit trip happen is higher than that of the contribution from passengers. Addressing this anomaly requires further research into this topic. One of the potential approaches is to have standardized values for service provision for a given geographical, economic context – for example a North American context– and can be introduced through a transit service manual such a TCQSM. This way, operating costs due to inefficiencies or special cases will not affect the accuracy of intended outcomes.

8.2. Conclusion

This study, in general, addresses two key concerns in the transit LOS literature: (1) incorporating both operator and user concerns to LOS measures; (2) developing a single measure assessing the overall level of service (LOS). Incorporation of operator and user concerns depends on the concept that the contribution of both the operator resources and user time in the form of generalized costs is required to make a transit trip happen. A balanced contribution from both parties provides an optimal service environment. Analytically modelling the total cost of contribution from both parties for a transit operation makes it possible to obtain the optimal operational parameters for a bus route. This model is used to derive an VoT for the passengers – one of the variables in the analytical model for optimum operation – that corresponds with parameters of an already operating bus route. The obtained implied VoT is compared with the VoT distribution of the passengers to decide on the LOS.

Three sub-LOS measures were developed: Headway and Crowding LOS; Access and Travel Time LOS; and Reliability LOS. An analytical model for the optimum headway from literature was extended to incorporate the crowding attribute of the service as well. Given that this study was undertaken in a period where the public transit was heavily affected by the COVID-19 pandemic, this model was extended to incorporate the COVID-19 health risks into transit schedule planning as part of the crowding cost (Devasurendra et al., 2022) - another important contribution of this study. An analytical model for obtaining the optimum bus stop spacing is used to derive the Access and Travel Time LOS. An analytical model to assess the optimum scheduled travel time is developed in this study that incorporates average passenger wait time with ride time and delay penalty – an important contribution for the reliability literature. This model is then used to derive the reliability LOS of a bus route. It is important to note that this study does not investigate any models associated with LOS measures deeply given the scope – passenger and operator concerns integration and developing an overall LOS measure. Finally, a weighted average approach for combining sub-LOS measures is introduced that uses the ratios of value of wait time and value of access time to the value of ride time to derive the weights associated with sub-LOS measures.

References

Abkowitz, M., Slavin, H., Waksman, R., Englisher, L. S., & Wilson, N. H. M. (1979). Transit service reliability. *Proceedings of the Twentieth Annual Meeting of the Transportation Research Forum*, 283–290. https://rosap.ntl.bts.gov/view/dot/9212

Alexander, E. R. (1993). Density measures: A review and analysis. *Journal of Architectural and Planning Research*, *10*(3), 181–202.

Andaleeb, S., Haq, M., & Ahmed, R. (2007). Reforming Innercity Bus Transportation in a Developing Country: A Passenger-Driven Model. *Journal of Public Transportation*, *10*(1), 1–25. https://doi.org/10.5038/2375-0901.10.1.1

Ansari Esfeh, M., Saidi, S., Wirasinghe, S. C., & Kattan, L. (2022). Waiting time and headway modeling considering unreliability in transit service. *Transportation Research Part A: Policy and Practice*, *155*, 219–233. https://doi.org/10.1016/J.TRA.2021.11.015

Ansari Esfeh, M., Wirasinghe, S. C., Saidi, S., & Kattan, L. (2020). Waiting time and headway modelling for urban transit systems – a critical review and proposed approach. *Transport Reviews*, 1–23. https://doi.org/10.1080/01441647.2020.1806942

Ansari, M., Wirasinghe, S. C., Saidi, S., & Kattan, L. (2017). Waiting Time Modeling for Urban Transit Systems – A Critical Review and Proposed Approach. *23rd International Symposium on Transportation and Traffic Theory*, *1*(i).

Apple. (2021). *COVID-19 - Mobility Trends Reports*. https://covid19.apple.com/mobility

Aronstein, R. H. (1976). A proposed Standard for Transportation Systems. *Proceedi'ngs of the Fourth Intersociety Conference on Transportation*, 1–9.

Balcombe, R., Mackett, R., Paulley, N., Preston, J., Shires, J., Titheridge, H., Wardman, M., & White, P. (2004). The demand for public transport: a practical guide. In *Transportation Research Laboratory Report TRL593* . Transportation Research Laboratory. http://discovery.ucl.ac.uk/1349/

Banister, D. (2018). *Inequality in Transport*. Alexandrine Press.

Banks, J., Driscoll, W., & Stanford, R. (1982). Design Methodology for Airport Limousine Service. *Transportation Science*, *16*(2), 127–148. https://doi.org/10.1287/trsc.16.2.127

Banks, J. H. (1990). Optimal headways for multiroute transit systems. *Journal of Advanced Transportation*, *24*(2), 127–155. https://doi.org/10.1002/atr.5670240205

Banks, J. H. (1991). Effect of time-value distributions on transit headway optimizations. *Journal of Advanced Transportation*, *25*(2), 161–186. https://doi.org/10.1002/atr.5670250204

Basu, D., & Hunt, J. D. (2012). Valuing of attributes influencing the attractiveness of suburban train service in Mumbai city: A stated preference approach. *Transportation Research Part A: Policy and Practice*, *46*(9), 1465–1476. https://doi.org/10.1016/j.tra.2012.05.010

Batarce, M., Muñoz, J. C., & Ortúzar, J. de D. (2016). Valuing crowding in public transport: Implications for cost-benefit analysis. *Transportation Research Part A: Policy and Practice*, *91*, 358–378. https://doi.org/10.1016/j.tra.2016.06.025

Bavel, J. J. V., Baicker, K., Boggio, P. S., Capraro, V., Cichocka, A., Cikara, M., Crockett, M. J., Crum, A. J., Douglas, K. M., Druckman, J. N., Drury, J., Dube, O., Ellemers, N., Finkel, E. J., Fowler, J. H., Gelfand, M., Han, S., Haslam, S. A., Jetten, J., … Willer, R. (2020). Using social and behavioural science to support COVID-19 pandemic response. *Nature Human Behaviour*, *4*(5), 460–471. https://doi.org/10.1038/s41562-020-0884-z

Beck, M. J., & Hensher, D. A. (2020). Insights into the impact of COVID-19 on household travel and activities in Australia – The early days under restrictions. *Transport Policy*, *96*(May), 76–93. https://doi.org/10.1016/j.tranpol.2020.07.001

Ben-Akiva, M., Bolduc, D., & Bradley, M. (1993). Estimation of Travel Choice Models With Randomly Distributed Values of Time. *Transportation Research Record*, *1413*, 88–97.

Benn, H. P. (1995). *TCRP Synthesis 10: Bus Route Evaluation Standards*. http://onlinepubs.trb.org/onlinepubs/tcrp/tsyn10.pdf

Bentes, S. R., & Menezes, R. (2012). Entropy: A new measure of stock market volatility? *Journal of Physics: Conference Series*, *394*(1). https://doi.org/10.1088/1742-6596/394/1/012033

Biba, S., Curtin, K. M., & Manca, G. (2010). A new method for determining the population with walking access to transit. *International Journal of Geographical Information Science*, *24*(3), 347–364. https://doi.org/10.1080/13658810802646679

Birago, D., Opoku Mensah, S., & Sharma, S. (2017). Level of service delivery of public transport and mode choice in Accra, Ghana. *Transportation Research Part F: Traffic Psychology and Behaviour*, *46*, 284–300. https://doi.org/10.1016/j.trf.2016.09.033

Boisjoly, G., & El-Geneidy, A. M. (2017). The insider: A planners' perspective on accessibility. *Journal of Transport Geography*, *64*, 33–43. https://doi.org/10.1016/J.JTRANGEO.2017.08.006

Boring, E. G. (1961). The Beginning and Growth of Measurement in Psychology on JSTOR. *Isis*, *52*(2), 238–257. http://www.jstor.org/stable/228681

Börjesson, M., & Rubensson, I. (2019). Satisfaction with crowding and other attributes in public transport. *Transport Policy*, *79*, 213–222. https://doi.org/10.1016/j.tranpol.2019.05.010

Botzow, H. (1974). A Level of Service Concept for Evaluating Public Transport. *53rd Annual Meeting of the Highway Research Board*.

Bree, S., Fuller, D., & Diab, E. (2020). Access to transit? Validating local transit accessibility measures using transit ridership. *Transportation Research Part A: Policy and Practice*, *141*(October), 430–442. https://doi.org/10.1016/j.tra.2020.09.019

Bunker, J. M. (2015). Assessment of Transit Quality of Service with Occupancy Load Factor and Passenger Travel Time Measures. *Transportation Research Record: Journal of the Transportation Research Board*, *2535*, 45–54. https://doi.org/10.3141/2535-05

Camus, R., Longo, G., & Macorini, C. (2005). Estimation of transit reliability level-of-service based on automatic vehicle location data. *Transportation Research Record*, *1927*, 277–286. https://doi.org/10.3141/1927-31

Canadian Urban Transit Association. (1993). *Canadian transit handbook.* Canadian Urban Transit Association. https://books.google.ca/books/about/Canadian_Transit_Handbook.html?id=pubZAAAAMAAJ&redir_esc=y

Canadian Urban Transit Association. (2001). *A Review of Canadian Transit Service Standards*.

Çelebi, D., & İmre, Ş. (2020). Measuring crowding-related comfort in public transport. *Https://Doi.Org/10.1080/03081060.2020.1805546*, *43*(7), 735–750. https://doi.org/10.1080/03081060.2020.1805546

Chang, J. S. (2010). Assessing travel time reliability in transport appraisal. *Journal of Transport Geography*, *18*(3), 419–425. https://doi.org/10.1016/J.JTRANGEO.2009.06.012

Chen, B. Y., Shi, C., Zhang, J., Lam, W. H. K., Li, Q., Xiang, S., Kong, H., & Fitts, E. P. (2017). Most reliable path-finding algorithm for maximizing on-time arrival probability. *Transportmetrica B: Transport Dynamics*, *5*(3), 248–264. https://doi.org/10.1080/21680566.2016.1169953

Chen, W., & Chen, Z. (2009). Service Reliability Analysis of High Frequency Transit Using Stochastic Simulation. *Journal of Transportation Systems Engineering and Information Technology*, *9*(5), 130–134. https://doi.org/10.1016/S1570-6672(08)60083-5

Cho, S.-H., & Park, H.-C. (2021). Exploring the Behaviour Change of Crowding Impedance on Public Transit due to COVID-19 Pandemic: Before and After Comparison. *Transportation Letters*, 1–8. https://doi.org/10.1080/19427867.2021.1897937

Clarens, G. C., & Hurdle, V. F. (1975). An Operating Strategy for a Commuter Bus System. *Transportation Science*, *9*(1), 1–20. https://doi.org/10.1287/trsc.9.1.1

Concas, S., & Kolpakov, A. (2009). *Synthesis of Research on Value of Time and Value of Reliability*.

Correia, A. R., & Wirasinghe, S. C. (2013). Modeling airport landside performance. In *Modelling and Managing Airport Performance* (pp. 1–26). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118535844.CH1

Cox, T., Houdmont, J., & Griffiths, A. (2006). Rail passenger crowding, stress, health and safety in Britain. *Transportation Research Part A: Policy and Practice*, *40*(3), 244–258. https://doi.org/10.1016/J.TRA.2005.07.001

Dai, T., & Taylor, B. D. (2020, October). When is Public Transit Too Crowded, and How Has This Changed During the Pandemic? Publication Date. *UC Office of the President Policy Briefs*. https://doi.org/10.17226/24766

Dai, Z., Ma, X., & Chen, X. (2019). Bus travel time modelling using GPS probe and smart card data: A probabilistic approach considering link travel time and station dwell time. *Journal of Intelligent Transportation Systems*, *23*(2), 175–190. https://doi.org/10.1080/15472450.2018.1470932

Dandapat, S., Bhattacharyya, K., Annam, S. K., Saysardar, K., & Maitra, B. (2020). Impact of COVID-19 Outbreak on Travel Behaviour: Evidences from early stages of the Pandemic in India. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3692923

Das, S., & Pandit, D. (2013). Importance of user perception in evaluating level of service for bus transit for a developing country like India: a review. *Transport Reviews*, *33*(4), 402–420. https://doi.org/10.1080/01441647.2013.789571

Das, S., & Pandit, D. (2015). Determination of level-of-service scale values for quantitative bus transit service attributes based on user perception. *Transportmetrica A: Transport Science*, *11*(1), 1–21. https://doi.org/10.1080/23249935.2014.910563

de Palma, A., Kilani, M., & Proost, S. (2015). Discomfort in mass transit and its implication for scheduling and pricing. *Transportation Research Part B: Methodological*, *71*, 1–18. https://doi.org/10.1016/j.trb.2014.10.001

dell'Olio, L., Ibeas, A., & Cecín, P. (2010). Modelling user perception of bus transit quality. *Transport Policy*, *17*(6), 388–397. https://doi.org/10.1016/j.tranpol.2010.04.006

dell'Olio, L., Ibeas, A., & Cecin, P. (2011). The quality of service desired by public transport users. *Transport Policy*, *18*(1), 217–227. https://doi.org/10.1016/J.TRANPOL.2010.08.005

Denuit, M., & Genest, C. (2001). An Extension of Osuna's Model for Stress Caused by Waiting. *Journal of Mathematical Psychology*, *45*(1), 115–130. https://doi.org/10.1006/jmps.1999.1297

Devasurendra, K. W., Saidi, S., Wirasinghe, S. C., & Kattan, L. (2022). Integrating COVID-19 health risks into crowding costs for transit schedule planning. *Transportation Research Interdisciplinary Perspectives*, *13*. https://doi.org/10.1016/J.TRIP.2021.100522

Diab, E. I., Badami, M. G., & El-Geneidy, A. M. (2015). Bus Transit Service Reliability and Improvement Strategies: Integrating the Perspectives of Passengers and Transit Agencies in

North America. *Transport Reviews*, *35*(3), 292–328. https://doi.org/10.1080/01441647.2015.1005034

Dong, H., Ma, S., Jia, N., & Tian, J. (2021). Understanding public transport satisfaction in post COVID-19 pandemic. *Transport Policy*, *101*, 81–88. https://doi.org/10.1016/j.tranpol.2020.12.004

Drabicki, A., Kucharski, R., & Cats, · Oded. (2022). Mitigating bus bunching with real-time crowding information. *Transportation 2022*, 1–28. https://doi.org/10.1007/S11116-022-10270-3

Eboli, L., Forciniti, C., & Mazzulla, G. (2014). ScienceDirect Service coverage factors affecting bus transit system availability. *Procedia - Social and Behavioral Sciences*, *111*, 984–993. https://doi.org/10.1016/j.sbspro.2014.01.133

Eboli, L., & Mazzulla, G. (2007). Service Quality Attributes Affecting Customer Satisfaction for Bus Transit. *Journal of Public Transportation*, *10*(3), 21–34. https://doi.org/10.5038/2375-0901.10.3.2

El-Geneidy, A., Grimsrud, M., Wasfi, R., Tétreault, P., & Surprenant-Legault, J. (2014). New evidence on walking distances to transit stops: Identifying redundancies and gaps using variable service areas. *Transportation*, *41*(1), 193–210. https://doi.org/10.1007/S11116-013-9508-Z/FIGURES/4

Elias, W., & Zatmeh-Kanj, S. (2021). Extent to which COVID-19 will affect future use of the train in Israel. *Transport Policy*, *110*(May), 215–224. https://doi.org/10.1016/j.tranpol.2021.06.008

Ettema, D., & Verschuren, L. (2007). Multitasking and value of travel time savings. *Transportation Research Record*, *2010*, 19–25. https://doi.org/10.3141/2010-03

FitzRoy, F., & Smith, I. (1998). Public transport demand in Freiburg: why did patronage double in a decade? *Transport Policy*, *5*(3), 163–173. https://doi.org/10.1016/S0967-070X(98)00024-9

Flanagan, J. C. (1982). Measurement of quality of life: current state of the art. *Archives of Physical Medicine and Rehabilitation*, *63*(2), 56–59. http://www.ncbi.nlm.nih.gov/pubmed/6460487

Fried, M. L., & Defazio, V. J. (2016). Territoriality and Boundary Conflicts in the Subway. *Psychiatry*, *37*(1), 47–59. https://doi.org/10.1080/00332747.1974.11023786

Fries, R. N., Dunning, A. E., & Chowdhury, M. A. (2011). University Traveler Value of Potential Real-Time Transit Information. In *Journal of Public Transportation* (Vol. 14, Issue 2).

Fu, L. (2007). A New Performance Index for Evaluating Transit Quality of Service. *Journal of Public Transportation*, *10*(3), 47–69. https://scholarcommons.usf.edu/cgi/viewcontent.cgi?article=1246&context=jpt

Furth, P. G., Hemily, B., Muller, T. H. J., & Strathman, J. G. (2006). *TCRP Report 113: Using Archived AVL-APC Data to Improve Transit Performance and Management*. Transportation Research Board. https://doi.org/10.17226/13907

Furth, P. G., & Muller, T. H. J. (2006). Service reliability and hidden waiting time insights from automatic vehicle location data. *Transportation Research Record*, *1955*, 79–87. https://doi.org/10.3141/1955-10

Furth, P. G., & Wilson, N. H. M. (1981). Setting Frequencies on Bus Routes: Theory and Practice. *Transportation Research Record*, *818*, 1–7. http://www1.coe.neu.edu/~pfurth/Furth%20papers/1981%20setting-frequencies.pdf

Garrido, R. A., & Ortuzar, J. D. D. (1994). Deriving Public Transport Level of Service Weights from a Multiple Comparison of Latent and Observable Variables. *The Journal of the Operational Research Society*, *45*(10), 1099. https://doi.org/10.2307/2584473

Gkiotsalitis, K., & Cats, O. (2020). Public transport planning adaption under the COVID-19 pandemic crisis: literature review of research needs and directions. *Transport Reviews*, 1–19. https://doi.org/10.1080/01441647.2020.1857886

Golob, T. F., Canty, E. T., Gustafson, R. L., & Vitt, J. E. (1972). An analysis of consumer preferences for a public transportation system. *Transportation Research*, *6*(1), 81–102. https://doi.org/10.1016/0041-1647(72)90113-X

Güner, S. (2018). Measuring the quality of public transportation systems and ranking the bus transit routes using multi-criteria decision making techniques. *Case Studies on Transport Policy*, *6*(2), 214–224. https://doi.org/10.1016/j.cstp.2018.05.005

Habib, K. M. N., Kattan, L., & Islam, T. (2011). Model of personal attitudes towards transit service quality. *Journal of Advanced Transportation*, *45*(4), 271–285. https://doi.org/10.1002/ATR.106

Hatcher, G., Hicks, D., Lowrance, C., Mercer, M., Brooks, M., Thompson, K., Lowman, A., Jacobi, A., Ostroff, R., Serulle, N. U., & Vargo, A. (2017). *Intelligent Transportation Systems Benefits , Costs , Deployment , and Lessons Learned 2008 Update*.

Hawkins, J., Wirasinghe, S. C., & Hunt, J. D. (2020). Bus headway optimization with consumer surplus as a measure of societal benefit. *International Journal of Urban Sciences*, *0*(0), 1–18. https://doi.org/10.1080/12265934.2020.1730224

Haywood, L., & Koning, M. (2015). The distribution of crowding costs in public transport: New evidence from Paris. *Transportation Research Part A: Policy and Practice*, *77*, 182–201. https://doi.org/10.1016/j.tra.2015.04.005

Heathington, K. W., & Brogan, J. D. (1975). *Fixed-Route, FixedSchedule Bus Systems*.

Hensher, D. A. (1998). The imbalance between car and public transport use in urban Australia: why does it exist? *Transport Policy*, *5*(4), 193–204. https://doi.org/10.1016/S0967-070X(98)00022-5

Hensher, D. A. (2014). The Relationship Between Bus Contract Costs, User Perceived Service Quality and Performance Assessment. *International Journal of Sustainable Transportation*, *8*(1), 5–27. https://doi.org/10.1080/15568318.2012.758454

Hensher, D. A., Mulley, C., & Yahya, N. (2010). Passenger experience with quality-enhanced bus service: the tyne and wear 'superoute' services. *Transportation*, *37*(2), 239–256. https://doi.org/10.1007/s11116-009-9240-x

Hensher, D. A., Rose, J. M., & Collins, A. T. (2011). Identifying commuter preferences for existing modes and a proposed Metro in Sydney, Australia with special reference to crowding. *Public Transport*, *3*(2), 109–147. https://doi.org/10.1007/s12469-010-0035-4

Hensher, D. A., Stopher, P., & Bullock, P. (2003). Service quality—developing a service quality index in the provision of commercial bus contracts. *Transportation Research Part A: Policy and Practice*, *37*(6), 499–517. https://doi.org/10.1016/S0965-8564(02)00075-7

Hess, D. B., Brown, J., & Shoup, D. (2004). Waiting for the Bus. *Journal of Public Transportation*, *7*(4), 67–84.

Higgins, D. L., & Ryan, W. M. (1983). *Levels of Service Concepts in Transport Planning: Review and Future Prospects*. http://atrf.info/papers/1983/1983_Higgins_Ryan.pdf

Highway Research Board. (1950). Highway Capacity Manual. In *Transportation Research Board, National Research Council*. National Academy of Sciences.

Hillman, R., & Pool, R. (1997). GIS-BASED INNOVATIONS FOR MODELLING PUBLIC TRANSPORT ACCESSIBILITY. *Traffic Engineering & Control*, *38*(10), 554–559.

Hodge, D. C., & Orell, J. D. (1995). *Measuring Los and Performance in Public Transportation.Pdf*.

Honey-Rosés, J., Anguelovski, I., Chireh, V. K., Daher, C., Konijnendijk van den Bosch, C., Litt, J. S., Mawani, V., McCall, M. K., Orellana, A., Oscilowicz, E., Sánchez, U., Senbel, M., Tan, X., Villagomez, E., Zapata, O., & Nieuwenhuijsen, M. J. (2020). The impact of COVID-19 on public space: an early review of the emerging questions – design, perceptions and inequities. *Cities & Health*, *00*, 1–17. https://doi.org/10.1080/23748834.2020.1780074

Hossain, M. S. (2019). *Out of Vehicle Time Values in Transit - personal communication*.

Hossain, M. S., Hunt, J. D., & Wirasinghe, S. C. (2015). Nature of influence of out-of-vehicle time-related attributes on transit attractiveness: a random parameters logit model analysis. *Journal of Advanced Transportation*, *49*, 648–662. https://doi.org/10.1002/atr.1297

Huo, Y., Zhao, J., Li, W., & Hu, X. (2014). Measuring Bus Service Reliability: An Example of Bus Rapid Transit in Changzhou. *Journal of Public Transportation*, *17*(2), 113–133. https://doi.org/http://doi.org/10.5038/2375-0901.17.2.6

Ibarra-Rojas, O. J., Giesen, R., & Rios-Solis, Y. A. (2014). An integrated approach for timetabling and vehicle scheduling problems to analyze the trade-off between level of service and operating costs of transit networks. *Transportation Research Part B: Methodological*, *70*, 35–46. https://doi.org/10.1016/J.TRB.2014.08.010

Jansson, J. O. (1980). A simple bus line model for optimisation of service frequency and bus size. *Journal of Transport Economics and Policy*, *14*(1), 53–80. https://doi.org/10.1099/jmm.0.46632-0

Jara-Diaz, S. R., & Gschwender, A. (2003). Towards a general microeconomic model for the operation of public transport. *Transport Reviews*, *23*(4), 453–469. https://doi.org/10.1080/0144164032000048922

Jaynes, E. T. (1957). Information Theory and Statistical Mechanics. *Physical Review*, *106*(4), 620–630.

Jordaan, I. (2005). Entropy. In *Decisions under Uncertainty* (pp. 272–316). Cambridge University Press. https://doi.org/10.1017/CBO9780511804861.007

Kamga, C., & Eickemeyer, P. (2021). Slowing the spread of COVID-19: Review of "Social distancing" interventions deployed by public transit in the United States and Canada. *Transport Policy*, *106*(March), 25–36. https://doi.org/10.1016/j.tranpol.2021.03.014

Kapur, J. N. (1993). *Maximum-Entropy Models in Science and Engineering* (Revised Edition). Wiley Eastern Limited.

Kendall, M. G. (1994). *Kendall's advanced theory of statistics.* (6th ed., Vol. 1). http://hdl.handle.net/2027/mdp.39015058965008

Kittelson & Associates, Inc., Urbitran, Inc., LKC Consulting Services, I., Morpace International, Inc., Queensland University of Technology, & Yuko Nakanishi. (2003). *TCRP Report 88: A Guidebook for Developing a Transit Performance-Measurement System*. www.TRB.org

Klumpenhouwer, W., & Wirasinghe, S. C. (2016). Cost-of-crowding model for light rail train and platform length. *Public Transport*, *8*(1), 85–101. https://doi.org/10.1007/s12469-015-0118-3

Klumpenhouwer, W., & Wirasinghe, S. C. (2018). Optimal time point configuration of a bus route - A Markovian approach. *Transportation Research Part B: Methodological*, *117*, 209–227. https://doi.org/10.1016/J.TRB.2018.08.021

Kocas, C. (2015). An extension of Osuna's model to observable queues. *Journal of Mathematical Psychology*, *66*, 53–58. https://doi.org/10.1016/j.jmp.2015.04.001

Kopsidas, A., Milioti, C., Kepaptsoglou, K., & Vlachogianni, E. I. (2021). How did the COVID-19 pandemic impact traveler behavior toward public transport? The case of Athens, Greece. *Transportation Letters*, 1–9. https://doi.org/10.1080/19427867.2021.1901029

Labonté-Lemoyne, É., Chen, S. L., Coursaris, C. K., Sénécal, S., & Léger, P. M. (2020). The unintended consequences of covid-19 mitigation measures on mass transit and car use. *Sustainability (Switzerland)*, *12*(23), 1–13. https://doi.org/10.3390/su12239892

Lam, W. H. K., Cheung, C. Y., & Poon, Y. F. (1999). Study of passenger discomfort measures at the Hong Kong Mass Transit Railway system. *Journal of Advanced Transportation*, *33*(3), 389–399. https://doi.org/10.1002/atr.5670330308

Langford, M., Fry, R., & Higgs, G. (2012). Measuring transit system accessibility using a modified two-step floating catchment technique. *International Journal of Geographical Information Science*, *26*(2), 193–214. https://doi.org/10.1080/13658816.2011.574140

Leake, G. R., & Huzayyin, A. (1979). ACCESSIBILITY MEASURES AND THEIR SUITABILITY FOR USE IN TRIP GENERATION MODELS. *Traffic Engineering & Control*, *20*(12), 566–572.

Leurent, F. (1994). Cost versus time equilibrium over a network. *Transportation Research Record*, *71*(1443), 84–91. https://doi.org/10.1016/0377-2217(93)90049-S

Li, Z., & Hensher, D. A. (2011). Crowding and public transport: A review of willingness to pay evidence and its relevance in project appraisal. *Transport Policy*, *18*(6), 880–887. https://doi.org/10.1016/j.tranpol.2011.06.003

Liu, G., & Wirasinghe, S. C. (1991). Extension of Newell's dispatching policy to the case of sensitive demand. *Journal of Advanced Transportation*, *25*(3), 309–328. https://doi.org/10.1002/atr.5670250305

Liu, G., & Wirasinghe, S. C. (2001). A simulation model of reliable schedule design for a fixed transit route. *Journal of Advanced Transportation*, *35*(2), 145–174. https://doi.org/10.1002/ATR.5670350206

Liu, Z., Yan, Y., Qu, X., & Zhang, Y. (2013). Bus stop-skipping scheme with random travel time. *Transportation Research Part C: Emerging Technologies*, *35*, 46–56. https://doi.org/10.1016/J.TRC.2013.06.004

Lu, H., Burge, P., Heywood, C., Sheldon, R., Lee, P., Barber, K., & Phillips, A. (2018). The impact of real-time information on passengers' value of bus waiting time. *Transportation Research Procedia*, *31*(2016), 18–34. https://doi.org/10.1016/j.trpro.2018.09.043

Lu, H., Fowkes, T., & Wardman, M. (2008). Amending the Incentive for Strategic Bias in Stated Preference Studies. *Transportation Research Record: Journal of the Transportation Research Board*, *2049*(1), 128–135. https://doi.org/10.3141/2049-15

Luo, W., & Wang, F. (2003). Measures of Spatial Accessibility to Healthcare in a GIS Environment: Synthesis and a Case Study in Chicago Region. *Environment and Planning. B, Planning & Design*, *30*(6), 865–884. https://doi.org/10.1068/b29120

Malekzadeh, A., & Chung, E. (2020). A review of transit accessibility models: Challenges in developing transit accessibility models. *International Journal of Sustainable Transportation*, *14*(10), 733–748. https://doi.org/10.1080/15568318.2019.1625087

Mallett, W. J. (2020). Emergency Funding for Public Transportation Agencies Due to COVID-19. In *CRS INSIGHT*.

Mamun, S. A., Lownes, N. E., Osleeb, J. P., & Bertolaccini, K. (2013). A method to define public transit opportunity space. *Journal of Transport Geography*, *28*, 144–154. https://doi.org/10.1016/J.JTRANGEO.2012.12.007

Manville, M., & Cummins, B. (2015). Why do voters support public transportation? Public choices and private behavior. *Transportation*, *42*(2), 303–332. https://doi.org/10.1007/s11116-014-9545-2

MathWorks. (2020). *MATLAB - MathWorks*. https://www.mathworks.com/products/matlab.html

METROLINX. (2020, March 18). *Metrolinx - Health care workers depend on GO Transit during COVID-19*. https://www.metrolinx.com/en/news/health-care-workers-depend-on-go-transit-during-covid-19

Mohammadi, M. (2023). *Impacts of Weather on Urban Bus Performance in the City of Calgary, Alberta* [Masters Thesis]. University of Calgary.

Mohring, H. (1972). Optimization and scale economies in urban bus transportation. *The American Economic Review*, *62*(4), 591–604.

Mohring, H., Schroeter, J., & Wiboonchutikula, P. (1987). The Values of Waiting Time , Travel Time , and a Seat on a Bus. *The RAND Journal of Economics*, *18*(1), 40–56.

Moriyama, M., Hugentobler, W. J., & Iwasaki, A. (2020). Seasonality of Respiratory Viral Infections. *Annual Review of Virology*, *7*(1), 83–101.

Morpace International Inc, & Cambridge Systematics Inc. (1999). TCRP Report 47: A Handbook for Measuring Customer Satisfaction and Service Quality. In *TRANSIT COOPERATIVE RESEARCH PROGRAM*. https://www.worldtransitresearch.info/research/2973

Müller, C., & Gosling, G. D. (1991). A framework for evaluating level of service for airport terminals. *Transportation Planning and Technology*, *16*(1), 45–61. https://doi.org/10.1080/03081069108717470

NAIT. (2021). *Alberta COVID-19 Dashboard*. Centre for Applied Disaster and Emergency Management. https://experience.arcgis.com/experience/d76b608cd17f432093664f322c4c2f8b

National Committee on Urban Transportation. (1958). Recommended Standards, Warrants, and Objectives for Transit Services and Facilities. In *Procedure Manual 8A*.

National Research Council (U.S.). (2010). *Highway Capacity Manual 6th Edition: A Guide for Multimodal Mobility Analysis*.

Ndoh, N. N., & Ashford, N. J. (1993). EVALUATION OF AIRPORT ACCESS LEVEL OF SERVICE. *Transportation Research Record*, *1423*, 34–39.

Newell, G. F. (1971). Dispatching Policies for a Transportation Route. *Transportation Science*, *5*(1), 91–105. https://doi.org/10.1287/trsc.5.1.91

Newell, G. F. (1977). Unstable Brownian Motion of a Bus Trip. In *Statistical Mechanics and Statistical Methods in Theory and Application* (pp. 645–667). Springer US. https://doi.org/10.1007/978-1-4613-4166-6_28

Newell, G. F. (1979). Some Issues Relating To the Optimal Design of Bus Routes. *Transp Sci*, *13*(1), 20–35. https://doi.org/10.1287/trsc.13.1.20

Nishiura, H., Linton, N. M., & Akhmetzhanov, A. R. (2020). Initial Cluster of Novel Coronavirus (2019-nCoV) Infections in Wuhan, China Is Consistent with Substantial Human-to-Human Transmission. *Journal of Clinical Medicine*, *9*(2), 488. https://doi.org/10.3390/jcm9020488

Noor, H. M., Nasrudin, N., & Foo, J. (2014). Determinants of Customer Satisfaction of Service Quality: City Bus Service in Kota Kinabalu, Malaysia. *Procedia - Social and Behavioral Sciences*, *153*, 595–605. https://doi.org/10.1016/J.SBSPRO.2014.10.092

NRG Research Group. (2017). *Calgary Transit Customer Satisfaction*. https://www.calgarytransit.com/content/transit/en/home/about-calgary-transit/facts-and-figures/reports-and-surveys.html

Ojeda-Cabral, M., Batley, R., & Hess, S. (2016). The value of travel time: random utility versus random valuation. *Transportmetrica A: Transport Science*, *12*(3), 230–248. https://doi.org/10.1080/23249935.2015.1125398

Ormos, M., & Zibriczky, D. (2014). Entropy-Based Financial Asset Pricing. *PLoS ONE*, *9*(12). https://doi.org/10.1371/JOURNAL.PONE.0115742

Orro, A., Novales, M., Monteagudo, Á., Pérez-López, J. B., & Bugarín, M. R. (2020). Impact on city bus transit services of the COVID-19 lockdown and return to the new normal: The case of A Coruña (Spain). *Sustainability (Switzerland)*, *12*(17). https://doi.org/10.3390/su12177206

O'Sullivan, S., & Morrall, J. (1996). Walking Distances to and from Light-Rail Transit Stations: *Transportation Research Record*, *1538*(1), 19–26. https://doi.org/10.1177/0361198196153800103

Osuna, E. E. (1985). The psychological cost of waiting. *Journal of Mathematical Psychology*, *29*(1), 82–105. https://doi.org/10.1016/0022-2496(85)90020-3

Pathak, P., Agrawal, K., Suman, H. K., & Bolia, N. B. (2020). Frequency optimization-based approach for reducing crowding discomfort in Delhi bus system. *Procedia Computer Science*, *170*, 265–272. https://doi.org/10.1016/j.procs.2020.03.039

Pollock, S. (2022). Parameters of the Value of ride time distribution of transit passengers in Calgary. In *Personal Communication*.

Pollock, S., Devasurendra, K. W., Weiss, A. M., De Barros, A., Wirasinghe, S. C., & Kattan, L. (2021). *Behaviour of transit passengers during COVID-19: A case study in the City of Calgary, Canada - Working Paper*.

Polydoropoulou, A., & Ben-Akiva, M. (2001). Combined Revealed and Stated Preference Nested Logit Access and Mode Choice Model for Multiple Mass Transit Technologies. *Transportation Research Record: Journal of the Transportation Research Board*, *1771*(1), 38–45. https://doi.org/10.3141/1771-05

Polzin, S. E., Pendyala, R. M., & Navari, S. (2002). Development of time-of-day-based transit accessibility analysis tool. *Transportation Research Record*, *1799*, 35–41. https://doi.org/10.3141/1799-05

Pratt, R. H., & Evans, J. E. (Jay). (2004). *Traveler Response to Transportation System Changes Handbook, Third Edition: Chapter 10, Bus Routing and Coverage* (Third Edit). Transportation Research Board. https://doi.org/10.17226/23330

Prioni, P., & Hensher, D. (2001). Measuring Service Quality in Scheduled Bus Services. *Journal of Public Transportation*, *3*(2), 51–74. https://doi.org/10.5038/2375-0901.3.2.4

Przybylowski, A., Stelmak, S., & Suchanek, M. (2021). Mobility Behaviour in View of the Impact of the COVID-19 Pandemic—Public Transport Users in Gdansk Case Study. *Sustainability*, *13*(1), 364. https://doi.org/10.3390/su13010364

Pushkarev, B. S., & Zupan, J. M. (1977). *Public transportation and land use policy*. Regional Plan Association. https://books.google.ca/books/about/Public_transportation_and_land_use_polic.html?id=PRtPAAAAMAAJ&redir_esc=y

Qin, F. (2014). *Investigating the In-Vehicle Crowding Cost Functions for Public Transit Modes*. https://doi.org/10.1155/2014/502708

Quality of Service Concepts. (2013). In *Transit Capacity and Quality of Service Manual* (3rd ed., pp. 4-i-4–48). National Academy of Sciences.

Quarmby D. A. (1967). Choice of Travel Mode for the Journey to Work: Some Findings on JSTOR. *Journal of Transport Economics and Policy*, *1*(3), 273–314.

Rahman, M., Kattan, L., & Wirasinghe, · S C. (2018). Trade-offs between headway, fare, and real-time bus information under different weather conditions. *Public Transport*, *10*, 217–240. https://doi.org/10.1007/s12469-018-0176-4

Rahman, M., Wirasinghe, S. C., & Kattan, L. (2018). Analysis of bus travel time distributions for varying horizons and real-time applications. *Transportation Research Part C: Emerging Technologies*, *86*, 453–466. https://doi.org/10.1016/J.TRC.2017.11.023

Richardson, A. J. (2002). Simulation Study of Estimation of Individual Specific Values of Time by Using Adaptive Stated-Preference Survey. *Transportation Research Record: Journal of the Transportation Research Board*, *1804*(1), 117–125. https://doi.org/10.3141/1804-16

Rizzi, L. I., Limonado, J. P., & Steimetz, S. S. C. (2012). The impact of traffic images on travel time valuation in stated-preference choice experiments. *Transportmetrica*, *8*(6), 427–442. https://doi.org/10.1080/18128602.2010.551524

Roebuck, C. S. (1979). An Interim Warrant for Bus Priority Lanes. In *Technical Manual K 39*.

Ryus, P., Ausman, J., Teaf, D., Cooper, M., & Knoblauch, M. (2000). Development of Florida's Transit Level-of-Service Indicator. *Transportation Research Record: Journal of the Transportation Research Board*, *1731*, 123–129. https://doi.org/10.3141/1731-15

Seneviratne, P. N. (1990). Analysis of on-time performance of bus services using simulation. *Journal of Transportation Engineering*, *116*(4), 517–531. https://doi.org/10.1061/(ASCE)0733-947X(1990)116:4(517)

Shamshiripour, A., Rahimi, E., Shabanpour, R., & Mohammadian, A. (Kouros). (2020). How is COVID-19 reshaping activity-travel behavior? Evidence from a comprehensive survey in Chicago. *Transportation Research Interdisciplinary Perspectives*, *7*. https://doi.org/10.1016/j.trip.2020.100216

Shelat, S., Cats, O., & Cranenburgh, S. V. (2020). Avoiding the crowd: How do passengers trade-off time and crowding in the age of COVID-19 (working paper). In *Preprint*. https://www.researchgate.net/publication/342392070_Avoiding_the_crowd_How_do_passengers_trade-off_time_and_crowding_in_the_age_of_COVID-19_working_paper

Shen, X., Feng, S., Li, Z., & Hu, B. (2016). Analysis of bus passenger comfort perception based on passenger load factor and in-vehicle time. *SpringerPlus*, *5*, 62. https://doi.org/10.1186/s40064-016-1694-7

Sheth, C., Triantis, K., & Teodorović, D. (2007). Performance evaluation of bus routes: A provider and passenger perspective. *Transportation Research Part E: Logistics and Transportation Review*, *43*(4), 453–478. https://doi.org/10.1016/J.TRE.2005.09.010

Shin, H., Kim, D.-K., Kho, S.-Y., & Cho, S.-H. (2021). Valuation of Metro Crowding Considering Heterogeneity of Route Choice Behaviors. *Transportation Research Record: Journal of the Transportation Research Board*, *2675*(2), 162–173. https://doi.org/10.1177/0361198120948862

Shires, J. D., & de Jong, G. C. (2009). An international meta-analysis of values of travel time savings. *Evaluation and Program Planning*, *32*(4), 315–325. https://doi.org/10.1016/j.evalprogplan.2009.06.010

Spring, G. S. (2000). Integration of Safety and the Highway Capacity Manual. In Werner Brilon (Ed.), *Transportation Research Circular E-C018: 4th International Symposium on Highway Capacity* (pp. 63–72). Transportation Research Board. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.602.5947&rep=rep1&type=pdf

Sun, H., Forsythe, W., & Waters, N. (2007). Modeling urban land use change and urban sprawl: Calgary, Alberta, Canada. *Networks and Spatial Economics*, *7*(4), 353–376. https://doi.org/10.1007/S11067-007-9030-Y/METRICS

Tan, L., & Ma, C. (2020). Choice behavior of commuters' rail transit mode during the COVID-19 pandemic based on logistic model. *Journal of Traffic and Transportation Engineering (English Edition)*. https://www-sciencedirect-com.ezproxy.lib.ucalgary.ca/science/article/pii/S2095756420301197?via%3Dihub

Taylor, W., & Brogan, J. (1978). *Level-of-Service Concepts in Urban Public Transportation*.

TCRP. (2013). *User's Guide - Transit Capacity and Quality of Service Manual, Third Edition*. National Academies Press. https://doi.org/10.17226/24766

TCRP Project A-15C. (1999). Transit Capacity and Quality of Service Manual 1st Edition. In *Transportation Research Board* (Issue 5).

TCRP Project A-15C. (2003a). Bus Transit Capacity. In *Transit Capacity and Quality of Service Manual, Third Edition* (Third). Transportation Research Board.

TCRP Project A-15C. (2003b). *Transit Capacity and Quality of Service Manual, 2nd Edition* (2nd Editio). Transportation Research Board.

TCRP Project A-15C. (2013). *Transit Capacity and Quality of Service Manual - Chapter 5, Third Edition*. Transportation Research Board. https://doi.org/10.17226/24766

Tirachini, A., & Cats, O. (2020). COVID-19 and public transportation: Current assessment, prospects, and research needs. *Journal of Public Transportation*, *22*(1), 1–34. https://doi.org/10.5038/2375-0901.22.1.1

Tirachini, A., Hensher, D. A., & Jara-Díaz, S. R. (2010). Comparing operator and users costs of light rail, heavy rail and bus rapid transit over a radial public transport network. *Research in Transportation Economics*, *29*(1), 231–242. https://doi.org/10.1016/j.retrec.2010.07.029

Transit Cooperative Research Program. (2003). *TCRP SYNTHESIS 48: Real-Time Bus Arrival Information Systems*. www.TRB.org

Tribby, C. P., & Zandbergen, P. A. (2012). High-resolution spatio-temporal modeling of public transit accessibility. *Applied Geography*, *34*, 345–355. https://doi.org/10.1016/J.APGEOG.2011.12.008

Turnquist, M. A. (1978). A MODEL FOR INVESTIGATING THE EFFECTS OF SERVICE FREQUENCY AND RELIABILITY ON BUS PASSENGER WAITING TIMES. *Transportation Research Record*, *663*, 70–73.

Tyrinopoulos, Y., & Antoniou, C. (2008). Public transit user satisfaction: Variability and policy implications. *Transport Policy*, *15*(4), 260–272. https://doi.org/10.1016/J.TRANPOL.2008.06.002

UITP. (2020). *Public Transport Authorities and COVID-19: Responses from the Front Line*. 1–7.

Uno, N., Kurauchi, F., Tamura, H., & Iida, Y. (2009). Using Bus Probe Data for Analysis of Travel Time Variability. *Journal of Intelligent Transportation Systems*, *13*(1), 2–15. https://doi.org/10.1080/15472450802644439

Verbas, İ. Ö., & Mahmassani, H. S. (2013). Optimal Allocation of Service Frequencies over Transit Network Routes and Time Periods. *Transportation Research Record: Journal of the Transportation Research Board*, *2334*(1), 50–59. https://doi.org/10.3141/2334-06

Vickrey, W. (1955). Some Implications of Marginal Cost Pricing for Public Utilities. *The American Economic Review*, *45*(2), 605–620.

Wardman, M. (2004). Public transport values of time. *Transport Policy*, *11*(4), 363–377. https://doi.org/10.1016/j.tranpol.2004.05.001

Wardman, M. (2013, January 17). Value of Time Multipliers: A Review and Meta-analysis of European-wide Evidence. *Transportation Research Board 92nd Annual Meeting*.

Wardman, M., Chintakayala, V. P. K., & de Jong, G. (2016). Values of travel time in Europe: Review and meta-analysis. *Transportation Research Part A: Policy and Practice*, *94*, 93–111. https://doi.org/10.1016/j.tra.2016.08.019

Watkins, K. E., Ferris, B., Borning, A., Rutherford, G. S., & Layton, D. (2011). Where Is My Bus? Impact of mobile real-time information on the perceived and actual wait time of transit riders. *Transportation Research Part A: Policy and Practice*, *45*(8), 839–848. https://doi.org/10.1016/j.tra.2011.06.010

Watson, L. S., Alvarez, D., Barker, J. B., Barnes, L. E. E., Barnes, R. L., Blair, G. L., Garber, C., Irwin, R. H., & Lee, D. A. (2002). TCRP REPORT 78 - Estimating the Benefits and Costs of Public Transit Projects: A Gudebook for Practitioners. In *TRANSIT COOPERATIVE RESEARCH PROGRAM*.

Whelan, G., & Crockett, J. (2009). Investigation of the Willingness to Pay to Reduce Rail Overcrowding. *Proceedings of the International Conference on Choice Modeling*.

WHO. (2020, March 11). *WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020*. https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020

Wilson, A. G. (1970). The Use of the Concept of Entropy in System Modelling. *Operational Research Quarterly*, *21*(2), 247–265. https://doi.org/https://doi.org/10.2307/3008157

Wirasinghe, S. C. (1990). Re-Examination of Newell's Dispatching Policy and Extension to a Public Bus Route with Many to Many Time-Varying Demand. In M. Koshi (Ed.), *Transportation and Traffic Theory* (pp. 363–377). Elsevier.

Wirasinghe, S. C. (1993). COST BASED APPROACH TO SCHEDULING TRAVEL TIME ON A PUBLIC TRANSPORTATION ROUTE. In C. F. Daganzo (Ed.), *12th International Symposium on The Theory f Traffic Flow and Transportation* (pp. 205–216). Elsevier Science Publishers B.V. . https://trid.trb.org/view/535554

Wirasinghe, S. C. (2021a). Bus Route and Stop Location. In *Transit Systems - Book in preparation* (1st ed., pp. 1–49). Department of CIvil Engineering, University of Calgary.

Wirasinghe, S. C. (2021b). Control and Reliability. In *Transit Systems - Book in preparation* (1st ed.). Department of Civil Engineering, University of Calgary.

Wirasinghe, S. C., & Ghoneim, N. S. (1981). Spacing of Bus-Stops for Many to Many Travel Demand. *Transportation Science*, *15*(3), 210–221. https://doi.org/10.1287/TRSC.15.3.210

Wirasinghe, S. C., & Liu, G. (1995a). Determination of the number and locations of time points in transit schedule design — Case of a single run. *Annals of Operations Research 1995 60:1*, *60*(1), 161–191. https://doi.org/10.1007/BF02031945

Wirasinghe, S. C., & Liu, G. (1995b). Optimal schedule design for a transit route with one intermediate time point. *Transportation Planning and Technology*, *19*(2), 121–145. https://doi.org/10.1080/03081069508717563

Wirasinghe, S. C., & Vandebona, U. (2011). Route layout analysis for express buses. *Transportation Research Part C: Emerging Technologies*, *19*(2), 374–385. https://doi.org/10.1016/j.trc.2010.05.021

Yang, H., Yan Kong, H., & Meng, Q. (2001). Value-of-time distributions and competitive bus services. *Transportation Research Part E: Logistics and Transportation Review*, *37*(6), 411–424. https://doi.org/10.1016/S1366-5545(01)00009-6

Yap, M., Cats, O., & van Arem, B. (2020). Crowding valuation in urban tram and bus transportation based on smart card data. *Transportmetrica A: Transport Science*, *16*(1), 23–42. https://doi.org/10.1080/23249935.2018.1537319

Yu, B., Yang, Z., Sun, X., Yao, B., Zeng, Q., & Jeppesen, E. (2011). Parallel genetic algorithm in bus route headway optimization. *Applied Soft Computing Journal*. https://doi.org/10.1016/j.asoc.2011.05.051

Appendix A

Two independent routes with time varying demand but equal mean VoWT distributions

Consider two independent routes with different demand profiles ($P_1(t)$ and $P_2(t)$) where mean VoWT distribution is the same and headways are equal (here also we assume headway stays constant in the evaluation period of T). Two different routes can get different headways. But we need to have the same headway for both the routes as if in an instance where an operator is trying to sync headways of two intersecting routes to ease transfers. The total cost can be expressed as follows;

$$\int_T Z(t)\,dt = \int_T \frac{1}{2}\gamma_w H P_1(t)dt + \int_T \frac{1}{2}\gamma_w H P_2(t)dt + \int_T \frac{\lambda_{D1}}{H}dt + \int_T \frac{\lambda_{D2}}{H}dt$$

$$= \frac{1}{2}\gamma_w H \int_T (P_1(t)+P_1(t))dt + \frac{T}{H}(\lambda_{D1} + \lambda_{D2})$$

Average total cost per line per unit time can be expressed as follows;

$$\frac{1}{2T}\int_T Z(t)\,dt = \frac{1}{2}\gamma_w H \frac{1}{T}\int_T \frac{(P_1(t)+P_1(t))}{2}dt + \frac{1}{H}\left(\frac{\lambda_{D1} + \lambda_{D2}}{2}\right)$$

$$= \frac{1}{2}\gamma_w H\bar{P} + \frac{1}{H}\overline{\lambda_D}$$

$$H^* = \left[\frac{2\overline{\lambda_D}}{\gamma_w \bar{P}}\right]^{\frac{1}{2}}$$

This is same for n number of routes as well where $n \geq 2$

$$\gamma'_w = \left[\frac{2\overline{\lambda_D}}{\bar{P}H^{*2}}\right]$$

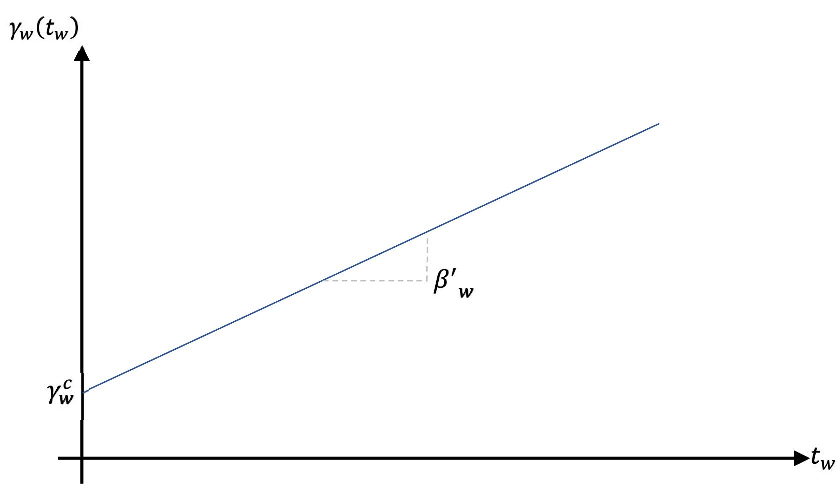
Where $\bar{P}$ is the mean passenger demand for the routes and $\overline{\lambda_D}$ is the mean dispatch cost for the routes.

Case of VoWT that is a Function of the Elapsed Wait Time

Apart from the fact that most of the studies in the field of transportation use a nominal value for the mean VoWT of a passenger(s), some studies have considered different functions of the elapsed wait time for the VoWT instead of a constant value. One school of literature points out that a passenger, without having any information about the potential time of receiving service as in the case of a frequency service where schedule is not known, accumulates stress and anxiety with time due to the sense of waste and uncertainty which will eventually lead him/her to increase their VoWT as he/she waits (Denuit & Genest, 2001; Kocas, 2015; Osuna, 1985). Therefore, it is shown that the 'psychological cost' per unit time - rate of accumulation of stress - increases with the time that a passenger waits.

Wirasinghe (1990) in his study extending Newell's dispatching policy for a time varying passenger demand introduces the total cost function for passenger wait time according to the assumption of a linearly increasing VoWT of a passenger with the elapsed wait time as follows.

$$\gamma_w(t_w) = \gamma_w^c + {\beta'}_w t_w$$



Where γ_w is the VoWT of a passenger, t_w is the elapsed wait time of a passenger at the bus station, γ_w^c is the VoWT of a passenger at the time he/she starting to wait ($t_w = 0$) and ${\beta'}_w$ is the rate at which the VoWT of a passenger is increasing with t_w (a passenger specific constant).

We are discussing here how to obtain a benchmark figure to represent the HLOS in case of a VoWT that is a linear function of the elapsed wait time. In the case of a high frequency (frequency based) service, it is reasonable to assume passengers are arriving uniformly to stations as there is no schedule posted and they are having linearly increasing VoWTs with elapsed wait time. However, this might not be the same for a low frequency service (scheduled service) where passengers plan their arrivals at bus stations to minimize their wait as well as to not miss the bus where passenger arrival distribution at the station will not be uniform. Also, as passengers are aware of the potential arrival time of the bus, there will not be stress building up as they wait as there is no uncertainty involved up until the scheduled bus arrival time. If the bus is late VoWT of the passengers will start to rise beyond scheduled arrival time as now the anxiety and stress will start building up due to uncertainty. But, in a frequency-based service, it is reasonable to disregard such complications allowing us to assume the VoWT of passengers to increase with wait time. Therefore, we will derive this relationship assuming a frequency-based service.

The initial VoWT ($at\ t_w = 0$) can be assumed to be 'zero' (no penalty for having to wait for the bus) as passengers know and accept the fact that they might have to wait. However, this will not be the case in a waiting situation at a transfer as there will be a penalty for having to transfer. The penalty for a transfer can be accounted for in the total cost function through multiplying it by the number of transfers. Due to these reasons, we assume the VoWT to be a linearly increasing function with elapsed wait time with the condition $\gamma_w^c = 0$ so that $\gamma_w(0) = 0$ which leaves us;

$$\gamma_w(t_w) = \beta'_w t_w$$

We can find the total cost of passenger wait time in a headway $h(t)$ by;

$$\int_0^{h(t)} P(t) \int_y^{h(t)} \{\beta_w(x - y)\}\, dx\, dy$$

where $P(t)$ is the passenger arrival rate, $(x - y)$ is the elapsed waiting time up to time x of the passengers arrived at time y and β_w is the average of β'_w values of passengers. Total cost of passengers per unit time and the cost of dispatching busses per unit time can be obtained by dividing the above expression and λ_D by $h(t)$ respectively and can be expressed as;

$$Z = \frac{P(t)\beta_w h(t)^2}{6} + \frac{\lambda_D}{h(t)}$$

Minimizing this in terms of $h(t)$ provides us the headway that minimizes the total cost $h^*(t)$

$$h^*(t) = \sqrt[3]{\frac{3\lambda_D}{P(t)\beta_w}}$$

By plugging in the values for $h(t)$, λ_D and $P(t)$ of an existing operation of a bus route, it is possible to obtain the value for the implied average rate of increasing of VoWT with elapsed wait time β_w^* for the passengers of that route.

$$\beta_w^* = \frac{3\lambda_D}{P(t)h(t)^3}$$

Given we have the distribution of ${\beta'}_w$ values of passengers in a route, it is possible to determine the HLOS as described in chapter 2.

Case of Low and High Frequency Services with Planning and Non-Planning Passengers

Normal practice in the field of transportation planning has been to assume half the headway as the mean waiting time of passengers for generalized cost functions. This is under the assumption that passengers wait only at the bus station to board the bus. However, in normal life, there is a significant amount of people who might incur a waiting time even at the destination of their trip. For example, for a person commuting to work, work starts at 9 am. But the bus he takes has a peak hour headway of 20 minutes and the normal arrival time of the bus at the nearest bus station to work is 8.45 am and 9.05 am. In this case, the passenger has to wait for around 15 minutes at his destination. (However, the VoWT at the destination for this person can be significantly lower than the VoWT at the origin. For this instance, we assume the passengers have the same VoWT both at origin and destination). Therefore, for a nominal commuter, there is a mean waiting time of half the headway at the destination as well. Since this person is planning his trip as he has a specific arrival time, we call them planning passengers and others (without specific arrival time) as non-planning passengers.

For this case, we identify the 'mean waiting time' of a passenger as the total mean waiting time (both at the origin and destination). The mean waiting time, however, changes for both planning and non-planning passengers depending on the type of service being low frequency (schedule-based) service or high frequency (frequency-based) service. Ansari Esfeh et al., (2020) expand on this issue in detail and derive the expressions for mean (expected) wait time, total cost (passenger wait and operator cost) and optimum frequency (reciprocal of the headway) for both the low and high frequency services (Ansari Esfeh et al., 2020). The expressions for the new mean wait times (for a reliable service) are as follows;

The mean wait time for a high frequency service $W_h = \left[\frac{1+\alpha}{2}\right] H$

Where H is the headway of the service and α is the proportion (percentage) of all planning passengers.

The mean wait time for low frequency service $W_l = \left[\frac{1-\alpha(1-\beta)}{2}\right] H$

where β is the proportion (percentage) of planning passengers with a fixed arrival time out of all planning passengers.

Accordingly, he derives the optimum operation frequency for a high frequency service ($g_H^*(t)$) and a low frequency service ($g_L^*(t)$), depending on modified mean wait times, as follows; (these expressions are assuming that the fleet size and bus capacity (seat capacity) are not constraints)

$$g_H^*(t) = \left[\frac{(1+\alpha)\gamma_w P(t)}{2\lambda_D}\right]^{1/2}$$

$$g_L^*(t) = \left[\frac{[1-\alpha(1-\beta)]\gamma_w P(t)}{2\lambda_D}\right]^{1/2}$$

Where γ_w is the mean value of waiting time of passengers.

Accordingly, by plugging in the values of an existing bus route and depending on the type of operation, it is possible to derive the implied VoWT for a high frequency service (γ_{wH}^*) and the implied VoWT for a low frequency service (γ_{wL}^*) as follows;

$$\gamma_{wH}^* = \frac{2\lambda_D}{(1+\alpha)P(t)h_H^2(t)}$$

$$\gamma_{wL}^* = \frac{2\lambda_D}{[1-\alpha(1-\beta)]P(t)h_L^2(t)}$$

where $h_H^2(t) = 1/g_H^2(t)$ and $h_L^2(t) = 1/g_L^2(t)$

Note that $g_H^*(t)$ is not necessarily equal to $g_H(t)$. The same is true for $g_L^*(t)$ and $g_L(t)$. $g_H^*(t)$ is the optimum frequency minimizing the total of passenger wait time cost and bus dispatch cost while $g_H(t)$ is the operational frequency of an existing bus route/operation (high frequency).

With the normal assumption of a half the headway for mean headway, the implied VoWT for both the low and high frequency services will be;

$$\gamma_w^* = \frac{2\lambda_D}{P(t)h(t)^2}$$

Compared to this value, with the new mean wait time values for a high frequency service, the denominator has been multiplied by a value of $(1+\alpha)$ which is greater than 1 higher the proportion of planning passengers, higher the value of $(1+\alpha)$. Hence, higher the value of $(1+\alpha)$, higher the value of denominator and lower the resulting value of the implied VoWT (γ_w^*). In the same manner, it can be shown that higher the proportions of planning passengers and passengers with a fixed arrival time out of planning passengers, higher the implied VoWT (γ_w^*) for a low frequency service.

Case of Stochastic Passenger Demands

If the passenger demand at a certain time of day varies day by day (e.g., passenger demand from 3 pm to 5 pm every weekday) and is in the form of a random variable, then the optimum headway given by the Equation (15) also becomes a function of a random variable and can be shown to have an approximate mean headway as follows (Wirasinghe, 1990);

The mean headway ($\bar{h}(t)$);

$$\bar{h}(t) = \frac{\sqrt{2\lambda_D / \gamma_w \bar{P}(t)}}{\left[1 - \frac{C_{P(t)}^2}{8}\right]}$$

where $\bar{P}(t)$ is the mean passenger demand at time t at any day over several days (e.g., weekdays) and $C_{P(t)}$ is the coefficient of variation of $P(t)$.

Deriving an expression for γ_w using the expression for $\bar{h}(t)$ and substituting values for an existing operation of a route and substituting the existing value of headway for $\bar{h}(t)$ will provide an implied VoWT by the operation. However, as the day-to-day demand variation in the peak hour and most of the off-peak times tend to be small leading to a small $C_{P(t)}$ value. This will lead the result of the expression for $\bar{h}(t)$ not to be significantly different than the result those can be obtained from the Equation (15).

Effect of headway variation on Headway LOS

In reality, the headways can deviate a few minutes from the scheduled headway. Due to the fact that more passengers are probable to arrive in longer headways the mean waiting time is distorted towards the longer headways. To specifically account this, the expected waiting time (mean wait time) can be expressed as follows (Furth & Muller, 2006);

$$E(W) = \frac{E(H)}{2}(1 + CV_H^2)$$

where $E(W)$ is the expected (mean) value of wait time, $E(H)$ is the expected value of headway and CV_H is the coefficient of variation of the headway. Most of the time, the expected headway is similar to the scheduled headway. Therefore, we will assume $E(H)$ is equal to the headway H.

Accordingly, the Equation (14) for the total cost of passenger waiting and bus dispatch can be modified as follows;

$$Z = \left(\frac{H}{2}\right)(1 + CV_H^2)\gamma_w P + {\lambda_D}/{H}$$

It can be shown that, the implied VoWT γ_w^* in the case of an unreliable service;

$$\gamma_w^* = \frac{2\lambda_D}{PH_\mu^2(1 + CV_H^2)}$$

Where, H_μ is the mean headway. Higher the unreliability, higher the value of CV_H^2 would be and lower the value of γ_w^*. Hence lower the LOS of the service.

Appendix B

Proof of the average loading expression for a many-to-many bus route

Proposition

In this section the expression by Qin, (2014) and Tirachini et al., (2010) is true for a many-to-many demand bus route. For this derivation, it is assumed that the speed of the bus is constant. In an actual case of a bus route, the buses would have approximately equal average speeds as the scheduled route travel time is constant over the day. For the calculations, the average speed can be used instead of the theoretical constant speed. Therefore, let us assume that the constant speed of the bus route be given by V.

Let the route trip travel time of the bus route be given by T' hours and the rate of boarding and alighting at any time t during T' be given by $B(t)$ and $A(t)$ passengers per hour respectively. Therefore, the number of passengers in the bus along the bus route at any given moment t – $N(t)$ – is given by

$$N(t) = \int_0^t B(t)dt - \int_0^t A(t)dt \qquad \text{B-1}$$

Consider a bus as a point system where passengers board and alight with time. Then the number of passengers inside the bus, hence the load of the bus, changes with time. Therefore, the average load of a bus over a time period of a single run can be given by the ratio of total passenger hours over a period of a run divided by the duration of a run, which is T' in this case.

Therefore, the average load of the passengers travelling in the bus over a run can be given by L

$$L = \frac{\int_0^{T'} N(t)dt}{T'} \qquad \text{B-2}$$

Equation (B-2) can be expanded using the Equation (B-1) as follows.

$$L = \frac{\int_0^{T'} \left(\int_0^t B(t)dt - \int_0^t A(t)dt\right) dt}{T'} \qquad \text{B-3}$$

In Equation (B-3), the expression is given as a function of time. This can be converted to a function of the distance x measured along the route from the start stop as follows. The length of the bus route is D.

$$L = \frac{\int_0^D \left(\int_0^x B(t) \frac{dt}{dx} dx - \int_0^x A(t) \frac{dt}{dx} dx \right) \frac{dt}{dx} dx}{T'} \qquad \text{B-4}$$

As t – the elapsed time from the start of the trip of a bus – reaches T', the x – elapsed distance from the start of the route – reaches D.

The round-trip time T' of the bus route can be expressed in terms of the speed and length of the route as follows.

$$T' = \frac{D}{V} \qquad \text{B-5}$$

In Equation (B-4), $\frac{dt}{dx}$ is the speed of the bus at time t and distance x. As the bus route has a constant speed, the $\frac{dt}{dx}$ can be replaced with V. Therefore, the Equation (B-4) can be modified as follows.

$$L = \frac{\int_0^D \left(\int_0^x B(t) \frac{dt}{dx} dx - \int_0^x A(t) \frac{dt}{dx} dx \right) dx}{T'V} \qquad \text{B-6}$$

Replacing T' in Equation (B-6) with Equation (B-5) provides,

$$L = \frac{\int_0^D \left(\int_0^x B(t) \frac{dt}{dx} dx - \int_0^x A(t) \frac{dt}{dx} dx \right) dx}{D} \qquad \text{B-7}$$

<u>Proof</u>

Expression for the average loading by Qin (2014) and Tirachini et. al, (2010) is shown below.

$$L = PH \frac{\bar{l}}{D} \qquad \text{B-8}$$

P is the demand for boarding the bus route as given in passengers per hour, H is the headway of the bus route given in hours, and $\bar{l}$ is the average trip distance of the passengers. It can be shown that the Equation (B-8) can be deduced to Equation (B-7) for a many-to-many demand bus route.

The following derivation is carried out for the many-to-many demand bus route described at the beginning of this section.

It is assumed that the average trip distance of the passengers during the time period T' is equal to that of the average trip distance of the passengers traveling in buses during time period T'. Average trip distance of the passengers can be obtained by dividing the total passenger kilometers travelled by all the passengers that boarded a certain bus along the route by the number of passengers boarded that bus along the route. Therefore,

$$\bar{l} = \frac{Total\ passenger\ km}{total\ number\ of\ passengers\ boarded\ the\ bus}$$

Number of passenger km can be obtained by multiplying the number of passengers traveling in the bus over an infinitely small distance dx by the distance dx along the entire route. Number of passengers traveling at a dx distance can be obtained using the expression below.

$$\int_0^x B(t)\frac{dt}{dx}dx - \int_0^x A(t)\frac{dt}{dx}dx$$

Also, the number of passengers that would use the bus – board the bus – during a run is given by PH.

Therefore,

$$\bar{l} = \frac{\int_0^D \left(\int_0^x B(t)\frac{dt}{dx}dx - \int_0^x A(t)\frac{dt}{dx}dx\right)dx}{PH} \qquad \text{B-9}$$

Equation (B-8) can be modified using Equation (B-9) as follows.

$$L = \frac{\int_0^D \left(\int_0^x B(t)\frac{dt}{dx}dx - \int_0^x A(t)\frac{dt}{dx}dx\right)dx}{D} \qquad \text{B-10}$$

As Equation (B-10) is equal to Equation (B-7), it can be deduced that the expression for the average load of a bus route – Equation (B-8) – as proposed by Qin, (2014) and Tirachini et al., (2010) for a many-to-one bus route is true and can be applied for a many-to-many bus route too.

Appendix C

The total cost function for optimum headway when CPF varies exponentially with loading factor

The exponential form of CPF with respect to the loading factor is investigated because it has been shown to be a better fit in some cases (Qin, 2014). The CPF can be expressed as

$$CPF = 1 + ae^{bL}$$

where 'a' and 'b' are route-specific constants.

Substituting for L from Equation (41), we have

$$CPF = 1 + ae^{b\frac{PH\bar{l}}{DS}}$$

Therefore, the total crowding cost (without riding cost) is

$$TC_{cr} = \frac{PT\bar{l}\gamma_r^o ae^{\frac{bPH\bar{l}}{DS}}}{D}$$

This equation shows the average crowding cost of passengers on a bus route per hour. The crowding cost now varies exponentially with headway.

Hence, the total cost Z is

$$Z = \left(\frac{H}{2}\right)\gamma_w P + \frac{PT\bar{l}\gamma_r^o ae^{\frac{bPH\bar{l}}{DS}}}{D} + {}^{\lambda_D}/_{H}$$

Taking the partial derivative with respect to H gives

$$\frac{\gamma_w P}{2} + \frac{abP^2\bar{l}^2T\gamma_r^o e^{\frac{bP\quad\bar{}}{DS}}}{D^2S} = {}^{\lambda_D}/_{H^2}$$

This equation can be solved using numerical methods. Although an approximation of the analytical solution of this equation can be obtained using the Taylor series expansion for the exponential CPF, it does not provide any useful insights as the solution turns out to be complex.

Appendix D

```
clc
clear all

M = 15/60; %Mean
sg = 5/60; %Sigma - standard deviation
lamda = 120;
P = 1000;
gp = 20;
gw = 20;
gr = 8;
a = 1.25;
b = 1.25;
theta = gw/gr;
n = 6;
N = 30;
D = 20;
l = 10;

E = (log(1+(sg/M)^2))^0.5; %eeta the shape parameter of the log normal distribution - sigma

L = log(M)-(0.5*E^2); % lamda the scale parameter of the log normal distribution - mean

S = (3/60):(0.5/60):1.5;

fun = @(x) ((1./(x.*E*sqrt(2*pi))).*exp(-1.*(((log(x)-L).^2))/(2*E^2)));

fun2 = @(x) ((1./(E*sqrt(2*pi))).*exp(-1.*((((log(x)-L).^2))/(2*E^2))));

for i=1:length(S)
```

```
        fs(i) = fun(S(i));

        F(i) = integral(@(x) fun(x),0,S(i));

        f(i) = integral(fun2,S(i),inf);

end

Q = fs.*(f-(S.*(1-F)))+(1-F).^2;

gammaR = ((gp.*fs) - (n*N*lamda/P))./((l*n/D)-(Q.*(a+(b*theta))));

figure (1)
plot(S*60,Q,'b'), grid on, grid minor
legend('Q(S)')
axis([15 30 -0.1 0.4])
xlabel('Scheduled Travel Time (Minutes)')
ylabel('Q(S)')
print(gcf, 'Q(S) with Scheduled Travel Time.png', '-dpng', '-r600');

figure (2)
plot(S*60,gammaR,'r'), grid on, grid minor
legend('gammaR')
axis([15 30 -10 60])
xlabel('Scheduled Travel Time (Minutes)')
ylabel('Implied VoRT ($/hr./pass.)')
print(gcf, 'gammaR with Scheduled Travel Time.png', '-dpng', '-r600');
```

Appendix E

Simplification of Equation (108) to Equation (109) – Chapter 6: Reliability LOS

$$Z(S) = \frac{P\bar{l}nS\gamma_r}{D} + P(a\gamma_r + b\gamma_w)\int_s^{\infty} f(t)dt \int_s^{\infty} (t - S)f(t)dt + P\gamma_p \int_s^{\infty} f(t)dt + (nS + T')\lambda N \quad \text{E - 1}$$

Equation (E - 1) is equal to Equation (108). What follows is the simplification of the Equation (E - 1).

As $\int_s^{\infty} f(t)dt = 1 - F(S)$,

$$Z(S) = \frac{P\bar{l}nS\gamma_r}{D} + (P(a\gamma_r + b\gamma_w))[1 - F(S)]\left[\int_s^{\infty} tf(t)dt - \int_s^{\infty} Sf(t)dt\right] + P\gamma_p[1 - F(S)] + (nS + T')\lambda N \quad \text{E - 2}$$

$$Z(S) = \frac{P\bar{l}nS\gamma_r}{D} + (P(a\gamma_r + b\gamma_w))[1 - F(S)]\left[E(t) - \int_0^{S} tf(t)dt - S[1 - F(S)]\right] + P\gamma_p[1 - F(S)] + (nS + T')\lambda N \quad \text{E - 3}$$

Equation (E - 3) is further expanded to be

$$Z(S) = \frac{P\bar{l}nS\gamma_r}{D} + (P(a\gamma_r + b\gamma_w))\left[E(t) - \int_0^{S} tf(t)dt - S + SF(S) - E(t)F(S) + F(S)\int_0^{S} tf(t)dt + SF(S) - SF^2(S)\right] + P\gamma_p[1 - F(S)] + (nS + T')\lambda N \quad \text{E - 4}$$

$$Z(S) = \frac{P\bar{l}nS\gamma_r}{D} + (P(a\gamma_r + b\gamma_w))\left[E(t) - \int_0^S tf(t)dt - S + 2SF(S) - E(t)F(S) + F(S)\int_0^S tf(t)dt - SF^2(S)\right] + P\gamma_p[1 - F(S)] + (nS + T')\lambda N \qquad \text{E - 5}$$

Taking the partial derivative of the $F(S)$ in terms of S.

$$\frac{\partial(F(S))}{\partial S} = f(S) \qquad \text{E - 6}$$

Taking the partial derivative of the Equation (E - 5) in terms of S.

$$\frac{\partial Z}{\partial S} = \frac{P\bar{l}n\gamma_r}{D} + (P(a\gamma_r + b\gamma_w))\left[0 - Sf(S) - 1 + 2\big(Sf(S) + F(S)\big) - E(t)f(S) + F(S)Sf(S) + f(S)\int_0^S tf(t)dt - S2F(S)f(S) - F^2(S)\right] - P\gamma_p f(S) + n\lambda N \qquad \text{E - 7}$$

$$\frac{\partial Z}{\partial S} = \frac{P\bar{l}n\gamma_r}{D} + (P(a\gamma_r + b\gamma_w))\left[f(S)\left(E(t) - \int_S^\infty tf(t)dt - E(t)\right) - \big(1 - 2F(S) + F^2(S)\big) + Sf(S)(1 - F(S))\right] - P\gamma_p f(S) + n\lambda N \qquad \text{E - 8}$$

$$\frac{\partial Z}{\partial S} = \frac{P\bar{l}n\gamma_r}{D} + (P(a\gamma_r + b\gamma_w))\left[Sf(S)\big(1 - F(S)\big) - (1 - F(S))^2 - f(S)\int_S^\infty tf(t)dt\right] - P\gamma_p f(S) + n\lambda N \qquad \text{E - 9}$$

$$\frac{\partial Z}{\partial S} = \frac{P\bar{l}n\gamma_r}{D} + (P(a\gamma_r + b\gamma_w))\left[f(S)\left(S(1 - F(S)) - \int_S^\infty tf(t)dt\right) - (1 - F(S))^2\right] - P\gamma_p f(S) + n\lambda N \qquad \text{E - 10}$$

$$\frac{\partial Z}{\partial S} = \frac{P\bar{l}n\gamma_r}{D} - \left[f(S)\left(\int_S^\infty tf(t)dt - S\big(1 - F(S)\big)\right) + (1 - F(S))^2\right](P(a\gamma_r + b\gamma_w)) + n\lambda N - P\gamma_p f(S) \quad \text{E - 11}$$

Equation (E - 11) is equal to the Equation (109)