PERFORMANCE ANALYSIS OF HUBLI-DHARWAD BUS RAPID TRANSIT SYSTEM

Thesis

Submitted in partial fulfilment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

By

SHIVARAJ HALYAL

(Register No. 187506CV505)

Under the guidance of

Dr. Raviraj H.Mulangi

DEPARTMENT OF CIVIL ENGINEERING

NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA,

SURATHKAL, MANGALORE – 575025

JUNE, 2023

DECLARATION

I hereby *declare* that the Research Thesis entitled "Performance Analysis of Hubli-Dharwad Bus Rapid Transit System" which is being submitted to the National Institute of Technology Karnataka, Surathkal in partial fulfilment of the requirements for the award of the Degree of Doctor of Philosophy in Department of Civil Engineering is a *bonafide report of the research work carried out by me*. The material contained in this Research Thesis has not been submitted to any University or Institution for the award of any degree.

Place: NITK Surathkal

Date: 08/06/2023

Shivaraj Halyal
Register Number: 187506CV505
Department of Civil Engineering

CERTIFICATE

This is to certify that the Research Thesis entitled "Performance Analysis of Hubli-Dharwad Bus Rapid Transit System" submitted by Shivaraj Halyal (Register Number: 187506CV505) as the record of the research work carried out by him, *is accepted as the Research Thesis submission* in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy.

Dr. Raviraj H. Mulangi
(Research Guide)
Assistant Professor
Department of Civil Engineering

NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA, SURATHKAL
DEPT. OF CIVIL ENGG.

Dr. Subhash C. Yaragal
(Chairman-DRPC)
Head of the Department
Department of Civil Engineering

Chairman (DRPC)
Department of Civil Engineering
National Institute of Technology Karnataka, Surathkal
Mangalore - 575 025, Karnataka, INDIA

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to **Dr. Raviraj H. Mulangi**, Assistant Professor, Department of Civil Engineering for his noble guidance, support with full encouragement and enthusiasm. Apart from the technical guidance, it was his constant encouragement, affection, support and solace during the moments of despair that have been behind the successful completion of this thesis. This research work would not have been possible without his guidance and invaluable suggestions.

I would like to thank the members of my Research Progress Assessment Committee **Dr. Ramesh M. R**, Department of Mechanical Engineering and **Dr. Gangadhar Mahesh**, Department of Civil Engineering, for their insightful suggestions, comments and encouragement which has helped me to improve my research work.

I am grateful to **Prof. B. R. Jayalekshmi**, Professor and Head, Department of Civil Engineering for providing necessary facilities for doing this work. I also take this opportunity to thank former Department Heads namely, **Prof. Varghese George** and **Prof. K. Swaminathan**, for their timely help during my entire research period. I also thank all the teaching and supporting staff of the Civil Engineering Department for their help and support provided during the research work.

I also immensely grateful and extend the deepest sense of gratitude to **Shri. P Rajendra Cholan** and **Shri. Bharath Selvan**, former and current Managing Director of North Western Karnataka Road Transport Corporation (NWKRTC) and Hubli-Dharwad Bus Rapid Transit System (HDBRTS), **Shri. Vivekananda V**, Divisional Controller of NWKRTC, **Shri. Ganesh Rathod** and **Shri. M. Rajkumar**, former and current General Manager of ITS, HDBRTS for providing the necessary data, information, guidance as well as humble support during current research work. In the same time, I take this opportunity to convey my greatest gratitude to all the **Technical**, **Administrative** and **Supporting Staff** of NWKRTC and HDBRTS who have helped me directly or indirectly in this research work. This research work would not have been possible without their support in constant way throughout the work during data collection, data validation as well as for clarification whenever needed.

I am also much indebted to **Dr. Ashok Shettar**, Voice-Chancellor, **Prof. P. G. Tewari**, Dean of Academics, former Registrar **Prof. B. L. Desai** and Registrar **Dr. Basavaraj Anami** of **KLE Technological University**, **Hubli** for providing the supreme research environment, research policy, support, and facility at the university to pursue the objectives under this able administration. I am also obliged to former Head Prof. **S. S. Quadri** and Current Head **Prof. M. V. Chitawadagi**, School of Civil Engineering, KLE Technological University for their support throughout the research work.

Last but not least, my **parents**, my **wife** , my **son**, and my research colleague **Harsha M.M** deserve special thanks for their continuous moral support and encouragement whenever I was in need throughout the current research work. Support from them was one of the driving forces to complete this work.

SHIVARAJ HALYAL

ABSTRACT

The concept of ‘Smart Mobility’ is one of the innovative solutions to tackle many urban transportation-related issues; that will connect various elements of technology and mobility, and Intelligent Transport System (ITS) is a step toward implementing it. The ITS integrates transportation system users with vehicles and infrastructure using information and communication technology. Bus Rapid Transit System (BRTS) is a state-of-the-art smart mobility system and is a boon for urbanized areas, which are affected by numerous transportation-connected glitches. The role of BRTS has now been recognized as essential for physically active, economically sound, and energy-efficient cities.

The BRTS has a combined structure of various exclusive features with a strong identity and distinctiveness. A dedicated lane of bus operation is the critical parameter of any BRTS, which will enhance its performance from all the perspectives. The interference of mixed traffic with the operation of BRTS buses, although it only occurs on a few road segments, can compromise the end-to-end travel time of the whole system due to congestion and contribute to reliability-related problems. Many internal and external factors will also influence the Travel Time Reliability (TTR) and Travel Time Variability (TTV) temporally as well as spatially and finally cause an impact on whole system performance.

The main motivation behind this research is to study the impact of such non-dedicated, and dedicated lanes of BRTS bus operation on its overall system performance from multiple perspectives by identifying bus stations, routes, and segments that are critical in nature. The current study used Automatic Vehicle Location (AVL) data and Automatic Passenger Count (APC) data from the recently implemented Hubli-Dharwad Bus Rapid Transit System (HDBRTS) as a case study.

HDBRTS buses operate as express and non-express routes along the single linear corridor between twin cities Hubli and Dharwad. Express route buses serve the limited bus stations, whereas non-express route buses serve all the bus stations. Most of the buses of both environments will run from terminal to terminal, such as the terminal

at the Hubli side to the terminal at the Dharwad side, which is named as UP direction, and the terminal at the Dharwad side to the terminal at the Hubli side which is named as a DOWN direction in the current study. Most of the length of this corridor has a dedicated nature for the bus operation, and a small part of it has non-dedicated nature, too; hence HDBRTS is considered as a hybrid-based BRT system. The BRT corridor from Hosur Circle of Hubli City to the Jubilee Circle of Dharwad is dedicated in nature, in UP and DOWN directions and the corridor from Hosur Circle Hubli to CBT Hubli is completely non-dedicated in nature. For the current research work, express routes and non-express routes were considered for the route level analysis, and one dedicated segment at the Dharwad side, one dedicated segment, and one non-dedicated segment towards Hubli were considered for the segment level analysis.

From the preliminary study carried out for the HDBRTS, it was understood that, higher dwell time, bus bunching at the stations, signal delays at intersections, peak, and off-peak traffic hours of the day were few of the general incidences that were actually influencing the travel time variability of the buses and further leading to the less travel time reliability of the system. Keeping all those points in observance, in the first part of the current study, systematic smart data-based end-to-end travel time variability and reliability analysis have been carried out for the HDBRTS.

Analysis has been done for two routes (express and non-express) and three segments exclusively (Two dedicated and one non-dedicated) in two stages. Travel time data points have been extracted for all the days of the week and different hours of the days as different aggregations. In the first stage, descriptive statistics and TTR analysis of the selected data points were done, whereas, in the second stage of the study, probability distribution fitting was carried out for both the routes and selected segments separately with seven potential continuous distributions to characterize the travel time. In the analysis, distribution parameters were extracted using the Maximum Likelihood Estimations (MLE) method. Kolmogorov-Smirnov (K-S) test was used to extract the distribution parameters and check for the goodness of the fit of each distribution. Hence based on the K-S p-value, the robustness of best-fit distribution was selected and ranked amongst all the choices, for describing the travel time data points under different conditions considered. In conclusion, as per the total number of cases passed by each selected distribution model, distribution performance was established at different ratios for all routes and segments. At the end of the probability distribution fitting with the travel time data points, the best fit distribution parameters were tried to compare with the passenger demand of that particular time stamp. From the analysis, it was found that peak and off-peak hours have a direct influence on the change in the characteristics of route and segment travel time and subsequent reliability indices. Except for the higher values of reliability indices during peak hours, the performance of the express routes seems to be more reliable. From the distribution study, the Generalized Extreme Value (GEV) distribution stood first on the best performance distributions list for the routes, dedicated segments, and even a non-dedicated segment. Hence it shows the robustness of GEV in explaining the heterogenous Travel Time (TT) characteristics. Based on TTR analysis with GEV distribution, it was inferenced that passenger demand and Buffer Time Index (BTI) have a direct correlation with the variations in the GEV shape parameter 'k'.

In the second part of the study, travel time reliability modelling was carried out with observed and unobserved independent variables obtained from HDBRTS operations. The travel time data points have been extracted according to the selected segments. Modelling was carried out with the Multiple Linear Regression (MLR) technique. Average travel time (ATT) and buffer time (BT) were the two dependent variables chosen from the operator's and passengers' point of view. Independent variables were selected based on permutation and combination of multiple covariables. Length of the segment, passenger demand, bus stop density, intersection density, peak and off-peak periods, and land use type were the finalized independent variables. Finally, two MLR models were developed in relation to the two dependent and eight independent variables. The performance of both models was examined with the adjusted R square values and t-statistics and significance values of individual covariables of both the developed models. With the higher adjusted R square values of

0.795 and 0.804, respectively, ATT and BT as dependent variables have shown superior explanatory power in describing the system's reliability.

In the third part of the current study, as passenger demand forecast for the public transit system is a crucial and inevitable step in keeping the public transit system in the direction of continuous upgrading mode in their performance; hence forecasting of passenger demand was done with Long Short-Term Memory (LSTM) using the three months Automatic Passenger Counter (APC). Then the forecasting of passenger demand was also done with Seasonal Autoregressive Integrated Moving Average (SARIMA) models, and the comparison of the forecasting accuracy of both methods was made using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). Furthermore, to validate the results, a novel approach has been adopted for the process, by following some more time series resampled with different time intervals. The study shows that LSTMs will be used satisfactorily in the traffic conditions present between Hubli- Dharwad, for forecasting passenger demand using APC data.

As the last objective, the travel time reliability-based Level of Service (LOS) of the HDBRTS has been established for three operating conditions, such as route, dedicated segment, and non-dedicated segment. Planning Time Index (PTI), Buffer Time Index (BTI), and Travel Time Index (TTI) were the three reliability indices used to establish LOS. K-mean clustering method was used to develop clusters, and silhouette analysis was carried out to validate the quality of the clusters. Most of the clusters were found to be reasonable and opt with an average silhouette coefficient of more than 0.5. Hence LOS development in the current study better suits with selected data points of travel time reliability indices.

Based on the analysis and obtained results of current research work, finally elaborated, three stages of recommendations were made to the operator for improving the performance of HDBRTS.

Keywords: Travel time variability, travel time reliability, Bus rapid transit system, Intelligent Transportation System , Passenger demand forecasting

CONTENTS	PAGE NO.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

BRTS – Bus rapid transit system

H-D – Hubli-Dharwad

TCQSM – Transit Capacity and Quality Service Manual

TT – Travel Time

TTT – Transit Travel Time

TTR – Travel Time Reliability

TTV – Travel Time Variability

BT – Buffer Time

FR – Failure Rate

DT – Dwell Time

LOS – Level of Service

ITDP - Institute of Transport Development and Policy

ML – Machine Learning

ITS - Intelligent Transportation System

ITMS – Intelligent Transit Management System

PPHPD - Passengers per hour per direction

SPV - Special Purpose Vehicle

TCRP - Transit Cooperative Research Program

CBD - Central Business District

DWD -Dharwad

HUB - Hubli

CBT - Central Bus Terminal

HDMC - Hubli-Dharwad Municipal Corporation

NWKRTC - North-Western Karnataka Road Transport Corporation

HDUDA - Hubli- Dharwad Urban Development Authority

HDBRTS - Hubli-Dharwad Bus Rapid Transit System

KMPH - Kilometre Per Hour

MT - Mixed Traffic

SPV - Special Purpose Vehicle

AVL - Automatic Vehicle Location

NUTP - National Urban Transport Policy

CHAPTER 1

INTRODUCTION

1.1 GENERAL

India, being one of the fastest-growing economies in the world, is progressing at the rate of 7.5% per annum (World Bank, 2022-23). With an exponential increase in the population, the country is also working towards transforming itself over the three successive decades. Any country's growth mainly depends on its cities, and its urban population. India's census record of 2011 clearly shows that around 31.2% of the total population (377 million) resides in urban. The UN estimated Indian census records will rise to 40% (590 million) by 2030, subsequently to 58% (875 million) by the year 2050. Even though around 31% of the total Indian population resides in urbanised zones; however, it contributes only 63% of India's Gross Domestic Product (GDP) (Prasad, D. R., 2017). Fast population growth has been seen in the past, particularly in most of the country's metropolitan cities such as Mumbai, Kolkata, and Delhi. It has been observed that these cities have more than 10 million individuals. Chennai, Hyderabad, Ahmadabad, and Bangalore each have more than 5 million population (Office of the Registrar General of India 2001).

Fast growth in the urbanised population has resulted in higher usage of the privatised vehicle, and their activities in the urban area are also growing rapidly, especially in most of the low-income-based countries of Asia. As a result, there is hasty progress in urbanization, increasing the urbanised income; with this, vehicle production is progressively moving towards the higher side, and markets are becoming saturated in the Organization for Economic Co-operation and Development (OECD) countries. In India, privatised vehicle ownership has reached around 87% of South Asia's privatised vehicles; it is seen that these numbers are doubling every four years over the last few decades (MORTH, 2017). It has also revealed by the Road Transport Year Book (RTYB) of the Ministry of Road Transport and Highways (MORTH) report in 2019, that privatised vehicle numbers have grown even more rapidly in the cities than countrywide. Indeed, urban privatised vehicles growth rates have far beaten urban population growth rates, which have been more astonishing. Figure 1.1 shows the trend in the vehicle population as per the RTYB of MoRTH, 2019.

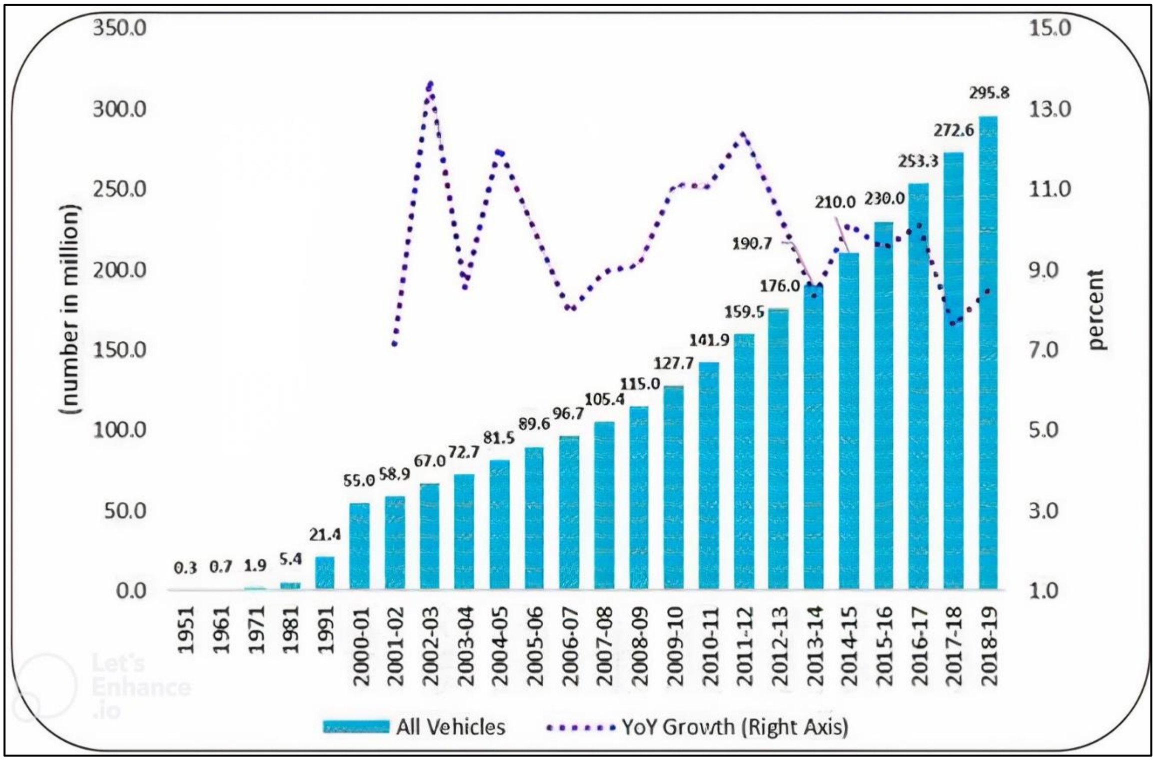


Figure 1.1 Vehicle Population in Past Years (RTYB, MoRTH, 2019)

The exponential growth of privatised vehicle ownership, including commercial vehicles, as shown in the figure 1.1, their increased activities in urbanised areas, and the increasing trend of India's population have brought a wide range of adverse effects on urban system management activities. Heavy traffic congestion is one of such adverse effect in the urbanised areas, and is leading to a loss in substantial time due to increased journey times and loss in the productivity of the people and resources. Because of the increased number of privatised vehicles and other energy-intensive activities in metropolitan centers like Delhi, Mumbai, etc., and the fact that until recently, these activities have been characterized by high pollution intensities, air quality has been poor in these centers since the late 1980s. Statistical studies and investigations carried out by MORTH in 2017 have also revealed the air quality values in terms of daily average suspended particulate levels, which are strongly associated with respiratory and cardiovascular diseases. Obtained figures were beyond the limits suggested by World Health Organization (WHO) (CPCB, 1996, 2004). Besides, increasing vehicle ownership in India is also causing higher accident rates; it has already reached the world's highest list, with accident severity expressed in terms of the number of persons killed per 100 accidents rose from 32.4 in 2018 to 33.7 in 2019 (Accident Report, MORTH, 2022).

The modal shift to the Public Transit System is an elementary solution to tackle adverse effects of increased population and subsequent privatised vehicles activities on urban system management such as severe traffic congestion, high rate of accidents, unhealthy environment, etc. As the public transportation industry is tightly integrated with the day-to-day life of majority of the population which constitutes a significant part of it. Therefore, any advancements and improvements in this sector will directly impact people's lives, urban area and almost all other industries (US Environmental Protection Agency, 2020). In the subsequent section of the current thesis, detailed explanation has been made on Public Transit System.

1.2 PUBLIC TRANSIT SYSTEM

Public transportation is a system that will offer travel in the neighbourhood that allows more public to travel together along chosen routes. Buses, trains, and trams are remarkable examples that many people commonly use. High-speed rails, airlines, and coaches govern public transportation in cities. Most public transport systems functions on specified timelines. Some transportation systems function on an occupied capacity basis, which means the vehicles will not start until fully occupied. However, several cities worldwide provide shared cabs when the time is main factor during journey made. Figure 1.2 shows the Progression of Public transit.

The share of bus-based transit systems is more than 90 percent of public transport in Indian cities. Indeed, many Indian cities have no rail transport at all and thus the public depends on the combination of buses, minivans, auto-rickshaws, cycle rickshaws, and taxis for their complete journey. In most of the India's largest cities, a rail-based transit system transmits less than a one-third of the total passengers of public transport. In this case, Mumbai is an exception; this city has the most extensive suburban rail network in India, carrying more than 5 million passengers daily. This was estimated as 58 % of total public transport passengers in the region (vs. 42% by bus) and 80 % of total passenger-km (vs. 20% by bus) (Brihanmumbai Electric Supply and Transport 2003: Indian Railways 2002).

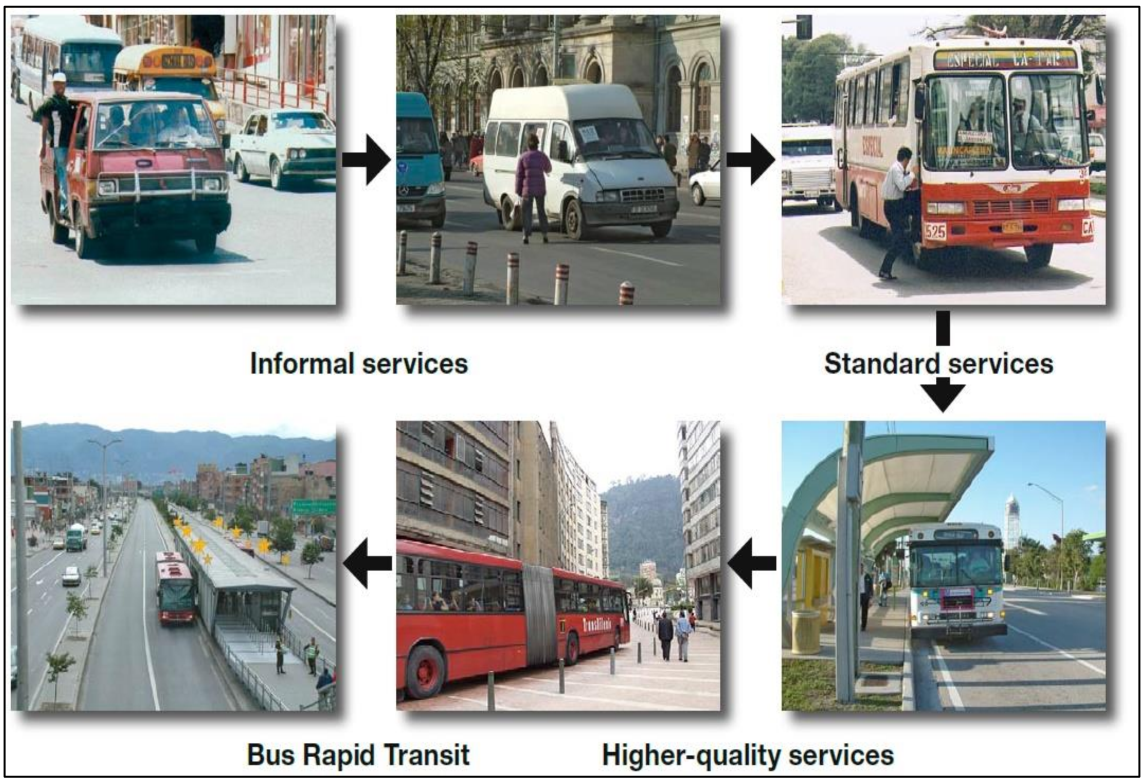


Figure 1.2 Progression of Public transit (Source: Sustainable Transport, 2003)

General statistics in India say, bigger the city size, more is the percentage of urban trips served by the public transport: this can be given as 30 percent in urban areas with a population between 1 and 2 million, 42 percent for urban areas with a population between

2 and 5 million and 63 percent for urban areas with a population over 5 million (Pucher, J., et al. 2004).

Due to increased economic activities, there has recently been an increase in the private vehicle movements in metropolitan cities of India and a rise in average income; in them, Delhi is leading the pack (CSIR-CRRI., 2012). The private vehicle remains to make expansions literally in every city. If this trend continues further, public transport might have a relatively uncertain future. As mentioned in previous paragraphs, due to the rise in the per capita income in developing counties, private vehicles are getting higher percentage usage while public transit systems user rate is almost unanimously declining as Figure 1.3 shows vehicle share in Past Years as per the MoRTH report 2019.

To curl the ongoing trend of private vehicles and look into the environmental issues the Government is trying to bring in many ways which can control the ill effects of the rise of private vehicles activities. The initiatives taken by the Government of India includes a metro system, enabling new road facilities, flyovers, etc. Since the road-based public transport system is the lifeline for easy access and reuse for a wide variety of land uses, the government of India is also trying its best to introduce new techniques for road-based public transport. But, The condition of conventional public transportation in India is the same as that of most other developing countries. Such as a study carried out by the Ministry of Urban Development in 2008 reveals that the stride of developments in public transport systems that are working or planning to commence operation in Indian cities is not able to pace with the rapid and significant rises in demand over the past few decades. Particularly, the bus-based transit systems have deteriorated in their service, and their modal share has been further reduced. It was mainly because the high rate of passengers shifted to the private and intermediate transport system, increasing traffic operations, and having bus operations crippled. In the same time, public transit operators are continuously attempting to overcome these impediments by implementing sophisticated public transit systems with improved infrastructure and management technologies such as Intelligent Transportation System (ITS), Bus Rapid Transit System (BRTS), Metro Rail, Mono Rail, etc.

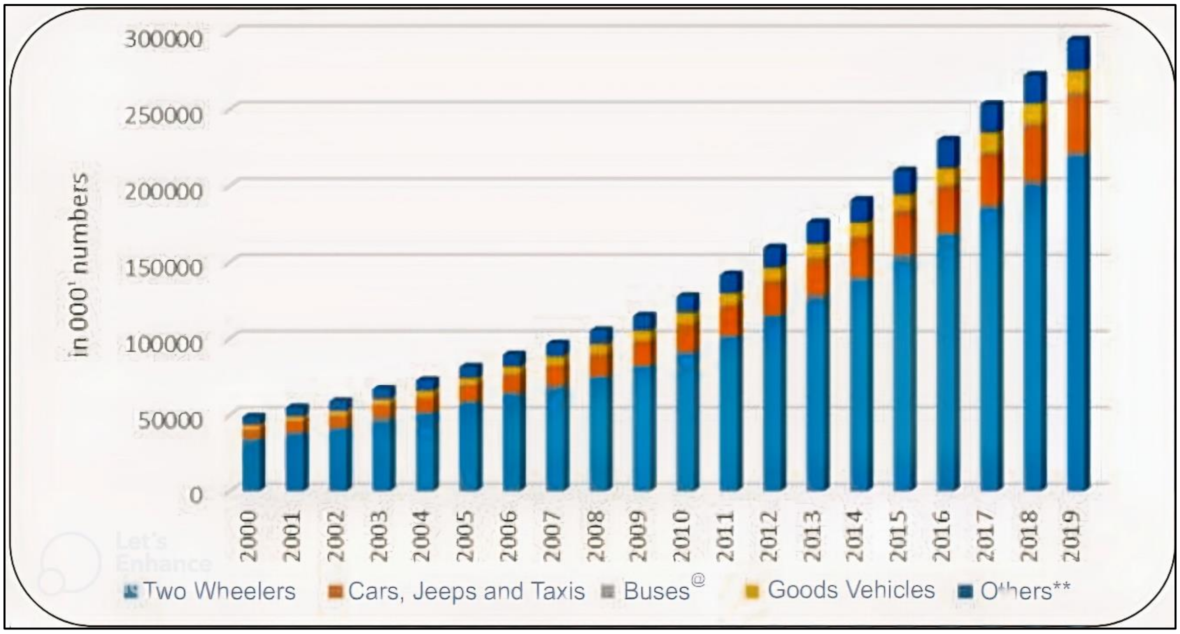


Figure 1.3 Vehicle Share in Past Years (Annual Report MoRTH, 2019)

1.3 BUS RAPID TRANSIT SYSTEM

The need for a sustainable mode of the public transit system in an urban area has been more evidenced by the increased demand for general transportation or travel. A well-organized, planned, and efficiently operating bus-based rapid transit system such as BRTS is an advantage for cities, which are ridden with numerous transportation-related urban issues due to an increased traffic density, population flareup, and the attentiveness of economic activities only in central business districts (CBD), and poor urban planning (Sharma., A, et al. 2015).

Bus Rapid Transit System (BRTS) is an advanced bus-based system with high passenger capacity, operates with lower cost public transport solution that can efficiently improve the issues related to urban mobility. It is also described as a flexible rubber-tiered system that will give speedy service to the people by joining the multiple stations, buses, facility areas, corridors, and ITS-based infrastructures in their operation between origin to the destination. Overall, BRTS transit has a combined structure of various special features with a strong identity and distinctiveness. It was observed that Public Transport Systems working in most of the Indian cities are speedily weakening because of the increasing passenger demand and, their inefficiency to handle those higher demands.

Numerous glitches are related to weak transport systems, such as a tremendous surge in the rate of accidents, unhealthy environment, traffic congestion, overcrowding, reduced frequency of service, and non-following the pre-schedules of the bus operations. Adverse impacts of the inefficient public transit system in urban areas have reached an alarming level for the transit system, and there comes in BRTS to tackle the unabated growth of the population -both people and motor vehicles. It thus becomes an appropriate choice for the purpose demanded, ensuring a clean, efficient, affordable, effective, and safe public transportation system.

Considering the passenger's perspective, BRTS gives improved features, more comfort in the service, frequent service, high accessibility and reliability, a decrease in travel time of the buses, and thus no delays. The capacity of the BRT system is an essential consideration in the operation that impacts the bus frequency, system reliability, and lower travel time. Hence, the success of a BRTS intensely depends on the capacity of the whole system (Sharma, A., et al. 2015).

As per Indian Road Congress guidelines (IRC:124-2017), the essential features of a BRT system in enhancing their overall performance are as follows (Kathuria, A., et al. 2016):

- Dedicated bus lanes
- Fast boarding and alighting facility at every station
- Easy and free transfers between routes
- Fare collection system and validation before actual boarding
- Safety and comfortable service at all bus stations and terminals.
- Clarifications on the route choice with accurate maps, signage, and information provided in real-time.
- GPS-based automatic vehicle location system for tracking bus movements at a lower frequency of time stamps.
- Facilitating the multi-modal integration at bus stations as well as at terminals

- Competitively-bid concessions for operations
- Effective reform of the existing institutional structures for public transit
- Spotless vehicle technologies
- Distinction in marketing with brand name and customer service

1.3.1 Types of BUS Rapid Transit System (BRTS)

As per IRC:124-2017, in Indian conditions, the BRT system is divided on the basis of its corridor type. There are mainly two types of BRT Systems implemented to date:

1. Closed system (Figure1.5)
2. Hybrid system (Figure 1.4)

A closed system has the following features,

- Dedicated busways on most of the system length.
- Convenient place of the transit station and, if possible, median-based busways.
- Best combination of the network with routes and transit feeder lines.
- Bus stops are safe, customers are satisfied, and they are given protected service during all kinds of weather.
- Provided with a pre-board fare collection system.
- Incorporation with the feeder system facility.
- Access to any other kind of bus rather than the prescribed one is restricted.
- Has an individual marketing identity comparable to MRT systems.

The hybrid BRT system has flexibility in features over the closed system. Apart from the mentioned advantages, below are the additional elements of the open BRT system's flexibility.

- System allows currently working bus routes to be contained within the new system.
- Bus stops provided at Kerbside will allow furnishing to the existing routes.
- All the buses are allowed to be served by the system.
- Tickets are issued to the passengers on -boarding the buses.

Figure 1.4 and Figure 1.5 show the BRT buses running on hybrid BRTS conditions (Surat BRTS) and closed BRTS conditions (Bogota BRTS), respectively.

Figure 1.4 Hybrid BRT System – Surat BRTS (Source: Google Images)

Figure 1.5 Closed BRT System – Bogota BRTS (Source: Google Images)

1.4 NEED FOR THE STUDY

- Urban development emphasizes providing innovative and exclusive infrastructure facilities that tackles the supply gaps in the public transit system.
- Improving the quality of service provided, facilitating better approachability to the service, and considering commuters perception on the service performance gets neglected once the public transit system is in place.
- There is no fixed approach or guiding standard that could measure whether transportation facilities given are fulfilling roles or not.
- This mandates the analysis of the service given and its maintenance to the required pace.
- To make the public transit system attractive and thereby increase its ridership, public transit systems need to be planned, operated, and marketed well. Apart from this they need to be measured and monitored continuously– i.e., Performance Analysis.

- Recently functioning public transit systems have been preinstalled with various ITS-based tools such as global positioning system (GPS), automatic fare collection systems (AFCS), etc., which helps in gathering varieties of intelligent data. The operators stores the data for many months to years, and more money is being spent on maintaining the same. Typically, operators use such data to control drivers and make an economic assessment of the system based on passenger demand; however, the application of such ITS-based data is vast. Hence, the availability of such data better utilized to improve the current system's performance and contribute to the overall societal improvements.
- Current research work considers, Hubli-Dharwad Bus Rapid Transit System (HDBRTS), which is working on Integrated Transit Management System (ITMS) as a case study. The primary aim of ITMS is to create an enterprise management system that would allow the company and its host of service providers to manage their activities in a highly coordinated manner leading to a high-productivity environment and reliable services to the users. The system also aims at creating a process-based system that continually allows the operations to be monitored against accepted service levels and provides improvement opportunities to transit managers to offer services at the best operational levels. Hence this has further motivated to frame the objectives of the current research work.

1.5 OBJECTIVES OF THE STUDY

The main aim of this research is to develop a generic approach to assess the performance of BRTS that fulfills the requirements of operators and commuters with the following objectives.

- To investigate the operational characteristics of HDBRTS by assessing the effect of Temporal and Spatial variation of the travel time on the end-to-end Travel Time Variability.
- To examine the impact of comprehensive variables on the Travel Time Reliability

- To analyse the performance of the passenger demand forecasting models with time-series resampled with different time-frames.
- To Develop the Level of Service (LOS) for the routes and segments based on the Travel Time Reliability Indices.
- To recommend strategies to enhance the performance of HDBRTS.

1.6 SCOPE OF THE STUDY

As per the case study considered such HDBRTS and according to the framed objectives, following is the scope of current research work.

Any sort of public transit system, more generally, the transportation industry is tightly integrated with the day-to-day life of the majority of the population, which constitutes a significant part of it. Therefore, any advancements and improvements in this sector will directly impact people's lives and almost all other industries. The transportation industry is one of the major contributing sectors to environmental pollution. Much of this can be attributed to the inefficient transportation system and the rapid increase in the number of vehicles. The increased purchasing power of people is leading to the ever-increasing number of vehicles on the road, which is also creating problems such as congestion. This congestion causes increased travel durations, increased waiting times and increased travel distances, all of which eventually results in an inefficient transportation system. This inefficient transportation leads to excessive consumption of valuable resources. The increased number of vehicles and congestion also causes an increased number of accidents. Hence, a complete overhaul of the whole industry becomes inevitable.

The concept of 'Smart Mobility' is an innovative solution to tackle above-mentioned problems efficiently. Smart mobility connects various elements of technology and mobility, and ITS is a step towards implementing it. The ITS integrates users of the transportation system with vehicles and infrastructure using information and communication technology.

One of the critical aspects of ITS is its robust endorsement of mass transportation's compelling performance. The mass transportation system of the ITS has to be designed considering many factors, passenger demand being an important one. However, with the current passenger demand, the future passenger demand should also be considered during its design and implementation stage. Demand forecast for the public transit system provides a realistic picture of its future usage and is essential for effective policy-making and planning. Thus, passenger demand forecasting is a crucial and inevitable step in keeping the smart mobility system such as BRTS, Metro Rail, Mono Rail, etc, in continuous upgrading mode in their performance.

Bus Rapid Transit System is a state-of-the-art smart mobility system; where an efficient BRTS is a boon for urbanised areas, which are affected by numerous transportation-connected glitches. Its role has now been recognized as essential for physically active, economically sound, and energy-efficient cities. From the transit user's perspective, BRT offers enhanced frequencies, increased system reliability, and reduced travel time and delays. In terms of the transit agency's perspective BRT system complements increased ridership and more fare box collection.

The BRTS has a combined structure of various exclusive features with a strong identity and distinctiveness. A dedicated lane of bus operation is the critical parameter of any BRTS, where buses makes their journey in an exclusive traffic environment and which enhances its performance from multiple perspectives. But the interference of mixed traffic with the operation of BRTS buses along non-dedicated lanes although it only occurs on a few road segments, can compromise the end-to-end travel time of the whole system due to congestion and contribute to reliability-related problems. Many internal and external factors are also influencing the Travel Time Reliability (TTR) and Travel Time Variability (TTV) temporally as well as spatially and finally cause an impact on the whole system's performance. Meantime it is well acknowledged that to make any sort of public transport facilities to be more attractive, and by this means, upsurge their ridership; public transit systems in the urban area not only need to be planned, operated, and marketed well, but they also needs to be measured and monitored on a regular manner.

The main motivation behind this research is to study the impact of such non-dedicated, and dedicated lanes of HDBRTS bus operation on its overall system performance from multiple perspectives by identifying bus stations, routes, and segments that are critical in nature. And utilizing the ITS data generated, such as Automatic Vehicle Location Data (AVL), Automatic Passenger Count Data (APC) Data, QR code Data, etc in a more comprehensive manner to analyse the identified performance perspectives of the system. Then finally, based on the obtained results, making the recommendation to the transit operators as an overall system performance enhancement strategy.

1.7 ORGANISATION OF THE REPORT

Chapter 1 introduces the background of the study. It sets the context with the need for a public transit system in the contemporary urbanisation of any region. An overview of the public transit system, more specifically, features of BRTS, have been discussed. In the subsequent sections of the chapter, various transit performance measures are considered and their importance in analysing the system's performance have been conferred in detail. The framed objectives and the subsequent scope of the present study have been deliberated. Finally, at the end of this chapter, significant contributions to the thesis have been discussed.

Chapter 2 explores the literature review of previous works considering the various public transit performance measures. Different transit performance measures considered for the literature review in this study are travel time reliability, travel variability, travel time reliability modelling, passenger demand forecasting, and transit capacity. At the end of the current section, the literature review is summarized through gap identification.

Chapter 3 elaborates the general overview of the study area considered in the current research by describing the routes and segments details with tabular and pictorial representation and subsequent explanation.

Chapter 4 explains the general overview of the chapter and slides into the source and type of data gathered along with an explanation on its extraction as per the research need. Then it describes the individual methodology adopted for accomplishing all the objectives in the current research work.

Chapter 5 emphasizes on the results and discussions of the current research work. The objective-wise results are shown in the form of tabular and graphical representations. The chapter opens to the results and discussions of travel time reliability of the system through descriptive statistics, travel time variability study through probability distribution of travel time, travel time reliability modelling considering operator's and passengers' perspectives, passenger demand forecasting using ITS-based data of the HDBRTS. Finally closes by establishing reliable transit service-based level of service (LOS) of HDBRTS.

Chapter 6 summarizes and concludes the results obtained. The chapter ends by giving performance-enhancing strategies and recommendations to HDBRTS and by defining or stating the scope of future study.

1.8 MAJOR CONTRIBUTION OF THE THESIS

The major contributions of the current research work are as follows,

- This study has made an attempt to comprehend the robustness of GEV distribution in explaining the heterogeneity in the travel time variability characteristics of the different routes, dedicated and non-dedicated segments.
- Proposed operator's and passengers' perspective travel time reliability models to capture correlations amongst segments' travel time and reliability features with underlying traffic and system states.
- The current study proposed a novel approach of explaining variation in the passenger perspective TTR measures and passenger density with the variation in the parameters of the best-fit distribution at different temporal aggregation levels.
- An attempt was made to explore the unexplored area of LSTM in passenger demand forecasting using APC data, and the results are compared with traditional

forecasting methods. And, there was confusion about the resampling time-interval of the time series to get the best-forecasted results, and studies on addressing such area were limited in the past literature. The current research explored the time series results, resampled with different time intervals, to find the most suitable time interval that gives the best forecasting results.

CHAPTER 2
LITERATURE REVIEW

2.1 GENERAL

Public Transport Performance Analysis can reflect various perspectives. Numerous performance indicators, or actions, will be used to analyse the transit service performance, such as reliability with its travel time data, the variability analysis with travel time data aggregating them to spatially and temporally, the effect of side friction on public transit performance, capacity estimation, speed estimation, etc. In the case of BRTS, these measures can be used to assess variation in the performance considering its operation on dedicated and non-dedicated lanes.

Current research work is intended to analyse performance indicators that work chiefly based on data collected through ITS of Bus Rapid Transit Systems such as Automatic Vehicle Location (AVL) and Automated Passengers Count (APC) data, Dwell Time Data at bus stations, etc. Accordingly, exhaustive literature review on performance analysis of public transit system, transit corridor based on TTV, TTR, TTR modelling, passenger demand forecasting, BRTS capacity and speed assessment, developing of LOS for the public transit systems has been carried out in the separate sections. At the end of the chapter, summary and gaps gathered from the literature review process have been given.

2.2 PERFORMANCE EVALUATION OF PUBLIC TRANSIT SYSTEM

Even though cities provided with high class public transit system, a declining trend was observed in the most of developing city’s modal share of public transit systems. (Institute for Transportation and Development Policy, 2014). It has mainly resulted due to weak in commitment to the service and/or lack of integration of different transport modes, e.g., public transport, feeder system and safe walking and cycling facilities, have forced many public transit system users to use the privatised vehicles as their daily mode of travel.

This shift has been interpreted into augmented traffic congestion, air, noise pollution, deterioration of public spaces, social exclusion, high emission of Greenhouse Gas (GHG), and many other undesirable externalities.

To make public transit transport facilities more attractive, and upsurge their ridership, the system not only needs to be planned, operated, and marketed well, but they also need to be measured and monitored regularly. It is well acknowledged that to make any improvements or changes and accomplish vital service from the system; one must first be able to measure and quantify it, that which in totality highlights there has to be as Performance of Public Transit System (Institute for Transportation and Development Policy, 2014).

Performance measures are beyond our senses – sight, hearing, touch, smell, and taste. Performance measures taken in the analysis act as direction-finding tools that aid a system operator in finding the path like, where it wants to go to, and how to get there. The performance measures can solve many issues in a practical sense, such as problem documentation, trend investigation, peer assessments, target setting, and finally, taking required steps for potential improvements in the system.

Evaluating transit performance, considering transit service reliability from different perspectives is one of the critical objectives for transit operators and policymakers. However, there are many perspectives through which many reliability measures can be checked through performance analysis. Like, commuters, operators, and drivers have diverse interests concerning systems operation, thus selecting a particular perspective's reliability measure is the main challenge.

The performance of the public transit system is evaluated on the basis of various perspectives, and many operators and researchers usually assess the public transit performance with indices such as load factor at the stations and cost-per-vehicle-kilo-meter measure operating efficiency. Some other indices used for analysis are commuters' comfort, speed of travel, reliability in the travel times, headway adherence, affordability, integration, and fulfilment, reflecting on the user experience. Passenger or commuters' perspective indices are the most significant in improving public transit systems performance that responds to passenger demands. Hence, transit systems can attract even more modal shifts from private vehicle users. Hence this area requires thoughtful attention from the operators, policymakers, and researchers' point of view in most developing cities today.

The transit industry uses more than 400 different categories, or perspective-based performance measures available in the transit industry today. Performance measures are given under the head of twelve goals and four population ranges of the area where the system is operating (Section 6, TCRP REPORT 88, 2003). As stated previously, every indicator is assessed based on its performance assessment, such as availability of buses, condition of service delivery, impact on the community, travel time of buses, the guarantee of safety and security at stations, terminals or buses, maintenance and construction, and economic/financial viability. Performance indicators also depend upon the type of the data, data collection methodology, analysis, and its potential strengths and weaknesses for specific applications (TCRP REPORT 26, 1997).

As per TCQSM 2013, based on operator's prospective performance of public transit can be assessed by considering,

- Reliability refers to how well the transit service can keep to its schedule.
- Speed (or, more accurately, travel time) refers to how fast the vehicles can operate.
- Capacity refers to the maximum number of transit vehicles, persons, or both that can travel past a particular location in a given period under specified conditions.

Based on user prospective performance of public transit can be assessed by considering, Customer satisfaction research depends on comfort and convenience measures from the transit service. Figure 2.1 shows the Public Transit stakeholder's interest areas and performance measure as given by the TCQMS, 2013.

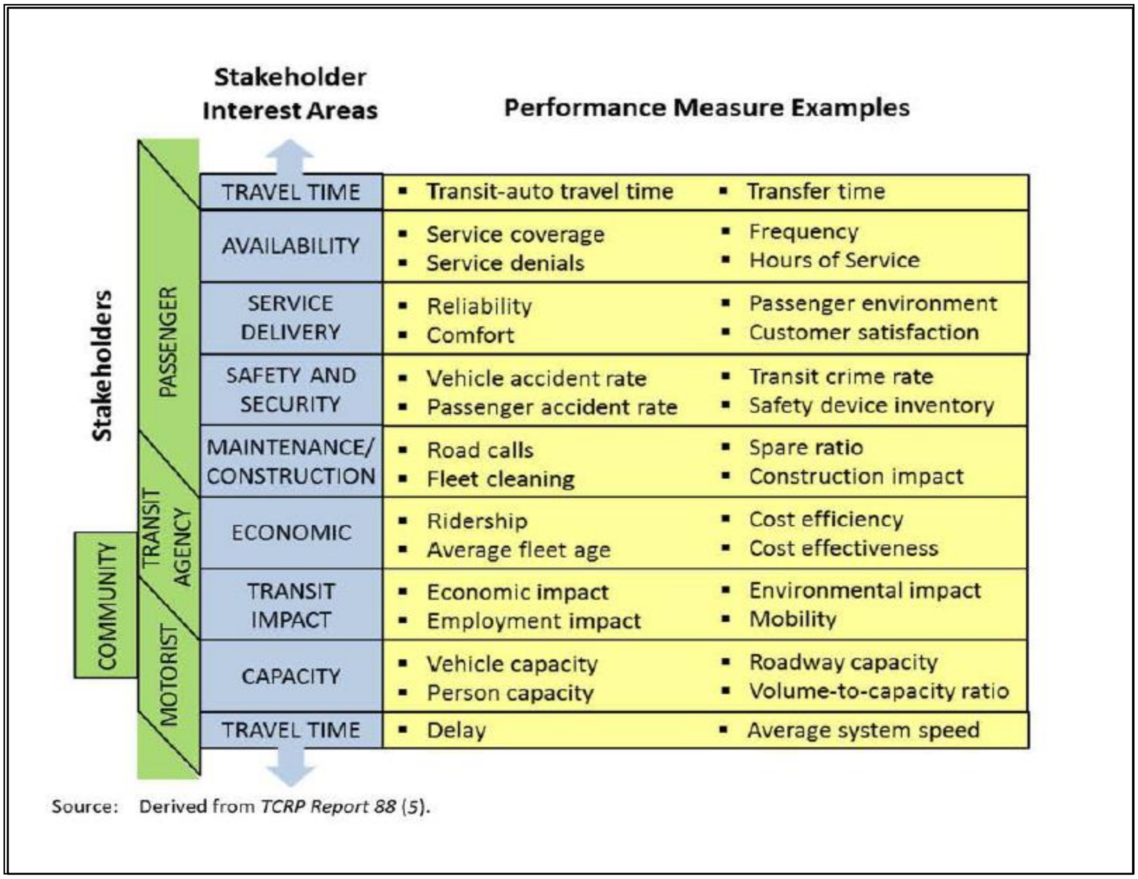


Figure 2.1 Public Transit Stakeholder's Interest Areas and Performance Measure (Source: TCQMS 2013)

2.3 TRAVEL TIME VARIABILITY (TTV), TRAVEL TIME RELIABILITY (TTR) MODELLING, AND LEVEL OF SERVICE (LOS) OF THE TRANSIT SYSTEM

Travel time reliability is the consistency of the transit system under different situations over time. It is also defined as the degree of steadiness of the system's quality that it usually can offer. Travel time variability concerns changes in the transit system's travel time during any time of day or changes from day to day. TTV mainly consists of free flow travel time and delay as its components.

Evaluating the transit performance based on reliability measures becomes vital from the passenger's point of view. Reliability indicators mainly gauge the commuters arriving at their destination on time and not having to wait too long at the bus station until the arrival of one's transit bus. In the view of operators, reliability measures mainly influence the schedule recovery factor of the total cycle time of the identified route. Thus, it cause higher operating costs when recovery time requires additional buses to run a route at a specified rate. Unreliability in transit operations results from more buses accumulating at the bus station at the same time, generally termed as bus bunching at the stations; with this, more travellers experience crowded onboard conditions.

The transit service reliability has been identified as an essential component of the system's performance. Liu, Ronghui., et al. (2007) deliberated three essential types of bus reliability measuring measures; such as Travel Time Variability, Travel Time Reliability, Waiting Time Reliability, and Headway Regularity of the buses in sequence.

BRT systems implemented in Indian conditions have median-based bus lanes that will distinguish the BRTS bus operations from the mixed traffic condition. With the literature background, dedicated BRTS bus operations will be well-thought-out to be 90% on schedule (Deng, T., et al., 2013), practically more than the conventional bus transport system. The hybrid or composite BRT systems are generally seen in India, where buses will operate on partially dedicated lanes up to some length and then merge their operation into mixed traffic conditions due to limited right-of-way (ROW) or raised stretch. This swift from dedicated to diverse traffic conditions reduces the system's overall reliability (Kathuria, A., et al. 2016). Out of the three aforementioned reliability measures, TTR and headway regularity use the ITS-based smart data gathered by the GPS installed in the individual buses.

Necessary TTR measures that are generally used in the transit industry are as shown in the next section:

2.3.1 Travel Time-based Reliability Measures Used in Transit Industry

TTR is the primary function of travel time variability (TTV) (Tu, H., et al. 2008). When the TTV of the transit is measured, it also gives an idea of transit TTR. Sekhar, C., et al. (2007) studied many reliability measures built on systems travel time data sets. This study also focuses on the fact that literature shows the reliability measures are mostly done on the average travel tome (ATT) or on the central tendency of the data set and working on the statistical distribution of travel time cases. Federal Highway Administration (2006) has also given many travel time reliability indices to enumerate the performance of the transportation systems. Consequently, the following paragraphs have discussed various statistical range and reliability measures used to assess the transit travel time reliability.

90th or 95th Percentile Travel Time: One of the most straightforward reliability measures is 90th or 95th percentile travel times for specific travel routes or trips. This reliability index clearly shows how bad delays will be on the heaviest travel days. The 90th or 95th percentile travel times are stated in terms of minutes and seconds and should be easily understood by travellers familiar with their trips. For this reason, this measure is perfectly suitable for traveller information.

Buffer Time Index (BTI): This index signifies the additional buffer time (or time cushion) maximum travellers add to their average travel time when planning trips to guarantee on-time arrival. This additional time also accounts for any unforeseen delay. This index is stated as a percentage, and its value increase as reliability gets inferior. The buffer index is calculated as the difference between the 95th percentile travel time and average travel time, divided by the average travel time.

Planning Time Index (PTI): The planning time index characterizes the entire travel time that should be prepared when an adequate buffer time is comprised. This index differs from the buffer index because it includes typical and unexpected delays. Hence, the planning time index likens near-worst case travel time to travel time in light or free-flow traffic flow conditions. The planning time index is beneficial as it can be unswervingly associated with the travel time index (a measure of average congestion) on similar numeric scales. The planning time index is calculated as the 95th percentile travel time divided by the free-flow travel time.

The Travel Time Index (TTI) is another crucial reliability measure that helps monitor the average congestion. The travel time index represents the average additional time required during peak times compared to light traffic times.

Many statistical indicators are available to enumerate the variation in the transit travel time, such as standard deviation of the travel time and coefficient of variation of the travel time are very commonly used measures. Whereas these indicators are more mathematical-based performance measures that are calculated based on the central tendency of the data set, sometimes they are not easily understood by non-technical audiences nor easily related to everyday commuting experiences.

Hence, principally to understand the dissimilarity in transit travel time following measures are preferred by many researchers and academicians.

- **T90–T10:** This index is calculated by taking the difference between the 90th percentile and the 10th percentile value of travel time. These values indicate the dispersion of the distribution. If this value is high, TTV gets high and hence low is the overall systems reliability.

- **Coefficient of variation (CV):** Is one of the most straightforward statistical tools to quantify the spread in the different data sets. This is calculated by taking the ratio of standard deviation and mean. Here, the CV value indicate the travel time data set distribution in this case.

- **λ_{var}:** It is calculated as the ratio of 90th percentile minus 10th percentile divided by 50th percentile of travel time. This gives an idea about the apparent width of the distribution concerning its median.

- **λ_{skew}**: One statistical measure to get an idea of the skew width of travel time variation (TTV) is the 90th percentile minus 50th percentile divided by 50th percentile minus 10th percentile.

2.3.2 Factors Affecting Travel Time Reliability

Various sources induce variability in travel time, and these sources are grouped into the following categories. (Kwon, et al. 2011) and also Figure 2.2 shows common factors influencing transit reliability.

• **Traffic incidents:** Traffic incidents are nothing but events occurring on the roads, such as crashes and other unplanned incidents which cause obstructions to the flow of traffic. Due to these incidents, there arises variability in travel time.

• **Work zone activity:** Sometimes, there are chances of construction work going on, on the part of the roadway; these repairs and other maintenance work reduce the actual width of the road and become a source of producing factors affecting the variability in travel time.

• **Weather:** The speed of vehicles is affected by the changes and adversity in weather conditions such as rain, fog, smog, snowfall, etc. Hence, these changes in environmental conditions become a cause of time delays.

• **Fluctuations in demand**: It is well known that the traffic demand is not constant, and thus varies day-to-day, monthly, and seasonally; when this variability increased and does not get managed correctly, it gives rise to travel time variability.

• Special events: Events such as rallies, strikes, and festivals cause changes in the traffic demand and act as temporary blockages of roadways. The changes due to those special reasons also induce travel time variability.

• Traffic-control devices: Control devices such as railroad grade crossings and poorly timed signals disrupt traffic flow and thus create variability in travel times.

• Inadequate base capacity: Roadways are designed for a specific capacity, and many times due to many factors such as inadequate shoulder width, on-street parking, movement of pedestrians on the roadway, presence of street vendors, curb side bus stops, etc. This reduces the capacity of the road and results in travel time variation.

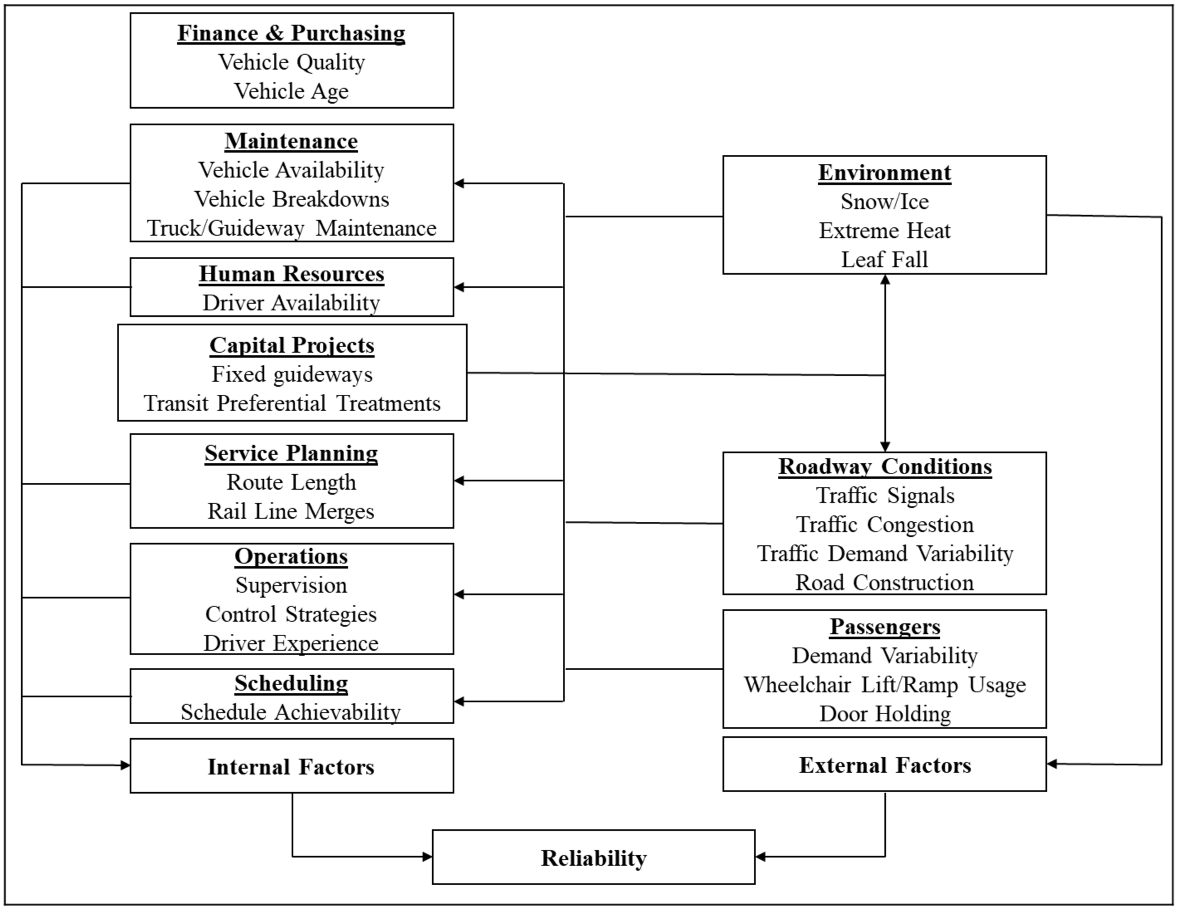


Figure 2.2 Factors Influencing Transit Reliability (Source: TCQSM 2013)

Some of the important studies carried out by researchers on TTV, TTR and TTR modelling are given below.

Camus, R., et al. (2005), have studied the transit quality of service and reliability using AVL data from four routes of the Trieste transit network in Italy.

The author has discussed many advantages and limitations of the TCQSM method of developing level-of-service (LOS). The weighted delay index (WDI) given by the author in this study overcomes the confines of the TCQSM methodology. A straightforward procedure for the estimation of WDI has been given. New LOS ranges and thresholds are used as a system reliability measurement tool.

Obtained results from the study have also been compared with the available data set. It indicates that the given methodology leads to a more graduated LOS estimation because of the higher number of parameters acquainted with the new transit service measure.

Li, R., et al. (2006) have conducted a study on a wide range analysis of the travel time distributions in terms of various temporal aggregations. Data has been obtained from extensive automatic vehicle identification data of City Link Toll way in Melbourne, Australia. The study focuses on two groups of data - travel times obtained from morning peak and afternoon peak that proves that they have distinctive sources of variation.

The study also inspects the components of travel time variability and discovers their associations. Statistical and multiple regression analysis adopted quantifies various factors affecting travel time variability and their interaction effects.

Results of peak travel times in the morning hours vary primarily because of driver-related factors, precisely lane choice decisions; however, 25% of the variation in the travel times in the afternoon peak corresponds to the system's capacity-related characteristics.

Mazloumi, E., et al. (2009), consider variation in the travel time of Melbourne’s public transport system on a day-to-day basis. Variability of travel time has been evaluated using GPS data obtained from the bus’s operations.

The study mainly uses linear regression to identify the factors causing travel time variability.

Obtained results shown that, if smaller the departure time windows, then normal distribution best characterises the travel time. In another case, lognormal or normal distributions better explains the morning peak traffic. From the linear regression modelling, the factors contributing the variation in the TT found in the study are land use, route length, number of traffic signals, number of bus stops, and departure delay relative to the scheduled departure time.

Schramm, L., et al. (2010) discuss various improved features of the BRT system and their effect on the travel time variability between peak and non-peak hours.

To understand the variations in scheduled end-to-end travel time, a ratio of average peak-hour travel time to average non-peak-hour travel time is compared across BRT systems considered in the study. 19 BRT systems in the United States have been taken for the survey and compared the variation in the TT based on seven identified features viz. condition of running ways of the buses, passing capability of the methods, bus station space, providing transit signal priority for the buses, frequency of buses, level boarding service, and the fare collection system.

Almost all the identified features of BRTS in the study have been found to improve the reliability and variability in the travel time of the buses. But features associated more closely with the bus stations of the BRT system are found to have less implication on the reliability of the service. The author has stated that BRT stations should be located on the far side of intersections to gain complete advantage of the signal priority system.

Susilawati, S., et al. (2010), assessed the reliability of Adelaide metropolitan road networks' travel time of ten critical corridors. The author has used buffer time and planning time indices as two measures to analyze the reliability of the selected corridors.

From the study, it is interpreted that obtained buffer time indices overestimate travel time reliability measurement; hence detailed assessment of travel time distribution is carried out in the study.

The results show that Probability Density Function (PDF) of the considered travel time distribution not follows the normal distribution pattern. Instead, for some corridors, the log-normal distribution is better suitable to explain the pattern. Authors have also suggested to try some other continuous distributions for future work.

El-Geneidy., et al. (2011) have used ITS-based data from Metro Transit in Minnesota. Types of data used are AVL and APC in 2005 to analyze the system's performance issues along the cross-town route.

Series of visual and analytical examinations to predict the running time in the study, schedule adherence, and reliability of the transit route at two scales have been focused. The time point is obtained from the segment and the route level demonstrates how to identify the causes of the deterioration in the system's reliability. The developed analytical models in the study show that while headways are maintained, schedule revisions are needed to improve run-time along with proper schedule adherence of the trips.

In conclusion, the study suggests that many scheduled stops along the selected route are underutilized and endorses stop consolidation as a tool to decrease the service variation by concentrating passenger demand along few stops.

Bharti, A. K., et al. (2013), planning Time Index and Buffer Time Index suggested by FHWA are used to analyze the performance of the urban arterial corridor in Delhi. Selected reliability measures in the study are linked to the volume-capacity ratio of the study section.

The results from the analysis show that the PTI and BTI takes higher values (2.63 and0.71) from 8.45 AM – 9.00 AM. Meantime, PTI is from 5.15 PM-5.30 PM hours; accordingly, the mean 95th % travel time for urban corridors varies. It is also suggested to investigate further the finding of the Level of Service (LOS) of the selected arterial corridor using Travel Time Reliability measures as future scope.

A study by **Huo, Y., et al. (2014),** conducted a bus service reliability study of Changzhou BRT. Travel time data gathered from the identified system and various value ranges of measures are considered, temporal and spatial distributions of travel time data points are made, and a comparison is established.

The number of times passengers need to wait, on average, and extra time passengers need to budget beyond typical wait time and journey time are some of the measures considered in the analysis by the authors. It is found that transit passengers regularly add extra time for bus waiting and for the entire journey more than typical waiting and journey times to guarantee arrival at their destinations on time at the highest likelihood. Specific suggestions are drawn for improving the bus service reliability of considered BRT, like; enhancing stop accessibility and educating passengers to board in an orderly manner, taking traffic control actions especially for peak hours journeys, such as police supervision of traffic to improve reliability and increasing service frequency at some locations.

Uno, N., et al. (2014) used bus probe data to evaluate the road network regarding travel time stability and reliability. For the study purpose, Hirakata City, Japan, has been selected as a case study.

The methodology has been proposed to estimate travel time distributions of arbitrary routes by statistically adding up the directly observed multiple travel time distributions. Framed methods have estimated travel time distributions of random routes covered by the BPS, they have also given an approach to evaluate the LOS of a road network based on the concept of travel time reliability. Two travel time reliability measures have been used in the study to establish the LOS. The average travel time for one kilometre is one of the indexes or measures used to assess the competence of the network. Then COV of the travel time is another measure used to evaluate the network's reliability. The cumulative distributions of considered measures are finally taken to propose the LOS of a network in the case study.

Godavarthi, G. R., et al. (2014), have assessed the capacity of Mixed Traffic (MT) lanes and BRTS lanes in Delhi and Ahmedabad based on calculating the V/C ratio of both lanes separately. Extensive data sets in the study are generated from VISSIM software.

The author has also performed the traffic volume survey (16-h classifies traffic vol.) and the speed-delay survey (Probe vehicle method) along the selected lanes. Roadway capacity is mainly calculated to understand the performance of the BRTS for MT and bus lanes.

The author has found a 0.688 V/C ratio from the study, which is the optimal flow value for BRT corridors. This indicated that up to 0.688, the mixed traffic lane users and bus lane users are appreciating reasonable travel speeds and more minor delays. In the case where the V/C ratio is exceeded mentioned value on either the BRT lane or mixed traffic lane(s), then it leads the BRT system to become untenable for the mixed traffic lane and BRT users, creating more traffic congestion-related issues.

Das, Shreya., et al. (2015), have attempted the to develop a level of service of the bus transit system operating in Kolkata city. User perception-based data has been gathered from the 'Law of Successive Interval Scaling' method. Obtained results are compared with LOS scales on the same perception developed in the TCQMS and MoUD, India, found that it is not logically correct to create the same LOS benchmarks for assessing transit performance across different cities. Development of LOS is purely dependent on the city's socioeconomic characteristics, land use patterns, traffic character's, etc. This study's successive interval scaling method proved to be the most suitable tool to solve the difficulties in gathering aggregate user rating data, mainly when an ordered categorical scale is used to collect user perception of transit service quality.

Kieu, L. M., et al. (2015), have worked on the distribution of Brisbane public transit travel times on its main corridor. Authors have tried all types of continuous distributions to understand the shape and nature of public transit travel times data sets. Results show that log-normal distribution is the best descriptor of bus travel time on urban corridors based on the Kolmogorov-Smirnov test and Bayesian information criterion technique. The current study gives the guiding principles in monitoring the public transit travel time variation, recovery time optimization, and statistical analysis of public transport travel time.

Ma, Z. L., et al. (2015) used AVL and AFCS data from Translink, Australia, and measured travel time reliability at the link level. Seemingly Unrelated Regression Equations (SURE) are used in the modelling process. Planning, operational, and related environmental factors are the three main types of unreliability contributing factors used very efficiently in their study.

The developed models give insights into these contributory factors that directly impact the bus travel time and cause variability. The most significant factors are the recurrent Congestion Index, traffic signals data, and passenger demand-related data at all stations. Finally, results are used effectively to address the reliability enhancement strategies aimed at reducing unreliability on multiple types of routes of the selected case study.

Chepuri, A., et al. (2015), have assessed the traffic flow characteristics obtained on the 1.8 kilometres of BRTS corridor in Surat city. The selected corridor of BRTS also has four intersections along its length.

Delays caused at those four intersections due to regular traffic are evaluated in the study using microscopic simulation software VISSIM 7.0. The travel time of the buses is obtained through performance-box GPS equipment, which gives real-time speed every 0.1 seconds. Traffic surveys are mainly carried out to collect data on road inventory, classified volume count (Videography method), rate and delay (V-box device), and spot speeds (Radar gun). From all the planned surveys in the study, a speed and delay survey is used for model building and validation in the VISSIM software.

Finally, it is found that on many roads stretches; the calculated V/C ratio was significantly less, and at many places, it was high; hence they tried to balance it by diverting some traffic from a higher to a lower V/C ratio corridor using VISSIM software.

Gunawan, F. E., (2015), have computed the consequences caused due to mixed traffic interference on the BRTS bus travel times. The author has used an empirical approach to measure the 11 corridors of TransJakarta BRT functions in the city of Jakarta, the capital of the Republic of Indonesia was carefully chosen as a case study. They have also recorded the travel time data from the station to station along the selected corridors for the case where the bus can travel smoothly without any interference and for the point where the bus journey was interfered with the mixed traffic and assessed this travel time variability.

Obtained results from the study show that Corridor 1 of the TransJakarta BRT has the best performance in terms of the travel time variation out of all selected corridors.

Kathuria, A., et al. (2016), have reviewed nine Indian BRT systems based on their design and operational characteristics in this study. For each selected BRTS, the primary active summary has been prepared along with their detailed features considering the regulatory context, cost models adopted for selected systems is shown in the form of systematic tables. Then, the carriageway concepts adopted, selection of the carriageway based upon the availability of ROW, and basic operation parameters have been assessed individually and compared. Finally, the system reliability review methodology is given by the author and is solely dependent on the travel time reliability (TTR). Obtaining the GPS-based travel time data from all the selected systems was difficult; hence, reliability evaluation is made only for the Ahmedabad BRTS.

Yan, Y., et al. (2016), have proposed and evaluate the performance of bus routes using systematic AVL-based data. Different percentile of travel times, coefficient of variation (COV) of travel times data points, and average commercial speed of traffic are some of the statistical measures considered along with the travel time distribution study. The distribution of the travel times is assessed spatially and temporally, and the impact of transit regulation indexes on the travel time variation analyzed effectively. Then the bus route with transit signal priority and a dedicated bus lane in Suzhou, China, is taken as a case study and validated the proposed methodology for the considered system.

The obtained result shows that the maximum influential travel time feature is their spatial and temporal aggregations, varying throughout the segments and all the time-of-day intervals. Bus lane violation and route repetition are the two significant parameters undermining the effectiveness of priority measures. The pre-schedule design of the system’s operation has an imperative impact on schedule adherence and headway uniformity in the bus’s operation.

Biswas, S., et al. (2016), have proposed a new methodology to evaluate the LOS of urban arterials. Percentage reduction in the speed (PSR) is obtained by comparing the actual speed to the free flow speed of the arterial roads (FFS). This has been recognized as an alternative performance measure for LOS assessment. Sixteen hours of traffic volume and rate data have been gathered through videography recorded at some important road segments of a six-lane divided urban arterial in the Kolkata metropolis.

FFS obtained from the individual categories of the vehicles are examined with the normal distribution curves tried on the speed data under free-flowing conditions. Kolmogorov-

Smirnov (K-S) test is used as a performance checking tool for the goodness-of-fit of those modelled curves, and they showed good compatibility with the experiential data. The k-mean clustering method is used to classify the observed PSR data into subgroups; the Silhouette method is used to validate the resulting clusters. Finally, six LOS classes circumscribed by threshold values of PSR have been proposed in the study.

Ma, Z., et al. (2017), have used the quantile regression method of model development to examine the effects of the fundamental factors on the travel time distribution characteristics instead of their central tendency values. AVL and AFCS-based supply and demand data are gathered from Translink, Brisbane, Australia, to carry out the intended study.

Results obtained from this study have revealed that the quantile regression model provides more indicative evidence than the conditional mean regression method. The findings give information on the impacts of planning, operational, and environmental factors on speed and variability in their values. Based on the obtained data, it is also suggested that transit designers and planners can design embattled strategies to efficiently and effectively improve travel time reliability.

Kathuria, A., et al. (2017), have considered transit operation and transit regulations as two perspectives concerning operators' and passengers' points of view for evaluating the operational performance of Ahmedabad BRTS. GPS-based automated vehicle location data of buses is effectively used in the evaluation approach adopted in the study. Operators-based transit operation measures considered percentile travel time, coefficient of variation (COV) of travel time, average journey speed, and travel time distribution. Whereas in the case of passenger-based transit regulation, schedule adherence and headway regularity were some of the measures in the analysis. Segment level analysis is mainly carried out to check variation in the bus travel time amongst all the segments taken in the study. As in the case of BRTS corridors, segments are partly dedicated and partly non-dedicated. Comparisons of the travel-time reliability-based performance of a BRT and a non-BRT route have been carried out by the authors. Subsequently, a thorough root level analysis has been done. Finally, LOS analysis is carried out based on two measures, average travel time per kilometre and travel-time coefficient of variation for a network level, to understand how the LOS of the network has changed from 2013 to 2016 with changing corridor length from 61km to 89km.

Bhana, P., et al. (2017), has evaluated the impact of BRT lane on MT lane with the help of traffic data like Classified Traffic Volume Counts at Mid-Block Section (four mid-block sections along the route of BRT), Spot Speed Study (segment of 20m length on MT lane and BRT lane), and Queue Length Survey (four selected intersections falling on the corridor). His study reveal that the MT lane's performance is affected by the implementation of BRTS. The traffic flow along the MT lane in each direction of travel in almost all the mid-block sections reaches its total capacity in the 60 m right of way. The queue lengths are high just because of low saturation and more normal flow along with the corridor approaches.

Jairam R., et al. (2018), has worked on developing a reliable structure for predicting public transit systems' travel times or arrival times under mixed traffic conditions that generally exist in Indian conditions. The required data in the study has been obtained from three important cities in India: Surat, Mysore, and Chennai. The obtained data is assessed considering various spatial and temporal patterns, and the same extracted data points are further employed in the prediction models developed. The performance of developed models has been examined with a k-NN classifier, Kalman Filter, and Auto-Regressive Integrated Moving Average (ARIMA) techniques (Based on ML-ANN).

In the case of Surat BRTS data, all models are performed the same. It is mainly due to less influence of BRTS operation by external traffic except intersections, whereas in the case of the bus lane gets exposed to the mixed traffic. But, in the case of mixed lane traffic, kNN+KFT performed better than the other models.

Chepuri, A., et al. (2018), have carried-out analysis in the variation in the bus travel times and their reliability using GPS-based travel time data. The author has taken intelligent data along a selected bus route in the city of Chennai in the southern part of India. Various reliability indices of travel time are used for analysis: buffer time index (BTI), planning time index (PTI), and a new reliability measure. Statistical indices such as ATT, CV of TT, SD of TT etc, are also used over different time frames. Variations in the travel times are studied with the various continuous distributions, and the generalized extreme value (GEV) distribution is found to be the best-fitted among others. The distributions' performance is checked with Kolmogorov-Smirnov (KS) test. Distribution results show that GEV explains the bus travel time variability sensibly well.

Then the study attempts to establish the procedure for developing a level-of-service (LOS) using reliability indicators. Segment-level travel time data, travel time coefficient of variation (COV), and volume-to-capacity ratio (V/C) are used to develop the LOS through clustering technique. Finally, the study concludes that 95^{th} percentile travel time and buffer time are the most effective performance indicators for examining travel time variability in the selected case study.

Bharti, A. K., et al. (2018), has conducted a systematic travel time reliability study on Delhi's urban arterial and intercity highway corridors. The automatic vehicle license plate matching method using IP-based video graphic data is the source of the data taken for estimating travel time and subsequent analysis of the transit reliability. The study considers various reliability measures for travel time reliability analysis; such as planning time (PT), buffer time (BT), planning time index (PTI), and buffer time index (BTI). Then LOS for selected corridor developed by establishing correlation of reliability measures with the obtained volume-to-capacity ratios. The LOS ranges are defined with the help of the K-mean clustering technique performed using MATLAB.

Obtained results indicate many insights, like; for obtained LOS B, the travel times of intercity highways are in the range of 40–46 sec/km, while it is 64–80 sec/km for urban corridors without interruption, and for the interrupted corridors, it is found to be in the range of 75–135 sec/km respectively. Reliability calculations such as PT, PTI, and BTI are also different for different volume-capacity ratios for identified corridors.

Chepuri, A., et al. (2019), have worked on the travel time reliability analysis of the public transit system of Mysore city. The author has obtained GPS-based AVL data for analysis. Initially travel time reliability of the selected system is examined spatially and temporally. Then relationships are established between speed, flow, and density, which helped develop the correlation models between journey speed, and stream speed and study the discrepancy in BTI and PTI concerning calculated V/C. Finally, the LOS of the selected system is developed concerning V/C and COV.

The identified bus route is divided into segments with 2 and 4 lanes. Then reliability indices ATT, BT, PT, PTI, BTI, and TTI are calculated for all the segments and examined for different hours of the day for two sections in min/km. A linear regression model is developed to establish a correlation between journey speed and V/C. Mean Absolute Percentage Error (MAPE) is used to validate the developed models between the derived and observed flow. Using k-mean clustering method is used to create LOS. The study concludes that BT might be lower and BTI may be higher during peak hours, mainly due to higher ATTs and 95th percentile travel times. Higher R^2 values of the developed model indicate that established relationships between journey speed and flow performed better. The study mentions that when the V/C value is greater than 0.5, slight variations in V/C will cause considerable variations in BT and PTI.

Chepuri, A., et al. (2020) have used sole reliability measure using GPS-based data gathered from certain bus routes in Mysuru city in southern India and aims to understand the variation in travel time and assess the reliability in travel time over different times and spaces. The data collected is processed for noise elimination, travel time calculation for the hour of the day, period of the day, day of the week, and bus stop wise. The AVL data is analysed for various reliability measures taken from FHWA 2006 viz. BTI, PTI, RBI, and descriptive statistics calculation at the Route and Segment level.

Travel time variability is also examined using GEV distribution. Obtained results show TTI, PTI, and RBI were decreasing as the trip distance is increasing, and shorter routes seemed to be unreliable in the study; amongst all the distributions tried, GEV distribution best descriptor of the travel time data and weekend and Monday are having higher travel time variation compared to other week days. The study has also suggested using the new performance measure (RBI) instead of assessing the performance with various measures. Finally, study recommends to develop the RBI-based LOS thresholds as the future scope of the study.

Kathuria, A., et al. (2020) have used GPS-based travel time data to study the travel time variability of the Ahmedabad Bus Rapid Transit System. The study's main objective is to investigate TTV from operators' perspectives using different reliability measures of travel time. The route, segment, and network-level analysis is carried out separately. Models are proposed to identify factors causing TTV and develop LOS threshold for the Network of BRTS in Ahmedabad, Gujarat state India.
Descriptive statistics such as AT, SD of TT, COV, T90-T50/T50 (%), PTI, average speed and are used for route level and segment level analysis. Travel time data points are also analyzed using statistical distributions and K-S hypothesis test is used to test the distribution models' goodness of fit. TTR based regression models are established and performance perceived through adjusted R2 value. Using the K-mean clustering technique, LOS is developed based on WDI and Headway adherence (COV of headway). Author has concluded that, TTR regression model with T90–T10 has the higher descriptive power as dependent variable compared to one more model with SD. The developed model's maximum impact on the travel time variation was observed from the length, number of intersections, and percentage segregated route as independent variables. Finally, it has recommended to provide Transit Signal Priority at intersections, making bus stop consolidation, increasing express buses services, and drivers to adherence to the timetable to reduces the TTV and improves TTR.

Harsha, M. M., et al. (2020), have chosen public transit of Mysore city as a case study for analysing travel time variation. In this study, the author has comprehensively analyzed the variability of travel times using travel time distribution. The procedure is implemented to accomplish the objective using normal, lognormal, logistic, log-logistic, Gamma, Weibull, Burr, and GEV distributions for the route and segment level (bus stop-wise and intersection-wise). Based on K-S goodness of fit test results all the distribution are ranked to assess their performance. For the chosen travel time cases, the top three distributions are identified and GEV distribution is the most suitable for explaining both cases' heterogeneous travel times with different traffic conditions.

Harsha, M. M., et al. (2021) have used a probability distribution to explain the travel time variability. AVL data has been obtained from four public transit routes of Mysore City in Karnataka. The author tries seven probability distributions in the analysis, namely Burr, GEV, Gamma, log-logistic, lognormal, normal, and Weibull. Distribution analysis is carried out for different temporal and spatial aggregation levels. The author has used the Anderson-Darling (AD) goodness of fit test to examine the performance of all the distributions tried in the analysis. It is found that GEV distribution best explains the variability of travel time and stood as the most suitable distribution out of all other distributions. The study have concluded that probability distributions best describe TT data points' characteristics.

2.4 ROLE OF PASSENGER DEMAND FORECASTING IN ENHANCING TRANSIT PERFORMANCE

As discussed in previous sections, the public transportation system is an apt solution for increased congestion in urban areas, excessive consumption of natural resources, hiked fuel and vehicle prices, unnecessary rise in trip distances and durations, and overall sustainable development in urban areas.

But the mere implementation of the public transportation system does not guarantee the results mentioned above. Unless, it is planned and implemented precisely, with sustainable and smart system. It must employ as few resources as possible, analysing the performance by considering many measures regularly. This concern has resulted in the proposal of a new concept, which is the refined form of public transportation system called ‘Smart Mobility’ (Benevolo, C., et al., 2016).

The concept of ‘Smart Mobility’ is an innovative solution to tackle this problem. Smart mobility connects various elements of technology and mobility, and Intelligent Transportation System (ITS) is a step towards implementing it. The ITS integrates users of the transportation system with vehicles and infrastructure using information and communication technology. The technology used can vary from a basic management system to more advanced applications that integrate real-time data and feedback from a number of other sources. Additionally, predictive techniques can also be used for advanced modelling, forecasting and comparison with historical baseline data (Satheesh Kumar, M., et al., 2014). ITS also helps in decreasing the travel durations and reducing congestion. With reduced congestion, the capacity of narrow roads will also be increased. One of the critical aspects of ITS is its strong endorsement of mass transportation.

The mass transportation system of ITS has to be designed considering many factors, passenger demand forecasting being an important one. However, with the current passenger demand, the future passenger demand should also be considered while designing through its forecasting to subsequent time horizon. Passenger demand forecasting is the term used to predict the future as accurately as possible and it provides practical picture of its future usage and is essential for effective policy making and planning (Nguyen, N. T., et al., 2020). In general, public transit operation conversion from heterogeneous to homogenous traffic condition, passenger demand management implementations, enhancing the main and feeder systems etc. More specifically, short-term forecasting of passenger demand is essential for hiring or scheduling carrier vehicles and personnel, maintenance of infrastructures, and allocation of other resources. Similarly, long-term forecasting of passenger demand is needed for the construction of permanent infrastructures such as roads or tracks, stops, stations, terminals, depots, administrative or departmental buildings, and procurement of carrier vehicles. Hence understandably, passenger demand forecasting is a crucial and an inevitable step.

Passenger demand forecasting is done using many methods, such as qualitative forecasting, cross-sectional data, and forecasting using time series data. However, it is found that Time Series Analysis is the most suitable way for forecasting passenger demand; where the passenger demand is represented through time series. Time series is a set of observations recorded sequentially over some time. The future values of a time series can be predicted based only on historical observations of the time series, external controlling factors, or a combination of both. Under time series data forecasting, the model-driven methods based on statistical models, data-driven methods based on machine learning networks, and deep learning strategies stand out for the effective use of massive time series data obtained from the ITS of the public transit systems to take effective strategies to improve the overall performance of the system (Zheng, J, et al. 2020).

A few essential model-driven methods for time series forecasting are exponential smoothing and autoregressive integrated moving average (ARIMA), whereas, seasonal ARIMA is used to handle the seasonality. The traditional methods are improved and tweaked only to a certain extent to get better results with the exponential growth of the processing power of computers, more advanced and complex methods are used for forecasting, which is data-driven. Such as Artificial Neural Networks (ANNs). Forecasting with Long Short-Term Memory (LSTM), a special case of Recurrent Neural Networks (RNN), is gaining popularity as it efficiently learns long-term dependencies. Accordingly, here in the current research work, a thorough literature review is done to closely understand and classify the works carried out previously on the passenger demand forecasting.

Noh, Y., et al. (2015), have carried fundamental research to propose a forecasting model that is used effectively for short-term railway passenger demand concentrating on significant routes in South Korea. The authors try to explore the potential application of SARIMA models. Features of the seasonal trip and daily mean forecasting models are individually built depending on weekdays/weekends. This has been mainly considered in the analysis to gather the features of weekday/weekend trips and the legal holiday trip data. The study confirms to have higher accurateness and consistency by verifying the forecasting values of developed models. The projected models of this study are expected to utilize for starting a plan for the short-term operation of the main rail lines of south Korea.

Kumar, S et al. (2015) have shown that limited data-based short-term forecasting of traffic flow is better done using Seasonal ARIMA (SARIMA) models than ARIMA. The study considers three urban lane roads in Chennai as a case study, and three successive days of limited traffic flow data has been used for model development using SARIMA. The performance of the developed model is then validated using 24 hrs testing data. In the study, forecasted flows are compared with the actual flow values and confirmed with less error. In the conclusion author has made a point that the ARIMA forecasting method proposed in this study for traffic flow prediction can be used when the database is the main limitation.

Li, L., et al. (2018), have proposed hybrid models to improve the prediction accuracy of Xi'an metro line passenger demand over conventional prediction models. This study combines hybrid models' symbolic regression and Autoregressive Integrated Moving Average Model (ARIMA).

In the model development process, the performance of the hybrid model, ARIMA model, and Back Propagation (BP) neural networks are compared. The results show that the hybrid model beats the other two models developed regarding their demand prediction accuracy. Mean Absolute Percentage Error (MAPE) is the tool used to measure the performance of all the models; hybrid models showed a 54.24%, 58.98% rise in the accuracy over the BP neural networks and an additional 64.44%, 68.27% rise over the conventional ARIMA models for entrance and exit respectively.

Pavlyuk, D., (2017), have reviewed the multivariate models systematically for their application in short-term traffic forecasting. Forecasting models which have been considered in the study are autoregressive integrated moving average models (ARIMA), error correction models (VECM and EC-VARMA), space-time ARMA models, and multivariate autoregressive space state models (MARSS). The author has discussed the basic principle of all the models and their importance in usage in the field of transportation engineering.

Gummadi, R., et al. (2018) used APSRTS passenger flow data to forecast using traditional ARIMA models.

The study focuses on, passenger demand for the Macherla to Chilakaluripet bus route, AP, as it is largely dependent on the public transit system instead of using privatised vehicles; so, it is essential to forecast the capacity percentage of the public transportation especially buses in a given specific period of the day for the convenience of the commuters.

This study has considered two traditional forecasting models, ARIMA and Seasonal ARIMA models, for future demand prediction purposes. Predicted results from the model helped operators improve service competence and decrease the commuters' waiting time due to the deficiency of the number of buses when there is high demand in the flow of commuters. Also, it is beneficial to plan the bus schedules per future demand and look at the available resources.

Cyril, A., et al. (2018) have carried out univariate time series forecasting analysis for the demand considering ARIMA as a model. The author has considered passenger data from 2010 to 2013 belonging to the public transport system buses operating between Trivandrum and five other Kerala districts. The forecasted accuracy of the selected models is checked by calculating MAPE between actual passenger demand observed in 2013 and forecasted values from the models. The study's results illustrate that the time series ARIMA model, developed using historical data of passenger demand, is precise for zones mainly dependent on each other and for short-term demand forecasting.

Astuti, S. W., (2018) has developed time series forecasting models to predict the number of passengers traveling between Surabaya- Jakarta rail lines to enhance PT's operational capabilities. State Owned Enterprise operates rail transport services covering passenger and freight transport, especially in East Java Province. All the necessary passenger demand data collected from the operators have been pre-processed and analysed using SARIMA models. Amongst various SARIMA-based models developed, the best one is finally selected. Then, the performance of all the models is assessed based on correspondence between the predicted and actual amounts and the calculation of the Sum Squared Residual.

Obtained results from the study show that the best time series model based on minimum Sum Squared Residual (SSR) is SARIMA Model (0,1,1) (1,10)12. Based on the best fit model, the forecasting of Surabaya – Jakarta train passengers from January 2018 until July 2018 ranged from 119,495 – 161,685. One of the inferences shows that, a higher volume of passenger flows in July 2018 resulted from school and college holidays. Increased trend of passengers during peak season helped the operators improve the performance of the train facility by increasing the number of trains serving higher demand during peaks.

Cyril, A., et al. (2019), have use seasonal different Holt-Winters' modelling methods and multiplicative models to predict the bus passenger demand from Thiruvananthapuram to five other districts of Kerala using the Electronic Ticketing Machine (ETM) data from 2010 to 2013 and compared the predictive accuracy of both the models using Mena Absolute Percentage Error (MAPE). Based on the obtained results, the author has concluded that better forecasting of seasonal data can be seen by the Holt-Winters' Damped Multiplicative method.

Chikkakrishna, N. K., et al. (2019), have used A PROPHET and SARIMA models to assess short-term traffic prediction. The author has taken the seven-day hourly traffic volume data of the National Highway 744, Tamil Nadu, as the base for the study.

SARIMA (1,0,0) (2,0,0) is developed to forecasts traffic volume along with PROPHET. The models' accuracy is cross-checked using mean absolute percentage error (MAPE) and root mean square error (RMSE). It is found that ANN-based PROPHET models are better performing than SARIMA models in terms of forecasted accuracy.

Gallo, M., et al. (2019) have worked on forecasting metro onboard passenger demand using Artificial Neural Networks (ANNs) based forecasting models. For the study purpose, passenger demand data have been gathered from passenger counts at Line 1 of Napel Metro station turnstiles. Numerical results from this study show that the projected ANN-based approach forecast passenger demand in metro sections with reasonable accuracy and identify the best performing ANN-based model to predict passenger flow.

Xiong, Z., et al. (2019) have considered two deep learning neural networks for forecasting suburban rail passenger demand time series, such as a long short-term memory neural network (LSTM NN) and a convolutional neural network (CNN). The analysis is carried out using the past passenger flow data of Beijing metro stations and lines. Meantime forecasting is also carried out using traditional time series models such as autoregressive integrated moving average (ARIMA) and seasonal autoregressive integrated moving average (SARIMA). The deep learning models' prediction results are compared with selected traditional ARIMA and SARIMA models.
It is found that both the deep learning models better capture the time or spatiotemporal features of the urban rail transit passenger flow and give precise results for the long-term and short-term forecasting of passenger flow. Authors have also mentioned that deep learning methods have strong data adaptability and robustness compared to traditional models. They are perfect for forecasting the passenger demand flow of stations during peaks and the passenger flow of lines during holidays.

Li, J., et al. (2019) have characterised the departure passenger demand of different stations based on the Wuhan-Guangzhou high-speed railway's onboard passenger flow data. The quantity of the data considered in the study is between January 2010 and

December 2015. The obtained data set of passenger flow is prepared as per the requirement of long short-term memory (LSTM) model processing. From the study it is found that LSTM model better captures the influences of significant parameters on the prediction accuracy. Obtained results of LSTM are also compared with other traditional passenger demand forecast models. The results show that the LSTM model has got valid information in a long passenger flow time series forecasting and achieved a better performance than other models.

Lai, Y., et al. (2019), have used recent taxi demand data and other related information of Chengdu and Xiamen cities, China, is used by the author to carry out forecast analysis based on LSTM models. More precisely, the author has used a Spatio-temporal component of time series data to capture the Spatio-temporal characteristics of the data set and used an attribute component to gain external information (e.g., weather, point of interest). Finally, these two components are used to make the final predictions about passenger demand. The study results show that the proposed LSTM-based demand forecasting approach beats other conventional times series methods.

Carmona-Benítez, R. B., et al. (2019) have given SARIMA Damp Trend Grey (DGT) Passenger demand predicting model (SDTGM) for the airline industry. The proposed model is specifically considered to gather all the varying behaviour of the time series data set. The author took the United States domestic air transport market data to compare the existing DTG model performance with the proposed SDTG model. Authors also show that a given model has less uncertainty in the prediction than ordinary DTG models. Both models are used to carry out the simulation, and it is found that the SDTG model is capturing the seasonality effect present in the data set and does not allow the forecast to rising exponentially. Author has concluded that the proposed model forecasts more reasonably with short lead times when having a massive data set than the DTG model.

Abbasimehr, H., et al. (2020) have conducted a forecasting analysis on the strategic product demand of furniture manufacturer companies using a deep learning technique based on Long Short-Term Memory (LSTM). The monthly sale quantity for a specific product between 2007-2017 of a furniture company in Iran is used as a data source for the study.

Analysis results have discovered that the deep learning-based LSTM model performed better than the traditional time series models. Also, the authors conclude that the proposed LSTM model is used effectively in other domains, especially passenger demand forecasting of transportation systems.

Li, X., et al. (2020) have used passenger data from hybrid ride-sharing facility transportation systems in DiDi Chuxing in Haikou, China, for forecasting future passenger demand. Initially, spatial and temporal features of the passenger demand for express and ridespliting services are related and assessed. The significant factors influencing the both the modes of passenger demands are identified. Based on historical order of demand, travel time rate, the demand of adjoining areas, day-of-week, time-of-day, weather conditions, and points of interest, a combined deep learning-based model such as WT-FCBF-LSTM (Wavelet Transform, Fast Correlation-based Filter, and Long Short-term Memory) is proposed. The model is used to predict passenger demand in different regions for various time intervals taken in the analysis.

Finally, results are validated and compared with LSTM, WT_LSTM, and FCBF_LSTM models; it is observed that WT-FCBF-LSTM improves the prediction precision and better captures the different Spatio-temporal features of express and ridespliting services.

Fuloria, S. (2020), has carried-out short-term forecasting of passenger demand of Uber service of New York City using three potential time series models such as exponential smoothing, multiple regression, and Long Short-Term Memory (LSTM) and finally prediction accuracy of all three models have been compared.

The main aim of study is to facilitate a platform to give information on the higher demand areas, so that drivers move from areas of low demand to areas of high demand. Current work shows that conventional time series models like exponential smoothing and multiple regression are more explainable. At the same time, deep learning methods like

LSTM, with their complex procedures, are more precise in some circumstances and conclude that LSTM models better predict training and testing datasets.

Lee, J., et al. (2020) have compared the air passenger demand predictive performance using various time series models. The data set is taken from Incheon Airport and considered the air passenger demand data between the years 2002 to 2019. The naïve method, the decomposition method, the exponential smoothing method, SARIMA, and PROPHET are some of the models used in the study.

Obtained data initially shows the trend and seasonality behaviour. The authors tries short-term, mid-term, and long-term forecasting of air passengers. The study concludes that exponential smoothing models are suitable for a short-term forecast, the SARIMA model best fit for medium-term forecast considering stationarity was excellent, and finally deep learning-based PROPHET model is the outstanding one for long-term forecasting.

Zheng, J., et al. (2020) established a traffic flow prediction model built on the long short-term memory (LSTM) network. The author has taken traffic flow time series data from the roads of Changsha, Central China's Hunan Province.

The planned model of LSTM is compared with the traditional forecasting model and one more neural network-based model, viz. autoregressive integrated moving average (ARIMA) model and backpropagation neural network (BPNN) model respectively. The analysis shows that the proposed LSTM forecasting models outperforms the two typical models in forecasting precision.

Rabbani, M. B. A., et al. (2021) have carried out accidental data forecasting through two potential time series models such as Seasonal Autoregressive Integrated Moving Average (SARIAMA) and Exponential Smoothing (ES), and the performance of both the models is compared based on different measures.

This work intends to clearly understand the pattern of accident rates at varying time points. Mean Absolute Error, Root Mean Square Error, Mean Absolute Percentage Error, and Bayesian Information Criterion are some of the study’s performance measures. Both models' forecasting accuracy is measured, and the study shows that ES model better forecasts the accident data than the SARIMA model.

Finally, the study provides information on future accident rates, which can be adopted effectively in designing roads to ensure the safety of end users. Analysis of the study helps policy-makers, design advisors, and accident prevention departments in taking many strategies or policies.

Kanavos A., et al. (2021) have forecasted the aviation passenger demand by developing time series and deep learning-based forecasting models. Forecasting analysis is done through two traditional models and one deep learning-based model, like, Autoregressive Integrated Moving Average methods (ARIMA), Seasonal ARIMA (SARIMA), and Deep Learning Neural Networks (DLNN).

The authors have used forecasted results to compare all three models, and hence optimal modelling approach can be further worked. The investigational results have shown that the deep learning-based DLNN method provides more significant support in forecasting air travel demand by giving precise and robust results. Consequently, the DLNN method can also be utilized to predict the air passenger demand reliably.

Li, W., et al. (2021) have carried out short-term passenger demand forecasting for urban rail transit in Beijing. Data is mainly gathered from three important subway rail stations in Beijing city.

Seasonal autoregressive integrated moving average model (SARIMA) and support vector machines (SVM) models are used to predict subway future passenger demand in the study. Both models show their robustness in adapting to the complicated, nonlinearity, and periodicity data obtained from urban rail transit. SARI-MA–SVM model show improved accuracy and reduced errors in predicted values.

So, the authors have concluded that the SARIMA– SVM model fully describe the variations in the traffic flow and is more apt for passenger demand prediction.

Nar, M., et al. (2022) have studied the significance of predicting the passenger's demand in COVID situations. The authors have taken railway passenger demand data for the study. Demand prediction work has been done in two stages; priorly, online-based passenger demand prediction is made using statistical techniques such as regression analysis and simple averages. Error in the forecast is measured using mean absolute percentage error.

In the second stage of the study, station-based passenger demand prediction is carried out using artificial neural networks and machine learning (ML) algorithms technique, and error is calculated and compared.

The study concludes that the most effective and consistent results for demand prediction on a station basis are obtained through the decision tree, which is one of the ML algorithms.

2.5 TRANSIT CAPACITY AND SPEED IN ANALYSING THE SYSTEM'S PEERFORMANCE

The utmost focus of the public transit system is to move individuals from one place to another. Accordingly, capacity concerning transit service is one of the critical performance measures, and it focuses more on the sum of persons served by the transit buses in a specified amount of time *(person capacity)* than on the number of transit buses aided by a transit system *(facility* or *line capacity).* Finding the vehicle capacity is habitually an essential primary stage in finding a person's capacity of transit system. Person Capacity of Transit System is defined as "The maximum number of individuals that can be moved from past a given place during a given period with definite operating circumstances; lacking arbitrary delay, risk, or constraint; and with realistic certainty" and Vehicle Capacity is defined as "The maximum quantity of transit vehicles (buses, trains, vessels, etc.) that can pass a given position during a given time at a definite level of consistency." (TCQSM 2013)

Various influential factors are identified that affect passenger capacities, such as carrying units present for the individual vehicles (e.g., cars per train), size of the vehicle, and how well the total area inside each transit vehicle is owed between a total number of seats and standees. Policies made by the policy makers or management regulations also control standees are to be allowed inside the buses and the providing number of wheelchair places. The operator's policy also regulates a required optimum space for each standing passenger, limiting the number of standees permitted or accommodated inside the buses.

The size of operating vehicles or buses, along with their inside layout, influence the dwell time at the stations. As vehicles sizes disturbs the probability of a bus incoming at a bus station that is already packed with passenger's demand as some of them will need to make their space to and out of the door(s) before co-passengers can board. The fare collection system adopted at the station, the platform height provided concerning the bus floor, the waiting location provided at the station for passenger's relation to boarding gates, and the number and width of boarding gates has an impact on the required average time for boarding the bus by each passenger. Lastly, several patterns adopted in land use of the particular station, pedestrian facilities provided, and transit service characteristics influence the passenger demand to use the service by the transit buses at a given stop or bus station. Thus, it can be inferred that dwell time at the bus station is the product of the number of boarding passengers at the critical (typically busiest) door reproduced by the time taken by passengers individually, along with that the time essential to serve to alight passengers through the similar door is also considered.

Vehicle capacity depends on the minimum headway planned for the buses in their effective scheduling and operation, which in turn influenced by number of parameters such as dwell time at the bus stations, characteristics of roadway such as dedicated lane or non-dedicated lane of bus operations, length of the platform provided at each bus station and transit as well as general traffic signal cycles. The factors that will affect the transit capacity values are the same factors that influence transit speed calculations and reliability estimation. Overall, passengers tend to get attracted to more reliable faster service.

The time essential for transit buses to travel along the route is reduced by improving overall transit speed. At the same time, a reduction in the scheduled recovery time is achieved with an enhancement in the system's overall reliability. In the best-case scenario for a transit operator, the combined reduction in running and recovery time would be greater than or equal to one headway. This result allows the route to be operated with one fewer bus or, alternatively, to be operated at a higher frequency than before at the same operating cost. To be precise, the time saved postpones the need to add more services to maintain a particular headway due to delays arising from traffic congestion. Figure 2.3 shows factors influencing transit capacity.

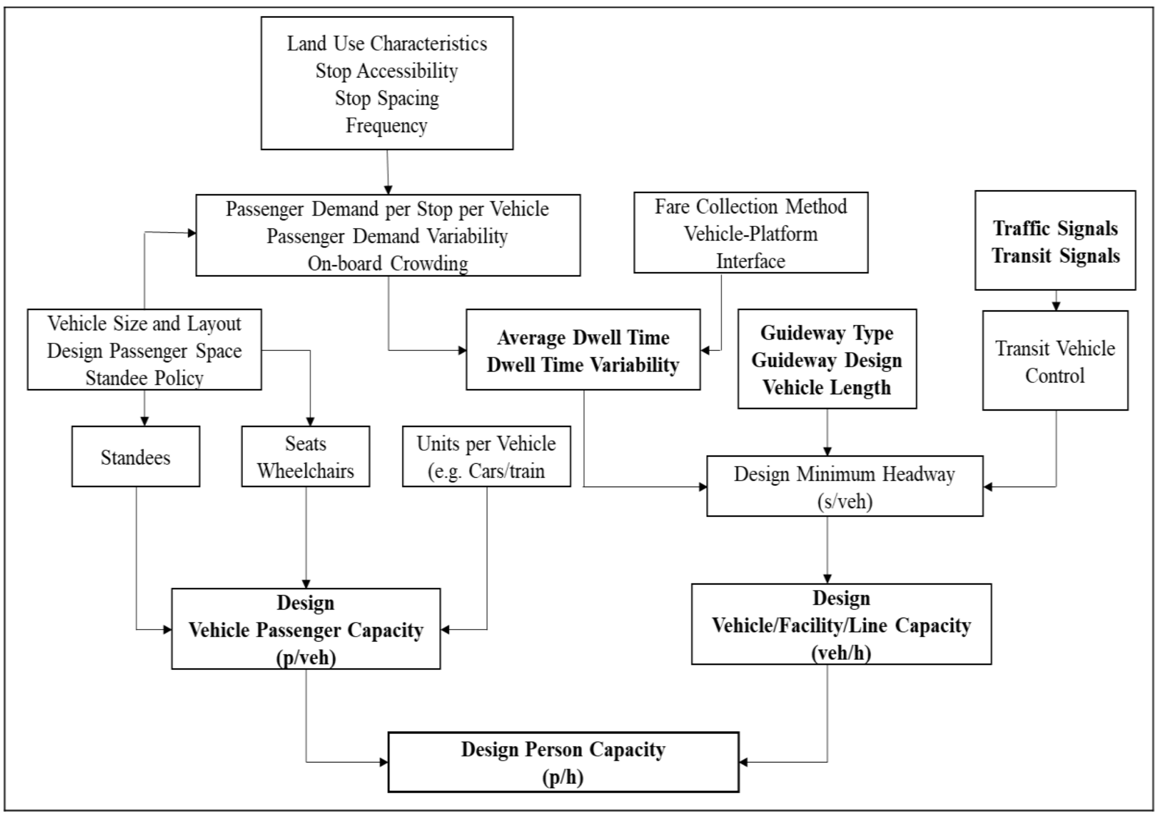


Figure 2.3 Factors Influencing Transit Capacity (Source: TCQSM 2013)

Transit speed is one of the significant indicators important to commuters; speed directly influences the travel time of individual buses while making each trip. The shorter the travel time taken by the transit buses compared with other modes of travel time, like, the two-wheelers, more is the potential passenger model shift to the transit service. Attracting higher ridership toward the public transit system is an important goal of any transit operator, but speed plays a vital role in deciding the cost of operating a route.

Determining a total number of buses required to meet the demand in the transit passenger at a particular frequency is subjected to many influential parameters on the entire route's cycle time, viz. the total time needed to complete a round-trip on the selected route, in addition to that driver's relaxation or break time or any other added schedule retrieval time required beyond break time. The total cycle time taken by the individual buses divided by the planned headway of the buses give rise to the necessary number of vehicles to serve the route. Reducing the route cycle time adequately mitigates the essential quantity of buses, resulting in cost savings. Then again, the saved number of buses can be effectively utilized to enhance the frequency of the buses on the same route or another route, with no total alteration in the system's operating costs (TCQSM 2013).

Generally, while working on speed estimation of transit buses, it is divided into three parts: *running time*, which is the time expended at a constant speed and with the subsequent acceleration of buses, and *passenger service time*, which is nothing but time spent in passenger boarding and alighting process, and *delay* is mainly due to external reasons that obstruct bus operation. Obtained bus times are expressed as the travel time rate of the bus (time required to travel a given distance); the inverse of the travel time rate is speed.

The number of bus stations between the route or line stimulate all three parts of transit speed mentioned previously. If buses often stop at the station, it leads to more time spent decelerating and accelerating the buses than during bus running. Similarly, if stops spread passenger demand among all the stops more, it reduce the average boarding volume and, in turn, reduces the dwell time at each station; however, acceleration and deceleration delays typically more than offset any dwell time benefits. Finally, when bus stops are more, transit buses never reach the maximum speed mainly because they must begin decelerating to the next bus stop.

Running time is characteristically controlled by the roadway design, such as dedicated and non-dedicated lanes, operational features of buses such as operated acceleration, maximum vehicle speed, etc., and frequency of stopping. The boarding and alighting time provided is directly related to the number of stops made to serve and the average dwell time taken at each stop. The total number of transit buses using the roadway concerning its capacity also impacts total delay. Transit vehicles operating on roads also face a general uncertainty at each intersection. Figure 2.4 shows the factors influencing transit speed. Transit preferential treatments can help offset some of mixed-traffic operation's negative impacts on transit speed.

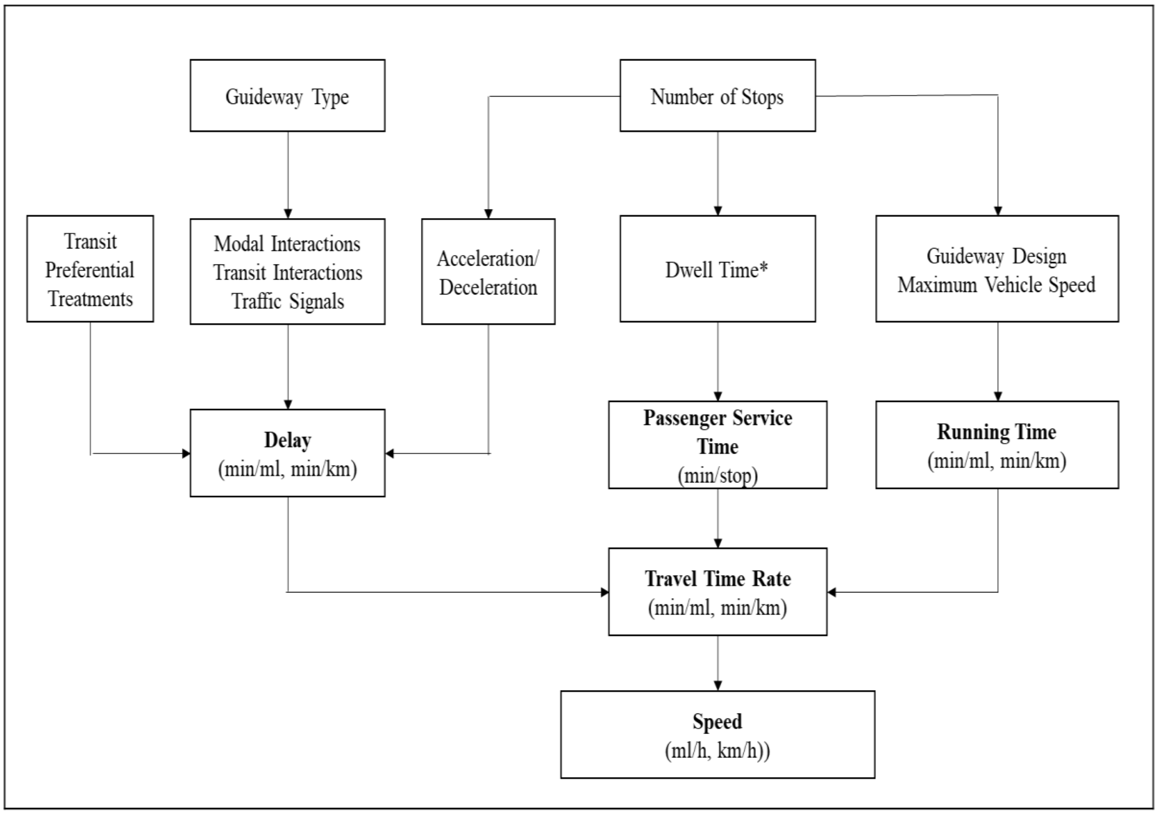


Figure 2.4 Factors Influencing Transit Speed (Source: TCQSM 2013)

Hidalgo, D., et al. (2012) have presented a systematic procedure for obtaining the passenger capacity for several BRT system features that drives beyond the limits of textbooks and manuals. The study comprises a theoretical and practical assessment of the maximum capacity of the bus lanes, signalized and unsignalized intersections, and bus stops. The framed methodology is applied and tested for the TransMilenio BRT system in Bogotá, Colombia.

Along with the proposed method, modified empirical models are given by TCQMS in view of the Saturation Rate instead of the failure Rate. As per the new formulas for assessing high-capacity BRT corridors, the identified critical segment of TransMilenio BRTS has a practical capacity of 48,000 passengers per hour per direction with its available infrastructure facilities. It is inferred from the study that improvements in the current infrastructures, such as providing supplementary platforms, operating high-capacity buses, non-grade facilities at critical intersections, and some other strategies, enhanced the capacity, speed, reliability, and service quality of the selected transit system.

It is identified that Bogota BRTS has many bottlenecks which can be addressed and solved with operational modifications, such as reallocating bus platforms within stations, refining the traffic signal timing to reduce inconsistency in vehicle arrivals, and reducing the number of vehicles that stop at bus stops with limited line-up capacity.

Widana Pathiranage, R., et al. (2013) have worked on assessing the consequence of non-stopping buses on the queuing capacity of critical bus stations. As part of the study, the author has developed a simulation model by taking the Buranda Busway stations as a case study. The station at Buranda is the important busway station, being fourth lengthiest of 10 stations along the southeast Busway (SEB), which is 16 km in length with 4.4 km south of the Brisbane CBD Queen Street Bus Station.

The microsimulation approach adopted in the study follows the procedure of the current deterministic practice mentioned in the TCQSM. It is found that the higher number of non-stopping bus percentages is causing an increased corridor demand with a minor decrease in the existing bus stop capacity. The simulation model developed shows better BRT line service capacity of the corridor due to heterogenous stopping and non-stopping buses percentage at the bus stops.

Gandhi, S., et al. (2013) have examined the multiple options of planning, operational and design strategies for the BRTS. The study quantifies the performance of all the possibilities with a spreadsheet tool, and a comparative analysis is carried out to shortlist the best alternate options available. A total of sixteen theoretical arrangements, two standard designs in variable settings, and two presently operational design variations are compared.

Finally, various BRTS design configurations are made using the spreadsheet technique. The tool is mainly based on standard motion equations, while default values and weights for indicators used in the tool are solely based on initial surveys conducted. The tool gives the required results with performance indicators considered in each category.

Results obtained from the tool are validated on three operational BRTS systems such as Ahmedabad BRTS, Bogota BRTS, and Delhi BRTS, and confirms with accuracy in the 94 to 99% range. Results are also show that operational bus speeds 25 % lesser in the case of open systems compared with closed BRTS systems.

However, it is observed that open systems provide more passenger speed than closed bus operations, with a length of the trips less than 10km. Limiting the bus speed to 40 KMPH, especially during peak hours, for safety considerations does not hinder the passenger or operational performance.

Sandeep., et al. (2013) have evaluated the different operational and design conditions of BRTS and established the relationships between various features of BRTS, such as spacing provided at the stations with the average speed and many more. Authors have again used a spreadsheet technique called a BEAD tool and effectively modelled various design configurations of BRTS. Performance measures provided by the BEAD tool in each category are validated on three working BRTS systems; Ahmedabad, Bogota, and Delhi.

A comparative study is carried out in three stages; a comparative assessment of sixteen theoretical design alternatives is done, then comparative analysis of two design alternatives with changing traffic conditions is done, and finally, a comparative analysis of two existing BRT systems is done.

The study has found that open systems offer more passenger speeds than closed bus operations for trip lengths less than 10km. The commercial speed of buses in BRTS operation has a linear relationship with average spacing between the stations.

Widana Pathiranage, R., et al. (2014) have stated that when a bus station area approaches its designed capacity, bus inflow to the particular bus station forms queuing. This situation is similar to the operation of a negligible vehicle's movement at unsignalized intersections. This concept is taken into consideration by the author to inspect BRT station operation and establish the relationship between bus queuing at each station and the capacity for the busway station. For the planned purpose, the author considered the South East Busway (SEB) as a case study in Brisbane, Australia. Different variables are used in this study: capacity of the station, degree of saturation, and queuing at the bus stations.

The authors developed two mathematical models in this study as per AIMSUM and ASB model to establish potential capacity of selected bus stations and results are compared with TSQSM models.

It is established that potential capacity, measured per AIMSUM and ASB model (maximum achievable outflow from the station), relatively matches the TCQSM deterministic model without operating margin. Finally, it is concluded by the author that queuing at the stations does not affect the overall capacity of all buses stopped at the station. However, this substantially enhances the present TCQSM procedure for assessing BRT station capacity.

Sharma, A., et al. (2015) have used empirical and simulation-based capacity models to calculate the bus lane capacity of Bhopal BRTS, India. Empirical models are taken from TCQSM, and the VISSIM tool is used to establish a simulation model in the study.
Through the empirical model, the bus lane capacity found to be 41 buses per hour. In the case of models established from simulation data, the bus lane capacity found to be 39 and 38 buses/h concerning to failure rate (FR) and speed reduction (SR) concepts respectively. Both the models are compared for their performances and the results show that the FR approach (error 4.8%) is closer to the actual bus lane capacity.

Singh, H., et al. (2017) have used Ahmedabad BRTS to estimate its capacity considering its operating condition as the hybrid bus rapid transit (BRT) corridor. The busiest route with the dedicated and non-dedicated lane bus operation is taken based on its boarding and alighting data.
After calculating the capacity of the hybrid BRTS corridor empirical approach, the effect of mixed traffic conditions is observed on overall corridor capacity. To compare empirical approach results, capacity estimation is also calculated using the conventional Greenshield model on a mid-block section.
The calculated capacity of the hybrid BRT corridor is found to be 101 buses/hour. Kalupur railway station bus stop is the most critical bus stop in a non-dedicated segment (mixed traffic environment). A dedicated lane-based Shivranjini bus stop has a capacity of 243 buses/h and roughly 42% more than a non-dedicated lane bus stop. It is also observed that buses faced more delays due to mixed traffic conditions, increasing headway between scheduled buses. This is the main reason for the lower capacity values in the non-dedicated segment. The Greenshield model, used in the study, overestimated the dedicated segment's capacity by around 19.34%.

Kathuria, A., et al. (2018) have hypothesised that BRTS bus stations gives higher failure rate (FR) values than a conventional bus transit system. With this hypothesis, BRTS station LOS ranges are developed based on Speed and FR values to calculate the maximum FR. The likelihood of queue formation at the stations reduces by adding every extra loading area for the selected station. Hence in this study, the authors have given FR-based level of service (LOS) ranges for BRTS stations with two loading areas using empirical models of TCQSM and a simulation model developed using VISSIM.

As a result, it is found that FR values suggested by the TCQSM manual result in underestimated capacity values of bus stations. A maximum FR of 29% is given in the current study by the author, which is around 4% more than the maximum FR of conventional bus stops given in the manual. The authors also stated that FR values suggested in the current study suits to estimate the accurate capacity of BRTS bus stations and corridors**.**

Chepuri., et al. (2015) traffic flow characteristics are analyzed on a 1.8 km BRTS corridor of Surat city. For characterization, they have used the microscopic simulation software VISSIM 7.0. As identified corridor has four intersections along its length; hence authors have evaluated the delay caused to the traffic flow along these intersections precisely. Travel time data has been obtained based on GPS equipment installed in the testing vehicles. Video graphics-based traffic surveys are conducted to have classified traffic volume count. Data on road inventory, speed and delay (V-box device), and spot speeds (Radar gun) are also gathered as a study. V/C ratios of the identified corridor are estimated and found to be very low, and at many places, it is high; hence authors tried to balance it by diverting some traffic from high ratio to low ratio, with the analysis done using VISSIM software itself.

2.6 SUMMARY OF THE LITERATURE REVIEW

Travel time reliability measures are relatively new and are evidenced as very effective tools in analysing the system performance. There are various measures through which TTR is accounted for, which are given in section 1.4.3. Reliability indices provided by FHWA, such as travel time index (TTI), buffer time index (BTI), planning time index

(PTI), 95^{th} percentile travel time, etc., and various descriptive statistics are some of the essential measures to quantify the TTR.

Frequent and arbitrary delays are the primary sources of the variation caused to transit travel time. Even though there are several methodologies to assess the TTV of the transit system, as travel time is random, random probability distribution is one of the practical and systematic methodologies. Many continuous distribution families are tried in the methodology to explain the transit travel time behaviours at various time stamps and spatially. Thus, Travel time variability studies for any transit system is significant in keeping its performance to the required level.

Several external factors cause unreliability in transit travel times, like; hour of the day, day of the week, condition of the bus operation, signalized and unsignalized intersection density, bus stop density, etc. Suppose operators want to make any implementation strategies related to advancement or improvement in the system. In that case, assessing the heterogeneity issue on the travel time and modelling covariate impacts on the TTR is significant. Hence, effective modelling of TTR with those affecting variables or factors by considering passenger and operator perspectives is needed in the transit industry today.

The mass transportation system of ITS has to be designed considering many factors, passenger demand being an important one. However, with the current passenger demand, the future passenger demand should also be considered while designing. Demand forecast for public transportation provides a realistic picture of its future usage and is essential for effective policy-making and planning for transit performance upgradation. Several methodologies exist to make this happen, ranging from conventional statistical methods to more advanced and complex techniques that work on artificial neural networks (ANN).

Transit capacity and speed estimations are empirical studies on transit systems to assess their performance. Vehicle capacity is generally controlled by the dwell time, bus lane characteristics such as mixed traffic manoeuvre, dedicated lane operations of the buses, length and width of a platform at stations, etc. Meanwhile, traffic and transit signals influence the minimum headway of buses; those operate on the roadway, which controls the vehicle capacity. Influential factors affecting the transit capacity also affect the transit speed. Transit speed directly correlates with the transit travel time; hence, assessing the travel time indirectly correlates with determining the transit speed. TCQSM 2013 has suggested a well-organized method to estimate the Transit capacity. Along the side, simulation-based methods can be tried as alternative tools.

Evaluating the Level of Service (LOS) is vital when transit systems are to be compared in their performance. This comparison can be made by establishing the LOS of any network, route, or segment, considering numerous travel time reliability measures; such as the TT coefficient of variation, Average TT, and other reliability indices. Cluster analysis and cluster validations are the significant stages to be followed during the LOS development.

2.7 GAPS IN LITERATURE REVIEW

The followings are some of the gaps found in the above literature review.

- Travel time variability studies of the BRTS corridor are less especially considering its variation spatially and temporally. Few works of literature have been found to use statistical distribution to assess TTV.
- Studies lack inferences of distribution model parameters considering different aggregation levels such as peak and off-peak hours, weekday and weekend days, etc.
- TTR studies of the transit systems considering distribution parameters are very few.
- TTR reliability modelling given operators' and passengers' perspective dependent variables, is less noticed in the Indian research community.
- Studies concentrating on modern forecasting techniques in developing countries like India are substantially less.
- Research carried out on passenger demand forecasting with deep learning techniques that use APC data is minimal.

- In the context of the existing BRTS corridor in India, studies on performance analysis considering capacity and speed are fewer.
- No such studies have so far defined the LOS of the BRTS considering dedicated and non-dedicated segments.

CHAPTER 3

STUDY AREA

3.1 GENERAL

This chapter focuses on the detailed background of the study area considered in the current research. The chapter presents the Hubli-Dharwad city profile in the first part in terms of the city's location, physical features, population, land use pattern, etc., and then explains the trend that has followed in the development of the transportation system between these cities to date. Finally, the chapters cover the elaborated description of the Hubli-Dharwad Bus Rapid Transit System and study stretches considered in the current research work.

3.2 HUBLI-DHARWAD CITY PROFILE

Hubli-Dharwad is located in Dharwad district, in the north-western part of Karnataka state of India. Hubli-Dharwad's location map is shown in figure 3.1. Dharwad district lies between 15°02' and 15°51' North latitudes and 73°43' and 75°35' East longitudes which comprises an area of 4230 km^2. In 1962, Hubli and Dharwad merged to form the Hubli-Dharwad Municipal Corporation (HDMC), which was incorporated as a sister or twin city. Hubli-Dharwad is today a substantial commercial, industrial, and educational centre of Karnataka State. According to the 2011 census, 9,43,857 people live in cities together.

Hubli is an important commercial centre of the region. It is also the headquarters of the Southwest Railway Branch. Dharwad is district and primarily an educational city. The two centres are 22 km apart and are connected by National Road (PB Road), National Road (NH4), and Mumbai-Bangalore Railway. Hubli Airport has daily and periodic based flights to Mumbai, Bangalore and to some other metropolitan cities of the country, respectively.

Three areas in the twin cities' land use plan combine residential, industrial, and institutional land use at a specific rate; Hubli, Dharwad, and Navanagar show different activity concentrations. Due to the industrial and tertiary activities in and around Hubli, Hubli is a centre of trade and commerce, Dharwad is an education centre, and Navanagar, located in the corridor between Hubli and Dharwad, was built in 1979 as a large well planned residential area. Because of this mixed land use, the twin cities attract large numbers of people every day (Directorate of Urban Land Transport, 2011-12).

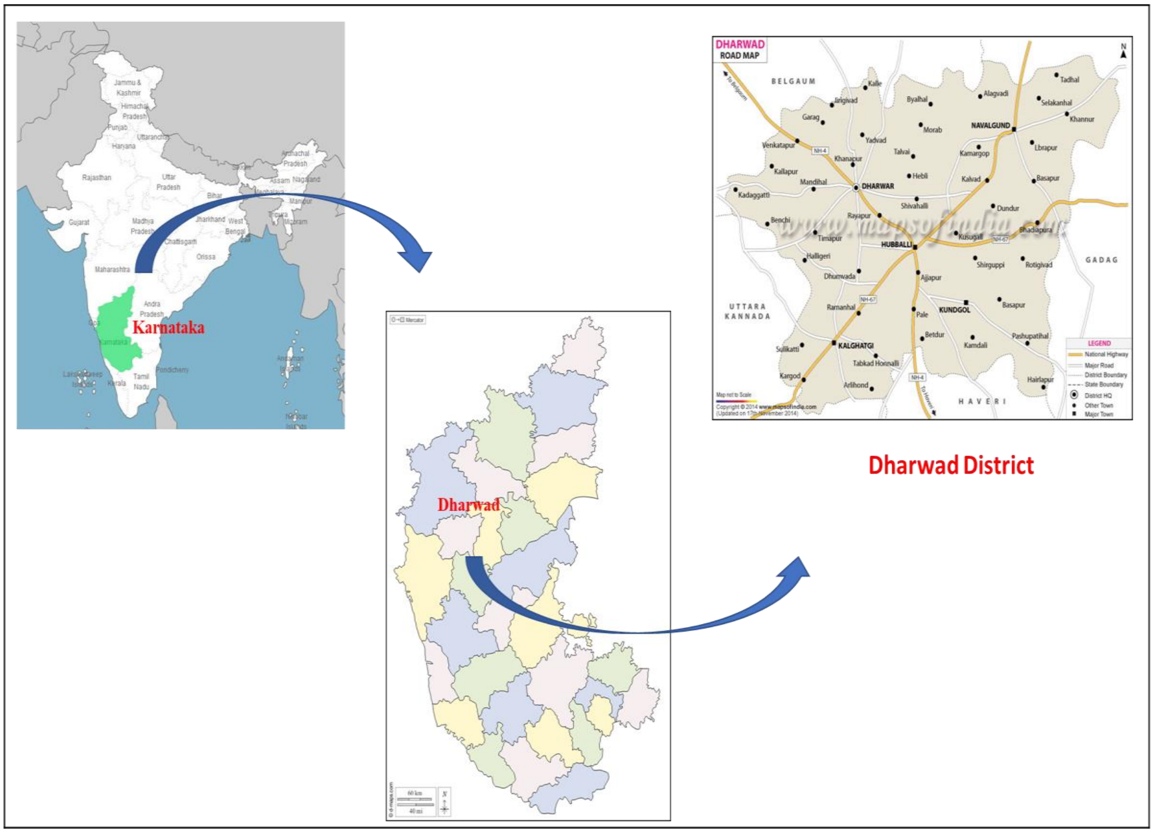


Figure 3.1 Hubli-Dharwad Location Map (Source: Google Images)

3.3 DEVELOPMENT OF PUBLIC TRANSIT SYSTEM IN BETWEEN HUBLI-DHARWAD

Hubli-Dharwad, a twin city, is well connected to other important cities in the country by road, rail, and air. The two towns are connected by a robust linear road (PB Road). The National Highway-4 was built to bypass the traffic on the PB Road and join the twin cities with Bangalore and Pune. Other highways passing through/connecting the city are NH-218 (to Solapur), NH- 63 (to Haliyal and Gadag), SH- 73, and SH- 28 (to Goa). The major district roads connect Kalghatai, Soundhatti, Haliyal, etc. Inside both cities, the road network is radial (Directorate of Urban Land Transport, 2011-12).

Based on typical land use characteristics of twin cities, work-based and education trips govern the total trips made by all the city's transit modes. Also, based on the passenger data characteristics, it is observed that more travel demand towards Dharwad is observed for education purposes, while travel demand towards Hubli is mainly work-driven. Before 2018, the public transport facility between the twin cities was majorly provided by NWKRTC, where Hubli City-2 Depot, routed from Hubli division, was managing those city buses. Along with the government undertaking public transit system, a consortium of private bus owners named Bendre Nagara Sarige has been giving healthy competition to cater to the large number of commuters between Hubli and Dharwad daily, till date.

In 2011, the Directorate of Urban Land Transport (DULT) carried out a modal share analysis as part of the Comprehensive Traffic & Transport Plan (CTTP) for the twin cities. It was observed that bus-based transit account only for 7% of the total passenger vehicles, but its share in terms of a passenger is over 70%. It was understood that the bus-based public transit system majorly carries the travel demand between twin cities. But even though there was higher passenger demand for public transit systems between the twin cities, existing systems lacked improper commuter service in terms of time, comfort, cost, safety, and, more importantly, reliability in the service provided. This established an intrinsic need for better bus-based public transit system between the twin cities in terms of space and operations. As a result of the need analysis by the DULT in 2011, The Hubli-Dharwad Bus Rapid Transit System (HDBRTS) was established in the year 2018 between twin cities to serve public transport as a part of the Sustainable Urban Transport Project (SUTP) and got funded by the Government of Karnataka, the Ministry of Housing and Urban Affairs (MHUA), World Bank, and Global Environment Facility (GEF).

3.4 HUBLI-DHARWAD BUS RAPID TRANSIT SYSTEM

The BRTS company has an authorized share capital of INR 20 Cr., out of which 70% is the share of the Government of Karnataka and Hubli-Dharwad Municipal Corporation (HDMC), North-Western Karnataka Road Transport Corporation (NWKRTC), and Hubli- Dharwad Urban Development Authority (HDUDA) share the remaining 30%**.** The BRTS corridor between Hubli (HUB) and Dharwad (DWD) is 22.25 km long, with the width of the cross-sections ranging from 44 m to 35 m. HDBRTS is a hybrid-based system where buses ply on dedicated and non-dedicated lanes for short lengths. The buses ply through a dedicated corridor from the Hosur circle in Hubli to the Jubilee circle in Dharwad. The BRT buses ply along with the mixed traffic beyond Hosur circle up to CBT, Hubli and beyond Jubilee circle up to CBT, Dharwad. The system has 35 stations, including both side terminals, out of which one station is yet to start operation (Station 25) effectively. Table showing HDBRTS stations considered in the study given in the Appendix A.3. Figure 3.2 shows the transit corridor map of the HDBRTS along with the route details. BRTS corridor ROW also includes mixed traffic lanes, footpaths, etc. Details are given as follows.

- Stretches from Hosur Circle to Naveen Hotel, Hubli – Total carriageway width 35 m
- Stretches from Naveen Hotel, Hubli to Gandhinagar, Dharwad – Total carriageway width 44m
- Stretches from Gandhinagar to Jubilee Circle, Dharwad: Total carriageway width 35m

The main corridor includes segregated busways, controlled bus stations, off-board ticketing through smart cards and tokens, and high-quality buses (standard and articulated). The corridor is designed for operating regular and express services. The BRT corridor consists of two lanes for BRTS buses on either side of the median bus station facilitating overtaking lanes for express services. Salient features of HDBRTS characteristics are listed in table 3.1 as follows. (Source: Directorate of Urban Land Transport, “HDBRTS DPR”).

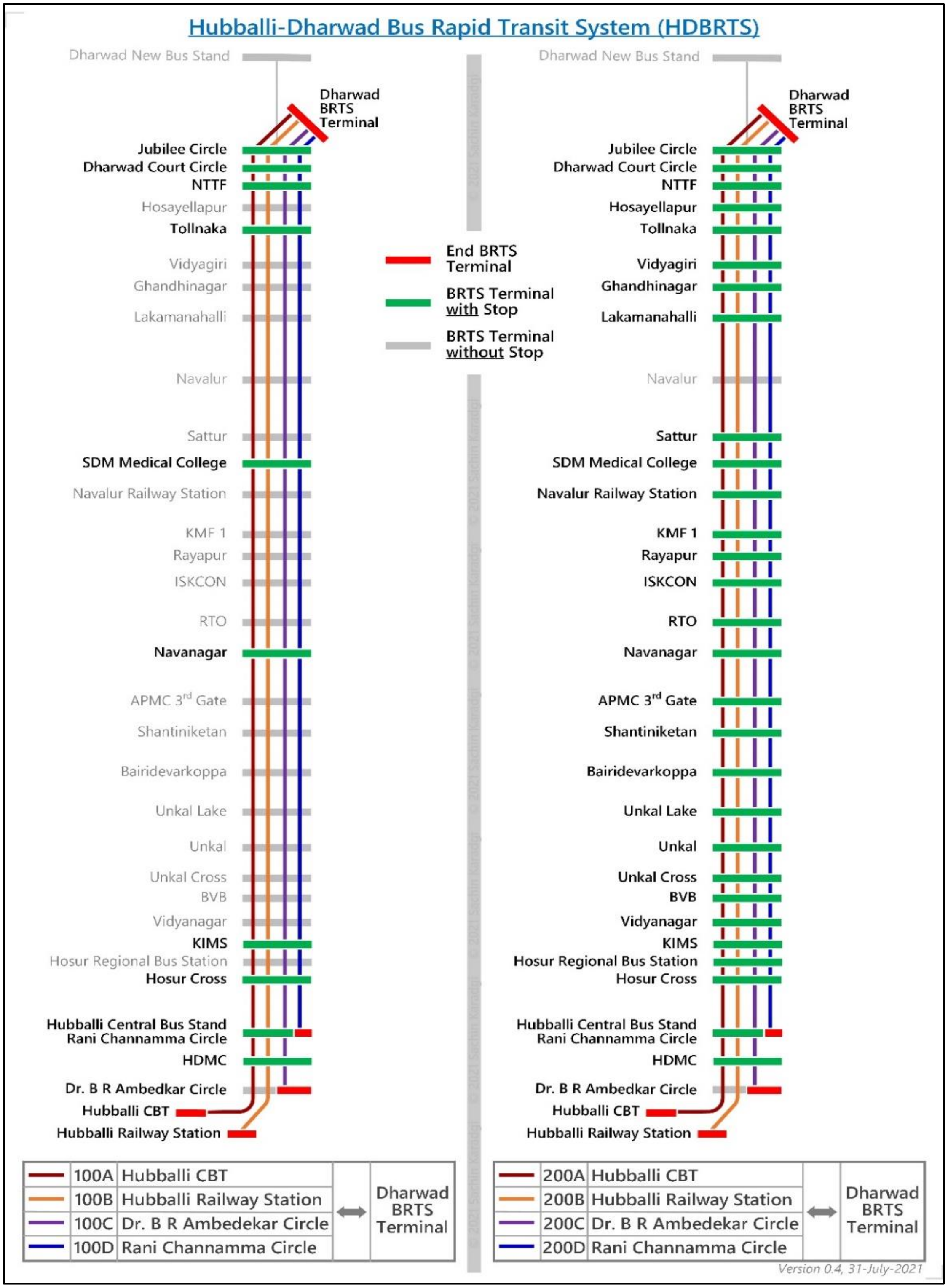


Figure 3.2 Transit Corridor Map of HDBRTS

Table 3.1 Salient features of HDBRTS

Features	Details
Began operation	October 2nd 2018
Type of system	Hybrid BRTS
Number and type of corridor	A single and straight corridor
Length of network	22.25kms
Number of stations	35
Total routes	Six all operational (as per march 2020 data)
Service time	5 am to 11 pm-midnight
Fleet size	100 (70 regular buses which stop at all the stations and 30 express buses stop at seven stations
Daily ridership	More than 60000
Average headway	Regular buses 3.0 min and Express buses 4.0 min.
Bus capacity	Total 54 in numbers
Type of bus station	Median type

As mentioned in the previous paragraph, HDBRTS buses operate through express and non-express routes along the single linear corridor. Express route buses serve the limited bus stations, whereas non-express route buses serve all the bus stations. Most of the buses of both environments run from terminal to terminal, such as the terminal at the Hubli side to the terminal at the Dharwad side, that is UP terminal which is from Hubli to Dharwad and the DOWN terminal is from Dharwad to Hubli. Bus routes are extended beyond the terminal at the Hubli side, up to Hubli CBT, and beyond the terminal at the Dharwad, up to Dharwad's new bus stand. But the frequencies of those routes are less compared to terminal-to-terminal bus operations. Each route operating in the UP and DOWN direction is assigned a unique route id for easy identification, and data regarding each route is gathered accordingly.

HDBRTS has a single BRT corridor of 22.25 km connecting the two twin cities. Most of this corridor has a dedicated nature for the bus operation, and a small part has a non-dedicated nature. The BRT corridor from Hosur Circle of Hubli City to the Jubilee Circle of Dharwad is dedicated in nature, in UP and DOWN directions and the corridor from Hosur Circle Hubli to CBT Hubli is completely non-dedicated in nature. Figure 3.3 show the google plots of the route, dedicated and non-dedicated segments of the HDBRTS corridor.

Figure 3.3 HDBRTS Selected Routes and Segments Plots on Google Map

For the current research work, express and non-express routes are considered for the route level analysis. One dedicated segment on the Dharwad side, one dedicated segment, and one non-dedicated segment on the Hubli side are considered for the segment-level analysis. Details of routes and segments are given in table 3.2 and table 3.3

Table 3.2 Details of HDBRTS Route

SI. No	Route Type	Route ID	Start	End	Length, km	Bus Stops Served
1	Express UP	100U, 100U1	Hubli BRTS terminal	Dharwad BRTS terminal	19.50	10
2	Express DOWN	100D, 100D1	Dharwad BRTS terminal	Hubli BRTS terminal	19.54	10
3	Non-express UP	200A -UP	CBT, Hubli	Jubilee Circle, Dharwad	21.3	29
4	Non-express DOWN	200A -DOWN	Jubilee Circle, Dharwad	CBT, Hubli	21.6	30

Table 3.3 Details of Dedicated and Non-dedicated Segments

SI. No	Segment Type	Origin	Destination	Distance (km)	Number of bus stops	Intersecti ons
1	Dedicated UP (DWD)	Lakhamanhalli	Jubilee circle	3.6	8	8
2	Dedicated DOWN (DWD)	Jubilee circle	Lakhamanahalli	3.6	8	8
3	Dedicated UP (HUB)	Hubli BRTS terminal	Bairidevarakopp a	5.5	10	9
4	Dedicated DOWN (HUB)	Bairidevarakop pa	Hubli BRTS terminal	5.5	10	9
5	Non-dedicated UP (HUB)	Hubli CBT	Hosur Circle	1.9	3	5
6	Non-dedicated DOWN (HUB)	Hosur Circle	Hubli CBT	2.1	3	5

Data collection for all the framed objectives and subsequent analysis has been done for selected routes (express and non-express) and three segments exclusively (Two dedicated and one non-dedicated). Data points have been extracted for all the days of the week, different hours and periods of the days. Hence, along with spatial aggregation patterns of routes and segments, different temporal aggregation patterns are also considered in the analysis. In detail discussion has been made in the subsequent chapter of the thesis. Further route details of HDBRTS given in the Appendix A.5.

3.5 SUMMARY

Hubli-Dharwad is located in Dharwad district, in the north-western part of Karnataka state of India. Both cities are popularly known as twin cities in the state of Karnataka. These twin cities are developing at a moderate pace by accommodating mixed land use patterns of residential, industrial, and institutional land uses in some proportion. Travel demand between two cities is majorly dependent on the public transit system, such as public transport vehicles between the twin cities account for only 7-11% of total vehicular movement between cities; but they carry an excessively high load of people movement around 70-80% of people on this corridor. With this background Government of Karnataka has taken the initiative to implement BRTS in these two cities. Subsequently, HDBRTS started its operation in the year 2018. The BRTS corridor between Hubli (HUB) and Dharwad (DWD) is 22.25 km long and provides services as express and non-express routes; and also, it is a hybrid-based system, where buses ply on dedicated and non-dedicated lanes for short lengths. In the current research work analysis has been carried out for two routes (express and non-express) and three segments exclusively (Two dedicated and one non-dedicated) of HDBRTS.

CHAPTER 4

DATA COLLECTION AND METHODOLOGY

4.1 GENERAL

As in the case of any other research on public transit system, in the current research work, data collection and analysis is another important step that will hold the effectiveness of formulated objectives and the working methodology. Accordingly, this chapter, firstly will explain the nature of data gathered and its pre-processing for the current research work. Then it focuses on explanation of the overall framework of the research work, followed by an explanation of the unique methodologies adopted for each objective planned in the current research work.

4.2 DATA COLLECTION AND PROCESSING

As mentioned in the section 1.4 of the chapter 1, Hubli-Dharwad BRTS has adopted an Integrated Transit Management System (ITMS) for systematic management of day-to-day transit operations. The primary aim of ITMS is to create an enterprise management system that allows the company and its host of service providers to manage their activities in a highly coordinated manner leading to a high-productivity environment and reliable services to the users. The system also aims at creating a process-based system that continually allows the operations to be monitored against accepted service levels and provides improvement opportunities to transit managers to offer services at the best operational levels.

ITMS comprises of many components and gathers a variety of smart data from the operation of a transit system. The system operation started on 2nd October 2018. As a part of current research work, official permission for ITS-based operational data sharing has been obtained from the administrative department of HDBRTS in the year 2019, and subsequently Automatic Vehicle Location data (AVL) from 8th December 2019 to 29th February 2020 (a total of 84 days) and Automatic Passenger Count Data (APC) from 1st December 2019 to 29th February 2020 (a total of 91 days) is considered for the further analysis. Here are the brief details of the data gathered as per the current research requirement.

4.2.1 Automatic Vehicle Location System (GPS-based vehicle tracking system - AVLS) Data

There are 100 AC buses of HDBRTS; equipped with Global Positioning System (GPS). The ITS control room continuously tracks the location and time stamps of the buses. Buses gather their location by updating their latitude, longitude, and time stamp at every six-second interval. The raw data from HDBRTS operators is obtained in the form of .csv format and includes detailed information; like, ‘asset_id,’ ‘created date,’ ‘direction’, ‘latitude,’ ‘longitude’, ‘route_id’, ‘speed’, ‘GPS_date’, ‘GPS_time’, ‘t1-time at when event generated at device’, ‘t2-event sent to server from device’, ‘t3-event received at server’, ‘t4-acknowledgment sent from server to device’ and ‘trip_id.’ This data underwent processing for further use as per research needs. A sample of the AVL data sheet obtained from HDBRTS operators is shown in the Appendix A.1.

4.2.2 Automated Passenger Count System (APC)

APC data is one more type of data gathered from the HDBRTS operators. The system is responsible for enabling individual passenger ticket data, which is stored at ITS regularly when passengers get their tickets from the bus stations. This data mainly helps in analysing the passenger-related reliability measures. The APC system was designed to get a ticket even if the passenger had the pass and which should be scanned both at the turnstiles installed at the boarding and alighting station to enter or exit the station. Hence, the ticket data can be reliably considered to count passengers. Again, similar to AVLS data, the raw APC data of the passenger obtained in the .csv format. Data sheets contain information such as ‘date issued,’ ‘time issued,’ ‘operator ID,’ ‘terminal ID,’ ‘device type,’ travel direction information such as ‘issued boarding station’ and ‘issued alighting station,’ ‘payment method,’ ‘ticket serial number,’ ‘rider type,’ ‘ridership’ and ‘total revenue.’ This data too was pre-processed and cleaned for outliers. A sample of APC data sheet obtained from HDBRTS operators is shown in the Appendix A.2.

4.2.3 Pre-processing of data obtained from AVL and APC

AVL and APC data used in the current research have been pre-processed before being used in the analysis. Pre-processing mainly deals with removing the outliers or inappropriate data points present in the raw data set.

All the 100 buses AVL data has been considered to process for getting final TT data points. In the case of AVL data, even though all the required information was in the obtained data set, it was observed that there were some issues with certain data sets where trip numbers were not changing once trips were completed regularly. Such trips have been considered outliers; some were removed, and some were modified with new unique trip numbers using algorithms and coding in python. The data set also had trial non-BRT trip details, which were removed, and only data about BRTS was kept.

While extracting the travel times from the AVL data for identified routes, most of the additional information in the data set was removed. For a particular BRT route, the data was filtered for a single asset id and single trip id and fed to QGIS to check the continuity and distribution of the data over the selected route. A sample of AVL data plotted on QGIS is shown in Appendix A.4. Trips that do not start at the origin and end at a destination were eliminated using Haversine's formula because these trips constitute less travel times and are incorrect. These short trips may result in due to non-functioning or error in the GPS device at some particular time stamp. Also, GPS points falling away from the corridor were removed, which may have also resulted due to the above-said reason. Once done with the pre-processing of the data set, Travel Times (TT) were extracted using python coding. Python coding used for TT extraction is given in the Appendix A.6. Very large and small travel times were removed from the extracted travel time data set using Median Absolute Deviation (MAD) technique. Finally, travel times were split temporally and spatially as per the research need for the selected routes and segments. For the final route and segment-level data sheets, it was ensured with more than 100 TT points existed for each hour and period respectively for the further analysis.

Python code used for temporal split is given in the Appendix A.7. and Sample of Hour-wise Split Travel Time Data points shown in the Appendix A.8.

In the case of APC data, the raw data contained 54,27,301 observations from 1st December 2019 to 29th February 2020. The extra information in all the other fields except ‘date issued,’ ‘time issued,’ ‘issued boarding station,’ and ‘ridership’ was removed. The total number of passengers boarded at each station in the 91 days was calculated.

Outliers are usually unexpected spikes or dips in the value of observations on the time-series graph and must be removed from the data. Some ways of dealing with outliers include modifying them after identifying their source, replacing them with the mean values, or neglecting them. There were hardly any outliers across all the time series here, and the ones present were simply due to technical or operational errors, so they were replaced by the mean values of passengers for that time. Moreover, it was observed that the passenger flow after 23:00 and before 5:00 was negligible and almost non-existent. So, the observations between that time period were removed, and data contained only between 05:00 and 23:00 was considered for further analysis using the APC data set.

4.3 METHODOLOGY

The devised methodology is presented in this report section to achieve the proposed objectives in this work. Different methods are framed for each objective mentioned in chapter 1, along with the flow chart.

Current research work deals with the performance analysis of HDBRTS from multiple perspectives. The work gets divided into four parts answering the work done to meet the four objectives of this research work. As discussed in the previous sections of chapter 3, Automatic Vehicle Location data and Automatic Passenger Count data are the two primary sources of the data used in this work. A comprehensive literature review of past work is made on the identified research area. These two became the basis through which the overall research framework and individual objectives methodologies have been finalized. The framework of the proposed research work is shown in figure 4.1.

As shown in figure 4.1, the first and second objectives, such as the travel time variability study of the system and TTR modelling are carried out based on AVL and APC data sets. Travel times are extracted at different temporal and spatial aggregations, and they are the primary sources of data points in the analysis, along with certain external conditions. External conditions considered in the study are the hour of the day, day of the week, intersection density, bus stop density, land use pattern, etc. Along with travel time data points, demand data points are extracted from the APC data set and used in carrying out those two objectives. The third objective, demand modelling of the system, is based only on the APC data set. The fourth objective is based entirely on the TTV and TTR analysis carried out from objective one. Again, the fifth objective is making strategic short-term, mid-term, and long-term recommendations based on the results obtained and conclusions drawn from all four objectives.

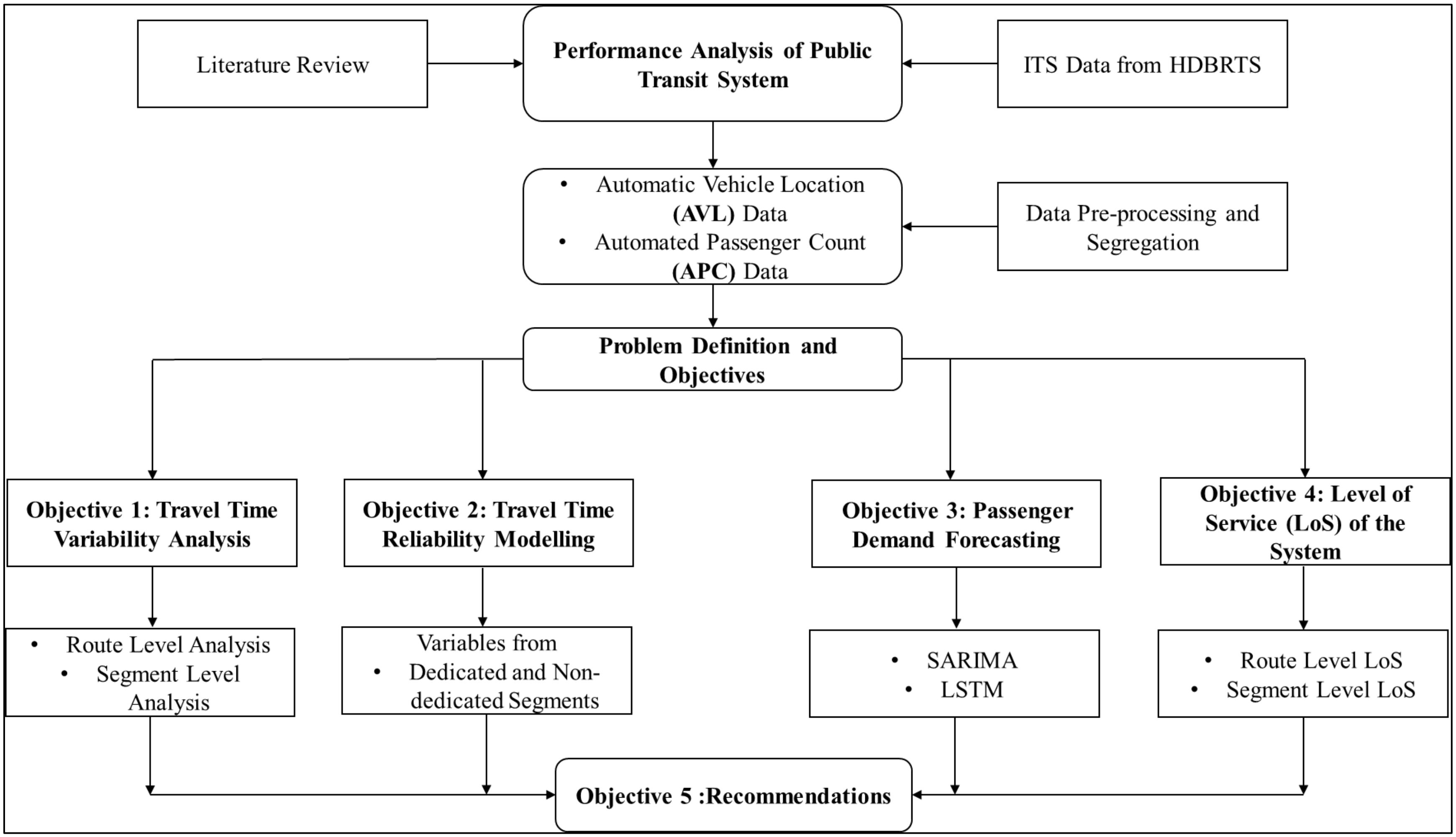


Figure 4.1 Framework of the proposed research work

4.3.1 Travel Time Variability Study of the System

The study's first objective aims to explore the travel time reliability and travel time variability concerning various aggregations attributes such as an hour of the day, day of the week, route-wise, and segments wise. The clear methodology is shown in Figure 4.2 to carry out this proposed objective.

Data collection being the significant step in this study, GPS-based AVL data and APC is taken from the ITMS of HDBRTS. More dwell time, bus bunching at the stations, signal delays at intersections, and peak and off-peak traffic hours of the day are a few of the general incidences which have influences on the TT reliability of the system. Keeping all these points in observation, end-to-end travel time variability and reliability analysis is carried out for the HDBRTS. Analysis is carried out for the two routes (express and non-express) and three segments exclusively (Two dedicated and one non-dedicated). Travel time data points are extracted for all the days of the week and different hours of the day. So, different temporal aggregation patterns are considered in the analysis. Extracted data points about the both UP and DOWN direction separately.

For the express and non-express routes, travel time data sheets are prepared for the hour-wise time frames between 05.00 to 22.00 for all the days of the week in both UP and DOWN directions. Segments travel time data sheet is prepared for the periods wise time frames such as morning off-peak (05:00 to 08:00), morning peak (08:00 to 11:00), Interpeak (11:00 to 14:00), afternoon off-peak (14:00 to 16:00), evening peak (16:00 to 20:00) and evening off-peak (20:00 to 22:00). Different period wise time frames are considered based upon the analysis of passenger demand from all the stations of the system. For the route level analysis, a total of 924 travel time cases are considered, whereas in the analysis carried out for the segment's a total number of 252 cases are considered.

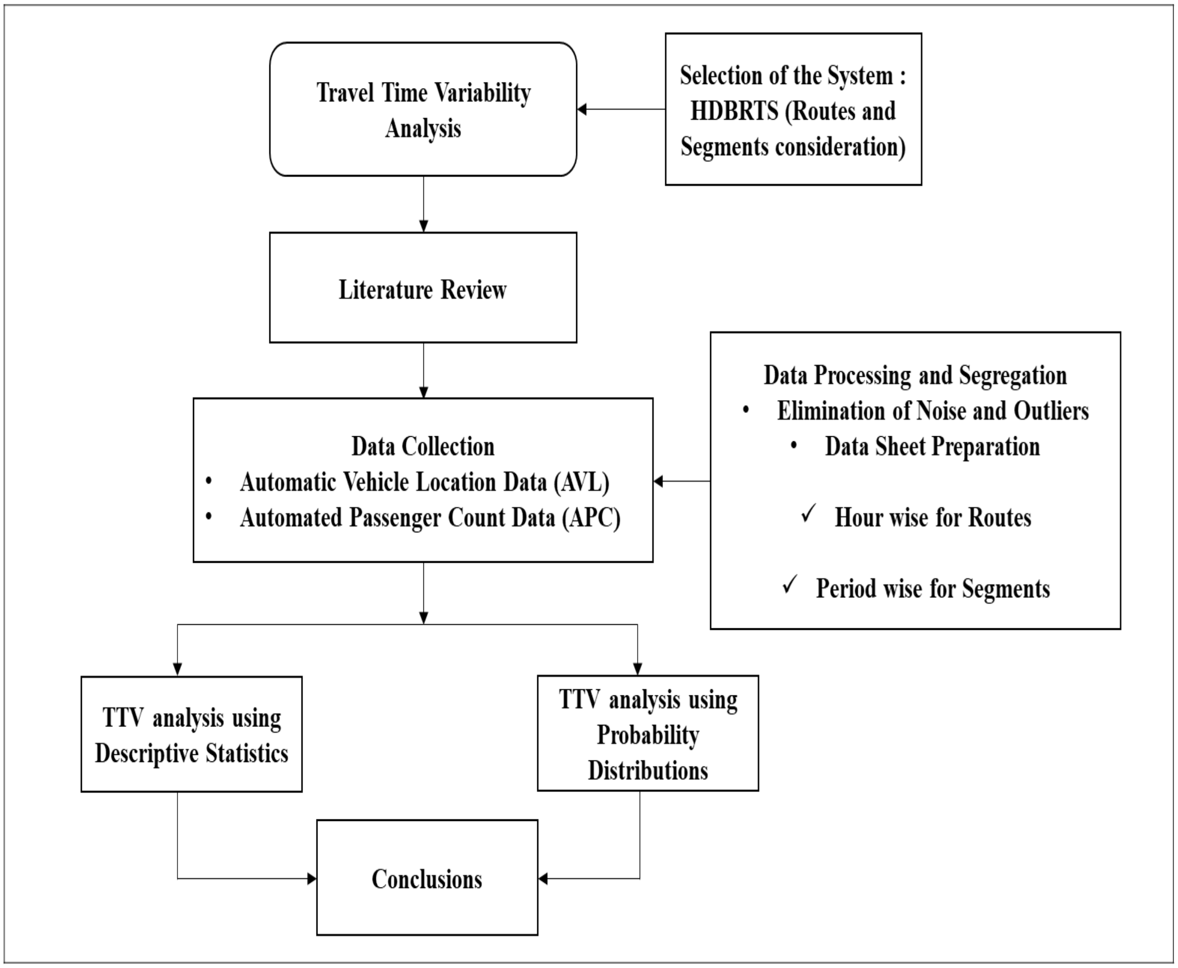


Figure 4.2 Methodology of TTV Study of the System

Travel time variability analysis for the route and segment levels are carried out in two stages. In the first stage, descriptive statistics and TTR analysis of the selected data points are done. To understand variations in the descriptive statistics and TTR values for weekdays and weekends and at the different hours of the day, analysis is carried out for all the selected hours of the day and all the days of the week. TTV and TTR measures considered in the study are average travel time (Avg. TT), Standard Deviation of the TT (SD of TT), Coefficient of Variation of TT (CV of TT), Planning Time Index (PTI), Buffer Time Index (BTI) and Travel Time Index (TTI). Equations to calculate the above measures defines by Federal Highway Administration (2006) and are given in equations (4.01) to (4.07).

- **Average Travel Time (ATT)**

$$ATT = \frac{Sum\ of\ travel\ time\ all\ the\ buses}{Number\ of\ buses} \tag{4.01}$$

- **Standard Deviation of TT (SD of TT)**

$$SD\ of\ TT = \sqrt{\frac{\sum_{i-1}^{n}(x-\bar{x})^2}{N-1}} \tag{4.02}$$

Whereas xi = Value of each data point

X̄ = Mean of the Travel Time

N = Number of data points of the Travel Time

- **Coefficient of Variation of TT (CV of TT)**

$$CV\ of\ TT = \frac{Standard\ deviation\ of\ travel\ time\ all\ the\ buses}{Mean\ of\ travel\ times\ of\ all\ the\ buses} \tag{4.03}$$

- **Buffer Time (BT)**

$$Buffer\ Time = 95th\ Percentile\ Travel\ Time - Average\ Travel\ Time \tag{4.04}$$

- **Planning Time Index (PTI)**

$$Planning\ Time\ Index = \frac{95th\ Percentiel\ Travel\ Time}{Free\ Flow\ Travel\ Time} \tag{4.05}$$

- **Buffer Time Index (BTI)**

$$Buffer\ Time\ Index = \frac{95th\ Percentiel\ Travel\ Time - ATT}{Average\ Travel\ Time}\quad X\ 100 \qquad (4.06)$$

- **Travel Time Index (TTI)**

$$Travel\ Time\ Index = \frac{Average\ Travel\ Time}{Free\ Flow\ Travel\ Time} \qquad (4.07)$$

An essential foundation for understanding travel time variability is characterizing the statistical distribution of travel time (Li, R., et al. 2006). Hence in the second stage of the study, probability distribution fitting is carried out for both the routes and selected segments separately. EasyFit software is used in the study to illustrate the distribution fit for the data set. The unimodality behaviour of the data points using the Hartigan dip test is confirmed before the distribution fit and selection of the best model (Hartigan, J. A., et al. 1985). Seven continuous distributions are tried, which were burr distribution, generalized extreme value (GEV), log-logistic distribution, logistic distribution, log-normal distribution, normal distribution, and Weibull distribution. The selection of the type of the distributions in the contemporary analysis is entirely based on the past literature review done. The Burr distribution is considered an essential applicant for travel time variability analysis and delivered an excellent overall depiction of the observed data (Susilawati et al.2011). Research has shown that even GEV distribution superlatively explains the TTV. GEV distribution is the best signifier for the public transit travel time variation. Meanwhile, transit performance is improved in terms of precision and robustness (M.M. Harsha., et al. 2021). Log-logistic distribution predicted the best estimate of the true conditional PDF of the travel time and generated the most accurate approximations of the expected secondary delays on the selected dataset (Ricard, L., et al. 2022). Weibull distribution has shown its efficiency as flexible in representing right-skew, left-skew, and symmetric travel time data (Kieu, L. M., et al. 2014). Some researchers have shown that a normal distribution is still a good fit for the peak, while a lognormal distribution is more appropriate for the off-peak (Mazloumi, E., et al. 2010).

The TTV analysis has extracted distribution parameters using the Maximum likelihood estimations (MLE) method. Kolmogorov-Smirnov (KS) test is used to extract the distribution parameters and check for the goodness of the fit of each distribution. The distribution can be well-thought-out as significantly suitable for observations made when the p-value is more than the significance level (0.05) and, in that case, fails to reject the null hypothesis Ho (M.M Harsha., et al. 2021). The data was not normally distributed at all hours of the day and in different conditions. Hence based on the K-S p-value, the robustness of best-fit distribution was selected and ranked amongst all the choices for describing the travel time data points under different conditions considered. In conclusion, as per the total number of cases passed by each selected distribution model, distribution performance was established at different ratios for all routes and segments.

Travel decision and transit mode selection of the commuter are considerably influenced by the travel time reliability due to traffic conditions and incidents on the corridor or network under a given roadway, traffic control, and environmental conditions. At an aggregate level, seven sources of variability in travel times have been identified: incidents, work zones, weather, demand fluctuations, special events, traffic control devices, and inadequate base capacity (Chepuri, A., et al. 2019). Demand fluctuation at each bus station is a common phenomenon that occurs in transit operations, and this fluctuation in passenger demand has a significant impact on causing variation in the travel time and subsequently lead to unreliable service from the transit system. Hence, inferencing the demand with the parameters obtained, the best-fit distribution model plays an important role in taking many strategic measures in the transit operations. These measures helps in improving the system's service condition, as this remains an unexplained study and have lies the core of this research work.

The probability distribution model's performance is decided based on the total sum of cases passed ratio, the ratio of the number of cases in the top three positions, and the ratio of the sum of cases there in the first position. Equations to calculate the above ratios have been given in the equation (4.08) to (4.1) (M.M. Harsha., et al. 2021)

$$Cases\ Passed\ Ratio = \frac{Number\ of\ Cases\ Passing\ the\ KS\ Test}{Total\ Number\ of\ Cases\ Generated} \quad (4.08)$$

$$Cases\ Top\ 3\ Ratio = \frac{Number\ of\ Cases\ Present\ in\ the\ Top\ 3\ Position\ based\ on\ p - value\ of\ KS\ Test}{Total\ Number\ of\ Cases\ Generated} \quad (4.09)$$

$$Cases\ First\ Ratio = \frac{Number\ of\ Cases\ Present\ in\ the\ First\ Position\ based\ on\ p - value\ of\ KS\ Test}{Total\ Number\ of\ Cases\ Generated} \quad (4.1)$$

Once after probability distribution fit with the travel time data points, the best-fit distribution parameters are attempted to compare with the passenger demand of that particular time stamp. Finally, insights are drawn from the shape of the best fit model's Probability Distribution Function (PDF) according to the demand variation.

4.3.2 Modelling Travel Time Reliability of the System

The second objective of the current research is to model the travel time reliability of the HDBRTS with observed and unobserved independent variables. Policymakers should be clear about the transit system's unreliable service when designing suitable strategies to improvise it. Effective TTR modelling determines the significance of the attributes by statistically testing the relationship strength of individual attributes with the overall transit service reliability (Ma, Z. L., et al. 2015). The methodology to deal with this objective is given in figure 4.3.

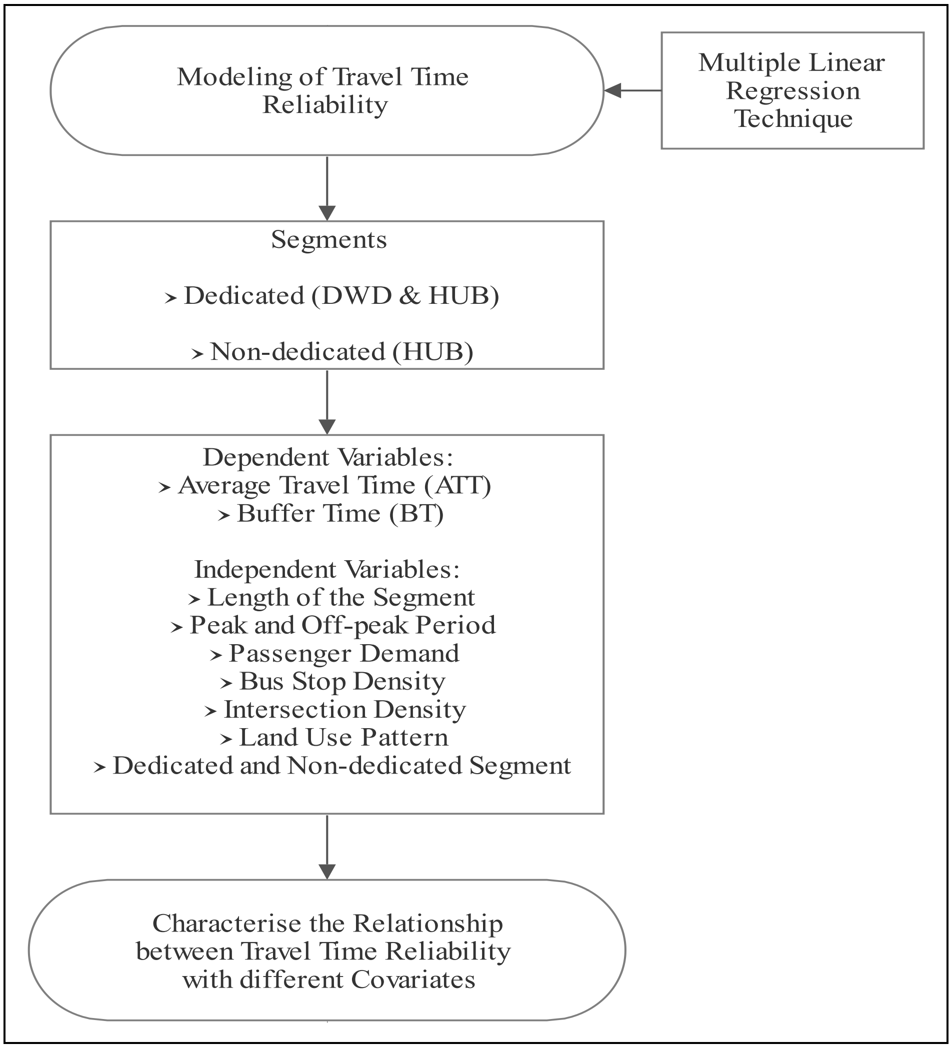


Figure 4.3 Methodology to Travel Time Reliability Modelling

Travel time reliability analysis (TTR) is vital in the transit industry. Despite substantial work in the national and international research community, less attention has been given to modelling the TTR with the heterogenicity of many covariates.

It is a well-known fact that BRTS bus operations in India are not entirely dedicated; part of it will carry out on the non-dedicated segment, and hence they are in hybrid mode (Kathuria, A. et al. 2020). Along with these non-dedicated segments, bus operations have mixed traffic conditions. The interference of mixed traffic, although occurs on a few road segments, considerably compromise the end-to-end travel time

(Gunawan, F. E., 2015). Hence, the segment-level TTR analysis and modelling become imperative for the planners and operators to give specific segment-level solutions rather than route-level solutions (Kathuria, A., et al. 2020).

HDBRTS is one such newly implemented hybrid modes of the transit system; hence TTR modelling is carried out at the segment level considering two dedicated and one non-dedicated segment in the UP and DOWN direction. Meanwhile, in the case of HDBRTS, no studies on TTV and TTR modelling have been carried out. Accordingly, results from current study helps in taking strategies that are essential to enhance the systems performance, which operators takes them precisely at the later stages.

Two dedicated and one non-dedicated segments mentioned in table 4.1 have been considered for developing the data set required for the modelling. The travel time data points have been extracted according to the selected segments. As express buses don't stop at some bus stations along segments, extraction of data points has considered only non-express route bus operations. Multiple Linear Regression (MLR) analysis is carried out for the TTR modelling, using SPSS is the tool for the analysis. Selecting the dependent variables and independent covariables is the critical phase in MLR modelling. Two models have been developed in the study, considering the operator's and commuter's perspectives.

Optimizing the average travel time is one of the challenging tasks for transit operators because the change in the average travel time significantly impacts service reliability (Ma, Z. L., et al. 2015). Hence, Average Travel Time (ATT) is considered one of the dependent variables.

Buffer time is extra time added to the commuter's journey for the guaranteed arrival at the destination, which causes unreliable service to influence on the planning behaviour of the commuters (Lomax, T., et al. 2003). Hence, Buffer Time (BT) is one more dependent variable considered in the study.

Independent variables are selected based on permutation and combination of multiple covariables. Pearson coefficient of individual independent variables is established priorly with the dependent variables. The Pearson coefficient yields a coefficient of determination of 0.72 or 72 % (Taylor, R., 1990) considers as 'high' correlation. Highly correlated independent variables with Pearson correlations of more than 0.65 to 0.70 are selected for further MLR modelling.

TTR is strongly affected by the number of bus stoppings and the length of the route (Mohamed, A. H., et al. 2021). All the three selected segments have variations in their length and different bus stop density as well as intersection densities. These are three variables are considered in the study. Passenger demand along the segment has shown a significant Pearson correlation with both independent variables of more than 0.75; hence, it is selected as one more covariable. Passenger demand and bus stop density are proxy variables for dwell times at this study's stations. Estimated travel time reliability parameters of the peak periods have directly related to the coefficient of variation of density as a measure of the spatial distribution of congestion across the network (Saedi, R., et al. 2020). Peak and off-peak periods are considered as dummy variables and assigned with Boolean variables 1 for peak and 0 for off-peak.

As HDBRTS is a hybrid-based system and one of the segments is non-dedicated in nature, to ascertain the impact of dedicated and non-dedicated lanes, Boolean variables are assigned 0 for dedicated segment and 1 for non-dedicated segment. Buses operating along the non-dedicated segment also experience the impact of CBD area. To analyse this impact systematically land use pattern has been also considered as one of the covariates by assigning Boolean variable of 0 for dedicated segment and 1 for non-dedicated segment in CBD area.

Finally, two MLR models were developed in relation to the two dependent and eight independent variables. The performance of both models was examined with the adjusted R^2 values, unstandardized coefficients, t-statistics, and statistical significance value of individual covariables of both the developed models. The following points are the basic knowledge that has been considered while interpreting generated results from MLR models.

Adjusted R square (R^2): One of the important parameters generated in the model summary to assess how well a multiple regression model fits the data is the adjusted R^2.

It compares the goodness-of-fit for regression models that contain different numbers of independent variables in building the model. Basically, adjusted R Square (R^2), as the name indicates, adjusts the apt number of independent variables in the efficient model. Notably, its value increases only when the new variable improves the model fit more than expected by chance alone. The adjusted R^2 value actually declines when the variable doesn't improve the overall model fit by a sufficient amount. The value varies from 0 to 1, if the obtained value is near 1, it shows that the model perfectly predicts value in the target field. On the other case if the obtained value that is less than or near to 0 indicates that the model that has no predictive value.

Unstandardized Coefficient: This parameter is obtained from the coefficients table generated in the MLR analysis. The unstandardized coefficient is one of the important parameters used in assessing the model developed. This parameter in the table is generated for all the independent variables taken in the model building. This measures the variation that happens in the selected dependent variable, for one unit of change in the independent variable corresponding to the raw values that are displayed in the original scale.

t-statistics: The value of t-statistics is again generated for all the independent variables in the coefficients table. Basically, its value measures the standard deviation of the coefficient obtained. In the interpretation of the MLR models, the t-statistics value for all the independent variables larger than **+2** or **-2** is considered acceptable.

Statistical Significance Value: This is one more significant parameter that can be seen in the coefficients table generated in the analysis. This value generated again for all the independent variables in the table. The significance value is also termed as p-value in the interpretation of the analysis. While in interpretation statistical significance for each independent variable where a p-value $\leq \mathbf{0.05}$ is considered acceptable at the 95^{th} percentile confidence interval.

4.3.3 Passenger Demand Forecasting of the System

Passenger demand forecasting with time series forecasting methodology is the third objective considered in the current research work. Figure 4.4 depicts the methodology to carry out this objective.

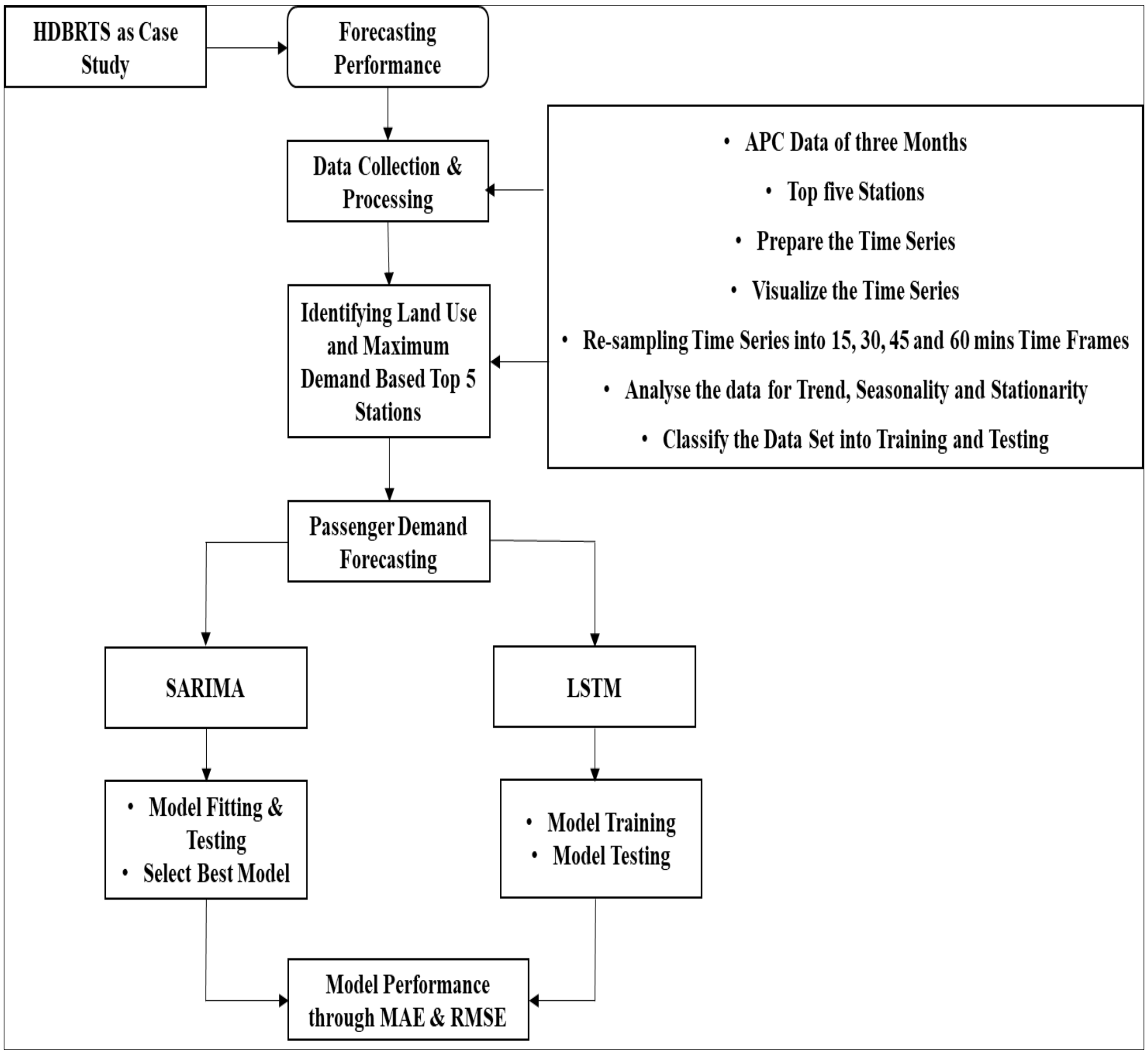


Figure 4.4 Methodology - Modelling Passenger Demand

The mass transportation system of ITS has to be designed considering many factors, passenger demand being an important one. However, with the current passenger demand, the future passenger demand should also be considered while designing (Halyal, S., et al. 2022). Demand forecast for public transportation provides practical picture of its future usage and is essential for effective policy making, planning and performance enhancement (Nguyen, N. T., et al. 2020).

Raw data in the analysis contained 54,27,301 passenger demand points from 1st December 2019 to 29th February 2020. Only required data points have been kept in the sheets after preliminary data processing. Top-5 stations with higher passenger demand were considered in the analysis; namely, station 05, station 28, station 33, station 34 and station 35, respectively.

Data sheets of individual stations are again resampled in the date time module of the python into time intervals of 15 minutes, 30 minutes, 45 minutes and 60 minutes. In this way, a total of 20 time-series – five stations with four time-frames per station – are prepared. Sample of ready time series is shown in table 4.1.

Table 4.1 Sample of Time Series

Issued at	Total Passengers
01-12-2019 06:00	19
01-12-2019 06:30	44
01-12-2019 07:00	142
01-12-2019 07:30	380
01-12-2019 08:00	467
01-12-2019 08:30	481
01-12-2019 09:00	500
01-12-2019 09:30	387
01-12-2019 10:00	480
01-12-2019 10:30	396
01-12-2019 11:00	490
01-12-2019 11:30	486
01-12-2019 12:00	631
01-12-2019 12:30	450
01-12-2019 13:00	535
01-12-2019 13:30	422
01-12-2019 14:00	466

The passenger data is plotted into time-series. A time series is an ordered sequence of values of a quantitative random variable at equally spaced time points that measures the status of some activity over time (Anvari, S., et al. 2016). Here, the total passenger count per time interval is plotted against the time. The time series, of 15 minutes, 30 minutes, 45 minutes and 60 minutes time intervals are plotted and observed for patterns such as trend, seasonality and outliers. Additionally, a day-wise time-series is also plotted and observed. Python code used to split the data accordingly given in the Appendix A.10.

Here, trend is a long-term increase or decrease in the value of time series observation. However, no apparent trend is observed in any of the time-series plotted. Seasonality is the presence of variations that occur at specific time intervals regularly over time. The seasonality can be of any time interval such as yearly, quarterly, weekly, daily or even hourly. Here, the data showed daily seasonality. As the data exhibits daily seasonality, the periodicity of the season considered as 17, 23, 34 and 68 for 60 minutes, 45 minutes, 30 minutes and 15 minutes time-series, respectively.

In the current study, obtained data is divided into training and testing sets. 80 % of the data is used for training, and the remaining 20 % is used for the testing (Medar, R., et al. 2017). Such as, the data from 1st December 2019 to 11th February 2020 (73 days) was used for training and the rest, from 12th February 2020 to 29th February 2020 (18 days), was kept for testing purpose in the analysis.

The time series forecasting can be done with traditional methods such as exponential smoothing and autoregressive integrated moving average (ARIMA). In addition, if seasonality is present, Seasonal ARIMA can be used to handle the seasonality. In addition to this, the traditional methods can be improved and tweaked only to a certain extent to get better results (Halyal, S., et al. 2022).

With the exponential growth of the processing power of computers, more advanced and complex methods can also be used for forecasting. One such method is the, use of Artificial Neural Networks (ANNs). Forecasting with Long Short-Term Memory

(LSTM), a special case of Recurrent Neural Networks (RNN), is gaining popularity as it can efficiently learn long term dependencies (Halyal, S., et al. 2022).

With this, the passenger demand forecasted is carried out using LSTM and seasonal ARIMA. Both models' forecasting accuracy are evaluated with Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). MAE is slightly more intuitive and easier to interpret and compute than RMSE. Optimising the forecasts for MAE results in forecasts of the median and optimising the forecasts for RMSE results in forecasts of the mean. Moreover, RMSE is more sensitive to outliers than MAE (Vandeput, N., 2021). Finally, the forecasted accuracy is compared between the methods, and the as a rough estimation of future demand seasonal naïve technique is used. MAE and RMSE are two of the most widely used measures to represent errors (Halyal S., et al. 2022). The detailed step-by-step methodology followed in both the passenger forecasting models has been given in subsequent paragraphs.

4.3.3.1 Forecasting by Seasonal ARIMA:

Autoregressive Integrated Moving Average (ARIMA) is a traditional forecasting technique that adds the lags of the differenced series (autoregressive (AR) terms) or lags of the forecast errors (moving average (MA) terms) or both to the prediction equation (Anvari, S., et al. 2016). An ARIMA model consists of three parameters – p, d, and q (ARIMA (p, d, q)) where p represents autoregressive (AR) lags, and q represents moving average (MA) lags. The parameter d is the integration order showing the number of times the time series must be differenced to make it stationary. Furthermore, seasonality in the data is also possible to handle including additional seasonal terms in the ARIMA models. The resulting model is termed as seasonal ARIMA model and is represented as,

SARIMA (p, d, q) × (P, D, Q) m

where m is the periodicity of the season or number of observations per season. Similar to ARIMA, the lowercase notations denote the non-seasonal part, whereas the uppercase notations denote the seasonal part. The model's seasonal part consists of terms similar to the model's non-seasonal components but there is lag throughout the periodicity.

Necessity for the Time Series to Be Stationary: It is utmost significant features of time series data. Stationary condition of time series is identified, when mean is constant along with variance; and the covariance is not time dependent. In simple terms, it can be defined as properties of time series do not depend on the time at which the series is recorded. Thus, time series with trends or with seasonality is not stationary.

If the series is non-stationary, they are very difficult to evaluate precisely. Non-Stationarity in the time series data results due to external actions too, such as promotion and campaigns, protests and strikes, holidays and festivals, which are required to be considered in the model development by distinctly representing existence of such incidences in the previous time series and then accordingly future events should be planned so that it can lead to a precise forecast for the future.

Otherwise, the finalised model may adopt the variations as part of the normal time series pattern and not as something produced by external actions and they transmit the impacts into the future values, finally cause false forecast values. Making time series stationary smoothens the variations resulting in more precise estimates and fewer errors.

The stationarity of a time series can be determined either by graphical and summary statistics or by statistical tests (Phillips, P. C., et al. 1988). The Augmented Dickey-Fuller (ADF.) Test for unit root is one such statistical test, which tests for the presence of a unit root that makes the time-series non-stationary. In this work, the ADF unit root test is used to determine the stationarity of the time-series and hence, effectively, decide the order of normal differencing, d, and the order of seasonal differencing, D (Dickey, D. A., et al. 2012). As the ADF test is basically a statistical significance test, more specifically hypothesis testing is used with null and alternate hypothesis, test statistics will be obtained as a part of result, and p-values are attained. An inference of p-value can be made as to whether a given series is stationary or not. As it is previously observed that the time series exhibits seasonality, the time series for all stations is made stationary by taking the first difference after the seasonal difference. Hence, the values of d and D are fixed as one.

The ADF test has its place in the group of 'Unit Root Test'. This test is basically used as the best method to test the stationarity of a time series. The unit root present in the time series is the typical characteristic of a time series, which generally makes data non-stationary. A unit root is supposed to occur in a time series of the value of $\alpha = 1$ in the below equation (4.2).

$$y_t = \alpha y_{t-1} + \beta \mathrm{X_e} + \epsilon \tag{4.2}$$

where y_t is the time series value at the time 't', and X_e is an exogenous variable (a separate explanatory variable, which is also a time series).

As mentioned before the existence of a unit root in the time series signifies the it as non-stationary. With respect to that, the total number of unit routes present in the series corresponds to the number of differencing are vital to make the series stationary.

A Dickey-Fuller test is a unit root test that tests the null hypothesis that $\alpha = 1$ in the following model equation (4.3). α is the coefficient of the first lag on y.

Null Hypothesis ($\mathrm{H_0}$): $\alpha = 1$

$$y_t = c + \beta_t + \alpha y_{t-1} + \phi \Delta Y_{t-2} + e_t \tag{4.3}$$

Whereas,

- $\mathrm{y_{t-1}}$ = lag 1 of time series
- ΔY_{t-2} = first difference of the series at time *(t-1)*

Fundamentally, it has a similar null hypothesis as the unit root test, i.e., the coefficient of Y_{t-1} is 1, implying the presence of a unit root. If not rejected, the series is considered to be non-stationary. The Augmented Dickey-Fuller test evolved based on the above equation and is one of the most common forms of the Unit Root Test.

As the name suggests, the ADF test is an 'augmented' version of the Dickey-Fuller Test. The ADF test expands the Dickey-Fuller Test equation to include a higher-order regressive process in the model shown in the equation (4.4).

$$y_t = c + \beta_t + \alpha y_{t-1} + \phi_1 \Delta Y_{t-1} + \phi_2 \Delta Y_{t-2} + \cdots + \phi_p \Delta Y_{t-p} + e_t \tag{4.4}$$

It is noticed, that only additional differencing terms, while the rest of the equation remains identical. This adds more diligence to the test carried out.

The null hypothesis is still the same as the Dickey-Fuller Test. However, since the null hypothesis assumes the presence of unit root, α=1, the p-value obtained should be less than the significance level (say 0.05) to reject the null hypothesis, thereby inferring that the series is stationary. If the p-value is greater than 0.05, the null hypothesis cannot be rejected, indicating the series is non-stationary. This study uses Python to conduct the ADF unit root test.

Identification of SARIMA (p, d, q) (P, D, Q) m model: Traditionally, the order of p, q, P and Q in SARIMA (p, d, q) × (P, D, Q) _m is obtained by observing Auto-Correlation Function (ACF) and Partial Auto-correlation Function (PACF) plots of the time-series, which are made stationary previously.

Auto-Correlation is used to know association between the time series data points at the present time spot and those at prior time spots. Just as association measures the degree of a linear relationship exists amongst two variables, hence autocorrelation function deals with establishing the linear relationship amongst lagged values of a time series.

Generally, while obtaining correlation amongst the data points at any two-time spots, observations of same data points at some other time spots is considered. For example, today's stock price correlated with yesterday and yesterdays with day before yesterdays. Hence, PACF of the yesterday is the real correlation between today and yesterday after taking out the influence of the day before yesterday.

However, with the increasing computing power of modern processors, many models can be fit in a minimal amount of time. Hence, all possible permutations of p, q, P, and Q are subjected to the range given in Table 4.2 and used to forecast from 12th February 2020 to 29th February 2020. The advantage of this method is that there is a better chance of finding the best model than the traditional method.

Table 4.2 Ranges of order of Parameters for Seasonal ARIMA

Time-frame	***p***	***q***	***P***	***Q***	***m***
15 minutes	0 to 68	0 to 68	0 to 2	0 to 2	68
30 minutes	0 to 34	0 to 34	0 to 2	0 to 2	34
45 minutes	0 to 23	0 to 23	0 to 2	0 to 2	23
60 minutes	0 to 17	0 to 17	0 to 2	0 to 2	17

Measures for Accuracy of Fitting: There are many parameters or criteria that measure this accuracy of fit and indicate which models are appropriate for the given time series. The important among those are mentioned below.

The fit models' values of Akaike Information Criteria (AIC) and Bayesian Information Criterion (BIC) and the errors for the forecasted model are obtained, and then MAE and RMSE are also calculated and tabulated. AIC, BIC, MAE, and RMSE are defined and explained in the subsequent paragraphs.

AIC and BIC are the measures that measure the accuracy of fit of the models. AIC is defined as in the equation (4.5),

$$AIC = T \log\left(\frac{SSE}{T}\right) + 2(k+2) \tag{4.5}$$

where T is the number of observations used for finding best fit model and k is the number of predictors in the model. Similarly, BIC is defined as in the equation (4.6),

$$BIC = T \log\left(\frac{SSE}{T}\right) + (k+2)\log(T) \tag{4.6}$$

Where T and k have the same meaning as in AIC. The model with minimum AIC and BIC values is the best forecasting model. If an actual underlying model is present, the

BIC tends to select that model given enough data. Still, if the number of observations included (T) is large enough, both AIC and BIC select the same models. Hence, the model chosen using the BIC is either the same as that chosen using the AIC or the one with fewer terms, as the BIC penalizes the number of parameters more heavily than the AIC (Hyndman, R. J., et al. 2018). However, it is preferable to use both AIC and BIC in combination, giving equal importance to both (Kuha, J., et al. 2004).

The difference between actual and forecasted values is known as forecast error. In this case, the term error does not imply a mistake but indicates an observation's unpredictable part (Hyndman, R. J., et al. 2018). It is given by equation (4.7),

$$e_{T+h} = y_{T+h} - \hat{y}_{T+h|T} \tag{4.7}$$

where the training data is given by $\{y_1, \dots, y_T\}$ and the test data is given by $\{y_{T+1}, y_{T+2}, \dots .\}$.

MAE and RMSE are two of the most widely used measures to represent errors. The definition of both MAE and RMSE are given below in the equation (4.8) and (4.9).

$$MAE = mean(|e_t|) = \frac{1}{n} \times \sum_{t=1}^{n} |e_t| \tag{4.8}$$

$$RMSE = \sqrt{mean(e_t^2)} = \sqrt{\frac{1}{n} \times \sum_{t=1}^{n} e_t^2} \tag{4.9}$$

MAE is slightly more intuitive, easier to interpret and compute than RMSE. Optimising the forecasts for MAE results in forecasts of the median and optimising the forecasts for RMSE results in mean forecasts. Moreover, RMSE is more sensitive to outliers than MAE (Vandeput, N., 2021). Suppose a model is selected solely based on the accuracy of fit measures such as AIC and BIC, it is possible that the selected model is underlying model and provides a reliable characterization of the sources of uncertainty and understands the underlying data-generating mechanism (Ding, J., et al. 2018). However, this does not guarantee the selection of a model with high forecasting accuracy.

Suppose a model is selected based only on the accuracy of forecasting measures such as MAE and RMSE tends to select a model with excellent predictive performance but may not select a model which captures the actual underlying characteristics. Therefore, the final model was selected, considering all of AIC, BIC, MAE, and RMSE, giving equal importance to all four measures to choose the robust models and have excellent forecasting accuracy.

The residual analysis was carried out on the selected final models to ensure that the model captures and utilizes most of the features, patterns, and information available in the data provided. The residuals in a time series model are what is left over after the model fit process (Hyndman, R. J., et al. 2018), or more precisely, residuals are the difference between the observations and the corresponding best fit values and shown in the equation (4.10):

$$e_t = y_t - \hat{y}_t \tag{4.10}$$

Where, e_t is the residual, y_t the actual value and $\hat{y}_t$ the corresponding best fit value. As a part of the residual analysis, the residual graph, ACF and PACF plots, histogram of residuals are plotted and observed for the following properties:

- The residual plots looked like white noise.
- There was no correlation between the residuals.
- The residuals were normally distributed.

Suppose the residual plot looks like the white noise (a series with zero mean, constant variance and zero correlation). In that case, it indicates that the model successfully captured all the information present in the data. The ACF and PACF plots and the histogram of residuals are plotted to reinforce the conclusion obtained by observing the residual plots. If the residual analysis results are satisfactory, then the selected final model is adopted, else a different model is chosen, and the same steps are iterated. Python code used for detailed SARIMA analysis is shown in Appendix A.11.

4.3.3.2 Forecasting by Long Short-Term Memory (LSTM):

Artificial Neural Networks (ANN) are a set of computing units that simulate how human brains analyse and process information. They have self-learning capabilities, i.e., there is no need to program everything. Recurrent Neural Networks (RNN) are a type of neural network suitable for processing sequential data such as time series, language, and speech. However, these RNNs have a drawback. Though in principle, RNNs are capable of learning long term dependencies, in practice, they cannot pick up these due to the vanishing gradient problem (Hochreiter, S., 1998).

LSTM is a special type of RNN, proposed to solve this problem of long-term dependency (Hochreiter, S., et al. 1997) and the vanishing gradient problem. All RNNs, including LSTMs, have a chain of repeating modules of neural networks. The repeating module in standard RNNs is a simple structure containing a single layer, such as a single tanh layer. However, in LSTMs, the repeating module has four layers interacting in a very special way instead of a single neural network layer. The structure of an LSTM unit is given in Figure 4.5 (Phi, M., 2018).

The sigmoid function here outputs a value between zero and one for all the input values. The output value describes the amount of information that must pass through the gate. More precisely, the value of zero indicates no information is passed through, and the value of one indicates everything is passed. As previously mentioned, the cell state is controlled by three gate layers: forget gate layer, input gate layer and output gate layer. The forget gate layer decides how much of the previous cell state should be kept. The previous hidden state and information from the current input are fed to the forget gate layer and passed through the sigmoid function present there. As described earlier, this sigmoid function outputs a value ranging from zero to one, indicating remembering completely to forgetting completely.

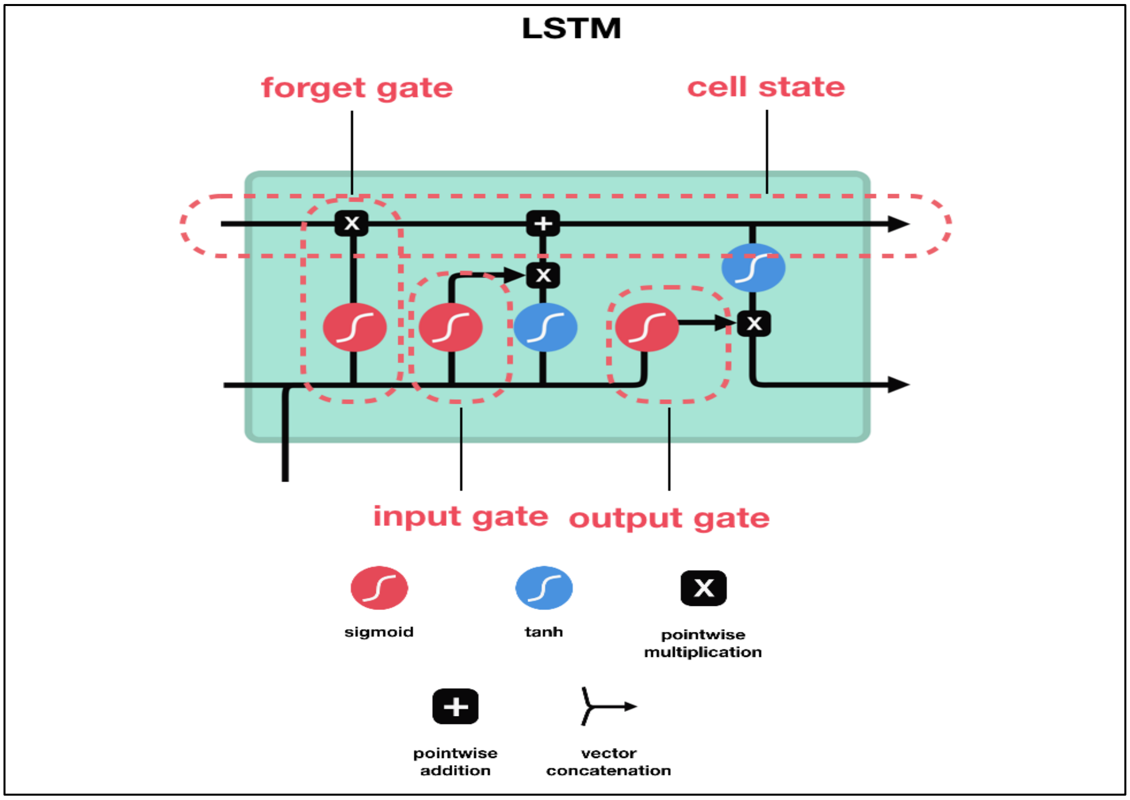


Figure 4.5 Structure of an LSTM unit

The previous hidden state and the information from the current input are again fed to the input gate. The input gate layer decides which values should be updated for the new cell state. Furthermore, the same previous hidden state and the information from the current input is fed to the nearby tanh layer too, which creates a vector of new candidate values that could be added to the cell states. The outputs from these two layers are combined using a pointwise multiplication operation. Then to update the cell state actually, the old cell state is multiplied by the output of the forget gate layer and added with the combined output from the input gate layer and the nearby tanh layer. Finally, the output gate layer decides the next hidden state. Again, the previous hidden state and the information from the current input are fed to the output gate, and the new (or updated) cell state obtained previously is passed through another nearby tanh layer. The outputs from these two layers are combined using a pointwise multiplication operation to obtain the new hidden state as output (colah's blog 2015).

In this work, the same passenger data that is used for forecasting using seasonal ARIMA earlier was used to forecast with LSTM. The training and testing sets used are also identical for both cases. The LSTM is implemented here using Anaconda Python distribution with Python 3.8.5, and the IDE (Integrated Development Environment) used is Jupyter Notebook 6.1.4 and Keras APIs (Application Program Interfaces). However, it is then slightly modified to suit the needs of LSTM. Then, time-series generator is used to organize the training data into a suitable format. A look-back of the last 15 days and a batch size of 16 is specified. A stacked LSTM model with two LSTM layers each consisting of 100 units and a dense layer with a single node, is used. The activation function used is tanh, and a learning rate of 0.0004 was specified with adam optimizer. The model is trained with mean absolute error as a loss function for 75 epochs. All these parameters and hyper-parameters are fixed based on a trial-and-error procedure.

The trained LSTM model was made to forecast the same as seasonal ARIMA in the testing period (from 12th February 2020 to 29th February 2020). The forecasted values are then transformed to the original scale using inverse transforming functions. The obtained values are then compared with the actual values in the testing set. Finally, MAE and RMSE (defined and explained in the previous section) are calculated. Python code used for the forecasting analysis using LSTM given in the Appendix A.12.

4.3.4 Level of Service (LOS) of the HDBRTS based on TTR

The development of a Level of Service (LOS) is the most vital requirement when there is a need to make a service comparison and establish thresholds for required reliable service from the transit system (Uno, N., et al. 2009). Thus, as the fourth objective of the current research work, service reliability indices-based LOS has been established for the HDBRTS. Figure 4.6 shows the detailed methodology adopted for LOS development in current research work.

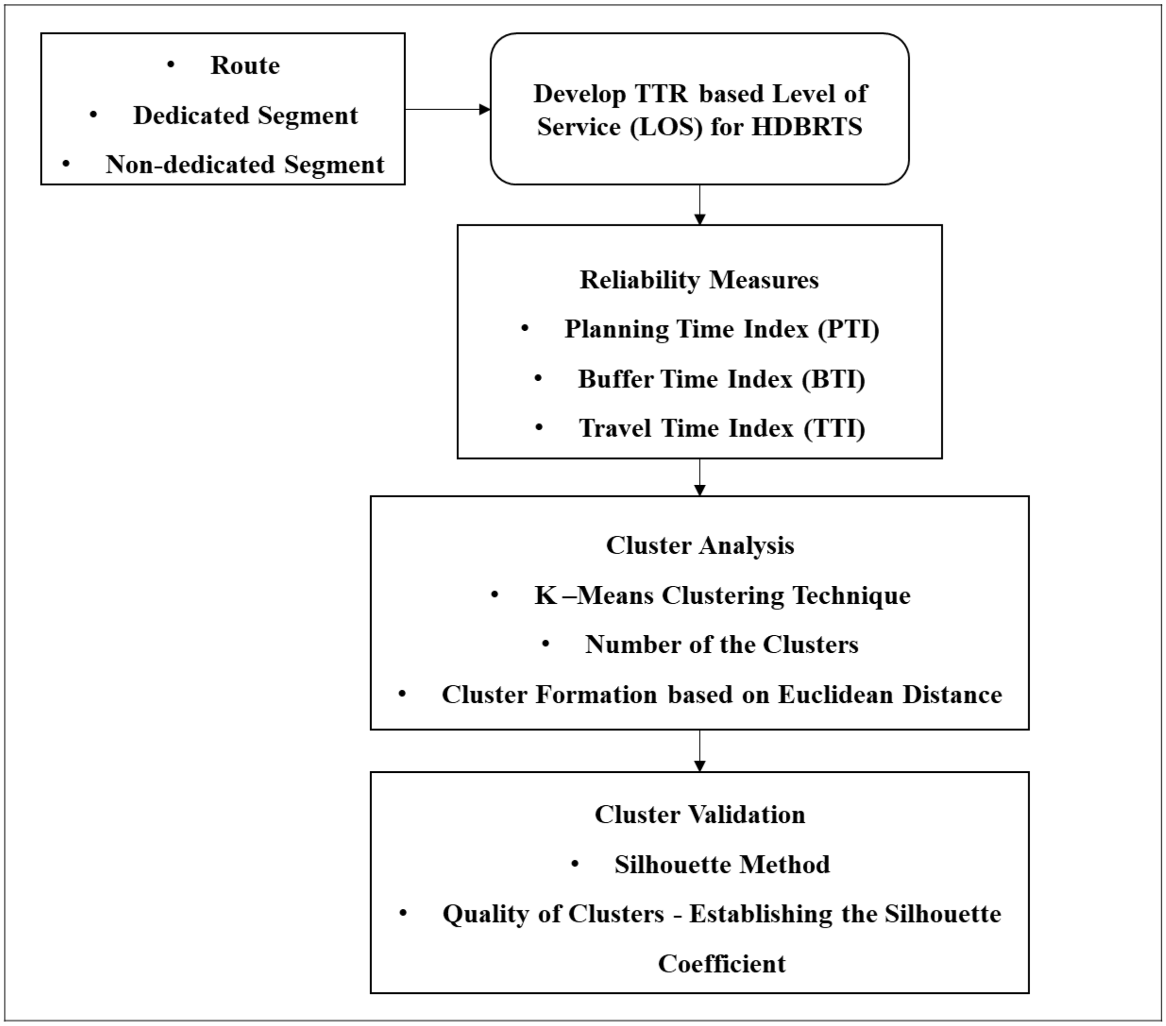


Figure 4.6 Methodology for Developing Level of Service (LOS)

In the processing of developing the LOS, three main operating conditions of the buses have been taken into the consideration. The first condition is at route level, which provided an operating facility for both the express and non-express buses. The second condition considered is the exclusive bus operating environment, which has a segment with a dedicated lane. In the third condition, the non-dedicated nature of bus operation is considered. Routes and segments which have been taken in the development of LOS are the same as what have been taken in the variability analysis of transit travel time. From the previous literature background as well as the variability study carried out on transit travel time as a part of current research work, it has been noticed that, various reliability measures exist for transit travel time analysis are the most appropriate for apprehending the travel time variations (Chepuri, A., et al. 2018). Hence, in this research work, LOS criteria has been developed for routes as well as segments founded on the three potential reliability indices of transit travel time viz. travel time index (TTI), planning time index (PTI) as well as buffer time index (BTI).

LOS development for the route and segments have been carried out in three stages. Such as route-wise and segment-wise agglomeration of the travel time reliability data points, cluster formation and analysis considering the K-mean clustering technique, and validation of framed clusters based on the Silhouette coefficient value.

In the first stage, all the three identified reliability indices belonging to different hours/periods and different days, have been grouped together as per the plan of analysis. Total 90 days of travel time data points route, dedicated segment, and non-dedicated segment have been utilized for LOS development. As the calculated data points are behaving the same in the UP and DOWN direction of route and segments, hence both the direction values are taken together in framing analysis-ready data sheets. Individual datasheets for all three reliability indices of routes, PTI, BTI, and TTI have been prepared and kept ready for the cluster formation analysis based on the K-mean clustering technique. Coding used in the analysis is given Appendix A.13.

4.3.4.1 Cluster Formation

K-mean clustering is used to categorize the PTI, BTI, and TTI individually for routes and segments and so each category represents different service levels. The concept of the working principle of the K-mean clustering technique is mainly dependent on reducing the distance of cluster mean value with individual data points taken in the analysis and then each cluster is formed (Chepuri, A., et al. 2018). This reduced distance is generally termed Euclidean distance. Algorithms for the analysis have been written using python programming and made the program run separately for all three datasheets of route and segments.

Hence, totally there are nine input and output files while carrying out this particular study. Based on the K-mean clustering analysis, all the three reliability indices are then grouped into six individual clusters for PTI, BTI, and TTI of the route, dedicated segment, and non-dedicated segment. Selected reliability indices are categorized into six clusters between LOS A to LOS F based on previous literature background and each cluster framed denotes each service level pertaining to the selected reliability indices in the current study. Six clusters that have been considered in the current analysis are universally approved service levels (Kathuria, A., et al. 2020). After the formation of successful clusters, they are validated for their quality of formation with different measures. Here the current study is done with Silhouette Coefficient-based technique and is explained in a subsequent section.

4.3.4.1 Cluster Validation

As mentioned in the previous paragraph, cluster validation is the process that is carried out to assess the quality of the already framed clusters after their formation stage. In the current research work, silhouette coefficient-based cluster validation is carried out. This method measures the deviation of all data points present in their particular cluster and in other clusters where those data points are absent. Basically, this method recognizes the dissimilarity of every data set and is represented by its coefficient value. A value obtained from every data group indicates the wellness of all data points allocated in their particular cluster formed and how those data points are unlike when compared to data points of some other cluster (Rousseeuw, P. J. 1987). The range of silhouette coefficient is generally taken in between -1 and +1. Higher the coefficient, better the cluster, and the lower the coefficient the slackly the framed cluster. Usually, if the average coefficient value of a particular cluster falls more than 0.5 then that cluster is taken as a realistic or reasonable structure. The formula to calculate the silhouette coefficient, S (i) is calculated as follows in equation (4.11) (Rousseeuw, P. J. 1987).

$$S(i) = \frac{e(i) - f(i)}{max\{e(i), f(i)\}} \tag{4.11}$$

$$S(i) = \begin{cases} 1 - \dfrac{e(i)}{f(i)}, & if\ e(i) < e(i) \\ O, & if\ e(i) = f(i) \\ \dfrac{f(i)}{e(o)} - 1, & if\ e(i) > f(i) \end{cases}$$

i = Individual data points

e (i) = average difference of i^{th} data point with other data points in their cluster

f (i) = average difference of i^{th} data point other data points of some other cluster.

As mentioned previously quality of the cluster evaluated based on average S(i) of whole individual cluster and different coefficient values make different interpretation of the obtained values. Table 4.3 shows indications of different ranges of silhouette values.

Table 4.3 Indications of different ranges of silhouette values

S(i) range	Indication
0.71 to 1.0	Robust Cluster
0.51 to 0.71	Reasonably strong Cluster
0.26 to 0.50	Weak and Artificial Cluster
<0.25	No substantial Cluster

4.4 SUMMARY

Current chapter has made detailed discussion on study area considered in the current study, kind of data used in the analysis, data processing strategies have been carried out and finally comprehensive methodologies implemented for all the framed objectives with relevance to performance analysis of Hubli-Dharwad Bus Rapid Transit System.

Study has taken the recently implemented and still in the expansion stage of the public transit system (HDBRTS) for study that is located in the northern part of

Karnataka. For the current research work Automatic Vehicle Location data (AVL) from 8th December 2019 to 29th February 2020 (Total 84 Days), Automatic Passenger Count Data (APC) from 1st December 2019 to 29th February 2020 (Total 91 Days) is considered as a data source. Later, the systematic travel time variability analysis has been carried out for the routes, segments at different temporal patterns and accordingly methodology has been shown in the chapter.

Operators and passenger perspective travel time reliability models have been developed for the segments using multiple linear regression (MLR) techniques, and the detailed methodology adopted is clearly given in this chapter. Attempt is made to explore the unexplored area of LSTM in passenger demand forecasting using APC data and obtained results are compared with results of traditional SARIMA method.

Finally reliable transit service-based level of service of HDBRTS have been established for the three different operational conditions of the buses. K-mean clustering for cluster formation and silhouette method for cluster validation are used and pertaining to that procedure has been given in this chapter. Connecting to all the methodologies of objectives, obtained results are showed with tables and figures in the subsequent chapter with detailed discussion.

CHAPTER 5

RESULT AND DISCUSSION

5.1 GENERAL

This chapter illustrates the results of all the objectives framed in the current research work, such as travel time variability study, modelling the travel time reliability, modelling the passenger's demand, and establishing the LOS of the system. To proceed systematically, a subsequent detailed discussion on the results of individual objectives is made as a continuation of the chapter.

5.2 TRAVEL TIME VARIABILITY STUDY OF THE SYSTEM

Travel Time variability analysis of the HDBRT is carried out in two stages. In the first stage, different descriptive statistics values of travel time and travel time reliability indices as per Federal Highway Administration 2006 are calculated and represented in the form of a table. Then in the second stage, travel time variability analysis is carried out through probability distribution analysis for routes and segments separately. Finally, results are shown in the form of tables and figures.

Prior to the actual analysis, it is essential to comprehend the behaviour of the travel time at different hours of the day or periods of the day. As in the current study, peak hours and peak periods are mainly selected based on the passenger demand data throughout the day along with these travel time patterns. Hence, the pattern of the travel time variability to varying hours of the day mainly helped in subsidizing the demand data in deciding the peak hours and off-peaks of the day. Figure 5.1 shows the within day variation in passenger demand of one of the important segments of the non-express route. In the figure X-axis represents the passenger demand and Y-axis represents the different hour of the day. Here it is very clear that, week days are having more passenger demand compared to weekend. And, early morning hours between 5:00 to 7:00 and evening hours between 20:00 to 22:00 are measured with minimal passenger demand. The remaining time of the day there is a higher fluctuating passenger demand both in weekdays and weekends.

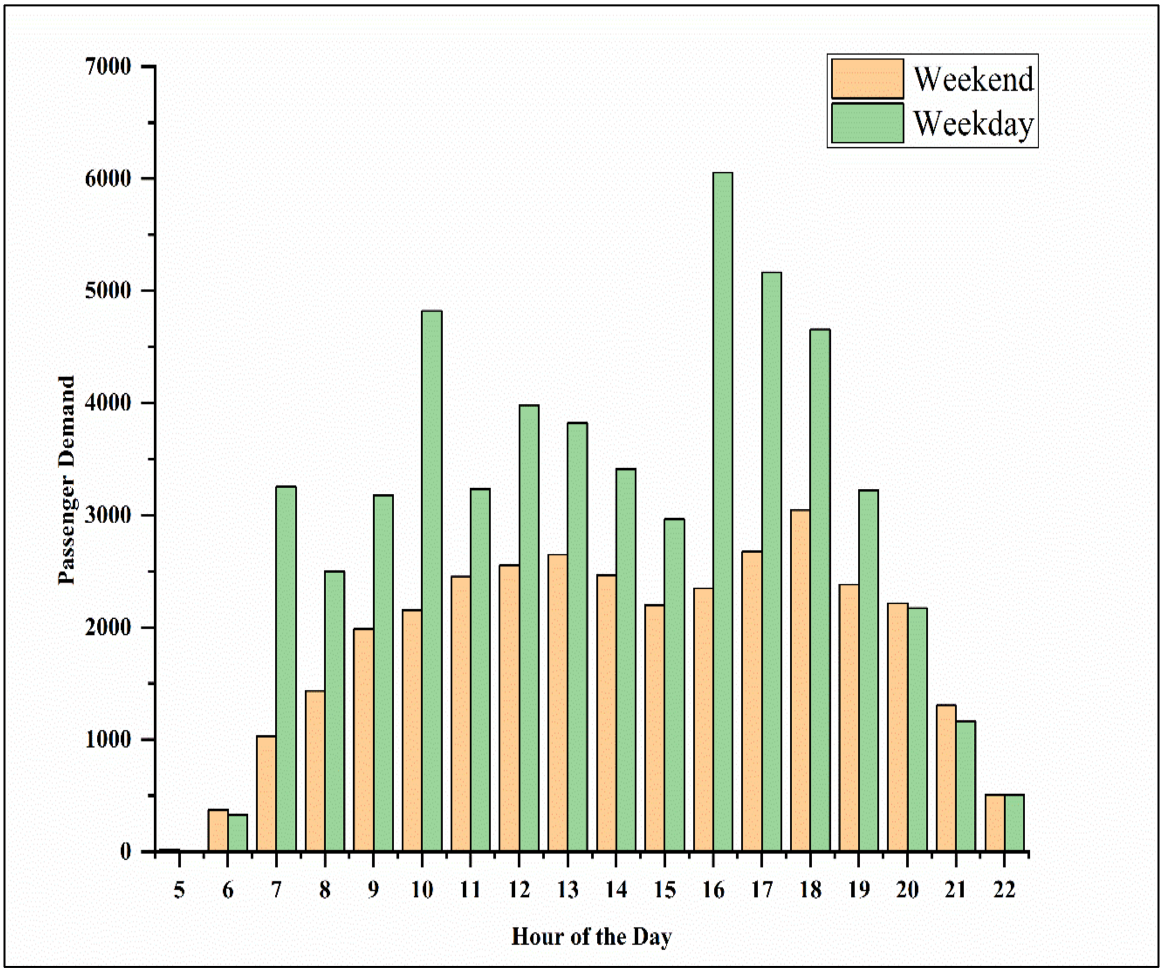


Figure 5.1 Within day Passenger Demand Variation

Figure 5.2 shows the travel time variability pattern along the day for the express and non-express routes. To study the travel time variation along the day, average travel times are plotted which are collected from routes wise trips. In the figure X-axis represent the different hour of the day from 5:00 to 22:00, 5:00 in the figure is the hour between 5 AM to 6 AM, and 22:00 is the hour between 10 PM to 11 PM, respectively. Y-axis represents the variation in the average travel time at different hours of the day. Travel time data in the analysis is considered both in the UP and DOWN directions; hence, the TT variation plot is also plotted accordingly.

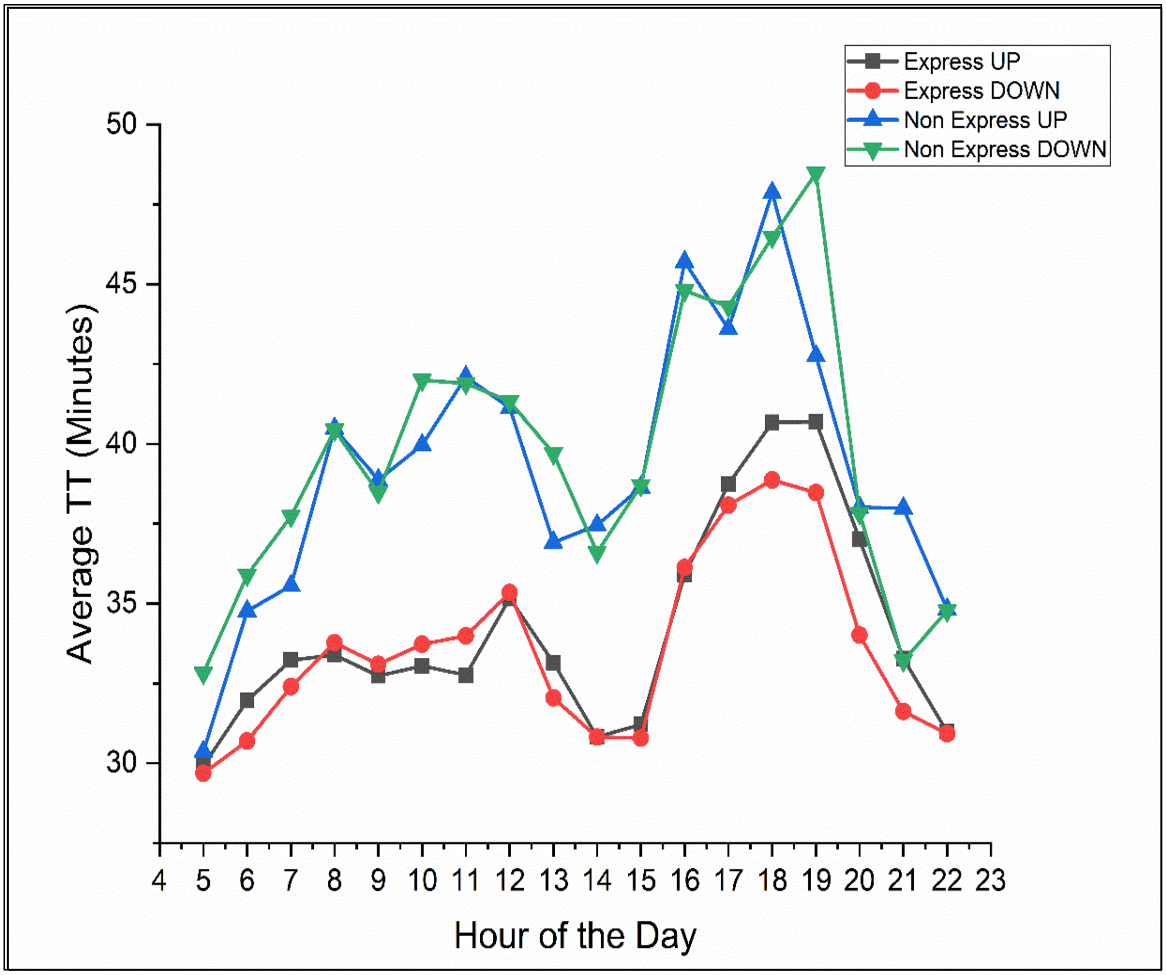


Figure 5.2 Travel Time Variation Pattern

In figure 5.2, it is clearly seen that travel time in the morning hour between 5:00 to 7:00 and in the evening hours between 20:00 to 22:00 are lower and hence considered morning and evening off-peak hours, respectively. Similarly, travel time between 8:00 to 10:00 and 16:00 to 19:00 seems higher; accordingly, they are identified as morning and evening peak hours correspondingly. But travel time variation between 11:00 to 15:00 is relatively lower than peak hours and higher than off-peak hours; hence these hours are identified as inter-peak hours. Travel time variation over the space and time for both the UP and DOWN directions follow a similar trend and have been analysed together.

Box plots have been plotted to study the distribution of travel time over all the weekdays and weekends. Figure 5.3 represents the box plot for the express route, and figure 5.4 represents the box plot of the non-express route.

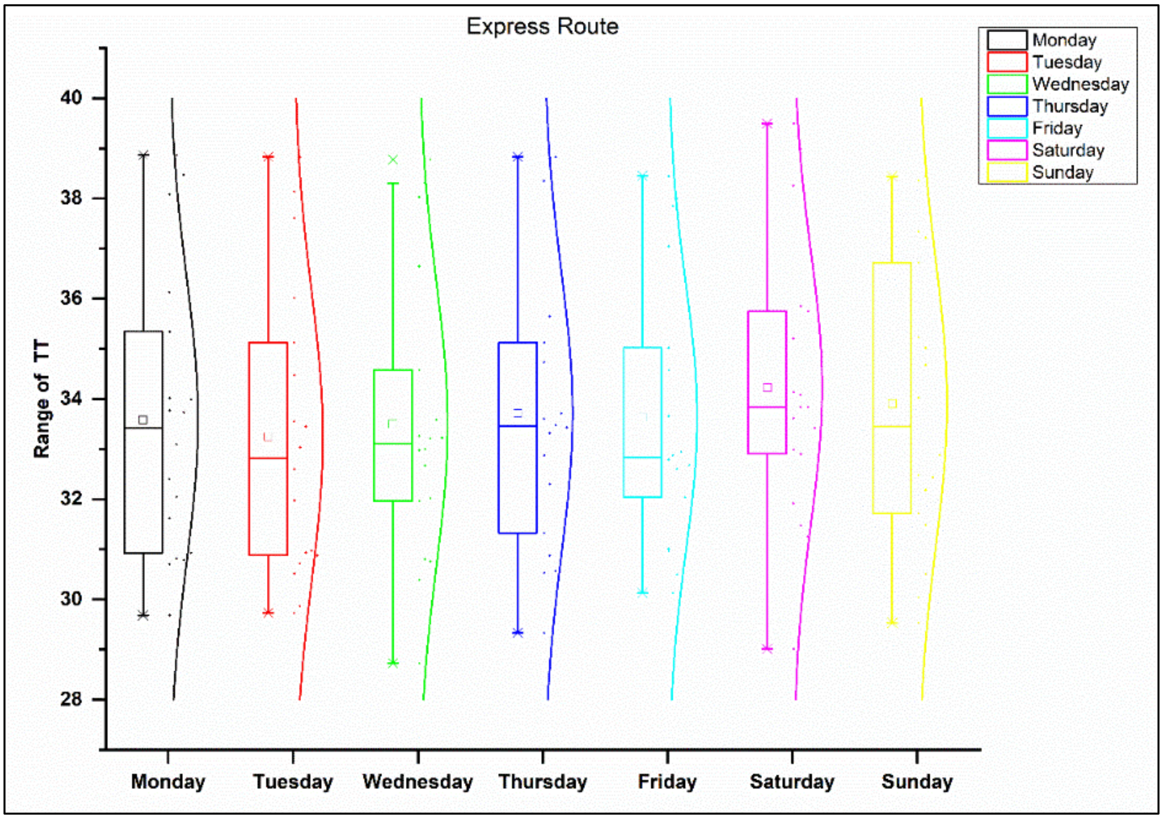


Figure 5.3 Travel Time of Spread over the Week Days and Weekends - Express Route

From figure 5.3, for the express routes, it is observed that the spread of travel time over all the weekdays and weekends is behaving similarly. In the case of non-express routes of figure 5.4, weekdays have a similar spread in the travel time, whereas only Sunday shows a lower spread in the travel time.

It is presumed in the study that Hubli known for many commercial activities in Karnataka. Hubli Railway Station Junction is also located in the South Western Railway district. It has excellent connections to all the cities in Karnataka and neighbouring state cities such as Mumbai, Hyderabad, Goa, and others. Consequently, similar passenger demand is observed on all the days of the week between terminal to terminal for the express routes. Hence travel times trends are also observed to be almost identical on all the days of the week as shown in the figure 5.3. However, in the case of non-express routes, it is seen that passenger demand in-between stations of terminals are less on weekends compared to weekdays and impact of which can also be noticed in the TT spread on Sunday in figure 5.4 is completely different than express route.

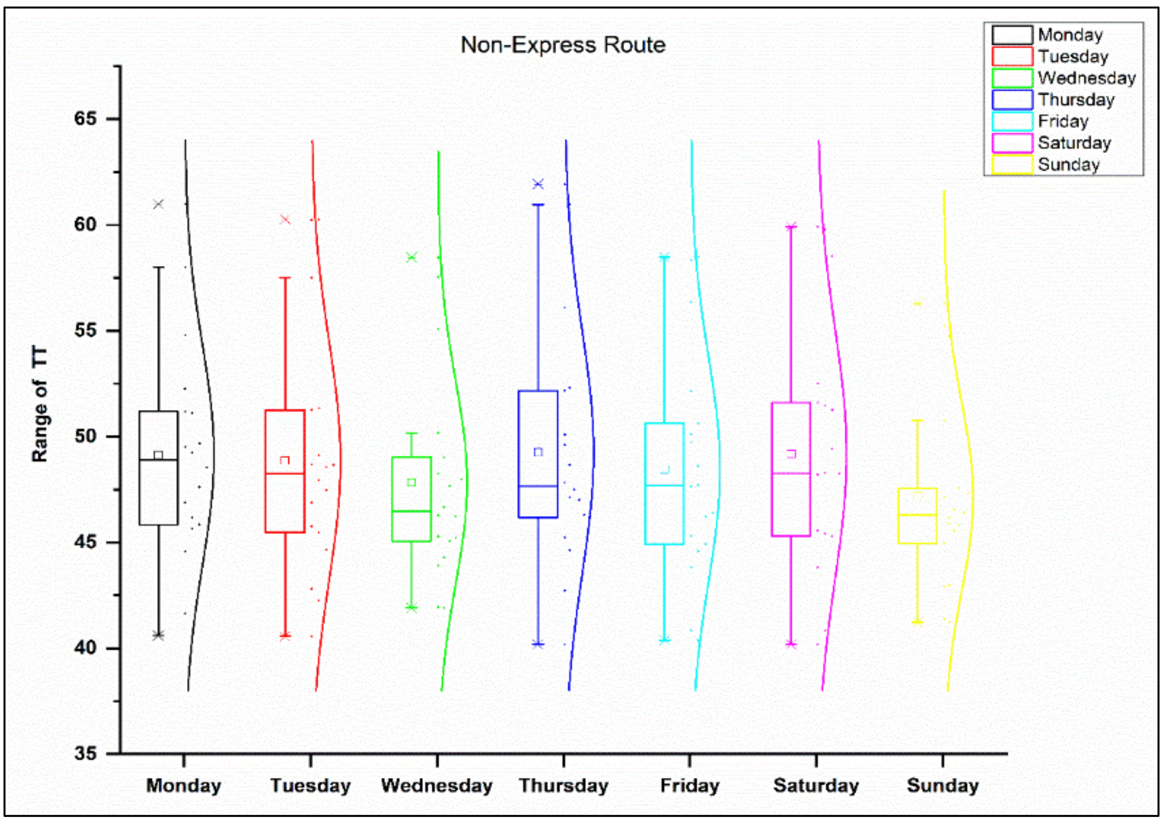


Figure 5.4 Travel Time of Spread over the Week Days and Weekends - Non-Express Route

One more important observation from figure 5.3 and 5.4 is that there is a dissimilar pattern of TT spread exists on Wednesday compared to all other weekdays. This mainly due to, as the CBD area in Hubli is closed on every Wednesday instead of any other weekend. Hubli is being second largest city in Karnataka next to Bangalore and serves as the corporate and commercial hub of Northern Karnataka. Most of the business centers, shopping centers, commercial streets are located in the CBD area of Hubli. Even many people from Dharwad, other side of the HDBRTS terminal, travel daily to run their business setup there in the Hubli city itself, or for any other activity related to shopping, recreation, etc. It is evident that, on Wednesday, most regular passengers will not travel between Hubli and Dharwad. Hence it is presumed that this may be a solid reason behind least spread in TT of Wednesday.

Travel time plots are subsidizing the passenger demand data in deciding the off-peak hours or periods, peak hours or periods and inter-peaks hours or periods. Also, they helped in clearly understanding the travel time pattern on all the days. Based on the overall interpretation gathered from the plots, the travel time data points of UP and DOWN directions have been grouped and analysed together. For the route level analysis, travel time data points are classified as peak hour, off-peak hours, and interpeak hours. Also travel time data points separately for weekdays and weekends. Three segments are considered for the segment-level analysis, and they had a dedicated and non-dedicated nature of bus operations. As obtained, travel time data points in the segments level analysis are relatively lesser than route level; hence they are classified into period-wise data points instead of hour-wise separately for weekdays and weekends and then analysed.

5.2.1 Travel Time Variability Study based on Descriptive Statistics

As mentioned previously TTV analysis has been carried out route level and segment level separately and it is as below.

5.2.1.1 Route Level Analysis

Descriptive statistical analysis for the express and non-express routes is done separately for the weekdays and weekends and represented in table 5.1, table 5.2, table 5.3 and table 5.4 respectively. Results of remaining week days given in the Appendix between A.15 to A.28. The main purpose of carrying out this analysis is to comprehend the variation in the parameters and indices of routes as well as segments with the change in the temporal and spatial aggregations. A total of 12510 travel time data points have been considered in establishing the descriptive statistics parameters and reliability indices. Average travel time, SD of travel time, CV of travel time, PTI, BTT, and TTI are the parameters and indices that have been considered; a detailed explanation has been made in section 4.4.1. Free flow travel time (FFTT) is the lowest travel time where the impact of various traffic incidences and passenger demand is less (FHA, 2006). Obtained FFTT for different routes and segments given in the Appendix A.8. In the case of HDBRTS, the travel time of buses between 5:00 and 6:00 is usually lower, and passenger demand is also lower. Hence, to calculate travel time indices, namely. TTI and, PTI, FFTI is essential and have been calculated considering travel time data points of buses that run early in the morning. For better visualisation in the variation of PTI, BTI and TTI values for express and non-express routes given in the form of figure 5.5 and figure 5.6.

From the tables 5.1, table 5.2 and figure 5.5 of the express route following inferences are made

- Minor variations are observed from the computed values of descriptive statistics parameters and reliability indices concerning weekdays and weekends. Average travel time values are lower during morning and evening hours of bus operation between 5:00 to 7:00 and 20:00 to 22:00. Subsequently, variation in the travel time of express buses is seen to be less, as it can clearly represented with the values of CV of travel times in the table, and those hours are referred previously as off-peak hours in the study. Meanwhile, remaining hours are considered as peak hours, viz. 8:00 to 11:00, 12:00 to 15 and 16:00 to 19:00 as those hours are having less variation in the CV of TT, meantime average TT values of peak hours moved towards the higher side.
- SD of travel time in the tables explains the dispersion of data points at different hours of the day. Off-peak hours have higher SD than peak hour values and are relatively on the lower side. With the lower value of SD and higher values of travel time during

peak hours, it is understood that impact of delay caused at intersections and bus stops are more during peak hours than off-peak hours.

- 95^{th} percentile travel time and average travel time together have shown significant impact through higher BTI values, where as individual effect of 95^{th} percentile TT and ATT is seen in the values of PTI and TTI respectively and variations in these two values is less compared to BTI.
- Variations in average travel time between the higher and lower sides have also impacted the values of TTI, BTI, and PPI. Travel Time Index is the average additional time required for a trip during peak times compared to that trip duration in no-traffic conditions, and PTI shows the total time needed for an on-time arrival in 95 percent of all trips.

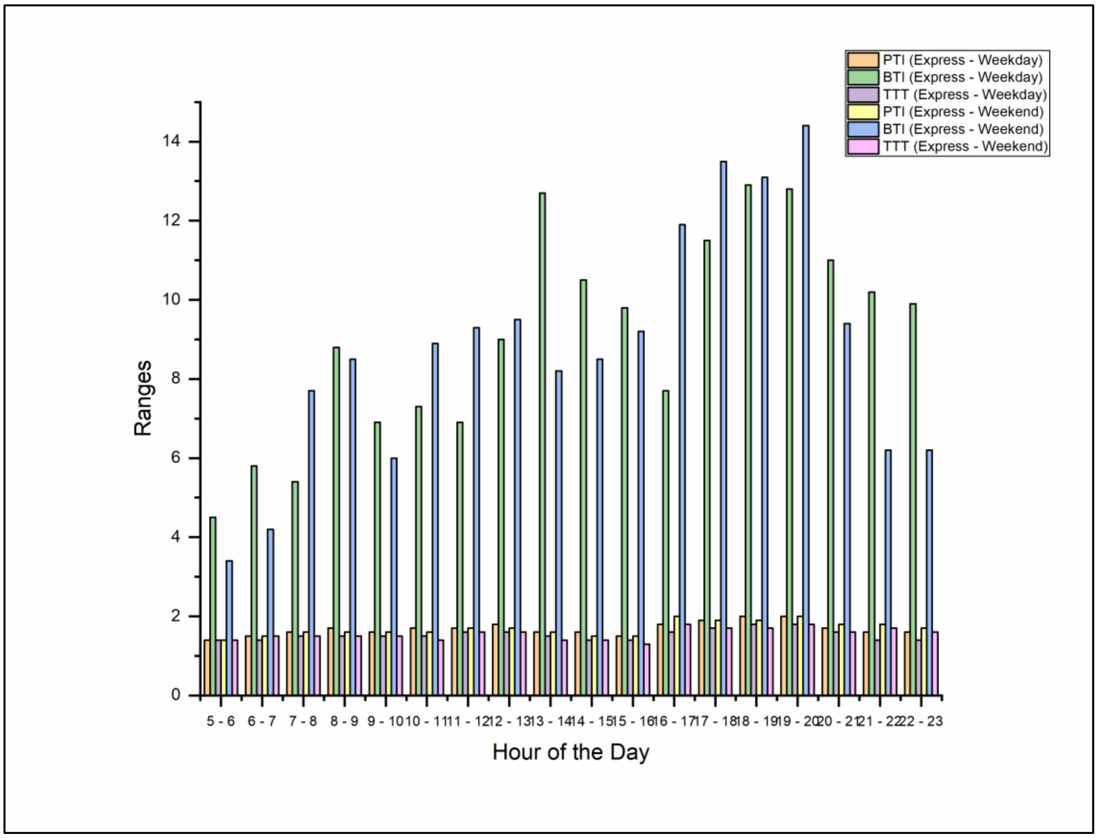


Figure 5.5 Variations in the Reliability Measures for Express Route

Table 5.1 Descriptive Statistics of Express Route – Weekday

Hour of the Day	Avg. TT (minute)	SD of TT	CV of TT	95th Percentile TT (minute)	PTI	BTI	TTI
5 - 6	29.7	2.61	8.79	31.0	1.4	4.5	1.4
6 - 7	30.7	2.55	8.30	32.5	1.5	5.8	1.4
7 - 8	32.4	2.16	6.68	34.1	1.6	5.4	1.5
8 - 9	33.8	2.46	7.27	36.7	1.7	8.8	1.5
9 - 10	33.1	1.54	4.66	35.4	1.6	6.9	1.5
10 - 11	33.7	1.33	3.94	36.2	1.7	7.3	1.5
11 - 12	34.0	1.02	3.00	36.3	1.7	6.9	1.6
12 - 13	35.3	1.62	4.58	38.5	1.8	9.0	1.6
13 - 14	32.0	2.19	6.83	36.1	1.6	12.7	1.5
14 - 15	30.8	2.73	8.86	34.0	1.6	10.5	1.4
15 - 16	30.8	2.34	7.61	33.8	1.5	9.8	1.4
16 - 17	36.1	1.85	5.11	38.9	1.8	7.7	1.6
17 - 18	38.1	1.42	3.74	42.5	1.9	11.5	1.7
18 - 19	38.9	1.43	3.69	43.9	2.0	12.9	1.8
19 -20	38.5	1.17	3.03	43.4	2.0	12.8	1.8
20 - 21	34.0	1.35	3.97	37.7	1.7	11.0	1.6
21 - 22	31.6	2.19	6.94	34.8	1.6	10.2	1.4
22 - 23	30.9	2.22	7.18	34.0	1.6	9.9	1.4

Table 5.2 Descriptive Statistics of Express Route – Weekend

Hour of the Day	Avg. TT (minute)	SD of TT	CV of TT	95th Percentile TT (minute)	PTI	BTI	TTI
5 - 6	30.0	2.62	8.72	31.1	1.4	3.4	1.4
6 - 7	32.5	2.57	7.92	33.9	1.5	4.2	1.5
7 - 8	32.2	2.77	8.61	34.7	1.6	7.7	1.5
8 - 9	32.4	1.78	5.48	35.2	1.6	8.5	1.5
9 - 10	32.9	1.94	5.89	34.9	1.6	6.0	1.5
10 - 11	31.7	1.20	3.79	34.5	1.6	8.9	1.4
11 - 12	34.0	0.98	2.89	37.2	1.7	9.3	1.6
12 - 13	34.7	1.03	2.97	38.0	1.7	9.5	1.6
13 - 14	31.5	1.13	3.58	34.1	1.6	8.2	1.4
14 - 15	30.5	2.55	8.35	33.1	1.5	8.5	1.4
15 - 16	29.5	2.04	6.91	32.2	1.5	9.2	1.3
16 - 17	38.4	1.31	3.41	42.9	2.0	11.9	1.8
17 - 18	37.3	1.28	3.41	42.4	1.9	13.5	1.7
18 - 19	37.2	1.75	4.69	42.1	1.9	13.1	1.7
19 -20	38.4	2.07	5.40	44.0	2.0	14.4	1.8
20 - 21	35.2	2.11	6.00	38.6	1.8	9.4	1.6
21 - 22	36.7	2.91	7.92	39.0	1.8	6.2	1.7
22 - 23	35.0	2.98	8.52	37.2	1.7	6.2	1.6

- Buffer Travel Index is the additional time (buffer) that most travellers add to their average travel time when planning their trips. Higher BTI values are observed during augmented average travel times; meanwhile, there are lower SD values of the travel times during peak hours there is less dispersion in the travel time of the buses.

Compared to off-peak hours**,** around 64.8 % increased buffer BTI values have been observed during peak hours. Also, some of the evening peak hours have shown higher side BTI values compared to morning off-peak hours, which is assumed mainly due to reduced bus frequencies and not due to increased travel times of buses.

- Overall, peak and off-peak hours have direct influence on the change in characteristics of travel time reliability indices considered in the current study.
- Except for the higher values of reliability indices during peak hours, performance of the express routes seems to be reliable considering total travel time that buses have taken to complete the whole length of the route.
- However, lower values of reliability indices give rise to scope for enhancing the system's performance.

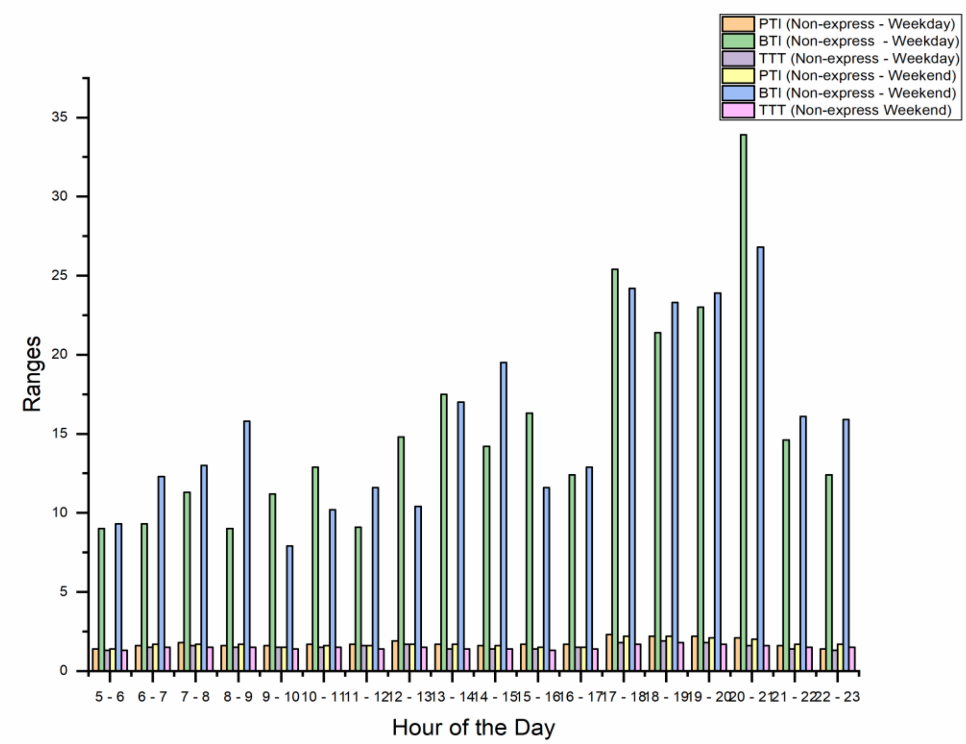


Figure 5.6 Variations in the Reliability Measures for Express Route

From the tables 5.3, table 5.4 and figure 5.6 of the non-express route following inferences are made,

- Variations in the descriptive statistics and calculated reliability indices values have been seen in the weekdays and weekend-wise analysis.
- Even though distinct behaviour in the values of data points of weekdays and weekends, the trend of obtained results remains the same as it is there in the case of express routes.
- Trend in the values of peak hours is referred to be the same as that of express routes; such as higher average travel times compared to off-peak hours, lower spread in the travel times seen by lower value of SD, and lower value of coefficient of variations.
- Weekend average travel time of buses is around 2 to 4 % less than weekdays travel time. This condition is seen mostly because of less passenger demand at each station during weekends.
- Reliability indices such as PTI, BTI, and TTI show almost the same trend in the weekdays and weekend analysis.
- As non-express buses serve all the stations between origin and destination, along with that distance it covers 2 kms more than express and hence it is evident that the average travel times of buses during the off-peak, as well as peak hours, are more than that of express routes.
- Compared with the express route, highest of around 27.21 % of increased average travel time has been seen in the case of non-express routes during morning off-peak hours. Subsequently, morning and evening peaks are 29.17 %, 29.66 %, and 33.51 % for the evening off-peak. This might be due to more distance coverage, a more number of bus stop served by the buses and impact of non-dedicated lane on the bus operations.

Table 5.3 Descriptive Statistics of Non- Express Route – Weekday

Hour of the Day	Avg. TT (minute)	SD of TT	CV of TT	95th Percentile TT (minute)	PTI	BTI	TTI
5 - 6	40.8	4.95	12.11	44.5	1.4	9.0	1.3
6 - 7	47.6	4.45	9.33	52.1	1.6	9.3	1.5
7 - 8	49.7	4.62	9.28	55.4	1.8	11.3	1.6
8 - 9	47.7	3.04	6.36	52.0	1.6	9.0	1.5
9 - 10	46.2	2.47	5.35	51.4	1.6	11.2	1.5
10 - 11	48.6	2.89	5.95	54.9	1.7	12.9	1.5
11 - 12	50.1	3.10	6.19	54.7	1.7	9.1	1.6
12 - 13	52.2	4.09	7.85	59.9	1.9	14.8	1.7
13 - 14	45.3	3.97	8.77	53.2	1.7	17.5	1.4
14 - 15	44.6	3.93	8.81	50.9	1.6	14.2	1.4
15 - 16	44.9	4.40	9.80	52.2	1.7	16.3	1.4
16 - 17	46.4	2.14	4.61	52.2	1.7	12.4	1.5
17 - 18	58.3	2.84	4.88	73.2	2.3	25.4	1.8
18 - 19	58.5	2.68	4.59	71.0	2.2	21.4	1.9
19 -20	56.3	3.00	5.32	69.3	2.2	23.0	1.8
20 - 21	50.6	3.85	7.60	67.8	2.1	33.9	1.6
21 - 22	43.8	5.37	12.26	50.2	1.6	14.6	1.4
22 - 23	40.4	5.12	11.02	45.4	1.4	12.4	1.3

Table 5.4 Descriptive Statistics of Non- Express Route – Weekend

Hour of the Day	Avg. TT (minute)	SD of TT	CV of TT	95th Percentile TT (minute)	PTI	BTI	TTI
5 - 6	41.4	5.53	13.37	45.2	1.4	9.3	1.3
6 - 7	47.1	5.28	11.21	52.9	1.7	12.3	1.5
7 - 8	46.2	5.18	11.21	52.2	1.7	13.0	1.5
8 - 9	46.6	4.89	10.51	53.9	1.7	15.8	1.5
9 - 10	45.0	3.10	6.90	48.5	1.5	7.9	1.4
10 - 11	45.9	3.11	6.77	50.6	1.6	10.2	1.5
11 - 12	45.5	2.79	6.12	50.8	1.6	11.6	1.4
12 - 13	47.6	3.27	6.87	52.5	1.7	10.4	1.5
13 - 14	45.8	3.27	7.13	53.6	1.7	17.0	1.4
14 - 15	42.9	5.10	11.89	51.3	1.6	19.5	1.4
15 - 16	41.2	2.90	7.04	46.0	1.5	11.6	1.3
16 - 17	43.0	2.68	6.23	48.5	1.5	12.9	1.4
17 - 18	55.2	2.54	4.60	68.6	2.2	24.2	1.7
18 - 19	56.3	2.74	4.88	69.4	2.2	23.3	1.8
19 -20	54.7	2.97	5.43	67.8	2.1	23.9	1.7
20 - 21	50.8	2.89	5.69	64.4	2.0	26.8	1.6
21 - 22	46.4	4.54	9.78	53.9	1.7	16.1	1.5
22 - 23	47.3	5.07	10.70	54.8	1.7	15.9	1.5

5.2.1.2 Segment Level Analysis

Descriptive statistics analysis is also conducted for the segment-level analysis. Here in the segments-level analysis different time periods are considered as temporal aggregations. As mentioned previously in section 4.3.1 of the methodology, three segments are selected for the TTV analysis. Out of them, two had dedicated nature of bus operation and one is non-dedicated. Segment level analysis helps in taking improvements strategies more effectively on specific segments, which has more impact on the travel time of the overall route. Here, dedicated and non-dedicated segments are purposefully considered in the analysis to assess the positive and negative impact of both types of segments on the overall travel time of the routes.

Analysis is carried out on weekdays and weekends separately for all three segments, and with respect to that, all the obtained results have shown in the form of tables, viz. table 5.5, table 5.6, table 5.7, table 5.8, table 5.9, and table 5.10 and for better visualisation of variations in the unit average TT of all three segments in both the weekday and weekend, plot has given in the figure 5.7.

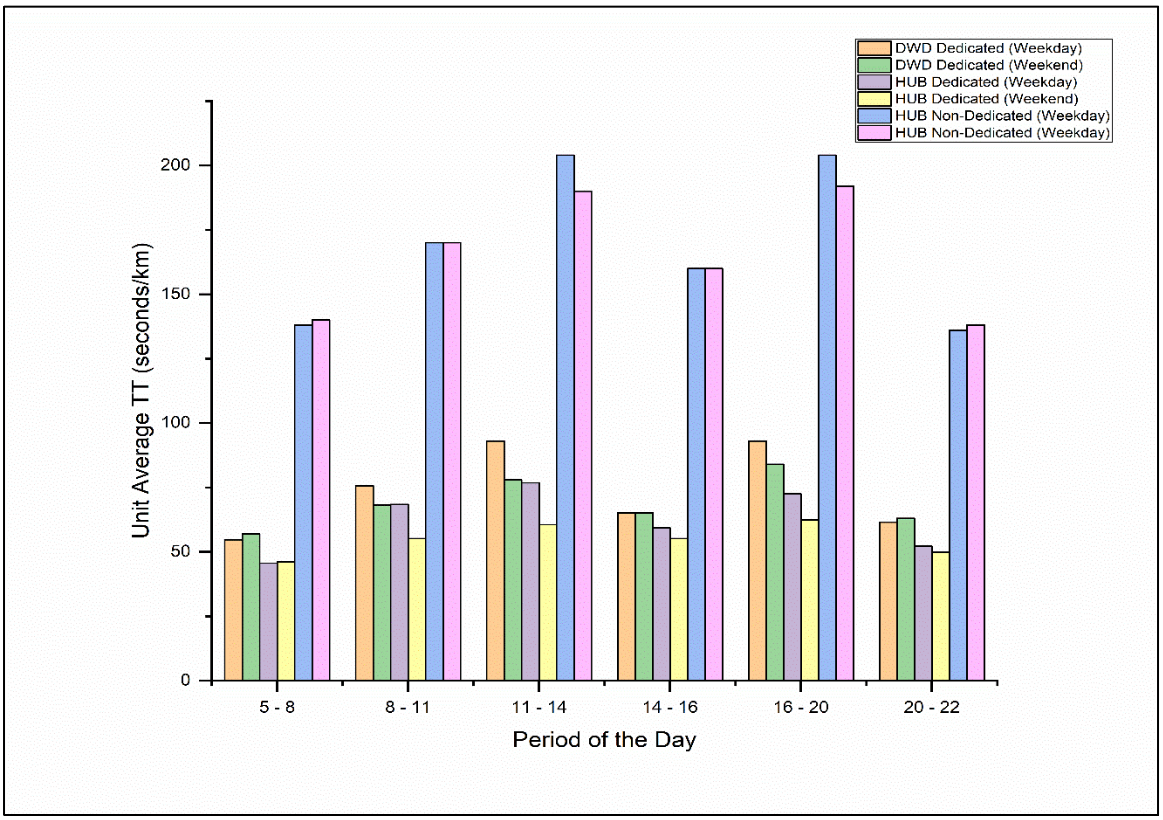


Figure 5.7 Variation in the unit average TT of segments

Table 5.5 Descriptive Statistics of Dedicated Segment (DWD) -Weekday

Hour of the Day	Average TT (minute)	SD of TT	CV of TT	95th Percentile TT (minute)	PTI	BTI	TTI
5 - 8	7.3	1.51	20.64	7.8	1.7	6.9	1.6
8 - 11	10.1	1.18	11.67	11.8	2.5	16.1	2.2
11 - 14	12.4	1.36	10.93	14.7	3.1	18.5	2.6
14 - 16	8.7	1.18	13.62	9.9	2.1	14.7	1.8
16 - 20	12.4	1.10	8.90	14.8	3.1	19.7	2.6
20 - 22	8.2	1.89	23.23	8.8	1.9	8.4	1.7

Table 5.6 Descriptive Statistics of Dedicated Segment (DWD) – Weekend

Hour of the Day	Average TT (minute)	SD of TT	CV of TT	95th Percentile TT (minute)	PTI	BTI	TTI
5 - 8	7.6	1.29	16.95	9.0	1.9	7.7	1.6
8 - 11	9.1	1.21	13.29	11.5	2.4	13.5	1.9
11 - 14	10.4	1.27	12.26	10.7	2.3	15.4	2.2
14 - 16	8.7	1.17	13.50	10.7	2.3	13.1	1.8
16 - 20	11.2	0.86	7.70	11.8	2.5	16.0	2.4
20 - 22	8.4	1.47	17.53	9.9	2.1	5.7	1.8

Table 5.7 Descriptive Statistics of Dedicated Segment (HUB) – Weekday

Hour of the Day	Average TT (minute)	SD of TT	CV of TT	95th Percentile TT (minute)	PTI	BTI	TTI
5 - 8	7.6	1.95	25.60	8.1	2.4	6.4	2.3
8 - 11	11.4	1.46	12.73	13.0	3.9	13.7	3.4
11 - 14	12.8	1.73	13.44	14.7	4.4	14.7	3.8
14 - 16	9.9	1.56	15.85	11.1	3.3	13.0	2.9
16 - 20	12.1	1.44	11.83	14.1	4.2	16.4	3.6
20 - 22	8.7	1.79	20.44	9.4	2.8	7.7	2.6

Table 5.8 Descriptive Statistics of Dedicated Segment (HUB) – Weekend

Hour of the Day	Average TT (minute)	SD of TT	CV of TT	95th Percentile TT (minute)	PTI	BTI	TTI
5 - 8	7.7	1.45	18.78	8.1	1.8	5.3	1.8
8 - 11	9.2	1.22	13.19	10.3	2.3	11.3	2.1
11 - 14	10.1	1.10	10.92	11.7	2.7	15.5	2.3
14 - 16	9.2	1.43	15.56	10.0	2.3	9.1	2.1
16 - 20	10.4	1.09	10.42	12.0	2.7	15.2	2.4
20 - 22	8.3	1.45	17.43	8.8	2.0	6.3	1.9

Table 5.9 Descriptive Statistics of Non-Dedicated Segment (HUB) – Weekday

Hour of the Day	Average TT (minute)	SD of TT	CV of TT	95^{th} Percentile TT (minute)	PTI	BTI	TTI
5 - 8	6.9	1.39	20.17	7.9	2.1	14.6	1.8
8 - 11	8.5	1.22	14.41	10.3	2.7	21.8	2.3
11 - 14	10.2	1.33	13.10	12.2	3.3	19.8	2.7
14 - 16	8.0	1.37	17.16	9.3	2.5	16.7	2.1
16 - 20	10.2	1.21	11.88	12.5	3.3	23.0	2.7
20 - 22	6.8	1.19	17.38	7.9	2.1	15.3	1.8

Table 5.10 Descriptive Statistics of Non-Dedicated Segment (HUB) – Weekend

Hour of the Day	Average TT (minute)	SD of TT	CV of TT	95^{th} Percentile TT (minute)	PTI	BTI	TTI
5 - 8	7.0	1.38	19.64	8.1	2.2	15.8	1.9
8 - 11	8.5	1.12	13.19	10.2	2.7	20.1	2.3
11 - 14	9.5	1.11	11.64	11.9	3.2	24.5	2.5
14 - 16	8.0	1.27	15.90	9.3	2.5	16.9	2.1
16 - 20	9.6	1.13	11.80	11.8	3.2	23.4	2.6
20 - 22	6.9	1.25	18.04	7.9	2.1	14.3	1.8

From the tables, 5.5 to table 5.10 and figure 5.7 of the segments following inferences are made,

- Distance covered by both the dedicated segments is more than the non-dedicated segment, as mentioned in table 3.3. But the average travel time taken by the buses to cover the dedicated segments distances is not much greater than non-dedicated average travel times. To elaborate, during peak hours of the weekdays, DWD dedicated segment and HUB dedicated segment are taking an additional 17.74 % and 5.55 % of the average travel time compared with the average travel time of the non-dedicated segment during those hours. As in the peak hours of weekends, they are taking an additional 23.69 % and 6.82 % more than the non-dedicated segment's average travel times. Hence, differences between average travel times of dedicated and non-dedicated travel times are smaller, and it clearly shows the impedance caused by the heterogeneous traffic conditions in the BRTS bus operations.
- Apart from the average travel times of dedicated segments, the non-dedicated segment also has higher side of TTI, BTI and PTI values during the peak hours of the weekdays and weekends compared to the dedicated segments. This signifies the impact of the unsegregated way of bus BRTS operations on the commuters.
- With the segment-level analysis in the current research work, the weekend effect is clearly seen, and it is nullified in the route-level analysis. All three segments are the busiest segments located in the prime areas of the Hubli-Dharwad cities. Whether, it is a week or a weekend; continuous passenger demand is there in most of the stations of the segments. Hence, the weekend results of all the segments show more dispersion in the travel times along with more CV values compared to the weekday's descriptive values.
- Overall, it is understood that the non-dedicated lane of bus operation is affecting the system's overall performance. Further, it gets intensified due to peak hour operations, high passenger demand, improper lane discipline followed by the mixed traffic, etc. On the contrary, dedicated lanes have good backing to the HDBRTS system in keeping the performance to the required level.

5.2.2 Travel Time Variability Study based on Probability Distributions

The travel time variability analysis, probability distributions are made fit to the travel time data points of the routes and segments. Detailed results and discussions connected to the analysis have been made in the subsequent sections of the current chapter.

5.2.2.1 Travel Time Aggregations for Route and Segment Level Analysis

Travel time variability analysis using probability distributions has been carried out for the routes considering the hour-wise temporal aggregations, whereas period-wise temporal aggregations have been considered in the case of segments. Table 5.11, table 5.12 and table 5.13 show the detailed descriptions of the route level and segment level data considered in the study. The total sample number of travel time data points considered in the study for each hour get the average value of that hour or period shown in the table.

Also, most of the data points in each hour show unimodal behaviour after conducting the Hartigan dip test and even period wise data points show the unimodal behaviours. The results have been given separately for the express route, non-express route and segments in the tables 5.11, 5.12 and 5.13 respectively. R Studio coding used in the Hartigan dip test is given in the Appendix A.14.

Once again, the same express and non-express routes are considered here in the TTV analysis using probability distributions, and for the segment level analysis, the same two dedicated segments and one non-dedicated segment have been considered. All the weekdays and weekend data points have been considered for the analysis. Segment-level analysis compensate ascertaining the effect of spatial variations in the reliability of bus routes (Chepuri, A., et al. 2019).

Hence segment level analysis here is mainly carried out to have the inclusive view of many unseen effects caused on the travel time variability, which are actually nullified in the route level analysis. Probability distribution analysis in the study considers seven potential statistical distributions, namely, Burr, GEV, Log-logistic, Logistic, Lognormal, Normal, and Weibull, based on the literature review and subsequent performance of considered distributions explained in the sections 5.2.2.2 and 5.2.2.3 of current chapter.

Table 5.11 Descriptive Summary for Travel Time Aggregations – Express Routes - Hour wise

Transit Path	Week	Time Period	Sample Size	Dip Test p-value	Week	Time Period	Sample Size	Dip Test p-value
Express Route	Week Days	5	806	0.966	Weekend Days	5	306	0.555
		6	1069	0.971		6	484	0.977
		7	906	0.838		7	410	0.824
		8	909	0.787		8	394	0.460
		9	1092	0.423		9	446	0.972
		10	1132	0.992		10	466	0.875
		11	1189	0.776		11	460	0.992
		12	1246	0.586		12	508	0.697
		13	1363	0.047		13	552	0.002
		14	1631	0.996		14	650	0.843
		15	1832	0.531		15	729	0.748
		16	1496	0.787		16	624	0.340
		17	1663	0.994		17	1468	0.930
		18	1490	0.790		18	1294	0.851
		19	1950	0.589		19	1562	0.598
		20	1498	0.656		20	1579	0.991
		21	1575	0.986		21	1601	0.865
		22	1411	0.822		22	801	0.707

Table 5.12 Descriptive Summary for Travel Time Aggregations Non-Express Routes – Hour wise

Transit Path	Week	Time Period	Sample Size	Dip Test p-value	Week	Time Period	Sample Size	Dip Test p-value
Non-Express Route	Week Days	5	1122	0.954	Weekend Days	5	377	0.994
		6	1184	0.779		6	639	0.563
		7	1069	0.989		7	577	0.722
		8	1034	0.966		8	568	0.845
		9	1542	0.686		9	645	0.961
		10	1109	0.775		10	672	0.753
		11	1115	0.956		11	642	0.955
		12	1229	0.482		12	684	0.735
		13	1911	0.142		13	775	0.112
		14	1212	0.982		14	881	0.991
		15	1151	0.202		15	848	0.592
		16	1619	0.945		16	675	0.946
		17	1482	0.645		17	566	0.867
		18	1232	0.564		18	435	0.558
		19	982	0.396		19	472	0.849
		20	917	0.631		20	470	0.602
		21	571	0.261		21	333	0.818
		22	289	0.855		22	234	0.928

Table 5.13 Descriptive Summary for Travel Time Aggregations Segments – Period wise

Transit Path	Week	Time Period	Sample Size	Dip Test p-value	Week	Time Period	Sample Size	Dip Test p-value
Dedicated Segment (DWD)	Week Days	Morning Off-peak	2403	0.986	Weekend Days	Morning Off-peak	1333	0.808
		Morning Peak	2765	0.660		Morning Peak	1594	0.492
		Inter Peak	2150	0.598		Inter Peak	1838	0.879
		Afternoon Off-peak	2869	0.991		Afternoon Off-peak	1525	0.991
		Evening Peak	2573	0.369		Evening Peak	943	0.815
		Evening Off-peak	223	0.487		Evening Off-peak	79	0.628
Dedicated Segment (HUB)	Week Days	Morning Off-peak	2305	0.689	Weekend Days	Morning Off-peak	1307	0.991
		Morning Peak	2629	0.691		Morning Peak	1579	0.896
		Inter Peak	2089	0.679		Inter Peak	1803	0.992
		Afternoon Off-peak	1650	0.822		Afternoon Off-peak	1441	0.505
		Evening Peak	2414	0.578		Evening Peak	888	0.546
		Evening Off-peak	336	0.137		Evening Off-peak	117	0.494
Non-Dedicated Segment (HUB)	Week Days	Morning Off-peak	2547	0.734	Weekend Days	Morning Off-peak	1000	0.664
		Morning Peak	2753	0.513		Morning Peak	1163	0.825
		Inter Peak	2088	0.977		Inter Peak	1318	0.961
		Afternoon Off-peak	2823	0.788		Afternoon Off-peak	1162	0.727
		Evening Peak	2492	0.972		Evening Peak	891	0.991
		Evening Off-peak	404	0.512		Evening Off-peak	172	0.547

5.2.2.2 Distribution Performance at the Route Level

The different distributions at different ratios are have better explained the total number of travel time cases considered in the analysis. KS test p-value has been used to check the goodness of fit of the chosen distributions. It has been called as the performance of individual distribution in the analysis. Based on the p-value of individual distributions, the number of cases passed by the distributions, how many of them are in the top first position, and top three positions have been obtained with their ratios calculated. In the meantime, each distribution's mean and SD p-values are obtained.

Their performance is analysed based on the three ratios and descriptive statistics of chosen distributions. Table 5.14 represents the chosen distribution's performance. The total cases considered are 924 for the analysis.

GEV distribution has passed 902 cases and stood first on the best performance list. It has been observed that the mean p-value of the GEV is 0.714 and is the highest amongst other distributions considered. SD value of GEV is 0.274 and is at the lower side among the other distributions. Burr distribution is the second best-performed distribution in the current study by passing a total of 790 cases with a mean p-value of 0.610 and SD of 0.253. Lognormal distribution is considered as the third best-fit distribution with a passing of 723 cases, a mean p-value of 0.485, and SD of 0.314. Weibull distribution is found to be the lowest performed one amongst others.

Sample of probability density functions (pdf) plots of route level analysis have shown in figures 5.8 (a) and (b). From the plots, it has been observed that express and non-express routes have the same nature plots. In the peak hours, a dispersion in the travel time is less; hence, they are normally distributed. Whereas in the off-peak hours, most of the travel times are towards the lower side, they are positively skewed in nature.

Table 5.14 Distributions Performance at Route Level

Type	p- Value		Cases			
	Mean	Standard deviation	Passed number	Passed ratio	Top three ratio	First ratio
Burr	0.610	0.253	790	85.5	63.5	28.6
GEV	0.714	0.274	902	97.6	81.5	40.1
Log-Logistic	0.459	0.348	720	78.0	29.8	6.0
Logistic	0.386	0.332	688	74.5	14.3	3.4
Lognormal	0.485	0.314	723	78.2	34.7	5.6
Normal	0.430	0.344	590	63.9	24.4	3.6
Weibull	0.268	0.387	558	60.4	16.5	3.6

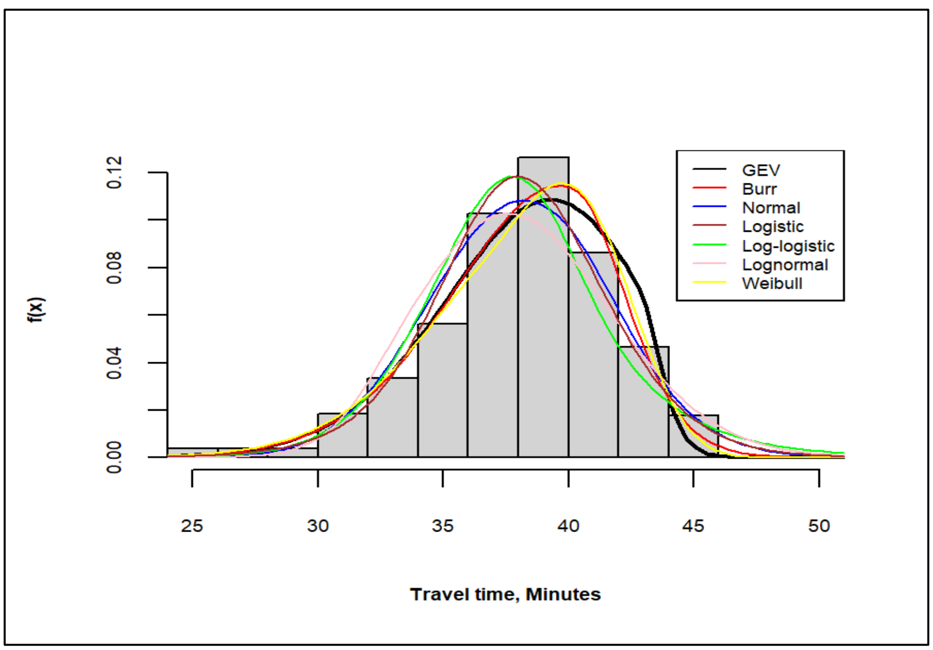


Figure 5.8 (a) Probability Density Function Plot (Route - Peak Hour-E)

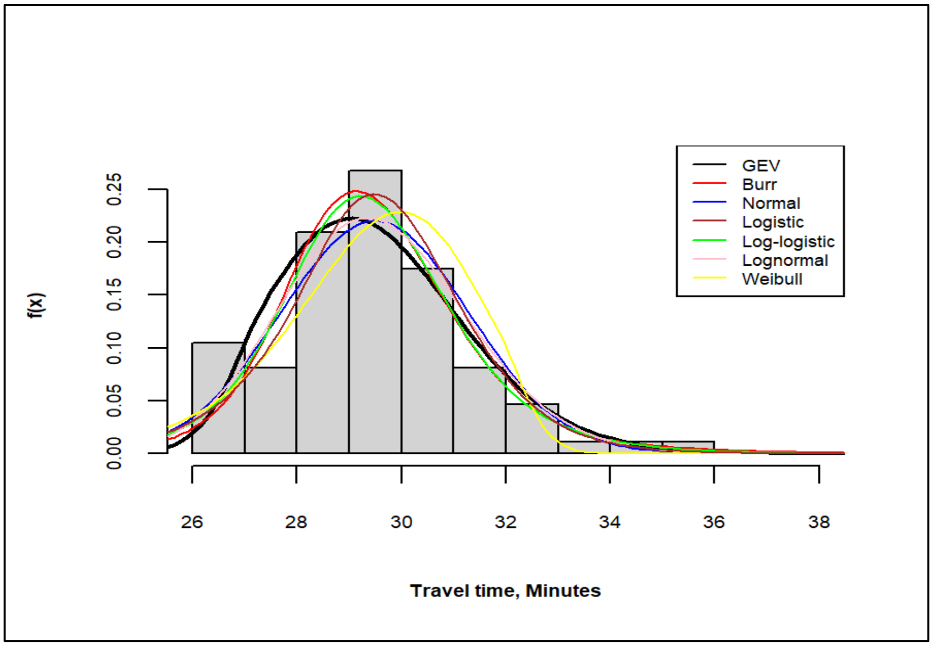


Figure 5.8 (b) Probability Density Function Plot (Route -Off-peak Hour-E)

5.2.2.3 Distribution Performance at the Segment Level

Distribution performance analysis at the segment level is shown in detail in table 5.15. Analysis is carried out individually for all the three segments. Based on the obtained results, it is inferenced that, in the dedicated and non-dedicated segments, GEV distributions have the highest mean p-value and lowest SD value of the KS test. Also, it has a higher ratio of cases passed, a higher ratio of cases in the top 3, and a higher ratio of cases in the top 1 position.

Again, based on next highest value of statistical p-value and lower SD value, Burr distribution stands in the second position of the best performance distribution. The third position is taken by log-logistics distribution both in the case of DWD dedicated segment and HUB non-dedicated segment with the 69.0 % and 66.7 % of passed cases ratio respectively. In contrast, lognormal has taken the third position in the case of HUB dedicated segment with 78.6 % of passed cases ratio.

Sample of probability density functions (pdf) plots of segment-level analysis have been shown in figures 5.9 (a) and (b) and 5.10 (a) and (b).

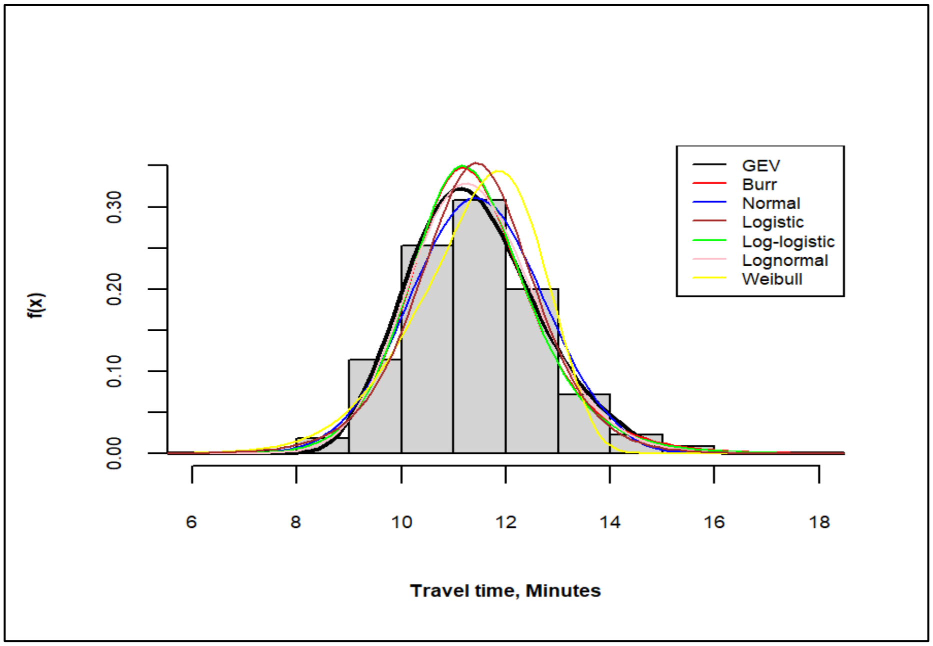


Figure 5.9 (a) Probability Density Function Plot
(Dedicated Segment Peak Period)

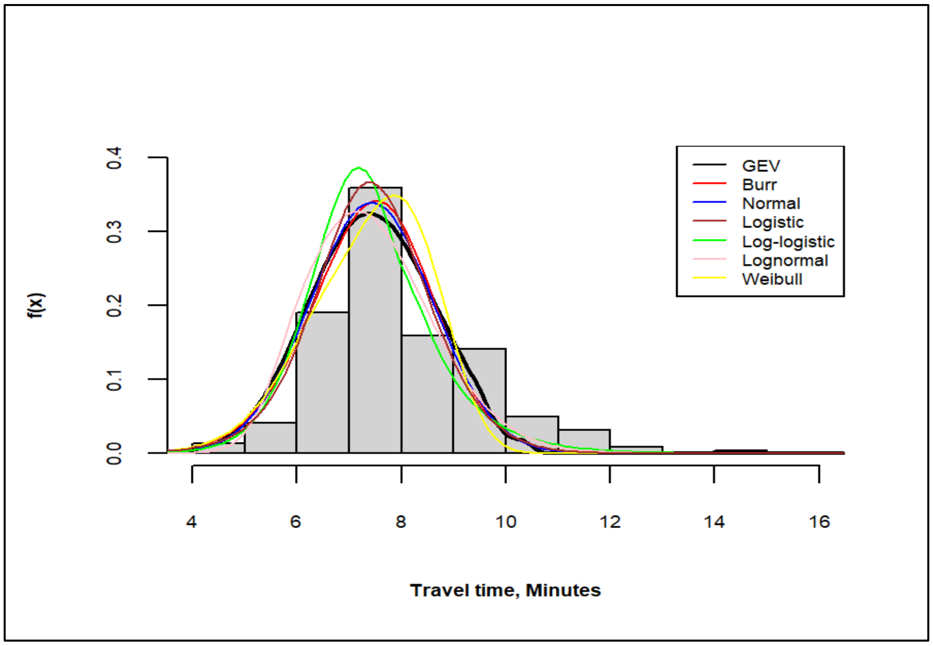


Figure 5.9 (b) Probability Density Function Plot
(Dedicated Segment Off-Peak Period)

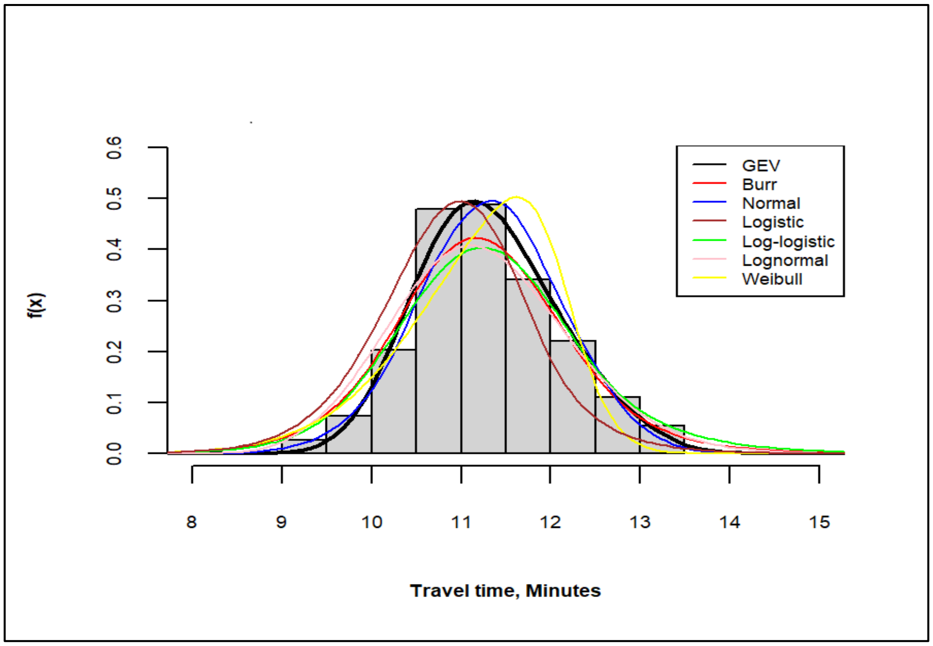


Figure 5.10 (a) Probability Density Function Plot (Non-dedicated Segment Peak Period)

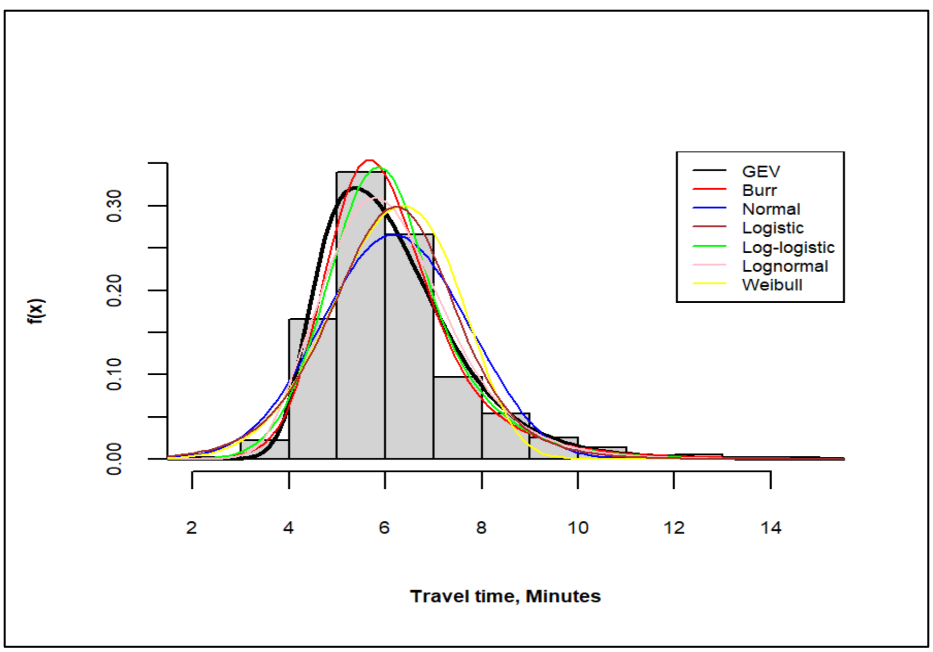


Figure 5.10 (b) Probability Density Function Plot (Non-dedicated Segment Off-peak Period)

Table 5.15 Distribution Performance at Segment Level

Segment	Type	p- Value		Cases		
		Mean	Standard deviation	Passed ratio	Top three ratio	First ratio
Dedicated Segment (DWD)	Burr	0.610	0.253	91.7	77.4	38.1
	GEV	0.675	0.231	95.2	89.3	47.6
	Log-Logistic	0.359	0.317	69.0	26.2	1.2
	Logistic	0.211	0.252	59.5	9.5	0.0
	Lognormal	0.404	0.362	64.3	35.7	4.8
	Normal	0.264	0.312	57.1	14.3	0.0
	Weibull	0.165	0.267	39.3	10.7	4.8
Dedicated Segment (HUB)	Burr	0.505	0.251	80.5	70.7	21.4
	GEV	0.626	0.230	90.9	85.7	41.7
	Log-Logistic	0.357	0.339	71.4	26.2	4.8
	Logistic	0.241	0.297	75.0	33.3	8.3
	Lognormal	0.422	0.337	78.6	47.6	8.3
	Normal	0.298	0.290	70.2	27.4	1.2
	Weibull	0.188	0.248	58.3	16.7	4.8
Non-Dedicated Segment (HUB)	Burr	0.525	0.319	82.1	75.0	34.5
	GEV	0.669	0.276	97.6	92.9	52.4
	Log-Logistic	0.282	0.300	66.7	29.8	3.6
	Logistic	0.173	0.257	48.8	14.3	2.4
	Lognormal	0.269	0.315	58.3	16.7	3.6
	Normal	0.182	0.292	36.9	8.3	1.2
	Weibull	0.146	0.304	35.7	14.3	0.0

5.2.3 Travel Time Reliability Analysis with GEV Distribution

Segment level travel time variability analysis is carried out with seven potential probability distributions. GEV distribution is well performed for all the travel time cases considered in the analysis in comparison with the other chosen distributions. The robustness and accuracy of the GEV distribution better helped out in capturing the characteristics of most of the travel time cases considered in the analysis. Hence, GEV is superior among other distributions in the case of segments and routes. GEV distribution belong to the family of continuous probability distribution and are developed within extreme value theory. It contains three parameters viz. location parameter represented with μ, scale parameter represented with σ, and shape parameter represented with $k \neq 0$. (Chepuri, A., et al. 2019). Variation in the travel time of the buses and subsequent changes in the reliability of the service mainly depends upon various incidences that takes place during the journey. Passenger demand is one such directly affecting factor on the travel time variability and reliability. BTI is an important factor that affects travellers' decisions (Yang, H., et al. 2020) and this reliability indices have high connections with passenger comfortability.

Herein, the current research work has tried to correlate the segment shape parameter of best fit GEV distribution with the passenger demand and Buffer Time Reliability Index of the segment. All the data points considered in this initiative belong to different periods of the day and hence within-day correlation and day-to-day correlation have been established.

From figures 5.9 (a) and (b) and 5.10 (a) and (b), for the dedicated and non-dedicated segments, it is observed that, during the peak periods, PDF of the distributions are normally distributed, whereas in the off-peak they are positively skewed as most of the travel times cases are on the lower side. The impact of higher and lower passenger demand along the segments is persistently seen on the peak and off-periods considered in the study. In the meantime, there is a variation in the value of the GEV shape parameter 'k' considered. During the peak period of the day, the 'k' value is observed to be on the higher negative side, and during off-peak periods it is at the lower negative or the positive side of the values. The identical nature of the variations in the considered parameters value is observed during all peak and off-peak periods in the analysis. As passenger demand also vary with peak and off-peak periods as higher and lower values respectively. Hence variations in the values of ‘k’ value with peak and off-peak period matches with the demand. such as higher negative ‘k’ value is observed with higher passenger demand, and a lower negative or positive value ‘k’ value is observed with lower passenger demand. Analysis is carried out separately on weekdays and weekends to comprehend the parameters' variations clearly. GEV shape parameters at different periods are also compared with variation in Buffer Time Index along with the segments.

During peak periods additional time spent by the passenger in their total journey time is more whereas during off-peak periods it is less, hence this variation is shown with the values of BTI. Again, this BTI has a direct correlation with shape parameter ‘k.’ Higher BTI is observed with higher negative ‘k’, whereas lower BTI is with a lower negative ‘k’ value. Figure 5.11 (a) and (b) and 5.12 (a) and (b) show the correlation between GEV shape parameter ‘k’, passenger demand, and BTI of the dedicated and non-dedicated segments, respectively.

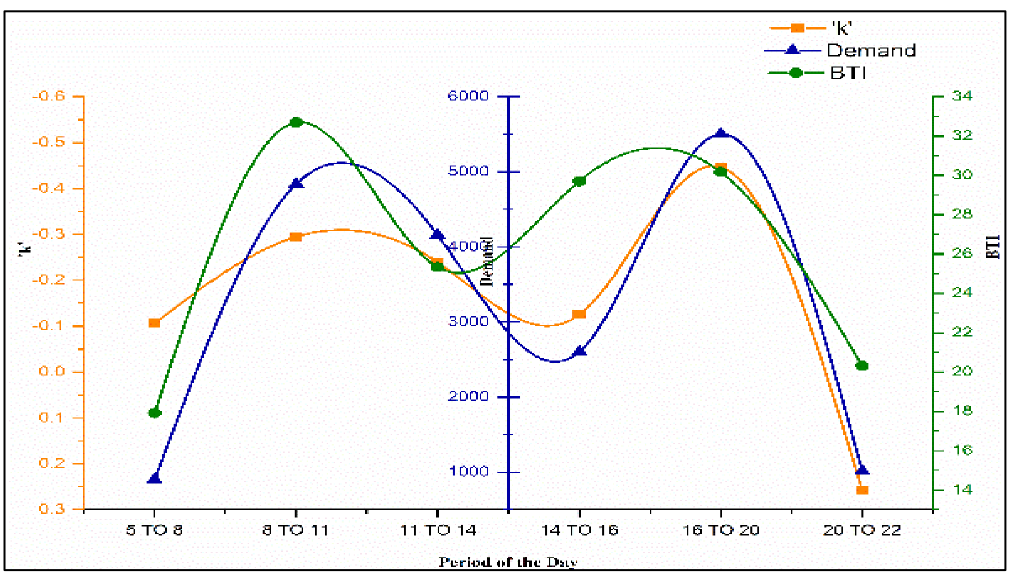


Figure 5.11 (a) Correlation Plot of ‘k’, Passenger Demand and BTI – Dedicated Segment - Weekday

Overall, when there is variation in the passenger demand and BTI, there is corresponding variations in the GEV shape parameter ‘k’ value. Hence, it is concluded that passenger demand and BTI have a direct correlation with the variations in the GEV shape parameter ‘k.’

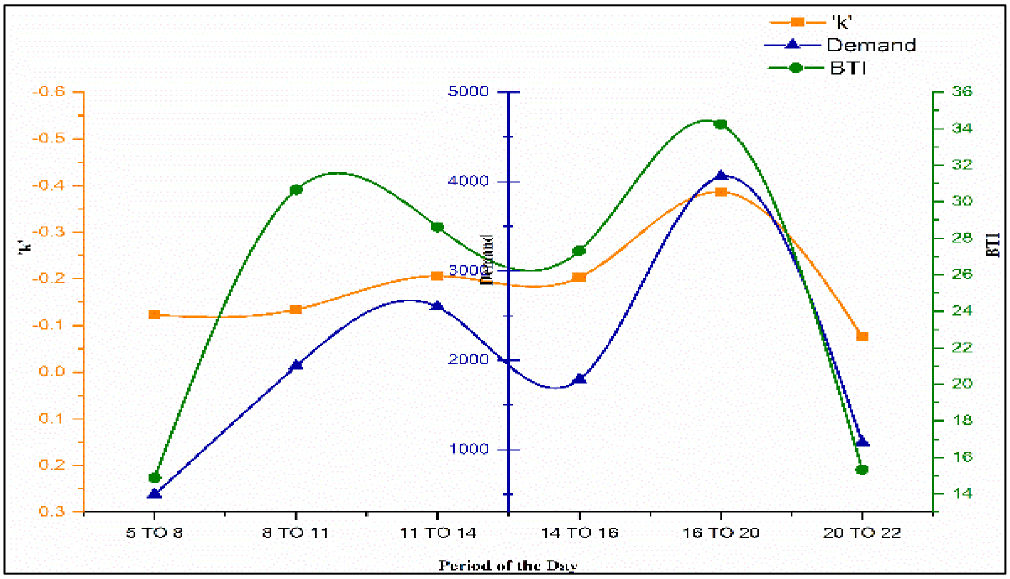


Figure 5.11 (b) Correlation Plot of ‘k’, Passenger Demand and BTI – Dedicated Segment - Weekend

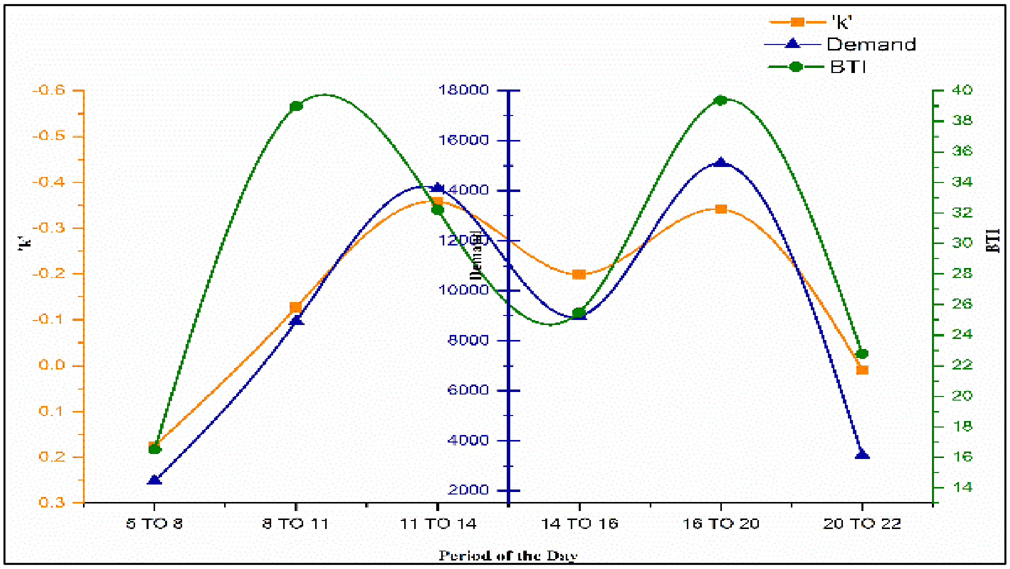


Figure 5.12 (a) Correlation Plot of ‘k’, Passenger Demand and BTI – Non-Dedicated Segment - Weekday

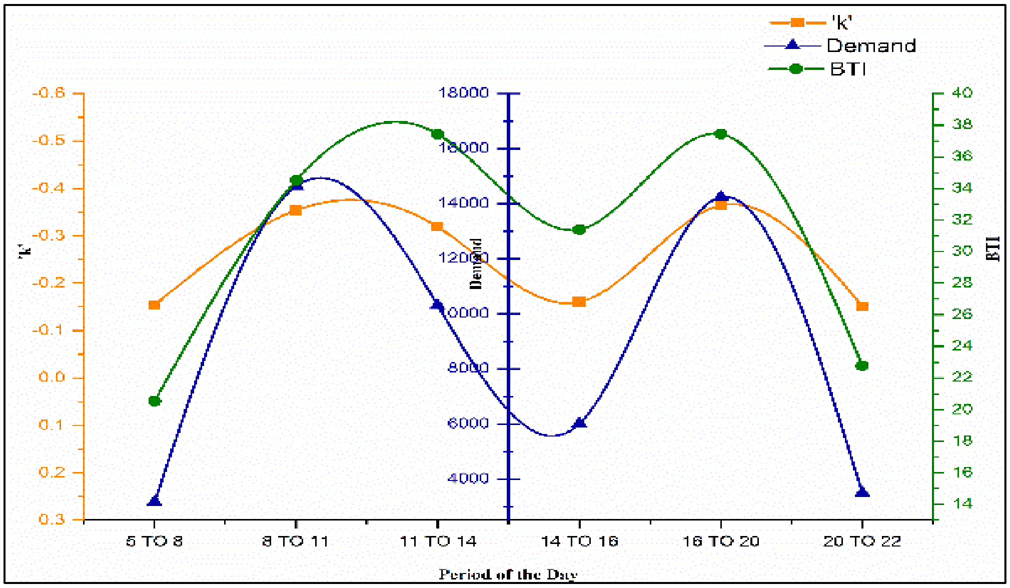


Figure 5.12 (b) Correlation Plot of 'k', Passenger Demand and BTI – Non-Dedicated Segment – Weekend

In the figures 5.11 (a) and (b) and 5.12 (a) and (b) X-axis represents the different periods of the day. Three Y-axis parameters have been used in plotting the figures, whereas the first Y-axis represents values of GEV shape parameter 'k', the second Y-axis represents passenger demand, and the third Y-axis represents the BTI along with the dedicated and non-dedicated segments, respectively.

It can be seen from the plots that both the segments have lower BTI and passenger demand during the off-peak periods of morning, afternoon, and evening concerning a lower negative or positive value of GEV shape parameter 'k' as seen in the plot. During the peak periods of the morning, evening, and interpeak BTI, passenger demand values are higher concerning a higher negative value of GEV shape parameter 'k' is seen. The parameter variations trend is same in both the dedicated and the non-dedicated segment plots corresponding to weekdays and weekends. But higher parameters value have been observed in the non-dedicated segment plot for weekdays and weekends. This condition evidently depicts the adverse impact of non-dedicated segments on the distribution and travel time reliability parameters.

5.3 MODELLING TRAVEL TIME RELIABILITY OF THE SYSTEM

To assess the statistically significant variables impacting the system performance, modelling the travel time reliability for the segments has been carried out. As mentioned in section 4.3.2, independent and dependent variables have been taken from the same two dedicated segments and one non-dedicated segment. Table 5.16 shows the description of statistically significant independent and two dependent variables considered in the analysis. As all covariables are selected based on individual Pearson coefficient values with the dependent variables, thus TTR modelling results in the study show that most of them are statistically significant in the models developed. Table 5.17 and Table 5.18 shows the obtained modelling results.

Table 5.16 Description of Dependent and Independent Variables

Variable	Description of Variable	Model 1	Model 2
Dependent variables			
Avg. travel time	Average time taken by the buses to travel between selected the segment (seconds).	D	-
Buffer time	Extra time budgeted to the guaranteed arrival at the destination with certain probability (seconds).	-	D
Independent variables			
Length	Length of the segment considered (km).	I	I
Peak and off-peak period	Dummy variable for the peak and off-peak periods in each day, 1 if peak and 0 if off-peak	I	I
Passenger demand	Total board and alight of passengers between the selected segment during different period of the day.	I	I
Bus stop density	Number of bus stops per kilometre of the segment considered, bus stops/km segment.	I	I
Intersection density	Number of signalised and unsignalised intersections per kilometre of the segment, intersections/km segment.	I	I
Land use pattern	Dummy variable considered if bus segment passes through the CBD area, 1 if passes through the CBD and 0 if not passes through the CBD.	I	I
Dedicated and non-dedicated	Dummy variable for the segment is dedicated or non-dedicated, 0 if dedicated segment and 1 if non-dedicated segment.	I	I

NOTE: D = Dependent variable, I = Independent variable
Off-peak periods = morning off-peak - 5:00 to 8:00, evening off-peak - 20:00 to 22:00
Peak periods = morning peak - 8:00 to 11:00, Interpeak - 11:00 to 14:00, afternoon peak - 14:00 to 16:00, evening peak - 16:00 to 20:00

Table 5.17 TTR Modelling Results – Average TT as Dependent Variable

Independent Variable	Standardized Coefficients	t-statistics	Significance
(Constant)	-	4.872	<0.001
Length	0.263	8.712	<0.001
Peak and off-peak period	0.295	11.715	<0.001
Passenger demand	0.380	12.653	<0.001
Bus stop density	0.288	6.117	<0.001
Intersection density	0.085	3.45	<0.001
Land use pattern	0.132	4.212	0.003
Dedicated and non-dedicated	-0.249	-4.621	<0.001
	Adjusted R square =	**0.795**	

Table 5.18 TTR Modelling Results – BT as Dependent Variable

Independent Variable	Standardized Coefficients	t-statistics	Significance
(Constant)	-	4.518	<0.001
Length	0.218	6.042	<0.001
Peak and off-peak period	0.279	10.174	<0.001
Passenger demand	0.326	13.403	<0.001
Bus stop density	0.249	5.621	<0.001
Intersection density	0.062	2.952	<0.001
Land use pattern	0.103	3.334	0.001
Dedicated and non-dedicated	-0.075	-3.573	<0.001
	Adjusted R square =	**0.804**	

MLR Models developed with ATT and BT as dependent variables have explained the statistical significance of all the variables in the models with adjusted R square values of 0.795 and 0.804, respectively. Passenger demand has the highest positive impact on both models, with t-statistics values of 12.653 and 13.403, respectively.

Passenger demand is the proxy variable of dwell time caused at the individual stations; hence it substantially impacted the travel time reliability of the system. It suggests that demand is to be managed with systematic scheduling of the bus frequency.

The impact of the peak and off-peak periods is considered in the analysis with the dummy variables, and these variables show the highest significance in both the models developed.

Different periods show the effect of general traffic conditions on bus operations. Hence properly managing the various traffic incidences, such as signal preferences for the buses, dwell time reduction at the stations, etc., may help in improvising the TTR of the system. Consistent with the past literature, the length of the segments has also shown a higher statistical significance in both the cases of models. Segment length is the proxy variable for the impedance caused due to pedestrian mid-block crossing and other side frictions effect (Kathuria, A., et al., 2020). Thus, as the length of the segment is more, then less becomes the TTR of the system.

Bus stop density and intersection density have also behaved positively as significant variables in the models as they act as the proxy variables concerning bus stop delays and intersections delays caused to the bus journeys. One of the segments in the current study is the non-dedicated and has mixed traffic conditions of bus operation throughout the length. The land use pattern of that segment is utterly in the CBD area. Therefore, the impact of both the conditions on the TTR of the system has been keenly observed in the models with obtained statistically significant values.

Making policies for improving the non-dedicated lane operation of buses is essential in the current system to tackle such effects on the TTV and TTR. The segment's performance represents the system's performance; hence reliable segments need to control major delays caused at the stops and intersections. Overall, outcomes of such systematic modelling of the TTR of the system with different covariates help in tackling their heterogeneity effect on the system performance. Both TTR models developed based on Average TT and BT as dependent variables have shown satisfactory performance; hence it is proposed to use both the models to assess the systems TTR considering operators and passengers perspectives.

5.4 PASSENGER DEMAND FORECASTING OF THE SYSTEM

Here in the current research work, the forecasting is also done using the seasonal naïve method, which is used for the seasonal data. This method is used as a control and as a benchmark in this work. Here, each forecasted value is equal to the last observed value from the same season. For example, in this case, as the data shows daily seasonality, the forecasted passenger demand for tomorrow 06:00 is the same as today's passenger count at 06:00. Mathematically, the forecast for any time, T+h is given by equation 5.01,

$$\hat{y}_{\mathrm{T+h|T}} = y_{T+h-m(k+1)} \tag{5.01}$$

where, $\hat{y}_{\mathrm{T+h|T}}$ denotes the forecast of y_{T+h} using the historical data $y_1, \dots, y_T$, m is the periodicity of the season, and k is the integer part of $(h-1)/m$.

The training and testing datasets used for seasonal naïve method are the same as that are used for forecasting with seasonal ARIMA and LSTM. For the forecasted values, MAE and RMSE are calculated. The forecasting performance of selected models, in the form of MAE and RMSE, of seasonal ARIMA, LSTM and seasonal naïve method for all 20 time-series, are tabulated in Table 5.19 (MAE) and Table 5.20 (RMSE).

The comparison of forecasting accuracy of seasonal ARIMA, LSTM and seasonal naïve method can be made by comparing the MAE % columns. It can be clearly observed in the graph given in Figure 5.13 that LSTM is more efficient than the seasonal ARIMA and seasonal naïve method. (In the graph, the x-axis is the 'SI No.' column of Table 5.19) The same comparison can also be made using RMSE % values instead of MAE % for further validation. The graph given in Figure 5.14 making such comparison shows that even based on RMSE %, LSTM outperforms seasonal ARIMA and seasonal naïve methods. (In the graph, the x-axis is the 'SI No.' column of Table 5.20)

Figure 5.19 Summary of Performance of Forecasting Models (MAE)

Sl. No	Station No.	Time-frame	Mean of Passenger	MAE % for Seasonal ARIMA	MAE % for LSTM	MAE % for Seasonal Naïve	% Reduction in MAE% with LSTM*
1.	5	60	707.85	13.59	8.93	15.73	34.27
2.	28	60	159.68	30.32	21.76	33.54	28.22
3.	33	60	164.80	22.37	19.45	22.41	13.02
4.	34	60	460.04	15.38	11.04	19.65	28.19
5.	35	60	275.43	18.09	11.72	20.67	35.24
6.	5	45	523.32	14.22	10.30	16.72	27.56
7.	28	45	118.02	31.60	22.02	36.65	30.32
8.	33	45	117.49	24.15	19.62	26.36	18.75
9.	34	45	340.04	16.12	12.03	20.58	25.39
10.	35	45	203.58	19.17	15.12	22.11	21.11
11.	5	30	353.92	14.66	10.91	17.64	25.58
12.	28	30	79.84	33.09	23.82	38.36	28.01
13.	33	30	79.43	26.80	21.89	29.29	18.32
14.	34	30	230.02	16.97	13.33	22.15	21.44
15.	35	30	137.71	21.52	18.84	24.74	12.42
16.	5	15	176.96	20.55	12.64	20.57	38.49
17.	28	15	39.92	49.65	24.30	42.73	51.06
18.	33	15	39.71	36.62	25.66	36.68	29.92
19.	34	15	115.01	24.90	14.66	24.91	41.13
20.	35	15	68.85	33.93	23.50	34.01	30.74

Figure 5.20 Summary of Performance of Forecasting Models (RMSE)

Sl. No.	Station No.	Time-frame	Mean Passengers	RMSE % for Seasonal ARIMA	RMSE % for LSTM	RMSE % for Seasonal Naïve	% Reduction in RMSE% with LSTM*
1.	5	60	707.85	20.59	11.32	23.23	45.03
2.	28	60	159.68	49.85	34.28	51.01	31.23
3.	33	60	164.80	29.18	24.92	29.77	14.60
4.	34	60	460.04	23.08	14.32	30.66	37.94
5.	35	60	275.43	26.82	16.15	30.10	39.78
6.	5	45	523.32	21.26	14.47	24.66	31.96
7.	28	45	118.02	51.70	37.51	55.36	27.45
8.	33	45	117.49	31.76	25.18	34.20	20.74
9.	34	45	340.04	24.35	17.02	32.14	30.09
10.	35	45	203.58	28.04	20.10	32.03	28.31
11.	5	30	353.92	20.67	14.01	25.40	32.24
12.	28	30	79.84	53.85	37.66	58.11	30.05
13.	33	30	79.43	35.21	28.19	37.87	19.95
14.	34	30	230.02	25.32	17.49	33.55	30.92
15.	35	30	137.71	30.14	24.86	35.16	17.49
16.	5	15	176.96	29.13	16.71	29.06	42.64
17.	28	15	39.92	64.48	39.10	64.54	39.36
18.	33	15	39.71	47.27	33.97	47.27	28.13
19.	34	15	115.01	36.48	20.25	36.47	44.48
20.	35	15	68.85	46.35	31.74	46.29	31.53

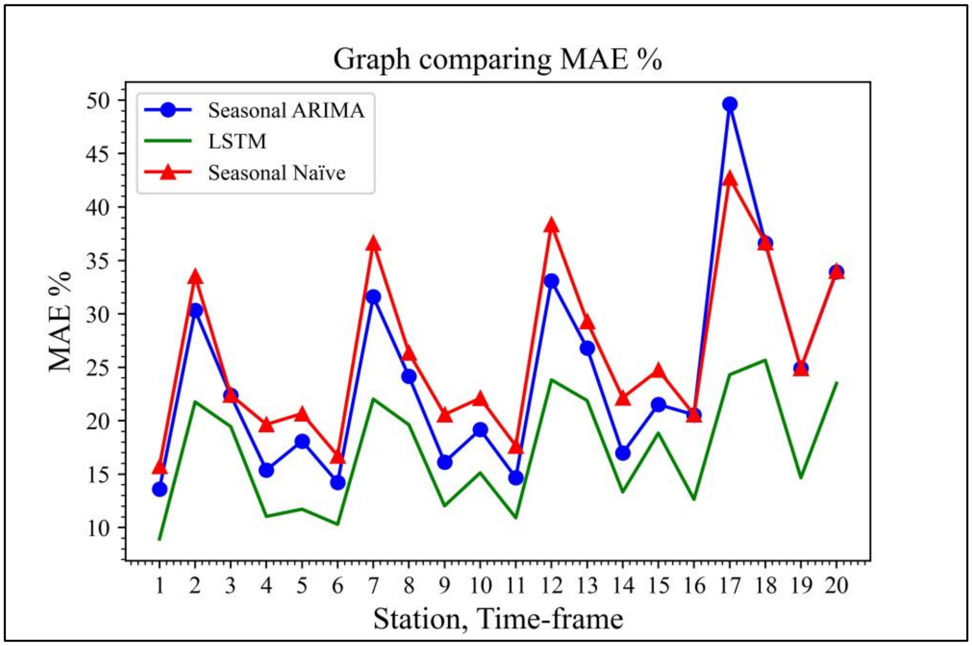


Figure 5.13 Forecasting Models Performance based on MAE %

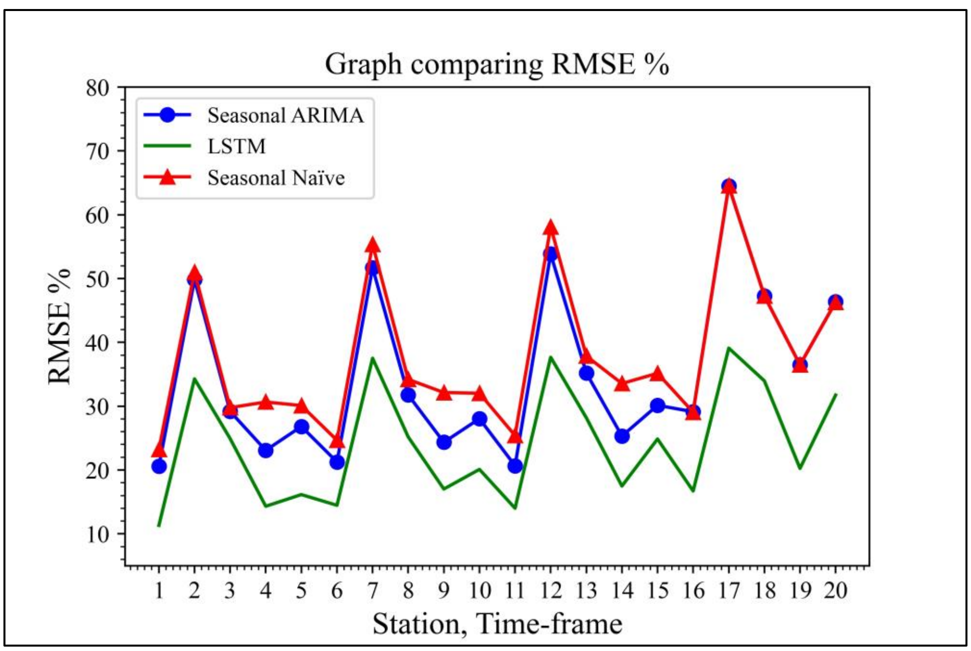


Figure 5.14 Forecasting Models Performance based on RMSE %

Thus, it is clear that, by all measures and for all time frames, LSTM is better than seasonal ARIMA. This can be attributed to the fact that LSTM can successfully learn long term dependencies and also non-linear relationships. The percentage reduction in both MAE % and RMSE % are given in the last columns of Table 5.19 and Table 5.20, respectively. They are also represented as graphs for easy visualisation and is given here in Figure 5.15. To summarize, the mean MAE % is reduced by 27.46 %, and the RMSE % is reduced by 31.08 %. (In the graph, the x-axis is the 'SI No.' column of Table 5.19 and Table 5.20).

The tables are also used to compare how the resampling time interval affects the forecasting accuracy. Firstly, considering MAE %, for LSTM, for all stations, it can be observed 60 minutes time interval gives better results than 15 minutes, 30 minutes and 45 minutes time intervals. For the visualisation purpose, the graph of this comparison is given in Figure 5.16 Similar results can be obtained with RMSE % for LSTM and also with MAE % and RMSE % for seasonal ARIMA. By this, it can be confirmed that the time-series resampled with 60-minute intervals performed better than other time intervals considered.

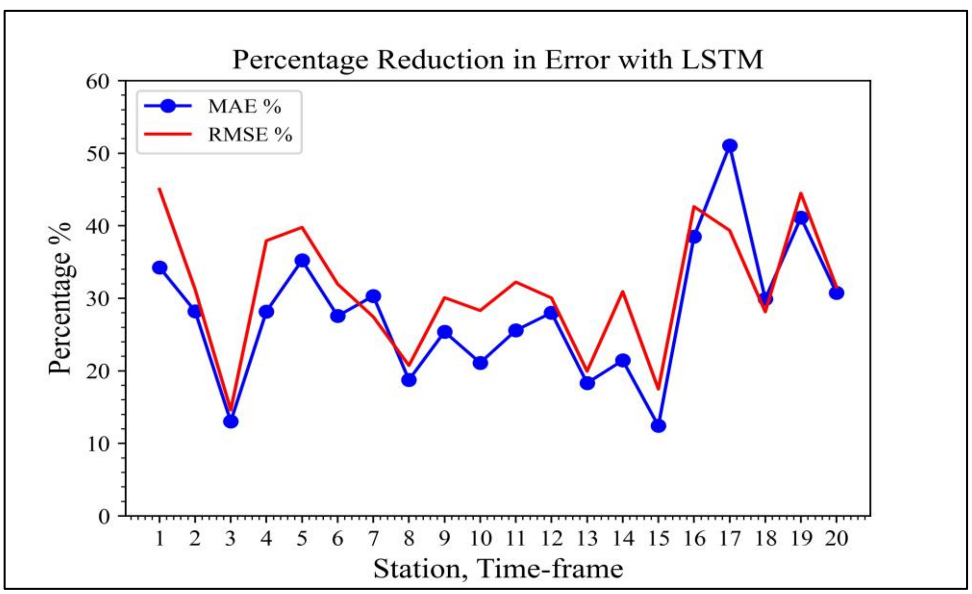


Figure 5.15 Percentage reduction in error with LSTM

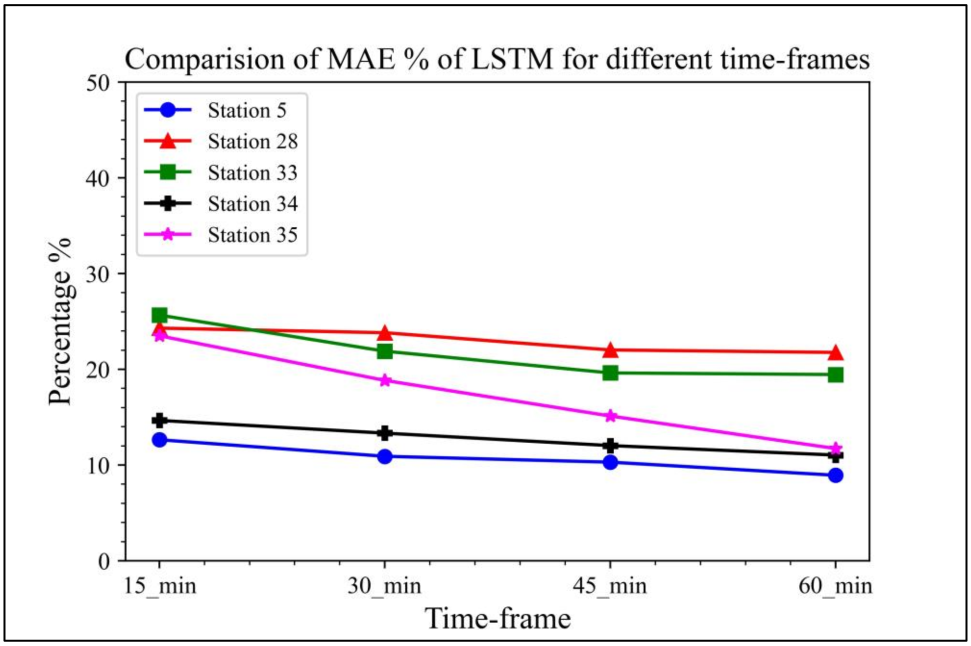


Figure 5.16 Comparison of MAE % of LSTM for different time-frames

5.5 LEVEL OF SERVICE (LOS) OF THE HDBRTS

Reliable transit service-based Level of Service (LOS) has been developed for HDBRTS. Travel time index, planning time index as well as buffer time index are the three-reliability measures of transit travel time considered in the analysis. The route which operates express, as well as non-express buses, dedicated segment, and non-dedicated, are three operational conditions of HDBRTS buses considered for the service establishment. Six service clusters are formed based upon analysis carried out with the K-mean clustering technique; such as service cluster A to service cluster F, which are universally recognized service levels. Six clusters have been framed separately for each identified reliability index for routes, dedicated segment, and non-dedicated segment.

As a benchmark to assess the quality of framed clusters as per K-mean clustering analysis, a subsequent cluster validation process has been carried out separately for all the clusters of routes and segments. For this purpose, in the current study, silhouette coefficients of all the clusters have been calculated and average values of each cluster taken as benchmarks and qualities are interpreted. Figure 5.17, 5.18, and 5.19 shows the obtained silhouette coefficient plots for all the reliability indices clusters of routes, dedicated segment, and non-dedicated segment.

From the plots, it is observed that, for most of the clusters of route and segments, calculated silhouette coefficient values fall in the range of 0.51 to 0.70. Hence, with the obtained values it is interpreted that, all the framed clusters are reasonably strong and hence all the clusters are acceptable. Once validating the quality, then reliable service-based clusters such as LOS A to LOS F are finalized at route and segment levels. For all the finalized LOS, thresholds have been fixed based 45^{o} plot. Here in this plot, sorted values with minimum to a maximum range of each reliability index of each cluster are plotted along both the axis, and then thresholds are added in the form of a table. Table 5.21, 5.22, and 5.23 show the thresholds of all the Level of Service (LOS) of PTI, BTI and TTI discretely at the route level, dedicated segment level and non-dedicated segment level. Overall to culminate the analysis carried out, the procedure recommended in this research work is useful for the operators and policymakers for evaluating the practical LOS of the HDBRTS, considering the operation of buses on routes, dedicated segments, and non-dedicated segments.

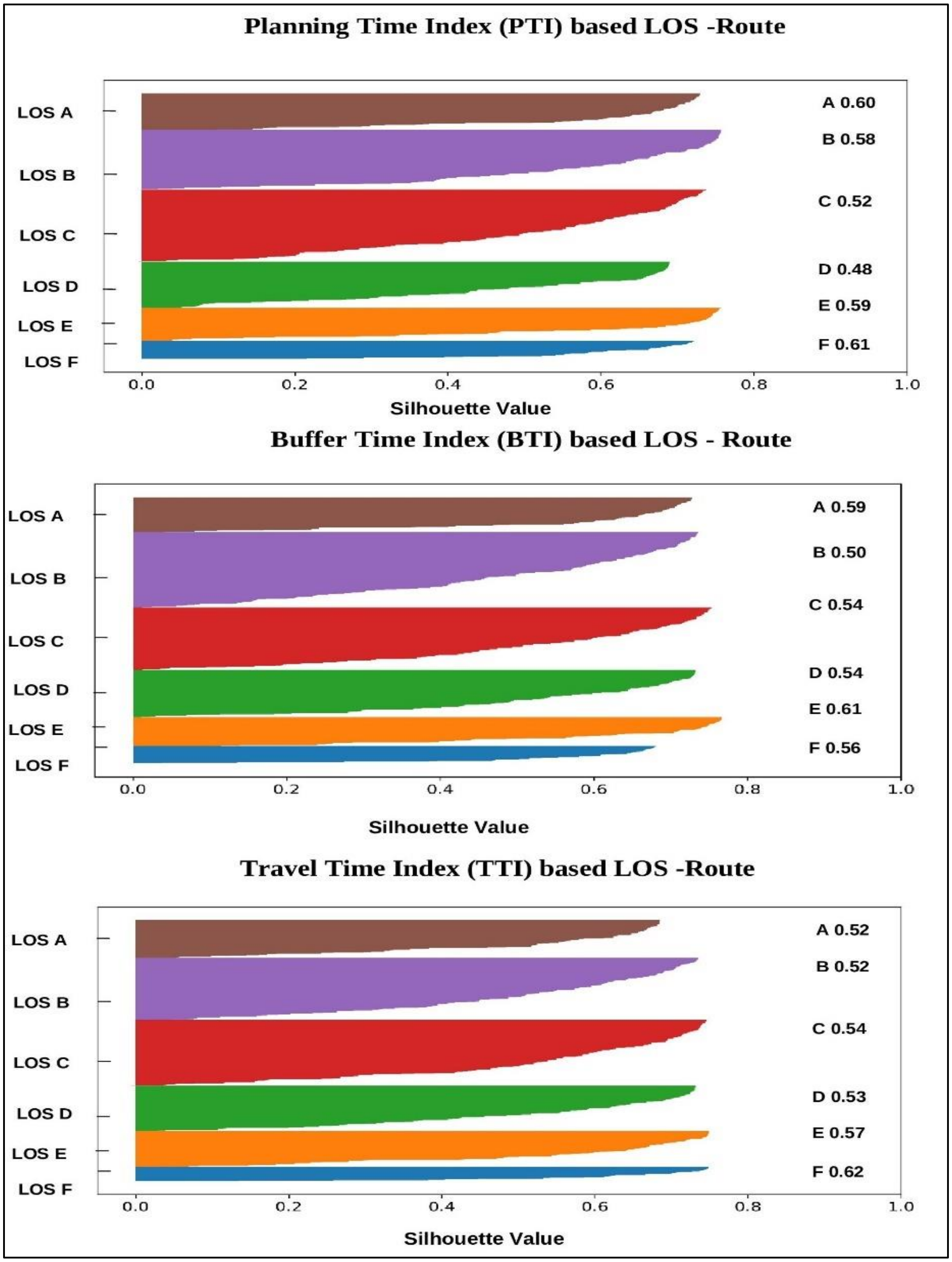


Figure 5.17 Silhouette Coefficient Plot for Route

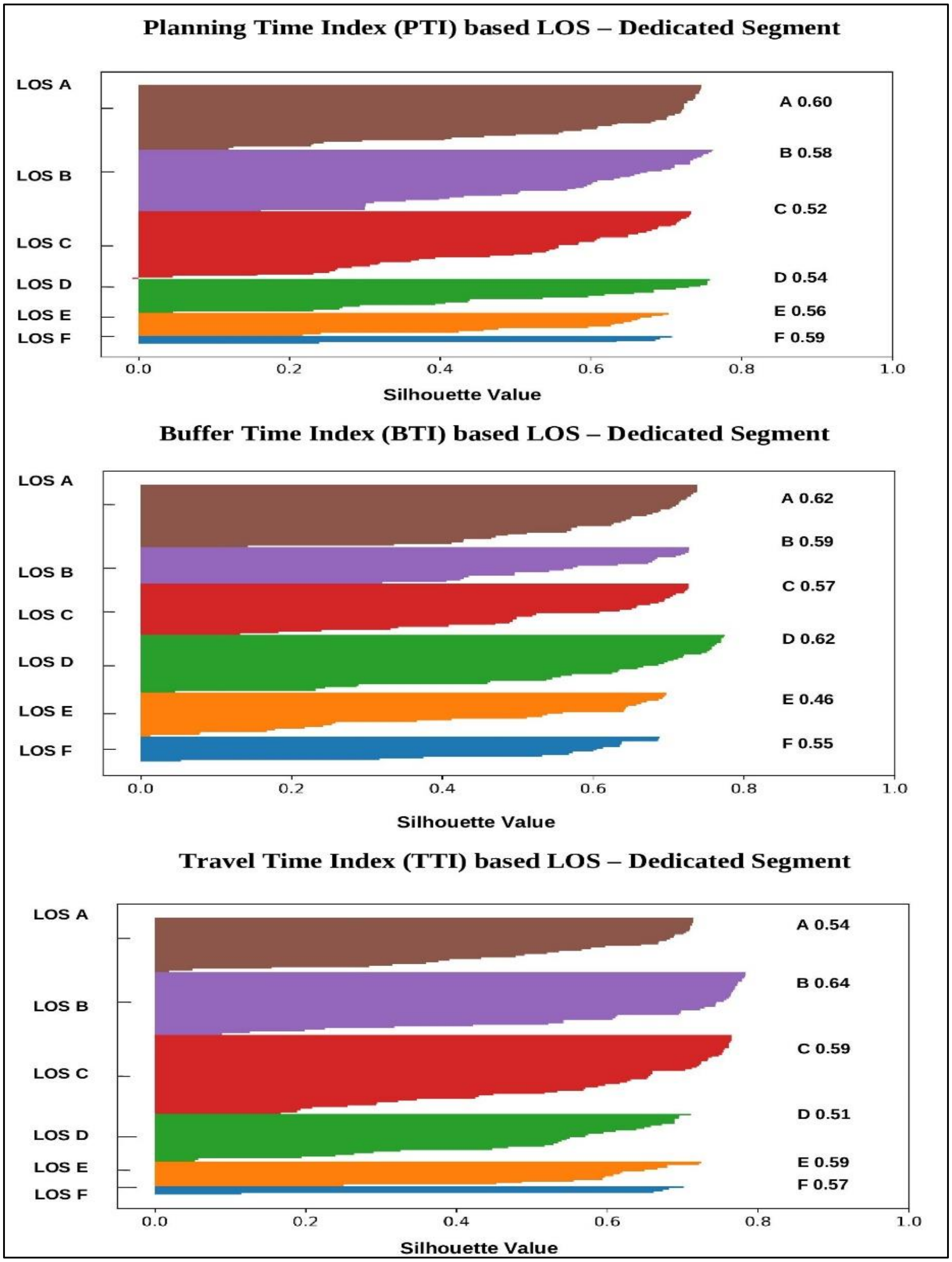


Figure 5.18 Silhouette Coefficient Plot for Dedicated Segment

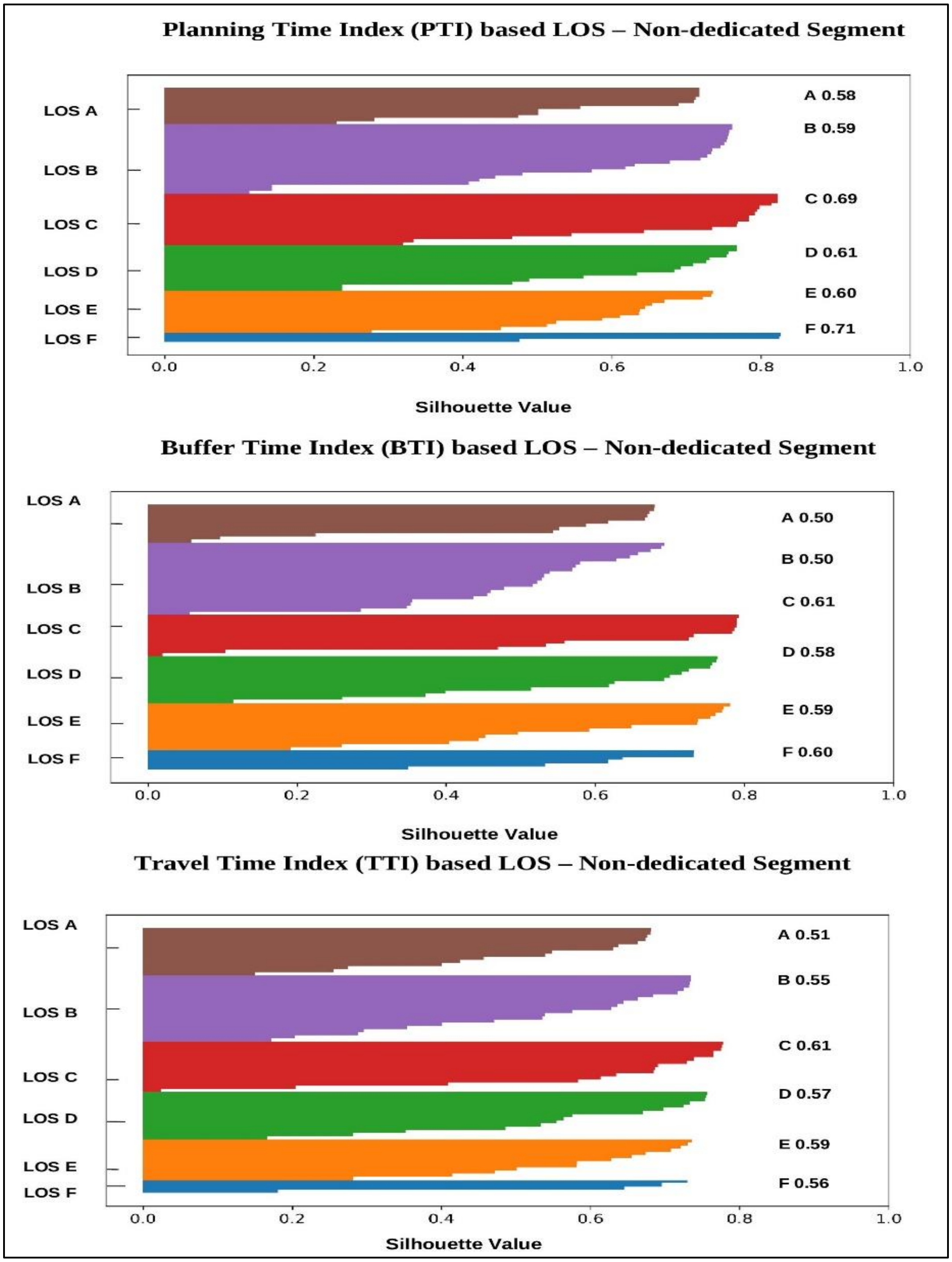


Figure 5.19 Silhouette Coefficient Plot for Dedicated Segment

Table 5.21 Level of Service (LOS) Threshold for Route

Level of Service (LOS)	Routes		
	PTI	BTI	TTI
LOS A	1.22 - 1.51	3.33 - 7.07	1.20 - 1.39
LOS B	>1.51 - 1.64	>7.07 - 10.00	>1.39 - 1.48
LOS C	> 1.64 - 1.75	>10.00 - 12.83	>1.48 - 1.57
LOS D	>1.75 - 1.92	>12.83 - 16.41	>1.48 - 1.57
LOS E	>1.92 - 2.19	>16.41 - 21.40	>1.57 - 1.69
LOS F	>2.19	>21.40	>1.69

Table 5.22 Level of Service (LOS) Threshold for Route for Dedicated Segment

Level of Service (LOS)	Dedicated Segment		
	PTI	BTI	TTI
LOS A	1.43 - 1.94	4.21 - 7.03	1.33 - 1.71
LOS B	>1.94 - 2.35	>7.03 - 10.04	>1.71 - 2.00
LOS C	>2.35 - 2.81	>10.04 - 13.69	>2.04 - 2.39
LOS D	>2.81 - 3.34	>13.69 - 15.90	>2.39 - 2.93
LOS E	>3.34 - 4.40	>15.90 - 18.14	>2.93 - 3.50
LOS F	>4.40	>18.14	>3.50

Table 5.23 Level of Service (LOS) Threshold for Non-dedicated Segment

Level of Service (LOS)	Non-dedicated Segment		
	PTI	BTI	TTI
LOS A	1.71 - 1.97	13.05 - 15.00	1.48 - 1.73
LOS B	>1.97 - 2.18	>15.00 - 16.84	>1.73 - 1.92
LOS C	>2.18 - 2.57	>16.84 - 18.51	>1.92 - 2.18
LOS D	>2.57 - 3.00	>18.51 - 20.57	>2.18 - 2.47
LOS E	>3.00 - 3.49	>20.57 - 22.96	>2.47 - 2.76
LOS F	>3.49	>22.96	>2.76

5.5 SUMMARY

This chapter attempts to accomplish all the methodologies framed for the objectives of the current research work. Travel time variability study at different temporal and spatial pattern have shown that GEV distributions have the highest mean p-value and lowest SD value of the KS test of all the cases. Also, it has a higher ratio of cases passed, a higher ratio of cases in the top 3 position, and a higher ratio of case in the top 1 position. Hence GEV distribution stands first on the best performance distributions list. It shows its robustness in explaining TT of express routes, non-express routes, dedicated segments and even in the case of non-dedicated segments. From the correlation plot, it is found that passenger demand and BTI have a direct correlation with the variations in the GEV shape parameter ‘k’.

Any change in terms of improving the variation in the travel time or worsening it makes following huge positive or negative impacts on the transit service to their commuters. With the higher adjusted R square values of 0.795 and 0.804, respectively, ATT and BT as dependent variables in the study have shown superior explanatory power in describing the system's reliability. Overall, results obtained from modelling of the TTR of the system with different covariates have shown good significance in both the models. The results showed that LSTM is better performed over the Seasonal ARIMA in modelling the passenger demand, the MAE % is reduced by 27.46 % with respect to the MAE % of seasonal ARIMA, and RMSE % is reduced by 31.08 % with respect to the RMSE % of seasonal ARIMA.

The LOS clusters are developed based on reliable transit service for the three operations conditions of the HDBRTS such routes, dedicated segment and non-dedicated segment. Cluster formation followed K-mean clustering technique and quality of framed clusters assessed through silhouette coefficient plots. From the plots, it is observed that, for most of the clusters of route and segments, calculated silhouette coefficient values fall in the range of 0.51 to 0.70. Hence, with the obtained values it is interpreted that, all the framed clusters are reasonably strong and hence all the clusters are acceptable.

CHAPTER 6

CONCLUSIONS AND RECOMMENDATIONS

6.1 GENERAL

The main aim of the current research work is the performance analysis of HDBRTS, considering its various types of operational data and the operational condition of the buses. AVL, APC, QR, and station location data are some of the operational data used in the analysis, and buses traveling along the dedicated and non-dedicated segments are some of the operational conditions considered.

The study is divided into different objectives: a literature review is done, methodologies are framed, analysis is carried out, results are obtained, and accordingly, the discussion are made in the previous chapters. Subsequently to the results and discussion in the last chapter, the conclusions are drawn in the current chapter and presented individually for each objective framed in the study.

6.2 TRAVEL TIME VARIABILITY STUDY OF THE SYSTEM - CONCLUSIONS

The travel time reliability of the transit system is the consequent measurement of its travel time variability. These variations in the transit travel time are caused due to various reasons such as headway irregularities, an impedance of the roadside frictions, bus station characteristics, land use pattern, conditions of the bus operation, improper passenger demand management, intersections density, etc. Thus, assessing the travel time variability of the system is the indirect measurement of all their effects on transit performance. It plays a vital role in enhancing the system's performance and has clarity on its reliable service to the commuters.

The current study considered the HDBRTS as a case study for the systematic travel time variability analysis. This system is recently implanted in the northern part of Karnataka and is still in the expansion stage. Hence, it is imperative to carry out such performance-enhancing and policy-making studies at the current stage of the system.

Analysis has been done for the two routes (express and non-express) and three segments exclusively (two dedicated and one non-dedicated). Data collection is the significant step in this study; 3 months of ITS-based AVL and APC data was collected from the ITMS of HDBRTS. Subsequently, travel time data points have been extracted for all the days of the week and different hours of the days for the route-level analysis and different periods of the day for segment-level analysis. Hence, the study shows the analysis of different spatial and temporal aggregation patterns.

Travel time variability analysis for both route and segment level is carried out in two stages. In the first stage, descriptive statistics analysis of the selected data points is carried out to understand variations in the descriptive statistics values about the different temporal and spatial aggregations. In the second stage, probability distribution fit is carried out for both the routes and selected segments separately to view the TTV comprehensively.

The current study attempts to adopt the approach of explaining variation in the TTR and passenger density with the variation in the parameters of the best fit distribution at different temporal aggregation levels. The most important conclusions drawn from the Objective 1 of the study are as follows,

- For the express routes, travel times trends are observed to be almost identical on all the days of the week. However, in the case of non-express routes, it is seen that passenger demand in-between stations of terminals are less on weekends compared to weekdays, and thus TT values on weekends are a bit lower than on weekdays.
- From the descriptive statistics analysis from route and segment level, it has been found that peak hours have less CV in the TT, and hence average TT values moved towards the higher side. This condition is seen vice-versa during off-peak hours. Thus, it is ensuring that off-peak hours have a lesser impact on the variation in the TT of the buses.

- Off-peak hours have higher SD than peak hour values and are relatively lower. With the values of SD on weekdays and weekends, it is interpreted that most of the travel time data points in the peak hours are on the higher side, and hence dispersion is less.
- TTI and PPI are directly proportional to the bus's average travel time in route and segment level analysis.
- Higher BTI values are observed during augmented average travel times; meanwhile, there are lower SD values of the travel times during peak hours compared to off-peak hours, and around 64.8 % increased buffer BTI values are observed during peak hours.
- Based on route and segment level analysis, it has been observed that peak and off-peak hours have a direct influence on the change in the characteristics of travel time and subsequent reliability indices considered in the current study.
- Except for the higher values of reliability indices during peak hours, the performance of the express routes seems more reliable considering the total travel time that buses have taken to complete the whole length of the route.
- Non-express routes have an average of around 29.4 % and 30.3 % increase in the TT values compared to express routes during the peak and off-peak hours, respectively; this significantly impacts passenger flow characteristics at all the stations along with its distance coverage.
- Even though there is a higher length of dedicated segments, differences between average travel times of dedicated and non-dedicated segments are smaller. It clearly shows the impedance caused by the heterogeneous traffic conditions in the non-dedicated lane of BRTS bus operations.
- Apart from being on par with the average travel times of dedicated segments, the non-dedicated segment also has higher side of TTI, BTI, and PTI values during the peak hours of the weekdays and weekends compared to the dedicated segments. This also signifies the impact of the unsegregated way of BRTS bus operations on the commuters.

- Based on the descriptive statistics analysis for the segments, it is found that the non-dedicated lane of bus operation is affecting the system's overall performance. Further, it is intensified due to peak hour operations, high passenger demand, improper lane discipline followed by mixed traffic, etc. On the contrary, dedicated lanes have good backing to the HDBRTS system in keeping the performance to the required level.
- Based on the obtained probability distribution results of routes and segments, it is inferenced that, GEV distributions have the highest mean p-value and lowest SD value of the KS test. Also, it has a higher ratio of cases passed, a higher ratio of cases in the top 3 positions, and a higher ratio of cases in the top 1 position. Hence, GEV distribution stood first on the best performance distributions list of the current study. It shows its robustness in explaining TT of express routes, non-express routes, dedicated segments, and even in the case of non-dedicated segments.
- In the current study, the impact of the peak and off-peak hours is seen on the PDF shape of the distributions. During peak hours all the TT data points are falling either in the mid-range or at higher range, hence its PDF shape resembles the normal distribution, whereas, during off-peak hours, most of the TT data points falling towards lower range and hence it is positively skewed for all the distributions considered in the analysis.
- Herein, it has tried to correlate the segment shape parameter of best fit GEV distribution with the passenger demand and Buffer Time Reliability Index of the segment at different weekday and weekend periods. A higher negative ‘k’ value is observed with higher passenger demand, and a lower negative or positive value ‘k’ value is observed with lower passenger demand. BTI has a direct correlation with shape parameter ‘k.’ Higher BTI is observed with higher negative ‘k,’ whereas lower BTI is with a lower negative ‘k’ value. Hence, it can be concluded that passenger demand and BTI directly correlate with the variations in the GEV shape parameter ‘k.’

- From this approach, it is also observed that the parameter variations trend is the same in both dedicated and non-dedicated segment plots for weekdays and weekends. But higher parameters value have been observed in the non-dedicated segment plot for weekdays and weekends. This condition evidently depicts the adverse impact of non-dedicated segments on the distribution and travel time reliability parameters.

6.3 MODELLING TRAVEL TIME RELIABILITY OF THE SYSTEM - CONCLUSIONS

Travel time reliability modelling of the HDBRTS is carried out considering operator- and passenger-based dependent variables. Both the dependent variables have shown superior explanatory power on all the covariables with good statistical significance values. Following are the conclusions drawn from this objective,

- With the higher adjusted R square values of 0.795 and 0.804, respectively, ATT and BT as dependent variables in the study have shown superior explanatory power in describing the system's reliability.
- All the covariables considered in the study have shown their significant impact on the model developed with acceptable statistical significance values.
- Passenger demand is the proxy variable of dwell time caused at the individual stations; hence it has shown greater statistical significance in both the models by impacting the travel time reliability of the system.
- The impact of passenger demand on the TTR could be an important observation for the policy-making with systematic scheduling of bus frequencies during peak hours.
- Dummy variables corresponding to peak and off-peak periods are shown to be significant with models; hence properly managing the various traffic incidences, such as signal preferences for the buses, dwell time reduction at the stations, etc., may help in improvising the TTR of the system.
- Consistent with the past literature, the length of the segments has the following higher impact on the TTR of the system. The shorter the segment, the lesser the variation in the TT, and less is the effect on the reliability.

- Concerning the statistical significance values of bus stop density and intersection density in the current models, as express bus service is not available for these segments, it can be suggested to increase the non-express frequency, especially during peak hours. So, that this strategy could handle the higher dwell times caused at these segment stations due to high passenger demand.
- Mixed traffic conditions and commercial land use patterns of the non-dedicated segment have a more significant impact on the overall TTR of the system. The segment's performance represents the system's performance; reliable segments need to control significant delays caused at the stops, during bus manoeuvring along the mixed traffic condition, and at the unsignalised and singnalised intersections. Hence, it is suggested to minimize the interaction of BRTS buses with mixed traffic by taking strategies like grade superiors, signal preferences etc.
- Overall, outcomes of such systematic modelling of the TTR of the system with different covariates will help in tackling their heterogeneity effect on the system performance.
- Both TTR models developed based on Average TT and BT as dependent variables have shown satisfactory performance; hence it is proposed to use both the models to assess the systems TTR considering operators and passengers perspectives.

6.4 PASSENGER DEMAND FORECASTING OF THE SYSTEM – CONCLUSIONS

In this research work, an attempt is made to explore the unexplored area of LSTM in passenger demand forecasting by considering HDBRTS as a case study. The study also focuses on addressing the confusion about the resampling time interval to get the best forecast results. As this area remains unexplored, this work explored the time series results, resampled with different time intervals, to find the most suitable time interval which gives the best forecasting results. Then, the passenger demand is forecasted using LSTM with the APC data for three months obtained from the HDBRTS. Furthermore, seasonal ARIMA is also used to forecast passenger demand, a comparison of the forecasting accuracy is made, and the seasonal naïve method is used as a benchmarking method. Following are the conclusions made out of this objective of the study,

- The results showed that LSTM is better performed over the Seasonal ARIMA in modelling the passenger demand, the MAE % is reduced by 27.46 % with respect to the MAE % of seasonal ARIMA, and RMSE % is reduced by 31.08 % with respect to the RMSE % of seasonal ARIMA.
- It is found that the time-series resampled with 60 minutes intervals gave better results than the 15 minutes, 30 minutes, and 45 minutes intervals.
- SARIMA is a traditional linear model which fails to describe the stochastic and non-linear nature of passenger demand. In such a state, deep learning-based models play their role effectively on passenger demand forecasting. These models will have good outcomes on spatial and temporal evolution of passenger demand flow. LSTM, is one of such deep learning-based models, which captures the characteristics of time series, combines underlying features and works on all modelling issues mentioned above efficiently.
- However, in the current study, it has also been understood that LSTM gives reliable results with larger datasets and requires more time. Nevertheless, if the necessary resources are available, one can always go with the deep learning-based model LSTM more effectively over traditional methods of forecasting.
- Finally, with this research work, it can be concluded that LSTM models can be satisfactorily used to forecast passenger demand with data obtained by APC of HDBRTS. The basic version of LSTM is used for forecasting in this study. Still, the research can be continued to understand and evaluate the forecasting accuracy of different variants of LSTMs and Gated Recurrent Units (GRUs) as a scope of future work.

6.5 LEVEL OF SERVICE (LOS) OF HDBRTS – CONCLUSIONS

Reliable transit service-based Level of Service (LOS) has been developed for HDBRTS. Travel time index, planning time index, and buffer time index are the three-reliability measures of transit travel time considered in the analysis. The route which operates express, non-express buses, dedicated segment, and non-dedicated buses are the three operational conditions of HDBRTS buses considered for the service establishment. Following are the conclusions made out of this objective of the study,

- Six service clusters are formed based on analysis with the K-mean clustering technique. Such as service cluster A to service cluster F, which are universally recognized service levels. Six clusters have been framed separately for each identified reliability index for routes, dedicated and non-dedicated segments.
- From the obtained ranges of each service level, it is interpreted that an upsurge in the values of selected indicators implies a reduction in the service level of all three operational conditions, such as routes and dedicated and non-dedicated segments.
- K-mean clustering method is adopted to frame the clusters, and then the quality of those clusters is efficiently checked by calculating the average silhouette coefficient value individually. Most of the clusters are reasonable and apt, with an average silhouette coefficient of more than 0.5. So, it is interpreted that the method adopted for LOS development in the current study better suits the selected reliability indicators data point.
- In the case of BTI-based LOS, the state of transit functioning on the non-dedicated segment seems to be at a lower level with higher range values of each service level. Ranges specified for each service level are more than dedicated segment-based LOS ranges. This is mainly due to the significant effect of mixed traffic conditions on the operation of BRTS buses.
- Overall, to culminate the analysis carried out, the procedure recommended in this research work is helpful for the operators and policymakers in evaluating the practical LOS of the HDBRTS, considering the operation of buses on routes, dedicated segments, and non-dedicated segments

6.6 MAJOR STRATEGIES OR RECOMMENDATIONS FROM THE RESEARCH WORK

The major recommendations from the current research work are as follows,

Short-term: Attainable in a short duration upto six months. It is based on making minor modifications to the currently running system.

- The performance of the express routes seems to be tremendously good considering the total travel time that buses have taken to complete the whole length of the route and lesser TTV. However, it has been observed that, these bus operations are affected by signal delays and bus bunching at major stations like BVB, Vidyagiri, and Vidyanagar due to the mixed types of demand. Hence "Providing Transit Signal Priority (TSP) for BRTS buses, at least at these major stations, will lead to a more enhanced end-to-end travel time including non-express route buses to some extent."
- Currently, most of the BRTS bus stations are located nearby signalised and unsignalised intersections, and hence intersection density has an impact on the reliability of specifically non-express bus operation, "Adopting the signal priority approaches at major intersections such as at BVB, Vidyanagar, Gandhinagar, Vidyagiri, KIMS (Karnataka Institute of Medical Science) will definitely enhance the end to end travel time of the buses along with providing reliable service to all the passengers at all the bus stations."
- The segment between the Hosur circle to Dr. B. R. Ambedkar circle is being non-dedicated in nature, affecting the overall system's performance as per the analysis carried out. Further, it is intensified due to peak hour operations, high passenger demand, improper lane discipline followed by mixed traffic, etc. "Complete dedicated lane operation is the need for eliminating this complexity in effect; however, this effect on the BRTS operation can be avoided at least during peak hour operations by taking short-term policy measures like diverting the mixed

traffic vehicles from Hosur circle - new cotton market - Channamma circle during peak hours and providing TSP for BRTS buses at Basava Vana Signal."

- It should be appreciated that the APC data of HDBRTS comprises a separate column called "rider type," which currently has 7 types. However, "more types could be added (ex: Teachers, Different Job Holders, etc.) which can help in increasing the accuracy of forecasting for taking decisions on many strategic policies."

Mid-term: Attainable in the time duration of six months to one year. Based on sufficient design backup of planned strategic policy.

- Based upon the user-oriented qualitative study, particularly at BRTS stations located near major junctions like Hosur Circle BRT station, Vidyanagar BRT station, BVB BRT station, and Jubilee Circle BRT station, have shown more interest in feeder service requirements. Also, HDBRTS is a single linear corridor; the feeder system and integration with other public transport services become crucial from an as futuristic point of view of the system. "Providing the feeder service at these stations may enhance the modal shift from private vehicles and other public transit services. Meanwhile, it will be more convenient for the passenger to start their journey at the doorstep."
- Being a single linear corridor, HDBRTS has a good direct connection to many of the important locations between Hubli-Dharwad. As a result of this, high ridership is commonly observed along the corridor. This is also seen in the TTR modelling as passenger demand for dedicated and non-dedicated stations has a higher positive effect. On the other hand, high ridership creates the leading impact on the HDBRTS performance, particularly from the user point of view, during peak hours as they meet with bus bunching situations and higher waiting times at stations like BVB, Vidyanagar, OCBS, Vidyagiri, Jubilee Circle and finally lead higher journey time. "Therefore, strategies like bus demand-based scheduling, schedule adherence

by the drivers and balancing the demand at the bus doors may improve the overall reliability."

- With Peak flow analysis of all the 35 stations of HDBRTS, the highest flow occurs between 09:00 to 11:00 in the morning, 12:00 to 1.30 noon, and 05:00 to 07:00 in the evening. Reliability-based LOS values of the route for the mentioned hours as well as segments are reaching the LOS D to LOS E, sometimes LOS F. Therefore, "preference-based scheduling can be done along with proper headway adherence, in such a way that the passenger demand is satisfied with less waiting as it will improve the LOS of routes and segments."
- Currently, there are no such guidelines nationally or internationally to establish the LOS criteria, particularly for BRTS. But Service Level Benchmarks (SLB) for Urban Transport have been given by the Ministry of Urban Development (MoUD) in 2012 is serves as guidelines for establishing the LOS of any urban public transit system in general. SLB for urban transport by MoUD considers ten service levels to establish the LOS of the transit system, such as pedestrian infrastructure, travel speed, road safety, pollution level, etc. It does not consider TTR measures to develop the LOS of the transit system, particularly for BRTS. "Due to this limitation, the SLB for urban transport by MoUD does not considered in the current research work to compare with obtained results. Hence, MoUD can consider this limitation for adding TTR measures as one of the Service Level Benchmarks for Urban Transit Systems."

Long-term: Attainable in the time duration of more than one year. Basically, based on making major changes in the existing infrastructure-related facilities of the system.

- The segment between Hosur circle to Dr. B. R. Ambedkar circle is also being used by the vehicles that are needed to reach Bangalore and Hyderabad highway. Usually, that segment capacity seems to be less and affects the BRTS bus operation. As this segment is located along the CBD area of Hubli, and has appropriated road space from the other traffic to the BRT system, further expansions of the BRT right-of-way may be complicated and unrealistic, Hence, “a long-term policy providing grade separator will definitely boost the overall performance of the HDBRTS buses, as out-going traffic use grade separator, and that can lead to the increased capacity of BRTS lane.”

- Considering the main question of the study, the HDBRTS can undoubtedly be considered a success, since it carries a significant number of passengers, providing them with significant time savings when compared to pre-BRTS conditions that existed between Hubli-Dharwad. In the future, if this corridor is redeveloped in the form of an integrated transit system with feeders and other modes of transit, the demand will surely meet with the higher capacity that those new systems are capable of providing.

6.7 FUTURE SCOPE OF THE RESEARCH WORK

Future scope of the study are as follows,

- Travel time variability analysis can be done at different departure time windows to understand the percentage differences in the TTV of all the time windows.

- Performance of the BRTS can also be done considering its Capacity and Speed as measures. Every BRT systems will have its own dynamically operating environments and based on that operational characteristics will also vary. In detail capacity and speed comparison studies can be taken up for the multiple Indian BRT systems.

- The basic version of LSTM is used for forecasting in this study but the research can be continued to understand and evaluate the forecasting accuracy of different variants of LSTMs and Gated Recurrent Units (GRU’s) as a scope of future work.

- LOS of the BRTS system can also be established using Service Level Benchmark for urban transport by MoUD and obtained results can be compared with guidelines.

APPENDIX

A.1. Sample AVL data obtained from HDBRTS Operators

1	id	asset_id	created_at	direction	driver_id	ignition	latitude	longitude	odometer	processed	route_id	route_latitude	route_longitude	speed	speed_source	t1	t2	t3	t4	trip_id
2	6.6953E+17	723	00:02.7	106.84		TRUE	15.351	75.1366	2048	FALSE	953	15	75	0	1	00:20.5	00:20.5	00:20.5	00:02.7	6.69521E+17
3	6.6953E+17	697	00:07.0	152.3	2596	TRUE	15.4462	75.0128	2048	FALSE	972	15	75	0	1	00:46.8	00:46.8	00:46.8	00:07.0	6.69526E+17
4	6.6953E+17	723	00:10.1	78.77		TRUE	15.351	75.1366	2048	FALSE	953	15	75	0	1	00:28.2	00:28.2	00:28.2	00:10.1	6.69521E+17
5	6.6953E+17	697	00:10.1	155.78	2596	TRUE	15.4461	75.0128	2048	FALSE	972	15	75	10	1	00:52.3	00:52.3	00:52.3	00:10.1	6.69526E+17
6	6.69531E+17	723	00:17.8	78.77		TRUE	15.351	75.1366	2048	FALSE	953	15	75	0	1	00:35.5	00:35.5	00:35.5	00:17.8	6.69521E+17
7	6.69531E+17	723	00:18.4	78.77		TRUE	15.351	75.1366	2048	FALSE	953	15	75	0	1	00:37.3	00:37.3	00:37.3	00:18.4	6.69521E+17
8	6.69531E+17	697	00:21.6	157.84	2596	TRUE	15.4457	75.013	2048	FALSE	972	15	75	30	1	00:58.3	00:58.3	00:58.3	00:21.6	6.69526E+17
9	6.69531E+17	697	00:22.8	157.69	2596	TRUE	15.4457	75.013	2048	FALSE	972	15	75	31	1	00:59.3	00:59.3	00:59.3	00:22.8	6.69526E+17
10	6.69531E+17	723	00:25.2	78.77		TRUE	15.3511	75.1366	2048	FALSE	953	15	75	0	1	00:43.3	00:43.3	00:43.3	00:25.2	6.69521E+17
11	6.69531E+17	697	00:31.1	154.7	2596	TRUE	15.4444	75.0135	2048	FALSE	972	15	75	45	1	01:12.9	01:12.9	01:12.9	00:31.1	6.69526E+17
12	6.69531E+17	697	00:40.7	132.5	2596	TRUE	15.4435	75.0143	2048	FALSE	972	15	75	53	1	01:22.3	01:22.3	01:22.3	00:40.7	6.69526E+17
13	6.69531E+17	697	00:41.7	133.16	2596	TRUE	15.4432	75.0146	2048	FALSE	972	15	75	53	1	01:26.1	01:26.1	01:26.1	00:41.7	6.69526E+17
14	6.69531E+17	723	00:47.1	101.39		TRUE	15.351	75.1368	2048	FALSE	953	15	75	5	1	01:03.2	01:03.2	01:03.2	00:47.1	6.69521E+17
15	6.69531E+17	697	00:49.6	136.65	2596	TRUE	15.4427	75.0152	2048	FALSE	972	15	75	38	1	01:32.1	01:32.1	01:32.1	00:49.6	6.69526E+17
16	6.69531E+17	723	00:54.8	124.4		TRUE	15.3507	75.137	2048	FALSE	953	15	75	4	1	01:12.5	01:12.5	01:12.5	00:54.8	6.69521E+17
17	6.69531E+17	723	00:55.5	123.13		TRUE	15.3507	75.137	2048	FALSE	953	15	75	3	1	01:14.3	01:14.3	01:14.3	00:55.5	6.69521E+17
18	6.69531E+17	697	01:00.4	142.7	2596	TRUE	15.4418	75.016	2048	FALSE	972	15	75	42	1	01:43.6	01:43.6	01:43.6	01:00.4	6.69526E+17
19	6.69531E+17	723	01:03.7	118.68		TRUE	15.3506	75.1371	2048	FALSE	953	15	75	12	1	01:20.3	01:20.3	01:20.3	01:03.7	6.69521E+17

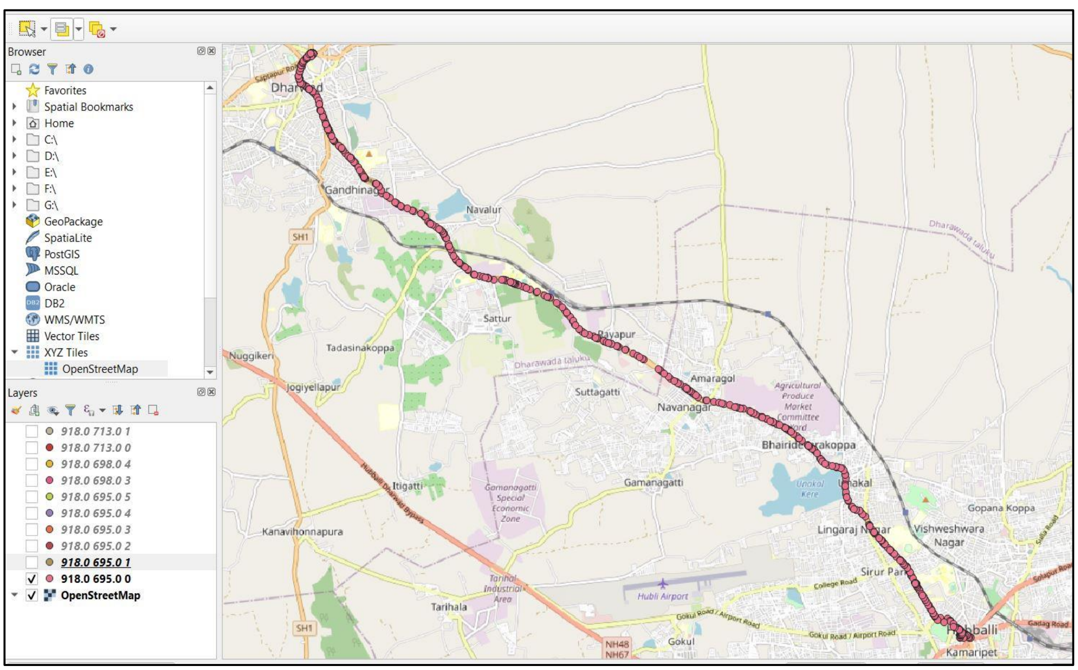
A.2. Sample APC data obtained from HDBRTS Operators

1	Date Issued	Time Issued	Operator Id	Device Type	Issued Boarding Station	Issued Alighting Station	Payment Method	Ticket Serial Number	Rider Type	Ridership	Total revenue
2	12/8/2019	23:05:49	P070	OFFICE_POS	HUBBALLI CENTRAL BUS TERMINAL - DH	RAILWAY STATION - DH	CASH	309C2330DF86230549-651759	Adult	1	10
3	12/8/2019	23:05:48	P070	OFFICE_POS	HUBBALLI CENTRAL BUS TERMINAL - DH	RAILWAY STATION - DH	CASH	309C2330DF86230548-651758	Adult	1	10
4	12/8/2019	23:05:13	P070	OFFICE_POS	HUBBALLI CENTRAL BUS TERMINAL - DH	RAILWAY STATION - DH	CASH	309C2330DF86230513-651757	Adult	1	10
5	12/8/2019	23:04:10	P070	OFFICE_POS	HUBBALLI CENTRAL BUS TERMINAL - HD	DHARWAD BRTS TERMINAL - HD	CASH	309C2330DF86230410-651756	Adult	1	22
6	12/8/2019	23:04:07	P070	OFFICE_POS	HUBBALLI CENTRAL BUS TERMINAL - HD	DHARWAD BRTS TERMINAL - HD	CASH	309C2330DF86230407-651755	Monthly Pass	1	2
7	12/8/2019	23:03:53	P070	OFFICE_POS	HUBBALLI CENTRAL BUS TERMINAL - DH	RAILWAY STATION - DH	CASH	309C2330DF86230353-651754	Adult	1	10
8	12/8/2019	23:03:52	P070	OFFICE_POS	HUBBALLI CENTRAL BUS TERMINAL - DH	RAILWAY STATION - DH	CASH	309C2330DF86230352-651753	Adult	1	10
9	12/8/2019	23:03:51	P070	OFFICE_POS	HUBBALLI CENTRAL BUS TERMINAL - DH	RAILWAY STATION - DH	CASH	309C2330DF86230351-651752	Adult	1	10
10	12/8/2019	23:03:50	P070	OFFICE_POS	HUBBALLI CENTRAL BUS TERMINAL - DH	RAILWAY STATION - DH	CASH	309C2330DF86230350-651751	Adult	1	10
11	12/8/2019	23:03:49	P070	OFFICE_POS	HUBBALLI CENTRAL BUS TERMINAL - DH	RAILWAY STATION - DH	CASH	309C2330DF86230349-651750	Adult	1	10
12	12/8/2019	23:02:21	P070	OFFICE_POS	HUBBALLI CENTRAL BUS TERMINAL - DH	RAILWAY STATION - DH	CASH	309C2330DF86230221-651749	Adult	1	10

A.3. Numbering of HDBRTS stations followed in the Study

Station No.	Station	Station No.	Name
1	CBT Hubballi	19	Iskcon Temple
2	Railway Station	20	Rayapur
3	Dr B R Ambedkar Circle	21	KMF 1
4	HDMC	22	Navalur Railway Station
5	Hubballi Central Bus Terminal	23	SDM Medical College
6	Hosur Cross	24	Sattur
7	Hosur Regional Terminal	25	Navalur (Yet to start)
8	KIMS	26	Lakhamanahalli
9	Vidyanagar	27	Gandhi Nagar
10	BVB College	28	Vidyagiri
11	Unakal Cross	29	Toll Naka
12	Unakal	30	Hosa Yallapur Cross
13	Unakal Lake	31	NTTF
14	Bairidevarakoppa	32	Court Circle
15	Shantinikethan	33	Jubilee Circle
16	APMC 3rd Gate	34	Dharwad BRTS Terminal
17	Navanagar	35	Dharwad New Bus Stand
18	RTO Office		

A.4. Sample AVL data plotted on QGIS



A.5. Route Details of HDBRTS

Route	Route Name	Route_id	Origin	Destination
200A	200A-U	1246	CBT Hubli	Dharwad BRTS terminal
	200A-D	937	Dharwad BRTS terminal	CBT Hubli
200D	200D-U	1428	Hubli BRTS terminal	Dharwad BRTS terminal
	200D-D	1429	Dharwad BRTS terminal	Hubli BRTS terminal
201B	201B-U	1248, 1430	Railway Station	Dharwad New Bus Stand
	201B-D	953, 1431	Dharwad New Bus Stand	Railway Station
100D	100D-U	918	Hubli BRTS terminal	Dharwad BRTS terminal
	100D-D	921	Dharwad BRTS terminal	Hubli BRTS terminal
100D1	100D-U1	958	Hubli BRTS terminal	Dharwad BRTS terminal
	100D-D1	926	Dharwad BRTS terminal	Hubli BRTS terminal

A.6. Python Code used for Travel Time Extraction

```
import pandas as pd

file_name='Jan 01-31 200D-D'

dataset=pd.read_csv('')

dataset.head()

dataset.tail()

dataset.columns

#dataset.sort_values(by=['GPS_time'],inplace=True)
```

```
Dataset

dataset.columns

import datetime

df_list=[]

trip_id_list=[]

for index,df_group in dataset.groupby('trip_id'):

    trip_id_list.append(index)

    df_group['Hours']=[(ele.split('.')[0]) for ele in df_group['GPS_time']]

    df_group.sort_values(by=['Hours'],inplace=True)

    end_time=list(df_group['GPS_time'])[-1]

    start_time=list(df_group['GPS_time'])[0]

    end_time_strf=datetime.datetime.strptime(end_time, '%H:%M:%S')

    start_time_strf=datetime.datetime.strptime(start_time, '%H:%M:%S')

    diff=end_time_strf-start_time_strf

    asset_id= df_group['asset_id'].iloc[0]

    route_id= df_group['route_id'].iloc[0]

    date=df_group['GPS_date'].iloc[0]

    df_list.append([asset_id,route_id,date,index,start_time,end_time,diff.seconds])

temp_df=pd.DataFrame(df_list,columns=['asset_id','route_id','date','trip_id','departure_time','arrival_time','travel_time'])
```

```
temp_df

temp_df.to_excel('TT '+file_name+'.xlsx')

temp_df.to_csv('TT '+file_name+'.csv')
```

A.7. Python Code used for Hour wise Split the Travel Times

```
import pandas as pd

filename='31.12.2019'

df=pd.read_excel(.xlsx')

df.head()

df['departure_hour']=[int(str(data).split(":")[0]) for data in list(df['Time Issued'])]

df.head()

flag=0

for hour,temp_df in df.groupby('departure_hour'):

    arrival_time=list(temp_df['Ridership'])

#     trip_id=list(temp_df['Unique_trip_id'])

    hour_n=hour

    if flag==0:

        new_df=pd.DataFrame(arrival_time,columns=[hour_n])

        flag=1

    else:

        index_df=pd.DataFrame(arrival_time,columns=[hour_n])
```

```
    new_df=pd.concat([new_df,index_df], axis=1)
new_df.head()
new_df.to_excel("Hourwise "+filename+".xlsx")
```

A.8. Free Flow Travel Times of Routes and Segments in Seconds (UP and DOWN)

Route	FFTT
100D-D	1319.00
100D-U	1351.00
200D-D	1410.00
200D-U	1454.00
200A-D	1896.00
200A-U	1873.00
201B-D	1657.00
201B-U	1601.00
Hubballi BRTS terminal to Jubilee Circle-D	1247.00
Hubballi BRTS terminal to Jubilee Circle-D	1249.00
Hosur to Dr.B.R. Ambedkar Circle D	225.00
Hosur to Dr.B.R. Ambedkar Circle U	240.00
Bairidevarakoppa to Hosur Circle D	283.00
Bairidevarakoppa to Hosur Circle U	294.00
Jubilee Circle to Lakhamanhalli D	292.00
Jubilee Circle to Lakhamanhalli U	289.00

A.9. Sample of Hour-wise Split Travel Time Data points in seconds

	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
0	1798	1806	1712	1825	1714	1831	1755	1871	1833	1659	1483	1449	1560	2759	2759	2756	2680	2750
1	1800	1851	1798	1829	1795	1832	1842	1895	1882	1700	1562	1506	1642	2753	2757	2728	2679	2734
2	1816	1867	1845	1850	1801	1875	1887	1949	1906	1708	1566	1509	2012	2753	2754	2706	2654	2681
3	1831	1871	1849	1858	1831	1881	1895	1957	1908	1756	1583	1509	2749	2750	2752	2688	2653	2572
4	1835	1880	1854	1864	1841	1887	1896	1957	1923	1769	1591	1533	2745	2750	2744	2677	2650	2485
5	1838	1884	1858	1867	1846	1887	1905	1962	1936	1811	1595	1537	2743	2745	2730	2654	2638	2376
6	1843	1890	1871	1874	1847	1893	1912	1963	1938	1815	1596	1548	2740	2744	2720	2653	2629	2336
7	1848	1896	1879	1882	1852	1900	1915	1974	1939	1835	1597	1551	2738	2744	2718	2647	2625	2336
8	1848	1899	1879	1885	1860	1903	1915	1982	1948	1889	1599	1552	2738	2744	2713	2622	2598	2331
9	1852	1899	1892	1891	1863	1907	1918	1988	1965	1891	1599	1556	2734	2742	2711	2574	2592	2311
10	1859	1905	1897	1901	1871	1908	1925	1991	1970	1894	1608	1556	2733	2736	2707	2563	2590	2309
11	1861	1907	1897	1918	1871	1909	1930	2011	1970	1894	1615	1559	2732	2734	2702	2563	2585	2301
12	1863	1908	1899	1937	1873	1913	1930	2012	1975	1894	1617	1562	2732	2733	2697	2551	2545	2256
13	1886	1909	1902	1938	1877	1915	1948	2014	1984	1902	1626	1563	2724	2729	2695	2549	2483	2227
14	1890	1919	1910	1942	1878	1916	1948	2014	1984	1910	1629	1565	2716	2728	2693	2549	2474	2201
15	1901	1920	1915	1943	1881	1918	1950	2015	2003	1911	1632	1583	2715	2728	2690	2536	2473	2194
16	1911	1920	1920	1943	1885	1919	1951	2023	2003	1918	1635	1587	2710	2726	2690	2528	2452	2192

A.10. Python Code used for Data Points Splitting for 15-minute, 30-minute ,45-minute, and 60-minute Time-frames for Passenger Demand Forecasting

```
import pandas as pd

import datetime

station = ['5','28','33','34','35']

month   = ['dec','jan','feb']

time    = ['15','30','45','60']

for k in range(len(station)):

    for j in range(len(month)):

        for t in range(len(time)):

            df = pd.read_csv('.csv')

            a = []

            # Date_Issued and Time_Issued in the original data are joined to get a single 
timestamp and stored under the column Date

            for i in range(len(df.Date_Issued)):

                dt_object = datetime.datetime.strptime(df.Date_Issued[i] +" "+ 
df.Time_Issued[i], "%d-%m-%Y %H:%M:%S")

                a.append(dt_object)

            df['Date_Issued'] = a

            df.rename(columns = {'Date_Issued':'Issued_at'}, inplace = True)

            df.drop(['Time_Issued'], axis = 1, inplace = True)

            # Sum of the ridership up to required time-period is calculated

            # 15T: 15 minutes, 30T: 30 minutes, 45T: 45 minutes, 60T: 60 minutes

            df = (df.resample(time[t]+'T', on = 'Issued_at').Ridership.sum())
```

```
        df = pd.DataFrame(df)

        df.rename(columns = {'Ridership':'TP'}, inplace = True)

df.to_csv('1.csv')

df = pd.read_csv(.csv')

   # Observations within 23:00:00 to 06:00:00 are removed

 df
=(df[~df.Issued_at.str.contains('23:|00:00:00|00:15:00|00:30:00|00:45:00|01:|02:|03:|04:|0
5:')])

   df.to_csv('D.csv', index = False)

for k in range(len(station)):

   for j in range(len(month)):

      for t in range(len(time)):

        df = pd.read_csv(.csv')

        a = []

        # Date_Issued and Time_Issued in the original data are joined to get a single
timestamp and stored under the column Date

        for i in range(len(df.Date_Issued)):

           dt_object = datetime.datetime.strptime(df.Date_Issued[i] +" "+
df.Time_Issued[i], "%d-%m-%Y %H:%M:%S")

           a.append(dt_object)

        df['Date_Issued'] = a

        df.rename(columns = {'Date_Issued':'Issued_at'}, inplace = True)

        df.drop(['Time_Issued'], axis = 1, inplace = True)

        # Sum of the ridership up to required time-period is calculated

        # 15T: 15 minutes, 30T: 30 minutes, 45T: 45 minutes, 60T: 60 minutes
```

```
df = (df.resample(time[t]+'T', on = 'Issued_at').Ridership.sum())
df = pd.DataFrame(df)
df.rename(columns = {'Ridership':'TP'}, inplace = True)
df.to_csv(.csv')
df = pd.read_csv(.csv')
# Observations within 23:00:00 to 06:00:00 are removed
df =
(df[~df.Issued_at.str.contains('23:|00:00:00|00:15:00|00:30:00|00:45:00|01:|02:|03:|04:|05:
')])
df.to_csv(.csv', index = False)
```

A.11. Python Code used for Passenger Demand Forecasting Analysis Using SARIMA

```
import numpy as np
import pandas as pd
import datetime
import statsmodels
import matplotlib.pyplot as plt
import pmdarima as pm
import statsmodels.api as sm
import math
from pandas.tseries.offsets import DateOffset
from statsmodels.tsa.arima_model import ARIMA
```

```
from statsmodels.tsa.statespace.sarimax import SARIMAX

from statsmodels.graphics.tsaplots import plot_acf, plot_pacf

from statsmodels.tsa.stattools import adfuller

station = 33; time = 45

df = pd.read_csv(.csv')

#df.index = pd.DatetimeIndex(df.index).to_period('30T')

df['Issued_at'] = pd.to_datetime(df['Issued_at'])

df.index = df['Issued_at']

df
```

Obtaining the time-series plot

```
df['TP'].plot(figsize = (14,6))

plt.title('Time series plot of passenger data for station 5, 30 minutes interval', font =
'times new roman', size = 14)

plt.ylabel('Passengers', font = 'times new roman', size = 14)

plt.xlabel('Time', font = 'times new roman', size = 14)

plt.xlim('2019-12-01 06:00:00','2020-02-29 22:30:00')

plt.xticks(font='times new roman', size = 11)

plt.yticks(font='times new roman', size = 11)

plt.minorticks_on()

plt.legend(['Passengers'], prop = 'times new roman')

plt.savefig("jpg")

plt.show()

# # for 15 min
```

```
# train = df[:4964]
# test = df[4964:]
# for 30 min
# train = df[:2482]
# test = df[2482:]
# # for 45 min
train = df[:1679]
test = df[1679:]
# # for 60 min
# train = df[:1241]
# test = df[1241:]
Name = []
AIC  = []
BIC  = []
MAE  = []
RMSE = []
AR  = 1
MA  = 1
SAR = 1
SMA = 1
d   = 1
D   = 1
m   = 34
```

```
for p in range(AR+1):

    for q in range(MA+1):

        for P in range(SAR+1):

            for Q in range(SMA+1):

                #Fitting

                model = SARIMAX(train['TP'].dropna() ,order=(p, d,
q),seasonal_order=(P,D,Q,m))

                results = model.fit()

                #Forecasting

                # # for 15 min

                # predictions = results.forecast(steps=1224).rename("TPF")

                # for 30 min

                predictions = results.forecast(steps=612).rename("TPF")

                # # for 45 min

                # predictions = results.forecast(steps=414).rename("TPF")

                # # for 60 min

                # predictions = results.forecast(steps=306).rename("TPF")

                #Finding MAE and RMSE

                error = list(predictions) - test.TP

                error = pd.DataFrame(error)

                error['squared'] = error.TP.pow(2)

                rmse = math.sqrt(error.squared.mean())

                error['absolute'] = error.TP.abs()
```

```
        mae = error.absolute.mean()

        #Create a summary table

        Name.append("SARIMA (" + str(p) + ", " + str(d) + ", " + str(q) + ") × " + "(" + str(P) + ", " + str(D) + ", " + str(Q) + ", " + str(m) + ")" )

        AIC.append(results.aic)

        BIC.append(results.bic)

        MAE.append(mae)

        RMSE.append(rmse)

        print ("Model Name = SARIMA (" + str(p) + ", " + str(d) + ", " + str(q) + ") × " + "(" + str(P) + ", " + str(D) + ", " + str(Q) + ", " + str(m) + ") | AIC = %.2f" % results.aic, "| BIC = %.2f" % results.bic, "| MAE = %.2f" % mae, "| RMSE = %.2f" % rmse)

summary = pd.DataFrame({'Model Name': Name, 'AIC': AIC, 'BIC': BIC, 'MAE': MAE, 'RMSE': RMSE})

summary
```

A.12. Python Code used for Passenger Demand Forecasting Analysis Using LSTM

```
import os

import tensorflow as tf

os.environ['CUDA_VISIBLE_DEVICES'] = '-1'

if tf.test.gpu_device_name():

   print('GPU found')

else:

   print("No GPU found")

mport numpy as np

import matplotlib.pyplot as plt
```

```
from pandas import read_csv
import math
import pandas as pd
from tensorflow import keras
from keras.models import Sequential
from keras.layers import Dense, SimpleRNN, LSTM
from sklearn.preprocessing import MinMaxScaler
from sklearn.preprocessing import StandardScaler
from sklearn.preprocessing import QuantileTransformer
from sklearn.metrics import mean_squared_error
from sklearn.metrics import mean_absolute_error
from sklearn.metrics import mean_absolute_percentage_error
import sklearn.metrics
from keras import callbacks
# load the dataset
station = 33; time = 60
dataframe = pd.read_csv(.csv', usecols=[1])
# plt.plot(dataframe)
dataset
# Normalization is optional but recommended for neural network as certain
# activation functions are sensitive to magnitude of numbers.
# normalize the dataset
# scaler = MinMaxScaler(feature_range=(0, 1))
```

```
# scaler = QuantileTransformer()

scaler = StandardScaler()

dataset = scaler.fit_transform(dataset)

# Take first 80 % values for train and the remaining 20 % for testing

# That is, 73 days (upto and including February 11 2020) for training.

# and 18 days (from February 12 2020) for testing.

if (time == 15):

    # for 15 minutes

    train, test = dataset[0:4964], dataset[4964-68*15:]

elif (time == 30):

    # for 30 minutes

    train, test = dataset[0:2482], dataset[2482-34*15:]

elif (time == 45):

    # for 45 minutes

    train, test = dataset[0:1679], dataset[1679-23*15:]

else:

    # for 60 minutes

    train, test = dataset[0:1241], dataset[1241-17*15:]

# for testing dataset[2482-length] or dataset[2482-lookback]. Here 15 days look back is
used.

# Use TimeseriesGenerator to organize training data into the right format

from keras.preprocessing.sequence import TimeseriesGenerator

if (time == 15):
```

```
    # for 15 minutes
    seq_size = length = 68*15
elif (time == 30):
    # for 30 minutes
    seq_size = length = 34*15
elif (time == 45):
    # for 45 minutes
    seq_size = length = 23*15
else:
    # for 60 minutes
    seq_size = length = 17*15
batch_size = 16
train_generator = TimeseriesGenerator(train,train,length=length,batch_size=batch_size)
print("Total number of samples in the original training data = ", len(train))
print("Total number of samples in the generated data = ", len(train_generator))
# Also generate validation data
validation_generator = TimeseriesGenerator(test, test, length=length
,batch_size=batch_size)
#Input dimensions are... (N x seq_size)
num_features = 1 # For univariate
# Stacked LSTM with 1 hidden dense layer
model = Sequential()
model.add(LSTM(100, activation='tanh', return_sequences=True, input_shape=(length,
num_features)))
```

```
model.add(LSTM(100, activation='tanh'))

model.add(Dense(1))

model.compile(optimizer = keras.optimizers.Adam(0.0004), loss =
'mean_absolute_error')

model.summary()

print('Train...')

history = model.fit_generator(generator = train_generator, epochs = 50, validation_data =
validation_generator)

# Forecasting using the trained model

trainPredict = model.predict(train_generator)

testPredict = model.predict(validation_generator)

# Inverting the normalization to get back original values

trainPredict = scaler.inverse_transform(trainPredict)

trainY_inverse = scaler.inverse_transform(train)

testPredict = scaler.inverse_transform(testPredict)

testY_inverse = scaler.inverse_transform(test)

# calculate mean absolute error

trainScore = mean_absolute_error(trainY_inverse[length:], trainPredict[:,0])

print('Train Score: %.2f MAE' % (trainScore))

testScore = mean_absolute_error(testY_inverse[length:], testPredict[:,0])

print('Test Score: %.2f MAE' % (testScore))

# calculate mean absolute error

trainScore = mean_absolute_error(trainY_inverse[length:], trainPredict[:,0])

print('Train Score: %.2f MAE' % (trainScore))
```

```
testScore = mean_absolute_error(testY_inverse[length:], testPredict[:,0])

print('Test Score: %.2f MAE' % (testScore))

# calculate root mean square error

trainScore = math.sqrt(mean_squared_error(trainY_inverse[length:], trainPredict[:,0]))

print('Train Score: %.2f RMSE' % (trainScore))

testScore = math.sqrt(mean_squared_error(testY_inverse[length:], testPredict[:,0]))

print('Test Score: %.2f RMSE' % (testScore))

# Plotting actual v/s forecasted

test_for_plot['TP'].plot(figsize = (14,6))

predictions_1['TPF'].plot(figsize = (14,6))

plt.title('Actual v/s forecasted using LSTM for station 5, 30 minutes interval', font =
'times new roman', size = 14)

plt.ylabel('Passengers', font = 'times new roman', size = 14)

plt.xlabel('Time', font = 'times new roman', size = 14)

#plt.xlim('2020-02-28 06:00:00','2020-02-29 22:30:00')

plt.xticks(font='times new roman', size = 11)

plt.yticks(font='times new roman', size = 11)

plt.minorticks_on()

plt.legend(['Actual Passengers','Forecasted Passengers'], prop = 'times new roman')

plt.savefig("jpg")

plt.show()
```

A.13. Python Code used for LOS Analysis

```
import matplotlib.pyplot as plt
```

```
from matplotlib.image import imread
import pandas as pd
import numpy as np
import seaborn as sns
from sklearn.cluster import KMeans, SpectralClustering
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import silhouette_samples, silhouette_score
import sklearn.metrics as metrics
dataset= pd.read_csv('BTI.csv')
X=dataset.iloc[:,[0,1]].values
kmeans_model = KMeans(n_clusters=6, random_state=1).fit(X)
cluster_labels = kmeans_model.labels_
cluster_cedntroids_BTI= kmeans_model.cluster_centers_
sample_silhouette_values = metrics.silhouette_samples(X, cluster_labels)
means_lst_BTI = []
for label in range(6):
means_lst_BTI.append(sample_silhouette_values[cluster_labels == label].mean())
avg_score = np.mean(sample_silhouette_values)
```

A.14. R Studio Code used for Hartigan dip test

```
#install.packages("magicfor")
```

```
#install.packages("diptest")
library(diptest)
library(writexl)
library(readxl)
library(magicfor)
test1 <- read_excel("Hourwise Monday U.xlsx")
#View(dip_test)
attach(test1)
magic_for(print, silent = TRUE)
for (p in test1) {
  dip.test(p)
  x <-dip.test(p, simulate.p.value = FALSE, B = 2000)
  print(x)
}
#x <- TT4
#plot(density(x)); rug(x)
#dip.test(x)
#dip.test(x, simulate.p.value = FALSE, B = 2000)
magic_result()
capture.output(magic_result(), file = "UNI_mOD_TRAVEL TIME PEAK HOURS.csv")
```

A.15. Descriptive Statistical Analysis for Express Route - Monday

Hour of the Day	Avg. TT (minute)	SD of TT	CV of TT	95th Percentile TT (minute)	PTI	BTI	TTT
5 - 6	29.7	2.61	8.79	31.0	1.4	4.5	1.4
6 - 7	30.7	2.55	8.30	32.5	1.5	5.8	1.4
7 - 8	32.4	2.16	6.68	34.1	1.6	5.4	1.5
8 - 9	33.8	2.46	7.27	36.7	1.7	8.8	1.5
9 - 10	33.1	1.54	4.66	35.4	1.6	6.9	1.5
10 - 11	33.7	1.33	3.94	36.2	1.7	7.3	1.5
11 - 12	34.0	1.02	3.00	36.3	1.7	6.9	1.6
12 - 13	35.3	1.62	4.58	38.5	1.8	9.0	1.6
13 - 14	32.0	2.19	6.83	36.1	1.6	12.7	1.5
14 - 15	30.8	2.73	8.86	34.0	1.6	10.5	1.4
15 - 16	30.8	2.34	7.61	33.8	1.5	9.8	1.4
16 - 17	36.1	1.85	5.11	38.9	1.8	7.7	1.6
17 - 18	38.1	1.42	3.74	42.5	1.9	11.5	1.7
18 - 19	38.9	1.43	3.69	43.9	2.0	12.9	1.8
19 -20	38.5	1.17	3.03	43.4	2.0	12.8	1.8
20 - 21	34.0	1.35	3.97	37.7	1.7	11.0	1.6
21 - 22	31.6	2.19	6.94	34.8	1.6	10.2	1.4
22 - 23	30.9	2.22	7.18	34.0	1.6	9.9	1.4

A.16. Descriptive Statistical Analysis for Express Route - Tuesday

Hour of the Day	Avg. TT (minute)	SD of TT	CV of TT	95th Percentile TT (minute)	PTI	BTI	TTT
5 - 6	29.7	2.51	8.45	31.1	1.4	4.7	1.4
6 - 7	30.5	2.48	8.12	32.4	1.5	6.3	1.4
7 - 8	30.7	2.67	8.70	32.9	1.5	7.3	1.4
8 - 9	33.6	2.10	6.27	36.1	1.6	7.6	1.5
9 - 10	32.6	1.31	4.01	34.9	1.6	7.1	1.5
10 - 11	33.0	1.50	4.54	35.4	1.6	7.3	1.5
11 - 12	33.4	1.27	3.80	36.3	1.7	8.6	1.5
12 - 13	34.5	1.52	4.40	36.9	1.7	7.0	1.6
13 - 14	32.0	2.68	8.38	35.1	1.6	9.9	1.5
14 - 15	30.9	2.85	9.23	34.8	1.6	12.4	1.4
15 - 16	31.0	2.40	7.75	34.7	1.6	12.1	1.4
16 - 17	36.0	1.45	4.02	40.0	1.8	11.0	1.6
17 - 18	37.6	1.33	3.54	41.7	1.9	11.0	1.7
18 - 19	38.1	1.53	4.00	42.8	2.0	12.1	1.7
19 -20	38.8	1.46	3.76	44.5	2.0	14.6	1.8
20 - 21	35.1	2.68	7.64	38.0	1.7	8.1	1.6
21 - 22	30.9	2.77	8.96	33.9	1.5	9.8	1.4
22 - 23	29.9	2.89	9.67	31.0	1.4	3.9	1.4

A.17. Descriptive Statistical Analysis for Express Route – Wednesday

Hour of the Day	Avg. TT (minute)	SD of TT	CV of TT	95th Percentile TT (minute)	PTI	BTI	TTT
5 - 6	28.7	2.61	9.09	30.0	1.4	4.4	1.3
6 - 7	30.4	2.57	8.46	32.4	1.5	6.6	1.4
7 - 8	33.0	2.91	8.83	35.4	1.6	7.3	1.5
8 - 9	33.3	1.54	4.62	36.4	1.7	9.5	1.5
9 - 10	33.0	1.43	4.32	35.6	1.6	7.8	1.5
10 - 11	33.2	1.28	3.86	36.1	1.6	8.7	1.5
11 - 12	33.6	1.42	4.23	36.9	1.7	9.7	1.5
12 - 13	34.6	1.32	3.82	37.9	1.7	9.7	1.6
13 - 14	32.7	1.52	4.65	36.5	1.7	11.8	1.5
14 - 15	30.8	2.10	6.83	34.6	1.6	12.2	1.4
15 - 16	30.8	2.32	7.54	34.0	1.6	10.5	1.4
16 - 17	36.6	2.02	5.51	39.6	1.8	8.2	1.7
17 - 18	38.0	1.77	4.66	42.6	1.9	12.2	1.7
18 - 19	38.3	1.42	3.71	42.7	2.0	11.6	1.7
19 -20	38.8	1.73	4.46	44.1	2.0	13.7	1.8
20 - 21	32.0	2.22	6.94	34.8	1.6	8.7	1.5
21 - 22	33.2	3.02	9.09	35.5	1.6	6.9	1.5
22 - 23	32.0	3.10	9.71	33.9	1.5	5.9	1.5

A.18. Descriptive Statistical Analysis for Express Route – Thursday

Hour of the Day	Avg. TT (minute)	SD of TT	CV of TT	95th Percentile TT (minute)	PTI	BTI	TTT
5 - 6	29.3	2.40	8.19	30.7	1.4	4.7	1.3
6 - 7	30.5	2.56	8.39	32.3	1.5	5.8	1.4
7 - 8	33.6	2.42	7.19	35.7	1.6	6.4	1.5
8 - 9	33.3	1.96	5.89	36.0	1.6	8.0	1.5
9 - 10	32.9	1.01	3.08	36.1	1.6	9.8	1.5
10 - 11	33.5	1.13	3.38	36.4	1.7	8.7	1.5
11 - 12	33.7	1.22	3.62	37.4	1.7	11.0	1.5
12 - 13	35.1	1.31	3.73	39.2	1.8	11.7	1.6
13 - 14	33.4	1.54	4.61	36.8	1.7	10.0	1.5
14 - 15	31.3	2.02	6.44	35.1	1.6	12.2	1.4
15 - 16	30.9	2.40	7.77	34.7	1.6	12.5	1.4
16 - 17	35.6	1.32	3.69	38.8	1.8	8.8	1.6
17 - 18	38.3	1.51	3.94	43.2	2.0	12.6	1.8
18 - 19	38.8	1.23	3.17	44.0	2.0	13.4	1.8
19 -20	38.8	1.71	4.41	44.2	2.0	13.8	1.8
20 - 21	34.7	2.38	6.85	37.4	1.7	7.6	1.6
21 - 22	32.3	2.55	7.89	34.3	1.6	6.1	1.5
22 - 23	30.6	3.02	9.87	32.2	1.5	5.3	1.4

A.19. Descriptive Statistical Analysis for Express Route – Friday

Hour of the Day	Avg. TT (minute)	SD of TT	CV of TT	95th Percentile TT (minute)	PTI	BTI	TTT
5 - 6	31.0	2.78	8.99	32.0	1.5	3.3	1.4
6 - 7	30.1	2.67	8.85	31.4	1.4	4.1	1.4
7 - 8	32.8	2.62	7.98	35.2	1.6	7.2	1.5
8 - 9	32.9	1.85	5.62	36.3	1.7	10.4	1.5
9 - 10	32.6	1.42	4.34	35.7	1.6	9.4	1.5
10 - 11	32.9	1.11	3.38	36.5	1.7	10.8	1.5
11 - 12	33.7	1.02	3.02	37.5	1.7	11.4	1.5
12 - 13	34.6	1.31	3.80	37.4	1.7	8.3	1.6
13 - 14	32.0	1.57	4.91	34.8	1.6	8.6	1.5
14 - 15	30.5	2.44	7.98	34.4	1.6	12.8	1.4
15 - 16	37.0	2.22	6.00	40.7	1.9	9.7	1.7
16 - 17	31.0	1.46	4.72	35.2	1.6	13.5	1.4
17 - 18	37.8	1.02	2.69	42.4	1.9	12.1	1.7
18 - 19	38.5	1.27	3.30	43.0	2.0	11.9	1.8
19 -20	38.0	1.32	3.48	42.1	1.9	10.9	1.7
20 - 21	35.0	2.22	6.34	38.9	1.8	11.1	1.6
21 - 22	32.7	2.43	7.44	35.2	1.6	7.7	1.5
22 - 23	32.2	2.99	9.30	34.2	1.6	6.2	1.5

A.20. Descriptive Statistical Analysis for Express Route – Saturday

Hour of the Day	Avg. TT (minute)	SD of TT	CV of TT	95th Percentile TT (minute)	PTI	BTI	TTT
5 - 6	29.0	2.68	9.24	30.1	1.4	3.9	1.3
6 - 7	31.9	2.80	8.77	33.4	1.5	4.6	1.5
7 - 8	34.1	2.90	8.48	36.0	1.6	5.4	1.6
8 - 9	34.1	1.64	4.81	36.6	1.7	7.4	1.6
9 - 10	33.6	1.56	4.65	36.7	1.7	9.1	1.5
10 - 11	33.8	1.28	3.79	36.5	1.7	8.0	1.5
11 - 12	33.8	1.04	3.06	37.3	1.7	10.2	1.5
12 - 13	35.2	1.46	4.15	39.4	1.8	12.0	1.6
13 - 14	33.4	2.10	6.30	36.5	1.7	9.3	1.5
14 - 15	31.5	2.66	8.46	35.4	1.6	12.4	1.4
15 - 16	31.3	2.38	7.61	34.9	1.6	11.6	1.4
16 - 17	35.8	1.28	3.56	39.0	1.8	8.7	1.6
17 - 18	38.3	1.20	3.14	41.9	1.9	9.6	1.7
18 - 19	39.0	1.24	3.18	43.9	2.0	12.6	1.8
19 -20	39.5	1.90	4.82	44.6	2.0	13.0	1.8
20 - 21	35.7	2.03	5.67	39.8	1.8	11.2	1.6
21 - 22	33.0	2.35	7.13	35.9	1.6	8.8	1.5
22 - 23	32.9	2.94	8.93	34.6	1.6	5.3	1.5

A.21. Descriptive Statistical Analysis for Express Route – Sunday

Hour of the Day	Avg. TT (minute)	SD of TT	CV of TT	95th Percentile TT (minute)	PTI	BTI	TTT
5 - 6	30.0	2.62	8.72	31.1	1.4	3.4	1.4
6 - 7	32.5	2.57	7.92	33.9	1.5	4.2	1.5
7 - 8	32.2	2.77	8.61	34.7	1.6	7.7	1.5
8 - 9	32.4	1.78	5.48	35.2	1.6	8.5	1.5
9 - 10	32.9	1.94	5.89	34.9	1.6	6.0	1.5
10 - 11	31.7	1.20	3.79	34.5	1.6	8.9	1.4
11 - 12	34.0	0.98	2.89	37.2	1.7	9.3	1.6
12 - 13	34.7	1.03	2.97	38.0	1.7	9.5	1.6
13 - 14	31.5	1.13	3.58	34.1	1.6	8.2	1.4
14 - 15	30.5	2.55	8.35	33.1	1.5	8.5	1.4
15 - 16	29.5	2.04	6.91	32.2	1.5	9.2	1.3
16 - 17	38.4	1.31	3.41	42.9	2.0	11.9	1.8
17 - 18	37.3	1.28	3.41	42.4	1.9	13.5	1.7
18 - 19	37.2	1.75	4.69	42.1	1.9	13.1	1.7
19 -20	38.4	2.07	5.40	44.0	2.0	14.4	1.8
20 - 21	35.2	2.11	6.00	38.6	1.8	9.4	1.6
21 - 22	36.7	2.91	7.92	39.0	1.8	6.2	1.7
22 - 23	35.0	2.98	8.52	37.2	1.7	6.2	1.6

A.22. Descriptive Statistical Analysis for Non-express Route – Monday

Hour of the Day	Avg. TT (minute)	SD of TT	CV of TT	95th Percentile TT (minute)	PTI	BTI	TTT
5 - 6	41.6	5.17	12.41	44.2	1.4	6.1	1.3
6 - 7	49.5	4.97	10.04	52.8	1.7	6.7	1.6
7 - 8	51.2	4.38	8.55	55.3	1.7	8.0	1.6
8 - 9	49.2	2.90	5.89	55.1	1.7	12.0	1.6
9 - 10	46.9	2.39	5.10	50.4	1.6	7.5	1.5
10 - 11	49.7	3.02	6.09	55.2	1.7	11.1	1.6
11 - 12	51.1	3.79	7.42	56.6	1.8	10.7	1.6
12 - 13	52.3	2.89	5.53	56.6	1.8	8.3	1.7
13 - 14	46.2	2.78	6.02	54.8	1.7	18.7	1.5
14 - 15	44.6	3.30	7.40	49.4	1.6	10.9	1.4
15 - 16	45.7	5.42	11.87	53.6	1.7	17.4	1.4
16 - 17	47.6	2.22	4.66	52.7	1.7	10.7	1.5
17 - 18	61.0	3.19	5.24	73.4	2.3	20.3	1.9
18 - 19	58.0	2.97	5.12	70.9	2.2	22.2	1.8
19 -20	54.8	3.51	6.41	66.4	2.1	21.2	1.7
20 - 21	48.5	5.87	12.09	55.8	1.8	15.1	1.5
21 - 22	45.8	5.32	11.60	50.5	1.6	10.0	1.5
22 - 23	40.6	5.12	10.72	46.7	1.5	9.6	1.2

A.23. Descriptive Statistical Analysis for Non-express Route – Tuesday

Hour of the Day	Avg. TT (minute)	SD of TT	CV of TT	95th Percentile TT (minute)	PTI	BTI	TTT
5 - 6	42.8	5.11	11.94	46.2	1.5	7.9	1.4
6 - 7	48.7	4.89	10.05	52.7	1.7	8.2	1.5
7 - 8	49.1	4.01	8.16	55.8	1.8	13.6	1.6
8 - 9	48.6	3.11	6.40	54.1	1.7	11.4	1.5
9 - 10	46.9	2.98	6.35	55.2	1.7	17.8	1.5
10 - 11	48.7	2.86	5.88	52.2	1.7	7.3	1.5
11 - 12	51.3	2.52	4.91	55.3	1.7	7.9	1.6
12 - 13	51.3	3.11	6.05	56.3	1.8	9.6	1.6
13 - 14	45.8	3.42	7.47	54.5	1.7	19.1	1.4
14 - 15	45.5	3.72	8.18	51.9	1.6	14.1	1.4
15 - 16	44.7	4.14	9.27	51.5	1.6	15.3	1.4
16 - 17	47.9	2.89	6.03	54.6	1.7	13.8	1.5
17 - 18	60.2	2.46	4.09	71.2	2.3	18.3	1.9
18 - 19	60.3	2.27	3.77	72.8	2.3	20.8	1.9
19 -20	57.5	2.26	3.92	66.2	2.1	15.1	1.8
20 - 21	47.5	3.10	6.54	52.4	1.7	10.5	1.5
21 - 22	42.3	5.6	13.2	46.5	1.5	10.0	1.3
22 - 23	40.6	5.77	14.23	46.5	1.5	14.5	1.3

A.24. Descriptive Statistical Analysis for Non-express Route – Wednesday

Hour of the Day	Avg. TT (minute)	SD of TT	CV of TT	95th Percentile TT (minute)	PTI	BTI	TTT
5 - 6	41.9	5.19	12.39	45.1	1.4	7.6	1.3
6 - 7	45.3	5.31	11.73	51.4	1.6	13.4	1.4
7 - 8	46.3	4.78	10.34	52.9	1.7	14.3	1.5
8 - 9	48.3	3.25	6.74	53.2	1.7	10.2	1.5
9 - 10	46.7	3.30	7.08	51.5	1.6	10.3	1.5
10 - 11	47.7	3.07	6.43	51.8	1.6	8.7	1.5
11 - 12	49.0	2.89	5.90	55.1	1.7	12.4	1.6
12 - 13	50.2	2.68	5.33	55.6	1.8	10.8	1.6
13 - 14	46.2	3.72	8.04	54.6	1.7	18.2	1.5
14 - 15	44.3	4.37	9.87	51.2	1.6	15.5	1.4
15 - 16	45.1	5.49	12.18	53.8	1.7	19.5	1.4
16 - 17	45.2	2.56	5.65	51.8	1.6	14.7	1.4
17 - 18	58.5	2.46	4.21	72.2	2.3	23.4	1.8
18 - 19	57.5	3.13	5.44	68.3	2.2	18.6	1.8
19 -20	55.1	3.76	6.83	66.5	2.1	20.7	1.7
20 - 21	48.0	3.53	7.35	52.6	1.7	9.6	1.5
21 - 22	43.9	5.22	11.88	48.1	1.5	9.6	1.4
22 - 23	41.9	4.8	9.1	45.9	1.5	9.5	1.3

A.25. Descriptive Statistical Analysis for Non-express Route – Thursday

Hour of the Day	Avg. TT (minute)	SD of TT	CV of TT	95th Percentile TT (minute)	PTI	BTI	TTT
5 - 6	40.2	5.37	13.37	44.2	1.4	10.0	1.3
6 - 7	47.8	4.10	8.58	53.2	1.7	11.2	1.5
7 - 8	47.1	3.98	8.44	53.9	1.7	14.3	1.5
8 - 9	49.6	3.17	6.39	56.3	1.8	13.5	1.6
9 - 10	47.5	2.56	5.39	50.9	1.6	7.1	1.5
10 - 11	48.7	2.79	5.73	55.8	1.8	14.8	1.5
11 - 12	52.2	2.20	4.21	58.6	1.9	12.3	1.7
12 - 13	52.3	2.89	5.53	57.3	1.8	9.6	1.7
13 - 14	47.0	3.38	7.20	56.0	1.8	19.1	1.5
14 - 15	45.2	3.28	7.25	52.3	1.7	15.6	1.4
15 - 16	44.6	5.14	11.52	53.3	1.7	19.4	1.4
16 - 17	46.3	2.15	4.65	53.1	1.7	14.8	1.5
17 - 18	61.9	2.55	4.12	76.0	2.4	22.8	2.0
18 - 19	61.0	3.67	6.03	75.3	2.4	23.4	1.9
19 -20	56.1	3.27	5.83	66.6	2.1	18.8	1.8
20 - 21	50.1	5.91	11.79	58.1	1.8	16.1	1.6
21 - 22	46.2	7.60	16.46	52.0	1.6	12.7	1.5
22 - 23	42.7	5.78	13.53	46.1	1.5	7.9	1.4

A.26. Descriptive Statistical Analysis for Non-express Route – Friday

Hour of the Day	Avg. TT (minute)	SD of TT	CV of TT	95th Percentile TT (minute)	PTI	BTI	TTT
5 - 6	40.8	4.95	12.11	44.5	1.4	9.0	1.3
6 - 7	47.6	4.45	9.33	52.1	1.6	9.3	1.5
7 - 8	49.7	4.62	9.28	55.4	1.8	11.3	1.6
8 - 9	47.7	3.04	6.36	52.0	1.6	9.0	1.5
9 - 10	46.2	2.47	5.35	51.4	1.6	11.2	1.5
10 - 11	48.6	2.89	5.95	54.9	1.7	12.9	1.5
11 - 12	50.1	3.10	6.19	54.7	1.7	9.1	1.6
12 - 13	52.2	4.09	7.85	59.9	1.9	14.8	1.7
13 - 14	45.3	3.97	8.77	53.2	1.7	17.5	1.4
14 - 15	44.6	3.93	8.81	50.9	1.6	14.2	1.4
15 - 16	44.9	4.40	9.80	52.2	1.7	16.3	1.4
16 - 17	46.4	2.14	4.61	52.2	1.7	12.4	1.5
17 - 18	58.3	2.84	4.88	73.2	2.3	25.4	1.8
18 - 19	58.5	2.68	4.59	71.0	2.2	21.4	1.9
19 -20	56.3	3.00	5.32	69.3	2.2	23.0	1.8
20 - 21	50.6	3.85	7.60	67.8	2.1	33.9	1.6
21 - 22	43.8	5.37	12.26	50.2	1.6	14.6	1.4
22 - 23	40.4	5.12	11.02	45.4	1.4	12.4	1.3

A.27. Descriptive Statistical Analysis for Non-express Route – Saturday

Hour of the Day	Avg. TT (minute)	SD of TT	CV of TT	95th Percentile TT (minute)	PTI	BTI	TTT
5 - 6	40.2	5.27	13.13	44.1	1.4	9.8	1.3
6 - 7	45.6	4.49	9.85	51.2	1.6	12.3	1.4
7 - 8	51.6	4.21	8.15	57.9	1.8	12.1	1.6
8 - 9	51.5	3.88	7.54	58.0	1.8	12.7	1.6
9 - 10	48.2	3.03	6.29	52.9	1.7	9.7	1.5
10 - 11	48.3	2.87	5.95	53.4	1.7	10.7	1.5
11 - 12	49.4	2.72	5.49	54.2	1.7	9.8	1.6
12 - 13	52.5	2.99	5.70	57.5	1.8	9.5	1.7
13 - 14	48.3	3.11	6.44	55.9	1.8	15.8	1.5
14 - 15	45.4	3.47	7.64	50.3	1.6	10.8	1.4
15 - 16	45.3	4.60	10.15	52.4	1.7	15.7	1.4
16 - 17	51.3	2.81	15.23	63.6	2.0	24.0	1.6
17 - 18	59.9	2.72	4.54	70.5	2.2	17.6	1.9
18 - 19	59.8	2.97	4.96	72.8	2.3	21.8	1.9
19 -20	58.5	3.46	5.91	72.0	2.3	22.9	1.9
20 - 21	44.8	4.32	9.64	50.4	1.6	12.4	1.4
21 - 22	43.8	5.92	13.51	50.0	1.6	11.4	1.4
22 - 23	40.8	5.0	12.3	47.9	1.2	8.7	1.3

A.28. Descriptive Statistical Analysis for Non-express Route – Sunday

Hour of the Day	Avg. TT (minute)	SD of TT	CV of TT	95th Percentile TT (minute)	PTI	BTI	TTT
5 - 6	41.4	5.53	13.37	45.2	1.4	9.3	1.3
6 - 7	47.1	5.28	11.21	52.9	1.7	12.3	1.5
7 - 8	46.2	5.18	11.21	52.2	1.7	13.0	1.5
8 - 9	46.6	4.89	10.51	53.9	1.7	15.8	1.5
9 - 10	45.0	3.10	6.90	48.5	1.5	7.9	1.4
10 - 11	45.9	3.11	6.77	50.6	1.6	10.2	1.5
11 - 12	45.5	2.79	6.12	50.8	1.6	11.6	1.4
12 - 13	47.6	3.27	6.87	52.5	1.7	10.4	1.5
13 - 14	45.8	3.27	7.13	53.6	1.7	17.0	1.4
14 - 15	42.9	5.10	11.89	51.3	1.6	19.5	1.4
15 - 16	41.2	2.90	7.04	46.0	1.5	11.6	1.3
16 - 17	43.0	2.68	6.23	48.5	1.5	12.9	1.4
17 - 18	55.2	2.54	4.60	68.6	2.2	24.2	1.7
18 - 19	56.3	2.74	4.88	69.4	2.2	23.3	1.8
19 -20	54.7	2.97	5.43	67.8	2.1	23.9	1.7
20 - 21	50.8	2.89	5.69	64.4	2.0	26.8	1.6
21 - 22	46.4	4.54	9.78	53.9	1.7	16.1	1.5
22 - 23	47.3	5.07	10.70	54.8	1.7	15.9	1.5

REFERENCES

1. Abbasimehr, H., and Khodizadeh Nahari, M. (2020). "Improving demand forecasting with LSTM by taking into account the seasonality of data." *Journal of applied research on industrial engineering*, *7*(2), 177-189.
2. Anvari, S., Tuna, S., Canci, M., and Turkay, M. (2016). "Automated Box–Jenkins forecasting tool with an application for passenger demand in urban rail systems." *Journal of Advanced Transportation*, *50*(1), 25-49.
3. Astuti, S. W. (2018, August). "Forecasting Surabaya–Jakarta Train Passengers with SARIMA model." In *IOP Conference Series: Materials Science and Engineering* (Vol. 407, No. 1, p. 012105). IOP Publishing.
4. Benevolo, C., Dameri, R. P., and D'auria, B. (2016). "Smart mobility in smart city." In *Empowering organizations*. Springer, Cham.13-28
5. Bharti, A. K., Chandra, S., and Chalumuri, R. S. (2013). "Performance evaluation of urban arterial in Delhi using travel time reliability." In *Proceedings of the Eastern Asia Society for transportation studies* (Vol. 9).
6. Bharti, A. K., Sekhar, C. R., and Chandra, S. (2018). "Travel time reliability as a level of service measure for urban and inter-urban corridors in India." *Current Science*, 1913-1922.
7. Biswas, S., Singh, B., and Saha, A. (2016). "Assessment of level-of-service on urban arterials: A case study in Kolkata metropolis." *Int. J. Traffic Transp. Eng*, *6*, 303-312.
8. Camus, R., Longo, G., and Macorini, C. (2005). "Estimation of transit reliability level-of-service based on automatic vehicle location data." *Transportation research record*, *1927*(1), 277-286.
9. Carmona-Benítez, R. B., and Nieto, M. R. (2020). "SARIMA damp trend grey forecasting model for airline industry." *Journal of air transport management*, *82*, 101736.

10. Chepuri, A., Joshi, S., Arkatkar, S., Joshi, G., and Bhaskar, A. (2020). “Development of new reliability measure for bus routes using trajectory data.” *Transportation Letters*, *12*(6), 363-374.
11. Chepuri, A., Kulkarni, R., Bains, M., Arkatkar, S., and Joshi, G. (2015). “Evaluation of BRTS corridor in India using microscopic simulation: a case study in Surat City.” In *Proceedings of the Eastern Asia society for transportation studies* (Vol. 10).
12. Chepuri, A., Kumar, C., Bhanegaonkar, P., Arkatkar, S. S., and Joshi, G. (2019). “Travel time reliability analysis on selected bus route of Mysore using GPS data.” *Transportation in Developing Economies*, *5*(2), 1-15.
13. Chepuri, A., Ramakrishnan, J., Arkatkar, S., Joshi, G., and Pulugurtha, S. S. (2018). Examining travel time reliability-based performance indicators for bus routes using GPS-based bus trajectory data in India. *Journal of Transportation Engineering, Part A: Systems*, *144*(5), 04018012.
14. Chepuri, A., Ramakrishnan, J., Arkatkar, S., Joshi, G., and Pulugurtha, S. S. (2018). “Examining travel time reliability-based performance indicators for bus routes using GPS-based bus trajectory data in India.” *Journal of Transportation Engineering, Part A: Systems*, *144*(5), 04018012.
15. Chikkakrishna, N. K., Hardik, C., Deepika, K., and Sparsha, N. (2019, December). “Short-term traffic prediction using sarima and FbPROPHET.” In *2019 IEEE 16th India Council International Conference (INDICON)* (pp. 1-4). IEEE.
16. CSIR – CRRI (2012). “Evaluating Bus Rapid Transit (BRT) Corridor Performance from Ambedkar Nagar to Mool Chand Intersection.” *-CSIR-CRRI.* https://crridom.gov.in/evaluating-bus-rapid-transit-brt-corridor-performance amebedkar-nagar-mool-chand.
17. Cyril, A., Mulangi, R. H., and George, V. (2018, August). “Modelling and forecasting bus passenger demand using time series method”. *7th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)* (pp. 460-466). IEEE.

18. Cyril, A., Mulangi, R. H., and George, V. (2019). “Bus passenger demand modelling using time-series techniquesbig data analytics.” *The Open Transportation Journal*, *13*(1).
19. Das, Shreya., and Debapratim, Pandit. (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.
20. Deng, T., and Nelson, J. D. (2013). “Bus Rapid Transit implementation in Beijing: An evaluation of performance and impacts.” *Research in Transportation Economics, 39*(1), 108-113.
21. Dickey, D. A., and Fuller, W. A. (1979). “Distribution of the estimators for autoregressive time series with a unit root.” *Journal of the American statistical association*, *74*(366a), 427-431.
22. Ding, J., Tarokh, V., and Yang, Y.: “Model selection techniques: An overview.” *IEEE Signal Processing Magazine, 35*(6), 16-34 (2018)
23. Directorate of Urban Land Transport, (2013). “Hubli-Dharwad Bus Rapid Transit System, Detailed Feasibility Report.” *CEPT University. Ahmedabad.*
24. El-Geneidy, A. M., Horning, J., and Krizek, K. J. (2011). “Analysing transit service reliability using detailed data from automatic vehicular locator systems.” *Journal of Advanced Transportation*, *45*(1), 66-79.
25. Federal Highway Administration. (2006). “Travel time reliability: Making it there on time, all the time.” U.S. Dept. of Transportation, Washington, DC.
26. Fuloria, S. (2020). “Passenger demand forecasting in the ridesharing context: a comparison of statistical and deep learning approaches.” *IUP Journal of Applied Economics, 19*(1), 57-71.
27. Gallo, M., De Luca, G., D’Acierno, L., and Botte, M. (2019). “Artificial neural networks for forecasting passenger flows on metro lines.” *Sensors, 19*(15), 3424.
28. GANDHI, S., TIWARI, G., and FAZIO, J. (2013). “Comparative evaluation of alternate bus rapid transit system (BRTS) planning, operation and design

options." *Journal of the Eastern Asia Society for Transportation Studies*, *10*, 1292-1310.

29. Godavarthi, G. R., Chalumuri, R. S., and Velmurugun, S. (2014). "Measuring the performance of bus rapid-transit corridors based on volume by capacity ratio." *Journal of Transportation Engineering*, *140*(10), 04014049.
30. Godavarthi, G. R., Chalumuri, R. S., and Velmurugun, S. (2014). "Measuring the performance of bus rapid-transit corridors based on volume by capacity ratio." *Journal of Transportation Engineering*, *140*(10), 04014049.
31. Gummadi, R., and Edara, S. R. (2018). "Analysis of passenger flow prediction of transit buses along a route based on time series." *Information and Decision Sciences* (pp. 31-37). Springer, Singapore.
32. Gunawan, F. E. (2015). "Empirical assessment on factors affecting travel time of bus rapid transit." *International Journal of Engineering and Technology*, *7*(1), 327-334.
33. Halyal, S., Mulangi, R. H., and Harsha, M. M. (2022). "Forecasting public transit passenger demand: With neural networks using APC data." *Case Studies on Transport Policy*.
34. Harsha, M. M., and Mulangi, R. H. (2021). "Probability distributions analysis of travel time variability for the public transit system." *International Journal of Transportation Science and Technology*.
35. Harsha, M. M., Mulangi, R. H., and Kumar, H. D. (2020). "Analysis of bus travel time variability using automatic vehicle location data." *Transportation Research Procedia*, *48*, 3283-3298.
36. Hartigan, J. A., and Hartigan, P. M. (1985). "The dip test of unimodality." *The annals of Statistics*, 70-84.
37. Hidalgo, D., Lleras, G., and Hernández, E. (2013). "Methodology for calculating passenger capacity in bus rapid transit systems: Application to the TransMilenio system in Bogotá, Colombia." *Research in Transportation Economics*, *39*(1), 139-142.

38. Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735-1780.

39. Hochreiter, S., (1998). “The vanishing gradient problem during learning recurrent neural nets and problem solutions.” *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems. 6*(02). 107-116 (1998).

40. Huo, Y., Zhao, J., Li, W., and Hu, X. (2014). Measuring bus service reliability: An example of bus rapid transit in changzhou. *Journal of Public Transportation*, *17*(2), 6.

41. Hyndman, R. J., and Athanasopoulos, G. (2018). “*Forecasting: principles and practice*.” OTexts.

42. Institute For Transportation and Development Policy, (2014). “The BRT Standard.” *ITDP*.

43. Institute For Transportation and Development Policy. (2017) “BRTS Planning Guide.” *ITDP*.

44. IRC: 124 (2017) “Bus Rapid Transit (BRT) Design Guideline for Indian Cities.” Indian Road Congress.

45. Jairam R, B., Anil, Kumar., Shriniwas S, Arkatkar. and Lelitha, Vanajakshi. (2018). “Performance Comparison of Bus Travel Time Prediction Models across Indian Cities” *TRB sagepub.com /journals Permissions.nav*.

46. Kanavos, A., Kounelis, F., Iliadis, L., and Makris, C. (2021). “Deep learning models for forecasting aviation demand time series.” *Neural Computing and Applications*, *33*(23), 16329-16343.

47. Kathuria, A., Parida, M., and Chalumuri, R. S. (2020). “Travel-time variability analysis of bus rapid transit system using GPS data.” *Journal of Transportation Engineering, Part A: Systems*, *146*(6), 05020003.

48. Kathuria, A., Parida, M., and Sekhar, C. (2020). “A review of service reliability measures for public transportation systems.” *International Journal of Intelligent Transportation Systems Research*, *18*(2), 243-255.

49. Kathuria, A., Parida, M., and Sekhar, C. R. (2017). "Route performance evaluation of a closed bus rapid transit system using GPS data." *Current Science*, 1642-1652.

50. Kathuria, A., Parida, M., Ravi Sekhar, C., and Sharma, A. (2016). "A review of bus rapid transit implementation in India." *Cogent Engineering*, *3*(1), 1241168.

51. Kathuria, A., Sekhar, C. R., and Parida, M. (2020). "Examining failure rate dynamics to estimate capacity of bus rapid transit system corridor." *Proceedings of the Institution of Civil Engineers-Transport* (Vol. 173, No. 5, pp. 333-344). Thomas Telford Ltd.

52. Kieu, L. M., Bhaskar, A., and Chung, E. (2015). "Public transport travel time variability definitions and monitoring." *Journal of Transportation Engineering*, *141*(1), Article-number.

53. Kittelson and Associates, Inc., KFH Group, Inc., Parsons Brinckerhoff Quade and Douglass, Inc., and K. Hunter-Zaworski. (2013). TCRP Report 100: "Transit Capacity and Quality of Service Manual (TCQSM)".3rd ed. *Transportation Research Board of the National Academies*. Washington, D.C.

54. Kuha, J., (2004). "AIC and BIC: Comparisons of assumptions and performance." *Sociological methods and research*. 33(2). 188-229.

55. Kumar, S. V., and Vanajakshi, L. (2015). "Short-term traffic flow prediction using seasonal ARIMA model with limited input data." *European Transport Research Review*, *7*(3), 1-9.

56. Kwon, J., Barkley, T., Hranac, R., Petty, K., and Compin, N. (2011). "Decomposition of travel time reliability into various sources: incidents, weather, work zones, special events, and base capacity". *Transportation research record*, *2229*(1), 28-33.

57. Lai, Y., Zhang, K., Lin, J., Yang, F., and Fan, Y. (2019, December). "Taxi demand prediction with LSTM-based combination model." In *2019 IEEE Intl Conf on Parallel and Distributed Processing with Applications, Big Data and Cloud Computing, Sustainable Computing and Communications, Social Computing and Networking (ISPA/BDCloud/SocialCom/SustainCom)* (pp. 944-950). IEEE.

58. Lee, J., Han, H., and Yoon, S. (2020). “Air passenger demand forecasting for the Incheon airport using time series models.” *Journal of Digital Convergence*, *18*(12), 87-95.
59. Li, J., Huang, P., Yang, Y., and Peng, Q. (2019, September). “Passenger flow prediction of high-speed railway based on lstm deep neural network.” In *RailNorrköping 2019. 8th International Conference on Railway Operations Modelling and Analysis (ICROMA), Norrköping, Sweden, June 17th–20th, 2019* (No. 069, pp. 723-739). Linköping University Electronic Press.
60. Li, L., Wang, Y., Zhong, G., Zhang, J., and Ran, B. (2018). “Short-to-medium term passenger flow forecasting for metro stations using a hybrid model.” *KSCE Journal of Civil Engineering*, *22*(5), 1937-1945.
61. Li, R., Rose, G., and Sarvi, M. (2006). “Using automatic vehicle identification data to gain insight into travel time variability and its causes.” *Transportation Research Record*, *1945*(1), 24-32.
62. Li, W., Sui, L., Zhou, M., and Dong, H. (2021). “Short-term passenger flow forecast for urban rail transit based on multi-source data.” *EURASIP Journal on Wireless Communications and Networking*, *2021*(1), 1-13.
63. Li, X., Zhang, Y., Du, M., and Yang, J. (2020). “The forecasting of passenger demand under hybrid ridesharing service modes: A combined model based on WT-FCBF-LSTM.” *Sustainable Cities and Society*, *62*, 102419.
64. Liu., Ronghui. and Sinha, Shalini. (2007). “Modelling Urban Bus Service and Passenger Reliability.” *In: The Third International Symposium on Transportation Network Reliability.19-20 July 2007*.The Hague. Netherlands.
65. Lomax, T., and Margiotta, R. (2003). “Selecting travel reliability measures.” *Texas Transportation Institute*. College Station.
66. Ma, Z. L., Ferreira, L., Mesbah, M., and Hojati, A. T. (2015). “Modelling bus travel time reliability with supply and demand data from automatic vehicle location and smart card systems.” *Transportation Research Record*, *2533*(1), 17-27.

67. Ma, Z., Ferreira, L., Mesbah, M., and Zhu, S. (2016). "Modelling distributions of travel time variability for bus operations." *Journal of Advanced Transportation*, *50*(1), 6-24.
68. Ma, Z., Zhu, S., Koutsopoulos, H. N., and Ferreira, L. (2017). "Quantile regression analysis of transit travel time reliability with automatic vehicle location and farecard data." *Transportation Research Record*, *2652*(1), 19-29.
69. Mazloumi, E., Currie, G., and Rose, G. (2010). "Using GPS data to gain insight into public transport travel time variability." *Journal of transportation engineering*, *136*(7), 623-631.
70. Medar, R., Rajpurohit, V. S., and Rashmi, B. (2017, August). "Impact of training and testing data splits on accuracy of time series forecasting in machine learning." In *2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA)* (pp. 1-6). IEEE.
71. Ministry of Road Transport and Highways, (2017). "Annual Report 2017-18." MoRTH, New Delhi.
72. Ministry of Road Transport and Highways, (2022). " Accident Annual Report 2021-22." MoRTH, New Delhi.
73. Ministry of Road Transport and Highways, (2019). "Road Transport Yearly Book 2018-19." MoRTH, New Delhi.
74. Ministry of Road Transport and Highways, (2022). "Annual Report 2021-22." MoRTH, New Delhi.
75. Mohamed, A. H., Adwan, I. A., Ahmeda, A. G., Hrtemih, H., and Al-MSari, H. (2021). "Identification of affecting factors on the travel time reliability for bus transportation." *Knowledge-Based Engineering and Sciences*, *2*(1), 19-30.
76. Nar, M., and Arslankaya, S. (2022). "Passenger demand forecasting for railway systems." *Open Chemistry*, *20*(1), 105-119.

77. Nguyen, N. T., Miwa, T., and Morikawa, T. (2020). "Demand forecast of public transportation considering positive and negative mass effects." *Transportation Research Part D: Transport and Environment, 85*, 102466.
78. Noh, Y., and Do, M. (2015). "Short-term Railway Passenger Demand Forecasting by SARIMA Model." *The Journal of The Korea Institute of Intelligent Transport Systems*, *14*(4), 18-26.
79. Olah, C., (2015). "Understanding LSTM Networks. colah's blog." https://colah.github.io/posts/2015-08-Understanding-LSTMs. Accessed July 2021.
80. Pavlyuk, D. (2017). "Short-term traffic forecasting using multivariate autoregressive models." *Procedia Engineering*, *178*, 57-66.
81. Phi, M., (2018). "Illustrated Guide to LSTM's and GRU's: A step by step explanation. Towards Data Science." https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21. Accessed July 2021.
82. Phillips, P. C., and Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, *75*(2), 335-346.
83. Prasad, D. R. (2017). "Urbanisation: Issues and Challenges." Avni Publications, New Delhi., Chapter 1, 1-34.
84. Pucher, J., Korattyswaroopam, N., and Ittyerah, N. (2004). "The crisis of public transport in India: overwhelming needs but limited resources." *Journal of public transportation*, *7*(4), 1.
85. R. Tolley., (2003). "Sustainable Transport." *Elsevier.* 29-Aug-2003
86. Rabbani, M. B. A., Musarat, M. A., Alaloul, W. S., Rabbani, M. S., Maqsoom, A., Ayub, S., ... and Altaf, M. (2021). "A comparison between seasonal autoregressive integrated moving average (SARIMA) and exponential smoothing (ES) based on time series model for forecasting road accidents." *Arabian Journal for Science and Engineering*, *46*(11), 11113-11138.

87. Ricard, L., Desaulniers, G., Lodi, A., and Rousseau, L. M. (2022). “Predicting the probability distribution of bus travel time to measure the reliability of public transport services.” *Transportation Research Part C: Emerging Technologies*, *138*, 103619.
88. Saedi, R., Saeedmanesh, M., Zockaie, A., Saberi, M., Geroliminis, N., and Mahmassani, H. S. (2020). “Estimating network travel time reliability with network partitioning.” *Transportation Research Part C: Emerging Technologies*, *112*, 46-61.
89. Satheesh Kumar, M., and Mali, R. Y. (2014). “Applications of Intelligent Transportation Systems using RFID Systems.” *International Journal of Engineering Research and General Science*,4 (1).
90. Schramm, L., Watkins, K., and Rutherford, S. (2010). “Features that affect variability of travel time on bus rapid transit systems.” *Transportation research record*, *2143*(1), 77-84.
91. Sekhar, C. R., and Yasuo, A. (2007). “Measuring travel time reliability of transportation system.” *Indian Highways*, *35*(11), 61.
92. Sharma, A., Parida, M., SEKHAR, C. R., and Kathuria, A. (2015). “Capacity analysis of Bhopal BRTS using empirical and simulation model.” *Journal of the Eastern Asia Society for Transportation Studies*, *11*, 1575-1593.
93. Singh, H., Kathuria, A., Sekhar, C. R., and Parida, M. (2017). “Estimating capacity of hybrid bus rapid transit corridor.” *Current Science*, 1586-1592.
94. Sources of Greenhouse Gas Emissions. US EPA. https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions. 2020
95. Susilawati, S., Taylor, M. A., and Somenahalli, S. V. (2010). “Travel time reliability measurement for selected corridors in the Adelaide Metropolitan area.” *Journal of the Eastern Asia Society for Transportation Studies*, *8*, 86-102.
96. Taylor, R. (1990). “Interpretation of the correlation coefficient: a basic review.” *Journal of diagnostic medical sonography*, *6*(1), 35-39.
97. TCRP REPORT 26., (1997). “Operational Analysis of Bus Lanes on Arterials.” *TRB.*
98. TCRP REPORT 88., (2003). “A Guidebook for Developing a Transit Performance Measurement System.” *TRB.*

99. Tu, H., van Lint, H., and van Zuylen, H. (2008, October). “The effects of traffic accidents on travel time reliability.” In *2008 11th International IEEE Conference on Intelligent Transportation Systems* (pp. 79-84). IEEE.

100. Uno, N., Kurauchi, F., Tamura, H., and Iida, Y. (2009). “Using bus probe data for analysis of travel time variability.” *Journal of Intelligent Transportation Systems*, *13*(1), 2-15.

101. Vandeput, N.: Data science for supply chain forecasting. De Gruyter. (2021).

102. Widana Pathiranage, R., Bunker, J. M., and Bhaskar, A. (2013, April). “A microscopic simulation model to estimate bus rapid transit (BRT) station service capacity with mixed stopping and non-stopping bus operation.” In *OPTIMUM 2013: International Symposium on Recent Advances in Transport Modelling*.

103. Widanapathiranage, R., Bunker, J. M., and Bhaskar, A. (2015). “Modelling the BRT station capacity and queuing for all stopping busway operation.” *Public Transport*, *7*(1), 21-38.

104. Xiong, Z., Zheng, J., Song, D., Zhong, S., and Huang, Q. (2019). “Passenger flow prediction of urban rail transit based on deep learning methods.” *Smart Cities*, *2*(3), 371-387.

105. Yan, Y., Liu, Z., and Bie, Y. (2016). “Performance evaluation of bus routes using automatic vehicle location data.” *Journal of Transportation Engineering*, *142*(8), 04016029.

106. Yang, H., Liang, X., Zhang, Z., Liu, Y., and Abid, M. M. (2020). “Statistical modelling of quartiles, standard deviation, and buffer time index of optimal tour in traveling salesman problem and implications for travel time reliability.” *Transportation Research Record*, 2674(12), 339-347.

107. Zheng, J., and Huang, M. (2020). “Traffic flow forecast through time series analysis based on deep learning.” *IEEE Access*, *8*, 82562-82570.

LIST OF PUBLICATIONS

Journals

1. **Halyal, S., Mulangi, R. H., and Harsha, M. M. (2022).** "Forecasting public transit passenger demand: With neural networks using APC data." *Case Studies on Transport Policy*,10 (2), 965-975. **https://doi.org/10.1016/j.cstp.2022.03.011** - (Published).

2. **Halyal, S., Mulangi, R. H., and Harsha, M. M. (2022).** "Short-term Passenger Demand Modelling Using Automatic Fare Collection Data: A Case Study of Hubli-Dharwad BRTS." *Advances in Transportation Studies*, 59. - **https://www.atsinternationaljournal.com** - (Published).

3. **Halyal, S., Mulangi, R. H., and Harsha, M. M. (2022).** " Investigation of travel time variability of bus rapid transit system; based on AVL-APC data." *Transport Policy*. – (Under Review).

International Conferences

1. **Halyal, S., Mulangi, R. H., Harsha, M. M., and Laddha, H. (2022).** "Comparative Assessment of Journey Time, Speed and Delay Characteristics of HDBRTS and Heterogeneous Traffic: Case Study on Speed and Delay Method." *In 13th International Conference on Transportation Planning and Implementation Methodologies for Developing Countries (TPMDC 2020)*, Mumbai, India.

2. **Halyal, S., Mulangi, R. H., Harsha, M. M., and Laddha, H. (2022**). "Study on Travel Time Characteristics of Hubli-Dharwad Bus Rapid Transit System in Comparison with Heterogeneous Traffic Lane." *In Recent Advancements in Civil Engineering (pp. 701-712). Springer Nature*, Singapore. **https://doi.org/10.1007/978-981-16-4396-5_61** - (Published).

3. **Halyal, S., Mulangi, R. H., & Harsha, M. M. (2022).** "Visualisation of Transit Passenger's Mobility from Automatic Fare Collection Data (AFC): Case Study of Hubli–Dharwad BRTS." *In Recent Advances in Transportation Systems Engineering and Management: Select Proceedings of CTSEM 2021 (pp. 431-447). Singapore: Springer Nature Singapore.* **https://doi.org/10.1007/978-981-19-2273-2_28** - (Published).

BIO DATA

Name	: Shivaraj Halyal
Department	: Civil Engineering
Specialization	: Transportation Engineering
Gender	: Male
Date of Birth	: 03/06/1989
Email Id	: shivaraj.halyal43@gmail.com
Mobile Number	: +91-9108646915
Address	: At Post. K. Junipet Tq. Ramadurg Dist. Belagavi, Karnataka, India Pin Code. 591114

Academic Profile:

Degree	Year of Completion	Institution	Board	Percentage/ CGPA
Ph.D. (Civil Engineering)	2023	National Institute of Technology Karnataka, Surathkal	Deemed University	6.77 %
MTECH (Highway Technology)	2014	RASTA Centre for Road Technology, Bangalore	VTU, Belagavi	81.00 %
BE (Civil Engineering)	2012	KLE's College of Engineering and Technology, Belagavi	VTU, Belagavi	72.86 %
PUC	2008	SMPU College, Dharwad	Department of Pre-University Education, Karnataka	48.66 %
SSLC	2005	C.D. Halyal High school, Ramadurg	Karnataka Secondary Education Examination Board	79.52 %