THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

Modelling Reliability and Distribution of Travel Times in Transit

Zhenliang Ma

B.Sc. (Hons.) and M.Sc. in Information Science and Technology

A thesis submitted for the degree of Doctor of Philosophy at

The University of Queensland in 2015

School of Civil Engineering

ABSTRACT

The performance of Transit Travel Time Reliability (TTR) influences service attractiveness, operating costs and system efficiency. Transit agencies have spent considerable effort on implementation of strategies related to advanced technologies capable of improving service reliability. Survey studies have shown that travelers tend to value a reduction in unreliability at least as important as a decrease in the average travel time. The increasing availability of data from automatic collection systems (e.g. automatic vehicle location, automatic fare collection, and etc.) provides opportunities in addressing transit TTR challenges. While most past studies estimate TTR for impact assessment of strategic and operational instruments, this research aims at developing generic models for TTR prediction that can fulfil different transit stakeholders' requirements (e.g. operators, unreliability causes identification; passengers, trip and departure planning). Three main issues are addressed, namely TTR quantification, TTR modelling and Travel Time Distribution (TTD) estimation. A unique integrated data warehouse was established for case studies of this research using different sources of data across six months of a year in Southeast Queensland area, Australia.

For TTR quantification, a set of TTR measures from the perspective of passengers using the operational AVL data was proposed, considering different perceptions of TTR under different traffic states. The results show that the proposed measure can provide consistent TTR assessments with high-level of details, while the conventional TTR measures may give inconsistent assessments. *For TTR modelling*, the underlying determinants of travel time unreliability were identified and quantified on links of different road types using Seemingly Unrelated Regression Equations (SURE) estimation to account for the cross-equation correlations across regression models caused by unobserved heterogeneity. Targeted strategies can be introduced to improve TTR under different scenarios. *For TTD estimation*, a novel evaluation approach was developed to assess the most appropriate probability distributions for travel time components (link running times and stop dwell times). The Gaussian Mixture Models (GMM) distribution was assessed to be superior to its alternatives, in terms of fitting accuracy, robustness and explanatory power. The correlation structures of travel time components were explored using both a global and a local correlation measures. On these basis, a generalized Markov chain model was proposed to estimate the trip TTDs for arbitrary origination-destination pairs at arbitrary times given the individual link TTDs, by considering their spatiotemporal correlations. The proposed approach is generalizable and computationally more efficient, while it provides a comparable performance with reported models in literature.

A major contribution of the research is the establishment of a generic TTD estimation methodology that can be applied for a comprehensive analysis and prediction of TTR to fulfill different requirements of operators and passengers in transit. The methodology is applicable under general

conditions as the link TTDs are derived conditional on the states of the current link and the transition probabilities are estimated as a function of explanatory covariates using logit models. The results of the research provide a better understanding on characterizing TTR from the perspective of passengers using the operational data, as well as the relationships between TTR and planning, operational, and environmental factors on different types of roads. In addition, the research demonstrates the existence of multiple traffic states for a given time period and the GMM distribution can well approximate the underlying characteristics of travel times, including symmetric, asymmetric and multimodal distributions.

In practice, the proposed TTD estimation methodology provides a generic tool to analyse and predict TTR that enables transit agencies to implement strategies to improve quality of service, as well as help transit users to make smart travel decisions (e.g. fast and reliable path). Given the complexity of problems and the constraint of available data, the empirical findings on the causes of travel time unreliability and the probability distributions of travel time components are valid within the range of the used data and should be used with caution beyond this range.

DECLARATION BY AUTHOR

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my research higher degree candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the policy and procedures of The University of Queensland, the thesis be made available for research and study in accordance with the Copyright Act 1968 unless a period of embargo has been approved by the Dean of the Graduate School.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate I have obtained copyright permission from the copyright holder to reproduce material in this thesis.

zhenliang Ma
Zhenliang Ma
School of Civil Engineering,
The University of Queensland

PUBLICATIONS DURING CANDIDATURE

Peer-Reviewed Journal Papers:

1. **Z. Ma,** J. Xing, M. Mesbah, L. Ferreira, “Predicting short-term bus passenger demand using a pattern hybrid approach,” *Transportation Research Part C: Emerging Technologies*, vol. 39, pp. 148-163, Jan. 2014. (IF: 2.82)
2. **Z. Ma,** L. Ferreira, M. Mesbah, S. Zhu, “Modelling distributions of travel time reliability for bus operations,” *Journal of Advanced Transportation*, Apr. 2015. (In press, IF: 1.88, Chapter 6)
3. **Z. Ma,** L. Ferreira, M. Mesbah, A. Hojati, “Modelling bus travel time reliability using supply and demand data from automatic vehicle location and smart card systems,” *Journal of Transportation Research Record*, Feb. 2015. (In press, IF: 0.55, Chapter 5)
4. **Z. Ma,** L. Ferreira, M. Mesbah, “Measuring service reliability using automatic vehicle location data”. *Mathematical Problems in Engineering*, vol. 2014, pp. 1-12, Apr. 2014. (IF: 1.08, Chapter 4)
5. **Z. Ma,** H. N. Koutsopoulos, L. Ferreira, M. Mesbah, “Trip travel time distribution estimation: A generalized Markov chain approach,” Submitted to *Transportation Research Part B: Methodological*, May. 2015. (Minor revision, Chapter 7)
6. N. Nassir, M. Hickman, **Z. Ma**, “Activity detection and transfer identification for public transit fare card data,” *Transportation*, vol. 42, pp. 683-705, Apr. 2015. (IF: 1.66)

Peer-Reviewed Conference Papers:

1. **Z. Ma,** L. Ferreira, M. Mesbah, A. Hojati, “Modelling bus travel time reliability using supply and demand data” In *Transportation Research Board (TRB) 94th Annual Meeting, Washington D.C. United States*, Jan. 2015.
2. **Z. Ma,** H. Koutsopoulos, L. Ferreira, M. Mesbah, Estimation of Traffic State Transition Probabilities and its Application to Travel Time Prediction. Submitted to *Transportation Research Board (TRB) 95th Annual Meeting 2016, Washington D.C. United States.* Aug. 2015 (Accepted, Chapter 7).
3. **Z. Ma,** L. Ferreira, M. Mesbah, “A framework for the development of bus service reliability measures,” In *36th Australasian Transport Research Forum (ATRF) Proceedings, Brisbane, Australia*, Oct. 2013.
4. N. Nassir, M. Hickman, **Z. Ma**, “Statistical inference of transit passenger boarding strategies from fare card data,” In *13th Conference on Advanced Systems in Public Transport (CASPT), Rotterdam, Netherlands*, Jul. 2015.

5. N. Nassir, M. Hickman, **Z. Ma**, "Behavioural findings of observed transit route choice strategies from the fare card data in Brisbane," Accepted by *37th Australasian Transport Research Forum (ATRF) Proceedings, Sydney, Australia*, Aug. 2015.

Publications included in this thesis

No publication included

Contributions by others to the thesis

No contribution by others

Statement of parts of the thesis submitted to qualify for the award of another degree

None

ACKNOWLEDGEMENTS

First of all, I would like to thank my supervisors Prof. Luis Ferreira and Dr. Mahmoud Mesbah at School of Civil Engineering of the University of Queensland (UQ), Brisbane, Australia, and visiting research advisor Prof. Haris N. Koutsopoulos at Department of Civil and Environmental Engineering of the Northeastern University (NEU), Boston, United States. I am grateful to Luis, Mahmoud and Haris for all their guidance, support and encouragement during my PhD studies. Together, they supported my research from both the theoretical and practical perspectives. I am always inspired by their knowledge and patience, and thus I have learned how to be a good researcher and instructor.

I acknowledge the efforts of the members of my PhD committee, thesis examiners (Prof. Peter Furth at Northeastern University and Dr. Jinhua Zhao at MIT, Boston, US) and appreciate their constructive remarks on my research. I would like to express my great appreciation to the committee chair, Prof. Mark Hickman, for his constructive comments on my research and financial supports on attending academic conferences, visiting research travel and professional trainings. I sincerely thank Prof. Phil Charles for supporting me to attend his series of blending training courses on Public Transport Professional Development. Many thanks to Prof. Fred Mannering at Purdue University for his suggestions on statistical analysis and Prof. Matthew Karlaftis (passed away) at National technical University for his comments on short-term demand prediction. Special thanks go to Prof. Nigel Wilson at MIT for his recommendations on my visiting research at US and my former supervisor Prof. Jianping Xing at Shandong University for his suggestions on future career plan. Also, I would like to acknowledge TransLink, division of Department of Transport and Main Roads in Brisbane, Australia for providing me necessary data to perform a realistic case study, and China Scholarship Council and UQ Graduate School for providing financial support for my PhD study.

At the UQ Transport Group, I want to give special thanks to my colleagues Dr. Ahmad Tavassoli Hojati, Dr. Neema Nassir, Dr. Inhi Kim, Dr. Ronald John S. Galiza, Dr. Sicong Zhu and many others I could not mention all by name. They were a great support during my entire research period. Furthermore, I would like to express many thanks to my friends and colleagues at UQ, CSIRO, NEU and MIT. It was really a fantastic and unforgettable time to talk, play and travel together, specially to Bing Guo, Helong Wu, Lang Liu, Svitlana Pyrohova, Lei Xiong, and Yue Wang.

Lastly, very special thanks to my parents (Changxing Ma and Xiangfen Wang) and elder sisters (Zhenfang Ma and Zhenchao Ma) for their love and tolerance throughout the time I am absent. I would like to dedicate this thesis to them for their support during my tough times.

马振良

August, 2015

Brisbane, Australia

Keywords

Travel time reliability, trip travel time distribution, Gaussian mixture models, Markov chain, spatial-temporal correlation, traffic state transition probability, automatic vehicle location and smart card data

Australian and New Zealand standard research classifications (ANZSRC)

ANZSRC code: 090507, Transport Engineering, 80%

ANZSRC code: 090599, Civil Engineering not elsewhere classified, 20%

Fields of research (for) classification

FoR code: 0905, Civil Engineering, 80%

FoR code: 0999, Other Engineering, 20%

Table of Contents

Table of Contents viii
List of Figures xi
List of Tables xiii
Abbreviations xiv
Chapter 1 Introduction 1
1.1 Background 1
1.2 Research aim and objectives 3
1.3 Thesis significance and contributions 3
1.4 Thesis outline 4
Chapter 2 Reliability and Distribution of Transit Travel Time: A Review of Past Work 5
2.1 Introduction 5
2.2 Travel time reliability definitions and measures 5
2.2.1 On-time performance 5
2.2.2 Headway regularity 6
2.2.3 Travel time 6
2.2.4 Waiting time 7
2.2.5 Transfer time 8
2.2.6 Buffer time 8
2.3 Travel time variability and unreliability causes 12
2.3.1 Sources of travel time variability 12
2.3.2 Travel time unreliability causes 13
2.3.3 Travel time reliability modelling 15
2.4 Travel time distribution fitting model 16
2.5 Travel time distribution estimation methodology 19
2.6 Summary of main findings and research gaps 21
Chapter 3 Data Description and Processing 24
3.1 Introduction 24
3.2 Data integration 24
3.3 Data processing 26
3.4 Case study area 27
3.5 Summary 28
Chapter 4 Travel Time Reliability Quantification 29

4.1 Introduction....29
4.2 Buffer time and its estimation....29
4.2.1 Passengers perspective on reliability....30
4.2.2 Buffer time estimation....31
4.3 Performance disaggregation....32
4.4 Measurement development....34
4.4.1 Reliability buffer time....34
4.4.2 Expected reliability buffer time (operator)....35
4.4.3 Trip planning time (passenger)....36
4.5 Case study....37
4.5.1 Probability distribution fitting....37
4.5.1.1 Single mode distribution....37
4.5.1.2 Mixture mode distribution....38
4.5.2 Assessment performance comparison....39
4.6 Discussions and applications....41
4.6.1 Strategy assessment (operators)....41
4.6.2 Trip planning (passenger)....42
4.7 Summary....43
Chapter 5 Travel Time Reliability Modelling....44
5.1 Introduction....44
5.2 Development of general models and alternative models....44
5.2.1 Dependent and independent variables....45
5.2.2 Seemingly unrelated regression equations (SURE) estimation....46
5.3 Case Study....47
5.3.1 Comparison between OLS and SURE estimations....48
5.3.2 General models for travel time reliability....49
5.3.2.1 SURE Model for Average Travel Time....49
5.3.2.2 SURE Model for Buffer Time....50
5.3.2.3 SURE Model for CV of Travel Time....51
5.3.3 Alternative models for travel time reliability....51
5.4 Main findings and practical implications....53
5.5 Summary....54
Chapter 6 Travel Time Distribution Modelling....56
6.1 Introduction....56
6.2 Distribution evaluation approach....57

6.3 Distribution evaluation measures....58
6.4 Case Study....59
6.4.1 Aggregation impacts on distribution....59
6.4.1.1 Temporal aggregation....59
6.4.1.2 Spatial aggregation....63
6.4.2 Distribution fitting performance evaluation....65
6.4.2.1 Route level distribution....65
6.4.2.2 Link level distribution....67
6.5 Discussions and applications....69
6.6 Summary....72
Chapter 7 Trip Travel Time Distribution Estimation....74
7.1 Introduction....74
7.2 Problem statement....74
7.3 Estimation framework....77
7.4 Methodology....78
7.4.1 State definition....78
7.4.2 Transition probabilities estimation....80
7.4.3 Probability distribution estimation....81
7.4.3.1 Generalized Markov Chain (GMC) approach....82
7.4.3.2 Moment Generating Function (MGF) Algorithm....83
7.5 Case Study....85
7.5.1 Trip TTD estimation....85
7.5.2 State definition....86
7.5.3 Transition probability model....88
7.5.4 Probability distribution estimation and performance analysis....92
7.5.4.1 Performance comparison....93
7.5.4.2 Sensitivity analysis....95
7.6 Discussions and applications....97
7.7 Summary....100
Chapter 8 Conclusions and Future Research....102
8.1 Summary of the thesis....103
8.2 Future research....105
References....106

List of Figures

Figure 1-1 Thesis outline 4
Figure 2-1: Flow diagram of reliability attributes concerned to the demand and supply sides 12
Figure 2-2: Time components for transit trip travel time.. 13
Figure 2-3: Interactions between demand and supply sides (adapted from (van Oort, 2011))......... 13
Figure 2-4: Time components of the dwell time at a stop 14
Figure 3-1 Overview of the integration scheme.. 25
Figure 3-2 Snapshot of the integrated data record 26
Figure 3-3: Data cleaning results and outliers identified 27
Figure 3-4 The used two transit routes in Brisbane, Australia.. 28
Figure 4-1: Journey departure decision and arrival time distribution.. 30
Figure 4-2: Travel time samples with mixture distributions.. 31
Figure 4-3: Illustration of possible service states and reliability buffer time 35
Figure 4-4: Travel time distribution fitting result using single and mixture models 39
Figure 4-5: CDFs and PDFs for different groups of GMM travel time samples 40
Figure 4-6: New designed trip planner for passengers 43
Figure 6-1: Distribution of travel times for routes 555 and 60 during (a) AM peak period and (b) inter peak period.. 60
Figure 6-2: Distribution of travel times with departure time window (DTW) 60 minutes and 15 minutes during AM peak period. 60
Figure 6-3: Travel times (actual and scheduled) and coefficient of variance (COV) between different stops along the route.. 64
Figure 6-4: Survivor function of Anderson-Darling (AD) test significance for alternative distribution models (a) route level and (b) link level. GMM Gaussian mixture models 66
Figure 6-5: Summary of the distribution of top 3 models for all cases (a) route level and (b) link level. GMM, Gaussian mixture models 68
Figure 6-6: Fitting results for a multimodality distribution (a) density and (b) cumulative probability. (Case: urban route, weekdays, eastbound, AM peak, travel time). GMM, Gaussian mixture models. 70
Figure 6-7: Fitting results for an asymmetric distribution (a) density and (b) cumulative probability. (Case: busway route, weekdays, inbound, inter peak, travel time). GMM, Gaussian mixture models. 70
Figure 6-8: Anderson–Darling (AD) test significance for distributions of hourly travel times over the whole day. (Case: urban route, weekdays, eastbound, hourly, travel time)................................ 71

Figure 7-1: Illustration of network components: segments, links, and trips 75
Figure 7-2: Two-dimensional diagram representing four groups of vehicles (F = Fast and S = Slow): (a) ideally uncorrelated observations within groups, (b) potentially correlated observations within groups. 77
Figure 7-3: Trip travel time distribution estimation framework 77
Figure 7-4: The description of GMMS clustering algorithm 79
Figure 7-5: The Markov chain structure to estimate trip travel time distribution 82
Figure 7-6: Global and local correlations of travel time components in AM peak period: (a) spatiotemporal autocorrelation function (ST-ACF) of unit running times between links with different spatial orders; and (b) cross-correlation function (CCF) between link unit running times and downstream stop dwell times on different types of roads. 85
Figure 7-7: The estimation approach for transit trip travel time distribution 86
Figure 7-8: State clustering results [Route 555, Weekday, Inbound, All Day] 87
Figure 7-9: Samples of correlated and uncorrelated running times [Route 60, Weekdays, Eastbound, 7-8 AM] 87
Figure 7-10: Mean and standard deviation of average silhouette width vs. number of clusters 88
Figure 7-11: Estimated transition probabilities with different congestion indexs of preceding time interval, (a) CI_PreL_CurT = 25%, RCI_CurL = 25% ; (b) CI_PreL_CurT = 25%, RCI_CurL = 75% ; (c) CI_PreL_CurT = 75%, RCI_CurL = 75% ; (d) CI_PreL_CurT = 75%, RCI_CurL = 45% 92
Figure 7-12: Probability density function and cumulative density function of the estimated distributions 94
Figure 7-13: KL performance metric as a function of trip distance and time of day (R555, inbound, weekdays). (a) Convolution, (b) GMC_GMMS_NC, (c) GMC_GMMS_CLN, and (d) MGF_CLN 96
Figure 7-14: The implementation of the proposed GMC structure for transit application 98
Figure 7-15: Predictions of means and intervals on a sample Motorway link for weekdays inbound trips over different times of day across six months period. [Model 1 uses the predicted probability; Model 2 uses the fixed probability; 95% conf. indicates 95% confidence intervals] 100

List of Tables

Table 2-1 General pool of reliability indicators....................10
Table 2-2 Selected studies on TTD fitting using single mode distribution model 18
Table 2-3: Summary of the main findings from literature review 22
Table 3-1: Description of Go card transaction errors 27
Table 4-1: Summary of fitting performance of single mode distribution models 38
Table 4-2: Parameters for the fitted single model and mixture models distributions....................38
Table 4-3: Parameters for different groups travel time samples....................40
Table 4-4: Assessment results for different groups travel time samples 41
Table 4-5: Assessment of service performance changes using different reliability measures....................42
Table 5-1: Description of Variables and Models....................46
Table 5-2: Descriptive Statistics of Dependent and Independent Variables....................48
Table 5-3: SURE Models for Average Travel Time, Buffer Time and CV Travel Time....................50
Table 5-4: SURE Models of Average Travel Time, Buffer Time and CV Travel Time on Different Types of Roads....................52
Table 6-1: Key descriptive statistics of travel times with different temporal aggregation level 62
Table 6-2: Characteristic of links and key descriptive statistics of TTDs [weekday inbound AM service 555] 63
Table 6-3: Characteristics of links and unimodal statistics of TTDs [weekday eastbound AM service 60] 64
Table 6-4: Descriptive summary of AD significance value and candidature distributions performance [route level]....................66
Table 6-5: Descriptive summary of significance value and candidature distributions performance [link level] 67
Table 6-6: Comparison of candidature models fitting performance for route level and link level travel times....................68
Table 7-1: Summary of dataset and descriptive statistics of variables 89
Table 7-2: MNL model estimation coefficients and performance....................90
Table 7-3: Specified MNL model coefficients and performance 91
Table 7-4: Performance comparison [R60, eastbound, 7:00-8:00, between stops 7 and 10]....................94
Table 7-5: Performance summary....................96
Table 7-6: Comparison of estimation performance for trips on different types of roads 97
Table 7-7: Summary of deterministic and interval predictions performance 99

Abbreviations

A-D	Anderson-Darling
AFC	Automatic Fare Collection
AIC	Akaike Information Criterion
ANPR	Automatic Number Plate Recognition
ATD	Average Trip Duration
AVL	Automatic Vehicle Location
ASW	Average Silhouette Width
AWT	Average Waiting Time
BoM	Bureau of Meteorology
BSTM	Brisbane Strategic Transport Management
BTI	Buffer Time Index
CBD	Central Business District
CCF	Cross-correlation Function
CDF	Cumulative Distribution Function
CI	Congestion Index
CV	Coefficient of Variation
DTMR	Department of Transport and Main Roads
DTW	Departure Time Window
EM	Expectation-Maximization
ERBT	Excess Reliability Buffer Time
ETC	Electronic Toll Collection
EWT	Excess Waiting Time
FCD	Floating Car Data
GMC	Generalized Markov Chain
GMM	Gaussian Mixture Models
GPS	Global Positioning System
GTFS	General Transit Feed Specification
KL	Kullback-Leibler
K-S	Kolmogorov-Smirnov
LTD	Latest Trip Duration
MAD	Median Absolute Deviation
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error

MGF	Moment Generating Function
MNL	Multinomial Logit
MVN	Multivariate Normal
OD	Origination-Destination
OLS	Ordinary Least Square
PDF	Probability Density Function
PTI	Planning Time Index
RBT	Reliability Buffer Time
RCI	Recurrent Congestion Index
RTI	Reliability Time Index
RV	Random Variable
SD	Standard Deviation
SEQ	Southeast Queensland
ST-ACF	Spatiotemporal Autocorrelation Function
STARIMA	Space Time Autoregressive Integrated Moving Average
SURE	Seemingly Unrelated Regression Equation
SW	Silhouette Width
SWT	Scheduled Waiting Time
TPM	Transition Probability Matrix
TTD	Travel Time Distribution
TTR	Travel Time Reliability
TTV	Travel Time Variability

Chapter 1 Introduction

1.1 Background

Transit agencies have spent considerable efforts on implementation of strategies related to advanced technologies capable of improving service reliability (Balcombe et al., 2004; Kittelson & Assoc et al., 2003). Survey studies have shown that travelers tend to value a reduction in unreliability at least as important as a decrease in the average travel time (Lam and Small, 2001). Reliability tends to be even more important in transit than in private car travel considering the transit passengers have only limited ability to adjust their departure times due to schedule constraints (Bates et al., 2001). Improving service reliability is believed to be a win-win situation for both operators and passengers (Abkowitz et al., 1978). Routes characterized by unreliable service may have difficulty in attracting potential riders and suffer patronage declines over time. Increased perceived burdens of waiting may ultimately impact mode choice decisions. Transit systems with poor reliability performance require extra fiscal resources due to higher operation costs (Kimpel, 2001).

Service reliability can be defined as the probability that a service can perform a required *function* under a given *condition* (recurrent and non-recurrent) for a stated *time period* (e.g. hourly, daily, monthly and yearly). The *function* can be connectivity reliability (Bell and Iida, 1997), capacity reliability (Chen et al., 2002) and travel time reliability (TTR) (Ng and Waller, 2010). This research is categorized as a study of TTR that focuses on daily recurrent unreliability caused by variations of traffic flow and demand when the infrastructure is fully available. Non-recurrent unreliability is not considered which are less frequent and relates to infrastructure failure (Tahmasseby, 2009).

The concern with the impacts of reliability on operation efficiency for operators and passengers brings about the need to identify and develop meaningful and consistent indicators of reliability. The workable and consistent reliability measurement can help to (Abkowitz et al., 1978): identify and understand problems in reliability; identify and measure actual improvements in reliability; relate such improvements to particular strategies; and modify strategies to obtain greater reliability improvements. At issue is that reliability has been defined in a variety of ways. Some studies associated reliability with on-time performance (Bates et al., 2001; Meyer, 2002), while others related it to travel time variability, headway regularity (Janos and Furth, 2002; Yu et al., 2010), waiting time (Fan and Machemehl, 2009; Furth and Muller, 2006). Integrated measures incorporating several service attributes were also reported (van Oort and van Nes, 2010). The emergence of automatic data collection technologies produces a wealth of accurate, continuous and automated point-to-point data that can be used to assess reliability more cost-effectively (Mesbah et al., 2012). *The framework of quantifying TTR from passengers' perspective using operational data needs investigation.*

To design appropriate strategies to improve service reliability, policy makers should be clear about the causes of unreliability. Various basic factors have been identified as affecting transit TTR (El-Geneidy et al., 2011; Mazloumi et al., 2010; Strathman et al., 2002; Tétreault and El-Geneidy, 2010). These factors include segment length, passenger activities (boardings and alightings), lift use, signalized intersections, number of scheduled stops, number of actual stops made, delay at the start, day of the week, time period of the day, service direction, weather conditions (rain and snow) and drivers experience. Accordingly, agencies implement strategies with expectations of improving service performance. Several researchers have investigated different strategies influencing running time and running time variability (Diab and El-Geneidy, 2012). These strategies include smart fare card collection system, reserved bus lanes, limited-stop bus services, stop consolidation, articulated buses and transit signal priority. *Constrained by the available data and the regression approach, the existing findings only provide partial understanding of unreliability causes impacts on TTR.*

Travel time distribution (TTD) contains maximum information that capture the stochastic characteristics of travel times (Du et al., 2012). Better understanding of the distribution of travel times is a prerequisite for analysing reliability and exploring the causes of unreliability (Sumalee et al., 2013). Many studies on TTR have attempted to fit mathematical distributions to travel times at different network levels (Clark and Watling, 2005; Fosgerau and Fukuda, 2012; Hollander and Liu, 2008). While some studies have considered symmetrical distribution models, for example, Normal (May et al., 1989), others have preferred skewed ones, for example, Lognormal (Emam and Ai-Deek, 2006). Recent studies have reported that a range of travel times could be found even for 5 min intervals (Zheng and Van Zuylen, 2010), and thus multimodal distributions could be more appropriate, for example, Gaussian Mixture Models (GMM) (Guo et al., 2010). *These inconsistencies clearly affect both the ability to gain insights into the nature of TTR and inhibit the ability to generalize findings to other applications.*

For many applications, e.g. trip planning, trip travel time information is of more interest (Bhat and Sardesai, 2006). The trip TTD can be derived or inferred using archived data of directed observations for the same origin and destination (OD) pairs under similar trip conditions, e.g. time period. One problem is that the archived database requires the full coverage of all OD pairs that travellers might take. Furthermore, with data from mobile sources, it is likely that for many OD pairs very few or no samples were observed. An effective approach for estimating trip TTDs between arbitrary OD pairs at arbitrary times is from individual link TTDs. Link travel times can be derived directly (e.g. transit AVL data) or estimated from the increasingly available but sparse opportunistic sensor data, e.g. vehicular GPS, Automatic Number Plate Recognition (ANPR), and mobile phone data (Hellinga et al., 2008; Hunter et al., 2009; Jenelius and Koutsopoulos, 2015; Rahmani and Koutsopoulos, 2013; Zheng and Van Zuylen, 2013). *The research on the estimation of trip TTDs from link TTDs is still evolving and insufficient.*

1.2 Research aim and objectives

Many researchers have highlighted the importance of TTR for both transit operators and passengers. While many studies estimate TTR for impact assessment of strategic and operational instruments, methods for prediction of TTR for decision making at all levels are still evolving and limited. Thus,

> ***The main aim of this research is to develop a generic approach to predict TTR that can fulfil different stakeholders' requirements.***

The following set of objectives with regard to TTR quantification, TTR modelling, and TTD estimation have been identified to accomplish the main aim:

1. Investigate the characteristics of current TTR indicators and develop new measures to quantify TTR from passengers' perspective using the operational data.
2. Develop a model to quantify and identify the influence of contributory factors on TTR.
3. Develop an approach to investigate spatiotemporal aggregation influence on TTD and specify the most appropriate link TTD model.
4. Propose a methodology to estimate trip TTD between arbitrary origin-destination (OD) pairs at arbitrary times from link TTDs.

1.3 Thesis significance and contributions

Accurate prediction of TTR can facilitate the implementations of proactive traffic management strategies and advanced traveler information system, which is a key component in addressing unban mobility issues. The main contributions of this research are:

1. New approaches have been developed to quantify and model TTR using AVL data.
2. A generalized methodology has been proposed to estimate trip TTDs from link TTDs.

In addition, the following outcomes are achieved during this research:

3. Development of an algorithm to integrate data from different databases, including AVL, Smart Card Transactions, General Transit Feed Specification (GTFS), Brisbane Strategic Transport Management (BSTM), and Bureau of Meteorology (BoM) data.
4. Development of TTR models for different types of roads using a Seemingly Unrelated Regression Equation (SURE) approach, as opposed to the Ordinary Least Square (OLS).
5. Investigation of spatiotemporal aggregation influence on TTD and development of an approach to specify the most appropriate probability distribution of travel times.
6. Development of a transition probability estimation model using a logit model formulation with the utilities being a function of link characteristic and trip conditions.
7. Development of a link TTD prediction method using TTDs conditional on states and logit model predicted transition probabilities.

1.4 Thesis outline

Figure 1-1 shows the thesis outline. Chapter 1 introduces the research background on TTR and TTD, establishes the research aim and objectives to be achieved, and describes the contributions and outline of this research. Chapter 2 reviews the relevant literature in the field of TTR and TTD, identifies the gaps in the existing knowledge of TTR measurement, TTR modelling and TTD estimation. An overview of various data sets, their processing and integration is presented in Chapter 3.

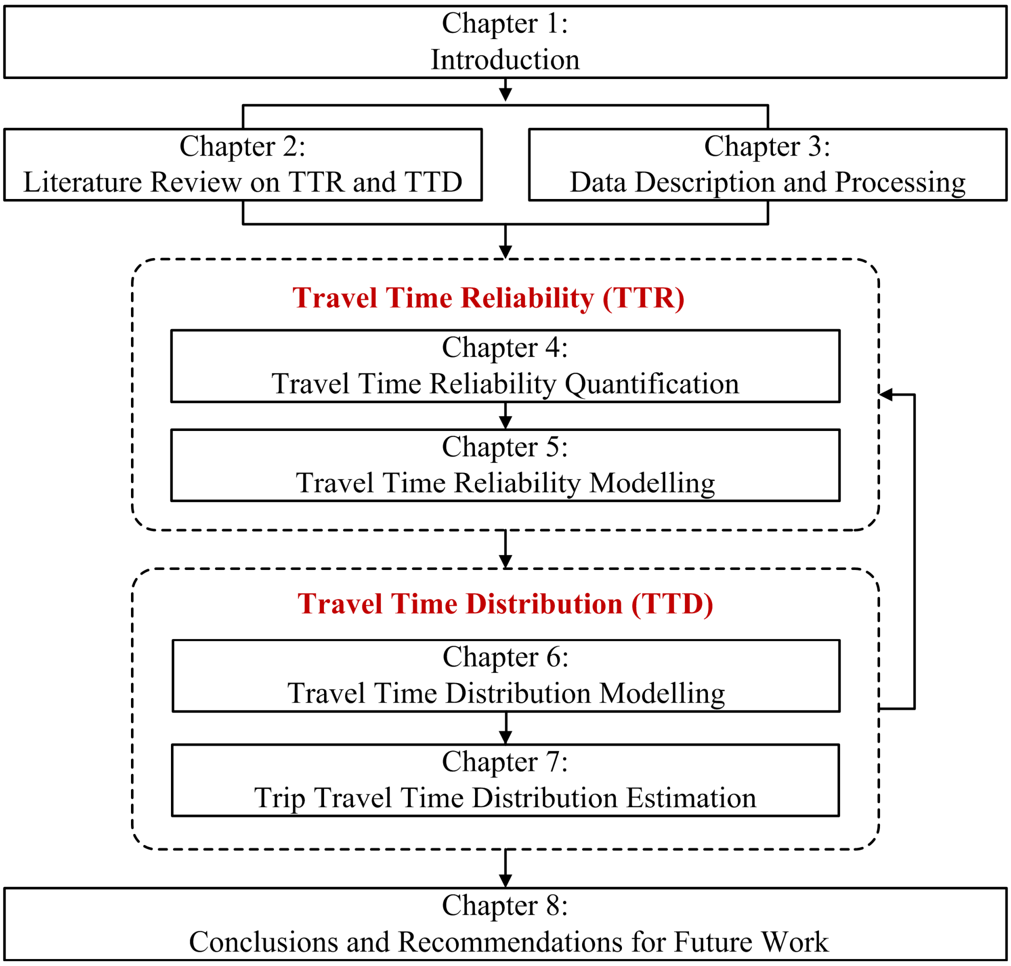

Figure 1-1 Thesis outline

The research consists of two main parts, namely TTR and TTD analysis as shown in Figure 1-1. Chapter 4 proposes a framework to quantify TTR from passengers' perspective using operational AVL data. A set of TTR models was then developed to identify and quantify the impact of unreliability factors on different types of roads in Chapter 5. The necessity to incorporate distribution information in TTR analysis and TTR prediction motivates the TTD related research. Chapter 6 specifies the most appropriate distribution models for link travel times. Based on these, a generalized approach is proposed to estimate the trip TTDs between arbitrary origination-destination pairs at arbitrary times from link TTDs in Chapter 7. Finally, the conclusions and recommendations from this research are given in Chapter 8.

Chapter 2 Reliability and Distribution of Transit Travel Time: A Review of Past Work

2.1 Introduction

This chapter reviews the relevant literature in the fields of TTR and TTD. Section 2.2 provides an overview of the definitions and measures of TTR. A general pool of TTR indicators is summarized, from which a sub-set can be selected according to different objectives and operational constraints. This is followed by discussions on sources of service variations and significant factors that affect TTR in Section 2.3. The following Section 2.4 provides insights into TTD fitting models, as well as spatiotemporal aggregation influence on TTD. The trip TTD estimation methodologies are then explored along with their major assumptions in Section 2.5. Finally, Section 2.6 summarizes the major findings from the literature review and identifies the gaps in the existing knowledge of TTR modelling and TTD estimation.

2.2 Travel time reliability definitions and measures

The reliability concept is interpreted and perceived diversely across groups of stakeholders and various studies have defined reliability from different aspects of transit service. While some studies associated reliability with travel time (Hollander, 2006; Mazloumi et al., 2008), others related it to maintain headway regularity (Janos and Furth, 2002; Yu et al., 2010), on-time performance (Bates et al., 2001; Meyer, 2002), and passenger waiting time at stops (Fan and Machemehl, 2009). Abkowitz et al. (1978) defined the reliability as the invariability of service attributes which influence the decisions of planners and travellers. It provides two key insights, consistency of the service attributes and distinct perspectives between demand-side and supply-side. Ceder (2007) identified six time-related service attributes concerned by demand-side and supply-side, namely, on-time performance, headway regularity, travel time, waiting time, transfer time and buffer time. A general pool of service reliability indicators based on which different sets of indicators can be selected for different objectives is summarized in Table 2-1.

2.2.1 On-time performance

For routes characterized by low frequency services, schedule adherence plays the most significant role, since passengers are expected to plan their arrivals to coordinate with the scheduled departures to minimize waiting time at stops with a tolerance probability of missing the trips.

On time performance is a commonly used schedule adherence measure in applied environments, defined as the percentage of trips that depart up to *m* minutes late and *n* minutes early from the scheduled departure time. The US Transportation Research Board presented a service delivery measure survey where zero minutes was the most common earliness threshold and 5 minutes was the most common lateness threshold (Kittelson & Associates et al., 2003). Camus et al. (2005) have proposed a weighted delay index, which is an interesting extension of an on time performance measure. Nakanishi (1997) has given a detailed discussion and potential improvements of on time performance indicators.

2.2.2 Headway regularity

For routes characterized by high frequency services, headway based measures become important (Currie et al., 2012). In these circumstances, passengers are prone to arrive at stops randomly, and the aggregate waiting time of passengers is minimized when services are evenly spaced (Osuna and Newell, 1972). Many indicators are proposed in this domain. Some indicators are defined by comparing with scheduled headway, such as service regularity, headway ratio (Strathman et al., 1999) and percentage regularity deviation mean (van Oort and van Nes, 2004), while others are defined based on headway distribution, such as standard deviation, coefficient of variance, average waiting time (Osuna and Newell, 1972) and probability-based headway regularity measure (Lin and Ruan, 2009). Additionally, two indicators are developed for specific purposes. The headway regularity index identifies the vehicle bunching problem while the irregularity index can effectively indicate long gaps between vehicles (Golshani, 1983).

On-time performance and headway regularity are schedule-based indicators. The main issue is that no universal benchmarking threshold can be found to mark the difference between frequent and infrequent services and define the on-time tolerance interval. Moreover, they cannot reflect demand-side perception of reliability. By altering the on time tolerance interval from 5 minutes to 10 minutes, the measured service performance improves without any changes perceived by passengers.

2.2.3 Travel time

According to Kaparias et al. (2008), most travel time reliability indicators use various features of the travel time distribution. Lomax et al. (2003) categorized them in three groups, namely statistical range measures, buffer measures and tardy trip indicators. When dealing with people’s perceptions, it appears to be more appealing to separate physical from psychometric performance indicators (Pronello and Camusso, 2012). For travel time reliability, physical indicators describe it as ‘it is what it is’, while psychometric indicators reflect it as ‘it is what it is perceived to be’. The following discusses physical indicators.

Statistical Range Indicators: This type of measure typically serves as an approximate estimate of the range of trip situations experienced by passengers, calculated on standard deviation statistics. Standard deviation of travel time represents reliability in such way that small values are considered reliable. Percentage variation of travel time, statistically known as the coefficient of variation, provides a clearer picture of the trends and performance characteristics than the standard deviation by eliminating route length from the calculation. Moreover, percentage variation is dimensionless thus enabling a comparison between links and routes to be made. The travel time window is defined as the average travel time plus or minus the standard deviation of travel time, and can provide the passenger with an idea of how much the travel time will vary (Lomax et al., 2003). The variability index is defined as a ratio of peak to off-peak variation in travel conditions, and is calculated as a ratio of the difference in the upper 95% and lower 95% confidence intervals between the peak period and the off-peak period.

Tardy Trip Indicators: Tardy trip measures are extreme values of travel time. The tardy trips are identified by setting unacceptable limit values in the form of additional minutes plus expected time or percentage over expectation. In most cases, these values are arbitrarily set. The Florida reliability measure (FRM) uses a percentage of the average travel time in the peak to estimate the limit of the tolerable travel time range. Travel time exceeding the expectations is termed a tardy trip. Extended FRM uses travel rate (travel time per unit distance) instead of travel time, so as to provide a length-neutral way of grading the service performance (Lomax et al., 2003). The misery index examines trip reliability by using the difference between the average travel rates of the worst trips and all trips.

Skew-Width Indicators: Skew and width of travel time distribution measures are based on percentiles (van Lint and van Zuylen, 2005). Skew of travel time distribution is defined as the ratio of the difference between the 90^{th} and 50^{th} percentile and the difference between the 50^{th} and 10^{th} percentile. Width of travel time distribution indicates the distribution compactness. The wider the distribution is, the lower the reliability will be.

2.2.4 Waiting time

Waiting time at a stop is, from the perspective of passengers, the most significant component of public transit travel and often cited as one of the most important factors hindering the usage of bus transit. Generally, waiting time indicators can be categorized into two groups, namely, mean-variance based and extreme-value based (van Oort and van Nes, 2004).

Mean-variance based : Excess waiting time (EWT) is defined as the difference between the average waiting time (AWT) and the scheduled waiting time (SWT) (Trompet et al., 2011). For frequent services, the SWT is defined as the average time passengers would wait when the service operates exactly as scheduled (Liu and Sinha, 2007).

For high frequency services, a commonly used AWT indicator is half the headway of successive buses, based on three assumptions: passenger arrives randomly, passenger catches the first bus that comes, and vehicles arrive regularly (Fan and Machemehl, 2009). Under irregular vehicle arrival condition, the AWT is calculated as $\mathrm{AWT} = \mu * \left(1 + s^2/\mu^2\right)/2$, where μ is mean headway and s^2 is headway variance (Osuna and Newell, 1972). Furthermore, under non-random passenger arrivals and irregular vehicle arrival conditions, empirical AWT models relate passenger waiting time with mean headway(Fan and Machemehl, 2009). Theoretical ones AWT models construct a relationship between "aware" passenger arrival patterns and service performance through an explicit behavioural mechanism.

Extreme-value based: Passengers are more concerned about extreme values in their perception of service performance when budgeting their arrival at stops. Budget waiting time is defined as 95th percentile waiting time for frequent services. It serves as the total waiting time that a passenger should budget for a trip to avoid missing expected services at a stop under certain probabilities. Potential waiting time, defined as the difference between budgeted waiting time and mean waiting time, serves as the buffer time that a passenger should plan for their arrival at stops (Furth and Muller, 2006). The concept of extreme-value based indicators separates the impact on operations from the impact on passenger planning. Extreme-value based waiting time is far more sensitive to service reliability than mean-variance based AWT.

2.2.5 Transfer time

Transfer time can be calculated from scheduled stops (Jang, 2010). Therefore, statistic indicators can be applied to measure transfer time reliability, such as the coefficient of variation of transfer delays (Turnquist and Bowman, 1980). However, day-to-day arrival time variations make the measurement rather difficult (Kittelson & Associates et al., 2003). Transfer waiting time usually serves as a transfer time reliability indicator (Ceder, 2007; Goverde, 1999). Goverde (1999) derived an expected transfer waiting time model, a function of arrival delays distribution, incorporating the risk and significance of missing connections.

2.2.6 Buffer time

The buffer time indicates extra travel time required to allow the passengers' on time arrival. Generally, it is defined as the difference between xx percentile and the average travel time. The planning time is defined as the xx percentile travel time. It indicates the total time that a passenger has to budget for the trip. Buffer time index is defined as the buffer time divided by the average travel time. These indicators associate closely with the way passengers make trip decisions (Lomax et al., 2003). Uniman et al. (2010) proposed the general form of an initial set of reliability buffer time measures under the 'percentile-based' and 'slack time' approach.

Reliability buffer time, defined as the difference between the upper percentile *xx*, and an intermediate or lower percentile *yy*, is the additional time that would be required to be *xx*-percent sure of arriving at the destination on time. Excess reliability buffer time (ERBT) is defined as the difference between the actual levels of reliability experienced by passengers and what they should have experienced had everything gone according to plan. The ERBT indicator can be used to capture the incident-caused additional unreliability above that was caused by recurrent factors.

Abkowitz et al. (1978) evaluated the typical service reliability measures in an applied environment and selected several criteria, including explicitness of definition, controllability, expense and accurate measurability, and independence. In defining summary statistics to assess the variability distribution impacts, three separate criteria were identified, including distribution compactness, likelihood of extremely long delays, and normalization of measures. Currie et al. (2012) developed a framework to assess reliability indicators based on four criteria. Summarizing the evaluation criteria mentioned above, several key effective indicators are identified: (1) passenger focused; (2) easy to understand; (3) consistent and objective; (4) easy to compare and aggregate; and (5) insights into unreliability causes provided.

Conceptually, buffer time based indicators fulfil the criteria described above. It is passenger focused, easy to understand, consistent and objective, comparable across different routes and time periods, easy to aggregate weighted values by passenger demand of each OD pair, and can also provide operators with insights into causes of unreliable service at different levels, such as route and network. Analytical and empirical studies have confirmed buffer time as a powerful tool in indicating and estimating service reliability (Pu, 2011).

Though buffer time is usually defined as buffer travel time, strictly speaking, it can be recognized as an extreme value based concept to evaluate reliability performance. It can be applied manifold: (a) buffer waiting time to indicate budgeted waiting time needed to catch the expected bus; (b) buffer transfer time to indicate additional time required to avoid missed connections; and (c) buffer travel time to indicate extra time necessary for on time arrival .

Table 2-1 General pool of reliability indicators

Attribute(1)	Indicators	Definitions (2)
On-Time Performance*	% On-Time Arrival/Departure	Percentage of arriving or departing a stop up to *m* minutes late and *n* minutes early
	Odds Ratio	$\frac{\%\,\text{On-Time Arrival}}{1\text{-}\%\,\text{On-Time Arrival}} \times 100$
	Weighted Delay Index	$\sum_{d=1}^{H} dP(d) \big/ H$
	On-Time Distribution	Distribution of difference between actual running and scheduled time
Headway Regularity*	Service Regularity	% of headways deviating within the predefined scheduled interval
	Percentage Regularity Deviation Mean	$\sum_i \left\lvert (h_{i,j} - H_{i,j}) / H_{i,j} \right\rvert \big/ n_j$
	Headway Regularity Probability	$P\left\{h_{i,j} \le Hmax_j\right\}$
	Standard Deviation of Headway	$\mathrm{SD_H} = \sqrt{\sum_{i=1}^{n_j} \left(h_{i,j} - \bar{h}_j\right)^2 \big/ (n_j - 1)}$
	Average Waiting Time	$\left(\bar{h}_j + \mathrm{SD_H^2} / \bar{h}_j\right) \big/ 2$
	Excess waiting time	$\left(\sum_i h_{i,j}^2 / \sum_i h_{i,j} - \sum_i H_{i,j}^2 / \sum_i H_{i,j}\right) \big/ 2$
	Coefficient of Variance of Headway	$\mathrm{CV_H} = \mathrm{SD_H} / \bar{h}_j \times 100$
	Headway Regularity Index	$1 - 2\left[\sum \left(h_r - \bar{h}_j\right) r\right] \big/ n_j^2 \bar{h}_j$
	Irregularity Index	$1 + \mathrm{CV_H^2}$
Travel Time*[#]	Standard Deviation of Travel Time	$\mathrm{SD_{TT}} = \sqrt{\frac{1}{N-1} \sum^{N} \left(TT_i - \overline{TT}\right)^2}$
	Travel Time Variability	$TT90 - TT10$
	Travel Time Window	$\overline{TT} \pm \mathrm{SD_{TT}}$
	Coefficient of Variation of TT	$\mathrm{SD_{TT}} \big/ \overline{TT}$
	Variability Index	$\left(UCL_{peak} - LCL_{peak}\right) \big/ \left(UCL_{off-peak} - LCL_{off-peak}\right)$
	Extended Florida Reliability Measure	$100\% - \left(count\,\vert_{TR_i > (1+p)\overline{TR}} \big/ count\,\vert_{TR_i}\right)$
	Misery Index	$\left(\overline{TR}\big\vert_{TR_i > TR80} - \overline{TR}\right) \big/ \overline{TR}$
	Travel Time Distribution Skew	$\left(TT90 - TT50\right) \big/ \left(TT50 - TT10\right)$
	Travel Time Distribution Width	$\left(TT90 - TT50\right) \big/ TT50$

Waiting Time#	Scheduled Waiting Time	Average scheduled headway during the analysis period
	Excess Waiting Time	Difference btw average waiting time and scheduled waiting time
	Empirical Average Waiting Time	$a\bar{h}_j + b$
	Theoretical Average Waiting Time	$(1-q)\left[pw_{min} + (1-p)w_{rand}\right]$
	Budget Waiting Time	$\mathrm{BWT}_{frequent} = W_{0.95}$ $\mathrm{BWT}_{infrequent} = V_{0.95}$
	Potential Waiting Time	Difference between budgeted waiting time and mean waiting time
Transfer Time#	CV of Transfer Delay	Coefficient of variation of transfer delays
	Transfer Waiting Time	Function of arrival delay of the feeder service
	Expected Transfer Waiting Time	Function of arrival delays distribution of the feeder service
Buffer Time#	Buffer Time	$\mathrm{BT} = TTxx - \bar{TT}$
	Buffer Time Index	$BT/\bar{TT}$
	Planning Time	$TT95$
	Reliability Factor	$TTxx - TT50$
	Reliability Buffer Time	$\mathrm{RBT} = TT95 - TT50$
	Excess Reliability Buffer Time	$RBT_{\text{overall}} - RBT_{\text{recurrent}}$
	% of Unreliable Journeys	$\mathrm{PUJ} = \left\{\begin{matrix} pecentage\ of\ overall\ journeys \\ with\ TT > RBT_{\text{recurrent}} \end{matrix}\right\}$
	% of Excess Unreliable Journeys	$PUJ - \left\{\begin{matrix} pecentage\ of\ journeys\ under\ recurrent \\ condition\ with\ TT > RBT_{\text{recurrent}} \end{matrix}\right\}$

Note: (1) * refers to operator-focused attribute; # refers to passenger-focused attribute.

(2) Term Definitions:

m,*n* -- Given time window limits, H -- Scheduled headway, d -- Delay value, $P(d)$ -- Probability for delay d, $h_{i,j}$, $\bar{h}_j$ -- Observed bus *i* and mean headway at stop *j*, $H_{i,j}$ -- Scheduled headway for bus *i* at stop j, $Hmax_j$ -- Expected max headway for stop j, n_j -- Number of buses at stop j, h_r -- Series of headways, r -- Ascending rank order of the headway, TT_i, $\bar{TT}$ -- Observed and average travel time, $TTxx$, $TRxx$ -- xx^{th} percentile of travel time and travel rate, TR_i, $\bar{TR}$ -- Observed and average travel rate, LCL_{peak}, UCL_{peak} ($LCL_{off-peak}$, $UCL_{off-peak}$)-- Lower and upper confidence limit for peak (off-peak) period, *a*,*b* -- Constant, p, q -- Predefined percentage level, $W_{0.95}$, $V_{0.95}$ -- 95th percentile of waiting time and scheduled headway deviation, RBT_{overall}, $RBT_{\text{recurrent}}$ -- Overall and recurrent reliability buffer time.

2.3 Travel time variability and unreliability causes

van Oort (2011) distinguished travel time variability (TTV) and TTR. TTV is the service variations on the supply side. TTR is defined as the matching degree of the supplied and the expected service (perceived by the demand side). TTR tends to vary in time and space impacted by different sources of variations from demand and supply sides, as well as interactions between both sides. Figure 2-1 shows the journey attributes concerned to the demand and supply sides. Conceptually, if the variations of all attributes are low, the service has a high reliability.

2.3.1 Sources of travel time variability

The sources of TTV can be generally categorized into two classes, namely, variations in passengers' behavior (demand-side), and operation performance (supply-side) (Tahmasseby, 2009).

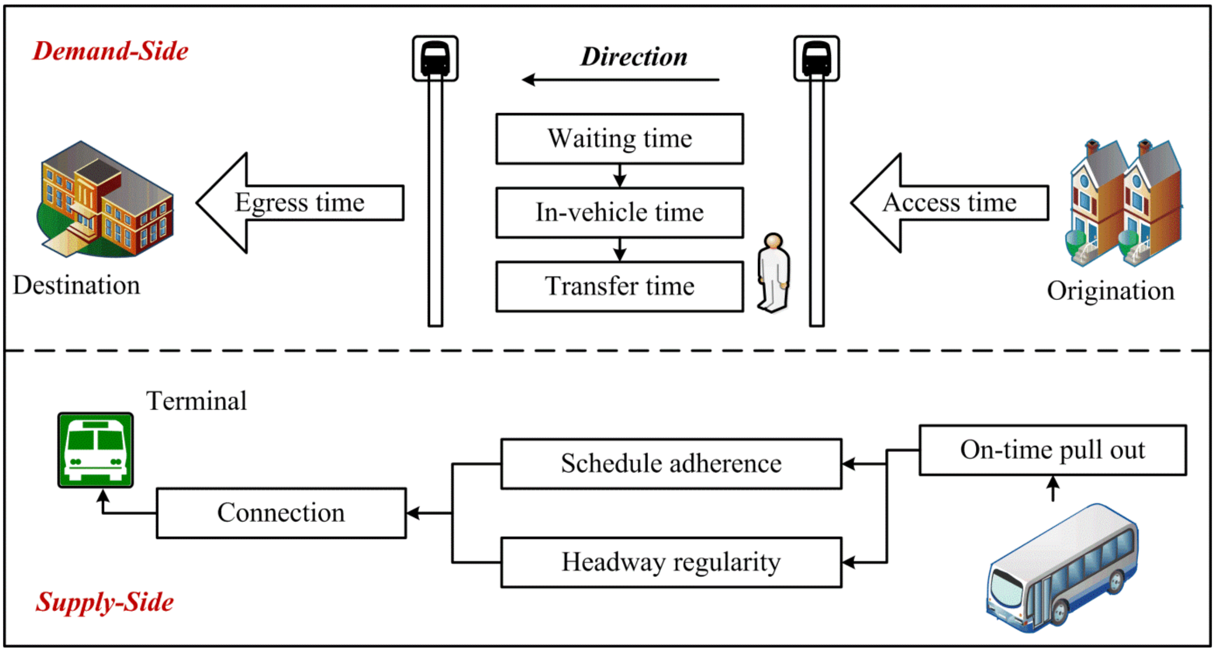

Figure 2-1: Flow diagram of reliability attributes concerned to the demand and supply sides

Demand-side: Access time is the time used from the origin to the boarding stop and egress time is the time used from the alighting stop to the destination. At the stop the waiting time occurs between passengers' arrival and the departure of vehicles. Passengers may arrive randomly or plan their arrival, and the budgeted waiting time may be preserved to avoid missing the expected vehicle at the stop (Furth and Muller, 2006). After successfully boarding a vehicle, the following component is in-vehicle time till the vehicle arrives at the destination stop. Passengers may transfer one or more times for a complete journey. All the time components are spatiotemporally stochastic.

Supply-side: For the fixed service, vehicle trips are scheduled in time and space resulting in on-route schedule adherence at all stops for infrequent service and headway regularity for frequent service. The supply variations includes terminal departure and trip time variations (van Oort, 2011).

The travel time of a transit trip consists of two components (shown in Figure 2-2), namely link running time between two consecutive stops and dwell time at a stop. Generally, running time is determined by the inherent network structure and link characteristics, speed profile, schedules and timetables, operational control strategies and weather (Sun et al., 2014). Dwell time mainly depends on passenger demand and various factors such as vehicle characteristics, crowding effects, and fare payment (Tirachini, 2013).

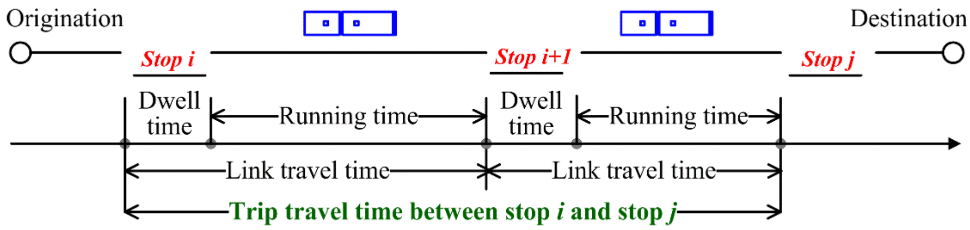

Figure 2-2: Time components for transit trip travel time

Interactions: Figure 2-3 shows the interactions of time components between the two sides. From the perspective of passengers, they are particularly concerned on the mean and variation of total travel times. The variation of travel times comes from both supply and demand sides. For instance, the combined impacts of passengers' arrival pattern, vehicle departure time and headway determine the variation of waiting time at a stop. From the perspective of operators, the dwell time at a stop is largely determined by passengers' activities (e.g. boarding, alighting, lift use, etc.).

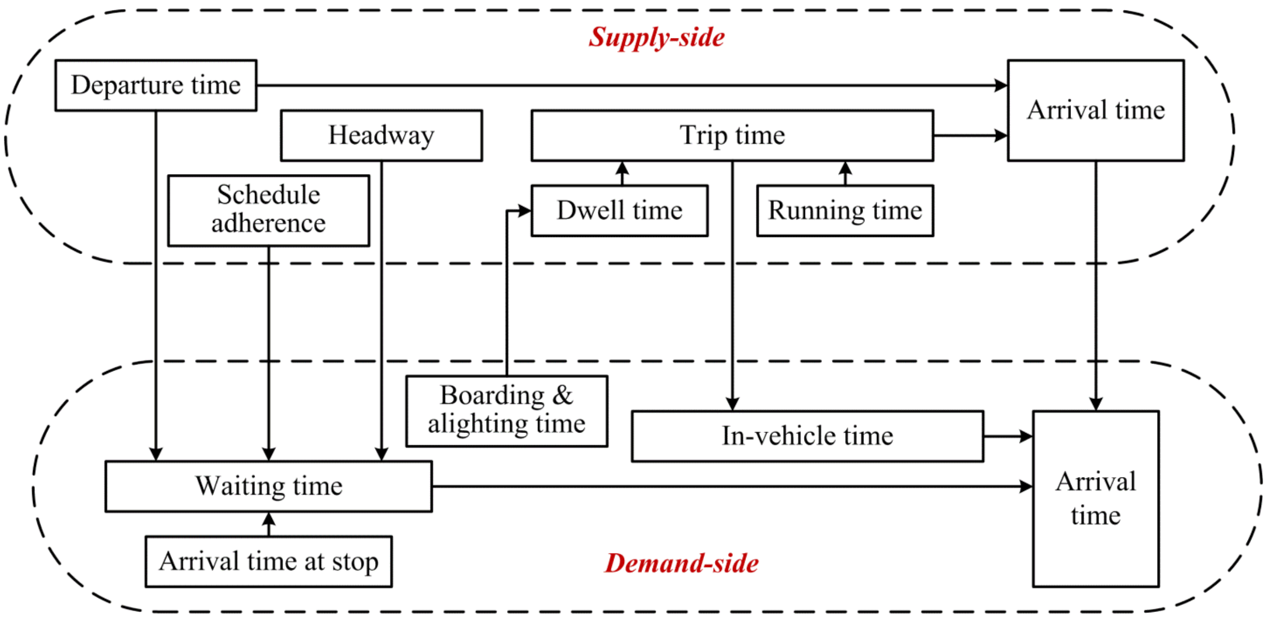

Figure 2-3: Interactions between demand and supply sides (adapted from (van Oort, 2011))

2.3.2 Travel time unreliability causes

Generally, the components of a transit trip travel time include departure delay from the first stop, dwell times at stops and link running times between adjacent stops. The causes of unreliability related to different trip time components are discussed separately.

Departure delay: this is the schedule deviation (early or late) of the actual departure at the terminal. Departure delay variation can introduce TTV and cause bunching at stops. In most cases, an early departure is regarded to be much worse than a late departure since passengers have to wait for a whole time interval between consecutive vehicles, especially for an infrequent service. The determinants of departure delay variation include: crew and vehicle availability, terminal infrastructure configuration (capacity, loading area, turning movements, etc.), timetable quality (slack of the layover time), driver behaviour (response to delay) (Kaas and Jacobsen, 2008; van Oort and van Nes, 2010) .

Stop dwell time: dwell time, the time a vehicle spends to load and unload passengers, is often the key determinant of speed and capacity (Dueker, 2004; Lin and Wilson, 1992; Tirachini, 2011). Most researches related dwell time with passenger demand, while others related dwell time with secondary factors such as fare collection methods, bus types, number of doors *et.al*. These factors may strongly influence the effectiveness of different strategies used to improve service (Milkovits, 2008). Among the determinants, passenger activity is recognized to be the principal determinant of dwell time and was studied most (Chen, 2012). Figure 2-4 shows the time components of the dwell time at a stop.

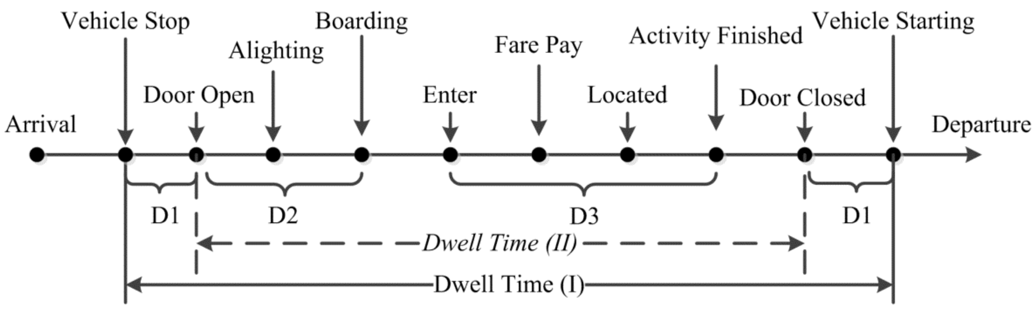

Figure 2-4: Time components of the dwell time at a stop

D1 (stop delay): Bus bunching, feeder route type, stop area condition and configuration, land use.
D2 (passenger demand): Boarding numbers, alighting numbers, max or sum of boarding & alighting, alighting by the front door or side door, time periods, passenger ages, platform crowding, cross town or radial, on-time, stop spacing.
D3 (passenger activity): Payment method, vehicle types, atypical passengers, lift operations, passengers friction, standee numbers, rank of boarding passengers, bus occupancy.

Link running time: this is composed of driving time and unplanned stopping time (caused by uncontrolled intersections excluding controlled intersection stop). Peng et al. (2009) classified the causes as environmental, planning, operational. Environmental factors include traffic conditions, number of signals, road work, on-street parking and demand variability. Planning factors include route length, schedules, and service frequencies. Operation factors include departure delays, vehicle conditions, field supervisor management and passenger behavior.

2.3.3 Travel time reliability modelling

In transit, various basic factors have been identified as affecting running time and associated variability (Abkowitz and Engelstein, 1983; Bertini and El-Geneidy, 2004; El-Geneidy et al., 2011; Mazloumi et al., 2010; Strathman et al., 2002; Tétreault and El-Geneidy, 2010). Running time is the amount of time that it takes for a bus to travel from point A to point B excluding recovery time at time points. These factors include segment length, passenger activities (boardings and alightings), lift use, signalized intersections, number of scheduled stops, number of actual stops made, delay at the start, day of the week, time period of the day, service direction, weather conditions (rain and snow) and drivers experience. Accordingly, agencies implement strategies with expectations of improving service performance. Several researchers have investigated different strategies influencing running time and running time variability (Diab and El-Geneidy, 2012; El-Geneidy and Vijayakumar, 2011; El-Geneidy et al., 2006; Kimpel et al., 2005; Surprenant-Legault and El-Geneidy, 2011; Tétreault and El-Geneidy, 2010). These strategies include smart fare card collection system, reserved bus lanes, limited-stop bus services, stop consolidation, articulated buses and transit signal priority. Diab and El-Geneidy (2013) further investigated the impact of the implementation of various strategies on service variations.

To understand the effects of general factors on running time variability, researchers have developed multivariate linear regression models through different measures of service variation (Strathman et al., 1999; Yetiskul and Senbil, 2012). Many studies have shown that the segment length can adversely influence service reliability, as well as number of scheduled stops, number of signalized intersections, variation of passenger activities, lift use, delay at first stop, variation of drivers experience (El-Geneidy et al., 2011; Strathman et al., 2002). The influence of adverse weather on reliability is controversial. Hofmann and O'Mahony (2005) found that rain reduced service unreliability since congested traffic flows result in a low variation of running times. Tu et al. (2007a) found that rain conditions make the travel time less reliable than under normal weather conditions. Mazloumi et al. (2010) concluded that the average and standard deviation of amount of rain did not significantly influence reliability. Dummy variables of time period and land use were used as proxies of general traffic flow conditions. A constant was usually considered in the model to approximate the omitted factors effect. To get deeper insights into the causes of unreliability, separate regression models were developed for different scenarios, such as models for different times of day (Li et al., 2006; Mazloumi et al., 2010), and models for different spatial-temporal dimensions and service characteristics (Yetiskul and Senbil, 2012). *Although traffic condition is believed to be one of the main factors affecting service reliability, only a small number of researchers have looked explicitly at the influence of traffic flow on TTV.* Tu et al. (2007b) showed that TTV is hardly related to the variability of flow in the free flow and hyper-congested regime, whereas it is positively correlated with flow variability in the congested regime.

Instead of using traffic flow to represent traffic conditions, some studies have defined congestion levels using actual travel times and free flow travel times (Gilliam et al., 2008; Peer et al., 2012). The congestion level could be calculated as the ratio or difference between actual and free flow travel times. Such measure can be comparable between links that differ in lengths and free flow traffic conditions. Intuitively, vehicles operating on different types of road would experience different travel times due to infrastructure configurations, traffic compositions and signal delays. *However, no study has been found to differentiate these factors in modelling bus travel time reliability. On this basis, an analysis of bus travel time reliability on Australian urban roads was undertaken to validate the factors arising in the literature, to uncover other potential factors that might influence the travel time reliability of bus services, and the lessons to be learnt from bus travel time reliability effects in the Australian context.*

Given a set of regression models, most studies estimated the coefficients equation-by-equation using the classical OLS (Diab and El-Geneidy, 2013; El-Geneidy et al., 2011). *However, Martchouk et al. (2010) argued that formulating separate ordinary and the standard deviation of travel time would leave out potentially important cross-equation correlation that would result in inefficient parameter estimates.* To address this problem, they used SURE estimation that can account for the correlation between the unobserved shared characteristics on travel time and travel time variability, since they were measured at the same time period on the same link. The SURE was first proposed by Zellner (1962). Previously, the SURE estimation was used by Mannering (2007) to study the effects of interstate speed limits on driving speeds and by Miller et al. (2009) to study the average and standard deviation of vehicle speeds in night time construction zones. SURE was also applied in studying the post rehabilitation performance of pavements with random parameters (Anastasopoulos et al., 2012). Detailed information on SURE can refer to (Washington et al., 2011).

2.4 Travel time distribution fitting model

TTV decreases passengers' confidence on perceived reliability by impacting the duration of travel time components (e.g. waiting time, in-vehicle time and transfer time) and causing uncertainty in making travel decisions (e.g. route, mode and departure time choices) (van Oort, 2011). A reduction in TTV has been found to be as or more valuable than a reduction in average travel time itself (Bates et al., 2001). TTV can be viewed from different perspectives (Noland and Polak, 2002):

- Vehicle-to-vehicle variability is the variability between travel times experienced by different vehicles travelling over the same route and the same time. This is caused by different delay time at signals, conflicts with pedestrians, differences in driving style and so on.

- Period-to-period (within day) variability is the variability between travel time of vehicles travelling over the same route on different times of a day, it is mainly attributed to demand levels, traffic incidents, weather conditions and so on.
- Day-to-day variability is the variability of travel time of the same route made at the same time on different days. It is caused by fluctuations in travel demand, driving behaviour, road side activity, weather conditions, incidents and so on.

TTD describes the nature and pattern of TTV. Better understanding of the distribution of travel times is a prerequisite for analyzing reliability and exploring the causes of unreliability (Sumalee et al., 2013). Knowledge of TTD is also an essential input for other analyses, such as micro-simulation of transit systems, travel time predictions, discrete route choices and timetable design (Mazloumi et al., 2010).

Research on fitting continuous distributions to empirical travel time data for private vehicles began many decades ago. For simple cases, symmetric distribution (e.g. normal) was initially believed to be appropriate to characterize vehicle travel time. However, statistical analysis identified TTDs to be asymmetric and significantly skewed to the right (Richardson and Taylor, 1978). Nowadays, the lognormal model is the most recommended TTD for its good fit and relative simplicity (Clark and Watling, 2005; Hollander and Liu, 2008). Faouzi and Maurin (2007) claimed lognormal was an attractive distribution that could be justified from an equivalent theorem derived from the central limit theorem. The theorem expresses that any product of independent identically distributed random variables will be distributed according to a lognormal model. Other reported models include, Gamma (Polus, 1979), Loglogistic (Chu, 2010), Weibull (Al-Deek and Emam, 2006), Burr (Taylor and Susilawati, 2012) and Stable distribution (Fosgerau and Fukuda, 2012).

Compared to studies on private vehicle TTD, transit TTD research began relatively late and there have been limited studies, mainly due to the unavailability of extensive travel time data over time and space. Taylor (1982) showed that bus travel times that started at 8:15 am every day over 15 successive days followed a normal distribution. Jordan and Turnquist (1979) showed that bus morning peak running times had a skewed distribution and Gamma distribution was a precise fit. However, day-to-day distributions of bus TTV have received increasing attention since the emergence of automatic data collection systems. Uno et al. (2009) showed that bus running times (excluding delay times at stops) on arterial roadways followed a skewed distribution but the lognormal distribution was rejected by 5 of the 12 routes. Xue et al. (2011) showed that bus peak hour travel times of stop pairs in urban areas followed Loglogistic distribution. Kieu et al. (2014) analyzed TTV using transit signal priority data and recommended lognormal distribution as the best descriptor of bus travel time on urban roads. Table 2-2 shows the selected studies on TTD fitting using single mode distribution models.

Table 2-2 Selected studies on TTD fitting using single mode distribution model

Studies	Data Collecting Methods	Time Periods	Departure Time Window	Distribution
Taylor (1982)	Manually (bus)	Morning peak	0	Normal
Jordan and Turnquist (1979)	NA (bus)	Morning peak	NA	Gamma
		AM/PM peak	2 hour	Normal
Mazloumi et al. (2010)	AVL (bus)	Off-peak	2.5 hour	Lognormal
		Peak & off-peak	5 min	Normal
Al-Deek and Emam (2006)	Theoretical	NA	NA	Weibull
Susilawati et al. (2013)	GPS probe cars	NA	NA	Burr
Fosgerau and Fukuda (2012)	AVL (bus)	6am-10pm	1 min	Stable
May et al. (1989)	Survey (bus)	NA	NA	Normal
Faouzi and Maurin (2007)	ETC (car)	NA	NA	Lognormal
Polus (1979)	Manually	Afternoon	NA	Gamma
Emam and Ai-Deek (2006)	Loop detector (car)	Evening peak	5min	Lognormal
Chu (2010)	GPS probe truck	MP/off-peak/AP	5min	Log-logistic

Note: NA = not applicable; AVL = Automatic vehicle location; GPS = Global positioning system; ETC = Electronic toll collection.

However, the above mentioned empirical studies of modelling distribution of TTV tend to give inconclusive and inconsistent overall results. One of the important reasons for such inconsistency is the empirical data used. Travel time data collected in different temporal-spatial scales would have different characteristics due to specific service areas and different traffic conditions. It is a common limitation of empirical modelling studies that the results are largely determined by the data. However, two other essential factors can also contribute to the inconsistencies across different studies, namely, data aggregation and the evaluation approach.

For data aggregation, Vlahogianni and Karlaftis (2011) verified that the temporal aggregation of traffic data can alter the underlying stochastic characteristics of traffic performance. Li et al. (2006) demonstrated that car TTDs on freeways follow a lognormal distribution when the Departure Time Window (DTW) is large (e.g. 1 hour). In a narrower DTW, the distribution tends to be normal. Mazloumi et al. (2010) showed that bus TTD tends toward a normal distribution in short DTWs. When the DTW increases, a normal distribution is still a good fit for the peak period, while a lognormal distribution is more appropriate for the off-peak period. However, recent researchers have shown that even in short DTWs (5min), TTDs are still skewed (Chu, 2010; Emam and Ai-Deek, 2006). Susilawati et al. (2013) showed that shorter links tended to have bimodal distributions and these phenomenon are broken up when they are merged into a longer link. *The influence of spatiotemporal aggregation levels of travel time on TTD is needed to be examined.*

For the evaluation approach, the majority of empirical studies usually evaluated the performance of different models solely based on fitting accuracy under some specific scenario, such as weekday AM peak inbound route TTD. *This is thought to be insufficient to claim an appropriate distribution model for characterizing TTV without considering other evaluation criteria (e.g. robustness and explanatory power) and travel time under different scenarios.*

The assumption of the above mentioned research is that, for a given time period, travel times are predominantly determined by a unimodal distribution. However, travel time can be impacted by various factors and multiple states can exist for a specific time period. van Lint and van Zuylen (2005) identified four phases that yielded distinctively different shapes of the day-to-day TTD, on the basis of empirical observations within 5 min intervals. Recent studies have proved the superior performance of multimodal distributions in fitting TTDs compared to its alternative models. In addition, those models provide a connection between the shape of TTDs and the underlying travel time states (Barkley et al., 2012; Guo et al., 2010; S. Park et al., 2010; Sangjun et al., 2011; Susilawati et al., 2013). Conceptually, bus trips can also experience different ranges of travel times due to stochastic traffic flow en-route, random delay at intersections and delay time at stops. *However, no study has been found which fits bus TTD using multimodal models. It is important to take into consideration all possible alternative models for specification of a distribution that can most appropriately characterize day-to-day variability of bus travel times.*

2.5 Travel time distribution estimation methodology

Knowledge of travel times is crucial at many levels of transportation planning and management. Network-wide travel time information provides inputs for impact assessment of strategic and operational instruments. Information of link travel times can reveal problematic locations where targeted strategies can be introduced to improve service reliability performance. In addition, disseminating information on travel time reliability to system users is a key component of addressing urban mobility issues, since it can aid travelers to make informed travel decisions (Kuhn et al., 2013).

Methods on the estimation and prediction of travel times can be generally classified into two categories, namely analytical and data-driven. Analytical models explore the physical relationship between travel times and other traffic variables (traffic flow, occupancy, signal phase plans and etc.) (Geroliminis and Skabardonis, 2011). Data-driven models estimate travel times by combining potential factors that can be easily implemented and show a promising performance in practice (Fei et al., 2011). Among the most applied data-driven techniques are parametric and nonparametric regression (Chang et al., 2010), Kalman filter (Cathey and Dailey, 2003), machine learning (Chun-Hsin et al., 2004; Van Lint et al., 2005; Yu et al., 2011), Bayesian (van Hinsbergen et al., 2009) and hybrid methods (Van Lint, 2008). Most of these studies focus on the estimation of expected travel times, which can be used as an indication of congestion levels once compared with free flow travel times for planning applications (van Hinsbergen et al., 2009) or to aid users in making smart travel decisions (Brakewood et al., 2015).

Travel time variability caused by the inherent network randomness in the context of supply, demand and service performance is important to consider (Jenelius, 2012). Reduction of travel time variability decreases commuting stress and uncertainty of making travel decisions, e.g. departure time choices (Fosgerau and Engelson, 2011). Many statistical scalar indexes have been used to characterize variability, including variance, percentiles and confidence intervals (Jenelius and Koutsopoulos, 2013; Abbas Khosravi et al., 2011; A. Khosravi et al., 2011; Li and Rose, 2011; Pattanamekar et al., 2003). A common limitation is that the scalar indexes cannot fully characterize the stochastic features of travel times without an assumption on the shape of distribution. They can only provide incomplete information since the features of distributions may be missed, e.g. skewness and multimodality (van Lint et al., 2008).

Probability distributions contain maximum information that captures the stochastic characteristics of travel times (Du et al., 2012). Many studies on TTR have attempted to fit mathematical distributions to travel times at different network levels (Clark and Watling, 2005; Fosgerau and Fukuda, 2012; Hollander and Liu, 2008). For many applications, e.g. trip planning, trip travel time information is of more interest (Bhat and Sardesai, 2006). The trip TTDs can be derived or inferred using archived data of direct observations for the same origin and destination (OD) pairs under similar trip conditions, e.g. time period. One problem is that the archived database requires the full coverage of all OD pairs that travelers might take. Note that the OD pairs are not restricted to the major planning zones, but can be any locations in the network. Furthermore, with data from mobile sources, it is likely that for many OD pairs very few or no samples were observed directly. An effective approach for estimating trip TTDs between arbitrary OD pairs at arbitrary times is from individual link TTDs. Link travel times can be derived directly (e.g. transit AVL data) or estimated from the increasingly available but sparse opportunistic sensor data, e.g. vehicular GPS, Automatic Number Plate Recognition (ANPR), and mobile phone data (Hellinga et al., 2008; Jenelius and Koutsopoulos, 2015; Kazagli and Koutsopoulos, 2013; Rahmani and Koutsopoulos, 2013; Zheng and Van Zuylen, 2013). Given the known link TTDs, the challenge is how to estimate trip TTDs by taking into consideration of the spatiotemporal correlations between link travel times.

For the estimation of the mean and variance of trip travel times, the Space Time Autoregressive Integrated Moving Average (STARIMA) model was proposed by Pfeifer and Deutrch (1980). The model can capture the spatiotemporal relationships within observations. Cheng et al. (2014) proposed a dynamic spatial weight matrix and a localized STARIMA approach to capture the heterogeneous and dynamic correlations between urban link travel times. Jenelius and Koutsopoulos (2013) used GPS taxi data to estimate the distribution of urban link travel times as a function of link characteristics and trip conditions and incorporate link correlations based on a spatial moving average structure. Yeon et al. (2008) considered the temporal correlation between freeway link travel times in 1 minute interval based on a Markov chain methodology.

For the estimation of the probability distribution of trip travel times, Hollander and Liu (2008) introduced the concept of estimating the TTDs using repeated simulations, but the approach is computationally expensive. One approach for the estimation of trip TTDs from link TTDs is by assuming that the link travel times are independent for a given time period (T. Hunter et al., 2013; Westgate et al., 2013). This assumption however, can lead to a considerable underestimation of the magnitude of variability. Correlations across links have been captured by using a Markov chain methodology (Timothy Hunter et al., 2013; Ramezani and Geroliminis, 2012; Woodard et al., 2015) and a dynamic Bayesian network model (Hofleitner,Herring,Abbeel, et al., 2012; Hofleitner,Herring and Bayen, 2012). Ramezani and Geroliminis (2012) used a Markov chain approach to estimate arterial trip TTDs by capturing the spatial correlations using a Transition Probability Matrix (TPM) calibrated from historical data, and assuming that the travel times are independent conditional on link states. Similar to the Markov chain approach, Hofleitner et al. (2012a,b) assumed that each link can be in a congested or uncongested state with its own independent and normal TTDs. The transition between link states is modelled using a dynamic Bayesian network approach. Fei et al. (2011) proposed a Bayesian inference based dynamic linear model for predicting travel times along with their confidence interval, by combining a priori distribution and real time traffic information on freeways. Recently, Srinivasan et al. (2014) approximated trip and link TTDs assuming shifted lognormal distributions and estimated trip TTDs by combining correlated link distributions using a Moment Generating Function (MGF) approach (Fenton–Wilkinson's approach).

Markov chains facilitate the modelling of the probabilistic nature of link travel times and their spatiotemporal correlations through TPMs. *The main assumption is that the travel times conditional on link states along a Markov path are uncorrelated (Timothy Hunter et al., 2013). However, the assumption was found to be violated in our case study. In addition, existing Markov chain based models usually compute TPMs based on the number of observations conditional on adjacent link states, which constraints their ability to generalize to a wide range of applications.* The MGF method has been widely applied in the wireless communication field to approximate the probability density functions of the sum of correlated normal or lognormal random variables (Mehta et al., 2007). The main limitation of the MGF method is that it needs prior assumptions on both link and trip distributions, usually unimodal. It is reasonable to use the MGF approach to estimate the Markov path TTD since the link TTDs conditional on states tend to be unimodal.

2.6 Summary of main findings and research gaps

A detailed review was carried out to establish past researches in TTR and TTD areas, including TTR quantification, TTR modelling and TTD estimation. The main findings from the review are summarized in Table 2-3.

Table 2-3: Summary of the main findings from literature review

Main findings	Section
Travel time reliability definitions and measures	
o No common agreement on TTR definition. Two key features: consistency of service attributes and distinction between supply & demand sides; o A general pool of indicators is summarized based on which different sets of indicators can be selected under different situations; o Buffer time concept based indicators are appropriate to capture passengers' experienced reliability by using operational data.	2.2
Travel time variability and unreliability causes	
o TTR varies in time and space impacted by different sources of variations from demand and supply sides, as well as interactions between them. o Basic factors have been identified as affecting TTR using regression analysis, including planning, operational and environmental characteristics; o Strategies have been identified as affecting TTR, including fare collection, reserved bus lanes, limited-stop bus services, stop consolidation, articulated buses and transit signal priority.	2.3
Travel time distribution fitting model	
o The distribution fitting model can be classified into two categories, namely single-mode and mixture-modes distributions; o Inconsistent distribution models were reported and the spatiotemporal aggregation of travel times largely influence TTD; o Mixture-modes distribution provides a better fitting performance, as well as a connection with the underlying traffic states.	2.4
Travel time distribution estimation methodology	
o Methods for the estimation of trip TTD between origination-destination pairs using the increasingly available data from mobile sources are still evolving and rather limited; o Previous studies on trip travel time distribution (TTD) estimation used a Markov chain methodology and are based on a number of important assumptions: independent conditional on states and constant transition probabilities for a given time period.	2.5

From the overview of the literature presented here, the following research gaps have been identified:

- Quantification of TTR from the perspective of passengers using operational AVL data;
- Investigating the impacts of unreliability causes with data from both demand and supply sides by taking into consideration cross-equation correlations caused by unobserved factors;
- Specifying the most appropriate probability distributions of day-to-day travel times at different spatiotemporal aggregation levels and network levels.
- Developing a generic model for trip TTD prediction between arbitrary OD pairs from link TTDs with consideration of their heterogeneous correlations.

To address these research gaps, Chapter 4 analyzes the limitations of current TTR measures and proposes a buffer time concept based measure to approximate passengers' experienced TTR using AVL data. To investigate the impact of unreliability causes, three TTR models with respect to main concerns by passengers and operators are developed using a seemingly unrelated regression equations method in Chapter 5. To build generic models for TTR prediction, Chapter 6 proposes a novel evaluation approach and set of performance measures to specify the most appropriate distribution model for the day-to-day travel time variability at stop, link and route levels. Chapter 7 proposes a generalized Markov chain approach for estimating the probability distribution of trip travel times from link travel time distributions and takes into consideration correlations in time and space. The case studies use data integrated from Automatic Vehicle Location (AVL), smart card transactions, General Transit Feed Specification (GTFS), Brisbane Strategic Transport Model (BSTM) and Bureau of Meteorology (BoM) systems.

Chapter 3 Data Description and Processing

3.1 Introduction

The increasing availability of dedicated sensors, such as AVL and AFC, is transforming a once data-starved transport field into one of the most data-rich. While these data can provide detailed traffic and operational information, methods for its processing for decision making at all levels (planning and policy, operations, control) is still evolving. Such methods can be complex and time consuming, especially in the collection of reliable and comprehensive data of travel times and the associated contributory factors in a large temporal and spatial scale. The current research establishes a travel time related data warehouse for TTR and TTD study. The unique performance database developed here integrates data from different databases, including AVL, smart card, GTFS, BSTM, BoM and STREAMS.

The remainder of the chapter is organized as follows: In Section 3.2, the description of different databases and the integrated data warehouse are presented. To minimize the possibility of erroneous data, the major procedure for data processing (error and outlier detection) is then briefly discussed in Section 3.3. The routes used for case studies in this research are described in Section 3.4. Finally, Section 3.5 concludes the chapter.

3.2 Data integration

The data used in this research was provided by TransLink, a division of the Department of Transport and Main Roads (DTMR) in Queensland, Australia. TransLink is responsible for coordinating and integrating Queensland's overall passenger transport system, including bus, ferry and rail services. TransLink operates an integrated smart card system which allows the use of one ticket on multiple services. The archived data covered a six months period from November 1, 2012 to April 30, 2013 across Southeast Queensland (SEQ) area. To complement the dataset, data from different sources were integrated, including AVL, Go card (69,194,428 records), GTFS, BSTM and BOM and STREAMs data. Figure 3-1 shows the databases integrated for this research. The integrated database reproduces what has happened for each operation run at stops (arrival, departure and passenger activities) and on links (route characteristics, traffic condition and weather) along the service route. It can provide detailed information of travel times, passenger demand, and operational environment.

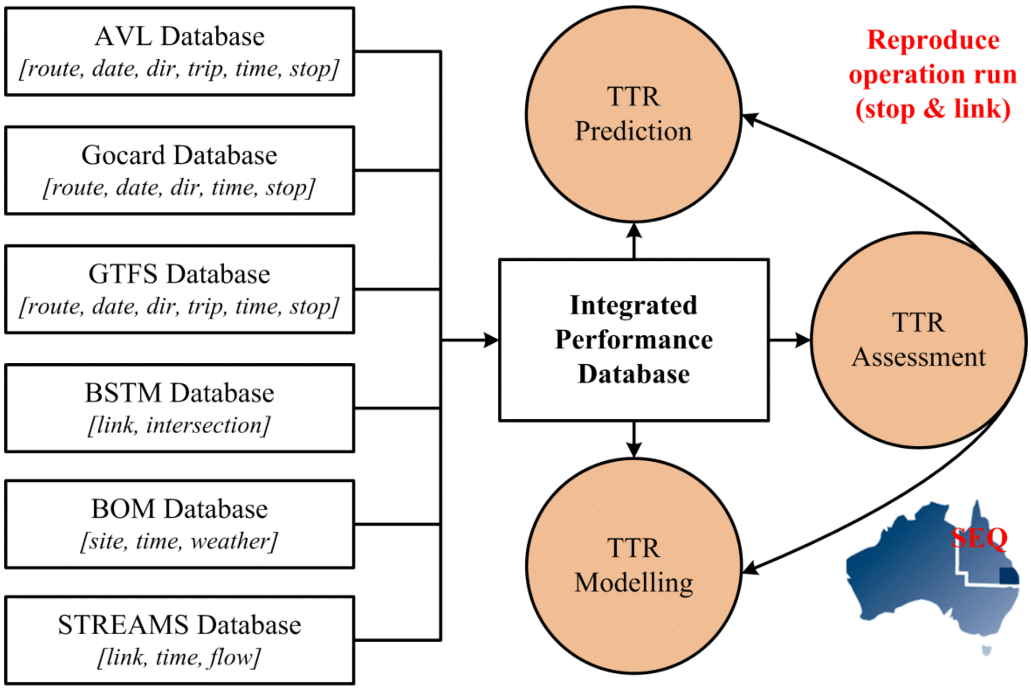

Figure 3-1 Overview of the integration scheme

The AVL system in Brisbane provides vehicle trip time information at a stop level, including operator, timestamp, route, direction, vehicle trip ID (unique), stop ID, vehicle arrival and departure time at a stop. The Go card system provides passengers trip transactions with both ‘Tap in’ and ‘Tap out’ information. The attributes for each transaction include operator, timestamp, route, run ID, direction, ticket number, Go card ID, boarding stop, alighting stop, boarding time, alighting time, passenger journey ID and passenger trip ID. A GTFS feed is composed of a series of text files and each file models a particular aspect of transit information: stops, routes, trips, and other schedule data (https://developers.google.com/transit/gtfs/). The BSTM is a four-step strategic transport model developed in the EMME/3 modelling platform. It provides information on road hierarchy, road type, lanes, road capacity, posted speed, and simulated volume over capacity (V/C). The weather data was obtained from BoM stations around SEQ, including rainfall, temperature, humidity, wind speed and wind direction, on a half hour basis. STREAM is an integrated intelligent transport system developed by Transmax in Australia. The link measure list in STREAM system provides information on link ID, timestamp, occupancy, level of service, flow. After examining the coverage of STREAM sensors, it is not used for the case study routes presented in Section 3.4.

In practice, the stop IDs are not necessarily consistent across different operators. The Go Card dataset does not include the latitude and longitude of the stops (x–y coordinates), nor do they match the region-wide SEQ transit schedule published by Translink. A stop-matching heuristic was developed and applied to the Go Card dataset to match the recorded stop IDs in the region-wide GTFS network. This heuristic is briefly presented in ‘‘Stop matching heuristic’’ section in the co-authored publication (Nassir et al., 2015). Finally, a mapping dictionary is generated to match stops across different systems.

Figure 3-2 shows a snapshot of the integrated outcomes. The integration procedure is implemented using MySql, C++ and Matlab software. For each record in the dataset, it possesses the trip information (Date, direction, link, scheduled departure time from the first stop), operational information (actual arrival time & departure time from AVL data, actual arrival time & departure time from Go card data, actual stop served), demand information (boarding, alighting and passenger load), and environmental information (route characteristics, length, number of lanes, speed limit, signals, land use, stop types and rainfall).

Trip information				Operation information							Demand information				Environmental information									
DATE	DIR	SCHDEP_RBI	LINK	SCHDEP_LBI	ACTARR_AVI	ACTDEP_AV	ACTARR_GI	ACTDEP_GI	SCHSTO	ACTSTO	BOARI	ALIGHT	LOAI	ROUTETY	LENGTH	LANE	SPEEDLI	SIGNAI	ROUNDA	OTHERIN	STOPNU	STATION	LANDUS	RAIN
20130201	1	52100	1_2	52100	52135	52135			0	0	0	0	0	6	0.5383	1	50	0	0	3	0	0	4	8.2
20130201	1	52100	2_3	52200	52234	52254	52238	52240	1	1	3	0	0	4	0.7633	1	60	0	0	7	1	0	4	8.2
20130201	1	52100	3_5	52400	52410	52435	52408	52411	1	1	3	0	3	4	0.7675	1	60	0	0	8	1	0	3	8.2
20130201	1	52100	5_6	52600	52537	52555	52535	52535	1	1	0	1	6	4	0.472	1	60	1	0	2	1	0	4	8.2
20130201	1	52100	6_7	52800	52649	52713	52647	52654	1	1	3	0	5	4	0.8855	1	60	6	0	1	1	0	2	8.2
20130201	1	52100	7_8	53100	53014	53046	53012	53015	1	1	1	1	8	5	0.7809	2	50	3	0	1	0	1	1	8.2
20130201	1	52100	8_9	53300	53222	53324	53220	53241	1	1	0	7	8	4	0.5556	1	50	3	0	0	1	0	1	8.2
20130201	1	52100	9_10	53500	53611	53630	53609	53615	1	1	0	2	1	1	1.2299	4	60	4	0	6	1	0	1	8.2
20130201	1	52100	10_11	54000	53951	54011	53949	53956	1	1	0	2	0	1	0.6761	4	60	4	0	3	1	0	5	8.2
20130201	1	52100	11_12	54300	54223	54242	54221	54221	1	1	0	1	0	1	0.5804	4	60	3	0	2	1	0	5	8.2
20130201	1	52100	12_13	54600	54433	54433			1	0	0	0	0	6	0.697	1	50	0	0	5	0	0	5	8.2
20130201	1	53600	1_2	53600	53627	53627			0	0	0	0	0	6	0.5383	1	50	0	0	3	0	0	4	8.2
20130201	1	53600	2_3	53700	53715	53730	53713	53713	1	1	1	0	1	4	0.7633	1	60	0	0	7	1	0	4	8.2
20130201	1	53600	3_5	53900	53829	53829	53827	53827	1	0	0	0	2	4	0.7675	1	60	0	0	8	0	0	3	8.2
20130201	1	53600	5_6	54100	53920	53920	53918	53918	1	0	0	0	2	4	0.472	1	60	1	0	2	0	0	4	8.2
20130201	1	53600	6_7	54300	54011	54143	54009	54009	1	1	1	0	2	4	0.8855	1	60	6	0	1	1	0	2	8.2
20130201	1	53600	7_8	54600	54402	54558	54400	54490	1	1	4	0	3	5	0.7809	2	50	3	0	1	0	1	1	8.2
20130201	1	53600	8_9	54800	54842	54900	54840	54840	1	1	0	1	7	4	0.5556	1	50	3	0	0	1	0	1	8.2
20130201	1	53600	9_10	55000	55101	55122	55099	55105	1	1	0	2	6	1	1.2299	4	60	4	0	6	1	0	1	8.2
20130201	1	53600	10_11	55500	55429	55446	55427	55427	1	1	1	0	4	1	0.6761	4	60	4	0	3	1	0	5	8.2
20130201	1	53600	11_12	55800	55713	55728	55711	55711	1	1	0	1	5	1	0.5804	4	60	3	0	2	1	0	5	8.2
20130201	1	53600	12_13	60100	55928	55928			1	0	0	0	4	6	0.697	1	50	0	0	5	0	0	5	8.2
20130201	1	60600	1_2	60600	60619	60619			0	0	0	0	0	6	0.5383	1	50	0	0	3	0	0	4	8.2
20130201	1	60600	2_3	60700	60715	60732	60713	60713	1	1	1	0	0	4	0.7633	1	60	0	0	7	1	0	4	8.2
20130201	1	60600	3_5	60900	60836	60855	60834	60834	1	1	1	0	1	4	0.7675	1	60	0	0	8	1	0	3	8.2
20130201	1	60600	5_6	61100	60957	61015	60955	60955	1	1	1	0	2	4	0.472	1	60	1	0	2	1	0	4	8.2
20130201	1	60600	6_7	61300	61122	61151	61120	61120	1	1	1	0	3	4	0.8855	1	60	6	0	1	1	0	2	8.2
20130201	1	60600	7_8	61600	61424	61445	61422	61422	1	1	0	1	4	5	0.7809	2	50	3	0	1	0	1	1	8.2
20130201	1	60600	8_9	61800	61705	61819	61703	61711	1	1	2	4	3	4	0.5556	1	50	3	0	0	1	0	1	8.2
20130201	1	60600	9_10	62000	61926	61926			1	0	0	0	1	1	1.2299	4	60	4	0	6	0	0	1	8.2
20130201	1	60600	10_11	62500	62408	62432	62406	62406	1	1	0	1	1	1	0.6761	4	60	4	0	3	1	0	5	8.2
20130201	1	60600	11_12	62800	62602	62649	62600	62600	1	1	0	1	0	1	0.5804	4	60	3	0	2	1	0	5	8.2
20130201	1	60600	12_13	63100	62744	62805	62742	62742	1	1	0	1	0	6	0.697	1	50	0	0	5	1	0	5	8.2
20130201	1	61600	1_2	61600	61642	61642			0	0	0	0	0	6	0.5383	1	50	0	0	3	0	0	4	8.2
20130201	1	61600	2_3	61700	61741	61823	61739	61739	1	0	0	0	0	4	0.7633	1	60	0	0	7	0	0	4	8.2
20130201	1	61600	3_5	61900	61942	62004	61940	61940	1	0	0	0	0	4	0.7675	1	60	0	0	8	0	0	3	8.2
20130201	1	61600	5_6	62100	62118	62118			1	0	0	0	0	4	0.472	1	60	1	0	2	0	0	4	8.2
20130201	1	61600	6_7	62300	62221	62245	62219	62219	1	0	0	0	0	4	0.8855	1	60	6	0	1	0	0	2	8.2
20130201	1	61600	7_8	62600	62547	62623	62545	62545	1	0	0	0	0	5	0.7809	2	50	3	0	1	0	0	1	8.2

Figure 3-2 Snapshot of the integrated data record

3.3 Data processing

Cleaning and refining the data are important steps in data processing. Raw data usually contain erroneous records caused by system failure or human faulty operation. The archived Go card data were screened to minimize the possibility of erroneous data by setting different rules and using schedule information from GTFS system. The results of the cleaned data indicated that 17% of records were excluded due to checking and fixing of erroneous data with different types of errors. These errors are summarized and described in Table 3-1.

The Median Absolute Deviation (MAD) technique was applied for outlier identification, that is, extremely long travel times. An item sample was considered as an outlier if it was outside the range of the lower and upper bound values determined by the MAD 3-delta criteria (Pearson, 2002). Figure 3-3 displays the cleaning results for weekday inbound travel times. The data were aggregated in 15 minutes interval. The MAD cleaning technique is promising with 3.2% outliers identified.

Table 3-1: Description of Go card transaction errors

Error types	Description (causes)
System failure	No boarding information, and no alighting information (not known)
Go card reload	An additional transaction record for the same passenger trip (top up go card in a vehicle)
Extremely large interval transaction	The difference between the alighting time and boarding time for a transaction is larger than 2 hours or across several days (forget to touch off from the last trip)
Ticket evasion	Boarding stop equals to alighting stop (tap in front door and tap again in the back door)
Driver faulty operation	The service direction in the transaction is wrong, e.g. the sequence of the inbound stops are actually the outbound stops (driver forget to change the route display information)
Abnormal stop	The boarding or alighting stops are not in the stop list of the recorded service route (the vehicle does not stop at the designated stop due to bus bunching)

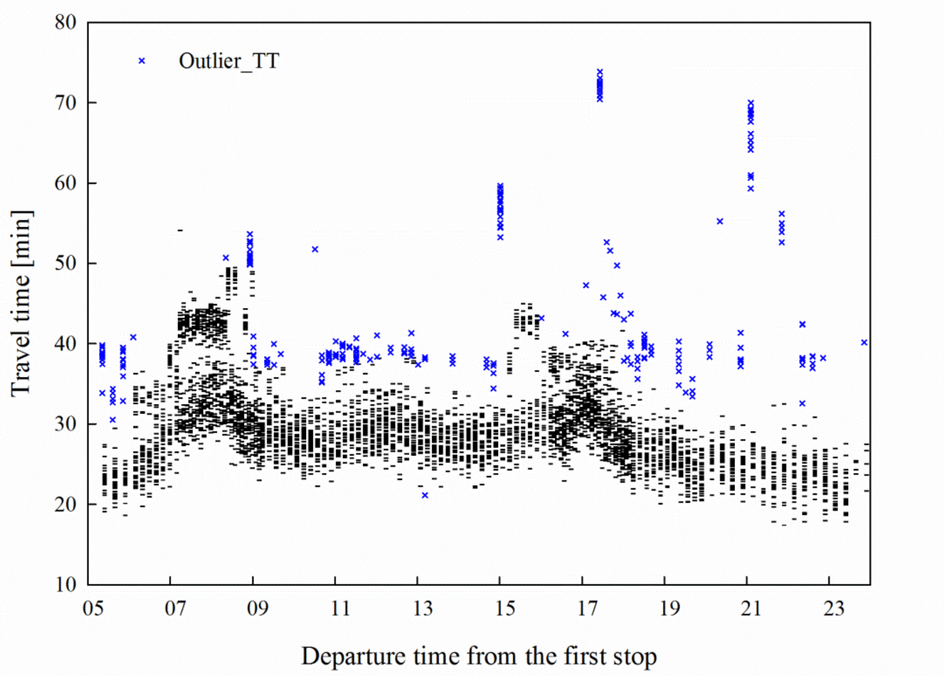

Figure 3-3: Data cleaning results and outliers identified

3.4 Case study area

The ideal data for empirical study would have large extent coverage of services that operating in different times and spaces. Translink implemented AVL system for only a limited number of bus routes (pilot test). Two bus routes equipped with AVL systems are used for case studies. The data used covers a six months' period with service operating from 5:30 to 23:30 every day. The two routes, which are shown in Figure 3-4, present diverse operating environments.

Route 60 is a cross-city route (mixed with local traffic) servicing two suburbs, West End and Fortitude Valley, as well as the CBD area of Brisbane. It operates along an arterial route of length 7.8 km and has 12 scheduled stops. It is one of the highest frequency bus services in Brisbane. It runs every 5 minutes between 7 and 9am and 4 and 6pm on weekdays and every 10 to 15 minutes during all other hours of operation. It operates from 5.30am until 11.30pm on Sunday to Thursday, and 24 hours on Friday and Saturday.

Route 555 is a radial route servicing Upper Mount Gravatt, Eight Mile Plains, Springwood and Logan as well as Brisbane City. For the inbound service to the City as shown in Figure 1, it operates first on the Pacific Motorway (mixed with local traffic) from Loganholme station to Eight Mile Plains station, then on the South East Busway (bus-only corridor) to the Cultural Centre station and continues to the City on an arterial road (mixed with local traffic). The route is 31km long and has 12 scheduled stops along the route. It runs every 15 minutes for services before 8pm on weekdays and Saturdays and before 6pm on Sundays and public holidays, and every 20 to 30 minutes for other hours.

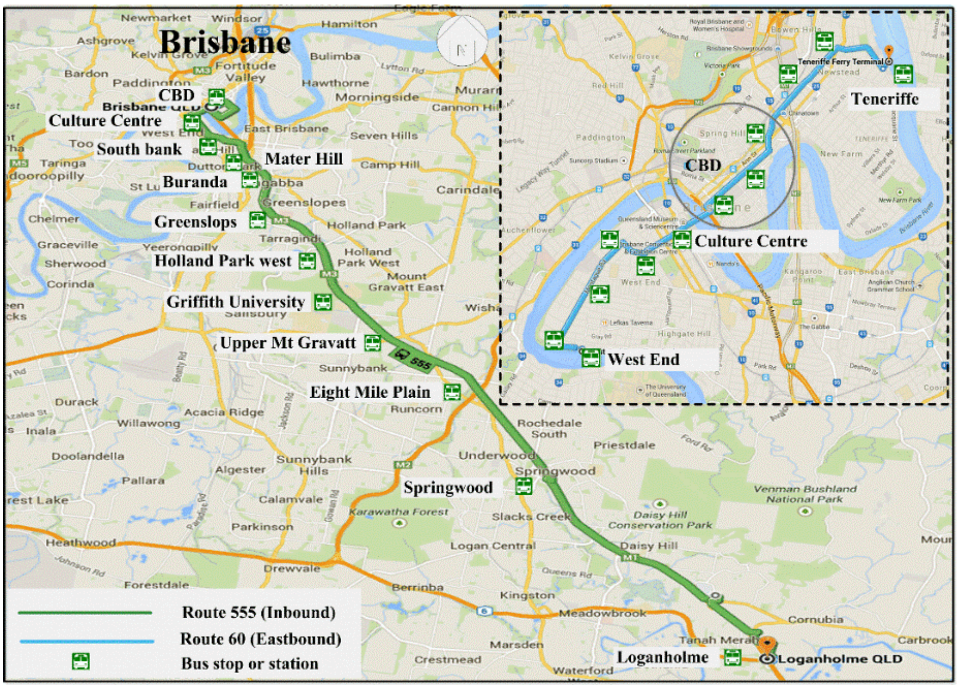

Figure 3-4 The used two transit routes in Brisbane, Australia

3.5 Summary

The research has developed a unique integrated dataset from different systems, including AVL, Go card, GTFS, BSTM, BOM and STREAMS. The integrated data warehouse provides detailed information on supply and demand information, as well as the associated environmental information. In this research, a stop matching heuristic algorithm has been put forward to build a stop mapping dictionary across different systems. The procedure to process the raw data is briefly discussed, including erroneous Go card transaction detection and abnormal travel time outlier identification. Two transit routes are used as case studies for this research mainly considering the diverse operating environments and data availability. These two routes cover diverse operating environments, including CBD area, residential area, major attraction area, suburban road, arterial road, motorway, and exclusive busway.

Chapter 4 Travel Time Reliability Quantification

4.1 Introduction

The concern with the impacts of reliability on operation efficiency (operators) and service effectiveness (passengers) brings about the need to identify and develop meaningful and consistent reliability measures. The automatic collection techniques, e.g. AVL and AFC, facilitate the gathering of enormous quantity and variety of spatial and temporal operational data that holds substantial promise for TTR analysis in a deep level. Leveraging on the AVL data, an investigation of the assessment performance of existing TTR measures is performed. Buffer time measures are believed to be appropriate to approximate passengers' experienced reliability. On this basis, the research proposes a set of TTR measures from the perspective of passengers using the operational AVL data considering different perceptions of TTR under different traffic states. The research findings are reported in a journal paper published in Ma, Ferreira, and Mesbah. (2014).

The remainder of the Chapter is structured as follows: In Section 4.2, two issues with regard to buffer time estimation are discussed, namely, performance disaggregation and capturing passengers' perspectives on reliability, followed by the detailed methodologies to address these issues presented in Sections 4.3 and 4.4, respectively. Case studies using both empirical data and numerical experiment are implemented in Section 4.5. The results show that the proposed reliability measure is capable of quantifying TTR consistently, while the conventional ones may provide inconsistent assessments. Then the potential applications of the TTR measure in reliability improvement and trip planning are briefly discussed in Section 4.6. Finally, Section 4.7 summarizes the main conclusions and highlights future researches.

4.2 Buffer time and its estimation

From literature review, there is no consensus on which attribute is capable of appropriately characterizing service reliability due to the heterogeneity of stakeholders' preferences and perceptions. Many studies have highlighted the importance to characterize TTR from the perspective of passengers (Cheng and Tsai, 2014; Hu and Jen, 2006). In this regard, buffer time measures are assessed to be more promising than other TTR measures (Currie et al., 2012) in two aspects: conceptually it can capture the influence of service unreliability on passengers travel decisions (Abkowitz et al., 1978); and mathematically, percentile-based buffer time is an indicator of compactness of TTD (Pu, 2011).

4.2.1 Passengers perspective on reliability

For a complete journey, excluding access time from the origination and egress time to the destination, a passenger is concerned of waiting time at the first stop, in-vehicle time during the trip and transfer time between different trips (Ceder, 2007). Generally, the unreliability can impact the duration and predictability of travel time which ultimately influence passengers' trip planning behaviours (van Oort, 2011). Due to the vehicle travel time variability, passengers may experience longer or shorter journey times which lead to early or late arrivals at their destination. These can be quantified using a measure of variability, such as standard deviation of travel time. In addition, unreliable service brings uncertainty to travel time which hinders passengers' ability to make optimal travel decisions to minimize disutility (Uniman et al., 2010).

For an infrequent service, passengers tend to arrive as close to their desired service departure time without missing the expected vehicle at the first stop. For a frequent service, passengers would be more interested in choosing a departure time that can minimize their late arrivals. If a passenger with a desired arrival time T_{des} travels in an ideal transport system without any variability, the departure time should be exactly the desired arrival time T_{des} minus the expected travel time of the trip TT_{exp}. However, in reality, a passenger can experience a stochastic arrival time distribution with non-zero probability of a late arrival for each departure time as shown in Figure 4-1.

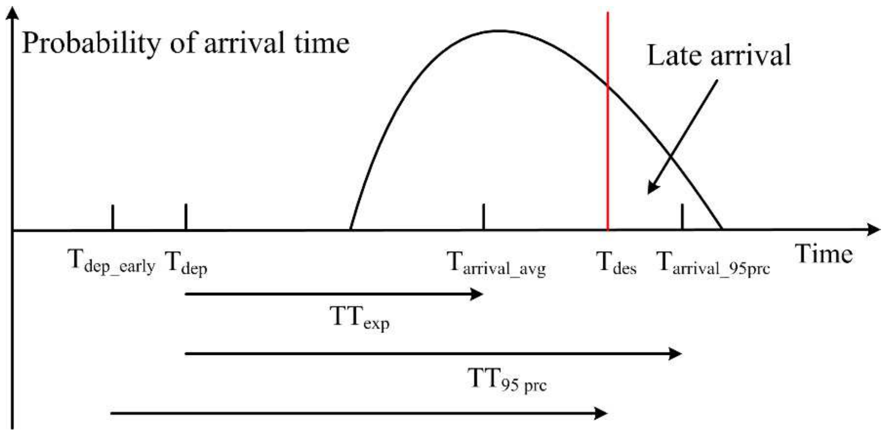

Figure 4-1: Journey departure decision and arrival time distribution

If a passenger highly valued an on-time arrival, he/she should shift the departure time earlier to reduce the probability of a late arrival. For example, to guarantee a late arrival probability no more than 5%, the passenger should leave before T_{dep_early} which is the difference between the desired arrival time T_{des} and 95th percentile travel time of a trip TT_{95prc}. This additional time budgeted by a traveller to increase the probability of an on-time arrival is regarded as buffer time. Intuitively, as the service reliability decreases, a passenger needs to budget more buffer time to avoid a late arrival.

Generally, buffer time is defined as the 95th percentile travel time minus average or median travel time which indicates the additional time that a passenger should budget to guarantee an on-time arrival under a given probability (Lomax et al., 2003). Although the term ‘buffer time’ usually denotes buffer travel time, it can be recognized as a concept of extreme-value based reliability evaluation measure, which can be applied manifoldly to: (a) buffer waiting time to indicate excess waiting time needed to catch an expected bus (Furth and Muller, 2007); (b) buffer transfer time to indicate additional time required to avoid missing connections (Goverde, 1999); and (c) buffer travel time to indicate extra time necessary for an on-time arrival.

4.2.2 Buffer time estimation

Two causes may reduce the usefulness of the existing buffer time measures in the context of performance evaluation when directly applying it to mixture distributions. One reason is that two different mixture TTDs could have the same buffer time value as shown in Figure 4-2. It shows that the service A and B have different probability density functions (PDFs), thus different reliability performance. However, they have exactly the same buffer time value (4.7 minutes) calculated using cumulative distribution functions (CDFs). This is conceptually unreasonable in reality. Further, any travel time samples with the same 95th percentile and median travel times would have the same buffer time value. It indicates by applying the buffer time measure directly on the source travel time profile could also lead to an inconsistent reliability assessment when mixture distributions exist.

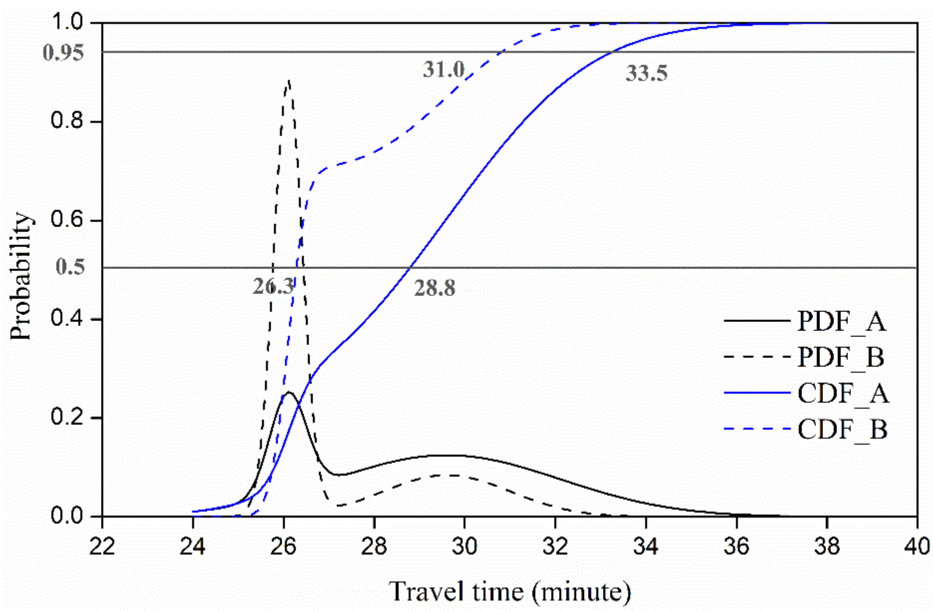

Figure 4-2: Travel time samples with mixture distributions

The other reason is, by considering the TTD as a whole, the buffer time measure could hide the sources of unobserved reliability changes, thus making it hard for the identification of unreliability factors. Many studies have claimed that TTD can be classified into recurrent and non-recurrent states for a time period (Barkley et al., 2012; Susilawati et al., 2013).

In addition, passengers would experience different arrival time distributions and thus have different departure decisions under different occasions. Conceptually, bus service state can be classified into three types, namely, fast, slow and non-recurrent service states. The former two states can be aggregated together as the recurrent state. Bus travel time under different states can have different characteristics. In a recurrent state, travel time is largely determined by traffic flow fluctuations and passenger demand characteristics. The difference between the fast and slow service states is mainly caused by stop delays (e.g. red light and queuing) and intersection delays (e.g. serving passengers, bus bunching, merging to the traffic flow). A vehicle in a fast service state may experience less intersection delays and stop delays than one in a slow service state. In this case, the passenger taking a fast service would plan less 'buffer time' than one taking a slow service, or even plan no 'buffer time' if the 95th percentile arrival time $TT_{arrival\_95prc}$ under a fast service state is already smaller than the desired arrival time T_{des} (Figure 4-1). In the non-recurrent state, the corresponding travel time unreliability will be higher than that in the recurrent state. In addition, the non-recurrent state can be further broke down to a more refined sub-set influenced by different factors, such as incidents, weather and extreme events (Barkley et al., 2012). In this case, passengers may not consider the non-recurrent state travel time in trip planning since the non-recurrent traffic condition is rare and cannot be predicted in practice.

In summary, buffer time measure can evaluate the reliability experienced by passengers in the context of departure planning using operational data. However, directly applying buffer time measure to a whole distribution of travel times may give inconsistent reliability assessments, hide unreliability causes and can not effectively capture passengers' departure behaviours. It is reasonable to develop a 'buffer time' concept based measure that can assess TTR under different states separately. Two issues will be addressed in the following sections, namely performance disaggregation and capturing passengers' perspectives on reliability.

4.3 Performance disaggregation

Based on the discussions, the primary task is to disaggregate the overall travel time performance for a trip origination-destination (OD) pair in a specific time period across different days into different states (or categories). A Gaussian Mixture Models (GMM) approach is applied to disaggregate the performance data.

Mixture models provide flexibility in modelling the underlying characteristics of the data. GMM is a special type of mixture models where the component distribution is Gaussian and is used as a clustering method that is more appropriate than *k*-means clustering, especially when clusters have different sizes and correlation within them (Yildirimoglu and Geroliminis, 2013). A GMM model with *K* components has the following PDF:

$$p\left(\mathbf{y} \mid \mathbf{\Theta}\right) = \sum_{k=1}^{K} w_k \mathcal{N}\left(\mathbf{y} \mid \mu_k, \Sigma_k\right) \tag{4.1}$$

where,

$\mathbf{y}$ = a vector of continuous-valued observations;

$\mathcal{N}\left(\mathbf{y} \mid \mu_k, \Sigma_k\right)$ = a Gaussian probability density for component k;

$\mathbf{\Theta} = \left(\boldsymbol{\omega}, \boldsymbol{\mu}, \boldsymbol{\Sigma}\right)$ the set of GMM parameters;

$\boldsymbol{\omega} = \left(w_1, \cdots, w_K\right)$ a vector of mixture coefficients such that $w_k \geq 0$ and $\sum w_k = 1$;

$\boldsymbol{\mu} = \left(\mu_\mathbf{1}, \cdots, \mu_K\right)$ parameter vector of the mean of Gaussian distribution;

$\boldsymbol{\Sigma} = \left(\Sigma_\mathbf{1}, \cdots, \Sigma_K\right)$ parameter vector of the variance of Gaussian distribution.

By changing the component distributions (e.g. Normal, Lognormal or Gamma) and the mixture coefficients, a mixture models is flexible to approximate a large range of different TTDs. In practice, the mixture coefficient w_k can be interpreted as the probability that a vehicle encounters state k (e.g. congested) and the component distribution indicates the TTD under such state. These connections provide an opportunity for analysing passengers experienced reliability under different states separately.

The GMM model can be estimated using the Expectation-maximization (EM) algorithm (Kazagli and Koutsopoulos, 2013; Yildirimoglu and Geroliminis, 2013). Considering a sequence of N training observation vectors $\mathbf{Y} = \left\{\mathbf{y_1}, \ldots, \mathbf{y_N}\right\}$ and assuming independence between vectors, the GMM likelihood can be written as:

$$L\left(\mathbf{Y} \mid \mathbf{\Theta}\right) = \prod_{n=1}^{N} \left[\sum_{k=1}^{K} w_k \mathcal{N}\left(\mathbf{y}_n \mid \mu_k, \Sigma_k\right)\right] \tag{4.2}$$

At each EM iteration, the parameters are updated to guarantee a monotonic increase in the model's likelihood value:

$$\overline{w}_k = \frac{1}{N} \sum_{n=1}^{N} p\left(k \mid \mathbf{y}_n, \mathbf{\Theta}\right) \tag{4.3}$$

$$\overline{\mu}_k = \sum_{n=1}^{N} p\left(k \mid \mathbf{y}_n, \mathbf{\Theta}\right) \mathbf{y}_n \Big/ \sum_{n=1}^{N} p\left(k \mid \mathbf{y}_n, \mathbf{\Theta}\right) \tag{4.4}$$

$$\overline{\Sigma}_k = \sum_{n=1}^{N} p\left(k \mid \mathbf{y}_n, \mathbf{\Theta}\right) \mathbf{y}_n^2 \Big/ \sum_{n=1}^{N} p\left(k \mid \mathbf{y}_n, \mathbf{\Theta}\right) - \overline{\mu}_k^{\,2} \tag{4.5}$$

where the parameter set $\mathbf{\Theta}$ is updated iteratively until the likelihood converges.

The probability of an observation y_i belonging to state k can be calculated as:

$$Prob\left(component\, k \mid y_i\right) = w_k \mathcal{N}\left(y_i \mid \mu_k, \Sigma_k\right) \Big/ \sum_{j=1}^{K} w_j \mathcal{N}\left(y_i \mid \mu_j, \Sigma_j\right) \tag{4.6}$$

An observation is assigned to the cluster having the largest $Prob\left(component\, k \mid y_i\right)$.

4.4 Measurement development

The concepts of service variability and reliability are different. Service variability is defined as the distribution of output values for the supply side, such as vehicle trip time, departure time and headways. It indicates the objective service performance provided by operators. Service reliability is defined as the degree of matching between the supplied service and the expected service. A service with high variability does not necessarily lead to poor reliability experienced by passengers. Given an expected trip travel time (e.g. 40 min), a passenger can perceive an early arrival time with large variability (e.g. range 34-39 min) as more reliable than a late arrival time with small variability (e.g. range 42-45 min). It is reasonable to capture passengers' different perspectives on reliability under different conditions given a certain expectation.

4.4.1 Reliability buffer time

Let $F_k(t)$ denote the CDF of the PDF $f_k(t)$ under state k. Let $F_k^{-1}(x)$ denote the inverse distribution function of $F_k(t)$. Then, the x^{th} percentile travel time TT_k^x under service state k is:

$$TT_k^x = F_k^{-1}(x) \tag{4.7}$$

Following the idea of the traditional buffer time definition, the reliability buffer time RBT_k for service state k is defined as the difference between M^{th} percentile travel time TT_k^{M} under service state k and N^{th} percentile travel time $\mathrm{TT}_{\mathrm{typical}}^{\mathrm{N}}$ under a typical condition. By the nature of the buffer time concept (additional time budgeted for a trip), the value of buffer time should be no less than zero. The RBT can be formulated as follows:

$$\mathrm{RBT}_k = \begin{cases} TT_k^{\mathrm{M}} - \mathrm{TT}_{\mathrm{typical}}^{\mathrm{N}}, \textit{if } TT_k^{\mathrm{M}} \geq \mathrm{TT}_{\mathrm{typical}}^{\mathrm{N}} \\ 0, \mathrm{otherwise} \end{cases} \tag{4.8}$$

The recurrent service state is chosen as the typical condition instead of using the whole service states, because in transit, the direct expectation of a trip travel time comes from the timetable published by operators which is usually designed according to the average travel time under the recurrent service state. In addition, it is meaningless to incorporate the unpredictable incident-influenced non-recurrent state in modelling a service expectation from the perspective of passengers, even though they might experience extremely long travel times. The selection of *M* and *N* depends on the usage purpose and transit passengers' preferences.

Wakabayashi and Matsumoto (2012) presented a detailed performance analysis of different percentile-based reliability measures and their relationships. Usually, *M* and *N* are chosen as 95 and 50, respectively. The 95$^{\text{th}}$ percentile refers to a traveller can be late for a work one time a month without getting in too much trouble (Lomax et al., 2003). The 50$^{\text{th}}$ percentile relates to the typical travel time under a certain condition.

For a given trip OD pair in a specific time period, TT_k^{M} represents the service variability performance provided by the operators under state k, and $\mathrm{TT}_{\mathrm{typical}}^{\mathrm{N}}$ represents the travel time expectation for such service by passengers. Figure 4-3 illustrates the possible service states for a trip and the calculation of RBTs under different service states. The definition of RBT for different service states can be regarded as an approximation of passengers experienced buffer times under different situations. It can be interpreted like this, if a passenger experiences a slow service state for his/her trip, the budgeted buffer time required to guarantee a late arrival possibility less than 5% is the difference between 95th percentile travel time ($\mathrm{TT}_{\mathrm{slow}}^{95}$) and his/her expectation of the service ($\mathrm{TT}_{\mathrm{recurrent}}^{50}$). However, if a passenger experiences a fast service state that 95th percentile travel time ($\mathrm{TT}_{\mathrm{fast}}^{95}$) is still in his/her expectation ($\mathrm{TT}_{\mathrm{recurrent}}^{50}$), there is no need to budgeted any buffer time no matter how variable the service performance is under such state.

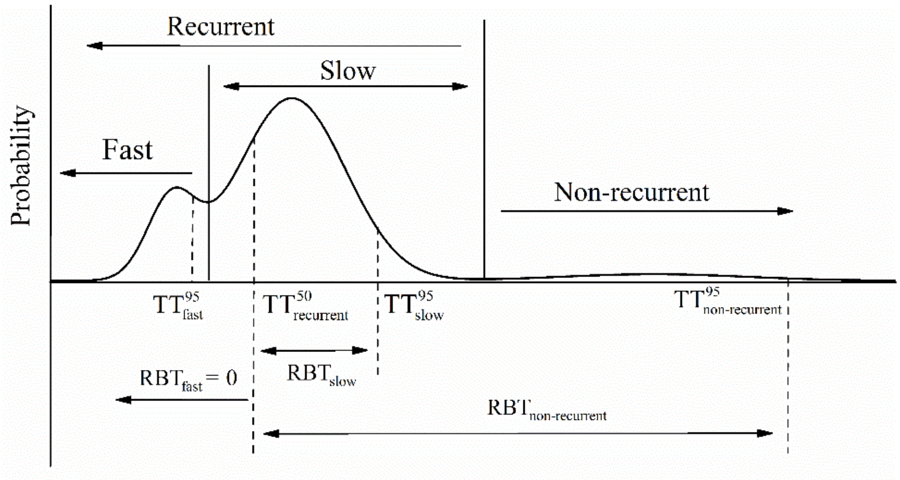

Figure 4-3: Illustration of possible service states and reliability buffer time

4.4.2 Expected reliability buffer time (operator)

As a service industry, transit operators are concern of providing satisfactory service to their passengers. Reliability measures are required to indicate current service reliability, identify causes of unreliability, assess different strategies effect on reliability and modify strategies to improve it. Based on different application objectives, a set of expected reliability buffer time measures (ERBT) are developed for operators by using different combinations of RBTs in Equation (4.8).

The OD-level ERBT (ERBT_{od}) is defined as the occurrence probabilities weighted RBT_k under different states for an OD trip pair, within a time period across different days.

$$\mathrm{ERBT}_{od} = \left(\sum\nolimits_{k=1}^{K} p_k \times \mathrm{RBT}_k\right)_{time\ period,days} \tag{4.9}$$

where,

p_k = the occurrence probability of the k^{th} service state.

The subscripts indicate different dimensions of aggregation including OD pairs, time period (e.g. morning peak, off peak or afternoon peak) and a span of time over many days (e.g. a 12 week-day sample). By changing the latter two dimensions allows for different levels of temporal analysis.

The Line-level ERBT (ERBT_{line}) spatially aggregates ERBT_{od} weighted by the passenger demand of an OD pair along a certain line during a studied time period.

$$\mathrm{ERBT}_{line} = \left(\sum_{od \in line} \left(d_{od} \times ERBT_{od} \right) \Big/ \sum_{od \in line} d_{od} \right)_{time\ period, days} \tag{4.10}$$

where,

d_{od} = the passenger demand of an OD pair along the concerned line.

The ERBT_{od} and ERBT_{line} measures indicate the average overall reliability performance for a given service for different temporal-spatial scales. The passengers' experienced RBTs for different states are aggregated together based on their contributions to the overall performance instead of treating them as equal. For example, the non-recurrent travel times can cause the highest RBT but if they rarely happen, the contribution of them to the overall service performance shouldn't be so much. In addition, as different factors contributing to different states, the RBTs in Equation 4.8 can be used to separate different factors influenced reliability apart which can make causes identification and strategies assessment more effectively. To implement fair reliability comparisons between two different services, the ERBT index (ERBTI) can be applied which is defined as the expected reliability buffer time divided by the expected travel time (e.g. travel time under a recurrent state).

4.4.3 Trip planning time (passenger)

Although the passengers experience different service states in their daily travels, it is hard for them to get an accurate and complete picture of the operational performance by themselves as shown in Figure 4-3. However, on-line applications of trip planner on websites and mobile phones provide a good way to convey such information to the public. Current trip planners provide a departure time calculated using average trip duration based on which a passenger can expect a low chance of on-time arrival. For a bus trip planning, passengers are concerned of deciding departure time to avoid late arrivals at their destinations and thus they are more interested in TTR than travel time itself. Considering travel mode choice and departure time planning, two types of times is interesting to a passenger, namely the average trip duration and the latest trip duration. These two measures can be used for mode choice and departure time planning, respectively.

The average trip duration (ATD) for a trip is defined as the 50th percentile travel time under the recurrent service state instead of the whole service states for a specific time period over different days. Non-recurrent service state is excluded because it is rare and unpredictable in reality, and including it would increase the ATD to a much high value, which is meaningless for a trip planning.

$$\mathrm{ATD}=\left(\mathrm{TT}_{\mathrm{recurrent}}^{50}\right)_{trip,\ time\ period,\ days} \tag{4.11}$$

where the subscript trip indicates a passenger travel from a boarding stop to a alighting stop along the same service route.

The latest trip duration (LTD) for a trip is defined as the 95th percentile travel time under the slow service state for a specific time period over different days. It indicates in 95% occasions, the vehicle would arrival at the destination using less than the LTD time. In other words, if a passenger plans a trip according to the LTD, he/she would encounter late arrival only once in a month (5% late arrival).

$$\mathrm{LTD}=\left(\mathrm{TT}_{\mathrm{slow}}^{95}\right)_{trip,\ time\ period,\ days} \tag{4.12}$$

4.5 Case study

4.5.1 Probability distribution fitting

Modelling travel time profile to a theoretical distribution is the prerequisite for the calculation of RBT and it can also provide the maximum information for reliability evaluation (Clark and Watling, 2005). Most importantly, by assigning travel time data points to different clusters based on distribution densities can disaggregate travel time data at a high-level of detail. Single and Mixture distribution models are tested to verify whether mixture states exist for peak period travel times.

4.5.1.1 Single mode distribution

The single distribution model assumes the travel time samples come from a single travel time state during a given time period. The Kolmogorov-Smirnov (K-S) and Anderson-Darling (A-D) were used to test the hypothesis that the AM-peak WD-IN travel time follows the potential theoretical distributions, including Burr, Exponential, Extreme value, Gamma, Log-normal, Logistic, Log-logistic, Normal and Weibull (Chu, 2010). The hypothesis test results are provided in Table 4-1. Parameters for the fitted travel time distribution are also provided and the goodness-of-fit is measured using Akaike Information Criterion (AIC) (Akaike, 1974). The test results illustrate that the TTD can come from any individual theoretical distribution presented here, except the Exponential model. However, the *p-values* for the accepted distributions are rather low and the largest p-value is only 0.241(Weibull). These indicate the limited ability of single distribution models in modelling the AM-peak WD-IN travel time data profile.

The AIC value indicates that the best distribution fitting model is Log-normal. And many studies have also claimed Log-normal as an appropriate travel time distribution model which can be justified from an equivalent theorem derived from central limit theorem (Faouzi and Maurin, 2007).

Therefore, the Log-normal model is selected as a representative of single model distribution for comparison purposes with the mixture models distribution.

Table 4-1: Summary of fitting performance of single mode distribution models

Single Models	K-S test (p-value) *	A-D test (p-value) *	Parameters[#]		Goodness-of-fit AIC
			Shape	Scale	
Burr	0.111 (0.089)	1.574 (0.160)	15.86/1.59	30.00	- 590
Exponential	0.563 (<0.05)	46.657 (<0.05)	28.89	N/A	- 1071
Extreme value	0.107 (0.110)	2.160 (0.075)	30.25	2.75	- 612
Gamma	0.111 (0.088)	1.364 (0.212)	118.19	0.24	- 585
Log-normal	0.110 (0.097)	**1.339 (0.220)**	3.36	0.09	**- 584**
Logistic	0.110 (0.096)	1.643 (0.146)	28.83	1.59	- 595
Log-logistic	0.111 (0.091)	1.625 (0.149)	3.359	0.055	- 593
Normal	0.112 (0.087)	1.372 (0.210)	28.89	2.68	- 587
Weibull	**0.091 (0.241)**	1.679 (0.139)	11.17	30.13	- 602

* P-value < 0.05 rejects the null hypothesis that the data come from the distribution
[#] The scale parameter indicates the degree of the spread for travel time distribution
The shape parameter indicates the shape and location of the travel time distribution

4.5.1.2 Mixture mode distribution

The source travel time profile is fitted using a two components GMM model. The parameters of the fitted distributions are shown in Table 4-2. From Figure 4-4, it can be seen that the mixture models is promising in capturing the bimodal characteristics of travel time distribution. The single model seems to have limited ability in tackling bimodal distribution. The Hartigan dip test confirms the existence of bimodal phenomenon in such distribution with p-value less than 0.05 (Hartigan and Hartigan, 1985). The goodness-of-fit AIC value in Table 4-2 also verifies the superiority of the GMM model when compared with the Log-normal model in modelling TTD. The first component of the GMM model can be regarded as the fast service state that encounters short stop delays and intersection delays and the second one is the slow service state that experiences long stop delays and intersection delays. The variance of the first component is much smaller than the second one due to the fact that the vehicles experiencing no stop and intersection delays would have much less variance than those experiencing stop and intersection delays.

Table 4-2: Parameters for the fitted single model and mixture models distributions

	Parameters			AIC
	Mixture coefficient	Shape	Scale	
Observed	N/A	N/A	N/A	N/A
GMM2	0.21	26.1	0.164	-542
	0.79	29.6	6.400	
Log-normal	N/A	3.360	0.090	-584

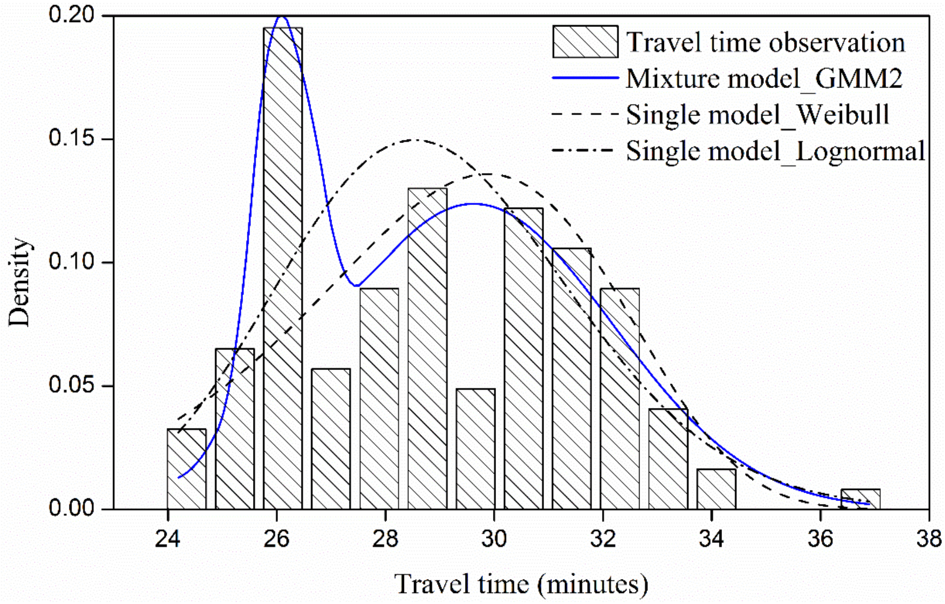

Figure 4-4: Travel time distribution fitting result using single and mixture models

4.5.2 Assessment performance comparison

The proposed ERBT is validated by comparing with the existing reliability measures using numerical travel time samples, including planning time index, buffer time index and reliability time index (Chu, 2010). The planning time index (PTI) is calculated as the 95th percentile travel time TT^{95} divided by average travel time TT^{avg}.

$$PTI = \frac{TT^{95}}{TT^{avg}} \times 100\% \tag{4.13}$$

The buffer time index (BTI) is calculated as the difference between the 95th percentile travel time and average travel time divided by average travel time.

$$BTI = \frac{TT^{95} - TT^{avg}}{TT^{avg}} \times 100\% \tag{4.14}$$

The reliability time index (RTI) is calculated as the difference between the 95th percentile travel time and median travel time TT^{50} divided by median travel time.

$$RTI = \frac{TT^{95} - TT^{50}}{TT^{50}} \times 100\% \tag{4.15}$$

Considering the fact that the BTI may be too conservative to incorporate random travel time fluctuations, the RTI is estimated by using median travel time instead of average travel time. To illustrate the current and proposed reliability measures of performance, different groups of travel time samples are generated using GMM models, with different parameters combination. To keep the central tendency the same for comparison, the parameters of means μ_1 and μ_2 for different groups are set equal to those of the empirical travel time samples.

The parameters of proportions p_1 and p_2 are set randomly in range [0, 1] with the constraint that $p_1 + p_2 = 1$. To illustrate the current measures limitations under some specific occasions, the parameters of sigmas σ_1 and σ_2 are calculated by solving the equation that $\mathrm{TT}^{95} - \mathrm{TT}^{50} = 4.4\,\mathrm{min}$. It indicates the fact that even with the same assessment result by the current measures, different shape of distributions can still exist, thus different actual reliability performance. Without loss of generality, five groups of travel time samples are presented to investigate different measures assessment ability. Table 4-3 shows the parameters for different groups travel time samples. The overall mean times, median times, and planning times are also calculated.

Table 4-3: Parameters for different groups travel time samples

Groups	mu		sigma		proportion		Mean time*	Median time*	Planning time*
	μ_1	μ_2	σ_1	σ_2	p_1	p_2			
Empirical	26.1	29.6	0.16	6.40	0.21	0.79	28.9	28.8	33.2
Group A	26.1	29.6	1.80	5.20	0.35	0.65	28.4	28.2	32.9
Group B	26.1	29.6	7.20	7.40	0.15	0.85	29.1	29.2	33.9
Group C	26.1	29.6	0.10	2.00	0.70	0.30	27.2	26.3	31.0
Group D	26.1	29.6	0.10	2.00	0.50	0.50	27.9	26.8	31.4
Group E	26.1	29.6	0.10	8.00	0.01	0.99	29.6	29.6	34.3

* Time unit is minutes

Figure 4-5 displays the CDFs and PDFs of different groups travel time samples. The order of the reliability performance can be identified by comparing the compactness of the distribution. That is, the more compact of a service travel time distribution, the more reliable of the service. The reliability performance order (from the best to the worst) is identified as Group C, Group D, Group A, Empirical, Group B and Group E.

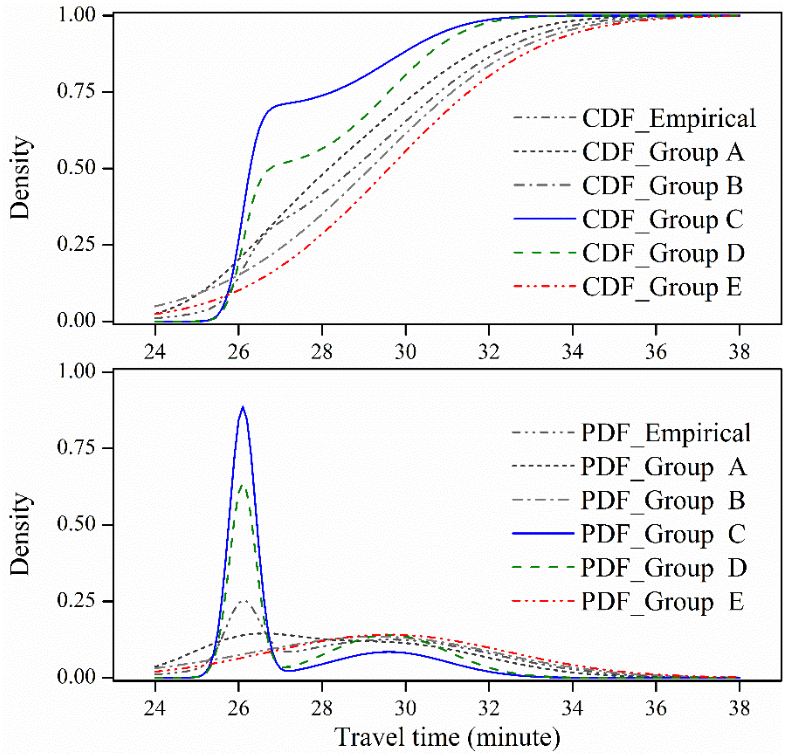

Figure 4-5: CDFs and PDFs for different groups of GMM travel time samples

Table 4-4 provides the reliability assessment results using PTI, BTI, RTI and ERBTI measures. It can be seen that the conventional PTI, BTI and RTI measures give inconsistent indications of reliability performances with the real ones identified from Figure 4-5. For example, the BTI measure indicates that Group D (0.125) has a better reliability than Group C (0.139), but oppositely, the latter group (blue line) should has a better reliability since its distribution is more compact than that of the former one (dark green line). In addition, the RT measure gives equally the same reliability time values (4.4 minutes) estimations for different groups travel time samples.

Table 4-4: Assessment results for different groups travel time samples

Groups	PTI (PT*)	BTI (BT*)	RTI (RT*)	ERBTI (ERBT)
Empirical	1.148 (33.2)	0.148 (4.30)	0.153 (4.40)	0.373 (10.7)
Group A	1.158 (32.9)	0.158 (4.50)	0.156 (4.40)	0.291 (8.20)
Group B	1.164 (33.9)	0.165 (4.80)	0.151 (4.40)	0.500 (14.6)
Group C	1.139 (31.0)	0.139 (3.80)	0.165 (4.40)	0.083 (2.19)
Group D	1.125 (31.4)	0.125 (3.50)	0.164 (4.40)	0.127 (3.40)
Group E	1.159 (34.3)	0.159 (4.70)	0.149 (4.40)	0.535 (15.8)

* PT = planning time, BT = buffer time, RT = reliability time

The proposed ERBTI measure can give a consistent reliability assessment with the reliability order identified using ERBTI is the same as that identified from Figure 4-5. Furthermore, the ERBTI measure can provide a significant identification of reliability differences for different groups. For example, it can be observed that Group C (blue line) and E (red line) have a much different reliability performance from Figure 4-5. The ERBTI values for such two groups are 0.083 and 0.535, respectively, which indicates a much different reliability performance between the two groups.

4.6 Discussions and applications

In transit, different stakeholders have different requirements. Operators are responsible for providing a reliable service to the public. They are concerned of reliability assessment to gain a deep insight into casual relationships between service inputs (service strategies) and outputs (reliability performance). Passengers are the recipient of bus services. They are concerned of deciding departure time to avoid late arrivals at their destinations (Kuhn et al., 2013). Potential applications for fulfilling different stakeholders' requirements are analysed.

4.6.1 Strategy assessment (operators)

Diab and El-Geneidy (2013) studied the impacts of various improvement strategies on service reliability and concluded that the strategies can decrease the standard deviation of travel time. Assuming the current service travel time distribution follows a 3-components GMM model (GMM3), the parameters for the GMM3 model are set as $\mu = [25, 30, 45]$, $p = [0.1, 0.8, 0.1]$ and $\sigma = [1, \sigma_2, 15]$. Different values of σ_2 indicate reliability performance changes after strategies are applied and set σ_2 decreasing from 6 minutes to 5 minutes and then to 4 minutes.

From Table 4-5, it can be seen that the proposed ERBTI measure can accurately reflect the service changes after applying a strategy, while the conventional reliability measures values counter-intuitively stay unchanged. Such phenomenon can be caused by the fact that the conventional measures are largely impacted by travel times under a non-recurrent state. If a non-recurrent state occupies more than 5% of the entire travel time profile, the 95th percentile travel time will remain constant no matter what improvements are made in other states. Furthermore, by considering service reliability under different states separately, different contributions of causes of service reliability are distinguished based on which efficient strategies can be made to improve performance.

Table 4-5: Assessment of service performance changes using different reliability measures

σ_2	PTI (PT*)	BTI (BT*)	RTI (RT*)	ERBTI (ERBT)
4	1.520 (45.0)	0.452 (14.0)	0.520 (15.4)	0.304 (9.0)
5	1.520 (45.0)	0.452 (14.0)	0.520 (15.4)	0.351 (10.4)
6	1.520 (45.0)	0.452 (14.0)	0.520 (15.4)	0.399 (11.8)

* PT = planning time, BT = buffer time, RT = reliability time

4.6.2 Trip planning (passenger)

As passengers are more concerned of TTR than travel time itself in mode choice and departure planning, a new trip planner design is presented to convey such information to passengers as shown in Figure 4-6. The new trip planner provides passengers with a trip summary and different departure options. Under a specific departure option, the trip travel information is presented using two different sections, namely SCHEDULED and EXPECTED. The SCHEDULED section is a brief summary of the scheduled travel information for a trip published by operators, including scheduled departure time, scheduled arrival time and scheduled total time. Usually, the scheduled time-table for the duration a trip is not necessarily equal to the actual operational travel time. The EXPECTED section displays the information of actual travel time and travel time reliability of a trip, including the expected arrival time, the latest arrival time and total expected travel time & total latest travel time based on service reliability.

The information shown in Figure 4-6 is given as an example. The total expected travel time & total latest travel time in the EXPECTED section are calculated using Equations 4.11 and 4.12. Three departure options are provided for different risk-aversion passengers for different trip purposes. For a passenger who needs high reliability, he/she might choose OP1, since the expected arrival and latest arrival time both occur before 8:45am. For a passenger who has less need for reliability, he/she might choose OP1 or OP2 since the expected arrival times are before 8:45am and the latest arrival time is within a tolerable time range.

TRIP PLANNER

Arrival Before : 8.45am on Friday 5 April 2013

From : Origination

To : Destination

Requirements : Bus, train or ferry

OPTIONS	SCHEDULED			EXPECTED		
	Depart	Arrival	Total	Arrival	Latest	Total
OP1: XXX	8:10	8:36	00:26	8:39	8:41	00:29 (00:31)
OP2: XXX	8:15	8:41	00:26	8:44	8:46	00:29 (00:31)
OP3: XXX	8:20	8:46	00:26	8:49	8:51	00:29 (00:31)

Figure 4-6: New designed trip planner for passengers

4.7 Summary

The concern with the impacts of reliability on operating efficiency for operators, as well as service effectiveness for passengers brings about the need to identify and develop meaningful and consistent measures of reliability. Buffer time measures are believed to be appropriate to quantify reliability experienced by passengers in the context of departure planning using operational data. Two issues related to buffer time estimation under mixture service states are addressed in the research, namely, performance disaggregation and capturing passengers' perspectives on reliability.

A GMM model based approach is applied to disaggregate the performance data which provides a great flexibility and precision in modelling the underlying characteristics of travel time. Based on the mixture distributions, a RBT measure is proposed to approximate passengers' experienced reliability by considering different perspectives on reliability under different operational service states. A set of ERBT measures is developed for operators by using different spatial-temporal levels combinations of RBTs. Average trip duration and the latest trip duration measures are proposed for passengers used to make a mode choice and determine the departure time for a trip.

Case studies verify the existence of multi-mode service states during a given time period. The proposed ERBT measures can provide consistent reliability assessment with a high-level detail, while the conventional reliability measures may give inconsistent assessment results. In addition, by considering different passengers' experienced reliability under different states, different contributions of causes can be evaluated based on which effective and efficient improvement strategies can be implemented. A new trip planner design is presented to convey reliability information to passengers. Different options for a trip are provided in the trip planner based on which a passenger can easily make a choice and plan a departure time according to their risk aversion preferences.

Chapter 5 Travel Time Reliability Modelling

5.1 Introduction

To design appropriate strategies to improve service reliability, policy makers should be clear about the causes of unreliability, as well as identify the causes that have the highest impact. Despite a significant body of research on TTR modelling, the conclusions are constrained by the data and approach used. On these basis, an analysis of bus TTR on Australian urban roads was undertaken to validate the factors arising in the literature, to uncover other potential factors that might influence the TTR of bus services, and the lessons to be learnt from bus TTR effects in the Australian context. The research focuses on the stop-to-stop link level reliability modelling, which can provide more insights into the impact of specific causes on service unreliability. In addition, the research used a unique dataset that was built to characterize unreliability by integrating different sources of data, including AVL, Go card, GTFS, BSTM and BoM. The research findings are reported in a journal paper published in Ma, Ferreira, Mesbah and Hojati (2015).

The remainder of the Chapter is structured as follows: In Section 5.2, three general TTR related models are developed with respect to main concerns by travellers and planners, namely, average travel time, buffer time and coefficient of variation of travel time. Five groups of alternative models have been developed to account for variations caused by different road types, including arterial road, motorway, busway, CBD and others. Seemingly Unrelated Regression Equations (SURE) estimation is applied to account for the cross-equation correlations across regression models caused by unobserved heterogeneity. Three main categories of unreliability contributory factors have been identified and tested in this study, namely: planning, operational and environmental. Section 5.3 presents the case studies on comparison between OLS and SURE models, and interpretations of the contributory factors' impacts on TTR at both aggregated and disaggregated levels. The main findings and implications in practice are summarized in Section 5.4. Finally, Section 5.5 concludes this chapter and highlights the future research direction.

5.2 Development of general models and alternative models

The main objective of the study was to identify and quantify the determinants of bus TTR at the link level. Three general models with respect to dependent variables have been developed, namely, travel time, buffer time and coefficient of variation (CV) of travel time. The dependent variables reflect the influence of bus TTR on service attributes of most concerned to passengers and operators.

5.2.1 Dependent and independent variables

The general models were developed using all of the dataset incorporating all types of roads. Five groups of alternative models have been developed to account for variations caused by different road types, including arterial road, motorway, busway, CBD and all other road types.

Optimizing travel time is challenging for transit agencies because changes in travel time have large and usually conflicting influences on service reliability and total operating costs. The general guideline for establishing optimal travel times is to set travel time between two stops equal to the average observed travel time (Kittelson & Assoc et al., 2003). Travel time is also an important factor that can impact passengers travel behaviors (Noland and Polak, 2002).

Buffer time is the additional budgeted time to guarantee arrival at a destination under a specific probability. It is usually defined as the 95^{th} percentile travel time minus the average or median travel time. The 95^{th} percentile travel time refers to a traveler could be late for a work only one time a month (Lomax et al., 2003). This deviation measure captures unreliable service influences on planning behavior of passengers in terms of departure decisions.

CV of travel time is the standard deviation of travel times divided by the mean of travel times. It captures the patterns of travel times in a way that allows direct comparisons across different times, routes and indicators. This variation measure provides a key piece of information for identifying unreliability causes and understanding impacts of various improved strategies on transit service reliability (Diab and El-Geneidy, 2013).

The recurrent congestion index (RCI) is defined as the ratio of mode speed to maximum or ‘free flow’ speed in Equation 5.1. Similar to the calculation approach of congestion level used in (Gilliam et al., 2008), the RCI is calculated using AVL data instead of using the simulated traffic flow from BSTM. To exclude stop delay influence, the link speed is calculated as length over running time.

$$RCI_{t,l} = \frac{V_{t,l}^{mode}}{V_l^{free}} \tag{5.1}$$

where

$RCI_{t,l}$ = recurrent congestion index for time t on link l;

$V_{t,l}^{mode}$ = mode speed for time t on link l;

V_l^{free} = the free flow speed on link l.

The mode speed is the speed that occurs most frequently under a given case. The reference speed is the speed that could theoretically be achieved when the traffic is free flowing. It is usually less than the speed limit in order to allow for slowing down at intersections, stops and other alignment features. The reference speed for each link has been derived from the minimum travel time using the cleaned dataset collected between 5:30 am and 23:30 pm.

The dependent and independent variables used in the models is provided (Table 5-1). The mean values of delay at first stop, boardings and alightings were not included in models II and III, because they were highly correlated with their SD values (Pearson correlations 0.90, 0.91 and 0.95).

Table 5-1: Description of Variables and Models

Variables	Descriptions	I	II	III
Travel time	The travel time between two consecutive stops (second).	◊		
Buffer time	The difference between 95th percentile and median travel times (second).		◊	
CV Travel time	The coefficient of variation of travel time between two consecutive stops.			◊
Planning variables				
Length	The length of the studied segment (kilometre).	√	√	√
Scheduled headway	The scheduled headway of the service along the studied segment (second).	√	√	√
Scheduled stop	The scheduled number of stops along the studied segment.	√	√	√
Weekday [1]	A dummy variable that equals 1 if the observed trip operated on Mon to Fri.	√	√	√
AM peak [2]	A dummy variable that equals 1 if the bus started during the morning peak.	√	√	√
PM peak [2]	A dummy variable that equals 1 if the bus started during the afternoon peak.	√	√	√
Inbound [3]	A dummy variable that equals 1 if the bus operated inbound to city.	√	√	√
Eastbound [3]	A dummy variable that equals 1 if the bus operated eastbound across city.	√	√	√
Operational variables				
Delay at first stop	The delay relative to schedule at the first stop along the studied segment.	√		
Actual stops served	The number of actual stops served by the bus along the studied segment.	√	√	√
SD delay at first stop	The standard deviation of the delay relative to the schedule at the first stop.		√	√
SD actual stops served	The standard deviation of the actual stops served by the bus.		√	√
Number of boardings	The number of passengers boarding the bus along the studied segment.	√		
Number of alightings	The number of passengers alighting the bus along the studied segment.	√		
Boardings squared	The number of boardings squared.	√		
Alightings squared	The number of alightings squared.	√		
SD boardings	The standard deviation of the number of passengers boarding the bus.		√	√
SD alightings	The standard deviation of the number of passengers alighting the bus.		√	√
Environmental variables				
Recurrent congestion index	The proxy index of recurrent traffic congestion state for different time of day.	√	√	√
Number of lanes	The number of lanes of the road link along the studied segment.	√	√	√
Speed limit	The post speed limit of the road link along the studied segment.	√	√	√
Number of signals	The number of signalized intersection along the studied segment.	√	√	√
Signals squared	The number of signalized intersection along the studied segment squared.	√	√	√
Light rain [4]	A dummy variable that equals 1 if the precipitation less than 2.5 mm/hour.	√	√	√
Rain [4]	A dummy variable that equals 1 if the precipitation larger than 2.5 mm/hour.	√	√	√
Motorway [5]	A dummy variable that equals 1 if the bus operated along a motorway road.	√	√	√
Busway [5]	A dummy variable that equals 1 if the bus operated along a busway road.	√	√	√
Arterial [5]	A dummy variable that equals 1 if the bus operated along an arterial road.	√	√	√
CBD [6]	A dummy variable that equals 1 if the bus operated in CBD area.	√	√	√

Notes:
◊ = dependent variables for model and √ = independent variables for model.
1. The referred day type is weekend. 2. The referred time period is off peak. Morning peak = 7:00 - 9:00 and afternoon peak = 16:00 - 19:00. 3. The referred direction is outbound and westbound for radial and cross city service, respectively. 4. The referred weather is good weather. 5. The referred road is local, district and suburban roads. 6. The referred land use is Non-CBD area.

5.2.2 Seemingly unrelated regression equations (SURE) estimation

The multivariate travel time reliability models can be written as,

$$TT = \mathbf{B_1 X_1} + \varepsilon_1 \tag{5.2}$$

$$BT = \mathbf{B_2 X_2} + \varepsilon_2 \tag{5.3}$$

$$CV = \mathbf{B_3X_3} + \varepsilon_3 \tag{5.4}$$

where,

TT, BT, CV = average travel time, buffer time and coefficient of variance of travel time;

$\mathbf{B_1, B_2, B_3}$ = vectors of estimated parameters;

$\mathbf{X_1, X_2, X_3}$ = vectors of independent predictors;

$\varepsilon_1, \varepsilon_2, \varepsilon_3$ = model regression disturbance terms.

In transportation-related studies, the models have been commonly treated separately and estimated equation-by-equation using the standard OLS method. Equations 5.2, 5.3 and 5.4 do not directly interact with each other as one would expect in a classic simultaneous equation system. That is, the travel time does not directly determine the buffer time, buffer time does not directly influence the CV of travel time, and so on. However, following Mannering (2007), the contemporaneous disturbance-term correlations are expected to exist across regression models, since they were measured during the same time period of day on the same link.

In this case, the equations are seemingly unrelated but actually shared common unobserved characteristics which should be considered as a group. Formulating separate OLS models would leave out potentially important contemporaneous correlations that result in consistent but inefficient parameter estimates. To address this problem, SURE estimation can be used to account for the correlation between the shared unobserved characteristics. For detailed information on estimation of SURE, refer to (Washington et al., 2011). Previously, SURE was used to study speed variability in construction zones and travel time variability on freeways (Martchouk et al., 2010; Miller et al., 2009).

5.3 Case Study

To build the dataset required for the analysis, the travel time observations on routes 555 and 60 were aggregated for each 15-min time interval on each link. Operations for weekdays or weekends and different directions were also considered for categorizing. Public holidays have been excluded from the analysis since they have different operation patterns. A sample size threshold of 30-trip observations was found to be the point when the analysis retains its robustness. Accordingly, any group with observations less than 30 was excluded from the analysis and 6535 categories of time-space observations were produced with sample size ranging from 30 to 167. Totally, 2681 categories were excluded with sample size less than 30. As bus bunching always occurs at cultural center station, any link connected to it was excluded from the analysis since such link has considerably different characteristics to other links. Finally, 42 links with 5393 cases were used in the analysis.

5.3.1 Comparison between OLS and SURE estimations

The independent variables (Table 5-1) were examined and pre-tested using statistical analysis. Some variables were excluded from the further analysis because they were either insignificant or collinear with other variables. In particular, the effect of incorporating time and direction dummy variables on model's explanation power was tested using a hierarchical multiple regression, including inbound, eastbound, weekday, AM peak and PM peak. These dummy variables were found to have negligible effects (effect size < 0.007) on improving models' explanatory power since the proposed RCI had already captured the within-day variation of traffic conditions. The descriptive statistics of dependent and independent variables that used in the analysis are provided (Table 5-2).

Table 5-2: Descriptive Statistics of Dependent and Independent Variables

Variables	Type of variable	Min	Max	Average	Std. dev.
Dependent variables					
Average travel time (s)	Continuous	55.75	692.3	178.06	114.89
Buffer time (s)	Continuous	2.5	649	39.25	33.07
CV of travel time[1]	Continuous	0.02	0.81	0.14	0.06
Independent variables					
Length between two stops (km)	Continuous	0.38	8.53	1.92	1.69
Delay at first stop of link (s)	Continuous	-377.37	1168.6	55.61	147.93
Number of actual stops	Continuous	0	1	0.74	0.24
Number of Boardings	Continuous	0	23.19	1.88	2.37
Boardings Squared	Continuous	0	537.66	9.15	25.85
Number of Alightings	Continuous	0	22.72	1.9	2.64
Alightings Squared	Continuous	0	516.2	10.55	34.34
SD delay at first stop of link[1]	Continuous	2.87	758.72	93.54	66.23
SD number of actual stops[1]	Continuous	0	0.52	0.34	0.17
SD number of boardings[1]	Continuous	0	10.92	1.63	1.4
SD number of alightings[1]	Continuous	0	13.9	1.52	1.48
Recurrent congestion index[2]	Continuous	7.87	105.81	74.61	18.74
Number of signals	Continuous	0	8	1.32	2.19
Signals squared	Continuous	0	64	6.57	15.41
Light rain versus good weather[3]	Dummy	0	1	0.08	0.27
Rain versus good weather[3]	Dummy	0	1	0.27	0.44
CBD versus Non-CBD area[1]	Dummy	0	1	0.19	0.39
Arterial versus Other roads[4]	Dummy	0	1	0.08	0.28
Busway versus Other roads[4]	Dummy	0	1	0.47	0.5
Motorway versus Other roads[4]	Dummy	0	1	0.1	0.3

Notes:
1. SD = standard deviation, CV = coefficient of variation, and CBD = central business district. 2. Recurrent congestion index = mode speed divided by free flow speed. 3. Good weather = no precipitation, Light rain = precipitation between 0 and 1.25 mm/30min, and rain = precipitation larger than 1.25 mm/30min. 4. Other roads = road types including local, district, and suburban roads.

To choose the appropriate regression models for Equations 5.2 to 5.4, SURE and OLS models were developed separately using the general dataset. The Pearson correlations of regression residuals were 0.56, 0.23 and 0.73 between Equations 5.2 and 5.3, Equations 5.2 and 5.4, and Equations 5.3 and 5.4, respectively. It indicates the existence of contemporaneous disturbance-term correlations across regression models. By comparing the results between SURE and OLS models, it was found that the standard errors of coefficients in SURE models were significantly smaller than those in OLS models, which highlights the more efficient estimation ability of SURE model.

No significant change of the adjusted R^2 was found but the coefficients of factors changed substantially ranging from 0 to 56%. One would have to accept the SURE estimates as more trustworthy since they have explicitly accounted for correlations of unobserved characteristics. Besides, from practical implications, the result from SURE model seems to make more sense. For example, the deviation of SD delay was not found significant in OLS model but it was found significant in SURE model. As this paper focuses on identifying and quantifying contributory factors, no detailed comparison of SURE and OLS models was provided here. All the regression results presented in the paper were from SURE models.

5.3.2 General models for travel time reliability

Table 5-3 shows the SURE models for average travel time, buffer time and CV of travel time using the general dataset. Overall, they can explain 95%, 46% and 40% of the variations in average travel time, buffer time and CV of travel time observations, respectively. The **bold** values highlight the top five important factors impacting average travel time, buffer time and CV of travel time.

5.3.2.1 SURE Model for Average Travel Time

Consistent with previous studies (Diab and El-Geneidy, 2013; El-Geneidy et al., 2011), travel time increases with an increase in route length, number of actual stops, number of boardings, number of alightings, and number of signals. Route length has the largest positive effect. Delay at first stop has a negative effect on travel time, which means bus drivers who have late departures have less travel times compared to those who depart on time or early. This could be explained by the fact that bus drivers aim to match a predefined timetables. As expected, travel time is adversely impacted by the RCI which means it takes less time to travel when traffic is less congested. Compared to a good weather, rain will increase travel time. This can be attributed to a decrease in driving speed and increase in the gaps between vehicles for safety. No significant difference in travel time was found between light rain and good weather.

The coefficient of boarding time (3.4 seconds) is relatively higher than alighting time (2.5 seconds), since passengers can only use the front door when boarding; while they can use both the front and back doors when alighting. The squared term for alighting indicates that the time associated with passenger alighting decreases with each additional passenger. It means that the first passenger takes an average of 2.5 seconds to alight, and the second passenger will take less time since they have already gotten their smart card and belongings ready. The test result by including the variable of boardings square showed that it had an unstandardized coefficient 0.032 with standard error 0.033. It indicates that the boarding time associated with each additional passenger could also increase, since the subsequent passengers may need more time to find a seat when the bus is crowded. The dummy variables of land use and road type suggest that these factors influence on average travel time are different under different environments.

Table 5-3: SURE Models for Average Travel Time, Buffer Time and CV Travel Time

Predictors	Average travel time B (β) ^{p}	Buffer time B (β) ^{p}	CV travel time[1] B (β) ^{p}
Length between two stops	**73.315 (1.077) ****	**14.806 (0.756) ****	-0.005 (-0.124) **
Delay at first stop of link	-0.015 (-0.020) **	NA	NA
Number of actual stops	20.296 (0.043) **	-8.252 (-0.06) **	**-0.044 (-0.174) ****
Number of Boardings	**3.4080 (0.070) ****	NA	NA
Boardings Squared	—	NA	NA
Number of Alightings	**2.5230 (0.058) ****	NA	NA
Alightings Squared	-0.077 (-0.023) **	NA	NA
SD delay at first stop of link[1]	NA	0.0180 (0.036) **	0.00003 (0.032) **
SD number of actual stops[1]	NA	**10.923 (0.055) ****	**0.1000 (0.272) ****
SD number of boardings[1]	NA	**2.2060 (0.094) ****	**0.006 (0.138) ****
SD number of alightings[1]	NA	1.1000 (0.049) **	0.002 (0.046) **
Recurrent congestion index[2]	**-2.4930 (-0.41) ****	**-1.142 (-0.647) ****	**-0.002 (-0.492) ****
Number of signals	**22.365 (0.426) ****	**3.7400 (0.248) ****	**0.026 (0.933) ****
Signals squared	—	—	-0.003 (-0.734) **
Light rain versus good weather[3]	—	—	0.006 (0.026) *
Rain versus good weather[3]	3.4840 (0.013) **	—	-0.005 (-0.036) **
CBD versus Non-CBD area[1]	-44.310 (-0.151) **	-9.790 (-0.116) **	-0.068 (-0.437) **
Arterial versus Other roads[4]	-108.04 (-0.259) **	-31.016 (-0.258) **	-0.048 (-0.217) **
Busway versus Other roads[4]	-29.675 (-0.129) **	-9.7630 (-0.147) **	-0.030 (-0.241) **
Motorway versus Other roads[4]	-80.580 (-0.209) **	-42.173 (-0.381) **	-0.050 (-0.242) **
Constant	207.36 **	99.607 **	0.27 **
Number of links	42		
Number of cases	5393		
Adjusted R^2	0.949	0.464	0.398

Notes:
The bold values highlight the top five important predictors with higher β than others.
The coefficients B (β) ^{p}: B = unstandardized coefficient, β = standardized coefficient and p = significance level.
t statistics significance ** = $p < 0.01$ and * = $p < 0.05$.
NA stands for Not Applicable information and the symbol '—' stands for insignificant variable with $p > 0.05$.
1. SD = standard deviation, CV = coefficient of variation, and CBD = central business district. 2. Recurrent congestion index = mode speed divided by free flow speed. 3. Good weather = no precipitation, Light rain = precipitation between 0 and 1.25 mm/30min, and rain = precipitation larger than 1.25 mm/30min. 4. Other roads = road types including local, district, and suburban roads.

5.3.2.2 SURE Model for Buffer Time

As anticipated, buffer time increases with an increase in route length, SD delay at first stop, SD actual stops, SD boardings and SD alightings, and number of signals. Route length has the largest positive effect on buffer time. Its impact can be related to factors that create friction, such as traffic entering the road and pedestrian crossings. SD boardings had a more important influence than SD alightings since boardings usually take more time than alightings. Each actual stop made along the route section decreased the buffer time by 8 seconds. It means that a link buffer time is less if the scheduled stop is served all the time, compared to the case of a link where the scheduled stop is served only occasionally. The RCI has a negative influence on buffer time, which indicates that passengers need to budget less buffer time when recurrent traffic congestions are relieved. No significant difference of buffer time was found between good and rainy weather.

5.3.2.3 SURE Model for CV of Travel Time

Consistent with previous study (El-Geneidy et al., 2011), CV of travel time increases with increasing SD delay at first stop, SD actual stops, SD boardings and SD alightings, number of signals. The signal has the largest positive effect on travel time variability. The squared term of signal indicates the influence of the signal on variability will decrease with each additional signal. CV travel time decreases with increasing of route length. The reason is that the variation of travel times could be made up by drivers to comply with the timetable. Each actual stop made along the route section decreases CV travel time. It indicates that the route section with the scheduled stops being served all the time is more reliable than the route section with the scheduled stops being served occasionally. The RCI has an adverse impact on reliability, since it is more probable for an incident to occur and takes longer time for clearance when congestion occurs. Compared to good weather, light rain increases CV of travel time by 0.006 while rain decreases it by 0.005. The combined effects of traffic condition and passenger demand could contribute to this.

5.3.3 Alternative models for travel time reliability

Table 5-4 shows the SURE models for AVG travel time, buffer time and CV travel time on different types of road. Effective strategies can be made after an improved understanding of unreliability factors for each link type. For average travel time, buses travelling on busway experience less running time (57 seconds) and stop loss time (18 seconds) than those travelling on other road types. It takes longer for vehicles travelling in the CBD area and stopping there, when all other variables are kept at their mean values. Vehicles in the CBD area are less influenced by signals (3 seconds), possibly due to less cycle length and coordination of traffic signals. Travelling in the CBD area is more sensitive to rain when compared to good weather. The proposed RCIs are all negatively significant and greatly important in explaining the variations in travel time observations. The boarding and alighting times will decrease with each additional passenger boarding and alighting. The signs of boardings and alightings in OTHERS and MOTORWAY models were unexpected.

Passengers travelling on busway need to budget less buffer time (7 seconds) than those travelling on other roads when all other variables are kept constant. Buffer time on busway is least sensitive to deviations of actual stops and most sensitive to deviation of passenger activities. Deviation of boardings was found to be more important than alightings except in the OTHERS model. The RCIs are all negatively significant and greatly important in explaining the variations in buffer time observations. Signals in CBD area would seem to decrease buffer time. Compared with good weather, light rain was not found significant in alternative models, while rain will increase buffer time on arterial and decrease it on busway.

Table 5-4: SURE Models of Average Travel Time, Buffer Time and CV Travel Time on Different Types of Roads

Predictors	ARTERIAL	MOTORWAY	BUSWAY	CBD[1]	OTHERS[4]
	B (β)	B (β)	B (β)	B (β)	B (β)
Average travel time (in second)					
Length between two stops	104.74 (.397)	67.444 (.764)	56.79 (1.234)	211.5 (1.116)	177.34 (.609)
Delay at first stop of link	-0.04 (-.092)	-0.016 (-.017)	-0.007 (-.037)	-0.070 (-.118)	-0.057 (-.17)
Number of actual stops	23.162 (.114)	52.861 (.106)	18.166 (.140)	79.968 (.200)	—
Number of Boardings	—	4.552 (.116)	4.3320 (.231)	4.5490 (.166)	-2.90 (-.140)
Boardings Squared	—	-0.118 (-.038)	-0.032 (-.012)	-0.207 (-.098)	0.3800 (.141)
Number of Alightings	6.919 (.839)	-5.199 (-.152)	3.205 (.157)	3.1740 (.119)	4.351 (.257)
Alightings Squared	-0.238 (-.47)	0.2710 (.124)	-0.105 (-.039)	-0.146 (-.076)	—
Recurrent congestion index[2]	-1.72 (-.518)	-4.646 (-.217)	-2.194 (-.639)	-3.149 (-.401)	-3.141 (-.99)
Number of signals	5.9420 (.247)	72.659 (.259)	8.3940 (.099)	3.3600 (.085)	NA
Light rain versus good weather[3]	—	—	—	9.209 (.029)	—
Rain versus good weather[3]	3.131 (.048)	—	—	10.037 (.054)	—
Constant	93.589	291.545	187.11	102.29	209.22
Adjusted R^2	*.802*	*.99*	*.938*	*.885*	*.640*
Buffer time (in second)					
Length between two stops	—	11.678 (.465)	6.9960 (.423)	50.889 (.682)	197.81 (.575)
Number of actual stops	—	—	-13.28 (-.285)	—	-45.87 (-.22)
SD delay at first stop of link[1]	—	0.053 (.099)	—	0.0480 (.087)	—
SD number of actual stops[1]	14.302 (.107)	32.427 (.145)	6.8340 (.078)	—	14.930 (.060)
SD number of boardings[1]	1.0430 (.100)	1.824 (.082)	4.939 (.576)	2.0550 (.100)	—
SD number of alightings[1]	—	—	1.1400 (.104)	—	7.6440 (.212)
Recurrent congestion index[2]	-0.31 (-.131)	-1.770 (-.291)	-0.474 (-.384)	-1.050 (-.338)	-3.589 (-.96)
Number of signals	4.1910 (.243)	16.831 (.211)	3.9950 (.131)	-2.180 (-.139)	NA
Light rain versus good weather[3]	—	—	—	—	—
Rain versus good weather[3]	5.179 (.111)	—	-1.714 (-.073)	—	—
Constant	46.545	105.77	50.156	74.576	187.053
Adjusted R^2	*.089*	*.476*	*.326*	*.268*	*.493*
CV travel time[1]					
Length between two stops	-0.266 (-.39)	-0.002 (-.081)	-0.024 (-.394)	-0.036 (-.263)	-0.166 (-.27)
Number of actual stops	—	—	-0.041 (-.239)	-0.070 (-.244)	-0.136 (-.36)
SD delay at first stop of link[1]	—	.00006 (.132)	—	—	—
SD number of actual stops[1]	0.1240 (.257)	0.0370 (.208)	0.1030 (.317)	0.0001 (.091)	0.1250 (.279)
SD number of boardings[1]	0.0050 (.090)	0.0020 (.107)	0.0140 (.440)	—	—
SD number of alightings[1]	—	—	—	—	0.0120 (.178)
Recurrent congestion index[2]	-0.002 (-.19)	-0.002 (-.388)	-0.001 (-.223)	-0.003 (-.563)	-0.006 (-.87)
Number of signals	0.0170 (.268)	0.0140 (.216)	0.0190 (.172)	0.0020 (.085)	NA
Signals squared	—	—	—	-0.001 (-.222)	NA
Light rain versus good weather[3]	—	—	—	—	0.0140 (.054)
Rain versus good weather[3]	—	—	-0.005 (-.058)	0.026 (.092)	—
Constant	0.197	0.223	0.129	0.256	0.533
Adjusted R^2	*0.196*	*0.256*	*0.450*	*0.110*	*0.630*
Number of links	4	4	14	10	10
Number of cases	445	533	2526	1024	865

Note:
The coefficients B (β) for each predictor B = unstandardized coefficient, and β = standardized coefficient.
Only significant variables with p <= 0.05 are presented in the model.
NA stands for Not Applicable information and the symbol '—' stands for insignificant variable with p > 0.05.
1. SD = standard deviation, CV = coefficient of variation, and CBD = central business district. 2. Recurrent congestion index = mode speed divided by free flow speed 3. Good weather = no precipitation, Light rain = precipitation in between 0 and 1.25 mm/30min, and rain = precipitation larger than 1.25 mm/30min. 4. OTHERS = road types including local, district, and suburban roads.

For CV travel time, travel time variability decreases if the buses are scheduled to serve all stops along the route on busway and CBD area. The deviations of actual stops and number of signals are significant in all the models. The squared term of signals in CBD area indicate that the travel time variability decreases with the increase of the number of signals. The deviation of boardings is more important in impacting variability than that of alightings, except in OTHERS model. The RCIs are negatively significant in explaining variations in travel time observations. Compared to good weather, the rain decreases variability on busway and increases it in CBD area.

In brief, the service performance of bus operating on different types of roads can be compared directly in alternative models by excluding the influence of other covariant factors. For example, busway can provide a faster and more reliable service than others in terms of average travel time and buffer time. Another insight can be obtained from alternative models is that the relative importance of factors would differ from different road types. Specific measures should be taken to efficiently improve reliability under different operating environments. For example, in BUSWAY average travel time model, the impact of the number of actual stops (.140) is less important than that of the number of boardings (.231). It implies that making strategies to speed the boarding can be more efficient than find ways to decrease stop delays.

5.4 Main findings and practical implications

The research aims to identify and quantify the underlying determinants of bus TTR on links of different road types using planning, operational and environmental data integrated from sources of AVL, smart card, GTFS, BSTM and BOM database. The main findings are summarized from perspectives of modelling approach and findings. Like any other empirical study, the conclusions are valid within the range of the used data and should be used with caution beyond this range.

1) Modelling Approach

Different from previous reliability studies in transit, the recurrent congestion index was defined to represent traffic conditions and SURE model was used to estimate coefficients.

- The RCI was found to be highly significant in reliability models. Further including dummy variables of time and direction had a negligible effect on models' explanatory power.
- Pearson test verified the existence of correlations between reliability models.
- The SURE model is capable to provide more efficient estimations than OLS model.

Implications: These findings offer a new perspective to model TTR in transit. The SURE model should be regarded as more trustworthy since it has explicitly accounted for cross-equation correlations of disturbance terms. The model could be generalized to other link since the RCI can well capture enroute traffic conditions. Other groups of reliability measures can be modelled using the approach proposed here.

2) General Models

From general models, the top five important factors for different attributes can be identified.

- For average travel time: link length, recurrent congestion index, signals, number of boardings, and alightings.
- For buffer time: link length, recurrent congestion index, signals, SD actual stops, and SD boardings.
- For the coefficient of variation of travel time: signals, recurrent congestion index, SD actual stops, number of actual stops and SD boardings.

Implications: These findings are supportive of general strategies, such as designated bus lanes, busway, signal priority, stop consolidation and smart card payment.

3) Alternative Models

Alternative models were developed to account for variations caused by different types of roads.

- From alternative models, service performance on different types of road can be compared by excluding the covariant factors impacts. For example, busway was found to provide a faster and reliable service than others.
- Alternative models can provide detailed insights into the influence of specific causes. For example, the influence of signal is not as important as that of actual stops made along route.

Implications: These findings can help to facilitate efficient strategies under different scenarios to improve service reliability and mitigate the impacts of unreliability for both travellers and operators. For example, in CBD areas, it may be more effective to introduce measures which reduce stop loss time rather than to implement signal priority.

5.5 Summary

The concern with making efficient and effective strategies to improve service reliability brings about the need to identify and quantify the impact of unreliability causes on TTR. Despite significant research in private vehicle reliability modelling, there has been much less emphasis on modelling link level transit TTR on different types of roads. The research identifies the most important factors that influence the service attributes for passengers and operators, to enable effective and efficient strategies to improve transit reliability performance.

A comprehensive set of reliability causes associated with planning, operational and environmental perspectives, has been estimated and tested using 6 months data from two bus routes in Brisbane, Australia. The data sources include AVL for stop level vehicle travel times, smart card for passenger demand, GTFS and BSTM for route characteristics and BOM for weather observations. A recurrent congestion index was developed here to reflect within-day variation of traffic conditions using historical travel time observations, instead of using dummy variables.

A SURE model was applied to address the inefficient coefficient estimation issue caused by the unobserved shared service characteristics across regression models for average travel time, buffer time and CV of travel time. The statistical tests suggest that the congestion index is highly significant in reliability models. Cross-equation correlations were found to exist between reliability models and the SURE provides more efficient estimation than the OLS model.

The model results provide insights into the causes that affect bus travel time, buffer time and the coefficient of variation of travel time. Targeted strategies are likely to be more effective and efficient after an improved understanding of the factors which impact on reliability for each link type. Due to the different characteristics of road links and within-day variation of traffic conditions captured by the recurrent congestion index, the results can be generalized to predict average travel time and its reliability on other bus routes with similar link types. Other reliability related dependent variables, such as headway regularity and schedule adherence, can be modelled using the approach proposed here.

Chapter 6 Travel Time Distribution Modelling

6.1 Introduction

Bus travel time reliability performance influences service attractiveness, operating costs and system efficiency. Better understanding of the distribution of travel time variability is a prerequisite for reliability analysis. A wide array of empirical studies has been conducted to model distribution of travel times in transport. However, depending on the data tested and approaches applied to examine the fitting performance, different conclusions have been reported. While some studies have considered symmetrical distribution models, others have preferred skewed and multimodal ones. These inconsistencies clearly affect both the ability to gain insights into the nature of TTV and inhibit the ability to generalize findings to other applications. The research aims to specify the most appropriate distribution model for the day-to-day travel time variability by using a novel evaluation approach and set of performance measures. Two important issues are addressed: 1) Data aggregation influence on the attributes of TTV, and 2) Evaluation of the alternative distribution models' performance. A novel evaluation approach and set of measures are developed to facilitate comprehensive comparison of alternative distribution models. The research findings are reported in a journal paper published in Ma, Ferreira, Mesbah and Zhu (2015).

The remainder of the Chapter is structured as follows: In Section 3, the evaluation approach focusing on finding the most appropriate distribution model is described. The data aggregation influence on the feature of TTV is investigated and the alternative distribution models' performance is evaluated in Section 4. The decrease of temporal aggregation of travel times tends to increase the normality of distributions. The spatial aggregation of link travel times would break up the link multimodality distributions for a busway route, but unlike for a non-busway route. The Gaussian Mixture Models is evaluated as superior to its alternatives in terms of fitting accuracy, robustness and explanatory power. The identified most appropriate model is further discussed in Section 5, including the reasons for its superior performance and its applications. The reported distribution model shows promise to fit travel times for other services with different operation environments considering its flexibility in fitting symmetric, asymmetric and multimodal distributions. Finally, Section 6 provides the main conclusions, as well as highlights potential future research.

6.2 Distribution evaluation approach

The proposed evaluation approach can provide comprehensive comparison of distribution fitting performance from three aspects, including case generation, alternative models selection and evaluation measures. The alternative distribution models are tested for different cases and evaluated comprehensively considering accuracy, robustness and explanatory power in order to specify the most appropriate distribution model in fitting day-to-day variability of bus travel time. The detailed evaluation approach is described as follows:

1) Case Generation

Test cases are generated by aggregating the pre-processed travel time data in combinations of different temporal-spatial scales and time components. The considered aggregation attributes are temporal scale (weekday or weekend, period, 60 min, 30 min, 15 min, and 5 min), spatial scale (directions, route level and link level) , and time components (travel time, running time and stop delay time). Accordingly, five distinct periods are used: AM off-peak (05:00-07:00), AM peak (07:00-09:00), Inter peak (09:00-15:00), PM peak (15:00-19:00) and PM off-peak (19:00-23:00). A case is a combination of the above aggregation attributes, such as weekday inbound AM peak route running time.

2) Distribution Fitting

For observations under each case, alternative distribution models are used to fit them. The single distribution models, including Burr, Normal, Lognormal, Gamma, Weibull, Logistic, and Loglogistic, are chosen from the literature that has been reported to be the best under specific testing environments. The PDF parameters are estimated using the Maximum Likelihood method. The GMM model, a special case of mixture distribution model, is also considered. The maximum components number K is set to be 3 considering the interpretation of the parameters in reality, that can be related to free flow, recurrent and non-recurrent service states .

3) Hypothesis Test

For fitted distributions under each case, the Anderson-Darling (AD) test is used to test if the alternative distribution models pass the null hypothesis H_0 that the observations comes from the alternative distributions (Anderson and Darling, 1954). A larger AD significance value highlights a better fitting performance of the model. The distribution model is rejected when the value of AD significance is smaller than 0.05. If accepted, the alternative distribution is placed into the candidature models pool. The candidate distributions fitting performance is then ranked by using the AD significance values in an ascending order. For example, if the GMM performs the best, it has the top mark of 1.

4) Performance Summary

The statistics, accuracy and robustness for each alternative distribution model, are summarized, and their explanatory power discussed. Here accuracy means the model can fit the observations with only a small fitting error. It can be measured using descriptive statistics of AD significance value and distribution mark. Robustness means the model itself can adjust to different cases with a tolerable fitting error, especially under complex situations. It can be measured using the proportion of cases that passed the hypothesis test. Explanatory power indicates the distribution model describes the reality in a useful way and is flexible enough to capture hidden patterns of travel times.

6.3 Distribution evaluation measures

Two groups of measures are chosen and calculated for aggregation influence analysis and performance evaluation. To explore the data aggregation influence on the shape of distribution, a set of measures of symmetry, normality and multimodality are selected. Skewness/se is a measure of the degree of asymmetry of a distribution while kurtosis/se is a measure of the 'flatness' (vs peakedness) of a distribution (Washington et al., 2011). The standard error $se = \sqrt{6/N}$, where N is the sample size. A Skewness (kurtosis) value of more than twice the corresponding standard error se is sufficient to reject a 0 value for skewness (kurtosis). A higher skewness/se (kurtosis/se) value highlights a more asymmetrical (peaked) distribution. The kurtosis/se can be used to indicate where the variance of data comes from. If the distribution is not peaky, the variance is distributed throughout. If the distribution is peaky, the variance of data close to the distribution centre is little and the variance mainly comes from tails. Normal significance and unimodal significance are measures of the degree of normality and unimodality, respectively. A higher significance value indicates a better normality or unimodality of TTD. The normal significance value is calculated using AD test and the unimodal significance value is calculated by the Hartigan dip test (Hartigan and Hartigan, 1985). The null hypothesis H_0 for the dip test is that the TTD is unimodal. The zero hypothesis H_0 cannot be rejected a distribution is unimodal with a significance value larger than 0.05.

To evaluate the alternative distribution models' performance in fitting day-to-day variability of bus travel times, a set of measures of accuracy and robustness are developed. Survivor function $\mathbf{Suv}(\bullet)$ is developed that can capture the probability that the distribution model will survive beyond a specified AD significance value.

$$\mathbf{Suv}(AD\_sig) = 1 - \mathbf{F}(AD\_sig) \tag{6.1}$$

where

AD_sig = the AD significance value;

$\mathbf{F}(AD\_sig)$ = the cumulative density function of AD significance value.

For a specified value of AD significance, a larger survivor probability highlights a more robust model. For a specified value of probability, a larger AD significance value indicates a more accurate model. Also, for accuracy measures, the mean of AD significance values and Cases_top3 ratio (ratio of cases marked with top 1 to top 3 to the total number of cases) are calculated. A model with a larger mean and higher Cases_top3 ratio highlights a more accurate model. For robustness measures, the standard deviation of AD significance values and Cases_pass ratio (ratio of cases with AD significance value larger than 0.05 to the total number of cases) are calculated. A model with a smaller standard deviation and higher Cases_pass ratio indicates a more robust model. The explanatory power is discussed by examining its model structure in fitting different types of distributions and the interpretation of its parameters in reality.

6.4 Case Study

The motivation for the research was to specify a type of distribution that can appropriately model the day-to-day TTV for public transport. The ideal data for empirical study would have a large extent coverage of services that operating in different times and spaces. The routes 555 and 60 data were used. Buses operating on the two routes were equipped with AVL systems that can provide travel time information in different time and space scales, which satisfied the data requirement of this study. The two typical routes cover diverse operating environments, including CBD area, residential area, major attraction area, suburban road, arterial road, motorway, and exclusive busway. The data used covers a six months' period with service operating from 5:30 to 23:30 every day. The two routes were used as a prototype to evaluate the most appropriate distribution models for other routes with similar operation environment. In total, 5,002 and 56,316 numbers of cases are identified for route and link levels times, respectively.

6.4.1 Aggregation impacts on distribution

Statistical tests were conducted to examine the symmetry, asymmetry, normality and multimodality of a distribution in order to explore the data aggregation influence on the characteristic of time distributions. Temporal and spatial aggregation impacts were investigated separately.

6.4.1.1 Temporal aggregation

Different levels of temporal aggregation that influence route and link level TTDs were examined. First, the distributions for different temporal aggregation levels of route travel times were visualized using the histogram for the weekday inbound service. Figure 6-1 shows the TTDs for AM peak and Inter peak time periods. For the AM peak period, route 60 travel times show a multimodality distribution while route 555 travel times shows an asymmetric and flat distribution with a short right tail.

The multimodality phenomenon may be caused by random vehicle delay times at signalized intersections and delay times at stops along the route. Accordingly, it is rather hard to use a uni-modal distribution model to fit such travel times with several peaks. For the Inter peak period, the TTDs on the two routes are rather symmetric with a small proportion of large travel times on the right compared to the distributions for the AM peak period. A normal distribution may appropriately characterize the Inter peak travel times.

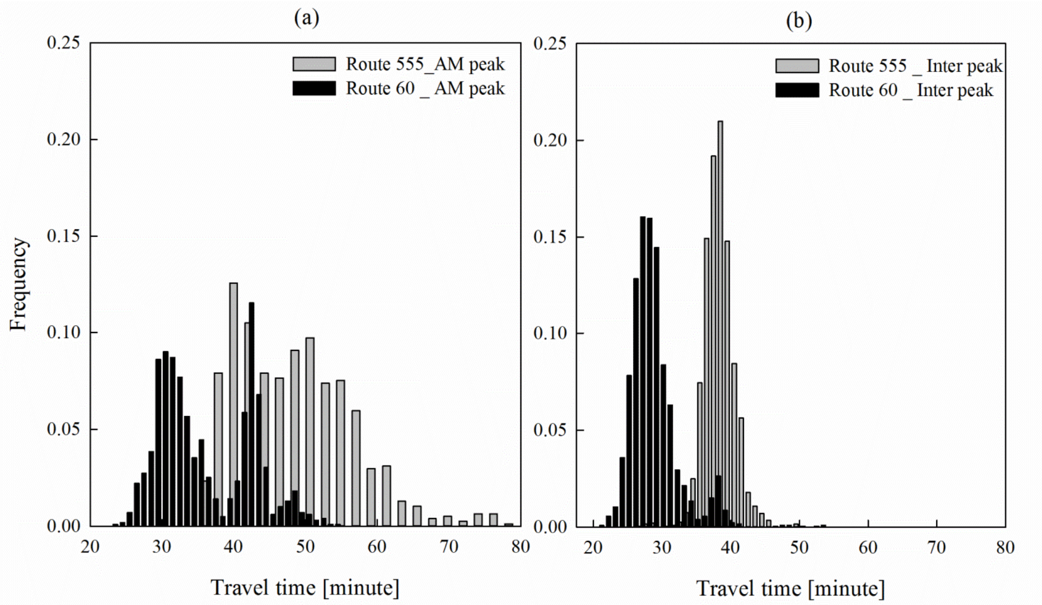

Figure 6-1: Distribution of travel times for routes 555 and 60 during (a) AM peak period and (b) inter peak period.

A different picture of TTDs emerges by decreasing the temporal aggregation level of travel times. Figure 6-2 shows the distribution for route 555 travel times in 60 min and 15 min departure time windows during the AM peak time period.

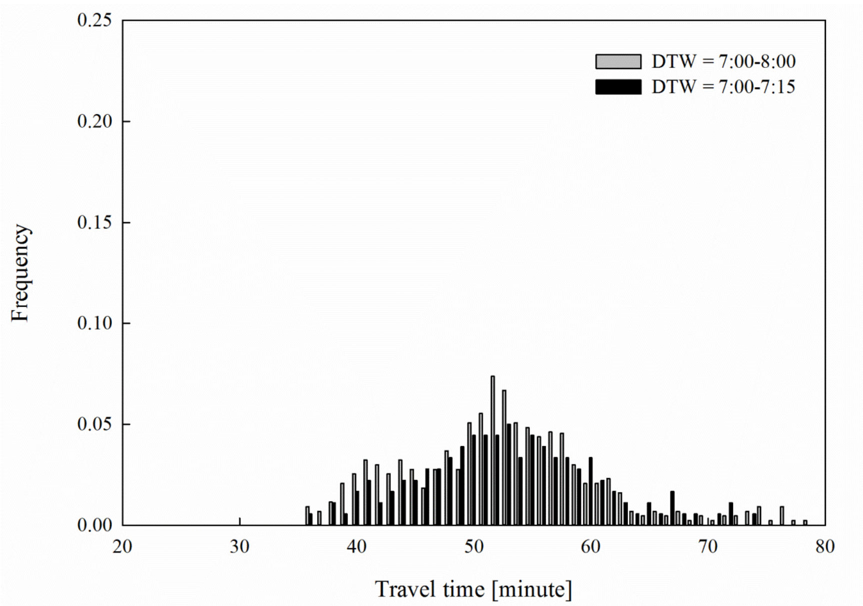

Figure 6-2: Distribution of travel times with departure time window (DTW) 60 minutes and 15 minutes during AM peak period.

In a comparative sense, both the TTDs are relatively more symmetrical than those shown in Figure 6-2 and TTD for the 15 min DTW is relatively more symmetrical than that for the 60 min DTW. A normal distribution would provide a promising fit for a short temporal aggregation level. To better understand the temporal data aggregation influence on time distributions, statistical tests were conducted to examine the characteristics of travel times for different cases.

Table 6-1 shows a series of key descriptive statistics of route level distributions for bus travel time components (travel time, running time and dwell time) with different aggregation levels for weekday inbound travel. Similar results were also found for other tested scenarios. No result of 5 min aggregation level was provided for route 555 since its minimum headway is 15 min. First, results of different statistical measures with a decrease of aggregation levels under each scenario were examined, such as route 555 peak travel time. No significant difference was found for the measure of COVs across different aggregation levels, except for route 60 Inter peak travel times and running times. The skewness/se and kurtosis/se values decreased under all scenarios which highlight a less skewed and more flat distribution. Accordingly, the normal sig values increased under all scenarios which indicate a more symmetric distribution. The normal distribution seems to be an appropriate model for travel times within a small aggregation level (e.g. 5 min) since the normal sig values are much larger than 0.05 for all cases. These findings are consistent with the visualized analysis above and the results reported by Mazloumi et al. (2010). Similarly, it seems rather unlikely that the distribution is multimodal for a very short time interval. The reason for those could be less variation of factors influencing travel times that exists for a shorter DTW. For an ideal assumption, if only one factor has a significant variance for a certain short DTW, the resultant travel times should follow a normal distribution.

Comparing peak and off-peak time periods for a same aggregation level, the COVs of the peak period tend to be larger than those of the off-peak period, since the travel conditions are more complicated for peak hours. However, by examining the values of skewness/se and normal sig measures, the off-peak travel times and running times show a more skewed and asymmetric distribution than those of the peak period while an opposite result was achieved for other levels aggregation. These seemingly conflicting results could be caused by the different time intervals for the AM peak period (2 hours) and off-peak periods (6 hours). The kurtosis/se measure indicates a more peaked distribution of travel times and running times during the off-peak period. The unimodal sig measure highlights that it is less likely for travel time and running time distributions to be multimodal during an off-peak period.

Table 6-1: Key descriptive statistics of travel times with different temporal aggregation level

	Route	Time period	Aggregation level	Sample size	COV[1]	Skewness /se[2]	Kurtosis /se[2]	Normal sig[3]	Unimodal sig[3]
TT*	555	Peak	AM peak	662	0.16	4.83	30.80	0.02	0.29
			60 min	331	0.15	3.16	22.29	0.10	0.69
			30 min	164	0.14	2.46	16.94	0.35	0.57
			15 min	83	0.14	1.19	11.00	0.63	0.75
		Off-peak	Period	2600	0.06	15.42	91.72	0.00	0.98
			60 min	372	0.05	1.43	23.85	0.37	0.76
			30 min	187	0.05	1.12	17.04	0.67	0.78
			15 min	93	0.05	0.99	11.87	0.74	0.64
	60	Peak	AM peak	972	0.17	4.34	23.51	0.00	0.00
			60 min	486	0.17	3.26	17.20	0.00	0.00
			30 min	243	0.17	2.90	13.94	0.00	0.04
			15 min	122	0.17	2.15	9.71	0.01	0.04
			05 min	42	0.15	0.80	5.09	0.25	0.20
		Off-peak	Inter peak	1688	0.12	31.67	134.62	0.00	0.95
			60 min	238	0.08	2.60	20.44	0.51	0.85
			30 min	119	0.08	1.81	14.39	0.60	0.70
			15 min	59	0.08	1.41	10.03	0.69	0.71
			05 min	40	0.08	1.34	8.22	0.72	0.74
RT*	555	Peak	AM peak	662	0.16	5.76	33.74	0.01	0.09
			60 min	331	0.15	3.88	23.65	0.08	0.83
			30 min	164	0.14	3.40	18.57	0.28	0.66
			15 min	83	0.15	1.70	11.87	0.52	0.68
		Off-peak	Inter peak	2600	0.06	14.83	113.89	0.00	0.87
			60 min	372	0.05	1.93	29.57	0.62	0.93
			30 min	187	0.05	1.98	19.84	0.63	0.94
			15 min	93	0.05	1.28	13.08	0.77	0.86
	60	Peak	AM peak	972	0.21	6.02	26.27	0.00	0.00
			60 min	486	0.21	3.98	18.98	0.00	0.00
			30 min	243	0.20	3.40	14.74	0.00	0.09
			15 min	122	0.20	2.35	10.04	0.01	0.12
			05 min	42	0.18	0.84	5.11	0.24	0.20
		Off-peak	Inter peak	1688	0.14	34.94	145.92	0.00	0.98
			60 min	238	0.09	3.43	24.35	0.50	0.77
			30 min	119	0.09	2.40	16.73	0.58	0.77
			15 min	59	0.09	1.79	11.10	0.63	0.71
			05 min	40	0.09	1.50	8.90	0.68	0.71
DT*	555	Peak	AM peak	662	0.34	18.85	137.70	0.00	0.82
			60 min	331	0.33	10.63	71.47	0.14	0.95
			30 min	164	0.33	6.48	42.49	0.38	0.94
			15 min	83	0.32	3.52	21.05	0.53	0.82
		Off-peak	Inter peak	2600	0.28	17.47	109.22	0.00	0.81
			60 min	372	0.27	5.19	35.16	0.29	0.61
			30 min	187	0.27	3.36	23.02	0.52	0.74
			15 min	93	0.27	2.11	14.74	0.67	0.77
	60	Peak	AM peak	972	0.22	9.02	95.42	0.04	0.96
			60 min	486	0.21	7.59	74.90	0.24	0.94
			30 min	243	0.20	6.47	55.27	0.16	0.80
			15 min	122	0.19	3.92	32.25	0.42	0.90
			05 min	42	0.19	2.21	13.17	0.61	0.79
		Off-peak	Inter peak	1688	0.23	8.55	68.90	0.00	0.88
			60 min	238	0.22	4.14	25.79	0.28	0.86
			30 min	119	0.22	2.86	17.69	0.46	0.89
			15 min	59	0.22	2.09	12.11	0.63	0.81
			05 min	40	0.21	1.69	9.78	0.67	0.72

* TT = Travel time, RT = Running time, DT = Delay time. 1. COV = Coefficient of Variance. 2. se = standard error. 3. sig = significance value of hypothesis test.

Comparing the same aggregation level across different time components shows that delay time tends to have a different patterns of changes with travel time and running time between the peak and off-peak time periods, such as during the peak hour, delay time have a more peaked distribution than during the off-peak hour. Also, delay times have a larger variability and a more skewed but peaked distribution. It indicates that the variance of stop delay time is largely influenced by some extremely small and large observations. These analyses have revealed that the temporal data aggregation could alter the nature of time distributions with different pattern change behaviour. An appropriate temporal aggregation level should be selected before distribution fitting.

6.4.1.2 Spatial aggregation

Different links along a bus route have different characteristics, such as road types, signalized intersections and land use (CBD, major attractors or residential area). These different characteristics can lead to different spatial time distributions. Figure 6-3 shows the actual and scheduled travel time and its COV of different links along route 555 for the weekday inbound AM peak service. Table 6-2 shows the characteristics of links and key descriptive statistics of travel time distributions.

Comparing COVs among the different links shows that, links 1, 7, 10 and 11 have relatively larger values than the others which could be caused by the combined effects of road type, road length, major attractors and signalized intersections. For links 1 and 11, the inbound traffic condition is usually congested during the AM peak period and the signalized intersections would further worsen the situation. For links 7 and 10, the large variability is mainly caused by bus bunching and high passenger demand at these stops. Compared to the measures for route level travel times in Table 1, the link travel times are more complicated with only 3 out of 11 links having normal sig values larger than 0.05. Conceptually, for a specific DTW, a link level TTD would be more viable than a route level TTD since the intersection and stop delay times will occupy a greater proportion of the travel times for the former.

Table 6-2: Characteristic of links and key descriptive statistics of TTDs [weekday inbound AM service 555]

Link number	Route type	Major attractor	Length (km)	Signal	COV	Skewness/se[1]	Kurtosis/se[1]	Normal sig[2]	Unimodal sig[2]
1	Motorway	Shopping	8.80	4	0.28	4.52	31.93	0.00	0.00
2	Motorway	-	5.60	2	0.19	6.72	34.59	0.00	0.78
3	Bus way	-	2.60	0	0.14	6.95	37.31	0.00	0.36
4	Bus way	Shopping	2.60	0	0.12	4.15	35.30	0.14	0.05
5	Bus way	University	2.20	0	0.14	4.35	35.00	0.07	0.00
6	Bus way	-	2.70	0	0.11	2.81	31.42	0.14	0.21
7	Bus way	Hospital	1.60	0	0.22	12.57	50.23	0.00	0.48
8	Bus way	-	1.90	0	0.15	6.77	38.88	0.00	0.77
9	Bus way	-	0.80	2	0.15	5.70	44.39	0.00	0.00
10	Bus way	Major stop	1.00	3	0.39	12.00	45.50	0.00	0.21
11	Suburban	-	0.75	3	0.30	10.59	44.08	0.00	0.94

1. se = standard error. 2. sig = significance value of hypothesis test.

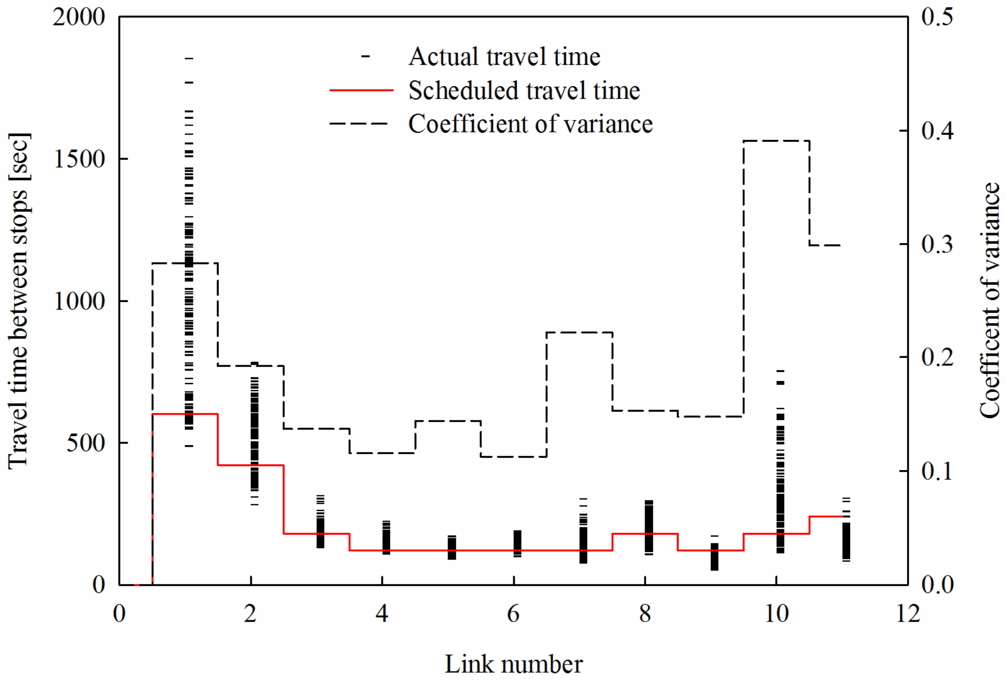

Figure 6-3: Travel times (actual and scheduled) and coefficient of variance (COV) between different stops along the route.

To examine the multimodality of distributions in Table 6-2, 4 out of 11 links have multi-modal distributions whereas the route level travel times for the AM peak have a unimodal distribution. The multimodal phenomenon on the links seems to be broken up when aggregated to a route level travel time which is consistent with the findings reported by Susilawati et al. (2013). The large distinction between different links travel times could be mutually made up by an increase of the spatial aggregation. For example, if a vehicle drives relatively slowly at the first link, the driver would speed up at the following stops to catch up with the time table.

To further explore the spatial aggregation level influence on the multimodality distributions, the unimodal sig values for link level travel times of route 60 in AM peak period and 60 min DTW are presented in Table 6-3, along with the characteristic of route 60.

Table 6-3: Characteristics of links and unimodal statistics of TTDs [weekday eastbound AM service 60]

Link number	Route type	Major attractor	Length (km)	Signal	Unimodal sig[1] (AM peak)	Unimodal sig[1] (60 min)
1	Local	Residential	0.54	0	0.00	0.01
2	District	Residential	0.76	0	0.01	0.04
3	District	-	0.77	0	0.01	0.12
4	District	-	0.47	1	0.17	0.40
5	District	Major stop	0.89	6	0.70	0.90
6	Suburban	CBD	0.78	3	0.43	0.82
7	District	CBD	0.56	3	0.00	0.17
8	Arterial	CBD	1.23	4	0.75	0.96
9	Arterial	-	0.68	4	0.78	0.68
10	Arterial	-	0.58	3	0.28	0.50
11	Local	-	0.70	0	0.02	0.20

1. sig = significance value of hypothesis test.

The results show that multimodality generally occurs at the first three links and not for the following links except the last link in the AM peak period. However, the multimodality of link TTDs cannot be made up by the spatial aggregation since the route level travel times still have a clear multimodal distribution as shown in Table 6-1. This reveals that the drivers would have limited flexibility to speed up on route 60 when constrained by the traffic conditions.

6.4.2 Distribution fitting performance evaluation

According to the above analysis, different data aggregation strategies could alter the time variability differently. To evaluate the alternative distribution models fitting performance and choose the most appropriate model, hypothesis AD tests were conducted for all cases with different combinations of temporal level, spatial level attributes and time components.

6.4.2.1Route level distribution

Figure 6-4 (a) displays the route level survivor function of the AD test significance value for alternative distribution models. The survivor curve highlights the probability that a model can provide a promising fitting performance for a specified significance value. For example, for a given significance value 0.4, the GMM model has a maximum probability to survive while the Weibull model has a minimum probability to survive. Under almost 95% of cases, the GMM can provide an AD significance value larger than 0.7 which highlights its superior performance to its alternatives in terms of accuracy and robustness. This is further discussed in the following section using two test cases. Comparing survivor function among the candidates, the Weibull model has the worst fitting performance with the fastest decrease rate as the AD significance increases. The Burr model and GMM model have a relatively similar steady survival AD significance value range from 0 to 0.7, which highlights their accurate fitting performance when they can converge to a solution. However, the initial drop of Burr model at the AD significance value 0 indicates that the Burr model cannot converge to a solution for almost 20% of cases. The failure cases of Burr model will decrease its application in reality even though it can provide a highly accurate fitting when it can converge. The performance of Loglogistic, Logistic, Gamma, Lognormal and Normal are similar although the Loglogistic model has a relatively better fit.

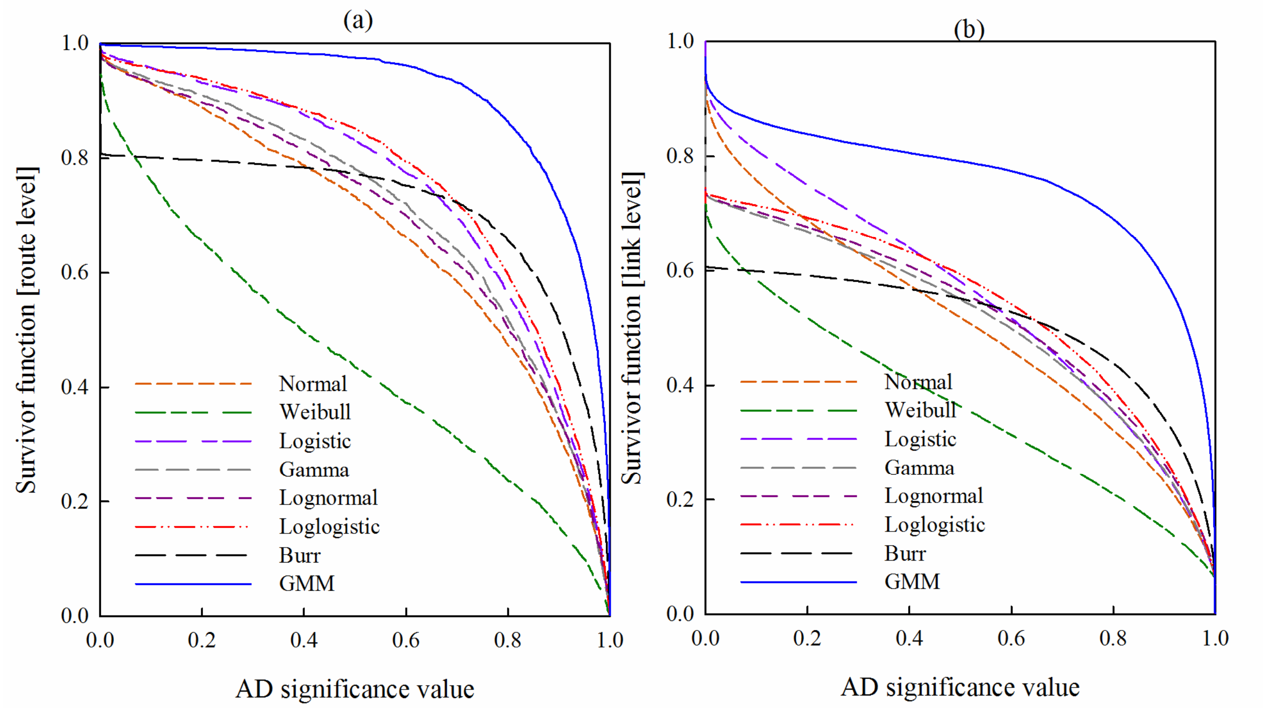

Figure 6-4: Survivor function of Anderson-Darling (AD) test significance for alternative distribution models (a) route level and (b) link level. GMM Gaussian mixture models

Table 6-4 shows the descriptive summary of the AD significance values and the alternative distributions performance. The results show that the GMM model has the largest AD significance mean and median values with the smallest standard deviation. This further highlights the relatively better accuracy and robust performance of the GMM model compared to its alternatives. The GMM model passes the AD test in 4,984 of 5,002 cases, which indicates its good flexibility to adjust to different situations for route level travel times. The GMM and Burr models are listed 3,092 and 3,106 times, respectively, in the top 3 best fitting distributions. However, the Burr model passes the AD test in only 4,022 out of 5,002 cases which is rather low compared to the GMM model. Figure 6-5(a) shows the distribution of the top 3 models for all cases. It can be clearly observed that the GMM model has a much larger proportion of the best fitting model compared with the Burr model.

Table 6-4: Descriptive summary of AD significance value and candidature distributions performance [route level]

Model	Mean_sig*	Median_sig*	SD_sig*	Cases_pass[1]	Cases_top3[2]
Normal	0.67	0.78	0.30	4751	1422
Weibull	0.45	0.40	0.35	4134	582
Logistic	0.75	0.84	0.26	4855	1459
Gamma	0.71	0.81	0.29	4769	1095
Lognormal	0.70	0.80	0.30	4740	1709
Loglogistic	0.76	0.86	0.26	4832	2359
Burr	0.71	0.91	0.38	4022	**3106**
GMM	**0.91**	**0.97**	**0.14**	**4984**	3092

* sig is the AD test significance value.
The bold value indicates the best model identified under each performance measure.
1. A passed distribution with AD *p value* > 0.05. The total number of cases is 5,002.
2. The total number of cased being listed as the top 3 and the total number of cases is 14,824.

6.4.2.2 Link level distribution

Figure 6-4(b) shows the link level survivor function of the AD test significance value for the alternative distributions. The results show that the GMM model has a better fitting performance than its alternatives for link level travel times. Comparing the alternative distributions performance between route level and link level travel times, all decrease when modelling link level travel times. Each model has a survival probability drop ranging from 8% to 40% at the AD significance value 0. These indicate the greater complexity of the distributions for link travel times than those of route travel times, which is consistent with the previous analysis in Section 6.4.2. Moreover, according to the Hartigan dip test, the proportions of multimodality cases for route level travel times and link level travel times are 2% and 16%, respectively. The increase of the multimodality proportions should be another factors worsen the alternative distributions performances. Relatively, the Weibull model provides the worst performance while the Normal, Logistic, Gamma, Lognormal, Loglogistic and Burr models have a similar and intertwined performance.

Table 6-5 shows the descriptive statistics of the AD significance value for the alternative distributions performance of link level travel times. The results show that the GMM model performs better than its alternatives with the largest mean and median significance values and the smallest standard deviation. It also has a better robustness characteristic than the other distributions with the maximum number of cases passing the AD test and being listed in the top 3 clusters.

Figure 6-5 (b) shows the distribution of the top 3 models for all cases and that the number of top 1 cases for the GMM model is more than the total number of the top 3 cases of the Burr model. This illustrates the greater flexibility of the GMM model for link level travel times. Obviously, the performance of all distribution models decreases largely in modelling the link level travel times compared with modelling the route level travel times, from the perspective of mean, median and SD of significance values.

Table 6-5: Descriptive summary of significance value and candidature distributions performance [link level]

Model	Mean_sig*	Median_sig*	SD_sig*	Cases_pass[1]	Cases_top3[2]
Normal	0.51	0.53	0.37	45347	19864
Weibull	0.36	0.23	0.37	35305	8677
Logistic	0.55	0.62	0.36	47772	18657
Gamma	0.51	0.60	0.40	40163	9383
Lognormal	0.52	0.62	0.40	40311	16121
Loglogistic	0.54	0.67	0.40	40737	16586
Burr	0.51	0.68	0.44	33946	22713
GMM	**0.75**	**0.94**	**0.35**	**49563**	**34246**

* sig is the AD test significance value.
The bold value indicates the best model identified under each performance measure.
1. A passed distribution with AD *p value* > 0.05. The total number of cases is 56,316.
2. The total number of cased being listed as the top 3 and the total number of cases is 146,247.

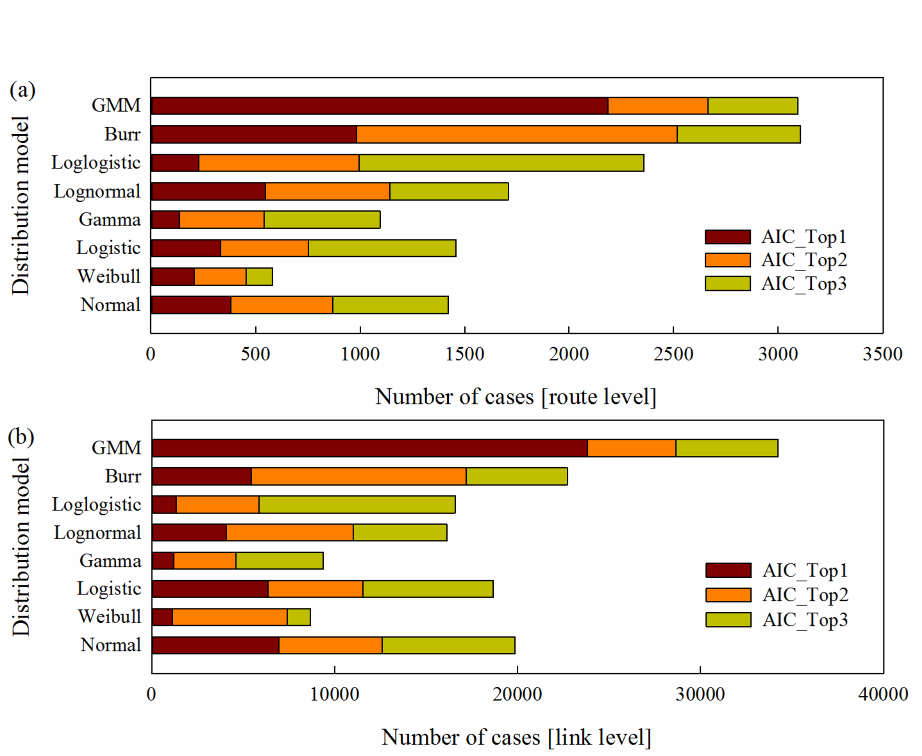

Figure 6-5: Summary of the distribution of top 3 models for all cases (a) route level and (b) link level. GMM, Gaussian mixture models

To make a direct comparison of the fitting performance of the alternative models for the route level and link level travel times, the COV of significance value, the passed cases ratio and the top 3 cases ratio are presented in Table 6-6. Clearly, all the distribution performances decrease in modelling link travel times with larger COV_sig values and smaller passed cases ratios. Comparatively, the GMM Cases_top 3 ratio increases from 21% to 23% while the Burr decreases from 21% to 16%. The passed_cases ratio for the GMM model in the link level scenario is still promising (88%) which indicates a relatively strong flexibility in modelling complex distributions of link travel times.

Table 6-6: Comparison of candidature models fitting performance for route level and link level travel times

Model	COV_sig[1]		Cases_pass ratio[2]		Cases_top3 ratio[3]	
	Route	Link	Route	Link	Route	Link
Normal	0.45	0.73	0.95	0.81	0.10	0.14
Weibull	0.78	1.03	0.83	0.63	0.04	0.06
Logistic	0.35	0.65	0.97	0.85	0.10	0.13
Gamma	0.40	0.77	0.95	0.71	0.07	0.06
Lognormal	0.43	0.76	0.95	0.72	0.12	0.11
Loglogistic	0.34	0.73	0.97	0.72	0.16	0.11
Burr	0.53	0.87	0.80	0.60	**0.21**	0.16
GMM	**0.16**	**0.47**	**1.00**	**0.88**	**0.21**	**0.23**

The bold value indicates the best model identified under each performance measure.
1. COV_sig is calculated as SD_sig divided by Mean_sig.
2. Cases_pass ratio is calculated as Cases_pass divided by the total number of cases under the corresponding scenario.
3. Cases_top3 ratio is calculated as Cases_top3 divided by the total number of cases under the corresponding scenario.

6.5 Discussions and applications

From the aforementioned analysis, the Mixture Models distribution can provide superior fitting performance than its alternatives. More than one mixture models with different distribution components (e.g. lognormal, Gamma or Log logistic) could be tested besides GMM, but not essential. The common limitation of empirical studies is that the findings are largely influenced by the data used and it may be not easy to generalize. Considering the diverse operating environments and the complete set of cases tested in the paper, the identified GMM model could be transferred to fit TTDs on other bus service routes to a large extent. Examination of more routes with different operating environments (e.g. rural area) using the methodology proposed here could further complement this research. However, like any other empirical study, the conclusions are valid within the range of the used data and should be used with caution beyond this range. The transferability of the GMM model to fit distributions of travel times in a more generalized manner is discussed below from perspectives of its mathematical characteristics, explanatory power and practical application.

1) Mathematical characteristics

The GMM is a special type of mixture models with Gaussian component distribution. GMM is flexible enough to fit a large range of distributions, by changing the mixture coefficients and component distributions. In a general sense, the basic shapes of distribution can be classified into symmetric, skewed and multimodal categories. A distribution in practical (e.g. TTD) could be regarded as the combination of these basic shapes. The ability of GMM in fitting these distributions is examined by visualizing and comparing with the candidate models.

Figure 6-6 shows that the AM peak travel times have an obvious bimodal distribution with two peaks. The first peak is relatively symmetric while the second peak is right skewed with a short tail. Under such a multimodal distribution case, the GMM model can properly capture the peaks of the TTD as well as the short tail in the second peak, which can be observed from the density and cumulative probability graphs. The Burr model is powerful in capturing the first peak and the second peak tail, but it fails to capture the second peak. Other models cannot capture any peak in the AM peak travel times.

Figure 6-7 shows that the Inter peak travel times have a skewed distribution with a long right tail (skewness/se = 31.7). Also the distribution is rather peaked than peak period travel times since the kurtosis/se is very large (kurtosis/se = 134.7). Under such a largely skewed distribution, the GMM model properly captures the peak of the TTD as well as the long tail on the right, while the Burr distribution can capture the peak well but fails to capture the long tail. Other models can generally fit the peak location well but could not fit the long tail. Figure 6-8 shows the AD test significance for distributions of hourly travel times over the whole day.

The result shows that the GMM model can pass the AD test with considerable significance values for travel times at different time periods of a day. In peak time periods (e.g. 08:00, 09:00 and 16:00), the GMM model can still perform well while other distribution models cease to fit the travel times, which further verifies the accuracy and robustness of the GMM model.

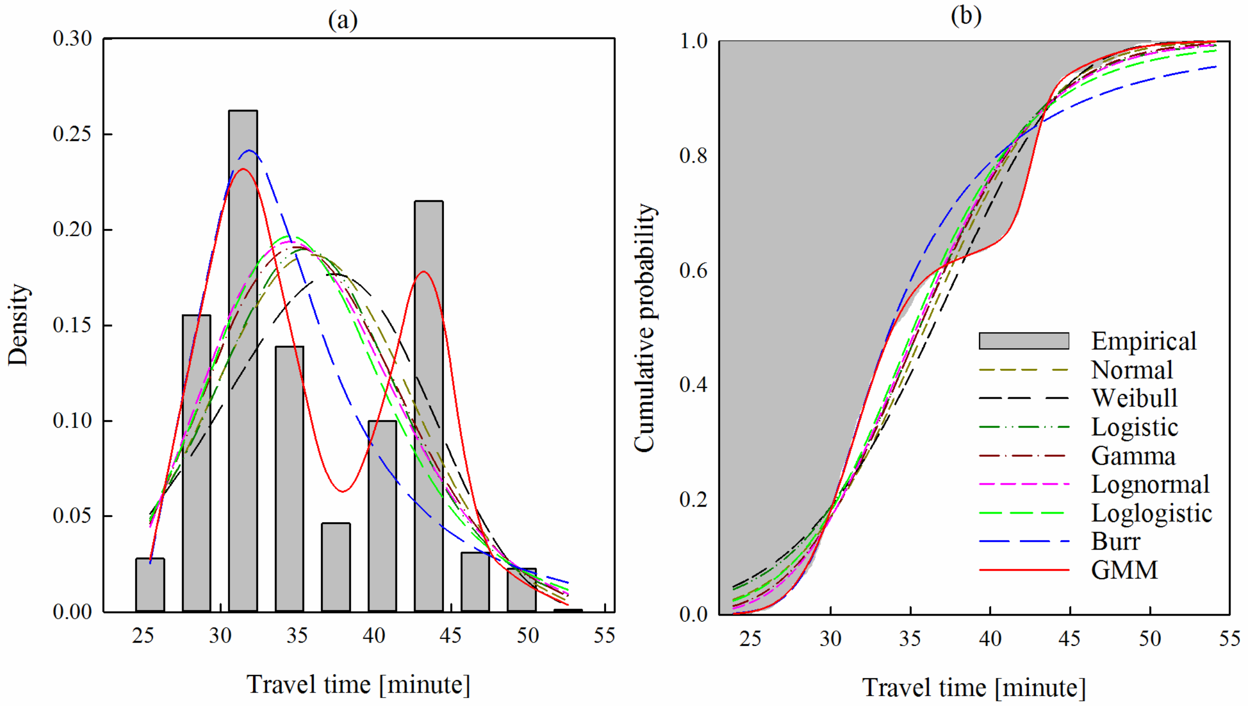

Figure 6-6: Fitting results for a multimodality distribution (a) density and (b) cumulative probability. (Case: urban route, weekdays, eastbound, AM peak, travel time). GMM, Gaussian mixture models.

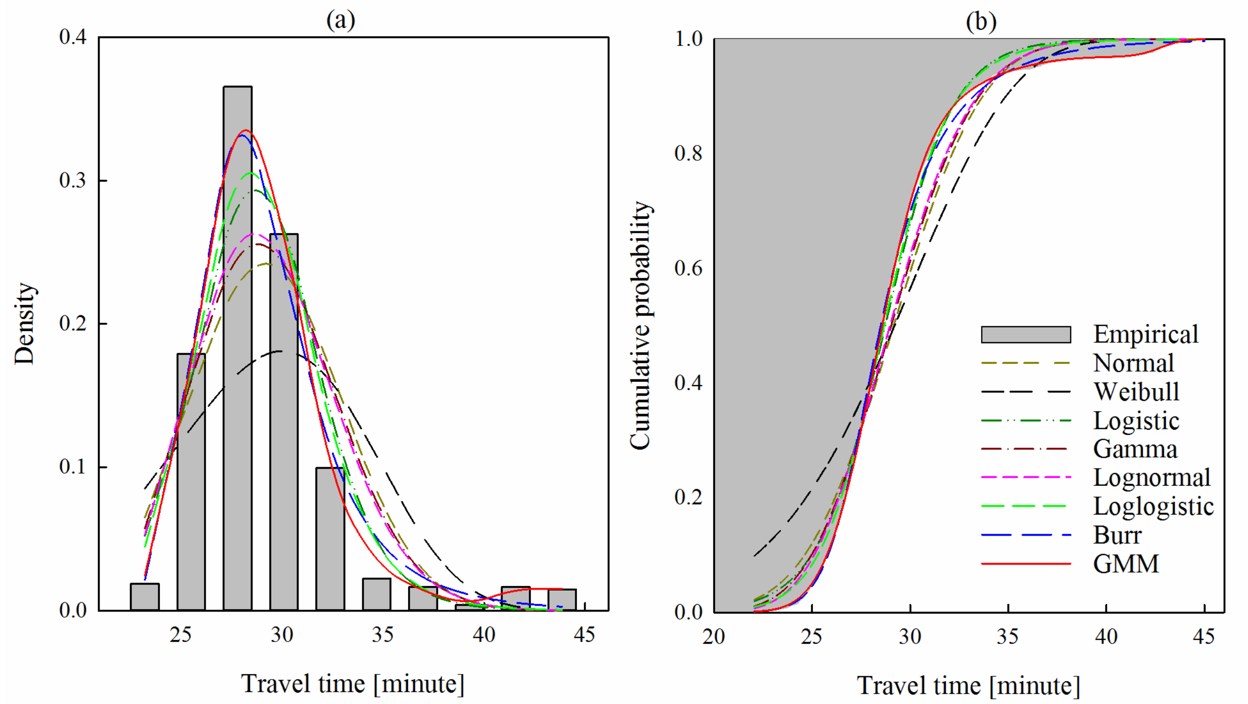

Figure 6-7: Fitting results for an asymmetric distribution (a) density and (b) cumulative probability. (Case: busway route, weekdays, inbound, inter peak, travel time). GMM, Gaussian mixture models.

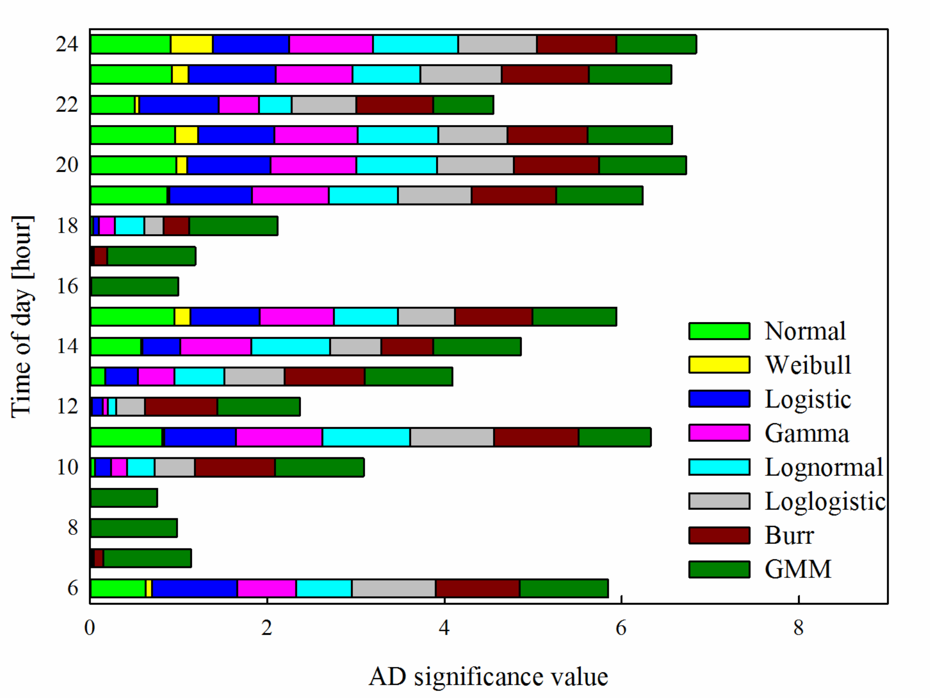

Figure 6-8: Anderson–Darling (AD) test significance for distributions of hourly travel times over the whole day. (Case: urban route, weekdays, eastbound, hourly, travel time).

2) Explanatory power

From a mathematical perspective, by changing the component distributions (e.g. Normal, Lognormal or Gamma) and the mixture coefficients, a mixture models is flexible enough to approximate a large range of different distributions. From a practical perspective, Guo et al. (2010) conducted a simulation and empirical study on freeway TTD and claimed the connection between GMM model parameters and the underlying traffic states. The premise of the GMM model is that travel times are dominated by complex stochastic traffic states rather than deterministic ones. Different travel time states could exist for a given time period, such as free flow and congested states. Two levels of uncertainty can be quantitatively assessed in the GMM model, namely concurrency probability of the state (mixture coefficient) and travel time variability under such a state (component distribution).

Compared to freeway travel times, bus travel times are more complicated and are mainly dominated by traffic flow, passenger demand and operational management (e.g. frequency, time schedule and time point). For a bus service travel time, different states may exist given the spatio-temporal aggregation levels, such as high speed service state, medium speed service state and low speed service state. The high speed and medium speed service states belong to a recurrent service state and a trip under the former state could experience relatively less total stops and intersection delays than that under the latter state. The slow speed service state is impacted by unexpected incidents or bad weather conditions. And the combined influence of other factors (delay from last time point, load in vehicle, drivers’ behaviour and et al.) could also contribute to different service states, even within a short aggregated time periods (7:00-7:15) across different days.

In practice, the mixture coefficient can be interpreted as the probability that travel times under a state (e.g. fast service state) and the component distribution indicates the distribution of travel times under such a state. The component in the mixture models could be symmetric or skewed distributions, depending on the definition of traffic states. Guo et al. (2012) have further done a fitting performance comparison between symmetric and skewed mixed model by fixing the number of components to be two. They concluded the multistate lognormal model is the optimal model for modelling freeway travel time under moderate to heavy traffic conditions. However, no evidence has been found on the performance of the alternative models if changing the number of components to be three or more. Theoretically, GMM can fit the skewed distribution well by regarding the skewed travel times coming from two or more different traffic states, and it has been verified in Figure 6-7. And from the interpretation perspective, the GMM model could be more promising than the skewed mixture models considering the simple form of normal distribution component.

3) Practical application

TTD fitting is the preliminary preparation for reliability analysis. The GMM model can provide much detailed travel time information for both management agencies and individual travellers. For agencies, the GMM model provides a flexible and superior distribution fitting than its alternatives, which enables accurate and effective assessment of the reliability performance of the system. Since distribution can provide the maximum information for reliability analysis, the improved statistical fitting can better support reliability analysis, especially considering passengers' different perspective on travel times under different service states. Also, Chapter 4 investigated the current reliability measures performance and concluded the shape of distribution plays a key role in service assessment. Moreover, the GMM model makes it possible to analyse travel time reliability and unreliability causes in a detailed disaggregated level under different states, and thus help policy makers to establish effective measures to improve reliability performance. For travellers, the GMM model enables a report of the reliability information analogy to a weather report which should be easily accepted by the general public and help passengers to plan their trip wisely (Guo et al., 2010). For example, for the AM peak travel, the probability of experiencing a fast service is 20%, and if that happens, the expected travel time for this trip would be 30 min.

6.6 Summary

The research focuses on the specification of distributions for day-to-day variability of bus travel times. The spatiotemporal data aggregation influence on distribution was investigated using six months AVL data on two typical service routes in Brisbane. The performance of alternative distribution models were examined under different cases considering fitting accuracy, robustness and explanatory power.

Consistent with previous studies, the decrease of temporal aggregation level results in a less asymmetric and flat distribution, and an increase of the normality of the distribution. The link level TTDs are more complicated than the route level TTDs, since the travel times of the former are more sensitive to intersection and stop delays. The spatial aggregation of link travel times breaks up the multimodality distribution for the busway service while it is not applicable for the non-busway service. The reason may be that the drivers have relatively more flexibility to adjust speed to catch up with schedules on a busway route. It is clear that the temporal-spatial aggregation of travel times could alter the hidden features of TTDs, and ultimately affect reliability analysis results. Better selecting the appropriate data aggregation level before reliability analysis needs further investigation.

The GMM model is evaluated as superior to its alternatives under different cases in terms of fitting accuracy, robustness and explanatory power. Under almost 95% cases, the GMM model provides an AD significance value larger than 0.7, which highlights its accurate and robust fitting performance. Its parameters can be connected to different states service performance, which is useful for identifying unreliability causes and reporting reliability information. The Burr model provides almost the same accurate fitting performance as GMM model in premise that it can converge to a solution which has a powerful ability in modelling extremely long tails of a distribution. However, the high ratio of Burr model failure to converge would largely decrease its usefulness in application. The Normal, Lognormal, Logistic, Loglogistic, and Gamma models have a relatively similar performance under the route level scenario and an intertwined performance under the link level scenario. The Weibull model has the worst performance under both scenarios.

Though constrained by the empirical data tested, the reported GMM distribution model remains promising for fitting travel times for other services with different operational environments. Mathematically, it is flexible enough to model different types of TTDs by changing the component distribution model and component numbers, including symmetric, asymmetric and multimodal distributions. A major limitation of GMM model is its lack of robustness to outliers, since the maximization of the likelihood function under an assumed Gaussian distribution is equivalent to finding the least-square solution. In the Bayesian model selection context, the presence of outliers often increases the number of mixture components employed in the model. Another limitation of the GMM model is its instability for each run of the algorithm, due to random initialization of the parameters, small sample size and inadequate number of components (B.-J. Park et al., 2010; Yildirimoglu and Geroliminis, 2013). It is important to properly clean the data and determine the optimal number of components for GMM model in practice. These limitations are further discussed and addressed in Chapter 7.

Chapter 7 Trip Travel Time Distribution Estimation

7.1 Introduction

Methods for the estimation of trip travel times between origination-destination pairs using the increasingly available data from mobile sources are still evolving and rather limited, especially in the context of probability distribution estimation. Previous studies on trip TTD estimation used a Markov chain methodology (Timothy Hunter et al., 2013; Ramezani and Geroliminis, 2012; Yeon et al., 2008) and are based on a number of important assumptions: conditional independence between link travel times (e.g. independent conditional on states); and constant transition probabilities for a given time period (e.g. 7:00-7:15am). However, empirical evidence suggests that the conditional independence assumption is not always appropriate. Furthermore, previous studies use constant transition probabilities for different environmental conditions and estimate them from empirical counts of transitions. This constraints their ability to generalize to a large range of applications. The research proposes a generalized Markov chain (GMC) approach for estimation of trip TTDs between arbitrary OD pairs at arbitrary times from link or segment TTDs. The research findings are reported in a journal paper (under review) in Ma, Koutsopoulos, Ferreira and Mahmoud (2015).

The remainder of the chapter is organized as follows: Section 7.2 defines the research problem. In Section 7.3, the framework for the distribution estimation is proposed, followed by the detailed methodology presented in Section 7.4. To address the methodological gaps, the Markov path TTD is approximated as a sum of correlated distributions using a moment generating function algorithm. The transition probabilities are estimated using a logit model formulation with the utilities being a function of explanatory covariates (link characteristic and trip conditions), as opposed from observation counts. The proposed approach is demonstrated in a case study for transit trip TTD estimation using AVL data that can provide both link and ground-truth trip TTDs in Section 7.5. The implementation of the proposed approach in transit is demonstrated in Section 7.6. Finally, Section 7.7 summarizes the main conclusions and highlights future research.

7.2 Problem statement

A road *segment* is a directed edge between two adjacent vertices (e.g. intersections) that is associated with edge identification (id), a starting point, an ending point and a set of intermediate points that describe the road segment using polyline.

A road network is a directed graph $G(V,E)$, where V is a set of vertices representing the terminal points of the road segments, and E is a set of edges representing road segments. A road *link* is a set of connected road segments between two adjacent setting points $link: e_1 \rightarrow e_2 \rightarrow \cdots \rightarrow e_I$, where e_i is road segment i and I is the number of road segments. These setting points may represent geometrical separators or sensor locations, e.g. signals, ANPR, bus stops, loop detectors, etc. A *trip* is a sequence of connected road segments or links traversed between an OD pair (O,D) in a road network. These are illustrated in Figure 7-1.

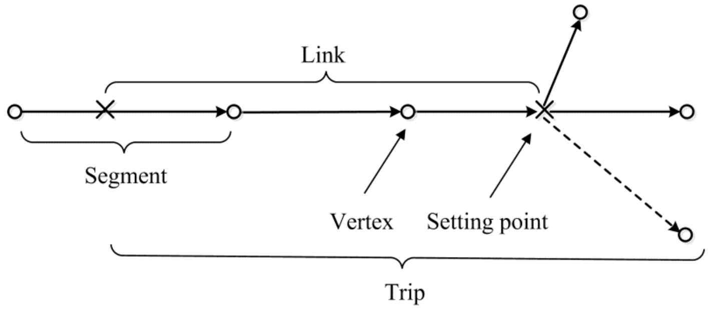

Figure 7-1: Illustration of network components: segments, links, and trips

An observation log on link l_n at time t consists of a link id lid, a vehicle id vid, distance traversed $d_{n,t}$, arrival timestamp t_n and travel time $T_{n,t}$, $obs(l_{n,t}) = (lid, vid, d_{n,t}, t_n, T_{n,t})$. This data can be obtained from fixed location sensor systems, for example ANPR, AVL, or loop detectors with vehicle re-identification (Coifman and Kim, 2009; Coifman and Krishnamurthy, 2007). Another promising source is from position-enabled fleets (Floating Car Data, FCD). In such systems, the two consecutive polled positions do not necessarily correspond to the starting and ending points of links. For high-frequency GPS measurements (e.g. every 5 seconds), the traversal times can be easily allocated on individual links (Timothy Hunter et al., 2013). For low-frequency measurements, many studies successfully decompose the traversal time to individual road links using a hybrid approach of physical and data-driven models (Hellinga et al., 2008; Hofleitner,Herring and Bayen, 2012; Rahmani et al., 2015; Zheng and Van Zuylen, 2013).

The problem of probability distribution estimation of trip travel times is defined as:

> *Given a set of probability distributions of link travel times, estimate the probability distribution of trip travel times for an arbitrary OD pair at an arbitrary time.*

Let $\{Dist(l_1,t), Dist(l_2,t), \ldots, Dist(l_N,t)\}$ denote link TTDs at time t. The simplest model for estimation of the trip TTD $Dist(trip_{od},t)$ at time t is by convoluting link TTDs.

$$Dist(trip_{od},t) = Dist(l_1,t) \otimes Dist(l_2,t) \otimes \ldots \otimes Dist(l_N,t) \tag{7.1}$$

where operator ($\otimes$) expresses convolution, and N is the number of links for a trip.

The convolution method postulates the independence among link travel times without considering correlation information. Markov chains have been successfully applied to estimate route travel times from link travel times on freeways and on arterials (Timothy Hunter et al., 2013; Ramezani and Geroliminis, 2012; Yeon et al., 2008). Markov chains assume a memory-less random process with transitions from one state to another among a finite number of states. Given a sequence of random variables $\left\{X_n\right\} = X_1, X_2, ..., X_N$, the conditional probability of the system moving to the next state x_{n+1} depends only on the current state x_n .

$$Prob\left(X_{n+1} = x_{n+1} \mid X_1 = x_1, X_2 = x_2, ..., X_n = x_n\right) = Prob\left(X_{n+1} = x_{n+1} \mid X_n = x_n\right) \tag{7.2}$$

Similar to a Markov chain process, the travel time of a vehicle on the current link depends only on the travel time on the previous link. Different from the definition of traffic propagation, the spatial Markov state progression designates how a vehicle experiences a series of travel time states on links along a route (Ramezani and Geroliminis, 2012). The state transition probabilities capture the combined influence of static and dynamic factors from the current and neighbour links at different temporal lags (e.g. upstream and downstream traffic conditions from preceding time periods, link geometric configurations, etc.). In previous studies, the main assumptions on TTD estimation using a Markov chain methodology are:

1. Conditional independence between link travel times (independent conditional on states);
2. Constant transition probabilities for a given time period (e.g. 7:00-7:15am).

The conditional independence assumption is illustrated in Figure 7-2(a). The grey rectangle represents the observations of travel times on adjacent links for the same vehicles. The red line approximates the linear relationship between adjacent link travel time observations. The increasing trend indicates a significant and positive correlation between adjacent link travel times. The correlation can be ideally eliminated when the observations are conditioned on link states as shown in dark grey groups (e.g. F-F stands for the group of vehicles that experience fast speed on both current link and next link). However, empirical evidence suggests that the conditional independence assumption is not always appropriate. The conditional observations could still be correlated as illustrated by the green-bound light-grey rectangle in Figure 2b. Therefore, such correlation should be incorporated in the model formulation.

Previous studies assume separate transition probabilities under different environmental conditions and estimate them from empirical counts of transitions. Alternatively, the transition probabilities can be estimated as a function of explanatory covariates with model parameters calibrated using historical data. This approach is more general and less sensitive to data availability, e.g. inadequate number of observations.

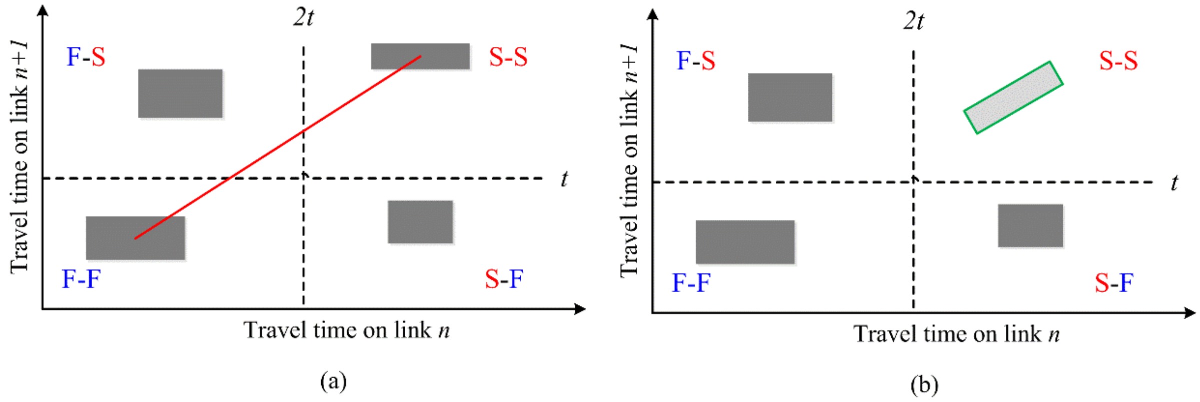

Figure 7-2: Two-dimensional diagram representing four groups of vehicles (F = Fast and S = Slow): (a) ideally uncorrelated observations within groups, (b) potentially correlated observations within groups.

7.3 Estimation framework

The framework of our proposed trip TTD estimation approach is shown in Figure 7-3. It is composed of two major parts: Markov Chain Identification, and Probability Distribution Estimation. The database provides information of link travel times, link characteristics and trip conditions.

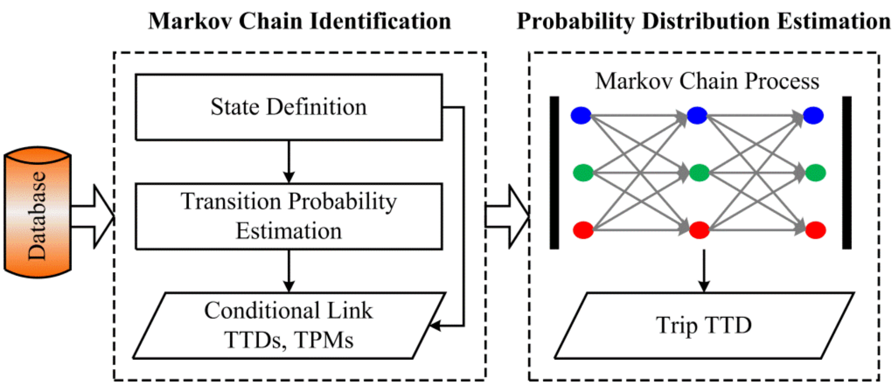

Figure 7-3: Trip travel time distribution estimation framework

Markov Chain Identification: this step aims to define traffic states and estimate the transition probability model. The outputs are probabilities of link traffic states, link TTDs (conditional on states), and time-space dependent transition probabilities.

- The state definition is performed using a Gaussian Mixture model (GMM) based clustering algorithm. The approach ensures homogeneity within each cluster; differentiation over space and time; large enough state that can characterize the underlying traffic conditions; and computational efficiency.
- The transition probabilities are estimated using a logit model formulation with the utilities being a function of explanatory covariates. These covariates include link characteristics, traffic conditions on neighbour links and from preceding time intervals.

Probability Distribution Estimation: this estimates the trip TTD using a Markov chain process. Markov paths are permutations of link states along a trip. The Markov path probability is estimated as the product of initial state probabilities and transition probabilities between links. The Markov path TTDs are estimated as the sum of correlated link TTDs conditional on states using a MGF approach. Finally, the trip TTD is estimated as the sum of Markov path TTDs weighted by their occurrence probabilities.

7.4 Methodology

7.4.1 State definition

A heuristic clustering algorithm based on GMM is developed to define states. The GMM approach identifies homogeneous clusters within which the observations are normally distributed. The optimal number of clusters can be determined by checking the accuracy and stability of the clustering results for different cluster numbers. For example, there could be one state for a link in a residential area, but there could be two or three states in a CBD area depending on link configuration and time-of-day. Following the GMM clustering, an additional step is proposed to convert the incorrect clustering outcomes. The Silhouette widths are examined and observations are reassigned if their Silhouette widths are negative. In addition, proportions of clusters are checked and the cluster with a proportion less than a predefined threshold is merged with its nearest cluster. Hence, an unreasonable large number of states can be avoided and the identified states are large enough to represent the underlying traffic conditions.

The optimal number of clusters is determined by fitting a set of mixture models with increasing number of clusters and using accuracy measures, including average silhouette width and information measures such as the Akaike Information Criterion (AIC). In addition to accuracy, the clustering method should return the same results when it is repeated several times to guarantee its stability. Two reasons could cause unstable GMM clustering results, namely random initialization and artificial cuts.

The EM algorithm may converge to a local maximum and hence its estimation performance may vary with randomly initialized parameters (B.-J. Park et al., 2010). The *k*-means++ algorithm can be used to identify initial GMM component means (Arthur and Vassilvitskii, 2007). *k*-means++ selects each centroid with a probability proportional to the distance from itself to the closest centre that has been already chosen. It has been used in the literature to find centroid seeds for the *k*-means algorithm. For each observation $i = 1,\ldots,N$ and already chosen centroid $c = 1,\ldots,k-1$, the new centroid k is chosen from all observations $\mathbf{Y}$ with a probability:

$$D\left(y_i, \mu_c\right) \Big/ \sum_{j, y_j \in X_c} D\left(y_j, \mu_c\right) \tag{7.3}$$

where,

$D(y_i, \mu_c)$ = the distance between observation i and centroid μ_c, and Y_c is the set of observations closest to centroid μ_c and y_i belongs to Y_c.

Artificial cuts relate to the fact that when some natural clusters may be classified in an artificial way to satisfy the given cluster number and thus can cause instability (Yildirimoglu and Geroliminis, 2013). The silhouette width (SW) for a point measures how similar the point is to points in its own cluster, compared to points in other clusters.

$$SW(i) = \frac{b(i) - a(i)}{\max\left\{a(i), b(i)\right\}} \tag{7.4}$$

where,

$a(i)$ = the average dissimilarity of i with all other points within the same cluster;

$b(i)$ = the lowest average dissimilarity of i to other clusters.

Distance metrics are usually used to measure dissimilarity, such as Euclidean, Mahalanobis distance, etc. The Mahalanobis distance is preferred as it takes into account the covariance within clusters. The average silhouette width (ASW) over all data is a measure of the overall clustering quality. The mean and standard deviation of the ASW measures can be used to assess the stability of the clustering results and determine the optimal number of clusters (Yildirimoglu and Geroliminis, 2013). The proposed heuristic algorithm (GMMS clustering algorithm) uses the GMM initial clustering results and the Silhouette values to determine the optimal number of clusters. The GMMS algorithm is described in Figure 7-4.

Initialization:

1. Set the maximum number of states ρ, proportion threshold α and replications r_1, r_2.

GMM clustering:

2. For $n = 1, 2, 3, \ldots, \rho$
   2.1 For $m = 1, 2, 3, \ldots, r_1$
      a. Initialize the GMM parameters using the *k*-means++ algorithm.
      b. Fit observations using the GMM algorithm with clusters n.
      c. Perform r_2 replications and select the GMM model with the smallest AIC.
      d. Cluster using posterior probability and calculate ASW.
   2.2 Calculate the mean and standard deviation of ASWs.

Posterior refinement:

3. Select the optimal number of states as the one that gives the larger mean and smaller standard deviation of ASWs, and return the component mean values $\mu_1, \mu_2, \ldots$ in ascending order and proportions $w_1, w_2, \ldots$.
4. If the cluster proportion w_i is smaller than α, merge this cluster to the nearest one.
5. Calculate the Silhouette width for each observation. If it is negative, re-assign the observation to a cluster that gives the largest silhouette value.

Outputs:

6. Return the number of states and state identification for all observations.

Figure 7-4: The description of GMMS clustering algorithm

7.4.2 Transition probabilities estimation

Let $X_n(t)$ denote the state on link n at time t. The state space $Q_n(t)$ on link n at time t is a set of values that $X_n(t)$ may take, $Q_n(t)=\left\{x_n^1(t),x_n^2(t),\ldots,x_n^{m_{n,t}}(t)\right\}$, where $x_n^i(t)$ is the state i on link n at time t, and $m_{n,t}$ is the number of states on link n at time t. Note that the state space $Q_n(t)$ vary across both links and time periods.

Count-based method

The majority of existing methods estimate the initial state probabilities, as well as the transition probabilities using available observations and calculating the corresponding frequencies. The initial state probabilities $\boldsymbol{\pi}(t)$ at time t are the probabilities of different states in $Q_1(t)$ on the first link which can be estimated by:

$$\hat{\boldsymbol{\pi}}(t)=\begin{bmatrix}\hat{\pi}_1^1(t)\\ \hat{\pi}_1^2(t)\\ \vdots \\ \hat{\pi}_1^{m_{1,t}}(t)\end{bmatrix}=\begin{bmatrix} Num\left(X_1(t)=x_1^1(t)\right)\Big/\sum_{k=1}^{m_{1,t}}Num\left(X_1(t)=x_1^k(t)\right)\\ Num\left(X_1(t)=x_1^2(t)\right)\Big/\sum_{k=1}^{m_{1,t}}Num\left(X_1(t)=x_1^k(t)\right)\\ \vdots \\ Num\left(X_1(t)=x_1^{m_{1,t}}(t)\right)\Big/\sum_{k=1}^{m_{1,t}}Num\left(X_1(t)=x_1^k(t)\right)\end{bmatrix} \tag{7.5}$$

where,

$\hat{\boldsymbol{\pi}}(t)$= an estimate of $\boldsymbol{\pi}(t)$;

$Num\left(X_1(t)=x_1^k(t)\right)$ = the number of observations of state k on link 1 at time t.

The TPM $\mathbf{P}_{n-1,n}(t)$ between a pair of successive links $n-1$ and n at time t can be represented as:

$$\mathbf{P}_{n-1,n}(t)=\begin{bmatrix} p_{1,1}(t) & \cdots & p_{1,m_{n,t}}(t)\\ \vdots & & \vdots \\ p_{m_{n-1,t},1}(t) & \cdots & p_{m_{n-1,t},m_{n,t}}(t)\end{bmatrix} \tag{7.6}$$

where,

$p_{i,j}(t)$ = the transition probability from state i to state j at time t between two successive links, with $\sum_{j=1}^{m_{n+1,t}}p_{i,j}(t)=1$.

Given a set of observations, the transition probabilities $p_{i,j}(t)$ can be estimated by:

$$\begin{aligned}\hat{p}_{i,j}(t)&=Prob\left\{X_n(t)=x_n^j(t)\mid X_{n-1}(t)=x_{n-1}^i(t)\right\}\\ &=\frac{Num\left(X_{n-1}(t)=x_{n-1}^i(t),X_n(t)=x_n^j(t)\right)}{\sum_{k=1}^{m_{n,t}}Num\left(X_{n-1}(t)=x_{n-1}^i(t),X_n(t)=x_n^k(t)\right)}\end{aligned} \tag{7.7}$$

where,

$\hat{p}_{i,j}(t)$ = the estimation of $p_{i,j}(t)$;

$Num\left(X_{n-1}(t)=x_{n-1}^i(t),X_n(t)=x_n^j(t)\right)$ = the number of observations of vehicles that encounter state i on link $n-1$ and state j on link n at time t.

Logit-based model

To capture the heterogeneous and dynamic link travel time correlations, a logit model is proposed that estimates transition probabilities as a function of link characteristics and trip conditions. To avoid inconsistent dimension of TPMs between different links at different times, an optimized number of clusters across all links is used. For convenience, the notation t is excluded in the following discussion.

Let the dimension of TPMs be $M \times M$. Let $p_{i,j}^{k}$ denote the probability that vehicle k experiences state j on link n, conditioned on having state i on link $n-1$. The transition probabilities vary with link characteristics and trip conditions. Let $\mathbf{Z}_k$ be a vector of the explanatory variables for vehicle k. The utility $V_{i,j}^{k}$ for vehicle k moving from state i to j can be expressed as:

$$V_{i,j}^{k} = \boldsymbol{\beta}_{ij}\mathbf{Z}_k + \varepsilon_{ij} \tag{7.8}$$

where,

$\boldsymbol{\beta}_{ij}$ = a vector of parameters and ε_{ij} is the error term.

Assuming the error terms are independently and identically distributed, the transition probability for vehicle k from state i on link $n-1$ to state j on link n is given by the logit model:

$$p_{i,j}^{k} = e^{V_{i,j}^{k}} \Big/ \sum_{j=1}^{M} e^{V_{i,j}^{k}} \tag{7.9}$$

The coefficients are estimated relative to a reference state (e.g. low speed state) using maximum likelihood. Since the sum of each row elements of TPM equals to 1, M logit models should be built conditional on the state of the previous link. The coefficients of these M conditional logit models are estimated separately (Madanat et al., 1995). Accordingly, M datasets need to be generated conditional on the state of the previous link. For each case, the dependent variable is the state (1, 2,…, M) and the independent variables are related to explanatory factors that influence the transition behaviour. For example, suppose that the number of states across all links is $M = 3$. Three datasets, namely, dataset 1 conditional on state 1 (fast speed) of the previous link, dataset 2 conditional on state 2 (medium speed) of the previous link, and dataset 3 conditional on state 3 (low speed) of the previous link, are used for model estimation.

7.4.3 Probability distribution estimation

Figure 7-5 illustrates the Markov chain process of probability distribution estimation for a route (trip). Each circle represents the state that a vehicle may encounter along a route. The edge between adjacent link states represents the transition probability to the state a vehicle encounters on the current link given the state on the previous link. Markov paths are permutations of link states along a route. If the number of states for link n is m_n, there are $\prod_{n=1}^{N} m_n$ Markov paths. Each represents a different realization of a set of TTDs on links.

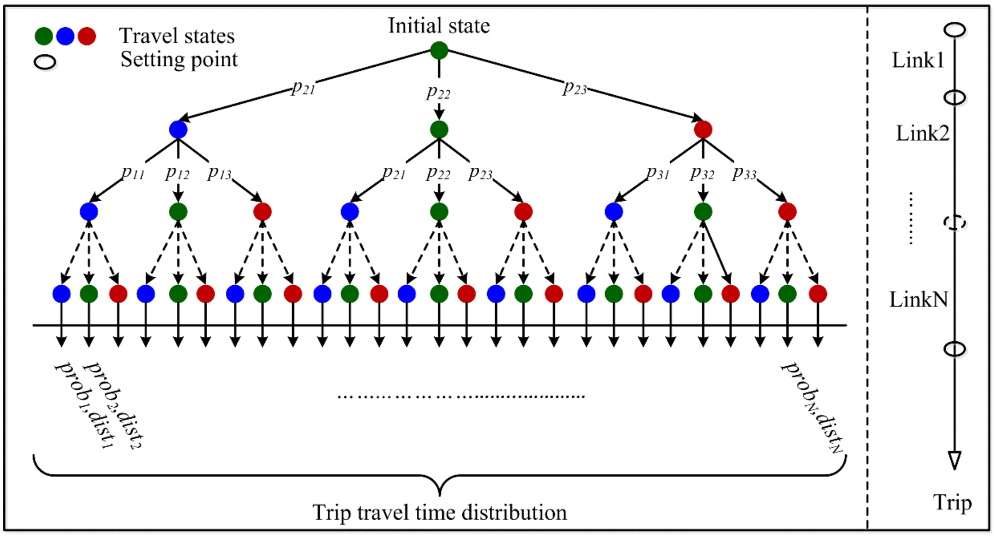

Figure 7-5: The Markov chain structure to estimate trip travel time distribution

7.4.3.1Generalized Markov Chain (GMC) approach

The probability of each Markov path is the product of the initial state probability and transition probabilities between adjacent links.

$$Prob\left(X_1 = x_1^{j_1}, X_2 = x_2^{j_2}, \cdots, X_N = x_N^{j_N}\right) = \pi_1^{j_1} \times p_{j_1, j_2} \times p_{j_2, j_3} \times \cdots \times p_{j_{N-1}, j_N} \tag{7.10}$$

The Markov path distribution is estimated as the sum of correlated conditional link TTDs using the MGF algorithm.

$$Dist\left(X_1 = x_1^{j_1}, X_2 = x_2^{j_2}, \cdots, X_N = x_N^{j_N}\right) = MGF\left\{Dist\left(x_1^{j_1}\right), Dist\left(x_2^{j_2}\right), \cdots, Dist\left(x_N^{j_N}\right)\right\} \tag{7.11}$$

where,

$Dist\left(x_n^{j_n}\right)$ = the probability distribution of travel times on link n conditional on the state of current link n;

$MGF\left\{\bullet\right\}$ = the moment generating function method to approximate the sum of the correlated distributions.

Finally, the probability distribution of trip travel times is estimated as the mixture of Markov path TTDs (Equation (17)) weighted by the corresponding probabilities (Equation (16)).

$$Dist\left(Trip\,travel\,time\right) = \sum_{q=1}^{Q}\left\{Prob\left(Markov\,path_q\right) \times Dist\left(Markov\,path_q\right)\right\} \tag{7.12}$$

Compared to the Markov chain approach proposed in Ramezani and Geroliminis (2012), the proposed approach takes into consideration of the correlations between link travel times conditional on states along a Markov path. It simplifies the calculation of link TTDs which is conditioned solely on the state of the current link (compared to conditioned on the states of upstream, current and downstream links), and makes it possible to model link TTDs and transition probabilities as a function of explanatory covariates, which are important from a practical point of view as it allows the general application of the model.

7.4.3.2 Moment Generating Function (MGF) Algorithm

MGF is an alternative approach to analytically work with pdfs. If two distributions have the same MGFs, then they have identical distributions. Srinivasan et al. (2014) used a MGF based approach to estimate the trip TTD as the sum of the correlated link TTDs assuming they have a unimodal distribution. However, the unimodal assumption is not always appropriate. It is more reasonable to assume a unimodal distribution (e.g. normal or log-normal) for link travel times along a Markov path since they are conditional on the underlying traffic states. In the approach proposed in this paper, the MGF method is adopted to approximate the sum of correlated random variables (RVs) along a Markov path by matching the MGF of the Markov path TTD with the MGF of the sum of the conditional link TTDs.

Let $\mathbf{X} = \left(X_1, X_2, \cdots, X_N\right)$ denote the link travel times conditional on states along a Markov path. The MGF of a vector of RVs $\mathbf{X}$ with a continuous joint distribution $f_{\mathbf{X}}\left(\mathbf{X}\right)$ can be written as:

$$M_{\mathbf{X}}\left(\mathbf{S}\right) = \mathrm{E}\left[\exp\left(-\mathbf{S}^T\mathbf{X}\right)\right] = \int_{-\infty}^{+\infty} \exp\left(-\mathbf{S}^T\mathbf{X}\right) f_{\mathbf{X}}\left(\mathbf{X}\right) d\mathbf{X} \tag{7.13}$$

where,

$\mathbf{S}^T$ = the transpose of $\mathbf{S}$, with $\mathbf{S} \in \mathbb{R}$.

Assuming that the conditional link travel times are normally distributed, the vector of RVs $\mathbf{X}$ follows approximately a multivariate normal (MVN) distribution (Mehta et al., 2007). The MVN distribution is given as:

$$f_{\mathbf{X}}\left(\mathbf{X}\right) = \frac{1}{\sqrt{\left(2\pi\right)^N \left|\mathbf{\Sigma}\right|}} \exp\left[-\frac{1}{2}\left(\mathbf{X}-\boldsymbol{\mu}\right)^{\mathrm{T}} \mathbf{\Sigma}^{-1}\left(\mathbf{X}-\boldsymbol{\mu}\right)\right] \tag{7.14}$$

where,

$\boldsymbol{\mu} = \left[\mathrm{E}\left(X_1\right), \mathrm{E}\left(X_2\right), \ldots, \mathrm{E}\left(X_N\right)\right]$ = the vector of mean values of the RVs $\mathbf{X} = \left(X_1, X_2, \cdots, X_N\right)$;

$\mathbf{\Sigma} = \left[\mathrm{Cov}\left(X_i, X_j\right)\right]$ = the covariance matrix and $\left|\mathbf{\Sigma}\right|$ its determinant.

The Markov path travel time $Y = \sum_{i=1}^{N} X_i$ is the sum of link travel times conditional on states. As the vector of RVs $\mathbf{X}$ follows the MVN distribution, any linear combination of its components is normally distributed, i.e. $\mathcal{N}\left(\mu_Y, \sigma_Y\right)$. The Markov path TTD parameters can be easily estimated as:

$$\mu_Y = \sum_{i=1}^{N} \mu_i, \ \sigma_Y^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} \Sigma_{ij} \tag{7.15}$$

where,

μ_i = the distribution mean on link i;

Σ_{ij} = the covariance between links i and j.

If the link travel times follow a lognormal distribution, the lognormal MGF does not have a general closed form. Mehta et al. (2007) applied a short Gauss-Hermite expansion to approximate it. Let RV Z follow a lognormal distribution:

$$f_Z\left(z\right) = \frac{1}{z\sigma_X\sqrt{\left(2\pi\right)}}\exp\left[-\frac{\left(\ln\left(z\right)-\mu_X\right)^2}{2\sigma_X^2}\right] \tag{7.16}$$

where,

μ_X, σ_X = mean and standard deviation of the normally distributed RV X, with $X = \ln\left(Z\right)$.

Then, the MGF of a lognormal distribution can be expressed by a series expansion based on the Gauss-Hermite integration.

$$\hat{M}_Z\left(s;\mu_X,\sigma_X\right) \triangleq \sum_{k=1}^{K}\frac{\omega_k}{\sqrt{\pi}}\exp\left[-s\exp\left(\sqrt{2}\sigma_X a_k + \mu_X\right)\right] \tag{7.17}$$

where,

$\hat{M}_Z\left(s;\mu_X,\sigma_X\right)$ = the approximation function of the MGF of Z;

K = the Hermite integration order, ω_n and a_n are K specific parameters (see Abramowitz and Stegun (1964) for values).

Following the same procedure, the MGF of the sum of correlated lognormal RVs Z, $Y = \sum_{i=1}^{N} Z_i$, is:

$$\hat{M}_Y\left(s;\boldsymbol{\mu},\boldsymbol{\Sigma}\right) \triangleq \sum_{k_1=1}^{K}\cdots\sum_{k_N=1}^{K}\left[\prod_{i=1}^{N}\frac{\omega_{n_i}}{\sqrt{\pi}}\right]\times\exp\left(-s\sum_{i=1}^{N}\left[\exp\left(\sqrt{2}\sum_{j=1}^{N}\Sigma'_{ij}a_{n_j} + \mu_i\right)\right]\right) \tag{7.18}$$

where,

$\hat{M}_Y\left(s;\boldsymbol{\mu},\boldsymbol{\Sigma}\right)$ = the approximation function of the MGF of Y .;

$\boldsymbol{\Sigma}'_{ij}$ = the $(i, j)^{th}$ element of the square root of the covariance matrix $\boldsymbol{\Sigma}$.

Finally, the sum of lognormal RVs is approximated by a lognormal RV with parameters given by the following two equations:

$$\hat{M}_Z\left(s_i;\mu_X,\sigma_X\right) = \hat{M}_Y\left(s_i;\boldsymbol{\mu},\boldsymbol{\Sigma}\right), \text{at } i = 1 \text{ and } 2 \tag{7.19}$$

The covariance matrix $\boldsymbol{\Sigma}$ can be estimated from historical observations for any two links at the same time period. However, computing the full covariance matrix is not practical due to observation constraints for large networks (Timothy Hunter et al., 2013). Typically, the covariance matrix is calibrated assuming only adjacent (or second-order, etc.) link travel time correlations (Srinivasan et al., 2014).

7.5 Case Study

7.5.1 Trip TTD estimation

The space-time autocorrelation and cross-correlation functions were used to explore the correlation structure of transit travel time components in the data (Figure 7-6). The link running times were strongly correlated with their first order neighbour links and insignificantly correlated with high order ones (Figure 7-6a). The correlations between link running times and downstream dwell times were not significant (Figure 7-6b). Longer link running times did not necessarily lead to larger headways downstream. The dwell times of first order adjacent stops were significantly correlated with little decrease when the spatial order increases. This makes sense because of the relatively stable travel patterns on weekdays at peak hours. The partial correlation test after controlling for the effect of demand at stops (boarding and alighting), indicates an insignificant correlation between stop dwell times.

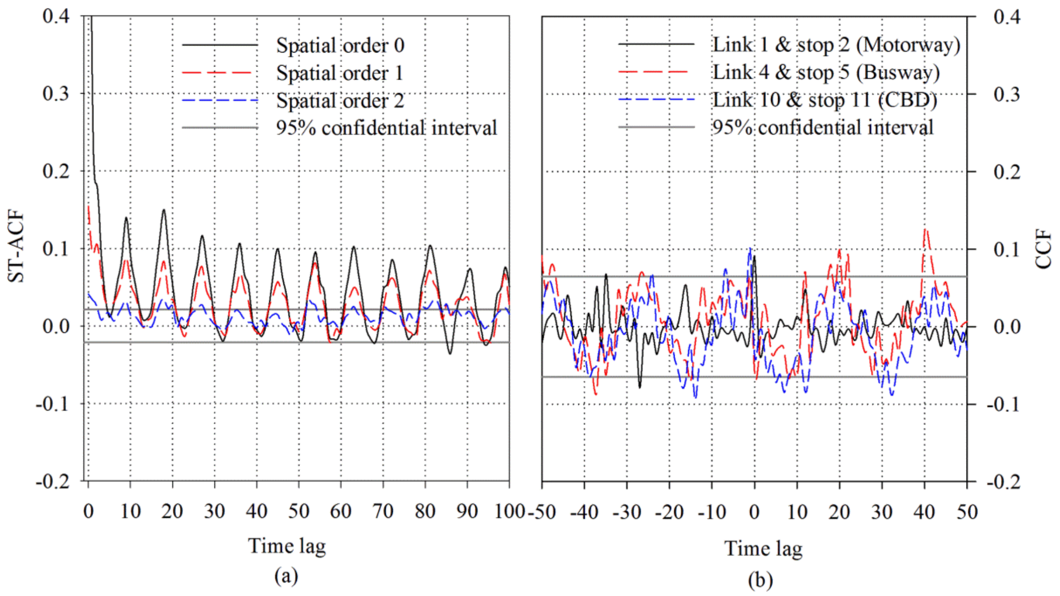

Figure 7-6: Global and local correlations of travel time components in AM peak period: (a) spatiotemporal autocorrelation function (ST-ACF) of unit running times between links with different spatial orders; and (b) cross-correlation function (CCF) between link unit running times and downstream stop dwell times on different types of roads.

Based on these results and in order to simplify the validation of the proposed method, dwell times are assumed to be independent for a specific time period. In this case, the probability distribution for transit trip travel times is estimated as shown in Figure 7-7. $f_i(rt_t)$ is the pdf of running times for link i at time period t and $s_i(dt_t)$ the pdf of dwell times for stop i at time period t. The trip running time distribution is estimated from link running time distributions $\{f_i(rt_t)\}$ using the proposed GMC approach. The trip dwell time distribution is estimated by convoluting stop dwell time distributions $\{s_i(dt_t)\}$. Finally, the trip TTD is estimated as the convolution of trip running time distribution and trip dwell time distribution.

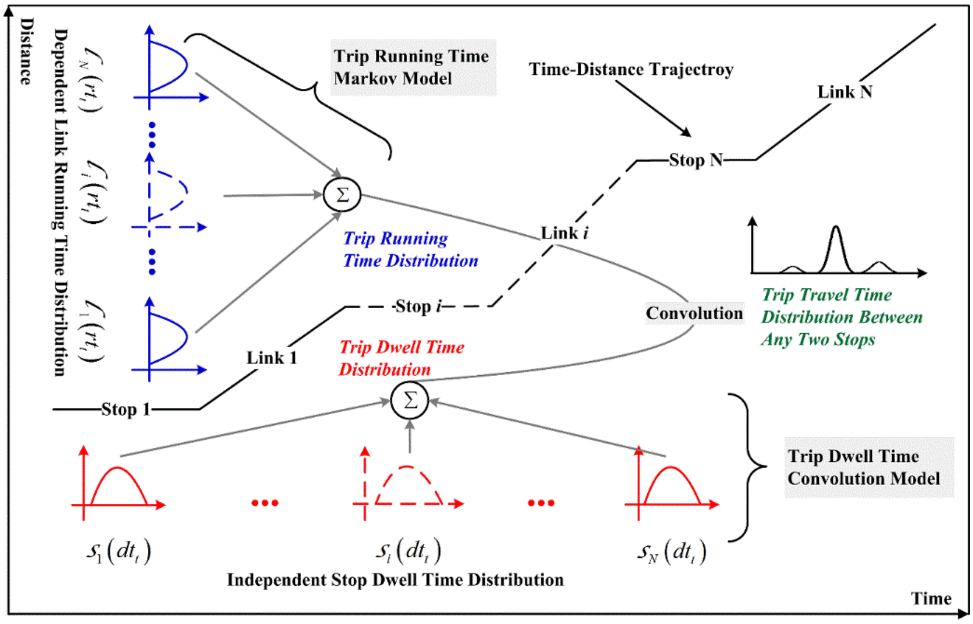

Figure 7-7: The estimation approach for transit trip travel time distribution

7.5.2 State definition

The GMMS algorithm parameters ρ , α , r_1 , r_2 in Figure 4 are selected to be 5, 0.001, 50 and 20, respectively. Figure 7-8 shows observations of unit running times. Each dot represents an observation of the unit running time of the same vehicle on two consecutive links. Group i_j indicates a vehicle experiencing state i on current link and state j on the next link. The results from the clustering steps are also illustrated in the figure by different colours. A standard GMM algorithm results in clusters with observations incorrectly clustered (rectangle area). The observations in the rectangle areas have negative silhouette widths. However, the application of the proposed GMMS algorithm re-assigns these observations to the clusters with the largest silhouette widths (left cluster). Different clusters on each link represent the underlying traffic conditions over different time periods of a day, since dwell times have been excluded from travel times.

In this case, the correlation between running times within each group is negligible. However, running times could potentially be correlated as illustrated in Figure 7-9 with observations from another bus route (route 60). The solid trend lines highlight the existence of significant correlations which cannot be neglected in the estimation of the Markov path running time distributions. Route 60 is a cross city route going through the CBD. Hence, these correlations could be caused by the combined effects of signals, stop activities, and driver behaviour.

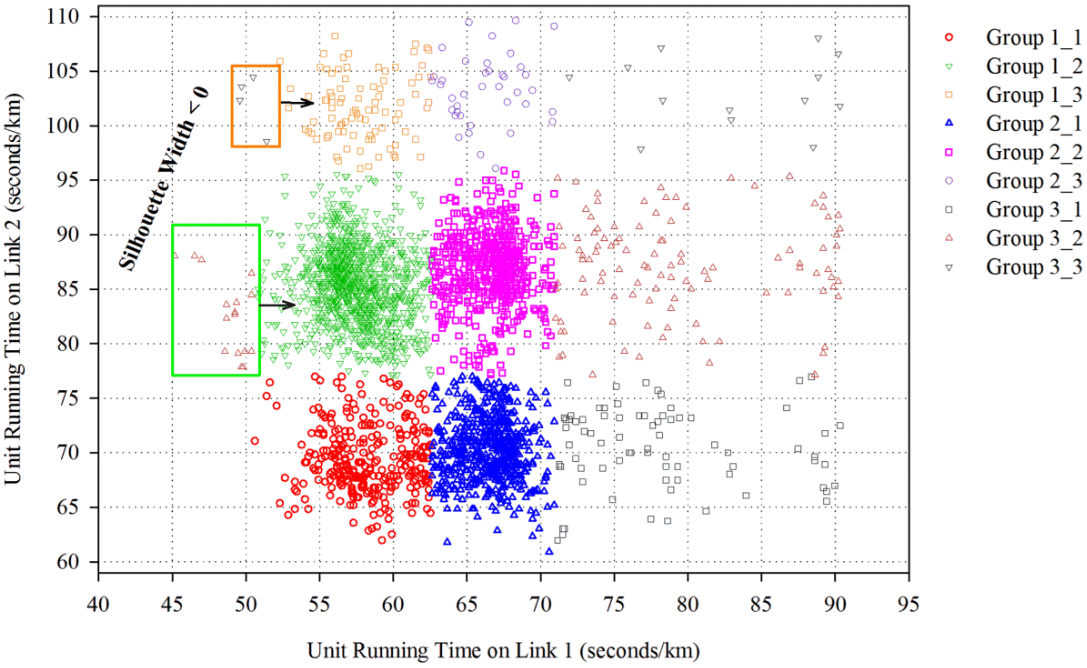

Figure 7-8: State clustering results [Route 555, Weekday, Inbound, All Day]

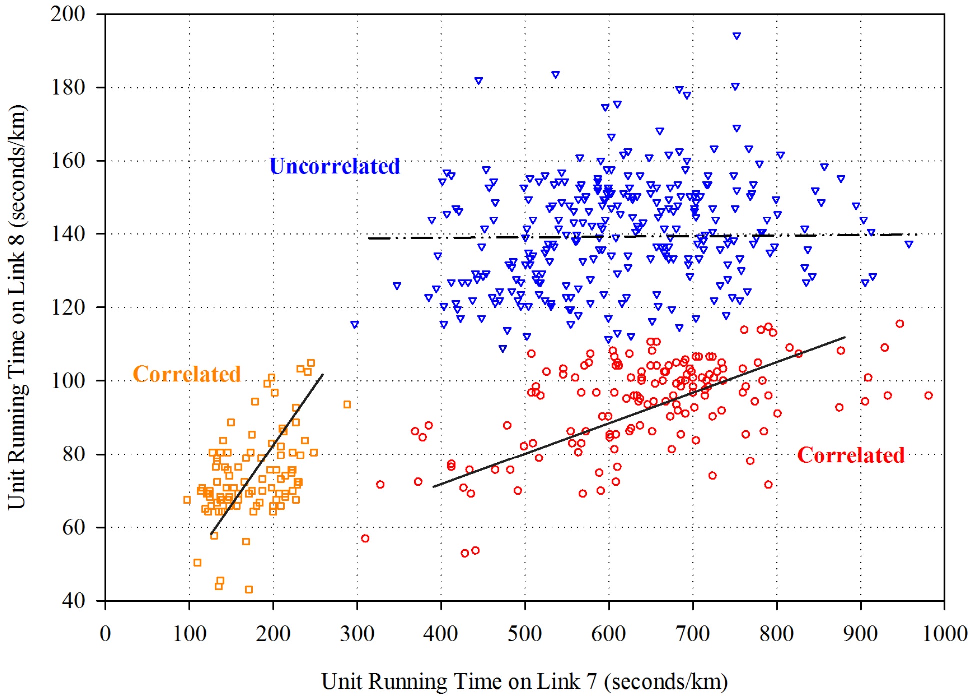

Figure 7-9: Samples of correlated and uncorrelated running times [Route 60, Weekdays, Eastbound, 7-8 AM]

Figure 7-10 shows the clustering results with different number of states for different temporal aggregation of travel times (intervals of 30 min, 1 hour, periods, and the whole day). The bars display the mean ASWs with error bars showing the average standard deviation of ASWs across all scenarios. The optimal number of states is 3 (where the clustering outcome is relatively accurate and stable).

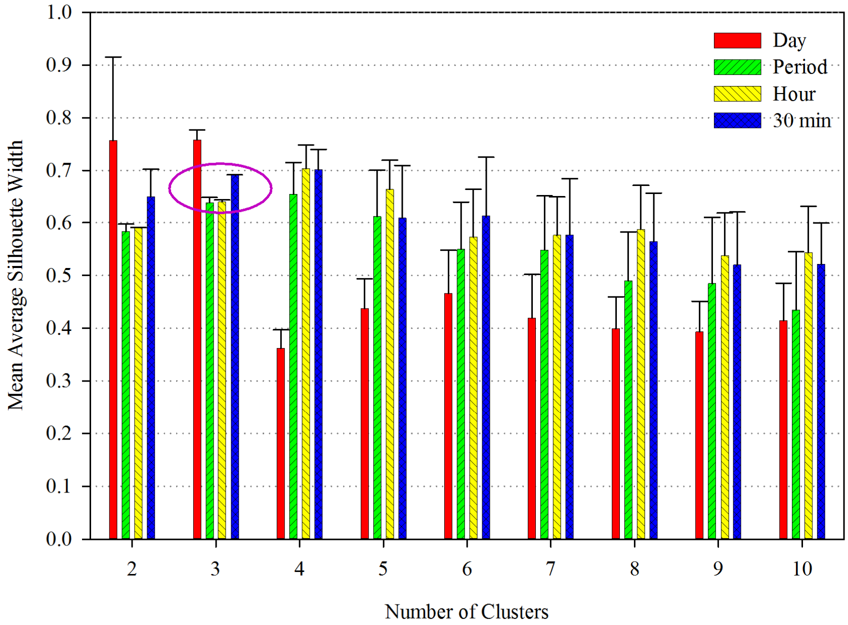

Figure 7-10: Mean and standard deviation of average silhouette width vs. number of clusters

7.5.3 Transition probability model

For the development of the transition probabilities model, in addition to AVL, data from other sources were used, including Smart Card Transactions, General Transit Feed Specification (GTFS), Brisbane Strategic Transport Management (BSTM), and Bureau of Meteorology (BoM) data. The data was partitioned into cases based on trip attributes (route id, direction, link id, and weekdays) in 30 minutes time intervals. Cases with sample size less than 150 were dropped as they are insufficient for multimodal distribution analysis. Finally, three separate datasets with a total of 113,426 cases were obtained, one for each state of the previous link (high, medium and low speed).

Table 7-1 summarizes the datasets and descriptive statistics of the associated variables. As expected, the percentage of cases in the high speed state on the current link decreases as the states on the previous link change from high to low. The recurrent congestion index (RCI) is defined as the mode speed over free flow speed (high values indicate good recurrent traffic condition, e.g. busway). It captures different characteristics of road links and within-day variation of traffic conditions. The free flow speed was derived from the running times observed between 5:30 am and 23:30 pm. The congestion index (CI) is calculated as speed over free flow speed (high value indicates a free flow traffic condition). The CI on current link at preceding time is calculated as the median CIs from the preceding 30 minutes. If observations within the preceding 30 minutes interval are not available, the RCI is used. The delay on the previous link is calculated as the actual arrival time at the stop minus the scheduled time. The actual stop indicates a vehicle skipping a stop or not by checking the difference between arrival and departure times.

Table 7-1: Summary of dataset and descriptive statistics of variables

Datasets*	Variables	Unit	Min	Max	Mean	Std
Given state H	CI on previous link at current time^	%	27.1	112.4	76.9	16.6
	CI on current link at preceding time^	%	5.29	107.2	69.1	18.7
Total: 44,503	RCI on current link^	%	12.0	88.5	67.9	18.3
H: 20,052 (45.1%)	Length of current link	km	0.38	8.53	2.04	1.73
M: 17,882 (40.2%)	Number of signals on current link	int	0.00	8.00	1.36	2.12
L: 6,569 (14.8%)	Delay on previous link	min	-9.38	14.5	1.55	4.16
	Actual stop on current link	(1,0)	0.00	1.00	0.80	0.40
	Passenger load on current link	int	0.00	86.0	15.5	12.4
	Precipitation per half hour	mm	0.00	1.80	0.05	0.23
	Route type on previous link#	[1-5]	na	na	na	na
	Route type on current link#	[1-5]	na	na	na	na
Given state M	CI on previous link at current time^	%	13.9	120.0	67.6	18.3
	CI on current link at preceding time ^	%	5.56	132.1	68.5	18.1
Total: 46,785	RCI on current link^	%	12.0	88.52	68.5	18.0
H: 15,902 (34.0%)	Length of current link	km	0.41	8.53	2.03	1.66
M: 22,159 (47.4%)	Number of signals on current link	int	0.00	8.00	1.23	2.00
L: 8,724 (18.6%)	Delay on previous link	min	-12.1	18.2	2.22	4.21
	Actual stop on current link	(1,0)	0.00	1.00	0.76	0.43
	Passenger load on current link	int	0.00	89.0	15.8	12.7
	Precipitation per half hour	mm	0.00	1.80	0.05	0.23
	Route type on current link#	[1-5]	na	na	na	na
	Route type on previous link#	[1-5]	na	na	na	na
Given state L	CI on previous link at current time^	%	9.67	87.10	56.7	20.0
	CI on current link at preceding time ^	%	6.09	107.5	66.1	18.7
Total: 22,138	RCI on current link^	%	12.0	88.5	66.6	19.2
H: 6,349 (28.7%)	Length of current link	km	0.38	8.53	1.98	1.70
M: 10,046 (45.4%)	Number of signals on current link	int	0.00	8.00	1.39	2.05
L: 5,743 (25.9%)	Delay on previous link	min	-8.27	24.9	2.21	4.41
	Actual stop on current link	(1,0)	0.00	1.00	0.74	0.44
	Passenger load on current link	int	0.00	90.0	14.8	12.1
	Precipitation per half hour	mm	0.00	1.80	0.05	0.25
	Route type on current link#	[1-5]	na	na	na	na
	Route type on previous link#	[1-5]	na	na	na	na

Notes: 'na' stands for not applicable.
* H, M, L stand for High, Medium and Low speed states, respectively. The marginal percentages are presented in parenthesis. The datasets are the observations on current link given previous link state (H , M, L).
^ CI is congestion index and RCI is recurrent congestion index.
# Route types 1-5 stand for Busway, Motorway, Arterial, Central business district (CBD) and others, respectively.

Note that the state transitions characterize how a vehicle experiences a series of link states along a trip rather than traffic propagations. The CI on current link at preceding time is used to capture the influence from both the upstream and downstream links at preceding time intervals. The CI on previous link at current time is used to reflect both the traffic and drivers' behavior impacts (e.g. schedule recovery, aggressiveness, etc.). In addition, the boundaries of states from the GMM clustering tend to be inconsistent for trips on different links over different time period of a day. For example, the boundary for a high speed state in the CBD area could be CI = 0.4 while on the Busway it could be CI = 0.7. The RCI is capable of capturing the recurrent traffic conditions for trips at different times and locations.

Table 7-2 summarizes the MNL models estimation results. Generally, the probabilities to be in high and medium speed states rather than a low speed state increase with the increase of the CI from the previous link and the CI on the current link from preceding time intervals.

Table 7-2: MNL model estimation coefficients and performance

State choice[1]	Variables	Given state H[2] B (β) ^{p}	Given state M[2] B (β) ^{p}	Given state L[2] B (β) ^{p}
H state[2]	CI on previous link at current time	0.034 (1.034)**	0.020 (1.020)**	0.017 (1.017)**
	CI on current link at preceding time	0.484 (1.622)**	0.462 (1.587)**	0.496 (1.642)**
	RCI on current link	-0.457 (0.633)**	-0.450 (0.638)**	-0.481 (0.618)**
	Length of current link	-0.141 (0.868)**	-0.321 (0.725)**	-0.397 (0.672)**
	Number of signals on current link	-0.06 (0.942)**	-0.053 (0.948)*	-0.177 (0.838)**
	Delay on previous link	0.032 (1.033)**	0.031 (1.031)**	0.048 (1.049)**
	Actual stop on current link	-0.068 (0.934)	-0.122 (0.885)**	-0.209 (0.811)**
	Passenger load on current link	-0.006 (0.994)**	0 (1)	0.001 (1.001)
	Precipitation at current time	0.076 (1.079)	0.084 (1.087)	0.064 (1.067)
	Route type on previous link[3]			
	Busway	-0.764 (0.466)**	-0.066 (0.936)	-0.204 (0.816)
	Motorway	-0.913 (0.401)**	0.706 (2.026)**	0.851 (2.342)**
	Arterial road	-0.102 (0.903)	0.04 (1.041)	0.334 (1.396)*
	Central business district	0.213 (1.238)*	1.159 (3.185)**	1.021 (2.775)**
	Route type on current link[3]			
	Busway	0.724 (2.063)**	0.144 (1.155)	0.877 (2.405)**
	Motorway	0.709 (2.033)**	1.008 (2.74)**	1.889 (6.61)**
	Arterial road	0.147 (1.159)	-0.849 (0.428)**	-0.176 (0.839)
	Central business district	0.631 (1.879)**	-0.704 (0.495)**	0.058 (1.06)
	Intercept	-1.79 (0)**	0.46 (0)*	0.518 (0)*
M state[2]	CI on previous link at current time	0.010 (1.010)**	0.012 (1.012)**	0.014 (1.014)**
	CI on current link at preceding time	0.197 (1.217)**	0.192 (1.211)**	0.201 (1.222)**
	RCI on current link	-0.172 (0.842)**	-0.193 (0.825)**	-0.202 (0.817)**
	Length of current link	-0.055 (0.946)*	-0.108 (0.898)**	-0.157 (0.855)**
	Number of signals on current link	-0.017 (0.983)	-0.013 (0.987)	-0.057 (0.945)**
	Delay on previous link	0.029 (1.029)**	0.024 (1.025)**	0.029 (1.029)**
	Actual stop on current link	-0.11 (0.896)*	-0.033 (0.968)	-0.076 (0.927)
	Passenger load on current link	-0.001 (0.999)	0.001 (1.001)	0.007 (1.007)**
	Precipitation at current time	-0.019 (0.981)	0.041 (1.042)	-0.168 (0.846)*
	Route type on current link[3]			
	Busway	-0.532 (0.587)**	0.027 (1.027)	-0.205 (0.814)
	Motorway	-0.863 (0.422)**	0.107 (1.113)	0.139 (1.149)
	Arterial road	-0.04 (0.961)	0.182 (1.199)*	0.423 (1.527)**
	Central business district	-0.177 (0.838)*	0.597 (1.817)**	0.357 (1.429)**
	Route type on previous link[3]			
	Busway	0.655 (1.925)**	0.019 (1.02)	0.628 (1.874)**
	Motorway	0.663 (1.94)**	0.511 (1.667)**	1.008 (2.74)**
	Arterial road	-0.106 (0.899)	-0.65 (0.522)**	0.071 (1.074)
	Central business district	0.581 (1.789)**	-0.384 (0.681)**	0.1 (1.105)
	Intercept	0.806 (0)**	1.599 (0)**	1.459 (0)**
L state[2]	Reference category (base outcome)			
L(0)		-44,862	-48,370	-23,615
L(B)		-33,515	-37,605	-17,840
rho-squared		0.253	0.223	0.245

Notes: The coefficients B (β) ^{p}: B = unstandardized coefficient, β = exponential coefficient ($\exp^B$) and *p* = significance level.
t statistics significance ** = $p < 0.01$ and * = $p < 0.05$.
1. The reference category is low speed state.
2. H, M, L = high, medium and low speed states, respectively.

The CI from a previous link has little effect as the models are already conditioned on the previous link state. The negative sign of RCI indicates that a vehicle operating on a link with a higher RCI (e.g. busway) would have a lower probability to be in a high speed state than on a link with a lower RCI (e.g. CBD) when all other factors are kept constant.

The probabilities to be in high and medium speed states compared to be in a low speed state decrease with the increase of link length (factors that create frictions, e.g. pedestrian crossing, traffic entering from side roads, etc.), number of signals (signalized intersection delay), and actual stop (stop delay). The delay on previous link has a positive effect due to the schedule recovery behaviour of drivers. If the vehicle is delayed on the previous link, it would speed up to catch up with timetables on current link. The passenger load has a negligible influence on transition probabilities. The precipitation has a significant and negative influence for the probability to be in a medium speed state compared to a low speed state conditional on a low speed state on the previous link. It indicates that the bad weather increases the probability for a vehicle to experience a low speed state on current link when the previous link is congested.

For prediction purpose, the MNL models in Table 7-2 were refined by taking into consideration of the relative importance of variables and the simplicity of data collection in practice. Table 7-3 summarizes the model specification and estimation results. CI on previous link at current time (*CI_PreL_CurT*), CI on current link at preceding time interval (*CI_CurL_PreT*), and RCI on current link (*RCI_CurL*) provide the best model specification.

Table 7-3: Specified MNL model coefficients and performance

Variables	Given state H #			Given state M #			Given state L #		
	H	M	L	H	M	L	H	M	L
CI_PreL_CurT	0.014*	0.006*		0.002*	0.001*		0.017*	0.007*	
CI_CurL_PreT	0.474**	0.189**		0.448**	0.181**		0.483**	0.193**	
RCI_CurL	-0.45**	-0.17**		-0.43**	-0.17**		-0.46**	-0.19**	
Intercept	-1.15*	0.734**		-0.27**	1.24**		-0.36**	1.38**	
L(0)		-44,810			-48,285			-23,605	
L(B)		-34,280			-38,773			-18,553	
rho-squared		0.235			0.197			0.214	

Notes: L state is the reference state in MNL models
Statistical t test significance ** $p < 0.001$ and * $p < 0.05$.
H, M, L represent high, medium and low speed states, respectively.

Figure 7-11illustrates the estimated state probabilities for CIs on the current link at the preceding time interval. Figure 7-11a and Figure 7-11b show the transition probabilities from the low speed state on the previous link to all states on current link having low (e.g. CBD area) and high (e.g. Busway) RCIs, respectively. Figure 7-11c and Figure 7-11d show the transition probabilities from a high speed state on the previous link to all states on current link having high and low RCIs, respectively. The probability to be in the high (low) speed state increases (decreases) monotonically with the increase of CIs at the preceding time interval. The probability to be in the medium speed state increases to a peak point and then decreases with the increase of CIs at the preceding time interval. The *x*-axis of the peak points are positively correlated with RCI values, which highlights the effectiveness of RCIs in differentiating the inconsistent boundaries of states on different types of roads.

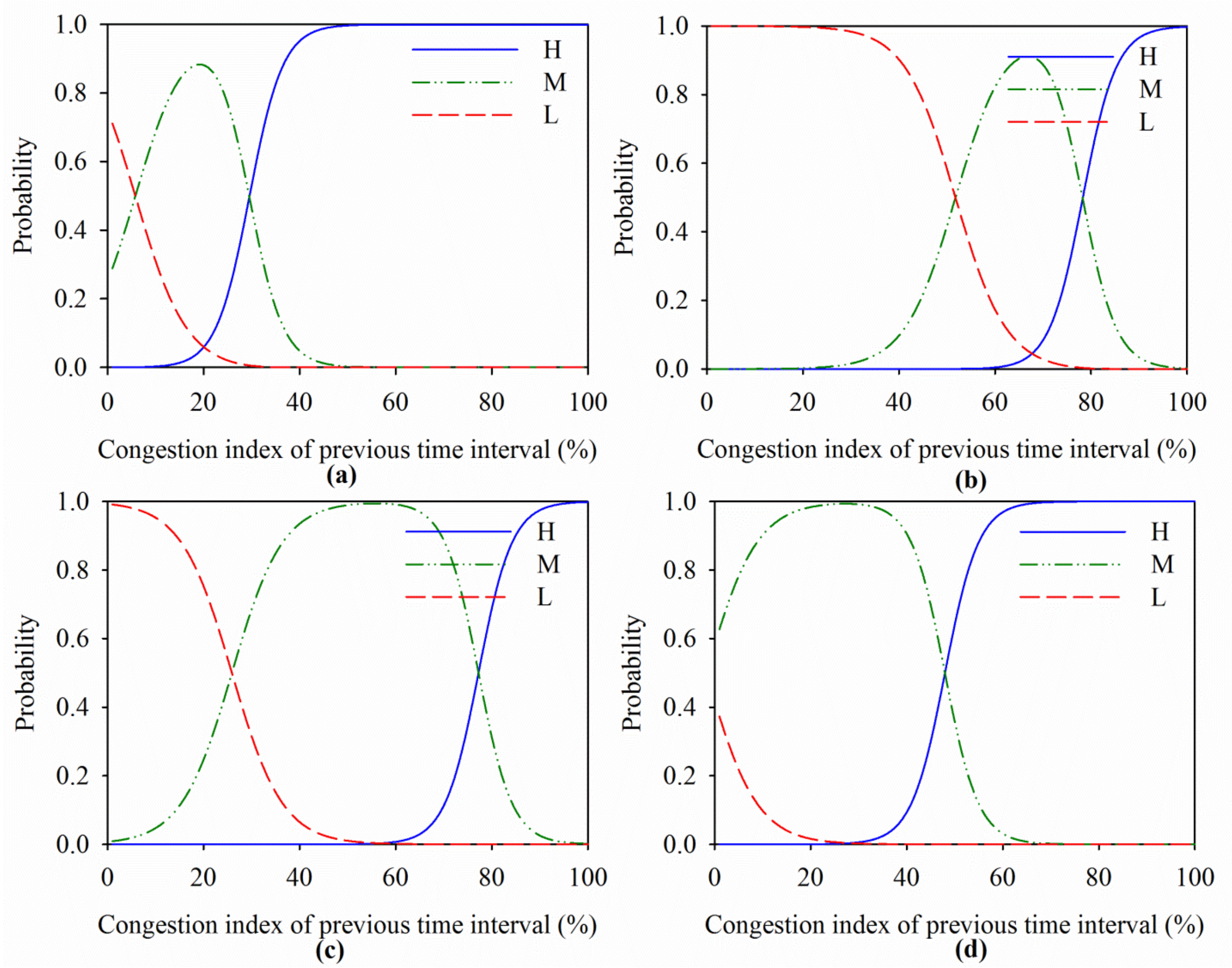

Figure 7-11: Estimated transition probabilities with different congestion indexs of preceding time interval, (a) CI_PreL_CurT = 25%, RCI_CurL = 25% ; (b) CI_PreL_CurT = 25%, RCI_CurL = 75% ; (c) CI_PreL_CurT = 75%, RCI_CurL = 75% ; (d) CI_PreL_CurT = 75%, RCI_CurL = 45% .

7.5.4 Probability distribution estimation and performance analysis

To assess the performance of the method, the Kullback-Leibler (KL) distance is calculated to compare the estimated distribution to the empirical (actual) one. For discrete distributions, the KL distance of the estimated distribution, TTD_{est} , from the empirical one, TTD_{emp} , is,

$$D_{\mathrm{KL}}\left(\mathrm{TTD}_{emp} \,||\, \mathrm{TTD}_{est}\right) = \sum_i p_{emp}\left(i\right)\log_2\left(p_{emp}\left(i\right)/p_{est}\left(i\right)\right) \tag{7.20}$$

where $p_{emp}\left(i\right)$ and $p_{est}\left(i\right)$ are the observed and estimated probabilities for an observation i.

The KL distance is a measure of the information lost when the estimated distribution is used to approximate the actual one. If the estimated distribution is equal to the actual one, the KL distance is 0. Statistical measures are also calculated to assess the accuracy of the estimation results, including mean, variance, and percentiles. The Kolmogorov-Smirnov (KS) test was used to test the null hypothesis that the estimated distribution equals the empirical one.

7.5.4.1 Performance comparison

The performance of the proposed GMC distribution estimation method is assessed by comparing it with the performance of alternative methods, including convolution, MC, and MGF models.

The ***convolution*** method assumes independence of link travel times.

Two MC models are implemented based on the route TTD estimation approach proposed by Ramezani and Geroliminis, (2012). One model is based on grid clustering (MC_Grid), and the second on GMMS clustering (MC_GMMS) for state definition. The transition probabilities for the MC models are estimated using the count-based method. The probability distributions of link running times and stop dwell times were estimated from empirical data using a nonparametric kernel distribution model. The MC_GMMS version is closer to the Ramezani and Geroliminis, (2012) method.

Two MGF models are implemented based on the path TTD estimation approach used by Srinivasan et al. (2014) and the MGF method described in Section 7.4.3. These two models are associated with different assumptions about the link TTDs, normal (MGF_CN), and lognormal (MGF_CLN). The covariance matrix $\boldsymbol{\Sigma}$ was estimated from historical observations. The nonlinear equation (25) is solved numerically using the standard trust-region-reflective algorithm in Matlab.

Three GMC models are implemented based on the proposed approach in Section 7.4.3. The states are clustered using the GMMS clustering algorithm in Section 4.1 and the transition probabilities are estimated from the logit models in Section 7.5.3. These three models are associated with different settings of link running time correlations along the Markov paths: no correlation (GMC_GMMS_NC), correlated normal distributions (GMC_GMMS_CN), and correlated lognormal distributions (GMC_GMMS_CLN). The covariance matrix $\boldsymbol{\Sigma}$ along each Markov path was estimated from historical observations conditional on link states between any link pairs at different time periods. The GMC_GMMS_NC is closer to Hofleitner,Herring,Abbeel, et al. (2012) method.

Table 7-4 presents the results for route 60 eastbound trips between stops 7 and 10 from 7:00-8:00 am on weekdays. Generally, the GMC and MC models provide more accurate estimations than the convolution and MGF models. The convolution model performs the worst since it fails to consider dependence between link travel times. Though the MGF models are capable of capturing spatial correlations, they fail the KS test due to their unimodal assumptions on link and trip TTDs. The MC_Grid model fails the KS test since the arbitrary state boundaries could not always reassure the independence as GMMS does. The MC_GMMS model provides the most accurate estimation even when conditional dependencies exist. The distributions conditional on states of upstream, current and downstream links can possibly minimize the impact of dependencies. The GMC_GMMS_CLN model performs well but relatively a bit worse than the MC_GMMS model. The major reason is the errors from the estimated transition probabilities. The GMC_GMMS_NC model performs worse than the GMC_GMMS_CN and GMC_GMMS_CLN models, which affirms the effectiveness of the MGF algorithm in capturing the correlations between travel times conditional on link states.

Table 7-4: Performance comparison [R60, eastbound, 7:00-8:00, between stops 7 and 10]

Models	KL dist.	KS test*	Mean	SD	Prc95	Prc75	Prc50	Prc25
Empirical	0.000	1	596	174	814	715	636	540
Convolution	0.444	0	597	136	798	689	612	515
MC_Grid	0.192	0	603	150	805	706	635	531
MC_GMMS	**0.036**	1	595	161	794	707	637	531
GMC_GMMS_NC	0.080	1	593	157	780	703	636	531
GMC_GMMS_CN	0.069	1	593	165	791	702	639	531
GMC_GMMS_CLN	**0.052**	1	594	162	797	702	637	531
MGF_CN	0.418	0	596	168	874	709	595	481
MGF_CLN	0.670	0	604	242	1060	729	561	432

Note: * The value 1 indicates that the model passes the KS test.

Figure 7-12 compares the estimated distributions with the empirical one. The GMC and MC models approximate the two peaks well, but the convolution and the MGF models fail to capture these. While the proposed GMC model provides a comparable performance with the MC_GMMS model, the proposed approach is generalizable (distributions conditional on states and transition probabilities can be modelled as functions of explanatory covariates) and computationally more efficient (distributions conditional only on current link state). It also requires less data for the calibration of the transition probabilities (compared to methods that estimate them based on link specific fractions from historical data. The MC_GMMS approach can only run with inputs of link travel time observations due to the introduction of intermediate stages in Markov paths (Ramezani and Geroliminis, 2012).

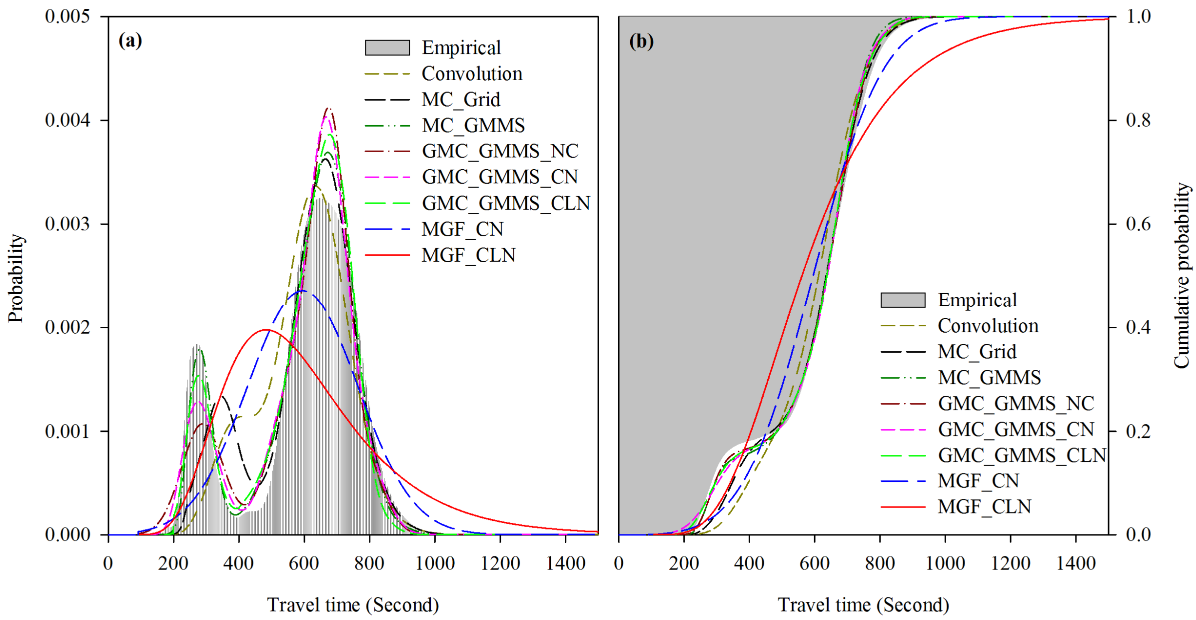

Figure 7-12: Probability density function and cumulative density function of the estimated distributions

7.5.4.2 Sensitivity analysis

The results in the previous section are from a path (trip) with three links. The sensitivity of the results to trip distances and road types is also of interest. Statistical tests, unimodality and spatiotemporal correlation, were conducted to examine the underlying characteristics of link TTDs. The unimodality of a distribution was tested using the Hartigan dip test (Hartigan and Hartigan, 1985). A high dip significance value provides support that the distribution is unimodal.

Trip distance

Figure 7-13 shows the KL performance contour maps of distribution estimation as a function of distance from the first stop (stop order) in different time 30 minutes time intervals (e.g. 7:00 = 7:00-7:30 am). Four models are presented: convolution, GMC_GMMS_NC (Hofleitner,Herring and Bayen, 2012), GMC_GMMS_CLN and MGF_CLN (Srinivasan et al., 2014). The MC_GMMS model is not shown as it was not feasible to estimate the distribution for trips with more than 9 links due to computational reasons which underlies one of the major limitations.

In general, the GMC models provide more accurate estimations than the convolution and MGF models for trips in peak hours, especially for trips between stops 1 to 3 on Motorway links. The reason for the lower performance of the convolution and MGF models are mainly due to the highly significant link correlations and multimodal distributions in peak hours on motorways. However, the MGF model performs well and relatively better than the GMC models for trips in off-peak hours (9:00-14:00), when the link and trip running times have unimodal distributions. The complementary performance between the GMC and MGF models points to the potential of developing a hybrid approach that can make full use of their advantages by balancing accuracy and computation burden. The GMC_GMMS_CLN model is more accurate and robust than the GMC_GMMS_NC model by further taking into consideration the spatial correlations along a Markov path. The superior performance of the former occurs for trips from stop 1 to stops 11 and 12 in peak hours. Stop 11 is a major spot near the CBD area where bus bunching happens frequently in peak hours. The correlation between its adjacent links was found to be significant. This could be explained by the fact that a bus having to stop would spend more time on both the adjacent upstream and downstream links than a bus driving through.

Table 7-5 summarizes the performance of the various models. The tested cases include trips on routes 555 and 60 with different distances from 7:00 to18:00 in 30 minutes intervals on weekdays. The results verify that the proposed GMC model provides more accurate and robust TTD estimation than the alternatives and it can fit the empirical distributions accurately for 88% of the tested cases.

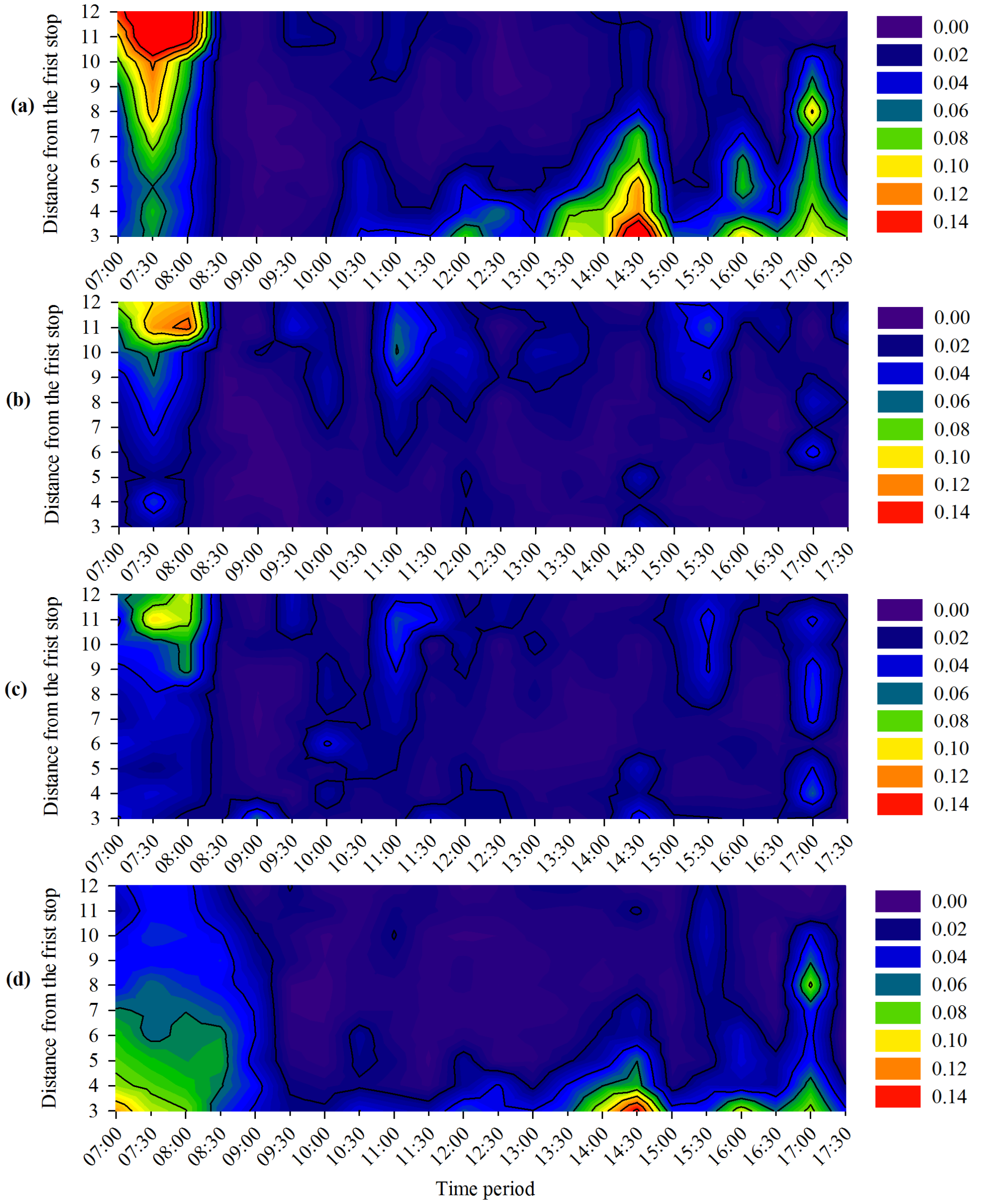

Figure 7-13: KL performance metric as a function of trip distance and time of day (R555, inbound, weekdays). (a) Convolution, (b) GMC_GMMS_NC, (c) GMC_GMMS_CLN, and (d) MGF_CLN

Table 7-5: Performance summary

Method	Average KL	Max KL	Percentage cases passed KS test
Convolution	0.083	0.190	61%
GMC_GMMS_NC	0.027	0.126	82%
GMC_GMMS_CLN	**0.014**	**0.089**	**88%**
MGF_CLN	0.044	0.156	69%

Road types

Table 7-6 compares the distribution estimations for trips on routes 555 and 60, on different types of roads on weekdays between 8:00-8:30 am. The facilities include Motorway, Busway, CBD, Arterial, and Residential roads. For Motorway, CBD, and Arterial trips, the link travel times have multimodal distributions (unimodal sig < 0.05) and are highly correlated (ST-ACF).

The convolution and MGF models fail to approximate the ground-truth empirical distributions. The proposed GMC model M3 incorporating link correlations along a Markov path performs relatively better than M2 assuming no correlation. For Busway trips, the travel times have an insignificant correlation and a unimodal distribution. In this case, the four models all approximate well the empirical distribution, with the MGF model M4 performing the best. For residential trips, the travel times have a significant correlation and a unimodal distribution. All models passed the KS test except the convolution model. If balancing simplicity and accuracy important, it is reasonable to use the convolution method for Busway trips. The MGF model can also be applied when the trip travel times have a unimodal distribution.

Table 7-6: Comparison of estimation performance for trips on different types of roads

Road types	Links*	Distance (km)	Unimodal sig[1]	ST-ACF (conf.)[2]	KL_M1 (pass)[3]	KL_M2 (pass)[3]	KL_M3 (pass)[3]	KL_M4 (pass)[3]
Motorway	1_3	14.4	0.00	0.25(.08)	0.140(0)	0.016(1)	0.011(1)	0.081(0)
Busway	3_10	14.4	0.82	0.03(.04)	0.018(1)	0.021(1)	0.014(1)	0.003(1)
CBD	6_9	2.57	0.01	0.21(.04)	0.273(0)	0.033(1)	0.023(1)	0.238(0)
Arterial	9_11	1.25	0.00	0.11(.04)	0.192(0)	0.041(1)	0.032(1)	0.139(0)
Residential	1_5	2.54	0.38	0.13(.06)	0.096(0)	0.014(1)	0.005(1)	0.011(1)

Note: * 1_3 indicates trips from stop 1 to stop 3.
1. sig = significance value. 2. conf. = 95% confidence boundary; ST-ACF = spatiotemporal autocorrelation function.
3. M1-M4 stands for Convolution, GMC_GMMS_NC, GMC_GMMS_CLN and MGF_CLN methods.

7.6 Discussions and applications

Figure 7-14 shows the implementation of the proposed GMC model for transit trip TTDs estimation. The data input is solely from the AVL system. For a trip between OD pair (i, j) at 7:00-7:30, the source data is partitioned according to trip attributes and then cleaned to exclude abnormal observations. The GMMS algorithm clusters the link running times and outputs the distributions conditional on current link states. The MNL TPM model estimates the transition probabilities with inputs of RCI on the current link, CI on the previous link at current time, and CI on the current link at the preceding time interval.

The Markov chain process constructs the Markov paths and calculates the probability of each path using the estimated transition probabilities. The Markov path distribution is approximated using the MGF algorithm with consideration of link correlations along the Markov paths. The covariance matrix conditional on the states between any two links can be derived from empirical data or estimated from a model. The trip running time distributions are estimated as the sum of Markov path distributions weighted by Markov path probabilities. The trip dwell time distribution is estimated as the convolution of stop dwell time distributions fitted using a non-parametric kernel model. Finally, the trip TTD is derived as the convolution of the trip running time and trip dwell time distributions.

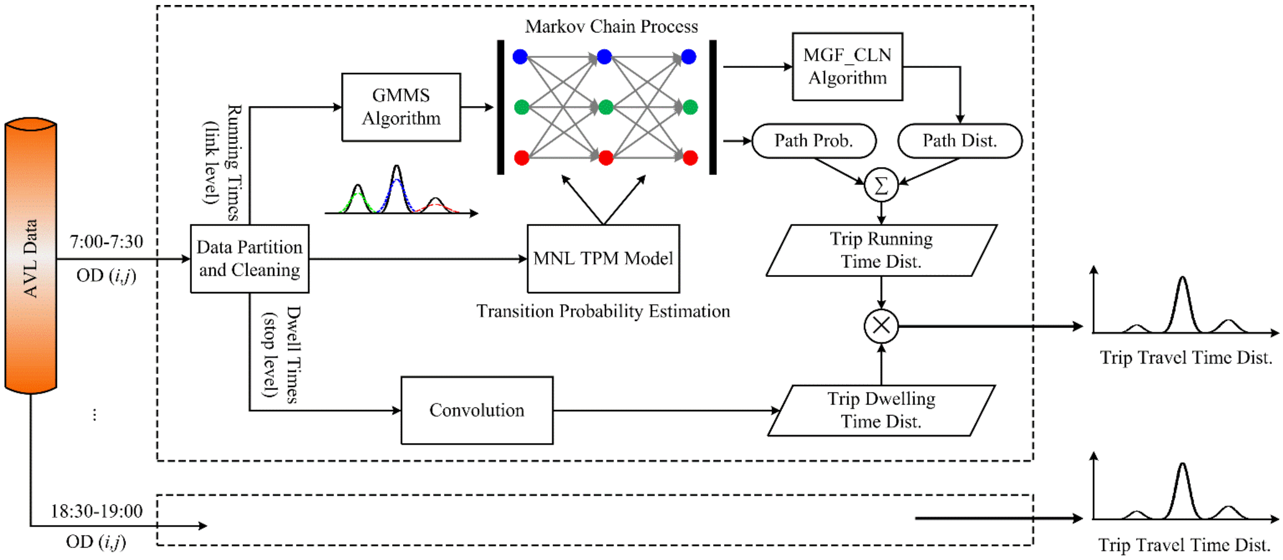

Figure 7-14: The implementation of the proposed GMC structure for transit application

As the proposed GMC structure is modular, it can be easily adapted for different applications. The upper part of the implementation structure can be readily used to estimate TTDs for car trips as well since all the inputs required are link based. The GMC structure can also be extended to provide real-time predictions of the bus arrival time distribution at downstream stops using a similar idea as in Noroozi and Hellinga (2014). For example, the state probabilities on the next link can be estimated using the MNL model given the current link state. The corresponding TTDs conditional on link states can be derived from historical data. Then the TTD on the next link can be estimated as the sum of these distributions weighted by the estimated probabilities.

To examine the effectiveness of the above mentioned approach in predicting the downstream link TTD, an alternative model with the fixed transition probabilities and a naïve historical data based model are developed for comparison. The fixed transition probabilities are calculated using the count-based approach for two successive links in 30 minutes under different cases (e.g. a case is weekday inbound from 7:00-7:30 am). The naïve model predicts the link TTD as the empirical distributions in 30 minutes under different cases. The lower and upper bounds of an interval prediction with theoretical coverage $100(1-\alpha)$ where $\alpha \in \left(0,1\right)$ are calculated as the $100(\alpha/2)$ and $100(1-\alpha/2)$ percentiles of the distribution, respectively (Woodard et al., 2015).

To measure the accuracy of the deterministic (mean) predictions, the mean absolute error (MAE) and the mean absolute percentage error (MAPE) are used. To measure the accuracy of interval predictions, the empirical coverage percentage and the average width of the interval are reported. The empirical coverage measures the percentage of test trips for which the observed travel time is inside the predicted interval with a specific theoretical coverage. If two methods can both predict the variability correctly, the average interval width is used to indicate which one is better (Woodard et al., 2015).

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|T_i^{act} - T_i^{pred}\right|$$
$$\text{MAPE}=\frac{1}{n}\sum_{i=1}^{n}\left|\frac{T_i^{act} - T_i^{pred}}{T_i^{act}}\right|\times 100\% \tag{7.21}$$

where,

T_i^{act} = actual travel time for observation i;

T_i^{pred} = predicted travel time for observation i;

n = number of observations.

Table 7-7 provides the summary of deterministic and interval predictions. The tested data are link running times for weekday inbound trips from 7:00-18:00 over a 6 months period. Generally, the models with the predicted and fixed TPMs perform better than the naïve historical data based model since they can both reflect the influence from the upstream link. The model with the predicted TPMs provides a relatively more accurate prediction performance than the one with the fixed TPMs in terms of both deterministic and interval predictions, since it further incorporates the real-time information from the preceding time intervals than can better adjust the predicted state probabilities on the successive link.

Table 7-7: Summary of deterministic and interval predictions performance

Accuracy measures	Model with predicted TPMs	Model with fixed TPMs	Naïve historical model
Mean absolute error*	**4.9**	5.4	8.5
Mean percentage error	**6.9%**	7.8%	13.2%
Empirical coverage#	92.9%	93.2%	**95.0%**
Average interval width#	**25.5**	28.7	40.9

Note: The bold values indicate the best prediction performance;
* the unit is seconds per kilometres;
# the theoretical coverage is 95% (the interval is between 2.5 percentile and 97.5 percentile of the distribution).

Figure 7-15 shows the mean and 95% confidence interval prediction results from the models with the predicted and fixed TPMs. The model with fixed TPMs tends to give relatively more constant mean predictions for all trips in the same time of a day which is counterintuitive in reality. The model with the predicted TPMs can reflect the various trip running times by taking advantage of the real-time information from the preceding time intervals, but still rather limited. The model with the predicted TPMs can provide relatively narrower interval predictions than the one with the fixed TPMs, especially in terms of the lower confidence boundary.

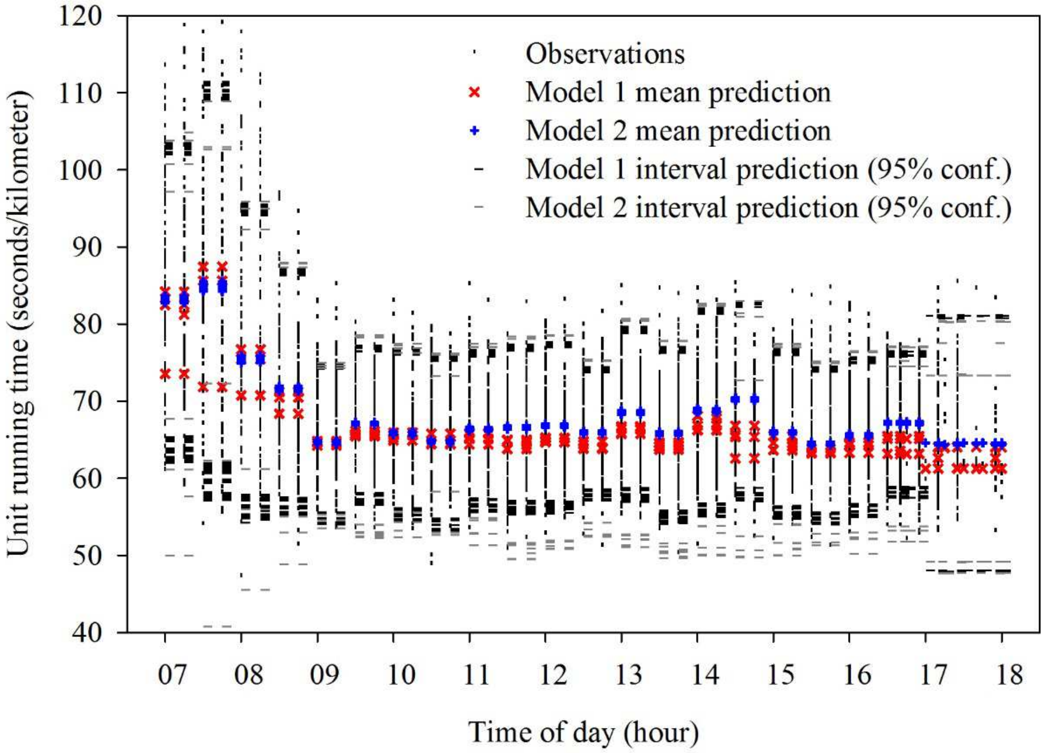

Figure 7-15: Predictions of means and intervals on a sample Motorway link for weekdays inbound trips over different times of day across six months period. [Model 1 uses the predicted probability; Model 2 uses the fixed probability; 95% conf. indicates 95% confidence intervals]

7.7 Summary

Travel time distribution along paths in transportation networks can provide comprehensive information for planning, operations monitoring and control, as well as travel planning. Previous studies focus on directly fitting the link or route travel times. However, the applicability of these methods may be limited due to the small number of direct observations at the Origin-Destination level. Furthermore, many of the methods assume link independence or correlated unimodal link distributions. The research proposes a GMC approach to estimate trip TTDs by aggregating link TTDs that takes into account the spatial-temporal correlations among link travel times. The proposed GMC approach captures the correlations among link travel times conditional on the underlying traffic states. The method is applicable under general conditions, as the link distributions are derived conditional on the states of the current link and the transition probabilities are estimated as a function of explanatory covariates using logit models.

The proposed approach has been demonstrated and validated in a transit case-study using AVL data. The results confirm that the GMC approach provides an effective and efficient way to estimate TTDs compared to other methods. The performance of the GMC method is promising specially when link correlations conditional on states and multimodal distributions exist. The method is also computationally more efficient than other methods proposed in the literature.

The use of transition probabilities as a function of explanatory variables makes the model general (and requires less data for calibration purposes than other methods). However, there is a small cost in estimation accuracy, as the estimated transition probabilities and the fitting of parametric link distribution models (required by the MGF algorithm), inevitably introduce errors that influence its estimation performance. The sensitivity analysis on trip distances and road types highlight the complementary performance of the GMC and MGF methods. Since dwell times are difficult to model at stop-level, being mainly determined by demand, the research has undertaken preliminary analysis on demand modelling and proposed an interactive multiple models (IMM) based pattern hybrid approach to predict short-term passenger demand at a route-level (Ma, Xing, et al., 2014).

Chapter 8 Conclusions and Future Research

Transit TTR performance influences service attractiveness, operating costs and system efficiency. Archived AVL-AFC data provides the potential for improving transit management and performance at all levels (planning and policy, operations, control). The research has reviewed the state-of-art in TTR and TTD analysis. Most studies estimate TTR for impact assessment of strategic and operational measures and methods for prediction of TTR are limited. TTD provides the comprehensive information for TTR analysis. While link-level TTD can be derived or inferred from dedicated or mobile sensors, methods for estimation of trip TTD between an origination and a destination pair are still evolving.

The thesis is divided into two main parts corresponding to TTR and TTD analyses. Leveraging on the AVL-AFC and the associated database of contributory factors, the first part quantified and modelled TTR in the Australian context. To develop the generic approach for analyzing and predicting TTR, the second part modelled TTDs at link and trip levels. The results can be used to derive TTR information that can be used to fulfill different transit stakeholders' requirements (impact assessment for operators and trip planning for passengers).

Following the analysis of existing studies, four main research gaps were identified, namely:

1. Current measures may provide inconsistent assessment with lack of distribution information and are insufficient to reflect passengers' perceptions under different occasions;
2. Analysis of unreliability causes is largely constrained by data availability and estimation approach. The traditional ordinary least square regression model can provide inefficient estimation of parameters due to cross-equation correlations caused by omitted covariates and unobserved heterogeneity;
3. Although TTD fitting is the prerequisite for TTR analysis, inconsistent results are reported to date. In assessing the performance of different distribution models, in addition to accuracy, the flexibility and explanatory power need to be considered;
4. For many applications, trip travel time information (e.g. trip planning) is important, but receives little attention in estimation of trip TTD from link TTDs with consideration of dynamic and heterogeneous spatiotemporal correlations.

The major contribution of this thesis is the proposed trip TTD estimation model using a generic Markov chain approach (Chapter 7). The proposed approach captures the correlations among link travel times conditional on the underlying traffic states. The method is applicable under general conditions as the link distributions are derived conditional on the states of the current link and the transition probabilities are estimated as a function of explanatory covariates using logit models.

8.1 Summary of the thesis

The empirical studies are commonly constrained by data availability. To establish a travel time related data warehouse for travel time reliability and travel time distribution studies, this thesis has developed a unique integrated dataset using different sources across six months of year, including AVL, Go card, GTFS, BSTM, BOM and STREAMS (Chapter 3). The integrated database provides detailed information on supply and demand, as well as the associated environmental information.

In the first part of this thesis (Chapter 4 and Chapter 5), TTR measures and models were developed to examine the impacts of unreliability factors, leading to the following contributions.

Chapter 4 proposes a set of TTR measures from the perspective of passengers using the operational AVL data considering different perceptions of TTR under different traffic states. Two issues with regard to buffer time estimation were discussed, namely, performance disaggregation and capturing passengers' perspectives on reliability. The main results are summarized as follows:

- The case studies verified the existence of mixture states during a given time period and the GMM model provides better fitting performance than single mode distributions.
- The proposed reliability measures provide consistent assessment with a high-level detail, while the conventional measures may give inconsistent assessment results.

Chapter 5 puts forward three general TTR related models with respect to main concerns by travellers and planners, namely, average travel time, buffer time and coefficient of variation of travel time. In addition, five groups of alternative models to account for variations caused by different road types, including arterial road, motorway, busway, CBD and others, were developed. Seemingly Unrelated Regression Equations (SURE) estimation is applied to account for the cross-equation correlations across regression models caused by unobserved heterogeneity. Three main categories of unreliability contributory factors have been identified and tested in this study, namely: planning, operational and environmental. The main results are summarized as follows:

- The defined recurrent congestion index captures different characteristics of road links and within-day variation of traffic conditions.
- Cross-equation correlations were found to exist between reliability models and the SURE provides more efficient estimation than the OLS model.
- The most important factors were found to be recurrent congestion index, traffic signals and passenger demand at stops.

In the second part of the thesis (Chapter 6 and Chapter 7), trip TTD estimation model from link TTDs was developed in order to predict TTR between arbitrary OD pairs at arbitrary times, leading to the following detailed contributions. The empirical findings (e.g. mixture distribution) of link and route TTDs affirm the importance of incorporating distribution information for TTR analysis in the first part of the thesis.

Chapter 6 specifies the most appropriate distribution model for the day-to-day travel time variability by using a novel evaluation approach and set of performance measures. We investigated the spatial and temporal aggregation influence on TTV. A novel evaluation approach and set of measures are developed to facilitate comprehensive comparison of alternative distribution models. The main results are summarized as follows:

- The decrease of temporal aggregation level results in a less asymmetric and flat distribution, and an increase of the normality of the distribution.
- The spatial aggregation of link travel times breaks up the multimodality distribution for the busway service while it is not applicable for the non-busway service.
- The GMM model is evaluated as superior to its alternatives under different cases in terms of fitting accuracy, robustness and explanatory power.
- Mathematically, GMM is flexible enough to model different types of TTDs, including symmetric, asymmetric and multimodal distributions

Chapter 7 proposes a generalized Markov chain (GMC) approach for estimation of trip TTDs between arbitrary OD pairs at arbitrary times from link or segment TTDs. The proposed approach consists of three major components, namely state definition, transition probabilities estimation and probability distribution estimation. A heuristic clustering method, based on Gaussian mixture models, has been developed to cluster link observations with regard to their homogeneity and underlying traffic conditions. A transition probability estimation model is developed as a function of link characteristics and trip conditions using a logit model. By applying a Markov chain procedure, the probability distribution of trip travel times is estimated as the combination of Markov path travel time distributions weighted by their corresponding occurrence probabilities. The link travel time distribution is conditioned on the traffic state of the current link that can be estimated from observations. A moment generating function based algorithm is used to approximate the Markov path travel time distribution as the sum of correlated link travel time distributions. The proposed approach is applied in a transit case study using automatic vehicle location data. The main results are summarized as follows:

- The proposed trip TTD estimation method is effective and efficient, especially when correlations and multimodal distributions exist and it is computationally more efficient than other methods proposed in the literature.
- The important factors for traffic state transitions are traffic condition from previous link, traffic condition on current link at preceding time interval and recurrent congestion index on current link (link characteristics).
- The proposed link TTD prediction method provides better deterministic and interval predictions of travel time than its alternatives.

8.2 Future research

This thesis has provided a detailed methodology for modelling TTR and TTD, as well as providing a deep insight into contributory factors. The main areas for future research are as follows:

1. In evaluating TTR from a passenger perspective, other important attributes should be incorporated, including waiting time, transfer time, budgeted waiting time, budgeted transfer time and schedule inconvenience. The calculations of these components under mixture mode distribution conditions need to be investigated.
2. The findings from TTR modelling are valid within the range of the used data and should be used with caution beyond this range. More bus routes with different operating characteristics can further complement the current findings. In addition, a refined reliability related dependent variables (Chapter 4) could potentially improve modelling performance and provide more insights of reliability contributors' impacts.
3. The modelling of TTDs with inputs of explanatory variables (link characteristics and trip conditions) needs to be investigated, especially for the prediction of mixture distributions (the states' occurrence probabilities and the corresponding distribution parameters).
4. The use of transition probabilities as a function of explanatory variables makes the model quite general. However, this comes at a small cost in estimation accuracy. More factors need to be examined and the performance will be evaluated for predictions of trip travel time distributions (a set of links) as well as link travel time distributions.
5. Developing a hybrid scheme based on different trip TTD estimation methods (corresponding to characteristics of the underlying traffic states) has the potential to improve estimation accuracy and decrease computation burden. This is a promising future research direction, specially for real-time analysis.
6. Since stop dwell times are difficult to model at stop-level, being mainly determined by demand, the research has proposed an interactive multiple models (IMM) based pattern hybrid approach to predict short-term passenger demand at a route-level (Ma, Xing, et al., 2014). Future work could involve the estimation of dwell time distributions given a predicted level of demand.
7. Intersection delay is an important part of travel time variability. To model intersections explicitly and regard them as separate type of trip TTD estimation entity from the links would be potential useful.

References

Abkowitz, M., Slavin, H., Waksman, R., Englisher, L., & Wilson, N. H. M. (1978). Transit service reliability *TSC Urban and Regional Research Series*. U.S. Department of Transportation (DOT): Cambridge.

Abkowitz, M. D., & Engelstein, I. (1983). Factors affecting running time on transit routes. *Transportation Research Part A-Policy and Practice, 17*(2), 107-113.

Abramowitz, M., & Stegun, I. A. (1964). *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*: Courier Corporation.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716-723.

Al-Deek, H., & Emam, E. B. (2006). New methodology for estimating reliability in transportation networks with degraded link capacities. *Journal of Intelligent Transportation Systems, 10*(3), 117-129.

Anastasopoulos, P. C., Mannering, F. L., & Haddock, J. E. (2012). Random Parameters Seemingly Unrelated Equations Approach to the Postrehabilitation Performance of Pavements. *Journal of Infrastructure Systems, 18*(3), 176-182.

Anderson, T. W., & Darling, D. A. (1954). A test of goodness of fit. *Journal of the American Statistical Association, 49*(268), 765-769.

Arthur, D., & Vassilvitskii, S. (2007). *k-means++: the advantages of careful seeding*. Paper presented at the Proceedings of the 18th annual ACM-SIAM symposium on Discrete algorithms, New Orleans, Louisiana.

Balcombe, R., Mackett, R., Paulley, N., Preston, J., Shires, J., Titheridge, H., Wardman, M., & White, P. (2004). The demand for public transport: a practical guide. *Crowthorne, UK*. TRL report, TRL593.

Barkley, T., Hranac, R., & Petty, K. (2012). Relating travel time reliability and nonrecurrent congestion with multistate models. *Transportation Research Record: Journal of the Transportation Research Board, 2278*, 13-20.

Bates, J., Polak, J., Jones, P., & Cook, A. (2001). The valuation of reliability for personal travel. *Transportation Research Part E: Logistics and Transportation Review, 37*(2–3), 191-229.

Bell, M. G., & Iida, Y. (1997). *Transportation network analysis*. Chichester, UK: John Wiley & Sons.

Bertini, R., & El-Geneidy, A. (2004). Modeling Transit Trip Time Using Archived Bus Dispatch System Data. *Journal of transportation engineering, 130*(1), 56-67.

Bhat, C. R., & Sardesai, R. (2006). The impact of stop-making and travel time reliability on commute mode choice. *Transportation Research Part B: Methodological, 40*(9), 709-730.

Brakewood, C., Macfarlane, G. S., & Watkins, K. (2015). The impact of real-time information on bus ridership in New York City. *Transportation Research Part C: Emerging Technologies, 53*, 59-75.

Camus, R., Longo, G., & Macorini, C. (2005). Estimation of Transit Reliability Level-of-Service Based on Automatic Vehicle Location Data. *Transportation Research Record: Journal of the Transportation Research Board, 1927*, 277-286.

Cathey, F. W., & Dailey, D. J. (2003). A prescription for transit arrival/departure prediction using automatic vehicle location data. *Transportation Research Part C: Emerging Technologies, 11*(3–4), 241-264.

Ceder, A. A. (2007). *Public Transit Planning and Operation:Theory, modelling and practice*. Oxford, UK: Elsevier, Butterworth-Heinemann.

Chang, H., Park, D., Lee, S., Lee, H., & Baek, S. (2010). Dynamic multi-interval bus travel time prediction using bus transit data. *Transportmetrica, 6*(1), 19-38.

Chen, A., Yang, H., Lo, H. K., & Tang, W. H. (2002). Capacity reliability of a road network: an assessment methodology and numerical results. *Transportation Research Part B: Methodological, 36*(3), 225-252.

Chen, G. T., Jing;Zhang, Dong;Yang, Xiaoguang;Yang, Xi-yu;. (2012). *Dwell Time Estimation with Consideration of Bus Bunching*. Paper presented at the Transportation Research Board 91st Annual Meeting, Washington D.C., United States.

Cheng, T., Wang, J., Haworth, J., Heydecker, B., & Chow, A. (2014). A Dynamic Spatial Weight Matrix and Localized Space–Time Autoregressive Integrated Moving Average for Network Modeling. *Geographical Analysis, 46*(1), 75-97.

Cheng, Y.-H., & Tsai, Y.-C. (2014). Train delay and perceived-wait time: passengers' perspective. *Transport Reviews, 34*(6), 710-729.

Chu, H. C. (2010). *Estimating travel time reliability on freight corridors.* Paper presented at the Transportation Research Board 89th Annual Meeting, Washington D.C., United States.

Chun-Hsin, W., Jan-Ming, H., & Lee, D. T. (2004). Travel-time prediction with support vector regression. *IEEE Transactions on Intelligent Transportation Systems, 5*(4), 276-281.

Clark, S., & Watling, D. (2005). Modelling network travel time reliability under stochastic demand. *Transportation Research Part B: Methodological, 39*(2), 119-140.

Coifman, B., & Kim, S. (2009). Speed estimation and length based vehicle classification from freeway single-loop detectors. *Transportation Research Part C: Emerging Technologies, 17*(4), 349-364.

Coifman, B., & Krishnamurthy, S. (2007). Vehicle reidentification and travel time measurement across freeway junctions using the existing detector infrastructure. *Transportation Research Part C: Emerging Technologies, 15*(3), 135-153.

Currie, G., Douglas, N. J., & Kearns, I. (2012). *An Assessment of Alternative Bus Reliability Indicators* Paper presented at the Australasian Transport Research Forum 2012 Proceedings, Perth, Australia.

Diab, E., & El-Geneidy, A. (2013). Variation in bus transit service: understanding the impacts of various improvement strategies on transit service reliability. *Public Transport, 4*(3), 209-231.

Diab, E. I., & El-Geneidy, A. M. (2012). Understanding the impacts of a combination of service improvement strategies on bus running time and passenger’s perception. *Transportation Research Part A: Policy and Practice, 46*(3), 614-625.

Du, L., Peeta, S., & Kim, Y. H. (2012). An adaptive information fusion model to predict the short-term link travel time distribution in dynamic traffic networks. *Transportation Research Part B: Methodological, 46*(1), 235-252.

Dueker, K. J. (2004). Determinants of bus dwell time. *Journal of public transportation, 7*(1), 21-40.

El-Geneidy, A., & Vijayakumar, N. (2011). The effects of articulated buses on dwell and running times. *Journal of public transportation, 14*(3), 63-86.

El-Geneidy, A. M., Horning, J., & Krizek, K. J. (2011). Analyzing transit service reliability using detailed data from automatic vehicular locator systems. *Journal of Advanced Transportation, 45*(1), 66-79.

El-Geneidy, A. M., Strathman, J. G., Kimpel, T. J., & Crout, D. T. (2006). Effects of bus stop consolidation on passenger activity and transit operations. *Transportation Research Record: Journal of the Transportation Research Board, 1971*, 32-41.

Emam, E. B., & Ai-Deek, H. (2006). Using real-life dual-loop detector data to develop new methodology for estimating freeway travel time reliability. *Transportation Research Record: Journal of the Transportation Research Board, 1959*, 140-150.

Fan, W. D., & Machemehl, R. B. (2009). Do transit users just wait for buses or wait with strategies. *Transportation Research Record: Journal of the Transportation Research Board, 2111*, 169-176.

Faouzi, N.-E. E., & Maurin, M. (2007). *Reliability of travel time under lognormal distribution.* Paper presented at the Transport Research Board 86th Annual Meeting, Washington D.C., United States.

Fei, X., Lu, C.-C., & Liu, K. (2011). A bayesian dynamic linear model approach for real-time short-term freeway travel time prediction. *Transportation Research Part C: Emerging Technologies, 19*(6), 1306-1318.

Fosgerau, M., & Engelson, L. (2011). The value of travel time variance. *Transportation Research Part B: Methodological, 45*(1), 1-8.

Fosgerau, M., & Fukuda, D. (2012). Valuing travel time variability: Characteristics of the travel time distribution on an urban road. *Transportation Research Part C: Emerging Technologies, 24*, 83-101.

Furth, P., & Muller, T. (2007). Service Reliability and Optimal Running Time Schedules. *Transportation Research Record: Journal of the Transportation Research Board, 2034*, 55-61.

Furth, P. G., & Muller, T. H. J. (2006). Service Reliability and Hidden Waiting Time: Insights from Automatic Vehicle Location Data. *Transportation Research Record: Journal of the Transportation Research Board, 1955*, 79-87.

Geroliminis, N., & Skabardonis, A. (2011). Identification and Analysis of Queue Spillovers in City Street Networks. *IEEE Transactions on Intelligent Transportation Systems, 12*(4), 1107-1115.

Gilliam, C., Chin, T., Black, I., & Fearon, J. (2008). *Forecasting and appraising travel time variability in urban areas.* Paper presented at the European Transport Conference, Noordwijkerhout, Netherlands.

Golshani, F. (1983). System Regularity and Overtaking Rules in Bus Services. *Journal of the Operational Research Society, 34*(7), 591-597.

Goverde, R. M. P. (1999). *Improving punctuality and transfer reliability by railway timetable optimization.* (Doctor of Philosopy), Delft University of Technology, Delft, Netherlands.

Guo, F., Li, Q., & Rakha, H. (2012). Multistate travel time reliability models with skewed component distributions. *Transportation Research Record: Journal of the Transportation Research Board, 2315*, 47-53.

Guo, F., Rakha, H., & Park, S. (2010). A multi-state travel time reliability model. *Transportation Research Record: Journal of the Transportation Research Board, 2188*, 46-54.

Hartigan, J. A., & Hartigan, P. (1985). The dip test of unimodality. *The Annals of Statistics, 13*(1), 70-84.

Hellinga, B., Izadpanah, P., Takada, H., & Fu, L. (2008). Decomposing travel times measured by probe-based traffic monitoring systems to individual road segments. *Transportation Research Part C: Emerging Technologies, 16*(6), 768-782.

Hofleitner, A., Herring, R., Abbeel, P., & Bayen, A. (2012). Learning the dynamics of arterial traffic from probe data using a dynamic Bayesian network. *IEEE Transactions on Intelligent Transportation Systems, 13*(4), 1679-1693.

Hofleitner, A., Herring, R., & Bayen, A. (2012). Arterial travel time forecast with streaming data: A hybrid approach of flow modeling and machine learning. *Transportation Research Part B: Methodological, 46*(9), 1097-1122.

Hofmann, M., & O'Mahony, M. (2005). *The impact of adverse weather conditions on urban bus performance measures*. Paper presented at the IEEE Conference on Intelligent Transportation Systems, Vienna, Austria.

Hollander, Y. (2006). *The Cost of Bus Travel Time Variability.* (Doctor of philosophy), The University of Leeds, Leeds, West Yorkshire, England.

Hollander, Y., & Liu, R. (2008). Estimation of the distribution of travel times by repeated simulation. *Transportation Research Part C: Emerging Technologies, 16*(2), 212-231.

Hu, K.-C., & Jen, W. (2006). Passengers' perceived service quality of city buses in Taipei: Scale development and measurement. *Transport Reviews, 26*(5), 645-662.

Hunter, T., Das, T., Zaharia, M., Abbeel, P., & Bayen, A. M. (2013). Large-Scale Estimation in Cyberphysical Systems Using Streaming Data: A Case Study With Arterial Traffic Estimation. *IEEE Transactions on Automation Science and Engineering, 10*(4), 884-898.

Hunter, T., Herring, R., Abbeel, P., & Bayen, A. (2009). *Path and travel time inference from GPS probe vehicle data*. Paper presented at the NIPS Analyzing Networks and Learning with Graphs, Whistler, Canada.

Hunter, T., Hofleitner, A., Reilly, J., Krichene, W., Thai, J., Kouvelas, A., Abbeel, P., & Bayen, A. (2013). Arriving on time: estimating travel time distributions on large-scale road networks. *arXiv preprint arXiv:1302.6617*.

Jang, W. (2010). Travel Time and Transfer Analysis Using Transit Smart Card Data. *Transportation Research Record: Journal of the Transportation Research Board, 2144*, 142-149.

Janos, M., & Furth, P. (2002). Bus Priority with Highly Interruptible Traffic Signal Control: Simulation of San Juan's Avenida Ponce de Leon. *Transportation Research Record: Journal of the Transportation Research Board, 1811*, 157-165.

Jenelius, E. (2012). The value of travel time variability with trip chains, flexible scheduling and correlated travel times. *Transportation Research Part B: Methodological, 46*(6), 762-780.

Jenelius, E., & Koutsopoulos, H. N. (2013). Travel time estimation for urban road networks using low frequency probe vehicle data. *Transportation Research Part B: Methodological, 53*(0), 64-81.

Jenelius, E., & Koutsopoulos, H. N. (2015). Probe vehicle data sampled by time or space: Consistent travel time allocation and estimation. *Transportation Research Part B: Methodological, 71*, 120-137.

Jordan, W. C., & Turnquist, M. A. (1979). Zone scheduling of bus routes to improve service reliability. *Transportation Science, 13*(3), 242-268.

Kaas, A., & Jacobsen, E. (2008). *Analysing the Metro Cityring in Copenhagen*. Paper presented at the 11th International Conference on Computer System Design and Operation in the Railway and Other Transit Systems, Toledo, Spain.

Kaparias, I., Bell, M. G. H., & Belzner, H. (2008). A new measure of travel time reliability for in-vehicle navigation systems. *Journal of Intelligent Transportation Systems, 12*(4), 202-211.

Kazagli, E., & Koutsopoulos, H. (2013). Estimation of Arterial Travel Time from Automatic Number Plate Recognition Data. *Transportation Research Record: Journal of the Transportation Research Board, 2391*, 22-31.

Khosravi, A., Mazloumi, E., Nahavandi, S., Creighton, D., & Van Lint, J. W. C. (2011). A genetic algorithm-based method for improving quality of travel time prediction intervals. *Transportation Research Part C: Emerging Technologies, 19*(6), 1364-1376.

Khosravi, A., Mazloumi, E., Nahavandi, S., Creighton, D., & Van Lint, J. W. C. (2011). Prediction Intervals to Account for Uncertainties in Travel Time Prediction. *IEEE Transactions on Intelligent Transportation Systems, 12*(2), 537-547.

Kieu, L.-M., Bhaskar, A., & Chung, E. (2014). *Establishing definitions and modeling public transport travel time variability*. Paper presented at the Transportation Research Board 93th Annual Meeting Washington D.C. United States.

Kimpel, T. J. (2001). *Time Point-Level Analysis of Transit Service Reliability and Passenger Demand*. (Doctor of Philosophy), Portland State University, Portland, United States.

Kimpel, T. J., Strathman, J., Bertini, R. L., & Callas, S. (2005). Analysis of transit signal priority using archived TriMet bus dispatch system data. *Transportation Research Record: Journal of the Transportation Research Board, 1925*, 156-166.

Kittelson & Assoc, I., Parsons Brinckerhoff, I., KFH Group, I., Institute, T. A. M. T., & Arup. (2003). Transit Capacity and Quality of Service Manual. Third Edition. Transit Cooperative Highway Research Program (TCRP) Report 165. Transportation Research Board, Washington D.C., United States.

Kittelson & Associates, Urbitran, LKC Consulting Services, MORPACE International, Technology, Q. U. o., & Nakanishi, Y. (2003). A Guidebook for Developing a Transit Performance-Measurement System. Transportation Research Board, Washington, D.C., United States.

Kuhn, B., Higgins, L., Nelson, A., Finley, M., Ullman, G., Chrysler, S., Wunderlich, K., Shah, V., & Dudek, C. (2013). Effectiveness of Different Approaches to Disseminating Traveler Information on Travel Time Reliability. Transportation Research Board, Washington D.C., United States.

Lam, T. C., & Small, K. A. (2001). The value of time and reliability: measurement from a value pricing experiment. *Transportation Research Part E: Logistics and Transportation Review, 37*(2–3), 231-251.

Li, R., & Rose, G. (2011). Incorporating uncertainty into short-term travel time predictions. *Transportation Research Part C: Emerging Technologies, 19*(6), 1006-1018.

Li, R., Rose, G., & Sarvi, M. (2006). Using automatic vehicle identification data to gain insight into travel time variability and its causes. *Transportation Research Record: Journal of the Transportation Research Board, 1945*, 24-32.

Lin, J., & Ruan, M. (2009). Probability-based bus headway regularity measure. *Intelligent Transport Systems, 3*(4), 400-408.

Lin, T., & Wilson, N. H. M. (1992). Dwell time relationships for light rail systems. *Transportation Research Record: Journal of the Transportation Research Board, 1361*.

Liu, R., & Sinha, S. (2007). *Modelling Urban Bus Service and Passenger Reliability*. Paper presented at the The 3rd International Symposium on Transportation Network Reliability, The Hague, Netherlands.

Lomax, T., Schrank, D., Turner, S., & Margiotta, R. (2003). Selecting travel reliability measures. Texas Transportation Institute and Cambridge Systematics Inc.

Ma, Z., Ferreira, L., & Mesbah, M. (2014). Measuring Service Reliability Using Automatic Vehicle Location Data. *Mathematical Problems in Engineering, 2014,* 12.

Ma, Z., Ferreira, L., Mesbah, M., & Hojati, A. T. (2015). Modelling Bus Travel Time Reliability Using Supply and Demand Data from Automatic Vehicle Location and Smart Card Systems. *Transportation Research Record: Journal of the Transportation Research Board (In press).*

Ma, Z., Ferreira, L., Mesbah, M., & Zhu, S. (2015). Modelling distributions of travel time variability for bus operations. *Journal of Advanced Transportation (In press).*

Ma, Z., Koutsopoulos, H. N., Ferreira, L., & Mesbah, M. (2015). Trip travel time distribution estimation: A generalized Markov chain approach. *Transportation Research Part B: Methodological (Under review).*

Ma, Z., Xing, J., Mesbah, M., & Ferreira, L. (2014). Predicting short-term bus passenger demand using a pattern hybrid approach. *Transportation Research Part C: Emerging Technologies, 39*, 148-163.

Madanat, S., Mishalani, R., & Ibrahim, W. H. W. (1995). Estimation of Infrastructure Transition Probabilities from Condition Rating Data. *Journal of Infrastructure Systems, 1*(2), 120-125.

Mannering, F. (2007). *Effects of Interstate Speed Limits on Driving Speeds: Some New Evidence*. Paper presented at the Transportation Research Board 86th Annual Meeting, Washington, D.C., United States.

Martchouk, M., Mannering, F., & Bullock, D. (2010). Analysis of freeway travel time variability using Bluetooth detection. *Journal of transportation engineering, 137*(10), 697-704.

May, A., Bonsall, P., & Marler, N. (1989). Travel time variability of a group of car commuters in north London: Institute of Transport Studies, University of Leeds, Leeds, UK.

Mazloumi, E., Currie, G., & Rose, G. (2008). *Causes of travel time unreliability–a Melbourne case study*. Paper presented at the 31st Australasian Transport Research Forum, Gold Coast, Australia.

Mazloumi, E., Currie, G., & Rose, G. (2010). Using GPS data to gain insight into public transport travel time variability. *Journal of Transportation Engineering, 136*(7), 623-631.

Mehta, N. B., Jingxian, W., Molisch, A. F., & Jin, Z. (2007). Approximating a Sum of Random Variables with a Lognormal. *IEEE Transactions on Wireless Communications, 6*(7), 2690-2699.

Mesbah, M., Currie, G., Lennon, C., & Northcott, T. (2012). Spatial and temporal visualization of transit operations performance data at a network level. *Journal of Transport Geography, 25*, 15-26.

Meyer, M. D. (2002). Measuring system performance: Key to establishing operations as a core agency mission. *Transportation Research Record: Journal of the Transportation Research Board, 1817*, 155-162.

Milkovits, M. (2008). Modeling the Factors Affecting Bus Stop Dwell Time: Use of Automatic Passenger Counting, Automatic Fare Counting, and Automatic Vehicle Location Data. *Transportation Research Record: Journal of the Transportation Research Board, 2072*, 125-130.

Miller, L., Mannering, F., & Abraham, D. M. (2009). Effectiveness of speed control measures on nighttime construction and maintenance projects. *Journal of Construction Engineering and Management, 135*(7), 614-619.

Nakanishi, Y. J. (1997). Bus Performance Indicators: On-Time Performance and Service Regularity. *Transportation Research Record: Journal of the Transportation Research Board, 1571*, 1-13.

Nassir, N., Hickman, M., & Ma, Z.-L. (2015). Activity detection and transfer identification for public transit fare card data. *Transportation, 42*(4), 683-705.

Ng, M., & Waller, S. T. (2010). A computationally efficient methodology to characterize travel time reliability using the fast Fourier transform. *Transportation Research Part B: Methodological, 44*(10), 1202-1219.

Noland, R. B., & Polak, J. W. (2002). Travel time variability: A review of theoretical and empirical issues. *Transport Reviews, 22*(1), 39-54.

Noroozi, R., & Hellinga, B. (2014). Real-Time Prediction of Near-Future Traffic States on Freeways Using a Markov Model. *Transportation Research Record: Journal of the Transportation Research Board, 2421*, 115-124.

Osuna, E. E., & Newell, G. F. (1972). Control Strategies for an Idealized Public Transportation System. *Transportation Science, 6*(1), 52.

Park, B.-J., Zhang, Y., & Lord, D. (2010). Bayesian mixture modeling approach to account for heterogeneity in speed data. *Transportation Research Part B: Methodological, 44*(5), 662-673.

Park, S., Rakha, H., & Guo, F. (2010). *Multistate travel time reliability model: Model calibration issues.* Paper presented at the Transportation Research Board 89th Annual Meeting, Washington, D.C., United States.

Pattanamekar, P., Park, D., Rilett, L. R., Lee, J., & Lee, C. (2003). Dynamic and stochastic shortest path in transportation networks with two components of travel time uncertainty. *Transportation Research Part C: Emerging Technologies, 11*(5), 331-354.

Pearson, R. K. (2002). Outliers in process modeling and identification. *IEEE Transactions on Control Systems Technology, 10*(1), 55-63.

Peer, S., Koopmans, C. C., & Verhoef, E. T. (2012). Prediction of travel time variability for cost-benefit analysis. *Transportation Research Part A: Policy and Practice, 46*(1), 79-90.

Peng, Z. R., Lynde, E., Chen, W. Y., & Kong, C. (2009). *Understanding Transit Service Gaps*. Paper presented at the Transportation Research Board 89th Annual Meeting, Washington, D.C., United States.

Pfeifer, P. E., & Deutrch, S. J. (1980). A Three-Stage Iterative Procedure for Space-Time Modeling Phillip. *Technometrics, 22*(1), 35-47.

Polus, A. (1979). A study of travel time and reliability on arterial routes. *Transportation, 8*(2), 141-151.

Pronello, C., & Camusso, C. (2012). A Review of Transport Noise Indicators. *Transport Reviews, 32*(5), 599-628.

Pu, W. (2011). Analytic Relationships Between Travel Time Reliability Measures. *Transportation Research Record: Journal of the Transportation Research Board, 2254*, 122-130.

Rahmani, M., Jenelius, E., & Koutsopoulos, H. N. (2015). Non-parametric estimation of route travel time distributions from low-frequency floating car data. *Transportation Research Part C: Emerging Technologies, (In press).*

Rahmani, M., & Koutsopoulos, H. N. (2013). Path inference from sparse floating car data for urban networks. *Transportation Research Part C-Emerging Technologies, 30*, 41-54.

Ramezani, M., & Geroliminis, N. (2012). On the estimation of arterial route travel time distribution with Markov chains. *Transportation Research Part B: Methodological, 46*(10), 1576-1590.

Richardson, A., & Taylor, M. (1978). Travel time variability on commuter journeys. *High Speed Ground Transportation Journal, 12*(1), 77-99.

Sangjun, P., Rakha, H., & Feng, G. (2011, 5-7 Oct. 2011). *Multi-state travel time reliability model: Impact of incidents on travel time reliability.* Paper presented at the 14th International IEEE Conference on Intelligent Transportation Systems, Washington, D.C., United States.

Srinivasan, K. K., Prakash, A. A., & Seshadri, R. (2014). Finding most reliable paths on networks with correlated and shifted log–normal travel times. *Transportation Research Part B: Methodological, 66*, 110-128.

Strathman, J. G., Dueker, K. J., Kimpel, T., Gerhart, R., Turner, K., Taylor, P., Callas, S., Griffin, D., & Hopper, J. (1999). Automated bus dispatching, operations control, and service reliability: Baseline analysis. *Transportation Research Record: Journal of the Transportation Research Board, 1666*, 28-36.

Strathman, J. G., Kimpel, T. J., Dueker, K. J., Gerhart, R. L., & Callas, S. (2002). Evaluation of transit operations: data applications of Tri-Met's automated Bus Dispatching System. *Transportation, 29*(3), 321-345.

Sumalee, A., Pan, T., Zhong, R., Uno, N., & Indra-Payoong, N. (2013). Dynamic stochastic journey time estimation and reliability analysis using stochastic cell transmission model: Algorithm and case studies. *Transportation Research Part C: Emerging Technologies, 35*, 263-285.

Sun, L., Tirachini, A., Axhausen, K. W., Erath, A., & Lee, D.-H. (2014). Models of bus boarding and alighting dynamics. *Transportation Research Part A: Policy and Practice, 69*, 447-460.

Surprenant-Legault, J., & El-Geneidy, A. M. (2011). Introduction of reserved bus lane: Impact on bus running time and on-time performance. *Transportation Research Record: Journal of the Transportation Research Board, 2218*, 10-18.

Susilawati, S., Taylor, M. A., & Somenahalli, S. V. (2013). Distributions of travel time variability on urban roads. *Journal of Advanced Transportation, 47*(8), 720-736.

Tahmasseby, S. (2009). *Reliability in Urban Public Transport Network Assessment and Design.* (Doctor of Philosophy), Delft University of Technology, Delft, Netherlands.

Taylor, M. (1982). Travel time variability-the case of two public modes. *Transportation Science, 16*(4), 507-521.

Taylor, M. A. P., & Susilawati. (2012). Modelling travel time reliability with the Burr distribution. *Procedia - Social and Behavioral Sciences, 54*, 75-83.

Tétreault, P. R., & El-Geneidy, A. M. (2010). Estimating bus run times for new limited-stop service using archived AVL and APC data. *Transportation Research Part A: Policy and Practice, 44*(6), 390-402.

Tirachini, A. (2011). Bus dwell time: the effect of different fare collection systems, bus floor level and age of passengers. *Transportmetrica A: Transport Science, 9*(1), 28-49.

Tirachini, A. (2013). Estimation of travel time and the benefits of upgrading the fare payment technology in urban bus services. *Transportation Research Part C: Emerging Technologies, 30*(0), 239-256.

Trompet, M., Liu, X., & Graham, D. (2011). Development of Key Performance Indicator to Compare Regularity of Service Between Urban Bus Operators. *Transportation Research Record: Journal of the Transportation Research Board, 2216*, 33-41.

Tu, H., van Lint, H. W., & van Zuylen, H. J. (2007a). *Impact of Adverse Weather on Travel Time Variability of Freeway Corridors*. Paper presented at the Transportation Research Board 86th Annual Meeting, Washington, D.C., United States.

Tu, H., van Lint, J. W., & van Zuylen, H. J. (2007b). Impact of traffic flow on travel time variability of freeway corridors. *Transportation Research Record: Journal of the Transportation Research Board, 1993*, 59-66.

Turnquist, M. A., & Bowman, L. A. (1980). The effects of network structure on reliability of transit service. *Transportation Research Part B: Methodological, 14*(1–2), 79-86.

Uniman, D., Attanucci, J., Mishalani, R., & Wilson, N. (2010). Service Reliability Measurement Using Automated Fare Card Data. *Transportation Research Record: Journal of the Transportation Research Board, 2143*, 92-99.

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

van Hinsbergen, C., Van Lint, J., & Van Zuylen, H. (2009). Bayesian committee of neural networks to predict travel times with confidence intervals. *Transportation Research Part C: Emerging Technologies, 17*(5), 498-509.

Van Lint, J. (2008). Online learning solutions for freeway travel time prediction. *IEEE Transactions on Intelligent Transportation Systems, 9*(1), 38-47.

Van Lint, J., Hoogendoorn, S., & van Zuylen, H. J. (2005). Accurate freeway travel time prediction with state-space neural networks under missing data. *Transportation Research Part C: Emerging Technologies, 13*(5), 347-369.

van Lint, J., & van Zuylen, H. (2005). Monitoring and predicting freeway travel time reliability: Using width and skew of day-to-day travel time distribution. *Transportation Research Record, 1917,* 54-62.

van Lint, J. W. C., van Zuylen, H. J., & Tu, H. (2008). Travel time unreliability on freeways: Why measures based on variance tell only half the story. *Transportation Research Part A: Policy and Practice, 42*(1), 258-277.

van Oort, N. (2011). *Service Reliability and Urban Public Transport Design.* (Doctor of Philosophy), Delft University of Technology, Delft, Netherlands

van Oort, N., & van Nes, R. (2004). *Service Regularity Analysis for Urban Transit Network Design*. Paper presented at the Transportation Research Board 83th Annual Meeting Washington, D.C., United States.

van Oort, N., & van Nes, R. (2010). Impact of Rail Terminal Design on Transit Service Reliability. *Transportation Research Record: Journal of the Transportation Research Board, 2146,* 109-118.

Vlahogianni, E. I., & Karlaftis, M. G. (2011). Temporal aggregation in traffic data: implications for statistical characteristics and model choice. *Transportation Letters, 3*(1), 37-49.

Wakabayashi, H., & Matsumoto, Y. (2012). Comparative study on travel time reliability indexes for highway users and operators. *Journal of Advanced Transportation, 46*(4), 318-339.

Washington, S. P., Karlaftis, M. G., & Mannering, F. L. (2011). *Statistical and econometric methods for transportation data analysis*. Boca Raton, FL.: Chapman and Hall/CRC.

Westgate, B. S., Woodard, D. B., Matteson, D. S., & Henderson, S. G. (2013). Travel time estimation for ambulances using Bayesian data augmentation. *The Annals of Applied Statistics, 7*(2), 1139-1161.

Woodard, D., Nogin, G., Koch, P., Racz, D., Goldszmidt, M., & Horvitz, E. (2015). *Predicting Travel Time Reliability using Mobile Phone GPS Data*. Retrieved from http://people.orie.cornell.edu/woodard/WoodNogiKoch15.pdf

Xue, Y., Jin, J., Lai, J., Ran, B., & Yang, D. (2011). *Empirical characteristics of transit travel time distribution for commuting routes.* Paper presented at the Transportation Research Board 90th Annual Meeting, Washington D.C., United States.

Yeon, J., Elefteriadou, L., & Lawphongpanich, S. (2008). Travel time estimation on a freeway using Discrete Time Markov Chains. *Transportation Research Part B: Methodological, 42*(4), 325-338.

Yetiskul, E., & Senbil, M. (2012). Public bus transit travel-time variability in Ankara *Transport Policy, 23,* 50-59.

Yildirimoglu, M., & Geroliminis, N. (2013). Experienced travel time prediction for congested freeways. *Transportation Research Part B: Methodological, 53*(0), 45-63.

Yu, B., Lam, W. H. K., & Tam, M. L. (2011). Bus arrival time prediction at bus stop with multiple routes. *Transportation Research Part C-Emerging Technologies, 19*(6), 1157-1170.

Yu, B., Yao, J., & Yang, Z. (2010). An improved headway-based holding strategy for bus transit. *Transportation Planning and Technology, 33*(3), 329-341.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association, 57*(298), 348-368.

Zheng, F., & Van Zuylen, H. (2010). Uncertainty and predictability of urban link travel time delay distribution-based analysis. *Transportation Research Record: Journal of the Transportation Research Board 2192*, 136-146.

Zheng, F., & Van Zuylen, H. (2013). Urban link travel time estimation based on sparse probe vehicle data. *Transportation Research Part C: Emerging Technologies, 31*, 145-157.