STUDIES IN COMPUTER SCIENCE NO. 41

MALMÖ UNIVERSITY

Shairyar Malik

Modelling and evaluating demand-responsive public transport

DISSERTATION IN FACULTY OF TECHNOLOGY AND SOCIETY

MODELLING AND EVALUATING
DEMAND-RESPONSIVE PUBLIC TRANSPORT

MODELLING AND EVALUATING DEMAND-RESPONSIVE PUBLIC TRANSPORT

Thesis for Licentiate Degree (Lic.)
By Shairyar Malik

The thesis will be defended in public in Room NI:A0311, Malmö University
Wednesday, November 04 2026 at 10.15

Principal Supervisor
Åse Jevinger
Malmö University
Department of Computer Science and Media Technology

Supervisor
Jan A. Persson
Malmö University
Department of Computer Science and Media Technology

Opponent
Miloš N. Mladenović
Aalto University
Department of Built Environment

Examinator
Bahtijar Vogel
Malmö University
Department of Computer Science and Media Technology

Malmö University, Faculty of Technology and Society
Licentiate Thesis 2026

© Shairyar Malik, 2026 [licensed under CC BY 4.0]

Paper 1 © The authors. Submitted for journal publication.
Paper 2 © 2026 The Authors. Published by IEEE (licensed under CC BY 4.0).
Paper 3 © 2026 The Authors. Published by IEEE (licensed under CC BY 4.0).

Cover design by Saleha & Tayyaba

DOI 10.24834/isbn.9789178778300
ISBN 978-91-7877-829-4 (print)
ISBN 978-91-7877-830-0 (pdf)

Printed in Sweden by Media-Tryck, Lund University,
Lund 2026

Media-Tryck is a Nordic Swan Ecolabel
certified provider of printed material.
Read more about our environmental
work at www.mediatryck.lu.se

MADE IN SWEDEN

SHAIRYAR MALIK

MODELLING AND EVALUATING DEMAND-RESPONSIVE PUBLIC TRANSPORT

Malmö University
Faculty of Technology and Society
Department of Computer Science and Media Technology

Studies in Computer Science

Faculty of Technology and Society
Malmö University

1. Jevinger, Åse. *Toward intelligent goods: characteristics, architectures and applications*, 2014, Doctoral dissertation.
2. Dahlskog, Steve. *Patterns and procedural content generation in digital games: automatic level generation for digital games using game design patterns*, 2016, Doctoral dissertation.
3. Fabijan, Aleksander. *Developing the right features: the role and impact of customer and product data in software product development*, 2016, Licentiate thesis.
4. Paraschakis, Dimitris. *Algorithmic and ethical aspects of recommender systems in e-commerce*, 2018, Licentiate thesis.
5. Hajinasab, Banafsheh. *A Dynamic Approach to Multi Agent Based Simulation in Urban Transportation Planning*, 2018, Doctoral dissertation.
6. Fabijan, Aleksander. *Data-Driven Software Development at Large Scale*, 2018, Doctoral dissertation.
7. Bugeja, Joseph. *Smart Connected Homes: Concepts, Risks, and Challenges*, 2018, Licentiate thesis.
8. Alkhabbas, Fahed. *Towards Emergent Configurations in the Internet of Things*, 2018, Licentiate thesis.
9. Paraschakis, Dimitris. *Sociotechnical Aspects of Automated Recommendations - Algorithms, Ethics, and Evaluation*, 2020, Doctoral dissertation.
10. Tegen, Agnes. *Approaches to Interactive Online Machine Learning*, 2020, Licentiate thesis.
11. Alvarez, Alberto. *Exploring the Dynamic Properties of Interaction in Mixed-Initiative Procedural Content Generation*, 2020, Licentiate thesis.
12. Alkhabbas, Fahed. *Realizing Emergent Configurations in the Internet of Things*, 2020, Doctoral dissertation.
13. Ashouri, Majid, *Towards Supporting IoT System Designers in Edge Computing Deployment Decisions*, 2021, Licentiate thesis.
14. Bugeja, Joseph, *On Privacy and Security in Smart Connected Homes*, 2021, Doctoral dissertation.
15. Azadvar, Ahmad, *Predictive Psychological Player Profiling*, 2021, Licentiate thesis.
16. Holmberg, Lars, *Human In Command Machine Learning*, 2021, Licentiate thesis.
17. John, Meenu Mary, *Design Methods and Processes for ML/DL models*, 2021, Licentiate thesis.
18. Tegen, Agnes, *Interactive Online Machine Learning*, 2022, Doctoral dissertation.
19. Ouhaichi, Hamza, *Towards designing a flexible multimodal learning analytics system*, 2022, Licentiate thesis.
20. Alvarez Uribe, Alberto Enrique, *Exploring Game Design through Human-AI Collaboration*, 2022, Doctoral dissertation.
21. Amouzad Mahdiraji, Saeid, *On the Use of Simulation and Optimization for the Analysis and Planning of Prehospital Stroke Care*, 2022, Licentiate thesis.
22. Holmberg, Lars, *Neural Networks in Context: Challenges and Opportunities*, 2023, Doctoral dissertation.
23. Pettersson, Mårten, *Getting Engaged in Cooperation: Design, Distance and Distributed Work*, 2023, Doctoral dissertation.
24. Engström, Jimmy, *Scaling Indoor Positioning - Methods for improving accuracy and privacy of indoor positioning*, 2023, Licentiate thesis.
25. Dytckov, Sergei, *Modelling and Simulating Demand-Responsive Transport*, 2023, Licentiate thesis.
26. Ouhaichi, Hamza, *A Framework for Designing and Analysing Multimodal Learning Analytics Systems*, 2024, Doctoral dissertation
27. Ymeri, Gent, *Machine Learning-Driven Analysis of Sensor Data for Objective Assessment of*

Parkinson´s Disease Motor Symptoms in Home Environments, 2024, Licentiate thesis.

28. Brötzner, Anna, *Visibility in Graphs and Polygons*, 2024, Licentiate thesis.
29. Bagheri, Sally, *The Role of Ethics in Smart Homes*, 2024, Licentiate thesis.
30. John, Meenu Mary, *Towards Continuous Development of MLOps Practices*, 2025, Doctoral dissertation
31. Belfrage, Michael, *Agent-based Social Simulation & Policy-Modelling: Facilitating Realistic and Credible Decision-making Support*, 2025, Licentiate thesis
32. Johansson, Emil, *Towards More Credible Agent-based Models for Policy Support*, 2025, Licentiate thesis
33. Abid, Muhammad Adil, *Artificial Intelligence for Enhanced Prehospital Stroke Care*, 2025, Licentiate thesis
34. Saleem, Hajira, *Towards improving localization for autonomous vehicles*, 2025, Licentiate thesis
35. Amouzad Mahdiraji, Saeid, *Optimization and simulation modeling for improved analysis and planning of prehospital stroke care*, 2025, Doctoral dissertation
36. Caramaschi, Sara, *Digital Health Technologies for Unsupervised Physical Activity Testing and Monitoring*, 2025, Licentiate thesis.
37. Jamali, Mahtab, *Context-aware learning for adaptive vision-based systems*, 2025, Licentiate thesis
38. Shokrollahi, Azad, *Occupancy information and people counting in smart buildings using noninvasive sensors and machine learning*, 2025, Licentiate thesis.
39. Soleimani, Ali, *Hybrid energy system optimization: towards intelligent and sustainable heating control*, 2025, Licentiate thesis
40. Ymeri, Gent, *Design, Development and Validation of Smartphone- and Wearable-Based Digital Biomarkers for Parkinson's Disease*, 2026, Doctoral thesis.
41. Malik, Shairyar, *Modelling and Evaluating Demand-Responsive Public Transport*, 2026, Licentiate thesis.

This publication is also electronically available at:
`mau.diva-portal.org`

To my parents, who made education our inheritance!

CONTENTS

Abstract 1

PREFACE 3
Author Contributions 4
Use of generative artificial intelligence 4

ACKNOWLEDGEMENTS 5

COMPREHENSIVE SUMMARY 7

1 Introduction 9
1.1 DRT in the Public Transport Context 9
1.2 Existing Research on DRT Design, Modelling, and Evaluation 10
1.3 Research Problem and Gaps 11
1.4 Thesis Organisation 12

2 Background and Analytical Concepts 13
2.1 DRT: Service Concept and Context 13
2.2 Representing Service Design and Passenger Demand 15
2.3 Performance Concepts for DRT Evaluation 16

3 Research Aim and Research Questions 17
3.1 Overall Aim and Main Research Question 17
3.2 Research Questions 17
3.3 Relationship Between the Research Questions and Papers 18

4 Research Design and Methods 19
4.1 Methodological Foundations and Research Design 19
4.2 Research Approaches in the Included Papers 20

5 Contribution 23
5.1 RQ1. Which service and operational design aspects are considered relevant in DRT modelling and evaluation, and how are design choices within these aspects represented and evaluated? 23
5.2 RQ2. How can selected patterns in empirical passenger demand be represented in synthetic demand-generation models, and how do these models affect simulated DRT performance? 25
5.3 RQ3. How can operational booking data and self-reported user data be used to analyse realised DRT performance, patterns of use, and the service's travel role? 28
5.4 Thesis-Level Synthesis and Contribution 30

6 Discussion 33
6.1 Representation, Demand, and Performance Evaluation 33
6.2 Implications for DRT Modelling and Evaluation 34
6.3 Limitations and Transferability 35

7 Conclusions and Future Research 37
7.1 Conclusions 37
7.2 Future Research 38

PAPERS 43

Abstract

Demand-responsive public transport (DRT) provides flexible, request-based mobility and can complement conventional public transport where demand is dispersed or difficult to serve efficiently with fixed routes and timetables. Its flexibility creates many possible service and operational configurations, making DRT design and evaluation challenging. Modelling and simulation can support such analysis, but their conclusions depend on how service characteristics and passenger demand are represented and on which performance measures are used. Strengthening DRT modelling and evaluation can therefore help researchers, modellers, and public transport authorities analyse service design choices and interpret DRT performance. This thesis aims to improve the methodological basis for modelling and evaluating selected aspects of DRT performance and use.

The thesis addresses this aim through three complementary studies. A systematic mapping study examines service and operational design aspects and performance metrics across 33 primary DRT modelling studies. A simulation study uses 518 days of empirical passenger-demand data to compare four synthetic demand-generation models with simulated empirical demand under common operating conditions. An observational empirical study analyses operational booking records and self-reported user data from a Swedish DRT pilot, Buss on Demand (BoD).

The mapping study shows that DRT modelling is organised around a strong operational core, particularly fleet management, route and operation, operational constraints, and scheduling and timing, while service-facing dimensions are less consistently represented. Performance evaluation is concentrated around passenger time and access, reliability and request fulfilment, cost and vehicle operation, and utilisation and pooling. The simulation study shows that synthetic demand models with similar aggregate demand can retain different temporal, spatial, within-day, and user-level structures, leading to different simulated passenger and vehicle outcomes. Among the models investigated, no single model is closest to the simulated empirical reference across all evaluated measures, indicating that model choice should reflect the intended evaluation purpose. In the observational study, trip distance showed the strongest association with realised pooling, while longer booking lead time was associated with more frequent pickup lateness; pooling was more frequent and pickup lateness less frequent later in the observed service period.

Users without a driving licence or household car used the service more intensively, while reported alternatives included private motorised transport, active modes, public or shared transport, and, for some respondents, not making the trip.

Taken together, the thesis shows that conclusions about DRT performance and use depend on how service design and passenger demand are represented, which outcomes are measured, and how the operational and user data are analysed. The thesis thereby provides methodological support for analysing DRT design and performance.

PREFACE

This thesis is based on the following three papers. These papers are referred to in the text by their Roman numerals I-III:

I. S. Malik, Å. Jevinger, and J. A. Persson, "Operational Design Aspects and Performance Metrics in Demand-Responsive Public Transport Modelling Studies: A Systematic Mapping Review," submitted for journal publication, 2026.

II. S. Malik, J. A. Persson, and Å. Jevinger, "A Modular Synthetic Demand Generation Framework for Demand-Responsive Public Transport," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 7, pp. 1520–1533, 2026. doi: 10.1109/OJITS.2026.3705486.

III. S. Malik, Å. Jevinger, and J. A. Persson, "Realised Pooling, Pickup Reliability, and Mobility Dependence in Rural Demand-Responsive Public Transport: Evidence From a Swedish DRT Pilot," *IEEE Access*, vol. 14, pp. 123678–123690, 2026. doi: 10.1109/ACCESS.2026.3722951.

Related publication not included in the thesis:

Å. Jevinger, J. A. Persson, and S. Malik, "Lessons from Modelling Demand Responsive Transport," submitted for journal publication, 2026.

Author Contributions

The research presented in this thesis was conducted in collaboration with the co-authors. For Papers I-III, Shairyar Malik was the first author and had primary responsibility for conducting the studies, developing and applying the methodological approaches, processing and analysing the data, interpreting the results, and preparing the manuscripts. He also contributed to empirical research conducted in collaboration with Västtrafik through the Policy Lab for Public Transport On-Demand project, including acquiring and curating the data and operational information required for the research.

Åse Jevinger and Jan A. Persson contributed to the conceptual and methodological development of the studies, interpretation of the results, supervision, and critical review and revision of the manuscripts.

The paper *Lessons from Modelling Demand Responsive Transport* was not included in the thesis because Shairyar Malik's contribution mainly involved gathering research papers and classifying them based on the framework developed in Paper I, building on his work presented there.

Use of generative artificial intelligence

Generative AI tools were used as writing support during the preparation of this thesis, restricted to language refinement and improving the clarity of the text. All substantive ideas, methodological decisions, analyses, interpretations, and conclusions are the work of the author and co-authors, as applicable.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisors, **Åse Jevinger** and **Jan A. Persson**, for their guidance, encouragement, constructive discussions, and continuous support throughout my doctoral journey to this point. I also thank my colleagues and collaborators, particularly **Camilla Ståhl** and **Frida Vernersdotter** at Västtrafik, and Policy Lab för kollektivtrafik on-demand at K2, for their valuable support.

My deepest gratitude goes to my family; their support made this journey possible long before I began my PhD.

I remember my late father with great respect. His own limited opportunities for formal education gave him a profound belief in its importance. Together with my mother, he made extraordinary sacrifices to ensure that all ten of their children could pursue higher education. In 2004, as one chapter of my life was beginning with engineering studies, another came to an end with the loss of my father. In many ways, I followed the path he had shown me; from modest beginnings and limited means, I kept moving forward through education. The milestones that followed were ones I wished I could have shared with him: graduating as an engineer, serving industry, completing my master's degree, moving to Sweden, and doing a PhD.

I am equally grateful to my mother, who stood beside him throughout and devoted herself to supporting our family. I am fortunate that she can see the continuation of what they started together. I also thank my brothers and sisters, who supported me after my father's death and helped me continue my education. Seeing one of them complete a PhD at KTH in Sweden also became an important source of inspiration for my own academic journey.

To my wife, I owe more than words can adequately express. Since our marriage in 2010, she has stood beside me along with my in-laws through financially demanding years, long periods of combining work with my master's studies, and, eventually, our move to Sweden. Her patience, sacrifices, encouragement, and care for our family have been essential throughout this journey.

Finally, I thank my children, **Saleha**, **Tayyaba**, and **Muhammad**. For many years, they have understood that *Baba is studying* and that work and studies sometimes continue into the evenings and follow me home. Their patience, love, and understanding have given me strength. Cheer up, bunnies! The game is on...

PART I

COMPREHENSIVE SUMMARY

1 Introduction

"An advanced city is not one where the poor own a car,
but one where the rich use public transport."
— Enrique Peñalosa

1.1 DRT in the Public Transport Context

Conventional public transport is most effective where travel demand is sufficiently concentrated, but maintaining frequent fixed-route services is harder when potential passengers and destinations are dispersed [1]. This challenge is especially relevant in rural, interurban, and other low-density settings, where demand may vary considerably across space and time [1]. Demand-responsive public transport (DRT) can address this challenge by adapting service provision to passenger requests [2].

DRT's role depends on how it is designed and positioned within the wider public transport system. A DRT service may provide local mobility, connect passengers to scheduled public transport, complement low-frequency fixed services, or, in some settings, substitute for them [3, 4]. European experience also shows substantial variation in how DRT services are organised and what purposes they serve [5]. DRT should therefore not be treated as one uniform service concept.

Flexibility alone does not ensure that a DRT service performs well or remains viable. Factors such as cost, patronage, operational complexity, service design, and local conditions can influence DRT performance and viability [6, 7]. This creates a practical and scientific challenge because DRT offers many possible service and operational choices. These choices affect both service operations and passenger-facing conditions, such as waiting time, travel time, and cost. Modelling and simulation are among the approaches that can support the analysis of such choices under specified assumptions [8]. Conclusions from such analyses can, however, depend on how service and passenger demand are represented, which outcomes are measured, and the context in which the results are produced.

1.2 Existing Research on DRT Design, Modelling, and Evaluation

DRT research has developed along several related but distinct strands. Service-oriented research classifies DRT according to characteristics such as flexibility, responsiveness, service pattern, and integration with conventional public transport, while adaptability studies examine the conditions under which different forms of DRT may be suitable [2, 3, 9]. A separate research tradition concerns the dial-a-ride problem (DARP), which studies routing, scheduling, request assignment, capacity, time windows, and ride-time constraints [10, 11]. Agent-based and other simulation approaches provide another way to examine DRT operations under specified assumptions [8]. Together, these strands provide substantial knowledge about DRT design and operation, but they address different analytical questions and use different service representations.

Simulation requires explicit representation of passenger demand. Demand may be derived from empirical observations or generated synthetically, and different approaches can preserve different temporal, spatial, within-day, and user-level structures [12, 13]. For example, researchers have used empirically derived activity patterns to construct structured transport demand, while others have developed synthetic commuting demand specifically from statistical data for rural DRT simulation [12, 13]. These examples show that passenger demand modelling is already an established research area. They also illustrate that different approaches to passenger demand modelling can produce different representations of demand for DRT simulation.

DRT performance is multidimensional, and conclusions depend partly on what is measured [7, 14]. Studies report a range of outcomes, including passenger-related measures, vehicle operation, cost, utilisation, and pooling, depending on the analysis purpose [7, 14]. Modelling studies also assess whether requests can feasibly be combined under conditions related to delay, detour, fare, and passenger acceptance [15]. Such results are meaningful within their modelling assumptions, but they do not automatically describe what is realised in an operating service.

Operating DRT services can provide empirical evidence on realised demand, service performance, patterns of use, and other aspects of service operation. Longitudinal booking records have been used to examine demand and realised pooling [16], while another study shows that interpreting service activity can depend on geographic and population context [7]. Operational records have also been combined with user surveys to examine public transport connections, private mobility resources, and reported travel alternatives [17]. These findings reflect live operation, but their interpretation remains tied to factors such as service design, observation period, available data, outcome definitions, and analytical unit.

Taken together, the literature is extensive, but the resulting findings are not directly interchangeable. Relevant differences include how modelling studies represent service and operational choices and passenger demand, which performance indicators

they use and how they define them, and how operating contexts and available data shape the interpretation of live-service results. Clarifying these differences matters scientifically when interpreting or comparing results across studies, and practically when modelling and operational analyses inform DRT service planning and evaluation. Understanding what a DRT result means therefore requires attention to what it represents, what it measures, and the context from which it is derived.

1.3 Research Problem and Gaps

Against this background, the thesis addresses a methodological problem in how DRT performance is modelled and evaluated. DRT performance can depend on a broad range of service, operational, and contextual characteristics, while conclusions also depend on how selected characteristics and passenger demand are represented, which outcomes are measured, and how realised operation and use are observed and analysed. The overall objective of the thesis is to improve the methodological basis for modelling and evaluating selected aspects of DRT performance and use. A clearer methodological basis can support researchers, modellers, and public transport authorities in interpreting DRT results and in analysing service and operational choices for planning and evaluation. The formal aim and research questions are presented in Chapter 3. This objective is motivated by three related gaps.

The first gap concerns the relation between DRT design and performance evaluation in modelling research. Existing reviews classify DRT systems and examine their adaptability, while the DARP literature provides detailed accounts of routing and scheduling problems [2, 9, 10, 11]. Less well structured is how service and operational design aspects are represented in primary DRT modelling studies, how individual design choices within these aspects are used, and which performance measures are directly linked to their evaluation. A clearer mapping of these relationships is needed to distinguish broad representation from direct evaluation and to identify which design-evaluation connections are well represented and which remain weakly covered.

The second gap concerns how simulation represents passenger demand. Existing studies demonstrate several ways to construct empirical and synthetic demand for DRT and related transport simulations [8, 12, 13]. Less well established is how selected characteristics of empirical passenger demand can be represented in alternative synthetic demand-generation models and how the resulting demand representations affect simulated DRT outcomes when operating conditions are held fixed. This matters because matching synthetic demand to empirical demand in aggregate volume alone does not ensure consistent representation of temporal, spatial, within-day, or user-level demand structures.

The third gap concerns what live operational and self-reported user data can establish about realised DRT performance and use. Operational studies already examine aspects such as demand, pooling, service context, user characteristics, and reported travel alternatives [4, 7, 16, 17]. However, studies address these dimensions in

different combinations across contexts, making integrated interpretation difficult. In particular, longitudinal analyses of DRT performance and use that consider realised pooling and pickup reliability together with user-level patterns of use and reported travel alternatives remain limited.

The thesis addresses selected parts of this broader problem through three complementary studies. These comprise a systematic mapping of DRT modelling research, a simulation study comparing alternative synthetic demand-generation models under fixed operating conditions, and an observational analysis of longitudinal operational and self-reported user data from a real-world DRT service. The empirical components of the studies were enabled by collaboration with Västtrafik through the Policy Lab for Public Transport On-Demand, providing access to operational data and contextual information needed to understand and interpret the service and its records. Together, the studies address how DRT is represented in modelling research, how passenger-demand representation affects simulated outcomes, and how performance and use can be analysed in live operation. The studies are not stages of a single experiment and do not validate one another; rather, each addresses a different part of the methodological problem using a method suited to that question.

Taken together, the thesis contributes a methodological basis for analysing selected aspects of DRT across modelling and live operation. First, it structures how service and operational design aspects and choices are represented and evaluated through performance metrics in modelling research. Second, it examines how alternative empirically grounded demand-generation models affect simulated outcomes under fixed operating conditions. Third, it shows how operational and self-reported data can support analysis of realised performance, patterns of use, and the travel role of a DRT service. The contribution is intentionally scoped to these aspects rather than intended as a universal framework for DRT evaluation.

1.4 Thesis Organisation

The remainder of Part I is organised as follows. Chapter 2, *Background and Analytical Concepts*, introduces the concepts required for the thesis. Chapter 3, *Research Aim and Research Questions*, presents the overall aim, the main research question, the three research questions, and their relationship to the included papers. Chapter 4, *Research Design and Methods*, describes the systematic mapping, simulation, and observational analyses. Chapter 5, *Contribution*, presents the contributions in relation to the research questions and the thesis-level synthesis. Chapter 6, *Discussion*, synthesises the findings from the three studies and discusses their implications, limitations, and transferability. Chapter 7, *Conclusions and Future Research*, presents the main conclusions and directions for future research. The references follow Part I, while Part II contains the three appended papers.

2 Background and Analytical Concepts

2.1 DRT: Service Concept and Context

In this thesis, demand-responsive public transport (DRT) is used as an umbrella term for shared, request-based public transport in which routes, schedules, or pickup and drop-off locations can be adapted to passenger demand [2]. Unlike conventional fixed-route public transport, its operation is not fully predetermined by a fixed route and timetable. The degree and form of flexibility vary between services, for example in how routes, schedules, and pickup and drop-off locations are determined. Services can also differ in service-area configuration and in the rules used to handle and assign passenger requests. Existing classifications therefore distinguish DRT services according to dimensions such as responsiveness, service type, and the flexibility of routes, timetables and stops [2].

Shared operation is an important characteristic of the DRT concept used here, but it does not imply that every realised trip is pooled. When demand is low, or passenger requests cannot be feasibly combined, a vehicle may serve a single passenger request, so an individual trip may operationally resemble a taxi-like service. This thesis therefore draws the distinction at the service level: DRT is organised as request-based public transport with the possibility of shared operation, even though sharing may not be realised on every trip.

DRT can also play different roles within the wider public transport system. Depending on its design and local context, it may provide local mobility, connect passengers to scheduled services, or complement fixed-route provision where coverage or frequency is limited [3, 4]. Figure 2.1 illustrates one possible stop-based DRT configuration, showing how DRT can support local mobility and connect with conventional fixed-route public transport.

Rural, interurban, and other low-density settings are important DRT applications because potential passengers and destinations may be spatially dispersed and demand may be insufficient to support frequent fixed-route service throughout the day [1, 16]. This provides a rationale for considering flexible service rather than evidence that DRT is invariably preferable; suitability assessments depend on the service context and the conditions under which it is intended to operate [9]. In this thesis, Buss on Demand (BoD) refers to the Swedish DRT pilot examined in Papers II and III, whereas DRT refers to the broader service concept.

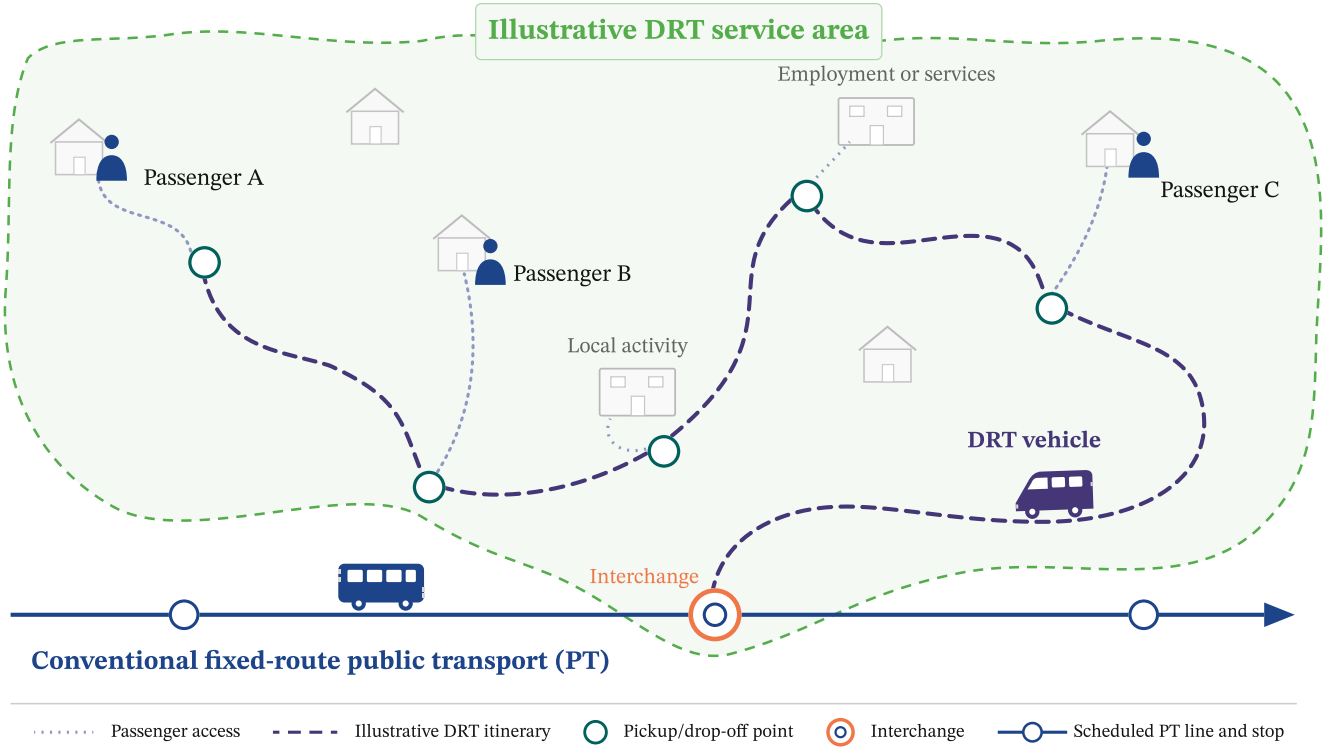

Figure 2.1: Illustrative stop-based DRT configuration in a low-density or dispersed setting, showing local mobility and a possible connection to fixed-route public transport.

2.2 Representing Service Design and Passenger Demand

Service and operational design aspects describe configurable features of a DRT service. Some influence how passengers can access and use the service, for instance, booking arrangements, pickup and drop-off options, and fares. Others concern how the service is operated and resourced, including routing and scheduling, service-area configuration, fleet resources, and integration with other transport services. These are examples of relevant decision areas, not an exhaustive taxonomy. Within the adjacent dial-a-ride problem (DARP) tradition, particular attention is given to assigning, routing, and scheduling passenger requests subject to conditions such as vehicle capacity, time windows, and ride-time limits [10, 11].

Simulation requires representing these service arrangements through explicit assumptions about demand, vehicles, network conditions, operating rules, and constraints. Agent-based approaches can explicitly represent individual travellers and vehicle operations, enabling the examination of their interaction under specified service conditions [8]. Simulation outcomes are therefore conditional on both the service configuration represented in the model and the demand supplied to it.

Passenger-demand representation concerns the characteristics of the travel requests supplied to a model or simulation. Demand is characterised not only by total request volume; relevant dimensions can include request timing, origin-destination structure, user or group composition, regularity, and dependence between trips made within the same day [12, 18]. For instance, representations that treat trips independently may omit sequential relationships such as trip chaining, in which one journey is related to subsequent travel within the same activity sequence [18]. Empirically derived activity and trip-chain data can preserve temporal and spatial structures that more aggregated demand representations may lose [12].

Demand inputs can be constructed in different ways depending on the purpose of the simulation. Empirical trip or activity data can be used to construct demand inputs, while synthetic approaches can generate new demand from statistical or other empirically derived information [12, 13]. Synthetic demand can be useful when scenario analysis requires alternative demand patterns or when direct replay of historical observations does not provide the flexibility needed for the intended experiment. Neither approach represents demand independently of the data and assumptions on which it is based; its suitability depends on which demand characteristics the simulation needs to represent. Paper II builds on these considerations by comparing alternative synthetic demand-generation models that represent different temporal, spatial, within-day, and user-level structures.

2.3 Performance Concepts for DRT Evaluation

For this thesis, DRT performance is treated as multidimensional because evaluation may concern passengers, service operation, resource use, and wider system outcomes. Following the performance-metric framework developed in Paper I, related measures are organised into five key performance indicator (KPI) domains: passenger time and access; reliability and request fulfilment; cost and vehicle operation; utilisation and pooling; and environment and energy. These domains provide a common analytical structure for organising and relating the different performance measures encountered across the thesis while retaining their study-specific definitions. They therefore support comparison and synthesis in this thesis rather than serving as a universal or exhaustive standard for DRT evaluation.

Within this framework, the passenger time and access domain covers measures such as waiting time, in-vehicle travel time, walking or access time or distance to pickup and drop-off points, and detour/excess travel, meaning additional passenger travel time or distance relative to a direct or shortest-path trip. Reliability and request fulfilment concern whether requested service is provided and whether service is delivered relative to stated temporal or service criteria. Pickup-related reliability measures used later in the thesis are therefore interpreted relative to their explicitly defined reference times and criteria. Cost and vehicle operation include monetary costs and vehicle-operation measures such as vehicle kilometres, dwell or stop time, and idle or empty waiting. Utilisation and pooling concern vehicle occupancy, shared travel, and related fleet-utilisation measures. The environment and energy domain covers outcomes such as emissions and energy use, although the present thesis does not empirically evaluate them.

Individual performance indicators must be interpreted alongside their definitions and analytical context. For example, a pooling rate alone does not fully capture operational performance, as its interpretation can depend on factors such as passenger distance, detours, occupancy, vehicle distance, and empty running [14]. The basis used for service comparison can also affect interpretation: service activity expressed relative to population, geographic area, or population density can lead to different conclusions [7]. The same performance concept may therefore be defined or measured differently across studies. Comparisons of DRT performance should therefore account for how measures are defined, what denominators they use, and whether they come from simulation or observed service operation.

3 Research Aim and Research Questions

3.1 Overall Aim and Main Research Question

DRT performance and use can be interpreted differently depending on how service and operational choices and passenger demand are represented, which performance outcomes are measured, and how realised operation and use are analysed. Against this background, the overall aim of the thesis is:

> *To improve the methodological basis for modelling and evaluating selected aspects of DRT performance and use.*

This aim is addressed through the following main research question:

Main RQ. ***How can modelling and evaluation of DRT support analysis of service and operational design choices, performance, and use?***

3.2 Research Questions

The Main RQ is addressed through three questions, each focusing on a different part of the overall methodological problem.

RQ1. ***Which service and operational design aspects are considered relevant in DRT modelling and evaluation, and how are design choices within these aspects represented and evaluated?***

Here, "relevant" refers to aspects identified and examined in the reviewed literature rather than to a normative assessment of which aspects should be prioritised in practice. "Represented and evaluated" distinguishes the inclusion of a design choice in a model from its assessment through performance measures.

RQ2. ***How can selected patterns in empirical passenger demand be represented in synthetic demand-generation models, and how do these models affect simulated DRT performance?***

Here, "selected patterns" refers to temporal, spatial, within-day trip-chaining, and user-level regularity represented in the empirical passenger demand. The question concerns the effects of complete demand-generation models under fixed DRT operating conditions rather than isolated model components.

RQ3. ***How can operational booking data and self-reported user data be used to analyse realised DRT performance, patterns of use, and the service's travel role?***

Here, "realised DRT performance" refers to realised pooling and pickup reliability, while "patterns of use" refers to BoD-use intensity in relation to private mobility resources; reported travel alternatives inform the service's travel role.

3.3 Relationship Between the Research Questions and Papers

Table 3.1 shows how each research question is addressed by the included papers.

Table 3.1: Relationship between the research questions and included papers

RQ	Paper	Research approach	Role in the thesis
RQ1	I	Systematic mapping review and cross-study analysis	Identifies service and operational design aspects represented in DRT modelling studies and examines how design choices within these aspects are represented and evaluated through performance measures.
RQ2	II	Synthetic-demand modelling and simulation	Examines how alternative synthetic demand-generation models represent selected patterns in empirical passenger demand and how these models affect simulated DRT performance under common operating conditions.
RQ3	III	Observational analysis of operational and self-reported user data	Characterises realised DRT performance, patterns of use, and the service's travel role using operational and self-reported user data.

4 Research Design and Methods

"Essentially, all models are wrong, but some are useful."
— George E. P. Box

4.1 Methodological Foundations and Research Design

This thesis sits within applied computing research, focusing on how computational models and empirical data can support modelling and evaluating DRT. Design and creation is one research strategy in computing in which artefacts can include models and methods, and it may be combined with other research strategies [19]. In this thesis, the design-and-creation component is most explicit in Paper II, where we construct alternative computational demand-generation models from empirical service data and evaluate them through simulation. Paper I uses a systematic mapping study to structure and synthesise previous DRT modelling research, while Paper III uses an observational case study of a real-world DRT service to examine realised operation and use.

The research approach used for each study follows from the corresponding research question. RQ1 requires a structured examination of how previous DRT modelling research represents and evaluates service and operational design choices. RQ2 requires constructing and comparing alternative passenger-demand representations while keeping the DRT operating conditions common. RQ3 requires analysis of realised performance and use in an operating DRT service. The thesis therefore combines systematic mapping, computational model construction and simulation, and observational analysis rather than relying on a single research strategy.

The empirical components of Papers II and III are grounded in Västtrafik's BoD pilot and were conducted in connection with the Policy Lab for Public Transport On-Demand. The collaboration provided access to operational data and contextual information used to understand the service, clarify data fields and records, and support data curation and interpretation. This interaction with service actors helped us understand the problem and interpret the empirical data.

Table 4.1 summarises the principal research approach, data source, and analytical role of each study.

Table 4.1: Methodological positioning of the included papers.

Paper	Research approach	Main data source	Analytical role
I	Systematic mapping and cross-study analysis	Published DRT modelling studies	Structured classification and mapping of design representation and performance evaluation.
II	Computational demand modelling and simulation	Empirical passenger-demand data and generated simulation outputs	Construction and comparison of alternative demand-generation models under common DRT operating conditions.
III	Observational case study of a real-world DRT pilot	Operational booking records and self-reported questionnaire data collected by Västtrafik	Characterisation of realised DRT performance and use and estimation of adjusted associations in observed service operation.

4.2 Research Approaches in the Included Papers

4.2.1 RQ1. Which service and operational design aspects are considered relevant in DRT modelling and evaluation, and how are design choices within these aspects represented and evaluated?

RQ1 concerns which service and operational design aspects are considered relevant in DRT modelling and evaluation and how design choices within these aspects are represented and evaluated in previous studies. A systematic mapping study is suitable for this purpose because it supports structured classification and mapping of a research area rather than quantitative effect synthesis [20]. Paper I therefore develops and applies two coding frameworks: one for service and operational design aspects and one for performance metrics.

Following the search and screening process described in Paper I, the quantitative mapping includes 33 primary modelling studies. Each study could receive multiple design-aspect and KPI codes. This coding helps us examine which aspects and performance measures the literature represents, and which design aspects and KPI domains occur within the same modelling or scenario contexts.

For the thesis-level synthesis, we additionally re-examined the 33 primary studies with respect to individual design choices. We classified choices by role as fixed assumptions, optimisation decision variables, varied scenario parameters, sensitivity parameters, constraints, or unclear cases. We recorded a direct choice-KPI link only when a performance measure was explicitly used as an objective, feasibility check,

or reported outcome in the corresponding choice-specific experiment or comparison. Therefore, the presence of a design choice and a KPI in the same study was not sufficient to establish a direct evaluative link.

The modelling study remains the unit of analysis for the systematic mapping presented in Paper I. The original design aspect-KPI relationships are therefore interpreted as study-level coverage and co-occurrence. The additional thesis-level analysis complements this mapping by identifying how individual design choices are used in the models and which performance measures directly evaluate them.

4.2.2 RQ2. How can selected patterns in empirical passenger demand be represented in synthetic demand-generation models, and how do these models affect simulated DRT performance?

RQ2 concerns how selected patterns in empirical passenger demand can be represented in synthetic demand-generation models and how these models affect simulated DRT performance. The study therefore combines constructing computational demand models with evaluating them through simulation. Simulation is appropriate for this question because it supports comparison of alternative model configurations under specified experimental conditions [21]. In Paper II, we keep the DRT operating configuration common across the demand-model variants.

Paper II uses empirical passenger requests from the Mark BoD service to construct the demand-generation models. Observed empirical demand (ED) represents the recorded passenger requests, while simulated empirical demand (SED) replays these requests in the common simulation environment and provides a simulation reference. Four synthetic demand-generation models, denoted V1-V4, represent different structures observed in the empirical demand: V1 uses independent sampling of trip characteristics; V2 introduces within-day trip dependence; V3 represents persistent user participation and user-level temporal preferences; and V4 extends this with repeated origin-destination behaviour and empirical within-day spacing. The four models are alternative demand-generation model variants rather than cumulative stages or a factorial decomposition of individual model components.

We evaluate all demand models using the same MATSim-based DRT operating configuration, holding the fleet and other operating conditions fixed while varying the passenger-demand representation. This allows us to examine differences in simulated passenger and vehicle outcomes under common operating conditions. Empirical service data therefore ground the construction of the demand representations, while simulation provides the environment in which we evaluate their consequences.

Constructing the synthetic demand-generation models constitutes the design-and-creation component identified in Section 4.1, while we evaluate them through simulation. Together, model construction and simulation provide the methodological basis for addressing RQ2.

4.2.3 RQ3. How can operational booking data and self-reported user data be used to analyse realised DRT performance, patterns of use, and the service's travel role?

RQ3 concerns how operational booking data and self-reported user data can be used to analyse realised DRT performance, patterns of use, and the service's travel role. Paper III addresses this question through an observational case study of the operating BoD services in Mark and Ulricehamn. The BoD pilot constitutes the empirical case, while the analysis is observational.

The study draws on 67,782 booking records, including 44,855 completed bookings, together with self-reported questionnaire data collected by Västtrafik. The operational records provide repeated observations of realised service operation across the two service areas and are used to characterise pooling and pickup reliability and to examine their associations with trip, demand, temporal, and contextual characteristics. The questionnaire data complement the operational records with information on private mobility resources, BoD-use intensity, and reported travel alternatives not available from booking records alone.

The analyses combine descriptive statistics with multivariable logistic regression and sensitivity analyses. The pooling and pickup-reliability models account for repeated observations from individual users, while sensitivity analyses examine alternative observation windows, lateness thresholds, and data specifications. For the user-level questionnaire analyses, we account for differences in observation time when measuring use intensity, and we use response weighting to assess sensitivity to observable differences between respondents and nonrespondents.

The observational case-study design enables examination of realised service performance and use within the context in which the BoD service operates. For instance, without experimental assignment, observed associations may reflect measured or unmeasured differences between observations in addition to the variables included in the analysis [22].

5 Contribution

5.1 RQ1. Which service and operational design aspects are considered relevant in DRT modelling and evaluation, and how are design choices within these aspects represented and evaluated?

Paper I maps 21 service and operational design aspects and 12 performance metrics grouped into five KPI domains across 33 primary modelling studies. Representation is concentrated around the operational core: fleet-management aspects occur in all 33 studies and route and operation in 31, while passenger time and access is the most widely represented KPI domain. Booking and fare, access and user-context aspects, and environmental and energy outcomes receive substantially less coverage.

Paper I maps relationships between design aspects and KPI domains at study level, while its supplementary coding also records which specific choices or parameters within the design aspects are represented. For the thesis-level synthesis, we re-examined these design-choice-level codes to classify how each choice was used in the model: as a fixed assumption, optimisation decision variable, varied scenario parameter, sensitivity parameter, constraint, or unclear case. We also identified a design-choice-specific KPI link when a performance measure evaluated that choice in the corresponding optimisation, scenario, or sensitivity analysis; the mere co-occurrence of the choice and KPI in the same study was not sufficient.

In total, 25 of the 33 studies deliberately varied at least one design choice through scenario comparison, sensitivity analysis, or both. When optimisation decision variables are also included, all 33 studies use at least one design choice as an optimisation decision variable, scenario parameter, or sensitivity parameter. Across all coded design choices in the 33 studies, we identified 33 instances of optimisation decision variables, 37 instances of scenario variation, and 10 instances of sensitivity analysis. A study can contribute more than one instance by evaluating several design choices or by using the same choice in more than one role. Table 5.1 shows the five design choices most frequently used in these roles.

Table 5.1: Most frequently used design choices or parameters and the performance measures used in their evaluation across the 33 primary modelling studies.

Design choice / parameter	Studies	Use of design choice			Performance measures used in evaluation						
	(n)	DV	Scen.	Sens.	KPI-1	KPI-3	KPI-4	KPI-6	KPI-7	KPI-8	KPI-11
Advance Scheduling (ST-1)	8	8	0	0	–	2	2	–	8	–	–
Dynamic Scheduling (ST-2)	10	5	5	1	6	2	3	4	3	3	2
Fleet Size (OC-1)	13	8	4	1	10	5	5	6	9	3	8
Vehicle Capacity (OC-2)	8	2	4	2	7	4	3	4	6	1	5
Ride Time / Detour Limit (OC-5)	6	0	2	4	2	4	2	2	3	1	4

DV = optimisation decision variable; *Scen.* = scenario parameter; *Sens.* = sensitivity parameter. The *Studies* column reports the number of unique studies in which the design choice was used in at least one of these three roles. A study may use the same design choice in more than one role; therefore, the DV, Scen., and Sens. counts need not sum to the number of studies.

Advance Scheduling (ST-1) refers to scheduling organised in advance using collected requests, whereas Dynamic Scheduling (ST-2) refers to scheduling adjusted in real time in response to incoming demand. These categories describe how the service schedule is formed or updated, not whether passengers can book in advance or near the time of travel.

Performance measures: KPI-1 = Waiting Time; KPI-3 = Detour / Excess Travel; KPI-4 = Travel / Ride Time; KPI-6 = Request Acceptance; KPI-7 = Cost of Service; KPI-8 = VKT (vehicle kilometres travelled); KPI-11 = Utilisation / Pooling. The values report the number of studies in which the corresponding performance measure was used to evaluate that specific design choice; a dash indicates that no such evaluation was identified.

The performance measures used to evaluate these design choices are concentrated on a relatively limited set of outcomes. Fleet size is most frequently evaluated through waiting time, cost of service, utilisation/pooling, and request acceptance, while vehicle capacity is evaluated mainly through waiting time, cost of service, and utilisation/pooling. Among studies using these scheduling configurations, dynamic scheduling is most frequently evaluated through waiting time and request acceptance, whereas advance scheduling is evaluated predominantly through cost of service. Ride time/detour limit is most frequently evaluated through detour/excess travel and utilisation/pooling.

The design-choice-level analysis also identifies dimensions that receive little direct evaluation. No optimised or deliberately varied design choice was evaluated using Dwell / Stop Time, and Emissions / Energy Use was used to evaluate such choices in only two studies. No choice under Accessibility and Inclusivity, Booking Method, or Booking Type was used as an optimisation decision variable, scenario parameter, or sensitivity parameter. Where represented, these characteristics may instead appear as fixed assumptions or contextual features of the model.

For RQ1, the results therefore provide two linked levels of analysis. At study level, the systematic mapping shows which service and operational design aspects and performance measures are represented together in DRT modelling research. At the design-choice level, the thesis-level synthesis shows which specific design choices are used and which performance measures evaluate them. The results indicate a strong emphasis on operational choices such as fleet size, scheduling, vehicle capacity, and operational constraints, with evaluation concentrated mainly on passenger-time, cost, request-acceptance, and utilisation outcomes. The reviewed literature evaluates booking, accessibility, and environmental dimensions less often through deliberately varied choices. This is particularly notable for booking arrangements, including booking methods (e.g., app or telephone) and booking types (e.g., on-demand or advance booking). Such choices may affect service accessibility and use, but the reviewed modelling studies do not directly establish their effects. These patterns characterise how the reviewed literature evaluates DRT design choices, rather than which choices are most important in practice.

5.2 RQ2. How can selected patterns in empirical passenger demand be represented in synthetic demand-generation models, and how do these models affect simulated DRT performance?

Paper II examines how selected patterns in empirical passenger demand can be represented in synthetic demand-generation models and how the resulting models affect simulated DRT performance. Four alternative demand-generation variants represent different demand structures. V1 independently samples trip attributes

from empirical marginal distributions and does not represent within-day trip dependence or user-level memory. V2 represents within-day trip dependence through trip chaining, empirical inter-trip gaps, and conditional spatial transitions. V3 represents persistent user participation and user-level temporal regularity through participation tendencies and preferred pickup times. V4 combines persistent user participation and temporal regularity with habitual origin-destination repetition and within-day temporal spacing. The variants are modular alternatives rather than cumulative stages, and therefore represent different combinations of demand characteristics.

To compare the simulated outcomes produced by these representations, we compared the four synthetic variants with simulated empirical demand (SED), in which we replayed the observed requests through the same MATSim DRT simulation. Although V1-V4 reproduced broadly comparable aggregate demand volumes, they did not produce equivalent DRT outcomes. The differences were most evident in passenger waiting and pickup performance and in vehicle-operation and utilisation measures. Thus, reproducing aggregate demand volume alone was insufficient to reproduce the simulated performance obtained from the empirical request structure.

Waiting-time results illustrate these differences. The independent-sampling representation (V1) reproduced empirical marginal demand distributions, but did not reproduce SED-like waiting performance. The representation that included within-day trip dependence through trip chaining (V2) was closest to SED for mean waiting on both weekdays and weekends and also aligned closely with mean in-vehicle time. Among the evaluated alternatives, the demand representation including within-day trip dependence was associated with closer alignment on several passenger-experience outcomes. The differences were larger in the upper tail: the 95th-percentile waiting time increased from 787 s for SED to 876-914 s across V1-V4 on weekdays, and from 481 s to 643-675 s on weekends. The larger weekend differences align with the demand-intensity analysis, which showed particularly pronounced waiting-time differences under lower demand. Figure 5.1 summarises these and other selected performance differences relative to SED.

The demand representations also affected vehicle-operation and utilisation outcomes. Across V1-V4, total vehicle kilometres travelled were about 5-9% higher than for SED, while passenger kilometres per vehicle kilometre (PKM/VKT) were about 2-6% lower. The representation with persistent user participation and user-level temporal regularity (V3) aligned most closely with SED on several distance-based vehicle-operation and utilisation measures, including PKM/VKT, particularly on weekends. For pooling, the variants representing persistent users and temporal regularity (V3 and V4) produced pooled-trip shares closer to SED than the independent-sampling and trip-chaining representations (V1 and V2), while their stronger temporal concentration coincided with less favourable waiting-time performance. No synthetic demand variant therefore aligned most closely with SED across all evaluated outcomes.

The comparison also shows that model choice depends on both the intended analysis and the available data. V1 can be constructed from empirical trip-level distributions

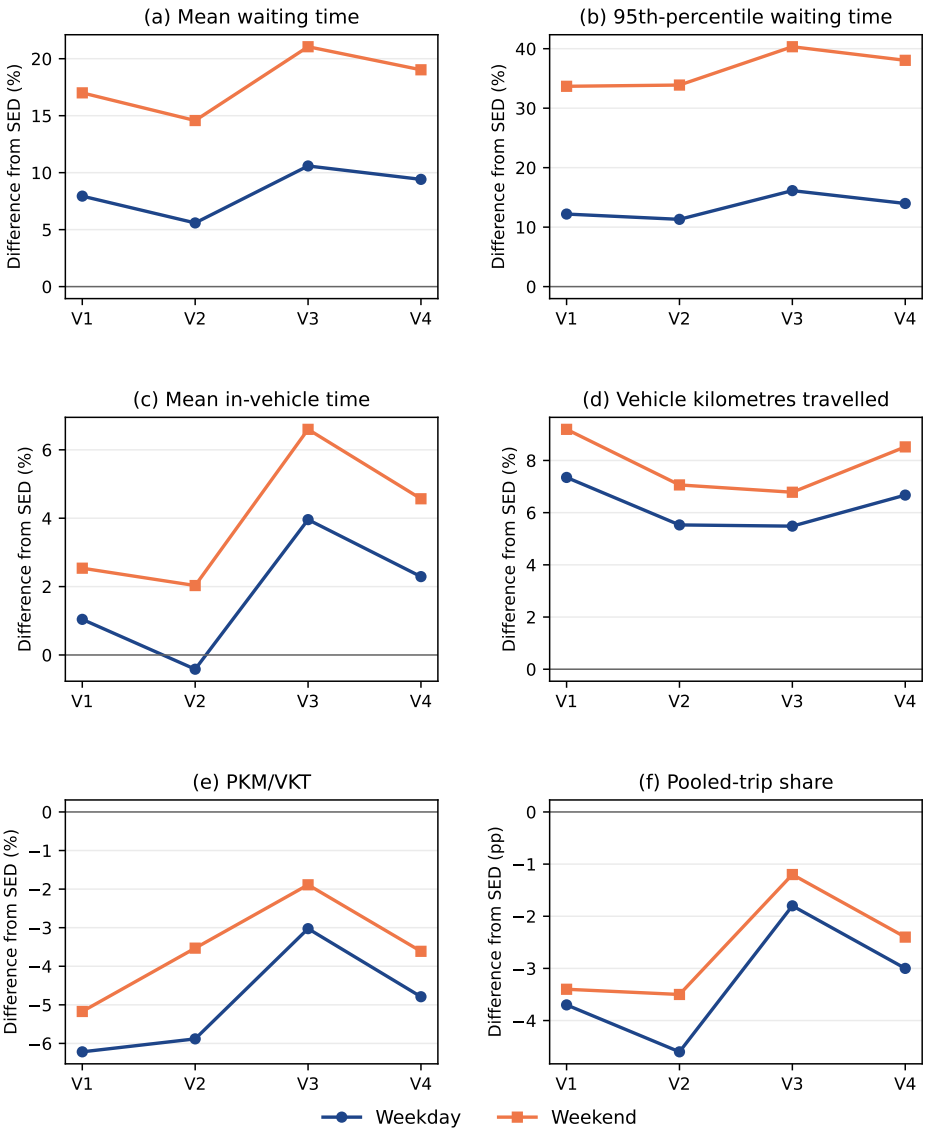

Figure 5.1: Differences between the synthetic demand models and SED for selected simulated outcomes. Panels (a)-(e) show relative differences; panel (f) shows pooled-trip share in percentage points. Zero denotes the SED reference.

without persistent user identities. V2 additionally requires identifiable users or trip-linkage information to derive within-day trip chains and inter-trip gaps. V3 and V4 require repeated user observations, while V4 additionally relies on user-level information about repeated origin-destination behaviour. The level of demand structure represented should therefore reflect both the simulation purpose and the information supported by the available data.

The comparison draws on 518 observed service days and 1,000 independently generated days per day type for each synthetic variant, rather than on individual demand realisations. Together with the broadly stable model patterns across low-, medium-, and high-demand conditions, this strengthens the robustness of the findings within the studied service context.

5.3 RQ3. How can operational booking data and self-reported user data be used to analyse realised DRT performance, patterns of use, and the service's travel role?

Paper III shows how operational booking records and self-reported user data can be used to examine different aspects of realised DRT performance and use. The booking records are used to analyse realised pooling and pickup reliability, while the questionnaire responses are used to compare BoD use across selected private mobility resources and to describe the travel alternatives respondents reported they would have used if BoD had not been available. Together, the two data sources address operational performance and user-related aspects of RQ3.

For realised pooling, the unadjusted pooling shares (i.e., without accounting for other measured characteristics) differed substantially between the two service areas. Across the full observation periods, 61.4% of eligible completed bookings in Mark and 42.1% in Ulricehamn were recorded as shared. During the common observation window, the corresponding shares were 52.7% and 45.4%, reducing the difference from 19.3 to 7.3 percentage points. The multivariable models then compared the service areas while accounting for trip distance, booking lead time, demand frequency, time of day, weekend, and operational age. Despite the lower unadjusted pooling share in Ulricehamn, the Ulricehamn indicator was associated with moderately higher adjusted pooling odds across all three model specifications. Operational age here refers to the number of calendar months since the first completed trip observed in the corresponding service area, not passenger age. The change between the unadjusted and adjusted comparisons shows why service-area pooling differences need to be interpreted together with observation period and measured trip and demand characteristics.

Trip distance showed the strongest and most consistent association with realised pooling. Across the model specifications, a one-standard-deviation increase in log-

transformed trip distance was associated with approximately 2.6-2.7 times higher pooling odds. In practical terms, longer trips were more likely to be recorded as shared than shorter trips after accounting for the other measured characteristics. Operational age was also consistently associated with pooling, with approximately 23% higher odds per one-standard-deviation increase, indicating that pooling was more likely at later stages of the observed service period. Exact origin-destination pair frequency was not consistently associated with pooling. Pickup-origin frequency, defined as the number of completed bookings sharing the same pickup stop, showed a smaller positive association, corresponding to about 9% higher pooling odds per one-standard-deviation increase.

Pickup reliability showed a different pattern. The primary analysis defined pickup lateness as an actual pickup occurring more than five minutes after the assigned pickup time; a 15-minute threshold was also examined as a sensitivity analysis. Booking lead time and operational age were the only factors with consistent associations across the different observation windows and the 5- and 15-minute lateness definitions. A one-standard-deviation increase in log-transformed booking lead time was associated with 32-43% higher odds of pickup lateness, meaning that bookings made further in advance were more likely to be picked up more than the specified time after their assigned pickup time. In contrast, a one-standard-deviation increase in operational age was associated with 20-34% lower lateness odds, indicating that pickup lateness was less likely at later stages of the observed service period. Associations involving service area, trip distance, and shared-booking status changed across the observation periods or lateness definitions and were therefore not consistent. The results consequently do not show that longer or shared trips were systematically more likely to experience pickup lateness.

The user-level results add information on how BoD use differed across selected private mobility resources. After accounting for unequal post-entry observation periods, respondents without a driving licence or household car used BoD more frequently than their respective comparison groups. After response weighting, mean use was 2.09 completed trips per 30 available days for respondents without a driving licence and 1.25 for licence holders. For household-car availability, the corresponding rates were 2.43 for respondents without a household car, 1.36 for those with one car, and 1.17 for those with multiple cars. The ordering therefore remained after weighting for observable questionnaire-response differences. These comparisons concern selected private mobility resources and do not establish that lacking a driving licence or household car caused higher BoD use.

The questionnaire responses also describe the travel alternatives respondents said they would have used if BoD had not been available. Among the valid responses, 33.9% reported a private motorised alternative, 28.9% an active-mode alternative, 26.2% public or shared transport, and 11.1% stated that the trip would not otherwise have occurred. Restricting the analysis to the first valid response from each user changed these grouped shares only modestly. The distribution therefore shows that respondents associated BoD with several different travel alternatives rather than

one dominant substitute. Because the question concerns what respondents said they would have done if BoD had not been available, rather than behaviour observed in that situation, these results represent stated travel alternatives rather than observed mode substitution or population-level travel shares.

For RQ3, the results show that operational booking records can support detailed analysis of realised pooling and pickup reliability when trip characteristics, observation periods, repeated users, and outcome definitions are considered. Self-reported user data add information that the booking records alone do not contain, particularly about private mobility resources and what respondents report they would have done if BoD had not been available. Such information is relevant because travel alternatives and related mobility-resource indicators are not consistently collected or defined across DRT services. Together, the two data sources therefore provide a broader account of realised performance and use than either source alone.

5.4 Thesis-Level Synthesis and Contribution

RQ1 concerns which service and operational choices modelling research represents and evaluates; RQ2 concerns how passenger-demand representation affects simulated performance; and RQ3 concerns what operational booking records and self-reported user data can reveal about realised performance and use. Together, they connect representation, evaluation, and empirical analysis without treating the three studies as sequential stages or as validation of one another.

Taken together, the findings show that conclusions about DRT performance and use depend on how service and operational choices are represented and evaluated, how passenger demand is represented in simulation, and how realised operation and use are defined and analysed from available data. The systematic mapping shows that modelling studies concentrate evaluation on particular operational choices and performance outcomes. The simulation results show that different representations of passenger demand can produce different performance conclusions even under the same operating conditions. The operational analysis further shows that interpretations of realised performance can change with observation period, measured trip characteristics, and statistical adjustment, while self-reported user data provide information about service use and travel alternatives that is not contained in booking records.

The thesis therefore makes a methodological contribution while also providing findings relevant to selected DRT design and operational considerations. Methodologically, it structures how service and operational design choices are represented and evaluated in DRT modelling research, demonstrates how empirically grounded demand representations affect simulated performance conclusions, and shows how operational booking records and self-reported user data can be analysed to characterise realised performance and use. The findings are also relevant to DRT design and operation by showing which design choices and performance measures receive attention in modelling research and how realised pooling, pickup reliability, and

patterns of use relate to trip, booking, operational, and user characteristics in the studied service. The thesis does not prescribe an optimal DRT design, but it provides a basis for more informed modelling, evaluation, and interpretation of selected design and operational considerations.

6 Discussion

Chapter 5 presented the contributions in relation to the research questions and brought them together at thesis level. This chapter discusses the relationships between those findings, their implications for DRT modelling and evaluation, and the boundaries within which they can be interpreted.

6.1 Representation, Demand, and Performance Evaluation

The cross-paper issue is not simply that service design, passenger demand, and performance measures all matter, but that each places a different boundary on what a DRT performance result means. Service and operational design choices define the operating setting and the alternatives under consideration, passenger-demand representation defines the requests presented to that setting, and performance measures determine which outcomes are evaluated. In operational analysis, the available data, observation period, and outcome definitions similarly shape the basis on which realised performance is interpreted.

Paper I clarifies this distinction at the service and operational design-choice level. A design aspect can be represented in a model even when a particular choice within that aspect is fixed or used only as a constraint, whereas direct evaluation requires the analysis to explicitly optimise, vary, or sensitivity-test the choice and link it to a performance measure. For design-oriented use of modelling studies, this does not mean that every possible design choice and KPI must be included. Instead, each analysis should consider relevant design aspects and select the particular choices and performance measures that correspond to its purpose. How frequently a choice is represented or evaluated in the literature does not by itself establish its practical importance or identify a preferred service design.

Paper II adds passenger-demand representation to this design-evaluation relationship. Under the same operating configuration and broadly comparable aggregate demand levels, alternative empirically grounded demand representations produced different simulated passenger and vehicle outcomes. This aligns with DRT simulation research using demand representations that retain different temporal, spatial, and traveller-level structures [8, 12, 13], while Paper II shows that these representation differences

can also affect performance conclusions under common operating conditions. Together with Paper I, this indicates that evaluating a DRT design choice requires attention not only to which choice and performance measure are selected, but also to how passenger demand is represented, because the same operating configuration can produce different performance outcomes under alternative demand structures.

A further boundary concerns traveller response to the service offered. In Paper II, the generated requests serve as inputs to the DRT simulation; travellers do not subsequently revise their travel decisions in response to the offered pickup time, drop-off time, or travel time. In practice, such service attributes may influence whether a traveller accepts or cancels the offered trip or chooses another mode. The simulated results therefore concern operational performance conditional on the generated request set rather than a feedback process in which the offered service also changes the realised demand.

Paper III shows a related, but analytically different, issue in live-service evaluation. Unadjusted service-area summaries can change in interpretation when observation periods are aligned and measured trip, demand, temporal, and operational characteristics are taken into account, while pickup-related conclusions also depend on how the outcome is defined. This is not the same mechanism as the demand-representation differences examined in Paper II. The common point is that an aggregate similarity or difference does not by itself explain the structure underlying a performance result: in simulation this concerns how demand is represented, whereas in operational analysis it concerns how observed outcomes are composed, defined, and compared.

Taken together, the studies therefore identify different requirements for drawing conclusions about DRT design and performance. Design-oriented conclusions require clarity about which service choices are actually evaluated and against which outcomes; simulation conclusions require passenger-demand representation to be treated as part of the model specification; and operational conclusions need to be interpreted in relation to the available data, observation conditions, and outcome definitions. These distinctions provide the basis for the implications discussed in the following section.

6.2 Implications for DRT Modelling and Evaluation

The implications primarily concern researchers and modellers who construct and evaluate DRT models, but they also matter for public transport authorities and other decision-makers who interpret model-based comparisons. The findings show where modelling choices enter an evaluation and which assumptions need to remain visible when results are used to compare service alternatives or interpret realised performance. In particular, Paper I shows the importance of distinguishing the service and operational choices being examined from the measures used to evaluate them, while Paper II adds passenger-demand representation as another modelling choice that can affect the resulting performance assessment.

Paper II gives a more specific implication for demand modelling. All four synthetic variants draw on the same historical operational dataset, but they retain different temporal, spatial, within-day, and user-level structures and consequently produce different performance patterns under identical operating conditions. Even the independent V1 model estimates its demand levels and trip-attribute distributions from historical records; the distinction is therefore not simply between using and not using empirical data. Rather, the modelling choice concerns which relationships present in those data are retained in the synthetic demand. This means researchers should select and report demand representations based on the analytical objective and the performance outcomes of interest, rather than treating empirical calibration alone as sufficient.

The findings also show why performance measures need explicit definitions when comparing results across models or operational studies. Paper II evaluates pooled-trip share within a controlled simulation, whereas Paper III uses operator-recorded shared-booking status to analyse realised sharing. These measures address a related performance domain but use different definitions and should therefore not be compared numerically as if they represented the same quantity. The same distinction applies to pickup-related performance: Paper II evaluates simulated waiting and pickup outcomes, whereas Paper III measures lateness relative to the assigned pickup time.

Paper III further shows what operational records and self-reported user data can add to model-based evaluation. Booking records support analysis of realised pooling and pickup reliability in relation to observed trip, demand, temporal, and operational characteristics, while the questionnaire data add information about private mobility resources and respondents' reported travel alternatives. For public transport authorities, combining these data sources can therefore broaden the basis for interpreting both service operation and the role the service plays for its users. The interpretation should still follow the type of data available: reported travel alternatives describe what respondents state they would have done without the service rather than observed mode substitution. Related empirical DRT studies likewise combine operational and user-oriented information when examining service performance and mobility context [7, 17].

6.3 Limitations and Transferability

The thesis deliberately addresses selected dimensions of DRT modelling and evaluation rather than all aspects of service design and performance. Paper I identifies comparatively limited modelling attention to areas such as booking and fare, broader accessibility and inclusivity, infrastructure and funding arrangements, and environmental outcomes. Papers II and III address other selected modelling, operational, and user-related questions and should therefore not be interpreted as closing all of the areas identified by the systematic mapping. This bounded scope allows the thesis to examine the selected relationships in greater depth while keeping the conclusions

aligned with the questions and data actually studied.

Each study also sets specific boundaries on interpretation. Paper I derives its findings from the defined review scope and the retained primary modelling studies, while the additional thesis-level design-choice and choice-KPI analysis depends on the applied coding frameworks. Paper II evaluates alternative demand-generation models developed from one empirical service context under a common operating configuration. The comparison supports conclusions about differences among the evaluated models within that simulation setting, rather than effects attributable to individual model components. Paper III uses observational operational and questionnaire data, so the adjusted associations do not establish causal service effects; the pooling and pickup-reliability results also depend on the available operational indicators and outcome definitions, while the questionnaire results remain subject to nonresponse and the stated nature of the reported travel alternatives.

Papers II and III also draw on related Västtrafik and Mark service contexts. This contextual continuity supports a coherent examination of demand representation and realised operation, but numerical findings may differ in services with other geographical settings, demand patterns, service designs, or operating conditions. The thesis's more transferable contribution lies in the methodological lessons across these studies: make explicit which service and operational choices are represented and evaluated, which passenger-demand structures are retained, how performance outcomes are defined, and what the available data support in interpretation. These principles provide a basis for applying the thesis approach in other DRT settings while adapting the specific models, measures, and assumptions to the context being studied.

7 Conclusions and Future Research

7.1 Conclusions

This thesis examined how modelling and evaluation can support analysis of selected aspects of DRT performance and use. Overall, such analysis depends on which service and operational choices are represented and evaluated, how passenger demand is represented, which performance outcomes are defined and measured, and how operational and user data are analysed. Making these choices explicit provides a clearer basis for interpreting modelling results and realised service performance.

The systematic mapping identifies a broad set of service and operational design aspects considered in DRT modelling, while the thesis-level analysis shows that active evaluation is concentrated around a smaller operational subset. Optimisation and variation focus on choices such as fleet size, scheduling, vehicle capacity, and operational constraints, while evaluation focuses mainly on passenger-time, cost and vehicle-operation, request-fulfilment, and utilisation outcomes. Service-facing and environmental dimensions receive less direct evaluative attention.

The controlled simulation study shows that synthetic demand models with broadly comparable aggregate demand can retain different temporal, spatial, within-day, and user-level structures and produce different simulated DRT outcomes under the same operating configuration. Among the evaluated alternatives, no single demand model is closest to the simulated empirical reference across all performance outcomes. The suitable demand representation therefore depends on the empirical structures relevant to the analytical purpose and on the performance outcomes being studied.

Operational booking records support analysis of realised pooling and pickup reliability in relation to trip, demand, temporal, and operational characteristics. Self-reported user data add information about service use in relation to private mobility resources and about what respondents report they would have done if BoD had not been available. Together, these data sources provide a broader account of realised performance and use than either source alone.

Taken together, the findings show that modelling and evaluation can support DRT analysis by structuring which design choices and performance outcomes are examined, testing how passenger-demand representation affects simulated performance, and analysing realised performance and use from operational and self-reported

data. The thesis therefore makes a methodological contribution while also providing findings relevant to selected DRT design and operational considerations. Rather than prescribing one preferred service design or evaluation procedure, it provides a clearer basis for modelling, evaluating, and interpreting selected aspects of DRT in relation to the analytical purpose and service context.

7.2 Future Research

A coherent direction for future research is to extend the present work towards comparative evaluation of DRT service and resource-design choices under varying demand and operating conditions. Ongoing work follows this direction by examining time-dependent fleet allocation, vehicle capacity, and the trade-offs between passenger service and operational efficiency. This extends the thesis from analysing how design choices are represented and evaluated to more directly comparing alternative operational configurations.

A further extension concerns heterogeneous and accessible fleet composition together with workforce constraints. Planned work will examine how combinations of conventional and wheelchair accessible vehicles interact with driver working hours, breaks, and feasible vehicle availability, and how these choices affect service reliability, passenger outcomes, and operator resource use. This will broaden the design space examined in the thesis while remaining closely connected to empirically grounded DRT operation.

The same research direction could then be extended to autonomous DRT. Removing driver-related constraints changes vehicle availability and relocation possibilities, allowing comparison of driver-based services with autonomous services using alternative idle-vehicle relocation strategies. Such work could examine whether conclusions about fleet configuration, service quality, resource efficiency, and demand representation remain stable across different demand levels and operating conditions.

REFERENCES

[1] P. Martí, J. Jordán, M. A. G. Arrieta, and V. Julian, "A survey on demand-responsive transportation for rural and interurban mobility," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 8, no. 3, p. 43, 2023.

[2] P. Vansteenwegen, L. Melis, D. Aktaş, B. D. G. Montenegro, F. S. Vieira, and K. Sörensen, "A survey on demand-responsive public bus systems," *Transportation Research Part C: Emerging Technologies*, vol. 137, p. 103573, 2022.

[3] A. Itani, W. Klumpenhouwer, A. Shalaby, and B. Hemily, "Guiding principles for integrating on-demand transit into conventional transit networks: A review of literature and practice," *Transport Policy*, vol. 147, p. 183–197, Mar. 2024.

[4] M. J. Alonso-González, T. Liu, O. Cats, N. Van Oort, and S. Hoogendoorn, "The potential of demand-responsive transport as a complement to public transport: An assessment framework and an empirical evaluation," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2672, p. 879–889, Aug. 2018.

[5] J. Mageean and J. D. Nelson, "The evaluation of demand responsive transport services in europe," *Journal of Transport Geography*, vol. 11, p. 255–270, Dec. 2003.

[6] G. Currie and N. Fournier, "Why most drt/micro-transits fail – what the survivors tell us about progress," *Research in Transportation Economics*, vol. 83, p. 100895, Nov. 2020.

[7] B. Kaufman, A. Leung, and M. Burke, "Evaluating demand responsive transit services using a density-based trip rate metric," *Journal of Transport and Land Use*, vol. 14, p. 499–519, Apr. 2021.

[8] N. Ronald, R. Thompson, and S. Winter, "Simulating demand-responsive transportation: A review of agent-based approaches," *Transport Reviews*, vol. 35, p. 404–421, Mar. 2015.

[9] H. Wang, J. Li, P. Wang, J. Teng, and B. P. Y. Loo, "Adaptability analysis methods of demand responsive transit: a review and future directions," *Transport Reviews*, vol. 43, p. 676–697, Jan. 2023.

[10] Y. Molenbruch, K. Braekers, and A. Caris, "Typology and literature review for dial-a-ride problems," *Annals of Operations Research*, vol. 259, p. 295–325, May 2017.

[11] S. C. Ho, W. Szeto, Y.-H. Kuo, J. M. Leung, M. Petering, and T. W. Tou, "A survey of dial-a-ride problems: Literature review and recent developments," *Transportation Research Part B: Methodological*, vol. 111, p. 395–421, May 2018.

[12] P. Franco, R. Johnston, and E. McCormick, "Demand responsive transport: Generation of activity patterns from mobile phone network data to support the operation of new mobility services," *Transportation Research Part A: Policy and Practice*, vol. 131, p. 244–266, Jan. 2020.

[13] S. Dytckov, F. Lorig, P. Davidsson, J. Holmgren, and J. Persson, "Modelling commuting activities for the simulation of demand responsive transport in rural areas," in *Proceedings of the 6th International Conference on Vehicle Technology and Intelligent Transport Systems*, p. 89–97, SCITEPRESS - Science and Technology Publications, 2020.

[14] C. Liebchen, M. Lehnert, C. Mehlert, and M. Schiefelbusch, "Ridepooling-effizienz messbar machen: Betriebliche effizienzgrößen für ridepooling-systeme," *Der Nahverkehr*, 2020. Finanzen and Recht section.

[15] R. Kucharski and O. Cats, "Exact matching of attractive shared rides (exmas) for system-wide strategic evaluations," *Transportation Research Part B: Methodological*, vol. 139, p. 285–310, Sept. 2020.

[16] L. Sörensen, A. Bossert, J.-P. Jokinen, and J. Schlüter, "How much flexibility does rural public transport need? – implications from a fully flexible drt system," *Transport Policy*, vol. 100, p. 5–20, Jan. 2021.

[17] V. T. Thao, S. Imhof, and W. von Arx, "Demand responsive transport: New insights from peri-urban experiences," *Travel Behaviour and Society*, vol. 31, p. 141–150, Apr. 2023.

[18] R. Kitamura, "Sequential, history-dependent approach to trip-chaining behavior," *Transportation Research Record*, vol. 944, pp. 13–22, 1983.

[19] B. Oates, M. Griffiths, and R. McLean, *Researching Information Systems and Computing*. SAGE Publications Ltd, 2022.

[20] K. Petersen, S. Vakkalanka, and L. Kuzniarz, "Guidelines for conducting systematic mapping studies in software engineering: An update," *Information and Software Technology*, vol. 64, p. 1–18, Aug. 2015.

[21] A. Law, *Simulation Modeling and Analysis*. McGraw-Hill series in industrial engineering and management science, McGraw-Hill Education, 2015.

[22] P. R. Rosenbaum, *Design of Observational Studies*. Springer International Publishing, 2020.

This thesis examines how modelling and evaluation can support the analysis of selected aspects of demand-responsive public transport (DRT) performance and use. It brings together a mapping of modelling research, simulations using alternative passenger-demand representations, and an empirical analysis of a Swedish Bus-on-Demand (BoD) pilot.

Tryckt av Media-Tryck, Lund 2026 Svanenmärkt trycksak, 3041 0903

MALMÖ UNIVERSITY PRESS
205 06 MALMÖ, SWEDEN
MAU.SE

ISBN 978-91-7877-829-4
DOI 10.24834/isbn.9789178778300