Imperial College London
Department of Civil & Environmental Engineering

Performance Optimisations for Urban Last-Mile Distribution using Autonomous Vehicle Fleets

Yubin Liu

Submitted in part fulfilment of the requirements for the degree of
Doctor of Philosophy at the Imperial College London, June 2024

I hereby declare that this thesis and the work reported herein was composed by and originated entirely from me. Information derived from the published and unpublished work of others has been acknowledged in the text and references are given in the list of sources.

Yubin Liu (2024)

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-NonCommercial 4.0 International Licence (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Abstract

Many logistics operators are considering deploying autonomous delivery fleets, such as Mobile Parcel Lockers (MPLs) and Autonomous Delivery Robots (ADRs), to address increasing delivery demands and sustainability challenges in cities. These solutions offer economical and efficient responses to urban delivery bottlenecks, such as time spent on parking and manual delivery modes. Regulatory issues, including requirements for using autonomous vehicles on roads and ensuring pedestrian safety with sidewalk ADRs, are major concerns due to varying laws across regions. Additionally, the feasibility, effectiveness, and capacity limits of deploying autonomous delivery modes in traffic-dense cities need validation before large-scale implementation. Considering the envisioned scale of autonomous delivery fleet deployment, conducting feasibility and capacity assessments becomes imperative.

This study developed a macroscopic assessment model incorporating cost, delivery efficiency, and pollution emissions to predict the performance of a collaborative delivery pattern integrating human-van delivery, MPL delivery, and ADR delivery in different London boroughs. A congestion-game model was proposed to explore customer selection behaviour between MPL and ADR services, providing parameters for future deployments. Mixed-integer programming models for MPL and ADR route planning were established, using reinforcement learning methods and graph theory to ensure solution quality and computational efficiency. Geographic datasets from Illinois, US, were used for the ADR route planning study.

Results indicate that MPL and ADR deployment in urban areas reduces operational costs and improves delivery efficiency compared to conventional manual delivery modes. Their lightweight and easy transportability can alleviate urban traffic congestion and parking space occupancy. Rational curb space planning and parking management strategies can facilitate the integration of autonomous delivery vehicles, promoting wider implementation. For logistics operators, differentiated pricing strategies and capacity deployment based on customer preferences, along with vehicle route planning according to customer time windows and activity trajectories, will reduce operational costs and enhance user convenience.

Acknowledgements

At the age of over twenty-seven, I have been devoted to pursuing knowledge for over two decades. Looking back on the path of my scholarly endeavours, though obscured by thorns and challenges, I have been fortunate to have benefactors supporting me. The years dedicated to learning in the UK are the period in which my knowledge is most abundant, and my gains are most significant. The insights gained along this educational journey are vast and profound, challenging to articulate in mere words.

“余年二十七有余，求于学道者亦二十载矣。回眸求学之路，虽隐荆棘以挑战，幸得贵人持之。于英期年，乃吾所学最为充盈，增益最多之时。学途感悟，万言难见。”

I am deeply fortunate to conduct research under the supervision of Dr. Panagiotis Angeloudis. I am grateful for his encouragement, meticulous instruction, and the trust he has placed in me. I distinctly remember his advice during moments of setbacks in my research, saying, ”Doing research is like practicing a dance; it requires daily repetition, and one will eventually grasp what is learned through consistent effort.”

“吾甚幸拜师于安杰卢迪斯博士门下。与师之缘，实乃奇妙。感恩其对吾之鼓励与细言深教，亦感恩其对吾之信任。吾固记其尝于吾研学遇挫时劝吾曰：研学如习舞，须日复一日而习之。而汝始学舞，因学日将得其所学。”

I also express my sincere gratitude to my co-supervisor, Dr. Jose Escribano-Macias. I am thankful for his esteemed academic insights and assistance in my research and paper writing, which have equipped me to analyse problems, clarify doubts, and think critically. His rigorous scholarly character serves as a model for my own learning endeavours.

“谨谢吾二导师埃斯克里瓦诺博士。感恩其尊吾学术之见解，予吾研学、论文创作之助，使吾具分析解惑、质疑思义之能。其治学严谨之品格，乃吾学习之楷模。”

Thanks to my colleagues at the Transport Systems and Logistics Laboratory (TSL) for providing research guidance and fruitful discussions. Their consistent support and encouragement have been a constant source of motivation for me. They praise my strengths and kindly pardon my weaknesses, always offering support and understanding. I am deeply touched by their concern for my well-being and their sincere kindness. I often reminisce about the shared

moments of discussion in research and life, which will remain unforgettable in the chronicles of my life.

“谨谢同门师兄师弟答疑解惑。其常支持，恒励吾。誉吾之长，厚吾之短。感其为吾排忧，对吾之深情厚意。余常忆共论交通之光阴，此将吾人生难忘之日月。”

I wish to express my gratitude to my mother, Mrs. Li Liao, and my father, Mr. Dewen Liu, who have played a crucial role in my upbringing. They have imparted profound teachings, consistently emphasising the importance of both understanding and action. Their strict guidance has been my steadfast support, providing a solid foundation and constantly encouraging me to move forward. I am deeply grateful for their kindness and encouragement, which have instilled in me the determination to face challenges head-on and navigate through hardships. These experiences have shaped my resilience and courage, surpassing that of my peers.

“谨谢吾母廖丽氏，吾父刘德稳氏，于吾成长之途，细言深教，恒持知行并重，严格待我，为吾坚实之靠山，常励吾前行。当下蒙他们厚意，感荷鼓励，使我有志独挑风雨，直面艰辛险阻。盖此等激荡，致我抗压之技艺、生存之胆略，远胜同侪。”

I express my heartfelt thanks to my wife, Mrs. Yu Lingnan, who has been a constant companion throughout the arduous journey of my research. She has provided unwavering support and understanding, going beyond words to convey the depth of her care. Moreover, our encounter in London was a serendipitous moment in my life, an enchanting experience that lingers in my heart, never to be forgotten.

“谨谢吾妻余聆楠氏在吾研究之艰辛途中，恒伴吾左右，以厚实之扶持与懂晓，余情非言。且与卿相会于伦敦，实乃人生奇遇，留连心香不忘。”

The years of seeking knowledge have truly been a period of sincere self-reflection, delving into the depths of my own heart, and diligently exploring the path of truth, goodness, and virtue. In the past, I regarded fame and fortune as the hard-earned rewards for my efforts. However, after experiencing the twists and turns of life, my heart still cherishes ideals and a desire for freedom, which, in essence, is the true reward for myself. Having received immeasurable kindness, I am determined to repay it in the future. Much like the phases of the moon, life may reach completeness, yet there might also be moments of regret. I remain embracing the present, gazing up at the vast sky, aspiring towards ideals, and pursuing the brilliance of academia.

顾盼求学之岁月，实属自我真诚，究求己心，悉心探求真、善、美之途。昔者吾将功名视为辛勤之报答，然经历坎坷曲折而犹心怀理想、保自由之志，方为自身真正之奖赏。受恩无量，他日当偿。月有阴晴圆缺，事可臻，亦或有所憾。吾将处于当下，仰望苍穹，瞩望理想，逐学术之华光。

Publications during PhD

Journal Articles:

- **Liu, Y.**, Ye, Q., Escribano-Macias, J., Feng, Y., Candela, E., & Angeloudis, P. (2023). Route planning for last-mile deliveries using mobile parcel lockers: A hybrid q-learning network approach. *Transportation Research Part E: Logistics and Transportation Review*, 177 , 103234.

Conference Papers:

- **Liu, Y.**, Escribano-Macias, J., Zhang, K., Ye, Q & Angeloudis, P., (2023). Customer Behaviour Analysis for Coordinated Delivery Service Using Mobile Parcel Lockers and Autonomous Robots: A Congestion-Game Model, in *'102nd Transportation Research Board Annual Meeting (TRBAM-2023)'*. Available at SSRN 4549554. **(Best Research Paper Award of Urban Freight Transportation Committee, Transportation Research Board, US)**.
- **Liu, Y.**, Zhang, K., Escribano-Macias, J. & Angeloudis, P., (2023). Quantifying the Capability of Autonomous Robot-Assisted Distribution in Last-Mile: Insight from Rural to Urban Configurations. in *'26th IEEE International Conference on Intelligent Transportation Systems ITSC 2023'*. Available at SSRN 4549465.
- **Liu, Y.**, Ye, Q., Feng, Y., Escribano-Macias, J & Angeloudis, P., (2022). Location-Routing Optimisation for Urban Logistics Using Mobile Parcel Locker Based on Hybrid Q-Learning Algorithm, in *'101st Transportation Research Board Annual Meeting (TRBAM-2022)'*. arXiv preprint arXiv:2110.15485.

Working Papers:

- **Liu, Y.**, Escribano-Macias, J., Zhang, K., Ye, Q & Angeloudis, P., (2024). Kernelised-Support-Vector-Machine- Based Congestion-Game Model for Coordinated Delivery Service Selection Using Mobile Parcel Lockers and Delivery Robots. *Transportation Research Part C: Emerging Technologies*. Status: Under Consideration. Available at: http://dx.doi.org/10.2139/ssrn.4383034.

- **Liu, Y.**, Zhang, K., Escribano-Macias, J. & Angeloudis, P., (2024). Quantifying the Capability of Autonomous Distribution in Last-Mile Market: A Cast Study of London. *Transportation Research Part A: Policy and Practice*. Status: Under Preparation.

Contents

Abstract **ii**

Acknowledgements **iii**

1 Introduction **1**

1.1 Last-Mile Deliveries 2

1.2 Autonomous Last-Mile Delivery 6

1.3 Overview of AV Technology Applications in Logistics 7

1.4 Business Model of Autonomous Last-mile Delivery 11

1.5 Regulations 12

1.6 Statement of the Problem 15

2 Research Objectives and Contributions **19**

2.1 Research Aims and Objectives 20

2.2 Contribution 22

2.3 Outline 24

3 Literature Review **28**

3.1 Literature Review Procedure 29

3.2 Macroscopic Evaluation for Autonomous Last-Mile Deliveries 29

3.2.1 Key Indicators of ALD Performance Evaluation 30

3.2.2 Techniques for ALD Performance Evaluation 32

3.3 Mesoscopic Decision for Autonomous Last-Mile Delivery Service Selection 34

3.3.1 Customer Decision-Making in Delivery Service Selection 34

3.3.2 Equilibrium Models of Last-Mile Delivery Service Selection 36

3.3.3 Applications of Machine Learning Techniques in Last-Mile Delivery Service 39

3.4 Operational Optimisation for Autonomous Last-Mile Deliveries 42

3.4.1 Route Planning Model for ALD 42

3.4.2 Comparison of Existing Route Planning Models 52

3.5 Concluding Remarks 56

4 Capability Evaluation for AV Fleets in Last-Mile Delivery 58

4.1 Scenario Definition for Delivery Options 59

4.2 Solution Method 63

4.2.1 Conventional Continuous Approximation Framework 64

4.2.2 Modelling Cooperative Deliveries Using Adjusted Continuous Approximation 66

4.2.3 Evaluation Matrix Modelling 73

4.2.4 Route Modelling for a Single Region 77

4.3 Case Study 78

4.4 Empirical Results 83

4.4.1 Performance Comparison of Different Demand Segmentation Strategies . 83

4.4.2 Performance Comparison of Different Boroughs 86

4.4.3 Performance Comparison of Different Delivery Options 89

4.4.4 Effect of service capacity 93

4.4.5 Implications and discussion 95

4.5 Concluding Remarks 99

5 Customer Decision Behaviour Modelling for Autonomous Delivery Service Selection 102

5.1 Problem Statement and Congestion-game Model 106

5.1.1 Assumptions and Notations 109

5.1.2 Agents and Action Space 110

5.1.3 Payoff Function of Delivery Services 111

5.1.4 Performance Evaluation Metrics 114

5.2 Existence and Uniqueness of Equilibrium Models and Inverse Game Process . . 116

5.2.1 Fully Rationality Equilibrium 117

5.2.2 Bounded Rationality Equilibrium 120

5.3 Solution Algorithms and Inverse Game Process 123

5.3.1 Solution Algorithm for Fully Rationality Model 124

5.3.2 Solution Algorithm for Bounded Rationality Model 124

5.3.3 Improved Algorithm with Global Information 125

5.3.4 Regression Prediction through KSVM 127

5.4 Computational Results and Discussions 130

5.4.1 Benchmark Result of the Game Process 131

5.4.2 Effect of Single Operating Parameter on System Performance 132

5.4.3 Effect on System Considering the Interaction of Homogeneous Operating Parameters . . . 135

5.4.4 Performance of Improved Algorithm with Global Information . . . 140

5.4.5 Price of Global Information in Different Rationality Models . . . 141

5.4.6 Performance of KSVM for Service Charge Prediction . . . 143

5.4.7 Managerial Implications and Discussions . . . 144

5.5 Concluding Remarks . . . 147

6 Route Planning Model for Last-Mile Deliveries using Mobile Parcel Lockers 150

6.1 Mobile Parcel Locker Problem (MPLP) . . . 152

6.2 Mathematical Model for MPLP . . . 155

6.2.1 Model Assumptions . . . 155

6.2.2 Model Structure . . . 157

6.2.3 Location Selection Model for Parking Spaces . . . 158

6.2.4 Route Planning Model for MPLs . . . 159

6.3 Solution Methods . . . 161

6.3.1 Task Generation Algorithm . . . 163

6.3.2 Structure Design of the State . . . 164

6.3.3 Hybrid Q-Learning Network Algorithm (HQM) . . . 165

6.4 Results and Discussion . . . 175

6.4.1 Generation of Problem Instances . . . 177

6.4.2 Performance Comparison between HQM and Exact Approach . . . 178

6.4.3 Performance Comparison between HQM and Heuristics Approach . . . 181

6.4.4 Effect of Critical Factors on Different Cost Units 186

6.5 Managerial Implications 191

6.6 Analysis and Discussion 194

6.7 Concluding Remarks 197

7 Route Planning Model for Robot-Assisted Deliveries in Last-Mile 200

7.1 ADR-Assisted Delivery Problem 202

7.2 Mathematical Model for AADP 203

7.2.1 Route Planning Model for AADP 204

7.3 Solution Methods 206

7.3.1 Road Impedance 208

7.3.2 Decision Rule for ADR Travelling 208

7.3.3 Simplification for ADR Service Set 209

7.3.4 Simplification for Decision Variable 210

7.3.5 Constraint Tightening 211

7.4 Results and Discussion 212

7.4.1 Performance Evaluation for Solution Techniques 213

7.4.2 Delivery Performance of AAD Scenario 213

7.4.3 Effect of Key Operating Parameters on Delivery Performance 215

7.5 Managerial Implications 218

7.6 Concluding Remarks 219

8 Conclusion **221**

8.1 Research Objectives Evaluation 221

8.2 Research Implications 225

8.2.1 Implications for Commerce 225

8.2.2 Implications for Academia 227

8.3 Future Work 228

Bibliography **229**

A Route Planning Model of Delivery Option **255**

A.1 Route Planning Model of HV Delivery Option 255

A.2 Route Planning Model of MPL Delivery Option 256

A.2.1 Location Model of MPL Stopovers 257

A.2.2 Route Planning Model for MPL Stopovers 258

A.3 Route Planning Model of ADR Delivery Option 259

B Equilibrium, POA, and POR of Rationality Models with Different Parameters **264**

C Correlation Analysis between MPLs and ADRs **267**

D POA & POR of Rationality Models with Different Service Charges **272**

E POA & POR of Rationality Models with Different Capacities **275**

F Implementation Result of KSVM Regression Prediction **278**

G Derivation for Auxiliary Variables **280**

G.1 Derivation for d_{ij} . . . 280

G.2 Derivation for t_{ij} . . . 280

G.3 Derivation for d_{in}^{k} . . . 281

G.4 Derivation for e_{ia}, l_{ia} . . . 281

G.5 Derivation for t_{ima} . . . 282

G.6 Derivation for q_{ia} . . . 282

H Decoding Algorithm for State **283**

I Performance Comparison between HQM and Gurobi **284**

I.1 Reward and Computation Time for HQM and Gurobi . . . 284

I.2 Cost Unit Comparisons between Gurobi and HQM . . . 286

J Code Listing of Genetic Algorithm (GA) **288**

K Performance Comparisons between HQM and GA **290**

K.1 Comparison of MPL Deployed and Travel Distance . . . 290

K.2 Parcel First-time Fulfilment of HQM and GA . . . 295

K.3 Reward Improvement of HQM and GA . . . 296

K.4 Reward and Computation Time Comparisons between HQM and GA . . . 296

L Proof for Proposition and Corollary **299**

L.1 Proof for Lemma 7.1 . . . 299

L.2 Proof for Proposition 7.1 . . . 299

L.3 Proof for Corollary 7.1 300

L.4 Proof for Proposition 7.2 and 7.3 301

L.5 Proof for Lemma 7.2 302

L.6 Proof for Lemma 7.3 302

L.7 Proof for Proposition 7.4 303

L.8 Proof for Proposition 7.5 303

M Implementation Results of the Service Set & Decision Variable Simplifications 304

List of Tables

1.1 Comparison table of autonomous and conventional delivery options. Based on Srinivas et al. (2022) 8

1.2 Properties of ADR products. Based on Li & Kunze (2023); Jennings & Figliozzi (2019) 10

1.3 Properties of MPL products. Based on Figliozzi & Jennings (2020); Jennings & Figliozzi (2019) 10

1.4 Regulations on delivery AVs in the US. Based on Jennings & Figliozzi (2019) . . 15

1.5 Definition of macroscopic, mesoscopic, and microscopic in the last-mile delivery settings. Based on Quora (2024) 16

3.1 Strings used for literature searching 30

3.2 Summary table of the studies focusing on delivery efficiency and cost indicators 31

3.3 Summary table of bounded rationality equilibrium (Adapted from Di & Liu (2016)) 37

3.4 Summary table of main supervised learning techniques for regression prediction 40

3.5 Summary table of MPLs-related studies. Based on Liu et al. (2023a) 46

3.6 Summary table of the existing AV-only and AV-assisted VRP studies. Based on Srinivas et al. (2022) 51

3.7 Summary table of the existing approaches for solving AV-only and AV-assisted VRPs. Based on Srinivas et al. (2022) 53

3.8 Classification of RL methods in solving VRP. Based on Liu et al. (2023a) 55

4.1 Different performance indicators for evaluation 74

4.2 Size, population, and estimated daily e-commerce parcels of different boroughs in London 79

4.3 Location of DC and UCC 79

4.4 Proportions of each service mode accounts for the total parcel volume of the borough 80

4.5 Parameter settings for the evaluation model 82

4.6 Emission comparison of different demand segmentation scenario 85

4.7 Total time consumption (hours) of delivery activities in different boroughs . . . 88

4.8 Gas emission of different boroughs 89

4.9 Total time consumption (hours) of delivery activities in different services 92

4.10 Gas emission of different delivery options 93

4.11 Fitness matrix of different autonomous delivery options against geographic regions and destination types 98

5.1 Notation and terminology of the model 108

5.2 Benchmark parameter setting of equilibrium models 131

6.1 Notation table for location selection model 158

6.2 Notation table for route planning model 160

6.3 Parameters setting in model validation 179

6.4 Result of Wilcoxon signed-rank test on best result obtained and computation time 181

6.5 Optimal gap between the local optimum of HQM and the truth optimum 186

7.1 Notation table for route planning model of AADP 207

7.2 Parameter Setting for Model 213

A.1 Notation table of route planning model of HV delivery option 255

A.2 Notation table of location model of MPL delivery option 257

A.3 Notation table of route planning model of MPL delivery option 258

A.4 Notation table for route planning model of ADR delivery option 261

C.1 POA correlations between MPLs and AGVs in service charge 268

C.2 POA correlations between MPLs and AGVs in capacity 269

C.3 POR correlations between MPLs and AGVs in service charge 270

C.4 POR correlations between MPLs and AGVs in capacity 271

F.1 MSE, EVRS, and R Square of different kernel functions of KSVM 279

I.1 Comparison of average computation time and reward under different problem instances 284

I.2 MPLs deployed, travel distance, and service delay of Gurobi and HQM 287

K.1 MPLs deployed and travel distance of HQM and GA 291

K.2 Parcel First-time Fulfilment of HQM and GA 295

K.3 Comparison of average computation time and reward under different problem instances 297

M.1 Implementation Results of the Solution Techniques Proposed in Section 7.3.3 and 7.3.4 305

List of Figures

1.1 Logistics network structure. Based on Macharis & Melo (2011); Escribano Macias (2021) 3

1.2 Effect of time window length. Based on Boyer et al. (2009) 4

1.3 Cost comparison between manual and unmanned delivery with and without signed information. Based on Punakivi & Saranen (2001) 4

1.4 Effect of geographical area and market penetration. Based on Boyer et al. (2009) 5

1.5 Illustration of MPLs and ADRs. Based on Liu et al. 9

1.6 Schematic operation of MPLs and ADRs in delivery market. Based on Liu et al. 11

2.1 Thesis content overview 27

3.1 Classification of VRPs in ALD. Based on Srinivas et al. (2022) 43

3.2 Different types of AV-assisted VRP. (a) two-echelon mode, (b) 'dispatch-wait-collect' mode, (c) 'dispatch-move-collect' mode, (d) 'dispatch-leave' mode, (e) platoon mode. Based on Srinivas et al. (2022) 49

4.1 Illustration of different delivery options for serving a single region 64

4.2 Illustration of conventional CA framework 65

4.3 Illustration of different route types of a region in CA framework 67

4.4 Illustration of the adjusted CA framework 71

4.5 Illustration of the case study areas 78

4.6 Cost comparison of different demand segmentation scenario 84

4.7 Time consumption comparison of different demand segmentation scenario 85

4.8 Cost comparison of different boroughs 87

4.9 Cost comparison of different delivery options 90

4.10 Effect of different MPL capacity settings on costs 94

4.11 Effect of different ADR service capacity settings on costs 95

4.12 Effect of different MPL capacity on delivery activities 96

4.13 Effect of different ADR service capacity on delivery activities 97

4.14 Effect of ADR service capacity on total emission 97

5.1 Implementation framework 124

5.2 Illustration of KSVM principle 129

5.3 Real-time queue length of MPL and ADR services 132

5.4 Equilibrium of varying customers' rationality under service charge properties . . 136

5.5 POA under different levels of customers' rationality 137

5.6 POR under different levels of customers' rationality 137

5.7 Equilibrium of varying customers' rationality under capacity properties 138

5.8 POA under different levels of customers' rationality 139

5.9 POR under different levels of customers' rationality 140

5.10 Performance improvement of different rationality models 142

5.11 POGI of different rationality models 143

6.1 Instance of MPLP 155

6.2 Illustration of the model relationship . . . 157

6.3 Solution framework of the model . . . 163

6.4 Illustration of a state . . . 165

6.5 Structure of the encoder . . . 167

6.6 Structure of the decoder . . . 171

6.7 Lookup table for state-action space . . . 173

6.8 Illustration of adjustment for variable x_2 . . . 175

6.9 Service delay of HQM and GA with different route adjustment polices applied . 182

6.10 Reward obtained under different problem instances . . . 184

6.11 Optimisation ability of HQM and GA . . . 185

6.12 Reward acquisition and convergence comparison . . . 185

6.13 Effects of critical factors against fleet size . . . 188

6.14 Effects of critical factors against travel distance . . . 189

6.15 Effects of critical factors against service delay . . . 191

7.1 Instance of conventional human-vehicle delivery and AADP . . . 204

7.2 Time consumption of different delivery activities in different regions . . . 214

7.3 Delivery route of vehicle (——) and ADR (——) in different regions . . . 214

7.4 Delivery time comparison between different distribution modes . . . 216

7.5 Effect of ADR capacity on completion time and activities . . . 216

7.6 Effect of parcel handling on completion time and activities . . . 217

7.7 Effect of battery capacity on completion time and activities . . . 218

B.1 Equilibrium, POA, and POR of rationality models with different parameters . . 266

D.1 POA under different service charges of ADRs and MPLs 274

D.2 POR under different service charges of ADRs and MPLs 274

E.1 POA under different capacities of ADRs and MPLs 277

E.2 POR under different capacities of ADRs and MPLs 277

K.1 Improvement rate of HQM and GA 297

Chapter 1

Introduction

In recent years, there has been a substantial surge in e-commerce and online transactional endeavours. Projections indicate that the online sales will exceed threefold the figures recorded in 2017 by the end of 2025 (Maas et al., 2023). In this context, megatrends, with urbanisation being a prominent example, are exerting a significant influence. The resulting growth in delivery traffic and the requirement of urban centres' utilisation are consequently giving rise to emerging demands in the last-mile logistics sector (Lemardelé et al., 2021).

The shift towards e-commerce and online shopping is reshaping consumers' purchasing preferences, leading to a heightened demand for a wider type of products (Boysen et al., 2021). This results in smaller but more frequent shipments and elevated unit expenses, primarily stemming from the absence of economies of scale in the most costly segment of the supply chain (Merkert et al., 2022; Joerss et al., 2016). The increasing expenditures and the need for adaptability are also impacted by increasing labour costs attributed to a driver shortage (Nguyen et al., 2023). Beyond the cost and delivery capacity, sustainability considerations introduce fresh requirements to the last-mile logistics. Issues such as shortage of parking spaces and inefficient transport utilisation impose supplementary responsibilities, including the mitigation of greenhouse gas emissions, traffic congestion, and noise pollution (Hu et al., 2019).

It is noteworthy that following the Covid-19 pandemic, the trend towards e-commerce has even significantly intensified, amplifying the demand for last-mile logistics (Kapser et al., 2021; Kapser & Abdelrahman, 2020). Concurrently, heightened attention has been directed towards

the concepts of Autonomous Last-mile Deliveries (ALD) where the compact and medium sized Autonomous Vehicles (AVs) are utilised to perform parcel or food handover, as they offer the potential to mitigate virus transmission through contactless delivery (Lim et al., 2024; Mishra et al., 2023; Srinivas et al., 2022; Pani et al., 2020). Moreover, they obviate the need for drivers and enable economically efficient transportation on a small scale (Yuen et al., 2022). In this manner, these technologies address challenges concerning heightened flexibility, escalating personnel expenses, and reduced bundling effects (Liu et al., 2023a; Bray & Cebon, 2022). From an urban sustainability perspective, the application of ALD serves to forestall second-row parking, diminish disruptions to traffic flow, and ultimately reduce carbon and noise emissions (Yuen et al., 2022; Koh & Yuen, 2023; Lemardelé et al., 2021).

Given the above concerns, ALDs exhibit significant promise in meeting the escalating demand in the years ahead since it allows various stakeholders such as customers, logistics operators, and urban planners benefit from this initiative. Nonetheless, given the high innovation and the operation of AVs in logistics within public domains, the broad deployment of ALD mandates a thorough assessment and planning within the large logistics network. This is crucial to ensure its effectiveness, safety, and alignment with the city's sustainability objectives.

1.1 Last-Mile Deliveries

Conventional last-mile delivery is defined as the final leg in a Business-to-Customer (B2C) delivery service wherein the consignment is delivered to the final recipient from local distribution centres through the human-dominant delivery fleets, such as vans, cargo bike, and motorcycle (Slabinac, 2015; Macharis & Melo, 2011).

Figure 1.1 illustrates a typical structure of the logistics network, which is a three-stage transportation process that consists of fist-mile, middle/second-mile, and last-mile delivery. The first-mile delivery encompasses the conveyance of goods from the manufacturing site to the warehouse (Macioszek, 2018). The goods are consolidated, packaged, and labelled in sizeable batches for more convenient transportation. This results in a large shipment size, leading to the realisation of economies of scale (Wang & Huang, 2021). The middle-mile delivery involves the

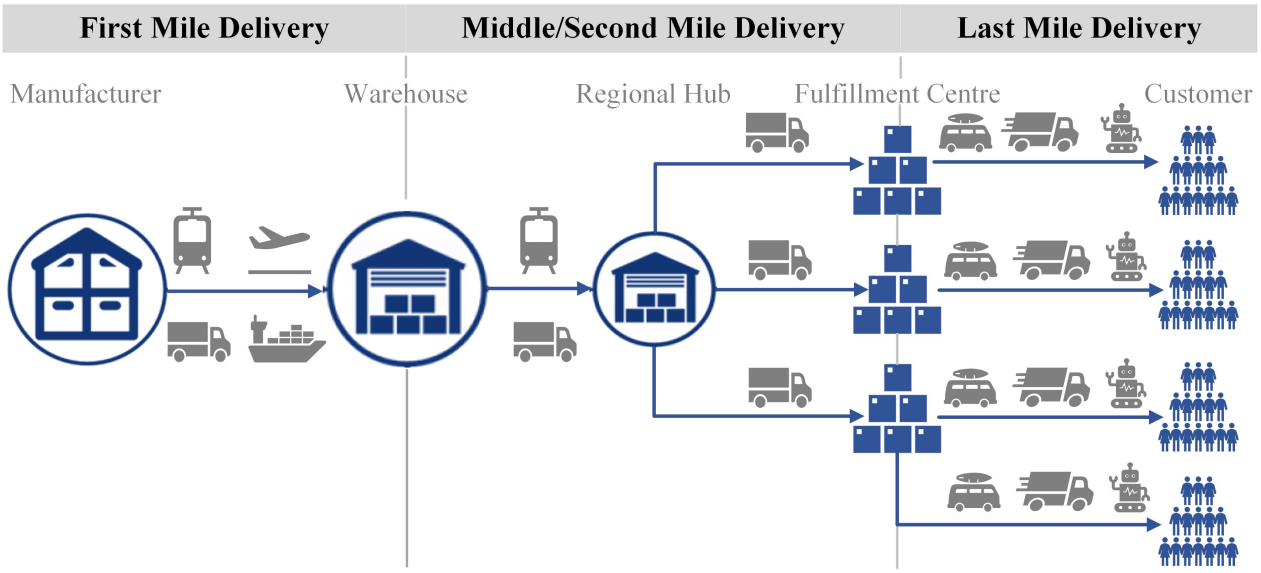


Figure 1.1: Logistics network structure. Based on Macharis & Melo (2011); Escribano Macias (2021)

division of goods into smaller segments, which are then transported to the local fulfilment center through regional hubs (Hinkamp & Ismael, 2023). In the last-mile delivery phase, individual orders are transported to specific destinations, resulting in smaller shipment sizes compared to the first and second mile delivery. It is a pivotal phase within the entire supply chain that entails the delivery route planning for the intricate and discrete distribution of demands, and thus accounts for a significant proportion of the total transportation cost (Wang et al., 2016). The cost of the last mile delivery activities could account for 13% to 75% of the total transportation expenditure (Aized & Srai, 2014).

There exist three pivotal factors originating from the demand and customer side that affects the cost and efficiency of the last-mile delivery operations: customer service levels, delivery type, and geographical area and market penetration (Macharis & Melo, 2011).

Customer service levels can significantly affect the efficiency of parcel fulfilment through the time-related variables, such as delivery time windows and delivery lead times. As illustrated in Figure 1.2, when the time window length allows for more flexibility, the distance a vehicle needs to cover to serve a customer is reduced (Boyer et al., 2009), thus the delivery efficiency improves. Consequently, incorporating narrower delivery time slots into the route schedule leads to an increase in costs (Gevaers et al., 2011) due to more delivery fleets that may be required to fulfil the time window constraints. It is estimated that a cost differential of 42%

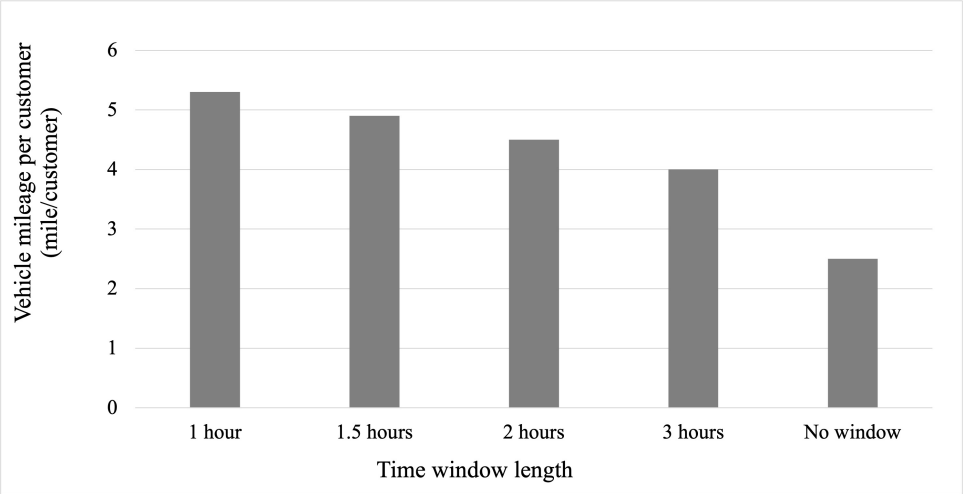


Figure 1.2: Effect of time window length. Based on Boyer et al. (2009)

between the deliveries to the reception or lockers compared to options that involve specified time windows (Kämäräinen et al., 2001). Nonetheless, in the absence of a predefined delivery window, the likelihood of unsuccessful deliveries due to customers not being present at home is bound to be considerable. Consequently, the parcel may need to undergo two or three delivery attempts before it is successfully handed over (Castillo et al., 2018).

The probability of missed delivery due to the absence of customers at home is significantly influenced by whether the consignee is obligated to provide a signature upon receipt, as well as the usage of delivery lockers or boxes. The missed delivery will yield extra costs due to the redelivery of packages. In cases where a signature is not required, the parcel may be placed in a secure locker on the customer's premises. Research has demonstrated that the cost differential ratio between deliveries with and without the recipient's signature can be as high as a factor of 2.5 (Punakivi & Saranen, 2001), as depicted in Figure 1.3. Delivery that necessitates a recipient's signature upon receipt is prone to experiencing a notable rate of missed deliveries, which is estimated to be approximately 30%, contingent on the type of the product (Fernie & Sparks, 2018).

The geographical expanse and market density play a crucial role in achieving a viable economic scale and operational efficiency in delivery. Market density is delineated by the average distance between different demand points and the potential for consolidating goods during the routing process (Macharis & Melo, 2011). Empirical studies demonstrate a clear inverse correlation between the distance a vehicle needs to cover per customer and the quantity

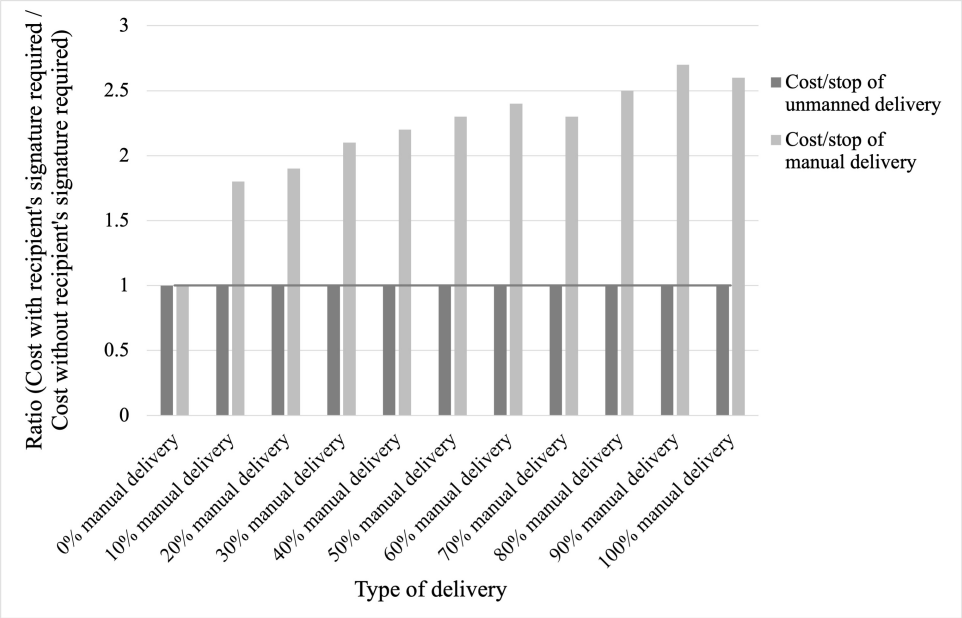


Figure 1.3: Cost comparison between manual and unmanned delivery with and without signed information. Based on Punakivi & Saranen (2001)

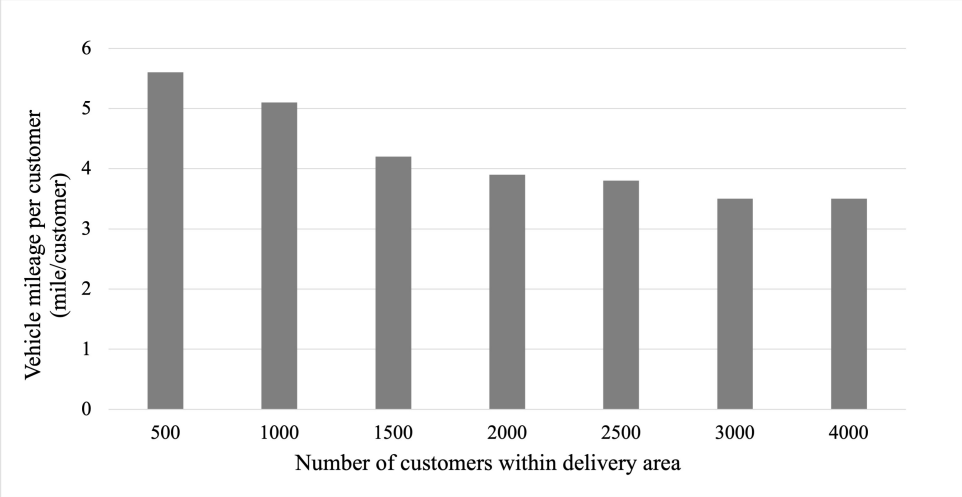


Figure 1.4: Effect of geographical area and market penetration. Based on Boyer et al. (2009)

of customers situated within the delivery zone (Boyer et al., 2009). As the number of customers within the delivery area decreases, there is a corresponding rise in associated costs. It is important to highlight that this correlation exhibits diminishing marginal miles per demand node (Figure 1.4).

From the supply perspectives, by incorporating the above factors, logistics operators have undertaken substantial investments in human-operated delivery fleets, skilled drivers, and urban fulfilment facilities to meet the rising demand in the past decades. While these incentives are effective for short-term demand fulfilment, they cause significant negative repercussions

on traffic efficiency, the environment, and urban land use. Their impact can be seen by the reduction in average vehicle speed within urban areas, which is limited to 25 km/hr (Kusumaningtyas & Lodewijks, 2008). Conventional urban freight transportation occupies around 40% of freight Vehicle Miles Travelled (VMT), contributing to roughly 6% of all greenhouse gas emissions from transportation (Suguna et al., 2021). Additionally, it is projected that between 3% and 5% of urban land is designated for logistics operations (von Schönfeld & Ferreira, 2021).

In response, several trials are carried out by logistics operators and retailers in cities worldwide to test their potential for urban freight distribution, notably in countries like the US, UK, and China (Legacy et al., 2019). These trials utilised AV-based delivery fleets comprising Autonomous Cargo Vans (ACVs), Autonomous Delivery Robots (ADRs), Mobile Parcel Lockers (MPLs), etc. Prominent logistics operators and retailers (e.g., Amazon, JD.com, DHL, and Cainiao) have conducted recent pilot initiatives to evaluate these autonomous freight transportation systems, specifically examining their potential for attended and unattended freight distribution. Research indicates that these compact and medium sized AVs are projected to dominate approximately 80% of Business-to-Consumer (B2C) deliveries by the end of 2035, given the rising expenses associated with traditional delivery options, including labor and fuel costs (Joerss et al., 2016). Furthermore, the market size of autonomous last-mile delivery in 2023 is $4.29 billion and will reach $29.13 billion by 2032 based on estimation (Precedence Research, 2023). Section 1.3 and 1.5 will introduce detailed description of the AV-related technologies and the corresponding pilot studies.

To sum up, the limitations imposed by conventional delivery options, characterised by their significant dependence on human operations and parking spaces, and lower delivery flexibility, contribute to hindering the expansion of the last-mile delivery market. In addition, a major concern within the urban logistics sector is the lack of human capital, with a limited pool of readily deployable skilled vehicle drivers, as well as the curbside space for vehicle parking (Ji-Hyland & Allen, 2022; Melo et al., 2019). As a result, the logistics operators are considering the use of compact and medium-sized AVs to assist delivery operations.

1.2 Autonomous Last-Mile Delivery

Autonomous Last-Mile Delivery (ALD) refers to the use of autonomous or self-driving compact and medium-sized vehicles and robots to complete the final leg of a product's journey from a distribution centre to its final destination, typically a customer's doorstep or a designated service point. This mode reduces the need for human drivers, allowing vehicles and robots to make independent decisions respond to surroundings and operate autonomously, relying on advanced technologies like artificial intelligence, computer vision, GPS, and other sensors to make safe and efficient delivery decisions. Table 1.1 represents the comparison analysis for autonomous and conventional delivery options.

ALD addresses the labour shortage in conventional delivery services, providing consumers with greater convenience and accessibility. Moreover, it offers a cost-effective business solution in the long run, as most delivery AVs are primarily electrically powered and do not consume fuel. Additionally, the characteristic of ALD primarily operating on sidewalks can reduce congestion on main roads. However, this necessitates their design to be compact and operate at lower speeds to meet sidewalk safety standards, resulting in a limited capacity. To enhance delivery capabilities, some compact robots (such as ADRs) often collaborate with traditional vans. For instance, vans may serve as mobile depots for storing packages, while ADRs are released from the van and deliver packages for customers (Reed et al., 2022b).

Although current delivery AVs are best suited for last-mile delivery on sidewalks or roads, they may ultimately face challenges in completing the final 50 feet of delivery, including climbing stairs and ringing doorbells (Cregger et al., 2020). A potential solution is to require customers to personally collect their packages from the delivery AV. However, this may be inconvenient for customers residing in high-rise apartments. Therefore, the adaptability of ALD needs to take into account the property characteristics of the final destination. For customers living in independent houses, ALD may provide greater potential for convenience. Another solution could be legged delivery robots. Their advantage lies in not requiring a recipient when delivering packages. However, legged delivery robot technology is more complex and can operate at much lower speeds. So far, legged sidewalk delivery robots have not been part of pilot research or

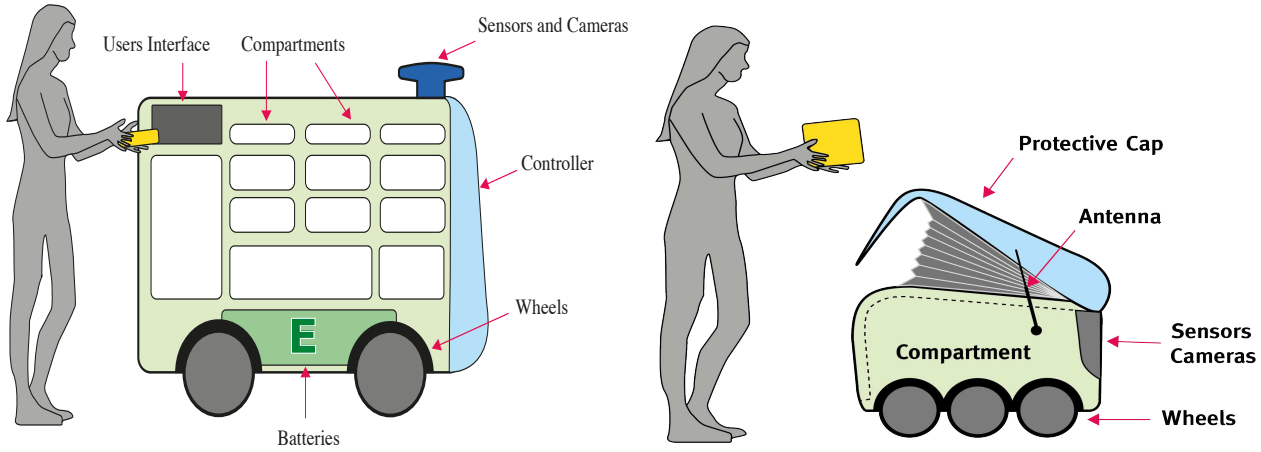
Table 1.1: Comparison table of autonomous and conventional delivery options. Based on Srinivas et al. (2022)

Class	**Autonomous**	**Conventional**
Driver Presence	No	Yes
Automation Level	High, based on sensory data	Relies on human operations
Payload Capacity	Low-Medium	High
Speed	Low	Traffic dependent
Noise	Low	High
Range	Low-Medium	High
Regulation	Actively being developed	Standards are well-established
Cost & Efficiency	Reduce costs in the long run, as it eliminates the need for human drivers and can optimise routes for efficiency	Relies on a labour force. Efficiency depends on factors like driver experience and traffic conditions
Flexibility	Mainly operate on sidewalk	Operate on vehicle lanes
Adaptability to Traffic	Adapt to traffic conditions using real-time data and advanced algorithms	Drivers rely on their experience and judgement to navigate traffic and road conditions
Vehicle Size & Type	Involve a range of compact and medium-sized vehicles, such as MPL & ADR	Involves a variety of vehicle types (e.g., bicycles, motorcycles, vans)
Public Perception	Public acceptance and trust in AV are still evolving and can be a factor in its adoption	Widely accepted by the public
Environmental Impact	Environmental friendly	Impact depends on vehicle type, fuel type, and driving habits of human drivers

deployment. The FedEx stair-climbing delivery robot serves as an example. Although AV technology holds promising potential and has recently emerged, relevant research is still scarce to date (Poeting et al., 2019a).

1.3 Overview of AV Technology Applications in Logistics

Delivery AVs refer to compact and medium sized ground vehicles devoid of an onboard pilot, capable of independent navigation within urban environments and remotely controlled. In contrast to traditional delivery vehicles that primarily operate on main roads, small and medium sized delivery AVs predominantly navigate along curbside and pedestrian pathways, exhibiting



(a) MPL (b) ADR

Figure 1.5: Illustration of MPLs and ADRs. Based on Liu et al.

the capacity to make responsive decisions based on their surroundings.

Recent strides in AV technologies and concurrent reductions in production expenses, coupled with their capability to cooperate with existing delivery personnel or vehicle fleets, render AVs especially well-suited for cargo delivery. Consequently, AVs are being contemplated as an enabler for enhancing last-mile distribution operations within commercial domains.

The two most representative and evaluated products of the existing AV pilot studies in the last-mile distribution sector are MPLs and ADRs, as illustrated in Figure 1.5a and 1.5b. Efforts to deploy AVs for last-mile delivery have been taking place since the foundation of Starship Technologies [1] in 2016 (Hoffmann & Prause, 2018; Figliozzi & Jennings, 2020). Several companies have announced or embarked on projects involving large-scale AV deployments including Amazon, Cainiao, DHL, Google, JD.com and UPS (Delfanti, 2021; Protasov et al., 2021; Schwerdfeger & Boysen, 2020). Moreover, these key players within the last-mile delivery domain and investment firms are significantly increasing their investments in research and market deployment of AV delivery. Taking Starship Technology as an illustrated example, since securing a Series A funding of $40 million led by Morpheus Ventures in 2019, the company further obtained funding of $17 million in 2021 from strategic backers TDK Ventures [2] and Goodyear Ventures (Lunden, 2022). In 2022, Starship Technology received approximately €57 million in financing from the European Investment Bank to promote the deployment

[1] Starship Technologies is a leading company in the ADR urban delivery industry
[2] TDK Ventures is an investment arm of the Japanese electronics giant

Table 1.2: Properties of ADR products. Based on Li & Kunze (2023); Jennings & Figliozzi (2019)

Companies	**Starship**	**Dispatch**	**Robby**	**Domino's**	**Twins Wheel**	**Thyssenkrupp**	**Marble**
Year	2016	2016	2019	2016	2019	2017	2017
Weight (kg)	23	N/A*	27	N/A	50	27	37
Payload (Kg)	10	45	N/A	9.5	50	35	N/A
Compartment	1	4	1	4**	1	1	1
Loading	manual	manual	manual	manual	manual	manual	manual
Offloading	manual	manual	manual	manual	manual	manual	manual
Velocity (km/hr)	6	6	6	19	3.6/7.2	56	6
Battery last (km)	6	77	32	19	15	16	N/A
Battery type	swappable	N/A	N/A	N/A	built-in	N/A	N/A

*: unknown, but it requires 2 people to lift the device; **: all four compartments are accessible at the same time.

of autonomous delivery. By the end of 2022, Starship Technology has raised $202 million. The company has now completed over 5 million commercial deliveries, covering a global travel distance exceeding 10 million miles, with its fleet making 10,000 deliveries per day (Starship Technologies, 2023). Given the substantial capital infusion from investment firms into the delivery AV companies, coupled with the escalating scale of AV deployments, the large-scale implementation of autonomous delivery is poised to materialise in the near future. The properties of different ADRs are presented in Table 1.2. The properties of different MPLs are presented in Table 1.3.

The properties summarised in Table 1.2 and 1.3 indicate that AV capability and performance are mainly influenced by the maximum battery capacity and the predefined number of compartments (Simoni et al., 2020). MPLs exhibit superior performance in terms of capacity, velocity, service coverage attributed to their larger vehicle structure which allows for a greater battery capacity, and may, in certain instances, share curbside space and roads with other motorised vehicles (Kötschau et al., 2023). Moreover, it was observed that they exerted similar effects on energy consumption per served customer and overall efficiency when compared to e-vans, albeit with lower initial investment costs (Li & Kunze, 2023). However, it should be

Table 1.3: Properties of MPL products. Based on Figliozzi & Jennings (2020); Jennings & Figliozzi (2019)

Companies	**Nuro**	**uDelv**	**AutoX**
Weight (kg)	680	1890	1769
Payload (Kg)	114	318	N/A
Capacity (parcels)	40*/12**	32	11 to 15
Loading	manual	manual	manual
Offloading	manual	manual	manual
Velocity (km/hr)	6	6	6
Arrox. Size (L × W × H in feet)	8×3.6×6	15×6×6	16×6×5
Range (km)	16	96	896
Battery type	built-in	built-in	built-in

*: the capacity for standard sized parcel; **: the capacity for large grocery bags

noted that this may fluctuate based on the density of demand fulfilment and the fleet size required. Conversely, ADRs exclusively utilise sidewalks or bicycle lanes, operating at speeds akin to pedestrians and offering greater flexibility concerning the safety and operation issue (Gehrke et al., 2023). Specifically, current advancements in ADR products primarily emphasise attributes such as a capacity spanning approximately 20 km, a velocity of under 10 km/hr, and a single compartment design. The constrained speed, as underscored by Baum et al. (2019), is regarded as a favourable safety feature due to its reduction in braking distance and satisfy the traffic regulation in most regions.

Figure 1.6 presents a classic schematic operation of MPLs and ADRs in last-mile deliveries. For MPL service, the locker can relocate its position based on customer schedules, thus exploiting mobility in the delivery network (Chen et al., 2024). However, the convenience and accessibility of MPLs are restricted because customers must travel to the MPL for parcel collection. To compensate for this inconvenience, MPL service is provided without extra charges if the parcel storage duration does not exceed the predefined threshold. As an alternative to MPL deliveries, ADR delivery options offer enhanced customer convenience by providing door-to-door service using the footpath and ramp (Boysen et al., 2018; Ostermeier et al., 2023). Nevertheless, such convenience is at the expense by the extra delivery fee charging, restricted service range due to battery limitations, compartment capacity constraints, and the mandatory parcel loading requirement before delivery. In certain situations, customers may experience extended waiting time because the ADRs encounter obstacles on the crowded sidewalk that

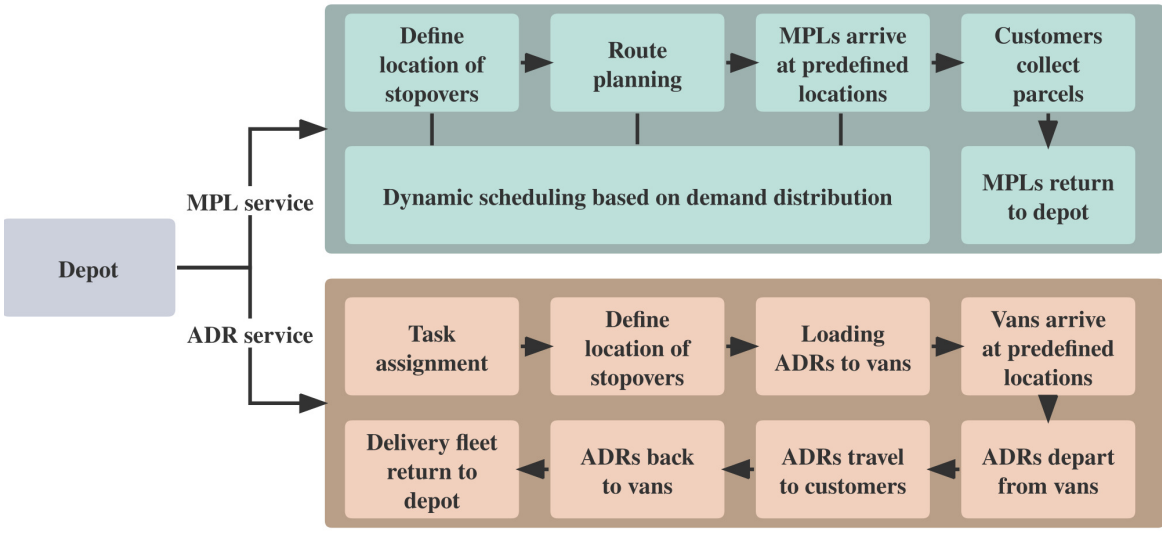


Figure 1.6: Schematic operation of MPLs and ADRs in delivery market. Based on Liu et al.

require extra time to avoid collisions (Pani et al., 2020).

1.4 Business Model of Autonomous Last-mile Delivery

The current business model in the ALD market operates in two primary ways: the MPL and ADR companies either run their own delivery services or partner with retailers such as *Tesco* and the *Co-op* (Woollacott, 2022). For customers, MPL and ADR companies have introduced the Robot-as-a-Service (RaaS) business model. In this model, customers can either pay per transaction performed by the robot or opt for a pay-as-you-go subscription based on time leased or a set of tasks performed (Biba, 2022).

Competition in the ALD market is currently limited, as most companies operate individually within a city or on a college campus. However, in some areas, multiple companies are present. As the number of autonomous delivery fleets from different companies increases and regulations are established, competition is expected to intensify. MPLs/ADRs companies will face competition from other MPLs/ADRs startups, drone delivery services, and self-driving cars, ultimately integrating autonomous delivery fleets into the community.

MPLs/ADRs companies typically earn commissions and tips for their services. Unlike some delivery service providers such as *UberEats*, which charge small delivery fees but also

take around 25% of the meal price from sellers, MPLs/ADRs services significantly reduce these commissions with lower additional costs and do not require tips (Hossain, 2022).

1.5 Regulations

The worldwide autonomous last mile delivery market had a valuation of USD 11.90 billion in 2021 (Straits Research, 2023). Projections indicate that it is poised to attain a value of USD 84.90 billion by 2030, with an anticipated compound annual growth rate of 24.4% over the forecast period (2022-2030). Among the regions, North America stands as the pre-eminent market for autonomous last-mile delivery, whereas Europe demonstrates the most rapid rate of expansion (Straits Research, 2023). The extensive application of the compact and medium sized AVs in the last-mile delivery sector necessitates the establishment of resilient regulatory frameworks that guarantee the safety and security of their operations.

The requirement of establishing regulatory framework for AVs in last-mile delivery focusing on resolving the following main concerns: i) the safety of other traffic participants, such as pedestrians, cyclists; ii) the safety of the AVs themselves and their payload; iii) liability issues in case of an accident; iv) obtaining the authorisations from different departments to operate, such as traffic authorities, city councils; v) privacy and data protection due to the use of cameras and customers' information (LMAD, 2022).

Currently, within the European Union, there exists a lack of explicit, overarching regulations pertaining to AVs operating in last-mile distribution sector in public spaces. Consequently, the opportunities for testing, experimentation, and implementation exhibit significant disparities across different countries (Hoffmann & Prause, 2018). However, the operation of AVs in last-mile deliveries should follow the general safety rules of AVs carried by the European Commission, including surrounding detections with camera or sensors, automated braking, data recorders, and emergency stop signals (European Commission, 2022).

In comparison to the US, the principal hindrance to the widespread implementation of large-scale autonomous last-mile solutions in the European Union lies in the legislative framework that governs experimentation and utilisation. This is particularly pertinent given

that a substantial portion of Europe's urban areas are densely populated. Nevertheless, there has been notable progress in this regard, particularly with the approval of the Mobility Orientation Law by the French government in 2020 (Pasquesoone, 2022). This legislation facilitated the commencement of experiments involving autonomous shuttles and, through a dedicated decree, accelerated the initiation of trials for delivery AVs. Subsequently, in 2022, the same framework was ratified to permit the operation of AVs and the commercial deployment of robot and drone-based delivery solutions (LMAD, 2022). In November 2023, the UK government introduced a new AV Bill, presenting one of the most comprehensive legal frameworks globally for AVs, with a primary focus on safety (GOV.UK, 2023). The new safety framework aims to establish clear responsibilities for users, set safety standards for legal autonomous driving, and institute an in-use regulatory scheme to monitor the continuous safety of these vehicles.

Certain nations, including Estonia, Austria, and Poland, permit the testing of autonomous vehicles under the condition that a person is present and capable of assuming control if necessary (Schwerdfeger & Boysen, 2022; Iwan et al., 2016). Conversely, Lithuania has embraced a more lenient approach where the testing of AVs is both feasible and actively encouraged, a policy aimed at enticing technology firms to conduct testing and implement solutions within its borders (LMAD, 2022). While MPLs and ADRs may not currently fall under the same regulatory framework, it is anticipated that they will gradually be incorporated as legal frameworks evolve to encompass the testing and operation of AVs.

In the US, some key players in the last-mile delivery sector have obtained authorisation to implement commercial delivery vehicle fleets, sometimes numbering in the thousands (LMAD, 2022). This underscores that the U.S. is leading ahead of Europe in this domain, owing to the earlier initiation of experimental and commercial endeavours in recent years. This advantage has been facilitated by regulations that offer greater flexibility and urban areas characterised by lower population density compared to those in Europe. To the best of author's knowledge, only a few states and cities have regulations in place. Table 1.4 provides an overview of diverse regulations pertaining to AVs in the context of last-mile delivery across various states and cities.

San Francisco maintains the most stringent regulatory framework for AVs in last-mile

delivery. This encompasses the imposition of not only speed and weight restrictions but also necessitates a permit for each individual device. Moreover, there exists a cap of nine permits for autonomous delivery devices allocated for the city as a whole (Jennings & Figliozzi, 2019). Furthermore, the device must emit an auditory alert to inform pedestrians and cyclists of its surroundings. In contrast, in other most states and municipalities like Arizona, regulations are notably less stringent. In these regions, the requirements encompass having an electric vehicle that operates at speeds below 16 km/hr, is under active control or supervision, adheres to pedestrian regulations, and yields the right of way to pedestrians.

Up to this point, regulations regarding the large-scale deployment of AVs in the last-mile delivery market have varied in different countries and regions, and governments are continuously refining the regulatory issues to promote the deployment of delivery AVs that consider vehicle and pedestrian safety. It can be anticipated that devices characterised by smaller design footprints and equipped with more advanced autonomous navigation systems and safety-assisting technologies will be pivotal for delivery AV companies in capturing the market in the future, thereby meeting increasingly stringent safety requirements. Therefore, existing delivery AV players are concentrating on the development of delivery AVs such as MPLs and ADRs, which are of small dimensions and can collaboratively operate with existing delivery vehicles. This strategic focus aims to facilitate the widespread adoption of these AVs in the delivery ecosystem in the near future.

1.6 Statement of the Problem

Given the involvement of various stakeholders with diverse objectives, as well as the intricate operational constraints and environments within urban logistics systems, erroneous judgements in the decision-making process emerge as a significant concern. Such errors can lead to pronounced negative effects on the efficiency and costs of the system (Bell, 2022). Consequently, mathematical optimisation serves as a fundamental tool in aiding decision-makers during this process, providing them with a palette of solutions evaluated based on agreed-upon metrics.

A substantial body of literature is dedicated to last-mile distribution and operations

Table 1.4: Regulations on delivery AVs in the US. Based on Jennings & Figliozzi (2019)

State or city	Date	Requirements				Legislation
		Sidewalk	Insurance	Braking system	Lights	
Arizona	3/28/2018	Yes, ped only	No	No	No	HB 2422[c]
Florida	7/1/2017	Yes	Yes	No	No	SB 460[d]
Idaho	7/1/2017	N/A	No	No	No	HB 204[e]
Ohio	9/29/2017	Yes, ped only	Yes	Yes	Yes	4511.513[f]
Utah	3/19/2018	Yes, ped only	Yes	Yes	Yes	HB 217[g]
Virginia	2/24/2017	Yes, ped only	Yes	Yes	Yes	46.2-908.1[h]
Wisconsin	6/21/2017	Yes	No	Yes	Yes	SB 148[i]
Austin	7/20/2017	Yes	Yes	No	No	20170810-012[j]
San Francisco	12/22/2017[a]	Yes, ped & Cyclists	No[b]	No	No	244-17[k]
Washington D.C.	9/15/2016	Yes	No	No	No	15C.50-1551 to 50-1555[l]

ped = pedestrian; [a]: the regulation was revised on 3/29/2018; [b]: must pay for caused damages; [c]: published by State of Arizona; [d]: published by the Florida Senate; [e]: published by Idaho State Affairs Committee; [f]: published by Ohio General Assembly; [g]: published by Utah General Session; [h]: Virginia Code; [i]: Wisconsin Legislature; [j]: published by Austin City Council; [k]: published by City and County of San Francisco; [l]: published by Council of the District of Columbia

research. These studies can be categorised into three types based on the system boundaries of the research scopes: i) macroscopic evaluation, which focuses on assessing the overall performance of different delivery methods in terms of cost, efficiency, and environmental impact; ii) mesoscopic decision-making, which explores the impact of decisions or actions taken by various stakeholders (e.g., customers, operators, government agencies) on the delivery system; iii) microscopic operational optimisation, which involves optimising operational aspects such as facility location selection, and optimal fleet routing and scheduling while considering realistic operational constraints. Table 1.5 defines the boundary of macroscopic, mesoscopic, and microscopic mentioned above. Chapter 3 set up a literature review based on each of these dimensions, joined with MPLs and ADRs delivery scenarios.

Nevertheless, the majority of this literature concentrates on detailed operational optimisation pertaining to delivery scenarios (Wang et al., 2024; Alverhed et al., 2024; Caunhye et al., 2012; Dekker et al., 2012). This includes aspects like route planning (Cattaruzza et al., 2017; Konstantakopoulos et al., 2020), the optimisation of fleet configurations (Gajda et al., 2024; Ensafian et al., 2023; Pedraza-Martinez & Van Wassenhove, 2012; Malladi et al.,

Table 1.5: Definition of macroscopic, mesoscopic, and microscopic in the last-mile delivery settings. Based on Quora (2024)

Scope	**Definition**	**Research Objectives**
Macroscopic	Large-scale system, focusing on average properties and performance without considering individual components.	Delivery system, networks
Mesoscopic	Intermediate-scale system, considering interactions between groups of individuals, which are used to study phenomena that exhibit collective behaviour	Group of customers, operators etc.
Microscopic	The smallest-scale system, focusing on the individual behaviour while considering interactions between individual entities and provide a detailed understanding of the system's underlying mechanisms.	Vehicle fleets, facilities etc.

2022), and the selection of optimal facility locations (Rao et al., 2015; Awasthi et al., 2011). These optimisations typically pertain to conventional transportation methods involving delivery personnel, vans, and cargo bikes.

While certain modelling techniques can be applied to the AV-based delivery system, they fall short in accounting for the intricacies specific to AV operations within the last-mile delivery domain. These include factors like limited battery capacity, standardised compartment size, and constrained service range. For example, the ADR is normally capable of transporting a maximum of four parcels per delivery round due to the predetermined compartment size (Jennings & Figliozzi, 2019), which is significantly lower than the capacity of conventional van delivery. Therefore, AV would transport high-timeliness low-volume cargo, such as groceries and parcels. Consequently, it is imperative to take into account the aforementioned attributes of AV delivery when defining the problem.

Based on the analysis of a extensive literature review, the following research gaps have been identified across macroscopic, mesoscopic, and microscopic systematic dimensions. These gaps will be addressed in the corresponding chapters of this study.

- Gap 1. From the macroscopic perspective, the systemic impact of delivery AVs on the existing nature of last-mile delivery has yet to be thoroughly explored. Considering the current limited scope of AV-based delivery implementations, it is imperative to assess the

feasibility and capability of diverse operational configurations of AVs on cost, efficiency, and environment. This determination establishes a prerequisite for subsequent endeavours aimed at optimising operations on a larger scale.

- Gap 2. From the mesoscopic perspective, equally important is the comprehension of the public's receptivity towards future AV adoption in the last-mile delivery market, along with their decision-making behaviour for different AV-based delivery services. As outlined in Section 1.1, AVs are projected to occupy 80% of the future B2C delivery market. Therefore, understanding the pivotal factors for successful AV delivery market entry (e.g., safety, operational configurations) and customer selection preference will serve as a guidance for future deployment and operation optimisation for logistics operators.

- Gap 3. Based on the statement of Gap 2, it is commonly assumed that customers make selections for delivery services based on complete rationality in the existing studies, aiming to minimise their costs. However, scenarios involving customers' bounded rational decisions due to imperfect information observation have not been extensively explored. Furthermore, the impact of incorporating global service information into customer decision models on system efficiency has yet to be determined.

- Gap 4. From the operation point of view, existing research lacks a thorough discussion on leveraging AV-based delivery mobility features to effectively address evolving customer demand patterns. To exemplify, customers may have multiple non-overlapping stopovers in a day, which can serve as potential pick-up locations. It is crucial to formulate efficient operation mechanisms that enhance customer convenience for parcel receipt at their preferred locations and time slots. Comprehensive research is also scarce on designing dynamic vehicle route adjustment strategies to mitigate service delays due to traffic congestion and time window conflicts, bearing substantial significance in ensuring delivery punctuality.

- Gap 5. Existing optimisation methods primarily focus on mathematical programming and heuristic algorithms, which often exhibit computational inefficiency or susceptibility to local optima in large-scale problem instances. There is a potential for the design of a

novel RL-based algorithm that strikes a balance between computational time and high-quality solutions.

Chapter 2

Research Objectives and Contributions

This chapter outlines the research aims and objectives, addressing the research gaps identified in Section 1.6. These objectives encompass the macroscopic evaluation of delivery method capabilities, the mesoscopic exploration of customer service selection behaviour, and the microscopic optimisation of delivery routes (will be elaborated in Section 2.1). Achieving these objectives will benefit various stakeholders within the sector: i) for government bodies, the systematic research framework, spanning from macroscopic delivery capability assessment to microscopic delivery vehicle operation optimisation, will enable the formulation of scientific curb space planning and the last-mile delivery market management guidelines, supporting the large-scale adoption of the proposed autonomous delivery services; ii) for logistics operators, by understanding customer decision patterns, logistics operators can develop more effective pricing strategies and configure resources more efficiently, thereby enhancing service capability and reducing operational costs; iii) for customers, a more rational and efficient delivery network increases customer convenience and accessibility when opts to these autonomous delivery options. Section 2.2 demonstrates the detailed research contributions, followed by the thesis framework presented in Section 2.3.

2.1 Research Aims and Objectives

This study aims to improve the delivery performance of urban freight transportation by using compact and medium-sized AV fleets, with specific considerations of the operational characteristics of different types of delivery AVs, the evolving customer demand patterns, the characteristics of customer destinations, and the suitability of these technologies for both urban and rural settings. Concretely, this study takes a structured approach, encompassing a comprehensive investigation spanning from macro-level feasibility assessment, and mesoscopic decisions of customers to micro-level operational optimisations. The following research objectives are defined to bridge the gaps mentioned in Section 1.6 point by point, and to attain the proposed research aim:

- Obj 1. Evaluate the suitability and effectiveness of implementing AV-based delivery solutions situated from urban to sub-urban of the London city, taking into account the cooperation among the proposed delivery modes (e.g., MPLs, ADRs, and conventional human-van delivery) and the types of customer destinations (e.g, apartments, houses). An improved continuous approximation approach is developed to estimate delivery route length and time consumption of delivery fleets, facilitating the quantification of associated operational costs, delivery efficiency, and gas emissions. Additionally, this study investigates the impact of each delivery option on pre-designed performance indicators when it occupies different market shares. The fulfilment of this objective resolves Gap 1 and will be introduced in Chapter 4.

- Obj 2. Develop a non-cooperative gaming framework to delineate the competitive dynamics between different autonomous delivery options (MPLs vs. ADRs). The operation process of delivery options is formulated as a queue system such that investigates the accumulative effect of customer decisions on system efficiency, wherein different settings of key operating parameters are applied for sensitivity analysis. The fulfilment of this objective resolves Gap 2 and will be introduced in Chapter 5.

- Obj 3. Apply a customised logit equilibrium model to capture different levels of customers' decision rationality for service selection. To improve the overall utility of the delivery

system, this study further develops an improved algorithm that incorporates global service information into customers' decision patterns. Finally, this study designs a Machine-learning-based (ML) regression prediction model to determine the optimal service pricing of the delivery options such that obtains the expected market share. The fulfilment of this objective resolves Gap 3 and will be introduced in Chapter 5.

- Obj 4. Develop a location-route planning model to achieve the optimal configuration of MPL deliveries, wherein customer stopovers and the corresponding time windows within the planning horizontal are considered in the model. This study further introduces dynamic prediction-based route adjustment strategies to reduce the service delay caused by time window conflicts. The detailed fulfilment will be introduced in Chapter 6. As an alternative to the MPL delivery option, we further establish a route planning model for ADR service to optimise the total delivery time consumption of ADR-assisted deliveries, wherein battery consumption constraints and auxiliary activities during the operation are incorporated into the model. The proposed route planning model will be validated through the real-world geographic dataset from rural to urban regions in Illinois, US, as introduced in Chapter 7. The accomplishment of this objective resolves Gap 4.

- Obj 5. Design an optimisation framework that utilises RL-based algorithms, game-theoretical algorithms, heuristics algorithms, and mathematical programming techniques to solve different sub-problems in accomplishing Obj 1 to 4 above. These solution techniques are specifically tailored to enhance the generalisation ability of the algorithms and solution quality in large problem instances. The proposed problem instances are applied in scale and scope, which allows us to evaluate the performance of the solution techniques effectively. The fulfilment of this objective resolves Gap 5.

Until now, there has been a notable absence of a comprehensive framework encompassing all the aforementioned facets of AV delivery operations. Chapter 3 provides a thorough review of pertinent literature in this domain.

2.2 Contribution

As outlined in Section 1.3 and 1.5, logistics operators have progressively expanded their deployment of AVs in the last-mile delivery market to enhance delivery efficiency and customer convenience. With technological advancements, the service range and delivery capabilities of small to medium-sized AVs are continuously improving.

Therefore, to maximise the contribution of this paper in the rapidly changing regulatory landscape and to provide a crucial reference for future large-scale deployments, the paper extensively investigates AV-based last-mile delivery from three perspectives: 1) From an urban planning standpoint, it evaluates the feasibility and capability of different AV delivery services in various geographical regions and compares them with the conventional human-led delivery approach. 2) From the perspective of end customers, it examines the influence of customer decision behaviour regarding different AV-based delivery services on system efficiency. 3) Operationally, it conducts delivery route planning considering the distinct operational characteristics of MPL and ADR such that it reduces delivery time and costs.

Through the accomplishment of the objectives outlined in Section 2.1, this research yields benefits for the market, urban planners, logistics service providers, and academia, which are summarised as follows. Each contribution corresponds to the accomplishment of the individual objective outlined in Section 2.1, respectively.

- Con 1. Capability quantification of autonomous delivery services based on costs, delivery efficiency, and emissions by decomposing the entire delivery process supports logistics operators to identify bottlenecks in delivery activities and implement operation improvements. The developed evaluation matrix can serve as an evaluation benchmark for future research endeavours. The improved continuous approximation framework introduced can be generalised to various delivery scenario assessments with necessary adjustments, given its flexibility for constructing evaluation functions. Additionally, the identified bottleneck of parking space dependence and parcel handover time disparities of conventional delivery service would facilitate the design of the curbside space to promote the successful popularisation of the proposed autonomous delivery modes.

- Con 2. Comprehensive insight into customer decision behaviour against different AV-based delivery service selections. The study identifies the principles guiding customer service selection based on the trade-off between service cost and convenience while exploring the system's equilibrium dynamics under various permutations of price and service efficiency. This can aid logistics operators in developing rational pricing strategies to achieve the desired market share. Additionally, the proposed non-cooperative game model considers the systemic impact of diverse configurations, encompassing service price, capacity, and efficiency, which empowers professionals in the AV-based delivery sector to deploy capacities based on the predictable reaction of customers towards different operation configurations, thereby increasing service levels.

- Con 3. The proposed customer decision algorithm has been demonstrated to support customers utilising global service information to reduce the effects of bounded rationality caused by incomplete information, thus guaranteeing the overall utility of the system when customers make self-interested decisions. This resolves the zero-sum game dilemma between customers and operators in most real-world operational scenarios. Furthermore, the developed realistic ML-based regression model shows effectiveness in predicting market share against emerging AV-based delivery services. This contribution ensures the sustainability of profitability for autonomous delivery services, as the prediction model can serve as a decision tool, enabling operators to adjust pricing and capacity deployment in response to market dynamics.

- Con 4. Development of a route planning model and adjustment strategy for AV fleets focus on minimising service delays from stochastic events and ensuring minimal delivery time consumption. The model integrates customer and parking time windows and customer trajectories, enabling practitioners to manage and deploy AVs in exploratory environments. By analysing the impact of key operational parameters (e.g., capacity, service radius) on cost and efficiency under various configurations, operators can adjust their setups to meet changing customer demands.

- Con 5. The proposed evaluation and optimisation framework has the generalisation

capability to address similar problems. The successful implementation of optimisation techniques (e.g., RL) detailed in Objective 5 shows their effectiveness in achieving a sound balance between solution quality and computational efficiency in large-scale instances, which overcomes the exploration and exploitation dilemma faced by traditional exact solutions and heuristic algorithms. This provides the potential to address models with more complex and multi-dimensional constraints in large urban networks.

2.3 Outline

Chapter 1 and 2 have outlined the motivation behind embarking on the proposed study and has furnished the essential contextual information required to grasp the significance, aims, and anticipated contributions. Figure 2.1 presents a structure of the topics researched in this thesis. The subsequent sections of this thesis are structured as follows:

Chapter 3 provides a survey pertaining to last-mile delivery within the context of AV fleet implementation in the distribution sector. This chapter further reviews the literature encompassing the following aspects: 1) the macroscopic assessment techniques for evaluating last-mile delivery operations, typically employed to establish the appropriate model and performance indicators for system evaluation; 2) the mesoscopic decision-making model for customer selection of last-mile delivery services, encompassing customer decision-making behaviour and equilibrium models for shaping service competition; 3) operational optimisation research centred on route planning for AV delivery fleets. Within this section, we encapsulate the primary constraints factored in problem formulation, development of solution techniques, and identifies the limitations within existing studies.

Chapter 4 serves as the foundation for the research conducted in Chapters 5 through 7. It is designed to validate the economic feasibility and operational effectiveness of the MPL and ADR delivery modes for future large-scale deployments from the perspectives of cost, delivery efficiency, and gas emissions. The implementation results serve as references for setting experimental parameters in the subsequent chapters. In Chapter 4, an evaluation model is developed to gauge the delivery capability of two distinct categories of AV fleets: MPLs and

ADRs, across various geographical areas in Greater London Region. A conventional mode of human-vehicle delivery is utilised as a benchmark for comparative analysis. Additionally, the chapter assesses the impact of diverse configurations, such as different demand segmentation strategies for various delivery services, on the overall performance of the system.

Having validated the effectiveness and capability of MPL and ADR delivery modes in Chapter 4, understanding consumer selection behaviour towards different autonomous delivery modes is crucial before determining specific service configuration strategies in the market. Therefore, Chapter 5 introduces a non-cooperative congestion-game model to formulate the service competition between MPL and ADR delivery services. This includes the establishment of Wardrop equilibrium and Logit equilibrium models to characterise customer service selection behaviour under fully and bounded rationality. Furthermore, an improved algorithm incorporating global service information into the customer decision pattern is proposed to enhance system performance. Lastly, this chapter presents an ML-based model to determine optimal service pricing, thereby achieving the desired market share.

Chapter 6 integrates the feasibility analysis of autonomous delivery modes from Chapter 4 and the customer service selection behaviour model derived in Chapter 5. It initiates the operational configuration planning for the MPL delivery option. Specifically, Chapter 6 proposed an MPL route planning model tailored to accommodate the dynamic demand patterns of customers. This model primarily considers a series of stopovers generated by customers within the planning horizon, along with constraints related to parking availability and allowable parking times. A RL-based algorithm is devised to solve the model, enabling the attainment of high-quality solutions in a shorter computation time for large-scale problem instances. Additionally, the proposed dynamic route adjustment strategy facilitates the modification of the initial fleet delivery routes, thereby reducing service delays caused by stochastic events.

As an alternative to the MPL delivery mode, it is equally crucial to undertake operational configuration planning for the ADR delivery mode. In contrast to the unmanned operation characteristics of MPL, the operational configuration planning for ADR focuses on how it can assist delivery personnel in parcel delivery, considering aspects such as manual intervention required for parcel loading. Specifically, Chapter 7 establishes a robot-assisted delivery model.

This model efficiently minimises parking space occupation and reduces the need for delivery personnel to make multiple round trips between the vehicle and the customers. It allows delivery personnel to focus more on productive activities, such as loading packages onto the robot. To reduce the computation complexity, this chapter further develops simplification techniques for decision variables and the determined service set. The proposed method and model are evaluated using geographical datasets of 9 counties in Illinois, which span from rural to urban regions. The effect of robot compartment, parcel handling, and battery capacity on delivery efficiency is also explored.

Chapter 8 encapsulates the findings, furnishes a compilation of conclusions and their implications for both the academic and industrial domains, as well as suggests avenues for future research to advance upon this endeavour.

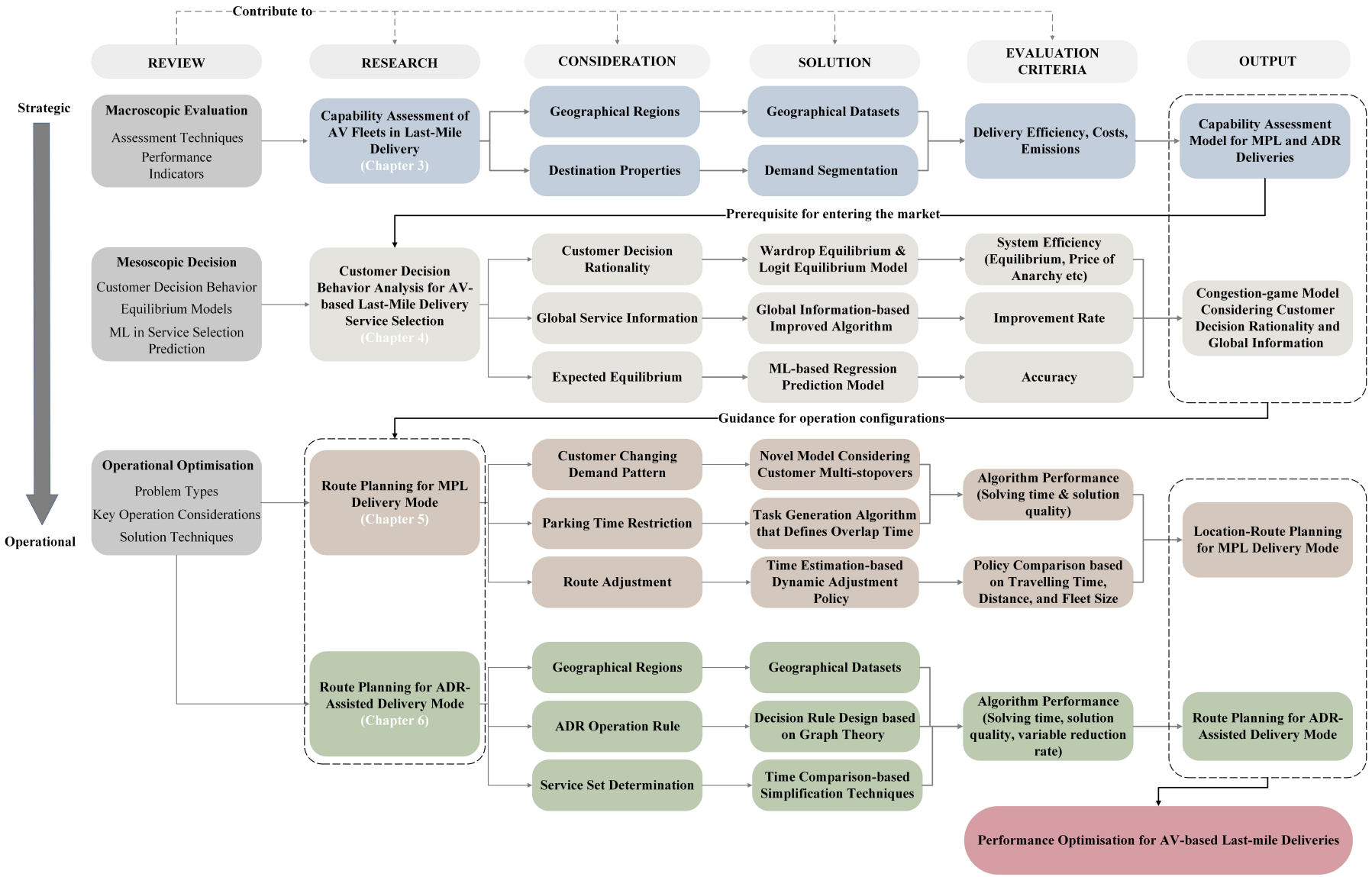


Figure 2.1: Thesis content overview

Chapter 3

Literature Review

Autonomous last-mile logistics seeks to optimise delivery efficiency and cost structure of the operation, holding the promise of revolutionising the manner in which goods are conveyed and distributed within urban settings. While the fundamental processes of autonomous last-mile logistics coincide with those of conventional delivery options, significant differences exist relevant to the operational objectives and constraints, and necessitate specific consideration when formulating the problems.

This chapter reviews the state-of-the-art of autonomous last-mile logistics operations. In particular, this review aims to:

R1. Review the off-shelve performance indicators and potential techniques that could be used for evaluating the feasibility and capability of autonomous deliveries in different regions.

R2. Provide a condensed overview of the key attributes influencing customer choice in various delivery services. Evaluate the applicability of diverse equilibrium models in formulating the competitive dynamics among distinct delivery services.

R3. Examine the body of research focused on optimising operations through the utilisation of MPL and ADR for parcel delivery, with a focus on route planning for the delivery fleets. Furthermore, review the prevailing solution technique for addressing autonomous delivery challenges.

This chapter is structured as follows: Section 3.1 introduces the strategy and keywords used for searching relevant literature reviews. Section 3.2 focuses on macroscopic assessment

techniques that consists of performance indicators determinations for evaluating autonomous deliveries (corresponding to R1). Section 3.3 demonstrates customer decision patterns and equilibrium models in the context of delivery service selections (corresponding to R2). Section 3.4 delves into route planning models for AVs (corresponding to R3). The section further summarises the solution approaches to solve the problems.

3.1 Literature Review Procedure

We utilised the *Google Scholar* database and implemented a string search strategy to comprehensively cover relevant literature under each research dimension, given *Google Scholar*'s extensive and up-to-date collection of publications. Additionally, *Google Scholar* database are not case-sensitive and allow the use of Boolean operators, which helps mitigate potential omissions caused by string formatting.

To better delineate the research boundaries in the field of last-mile delivery, the initial stage of our review concentrates on those literature providing systematic literature reviews, which offer extensive and in-depth research insights. The literature on the last-mile delivery can be classified according to the system boundaries of the research into macroscopic evaluations, mesoscopic decision-making, and microscopic operational optimisation. The second stage of the literature review involves conducting keyword searches for each research dimension. This stage integrates the proposed MPL and ADR delivery scenarios, ensuring that the literature review is more aligned with the theme of this study. Table 3.1 demonstrates the strings used for searching the relevant literature.

3.2 Macroscopic Evaluation for Autonomous Last-Mile Deliveries

This section reviews the optimisation indicators addressed in existing ALD research, providing benchmark indicators for the proposed assessment of feasibility and capability in ALD. Furthermore, this section summarises the mainstream methods available for evaluating delivery

Table 3.1: Strings used for literature searching

Dimensions	Strings
1st-stage review	
Systematic	("logistics" OR "last-mile" OR "delivery" AND "literature" OR "review")
2nd-stage review	
Macroscopic	("logistics" OR "last-mile" OR "delivery" AND "evaluat*" OR "assess*" OR "performance")
Mesoscopic	("logistics" OR "last-mile" OR "delivery" AND "service selection" OR "decision making" OR "machine learning" OR "competition" OR "game" OR "equilibrium")
Microscopic	("logistics" OR "last-mile" OR "delivery" AND "mobile" OR "autonomous" OR "moving" AND "locker*" OR "robot*") ("vehicle routing problem*" OR "fleet schedul*" OR "vehicle dispatch*") ("logistics" OR "last-mile" OR "delivery" OR "vehicle routing problem*" AND "reinforcement learning")

Notes: The asterisk (*) represents wildcard searches within the same word category, thus preventing exclusion due to mismatched word suffixes. For example, "evaluate" and "evaluation" should be treated as the same meaning for the literature search. Considering the *Google Scholar* could not recognise the asterisk directly, we enumerate all possible suffixes, such as "s", "tion", "te", "ment", and "ing", for the corresponding word.

services.

3.2.1 Key Indicators of ALD Performance Evaluation

The evaluation indicators associated with ALD can be categorised into four classes: operational safety, environmental impact, traffic impact, delivery efficiency, and operational costs (Srinivas et al., 2022). Regarding operational safety, relevant standards can be referenced from the regulatory requirements outlined in Section 1.5 concerning factors such as payload capacity, operating velocity, and robot design (Jennings & Figliozzi, 2019). These standards are legislative and mandatory, thus most studies assume that logistics operators must comply with them when deploying ALD in the delivery market. Some more detailed safety standards, such as the safe distance between AVs and pedestrians and accident rate, have not been uniformly established. Consequently, existing research have difficulty conducting detailed operational safety assessments for scenarios involving multiple groups of AVs in ALD from a macroscopic perspective, and can only impose constraints on individual AVs.

Environmental impact indicators primarily encompass greenhouse gas emissions and noise

pollution (Figliozzi & Jennings, 2020). Given that the majority of delivery AVs are electrically powered, they do not generate gas emissions. However, if delivery AVs such as ADRs are used in conjunction with conventional gasoline vans, the gasoline vans will produce greenhouse gases like carbon dioxide, impacting the environment. Furthermore, compared to conventional delivery vehicles that rely on internal combustion engines, delivery AVs generate lower levels of noise pollution.

In terms of traffic impact indicators, to the best of the author's knowledge, Fang et al. (2013) and Wu et al. (2017) are the only two studies that focus on the impact of ALD on traffic flow. Specifically, for those large sized delivery AVs that need to share the vehicle lanes with passenger vehicles, such as autonomous trucks, their impact on traffic flow is primarily reflected in lane occupation and reservation. This is aimed at minimising interference with human-droved vehicles. It requires decision-makers to engage in judicious lane allocation and design effective reservation mechanisms, such as operable time slots. As for compact or mid-sized AVs operating on sidewalks, their impact on the main road traffic flow is minimal.

It is noteworthy that most existing ALD-relevant studies have used delivery efficiency and operational costs as evaluation indicators (as shown in Table 3.2), which are also of primary concern to logistics operators (Simoni et al., 2020; Bakach et al., 2021). The performance of ALD in terms of delivery efficiency and cost is contingent upon the effective route planning of the delivery fleet (Lemardelé et al., 2021). Therefore, many studies focus on the development of novel AV-based delivery models and route planning algorithms with the objective of minimising delivery time and costs when formulating the problems.

3.2.2 Techniques for ALD Performance Evaluation

Based on the findings in Section 3.2.1, the evaluation of delivery efficiency and cost, which are critical factors influencing the potential success of ALD deployment, is of paramount importance to operators. As autonomous delivery concepts are at an early stage of development, there is a lack of extensive and precise real-world data and widespread deployment of large-scale autonomous delivery systems. This absence makes it challenging to conduct a comprehensive evaluation of the various models in terms of their representativeness (Ansari

Table 3.2: Summary table of the studies focusing on delivery efficiency and cost indicators

Indicators	**Sub-indicators**	**Citations**
Delivery Efficiency	Delivery Time	Ulmer & Streng (2019); Reed et al. (2022b); Chen et al. (2021a,b); Simoni et al. (2020); Liu et al. (2023a)
	Service Delay	Alfandari et al. (2022); Boysen et al. (2018); Gnegel et al. (2021); Liu et al. (2023a)
Delivery Costs	Operation Cost	Sonneberg et al. (2019); Lemardelé et al. (2021); Bakach et al. (2021); Jennings & Figliozzi (2019); Yu et al. (2020b, 2022a); Ostermeier et al. (2022); Scherr et al. (2020, 2019)
	Route Costs	Poeting et al. (2019a,b)

et al., 2018). Additionally, most evaluation problems fall under the category of NP-hard problems. Consequently, tackling large-scale instances requires significant computational powers, with the required effort likely escalating exponentially with the size of the problem. As a result, achieving optimal solutions for large-scale logistics problems is often not feasible in practice.

The relevant evaluation techniques discussed by the literature primarily deals with simulation-based modelling, Continuous Approximation (CA) modelling, and theoretical discrete route optimisation, as described by Daganzo (2005).

Simulation-based modelling includes Discrete Event Simulation (DES) (Balassa et al., 2023), Agent-based Simulation (ABS) (Poeting et al., 2019b), and System Dynamics (SD) (Serrano-Hernandez et al., 2021). DES models systems as a series of events occurring in continuous time, with each event altering the system's state. It is especially useful for processes with distinct events and discrete changes. However, DES may not adequately capture the complexity of individual behaviours or interactions between agents within the system (Guo et al., 2019). ABS models individual agents that act and interact autonomously within an environment. This approach can capture variability at the individual level, as well as interactions and feedback loops within the system. However, ABS can become computationally demanding as the number of agents and interactions grows, and defining realistic behaviours and interactions for agents may require detailed data and expert knowledge (Ballano et al., 2023). SD models systems using interconnected feedback loops and resource flows, focusing

on understanding system behaviour over time and the impact of feedback mechanisms. Nevertheless, SD may oversimplify individual interactions and behaviours and may not capture the detailed processes and events that DES or ABS can model (Heinen, 2021).

Simulation-based approaches offer a mathematical model-free evaluation framework, providing flexibility in constructing delivery networks and generating delivery scenarios. However, the simulators are typically customised according to the specific operation conditions and lack generalisation capability. Therefore, adapting them to similar problems often requires significant modifications, which can be time-consuming. Furthermore, the computational efficiency of simulators significantly decreases as the size of the problem instances increases (Raj et al., 2024).

Continuous Approximation (CA) is characterised by utilising continuous representations of input data and decision variables as density functions in time and space (Langevin et al., 1996). Its core concept involves approximating the objective function as a formulation of localised functions that can be optimised through relatively simple analytical operations. Each localised function approximates the cost structure of a local neighbourhood with approximately homogeneous settings. This homogeneous approximation allows mapping other high-dimensional decision variables to a lower-dimensional space. Even in the presence of spatial randomness and temporal dynamics, the optimal design of this neighbourhood can be obtained through straightforward calculations (Huang et al., 2013). Compared to discrete methods, CA imposes a lighter computational burden, is less reliant on precise input data, and offers managerial insights for large-scale problems.

Furthermore, CA is commonly employed to assess large-scale macroscopic delivery problems related to transportation routes, aiming to ascertain the optimal routes for vehicles to transport or pick up packages in a continuous space. A fundamental problem of route selection based on CA involves formatting or partitioning space into specific geometric shapes conducive to constructing approximately optimal vehicle routes (Ansari et al., 2018).

Although CA significantly simplifies the complexity of assessing large-scale delivery problems by estimating the optimal vehicle routes, it does so at the expense of the gap between the predicted optimal route and the truth optimal vehicle routes, and consideration

for finer operational details, such as assessing time consumption for different delivery activities. Therefore, striking a balance between evaluation complexity and the accuracy of predicted optimal routes is crucial.

A markedly distinct approach from CA is the route-based evaluation technique, which entails the precise solution of all routes within the system, i.e., solving the VRP. This technique permits the calculation of various cost components generated within each sub-route, offering more detailed managerial insights (Reed et al., 2022b). However, the route-based evaluation technique is prone to significantly increased computing times with the expansion of the problem scale.

As mentioned in Obj 1 of Section 2.1, given that the proposed macroscopic evaluation area encompasses Greater London and involves various delivery services and large-scale delivery fleets, simulation-based evaluation methods may lead to computational difficulties. Hence, integrating CA with a route-based evaluation approach achieves an excellent compromise between implementation feasibility and evaluation accuracy, and shows effectiveness in evaluating emerging autonomous delivery services such as drone delivery (Zhang et al., 2024). A potential approach is to design a mechanism that enhances the credibility of the evaluation by mapping a portion of the continuous space that has obtained the precise optimal routes through solving the exact VRP model to the remaining parts, given that all portions within the continuous space are homogeneous.

3.3 Mesoscopic Decision for Autonomous Last-Mile Delivery Service Selection

This section reviews three topics relevant to the the service selection of last-mile delivery: i) customer decision-making in last-mile delivery service selection; ii) equilibrium models of last-mile delivery service selection; and iii) the application of ML techniques in last-mile delivery service.

3.3.1 Customer Decision-Making in Delivery Service Selection

In the B2C model, recipients are assumed to choose from a range of delivery services with the aim of either maximising their utility or minimising associated expenses (Ban et al., 2019). These decisions are subject to influence from diverse operational variables, including factors like service charges, customers' travel distance for parcel retrieval, anticipated delivery duration, and overall delivery efficiency. The specific factors taken into consideration depend on the attributes of the employed delivery service.

In relation to the delivery service provided by stationary or mobile lockers, Weltevreden (2008), and Lagorio & Pinto (2020) have highlighted the significance of factors such as locker location and customer's travel distance for parcel pick-up, which greatly impact customer acceptance. Additionally, the availability of locker compartments also influences customer decisions (Lemke et al., 2016), meaning that the locker capacity is desired to be sufficient to increase customer usage under the premise of cost-savings and convenient handling. For ADR delivery services, the primary concerns of customers revolve around the efficiency of deliveries and the associated service charges (Pani et al., 2020; Boysen et al., 2018). In practical operational scenarios, it is anticipated that ADRs will be responsible for delivering a single parcel per trip and may encounter delays due to sidewalk collisions, resulting in extended waiting times for customers.

Existing studies on the last-mile delivery optimisation frequently assume that customers will either make arbitrary or predictable choices regarding the mode of parcel delivery or that logistics operators will make these choices on behalf of the customers (Deutsch & Golany, 2018; Schwerdfeger & Boysen, 2020; Orenstein et al., 2019). Customers, however, can choose their preferred delivery option in real-world operations, which is formulated as a probability distribution (Gul et al., 2014; Rusmevichientong et al., 2010; Gunawan et al., 2020). Different studies have attempted to characterise the probabilistic decision behaviour of customers using various utility models. For instance, Arentze et al. (2015) introduced an integrated random framework extended from discrete choice models to maximise utility function. As another example, Papola (2016) proposed a random utility model for service selection and routing selection problems. Although the aforementioned utility models show effectiveness in capturing

customers' probabilistic decision behaviours. Both of these patterns, however, ignore the impact of correlations among alternatives on customer decisions.

To resolve this limitation, Lin et al. (2020) and Lyu & Teo (2022) applied a Multi-Nomial Logit (MNL) model to predict customer selection on locker service usage. The findings demonstrated that the MNL model precisely interprets the model estimation outcomes through correlation analysis. This ensures that the model aligns effectively with the expected operational mechanism. Towards this end, the MNL model is a natural option for formulating customers' behaviour in the last-mile delivery service selection problem and will be tailored to suit the scenario in this study.

3.3.2 Equilibrium Models of Last-Mile Delivery Service Selection

To model the competitive dynamics of different delivery services, the majority of existing studies employ game theoretical approaches, where the ideology is interacting choices of agents/players to produce outcomes with respect to the preferences (or utilities) of those agents/players (Von Neumann & Morgenstern, 2007). Equilibrium refers to a situation where each player's strategy is optimal, given the strategies chosen by all the other players. In other words, no player has an incentive to unilaterally deviate from their chosen strategy, as doing so would not lead to a better outcome for them, this is also known as Nash equilibrium (Fudenberg & Tirole, 1991). Based on the distinctive game characteristics of various systems, there are different types of equilibrium models.

The Wardrop equilibrium (WE) is a frequently employed solution concept in network games, particularly in the modeling of transportation and telecommunication networks that experience congestion (Acemoglu et al., 2018). It assumes that a single player's action has a negligible impact on other participants, whilst deviations among a large group of players affect the utility of others (Young & Zamir, 2014; Hayel et al., 2016). WE model follows two principles: i) each player is considered an independent decision maker and minimises its own cost; ii) deviations from equilibrium do not benefit any player. Following these characteristics, WE model is capable to formulate customers' fully rational behaviour.

Table 3.3: Summary table of bounded rationality equilibrium (Adapted from Di & Liu (2016))

Type	Features	Limitations	Conditions	Citation
BRUE	Capture flow distribution patterns that flow cost is within an indifference band of the shortest path cost.	Difficulty in solving and may be mathematically intractable.	Increasing link cost functions (non-unique)	[1, 2, 3, 4]
IUE	Manifest the delay in people's response to information.	Cumbersome enumeration of all actions which require nontrivial estimation.	(Strictly) monotone link cost functions (non-unique)	[5, 6]
SUE	Capture players' perception errors and others' unknown choices.	Not statistical and thus not for parameter estimation.	Monotonically increasing functions (unique exist)	[7, 8, 9]
BUE	Captures searching behaviour using a positive behavioural model	Not guarantee an equilibrium exists.	N/A	[10, 11, 12]
QRE	Non-falsifiability; players are assumed to make errors in the decision.	Loss of information; Variation source of rationality parameters is unclear.	Monotonically increasing functions (unique exist)	[13, 14, 15, 16, 17]

Notes: BRUE - Boundedly Rational User Equilibrium; IUE - Initial User Equilibrium; SUE - Stochastic User Equilibrium; BUE - Bi-objective User Equilibrium; QRE - Quantal Response Equilibrium; [1] - Guo (2013); [2] - Lou et al. (2010); [3] - van Essen et al. (2016); [4] - Ye & Yang (2017); [5] - Zhang (2011); [6] = Zhang & Yang (2015); [7] - Daskin (1985); [8] - Di & Liu (2016); [9] - Riemann et al. (2015); [10] - Wang et al. (2010); [11] - Zhu et al. (2023); [12] - Xu et al. (2023); [13] - Hayel et al. (2016); [14] - Rogers et al. (2009); [15] - Goeree et al. (2016); [16] - Friedman & Mauersberger (2022); [17] - Wang et al. (2022)

To model customers' bounded rational decision behaviour regarding delivery service selection, prior studies have developed various equilibrium models, as show in Table 3.3, such as the Bounded Rational User Equilibrium (BRUE), Initial User Equilibrium (IUE), Stochastic User Equilibrium (SUE), Bi-objective User Equilibrium (BUE), and Quantal Response Equilibrium (QRE) models.

BRUE model is applicable for situations where none of the players can significantly reduce their costs through individual action within a specific range (Lou et al., 2010). However, the introduction of new flows into the system causes the non-uniqueness of the equilibrium due to the presence of indifference bands in the BRUE model. Such non-uniqueness creates uncertainty

within the system. This weakness limits the applicability of the BRUE model in this study, as the competition between MPLs and ADRs forms a unique equilibrium, rather than the result of multiple equilibrium patterns.

IUE is a variant of user equilibrium that takes into account a subset of all available alternatives (Zhang, 2011; Zhang & Yang, 2015). However, the presence of inertia implies that the performance of IUE cannot be improved, even when more information is made available to the players (Guo, 2013). This attribute contradicts the hypothesis of this research, which proposes that the system's performance can be enhanced by offering global information, thereby promoting customers to take action that benefits both the system and themselves.

SUE model is used to model continuous games in cases where the number of players goes to infinity (Di & Liu, 2016). It assumes that players' utility is random due to their perception error, and none of the players can improve their utility through individual action (Riemann et al., 2015). Nevertheless, the SUE model primarily considers the system's overall equilibrium and overlooks the diverse response strategies adopted by individual players throughout the game process (Daskin, 1985). In the proposed study, we aim to examine the mixed strategies employed by customers in response to the choices made by other players. This entails investigating individual behaviour and response strategies to understand how customers may opt for a different service if all other players are selecting the same one. Therefore, the SUE model is not suitable for our case.

QRE is a solution to the limitations of previous equilibrium models. It seeks to formulate the mixed strategies that individuals may use in response to the decisions of other players, under the assumption that all players are homogeneous (Rogers et al., 2009; Friedman & Mauersberger, 2022; Wang et al., 2022). A typical specification for QRE is Logit equilibrium, in which the player's strategies are performed according to a probability distribution, and its errors follow a bell distribution (Anderson et al., 2002).

Different studies have validated the effectiveness of QRE in capturing customers' bounded rational behaviour in service selection that are relevant to this study. For instance, Hayel et al. (2016) and Chen et al. (2016) applied Wardrop and Logit equilibrium from congestion/routing games to incorporate customers' varied bounded rational behaviour of last-mile delivery service

selection between the relay station and delivery at home. This mimics the delivery service options introduced in the proposed study (see Chapter 4), and it can be concluded that the Logit equilibrium is the most appropriate model to formulate customers' bounded rational behaviour. To this point, three main limitations have been identified during the review and will be resolved in Chapter 4:

- The source of customers' bounded rationality in relation to various delivery options was not disclosed by Hayel et al. (2016). Moreover, the customer utility function did not incorporate crucial factors that affect the error estimation of the payoff associated with different service options. To tackle these limitations, the proposed study has incorporated critical parameters related to various ALD services into the payoff function and examined their influence on the system.
- The variance in customer tolerance towards different levels of service convenience and efficiency has not been considered in the customer utility function, which does not align with real-world operational scenarios. To resolve this limitation, the proposed study introduces piece-wise functions to capture different levels of customer tolerance for service convenience and efficiency.
- The existing model assumes that customers make decisions based solely on their individual utility in the present state. Consequently, the influence of global information on customer decision-making has not been investigated. Global information encompasses factors such as congestion levels or service occupancy, which can be considered as observable strategic cases in a queuing system. To address this limitation, this study proposes an improved algorithm that leverage global information to enhance customer decision-making patterns, thereby promoting system efficiency.

3.3.3 Applications of Machine Learning Techniques in Last-Mile Delivery Service

The machine learning techniques include supervised learning (with truth labels) and unsupervised learning (without truth labels), classified by whether the true label is applied

in the training set. In the logistics field, unsupervised learning techniques such as clustering are normally applied to address location selection problems due to the absence of truth facility location during training (Giuffrida et al., 2022). Conversely, supervised learning approaches such as Kernalised Support Vector Machine (KSVM) and decision trees are commonly utilised for classification or prediction tasks, as they rely on labelled training data to make accurate predictions (Plissonneau et al., 2022). Given that the proposed study involves making predictions for the service pricing, this section focuses on supervised learning techniques for the review purpose.

Table 3.4 summarises the characteristics of different supervised learning techniques for regression predictions. The review highlights that the selection of the appropriate supervised learning model mainly depends on the following considerations: i) nature of the data; ii) number of features and feature scaling; iii) interpretability; iv) overfitting and generalisation; and v) time complexity (Jordan & Mitchell, 2015).

Regarding the nature of the data, it should be considered whether the relationship between the independent and dependent variables is anticipated to be linear or nonlinear. In cases of linear relationships, linear regression proves to be apt, whereas for non-linear relationship, models such as decision trees or polynomial regression may be requisite (El Naqa & Murphy, 2015).

When dealing with data characterised by an extensive array of features, techniques such as Lasso or Ridge regression can assist in feature selection and mitigating multicollinearity (Zheng & Casari, 2018). Additionally, it's crucial to contemplate the magnitude of each feature. Some models, like KSVM or K-Nearest Neighbors, are sensitive to feature scales, necessitating the implementation of feature engineering such as normalising feature scales.

For the circumstance that needs a clear understanding of how individual features influence the outcome (interpretability), opting for models such as linear regression or KSVM might be favoured over more intricate approaches like neural networks or gradient boosting (Lipton, 2018).

Complex models such as random forests and gradient boosting have the capacity to grasp intricate relationships among different features, yet they are susceptible to overfitting. To

Table 3.4: Summary table of main supervised learning techniques for regression prediction

Method	**Use Case**	**Pros**	**Cons**
Linear Regression	Predict a continuous value	Easy to understand; Interpretable coefficients; Handle many features	Assumes a linear relationship between variables; Sensitive to outliers
Decision Trees	Regression with categorical and numerical features	Capture non-linear relationships in the data; Interpretable; Handle both numerical and categorical features	Prone to overfitting; Can be unstable
Random Forest	Regression with complex interactions	Combines multiple decision trees to reduce overfitting; Handles non-linearity and interactions well; Robust to outliers	Computationally expensive; Less interpretable compared to individual trees
KSVM	Regression with complex relationships	Effective in high-dimensional spaces; Handle non-linear relationships using kernels; Robust to overfitting	Sensitive to the choice of kernel and parameters; Computationally intensive for large datasets
K-Nearest Neighbors	Regression based on similarity measures	No training phase; Capture local patterns in the data	Sensitive to the choice of neighbors size; Computationally expensive for large datasets
Gradient Boosting	High-performance regression	Handles complex relationships; Robust to outliers and noisy data	Computationally expensive; Requires tuning of hyperparameters
Ridge	Regression with multicollinearity	Handles multicollinearity well; Prevent overfitting	Does not perform feature selection
Lasso	Feature selection	Performs feature selection; Handles multicollinearity	Sensitive to the scaling of features; Can only select features up to a certain extent
ElasticNet	Regression with many features	Handle multicollinearity and perform feature selection	Adds additional hyperparameter to tune
Huber	Regression with potential outliers	Robust to outliers	Require more tuning

mitigate this risk, employing regularised regression techniques like Ridge, Lasso, or ElasticNet can be instrumental (Ying, 2019). For time complexity, linear regression and KSVM exhibit the capability to be trained swiftly, even when handling sizeable datasets (Cai et al., 2018).

Among the aforementioned techniques, KSVM has seen a growing application in issues pertaining to last-mile delivery and the demand prediction (Liu & Li, 2016; Pan et al., 2017; Hughes et al., 2019; Yu et al., 2020a; Cameron, 2021). The KSVM offers an effective solution

to the regression problem by demonstrating high prediction accuracy, even when the number of features exceeds the sample size. Unlike utilising the entire dataset, KSVM relies on a subset of support vectors to make predictions (Liu et al., 2017). This property proves particularly valuable in predicting service selections, where obtaining comprehensive customer data can be challenging due to privacy concerns (Pan et al., 2017).

The use of KSVM in modelling distribution services selection involves inputting predetermined incentives for a specific service option and outputting a parameter that indicates whether the company can persuade consumers to switch their prior selection to accept a new service by using these incentives that contribute most to the profit (Cameron, 2021). However, accurately determining representative feature vectors that capture the impact of these incentives can be challenging, mainly due to the requirement for appropriate kernel selection. As a result, the preprocessing of the feature vectors (e.g., using feature engineering) can significantly improve the performance of the KSVM (Hughes et al., 2019).

Another major issue of implementing KSVM is selecting an appropriate kernel that can effectively fit the characteristics of the sample data. In Chapter 5, we evaluate the performance of three different kernels: Linear, Radial Basis Function (RBF), and Sigmoid on the dataset to determine the optimal kernel. Additionally, we employ data scaling and polynomial interaction to enhance the prediction accuracy of the KSVM, as the number and interaction of the feature vectors can impact the performance of the model.

3.4 Operational Optimisation for Autonomous Last-Mile Deliveries

The operational optimisation for ALD revolves around route planning (Ostermeier et al., 2022; Chen et al., 2021a,b; Ulmer & Streng, 2019), fleet and infrastructure planning (Poeting et al., 2019a,b), and investigating factors impacting efficiency (Alfandari et al., 2022; Simoni et al., 2020), while route planning receives the most attention from the existing studies (Srinivas et al., 2022). This encompasses the use of delivery fleets with different capacities to cater to a set of customers spread out in various locations around depots (Clarke & Wright, 1964; Braekers et al.,

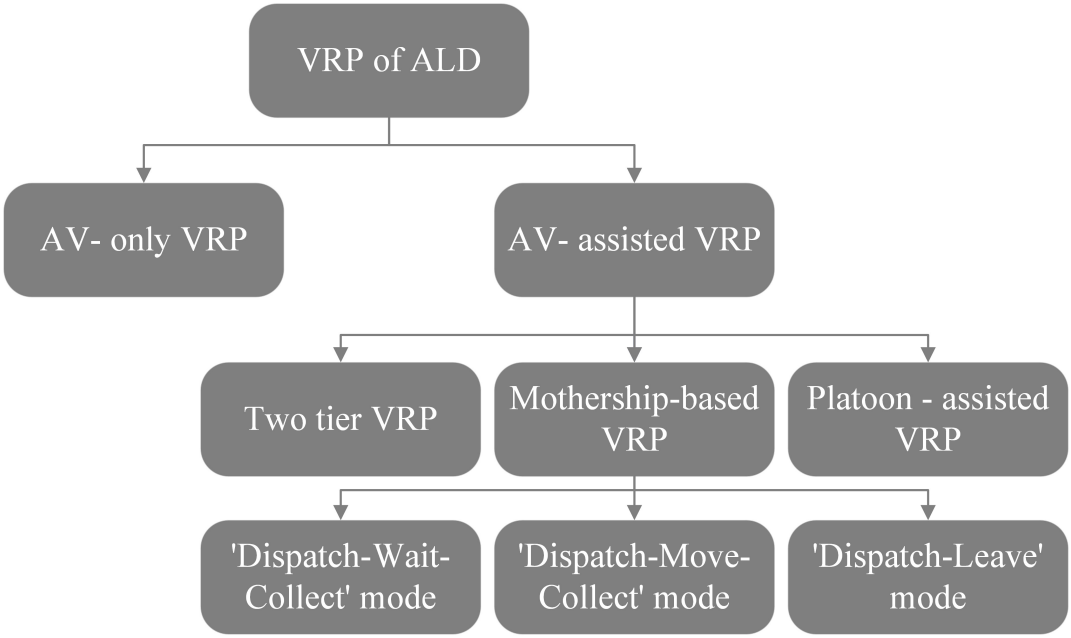


Figure 3.1: Classification of VRPs in ALD. Based on Srinivas et al. (2022)

2016). A comparable issue is the LRP, which entails not only determining the routes for vehicles but also making concurrent choices regarding the placement of facilities such as depots, hubs, and plants (Drexl & Schneider, 2015). In contrast to VRP, which addresses short-term and daily decisions, LRP encompasses both tactical and operational decision-making (Sonneberg et al., 2019).

VRPs have grown intricate in their efforts to encompass real-world intricacies such as time-dependent traffic congestion or travel durations, as well as accommodating customer preferences like specified time windows for deliveries and pick-ups, along with dynamically fluctuating demand (Braekers et al., 2016). The relevant literature on VRP in the ALD domain can be classified based on the problem characteristics and application areas, as shown in Figure 3.1.

The contents of this section are structured as follows: Section 3.4.1 reviews different types and the corresponding characteristics of VRP in the ALD domain. Section 3.4.2 reviews the existing solution techniques for VRPs.

3.4.1 Route Planning Model for ALD

As depicted in Figure 3.1, the categorisation reveals two VRP variants within the ALD context: AV-only VRP and AV-assisted VRP.

AV-only VRP

The AV-only VRP aims to fulfil a set of customers soley utilising AV fleets that independ from conventional vehicles, stationed at various urban consolidation centres (Srinivas et al., 2022). Each urban consolidation centre imposes an upper limit on both the quantity of AVs and parcels it can accommodate. Additionally, each AV possesses specific attributes including payload capacity, compartment count, and operational range constraints (Sonneberg et al., 2019). These AV features constrain the demand nodes accessible from a particular urban consolidation centre and set the maximum number of customers that can be serviced in a single trip. Typically, AVs commence and conclude their routes at a consolidation center, with their routes optimised to minimise (or maximise) a particular objective (e.g., operational expenses or delivery duration). Furthermore, each customer is attended to precisely once by an AV (Ulmer & Streng, 2019).

Considering that the AV-only delivery scenario operates independently of external vehicles and requires superior adaptability to urban traffic conditions and delivery capacity to ensure efficient deliveries, the AV-only scenario is applicable for modelling VRP that utilise medium-sized AVs, such as MPLs. In contrast to alternative autonomous delivery technologies like ADRs, MPLs are employed for unattended deliveries, obviating the necessity for customers to be present at the delivery site when parcels are dispatched from the MPLs.

Table 3.5 summarises the prior studies that focus on utilising MPLs for delivery. Orenstein et al. (2019) proposed a flexible parcel delivery problem that allows customers to provide multiple delivery locations to couriers. However, this model does not account for the scenario that customers may change their preferred destinations if initial delivery attempts fail. This is a significant consideration, as it is estimated that 42% of orders delivered to consumers without specific delivery times will result in repeated delivery attempts (Allen et al., 2017). To address this issue, Chapter 6 proposed a model enumerates all potential customer stopovers and considers the overlap of the corresponding time windows when formulating the delivery scenario to ensure that all demands can be met at specific parking spaces.

Schwerdfeger & Boysen (2020) introduced a mixed-integer programming model that

considers customer coverage and minimises the MPL fleet size by determining MPL stopovers and their corresponding routes. While their exact solution approach demonstrates significant improvements compared to stationary locker systems, several key limitations arise: (i) it does not consider the dynamic scenario (e.g., parcel resupply) and stochastic factors, such as service delays, which are present in real-world operations and can affect the dynamic adjustment of scheduling strategies; (ii) it does not track the operational details of a single MPL, such as starting and ending times and duration at parking spaces, which limits in-depth route optimisation for a single MPL and leads to a larger gap from the optimal solution; (iii) from a computational perspective, the efficiency of the proposed exact solution approach is significantly lower (computing time $\approx$ 60 min) than most evolutionary algorithms in large networks (nodes $\geq$ 100).

To address the aforementioned limitation (i), Chapter 6 incorporated service delay and travel distance into the objective function and implementing route adjustment strategies to address latency considering parcel resupply. To address limitation (ii), Chapter 6 designed a RL-based algorithm that uses a state-based structure for solution extraction, which allows the delivery scheme of each MPL to be interpreted and revised by the proposed algorithm. To address limitation (iii), Chapter 6 further developed global and local search mechanisms to achieve a better balance between solution quality and computational efficiency compared to traditional exact and heuristic approaches.

Wang et al. (2020) integrated the LRPs with MPLs and established a non-linear integer programming model to minimise operating costs. The results show that the policy that considers demand aggregation can significantly reduce delivery times while a scheme without demand aggregation would reduce the number of vehicles required. Hence, considering demand aggregation to reduce fleet size and travel distances is an option when formulating the model.

Li et al. (2021a) introduced a two-echelon delivery network, in which MPLs travel to couriers in the field and transfer parcels between the couriers and the depot, while couriers perform the final delivery without making repeated returns to the depot. A hybrid Clark and Wright heuristic was developed to solve the mixed-integer programming model. The common weakness of the model introduced by Li et al. (2021a) and Wang et al. (2020) is that all

customer demand patterns are assumed to be static in contrast to the stochastic nature of real-world customer requests, and the proposed method becomes computationally infeasible for larger problem instances.

Schwerdfeger & Boysen (2022) evaluated the delivery performance of different concepts of MPLs, which are classified by whether the locker is mounted or loaded on the vehicle body, and whether the vehicle is autonomous or human-drive. The findings indicate that if customers prioritise convenience and prefer minimal walking distances and extended time windows for retrieving parcels from a designated location, MPLs can meet these service objectives without requiring substantial investment in an extensive network of stationary lockers. On the other hand, MPLs may generate higher traffic volume than stationary lockers, which only require intermittent parcel storage. However, it is important to note that this observation holds validity only when customers physically visit their designated parcel lockers. Therefore, it is imperative to devise an effective route adjustment mechanism to mitigate the impact of MPLs on traffic flow and address the influence of external traffic flow on MPL operations. This can be achieved by, for example, avoiding the repositioning of MPLs during peak commuting hours.

AV-assisted VRP

While the AV-only delivery scenario is becoming a reality, most of its current applications are associated with facilities (such as curbside drop-off points) and outdoor locations (such as university campuses and business districts). It is worth noting that there is limited research on pure AV-only delivery scenarios, attributed to specific AV characteristics such as slower speeds, limited service range, and payload capacity. However, recent advancements have greatly been made by proposing delivery concepts involving collaboration between AVs and conventional vehicles (e.g., trucks/vans) to overcome these constraints. It has been reported that in certain scenarios, these hybrid approaches can be more beneficial compared to traditional or AV-only delivery scenarios (Srinivas et al., 2022). Furthermore, in alignment with these advantages, most AV-related route planning studies have focused on these hybrid AV-assisted delivery systems in the context of last-mile logistics. Particularly, most studies focus on the application of small robots (like Amazon delivery bots) accompanied by conventional vans.

Table 3.5: Summary table of MPLs-related studies. Based on Liu et al. (2023a)

Citation	Model/Method	Network Type/Size	Objectives	Key Constraints
Orenstein et al. (2019)	CVRP/Saving heuristics & tabu search, LNS	CDPs with alternative delivery points/Up to 50 service points	Minimise fixed and travel costs, penalty of undelivered parcels	Time windows, locker capacity, demand fulfilment, split deliveries are not allowed, fleet size at SP
Schwerdfeger & Boysen (2020)	Dynamic LRP/Domain reduction, MIPs, exact approach	CDPs/Up to 300 customers, 40 parking locations	Minimise fleet size	Time windows, locker capacity, service radius
Wang et al. (2020)	CVRP with demand aggregation/MIPs, Multilayer GA	CDPs with alternative collection points/32 nodes	Minimise route and maintenance costs	Time windows, locker capacity, service range, demand fulfilment, fleet size at SP
Li et al. (2021a)	LRP/Hybrid Clark & Wright heuristic	CDPs with 2-Echelon/Up to 4 depots, 200 customers	Minimise depot rental, locker deployment and the operational costs	Locker capacity, service radius, demand fulfilment, visiting times at SP
Serrano-Hernandez et al. (2021)	FLP/Agent-based modelling	Population size is 203,944	Minimise operational and service costs	Set up & retire rate of MPLs, demand fulfilment, utilisation rate of MPLs
Schwerdfeger & Boysen (2022)	Dynamic LRP/Heuristic multi-stage optimisation	CDPs/Up to 50 customers, 15 parking locations	Minimise fixed and variable costs for MPLs' relocation and driver assignment	Time windows, locker capacity, service radius, travel path
Kötschau et al. (2023)	FLP/MIPs	Heterogeneous locker network/Up to 400 customers	Maximise the number of customers served	Customer fulfilment, MPL capacity, flow conditions, customer preferences

1) FLP = Facility location problem; 2) MIPs = Mixed integer programming; 3) CVRP = Capacitated vehicle routing problem; 4) LNS = Large neighbourhood search; 5) CDPs = Collection and delivery problems; 6) SP = Service point.

The existing AV-assisted delivery model can be classified into the three categories: two-echelon model (Figure 3.2a), mothership model (Figure 3.2b-d), and platoon model (Figure 3.2e). In two-echelon model, traditional vans function in the first echelon, conveying a substantial volume of packages from a central depot to nearby AV depots. Subsequently, AV

fleets are operating in the secondary echelon, and handle the distribution of these parcels to final recipients (Poeting et al., 2019a,b; Bakach et al., 2021; Alfandari et al., 2022).

The mothership model entails the modification of conventional trucks or vans into a centralised system referred to as a "mothership". This mothership is equipped with several compact ADRs and has the capacity to store multiple packages. Under this model, the mothership navigates through various neighbourhoods and dispatches its ADRs, loaded with parcels, to conduct deliveries to nearby customers (Simoni et al., 2020; Ostermeier et al., 2022; Chen et al., 2021a,b; Yu et al., 2022a; Jennings & Figliozzi, 2019; Boysen et al., 2018). The current research on VRP employing the mothership concept primarily varies in the strategy adopted by the mothership van after deploying its ADRs:

The first type of mothership model involves using a "Dispatch–Wait–Collect" mode (Figure 3.2b), where the mothership van transports parcels to customers with available parking spaces for service. While the van is attending to a customer, ADR fleets are loaded with parcels and are deployed from the van for collective delivery. During this process, the van must remain at the customer parking location until it finishes attending to its customer and all the ADRs return after completing their deliveries. Significantly, the existing studies include a stipulation that some customers are excluded from ADR deliveries (Chen et al., 2021a,b). This inclusion enhances the feasibility of the solution, accounting for situations where an ADR may be unable to fulfil a specific customer due to factors such as customer preference and restricted payload capacity. Nevertheless, the studies restrict a single ADR to serve only one customer per delivery trip.

The second type of mothership mode entails a "Dispatch–Move–Collect" policy (Figure 3.2c), where a mothership van proceeds with its route after releasing the ADRs, without pausing at the initial drop-off location waiting for the slower ADRs to come back. Nevertheless, it remains the mothership's duty to retrieve ADRs at a later stage (Simoni et al., 2020; Jennings & Figliozzi, 2019; Yu et al., 2022a). Compared to the "Dispatch–Wait–Collect" mode, "Dispatch–Move–Collect" policy may achieve greater time savings since the van do not need to wait at the initial drop-off locations. Moreover, it enables ADRs to cater to multiple customers in a single journey, making use of the multi-compartment configuration within the ADRs.

In contrast to the majority of studies where the van is designated with the task of retrieving the dispatched ADRs, a limited number of studies introduce a "Dispatch-Leave" policy (Ostermeier et al., 2022; Boysen et al., 2018), as shown in Figure 3.2d. The urban AV stations serve as designated service points where ADRs can autonomously return after completing their delivery tasks. Unlike ADR depots shown in Figure 3.2b, these ADR stations do not store parcels, but rather, they function akin to compact garages. Moreover, they aid vans in replenishing parcels of ADRs during their routes, eliminating the need to revisit the central depot.

Recent research indicates that there are instances where ADRs necessitate transportation to the AV zone for the execution of autonomous delivery tasks. This implies that certain ADRs may require guidance during transit between AV zones, as the routes in between may not be conducive for autonomous driving. To address this, a platoon model is introduced where a human-driven vehicle can lead a group of ADRs in a platoon formation (Scherr et al., 2019). This entails a fleet of ADRs closely following one another (i.e., in a tail-to-head configuration) while replicating the driving actions of the leading vehicle. The ADRs have the capability to join these platoons whenever they need to traverse between AV zones.

Table 3.6 summarises the model characteristics of all existing AV-only and AV-assisted VRP studies. Much like research in the AV-only scenario, the primary objective in most AV-assisted studies is to minimise delivery time, travel distance, and associated costs. However, a limited number of AV-assisted studies have given specific attention to the minimisation of tardiness or service delays (Boysen et al., 2018; Alfandari et al., 2022). This is especially crucial since ADRs in the AV-assisted models currently rely on attended deliveries. Furthermore, only a minority of ADR-based routing studies include strict delivery time windows (Ostermeier et al., 2022; Chen et al., 2021a,b; Yu et al., 2020b). Nevertheless, attended deliveries necessitate the avoidance of both premature deliveries and delays.

Additionally, the reviewed literature reflects that higher delivery efficiency can be achieved by accommodating multiple customer orders in separate compartments of a single AV, which has been integrated into routing problems in various studies (Simoni et al., 2020; Jennings & Figliozzi, 2019; Figliozzi, 2020; Sonneberg et al., 2019; Ulmer & Streng, 2019; Yu et al.,

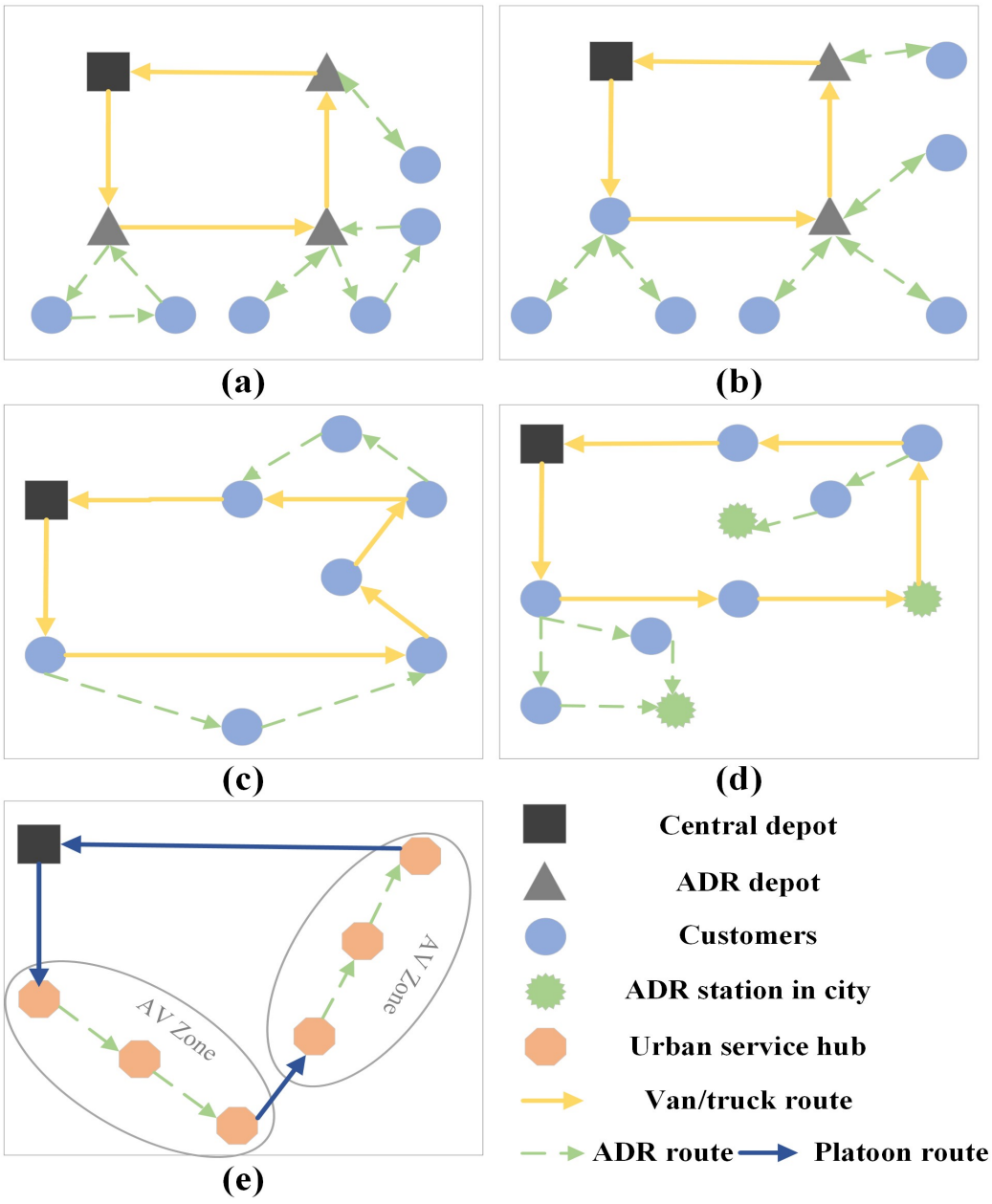


Figure 3.2: Different types of AV-assisted VRP. (a) two-echelon mode, (b) 'dispatch-wait-collect' mode, (c) 'dispatch-move-collect' mode, (d) 'dispatch-leave' mode, (e) platoon mode. Based on Srinivas et al. (2022)

2020b). However, the study pertaining to two-tier AV-assisted deliveries and the mothership model with a "Dispatch-Wait-Collect" mode have yet to capitalise on this feature. Therefore, a single-echelon AV-assisted delivery model that utilises the feature of multiple compartments of the AV has the potential to reduce delivery time.

In addition to leveraging the above advantages, only a limited number of studies have considered real-world constraints, such as the accessibility of final destinations. Among the studies where vans are involved in parcel deliveries, Yu et al. (2020b) incorporate destinations

that are restricted for van access due to factors like narrow streets, campus areas, or restricted regions, necessitating service by AVs.

Similarly, Chen et al. (2021b) and Chen et al. (2021a) account for destinations where AVs face accessibility challenges. In platoon-based AV-assisted deliveries, Scherr et al. (2019) and Scherr et al. (2020) address delivery locations outside AV zones, which AV alone may not be able to serve. However, most of the existing literature has yet to incorporate such practical considerations. Finally, while current research on AV routing predominantly focuses on delivery operations, there is a scarcity of studies on hybrid pick-up and delivery VRPs.

Table 3.6: Summary table of the existing AV-only and AV-assisted VRP studies. Based on Srinivas et al. (2022)

Citations	**AV-only**	**Dest**	**No. of AVs**	**AV Capacity**	**No. of Vans**	**No. of AVs in Van**	**Van Perform Delivery**	**Time Window**	**Hybrid Pick-up & Delivery**	**Van cannot access door**	**AVs cannot access door**
Sonneberg et al. (2019)	Y	E	M	M	-	-	N	Y	N	Y(AV)	-
Gnegel et al. (2021)	Y	E	M	1	-	-	N	Y	N	Y(AV)	-
Ulmer & Streng (2019)	Y	S	M	M	-	-	N	Y(D)	N	Y(AV)	-
Reed et al. (2022b)	Y	E	1	M	-	-	N	Y	N	-	-
Fang et al. (2013)	Y	S	1	M	-	-	N	Y	N	-	-
Wu et al. (2017)	Y	S	1	M	-	-	N	Y	N	-	-
Scherr et al. (2019)	N	S	M	M	M	M(P)	Y	Y(A)	N	N	Y
Scherr et al. (2020)	N	S	M	M	M	M(P)	Y	Y(A)	N	N	Y
Lemardelé et al. (2021)	N	E	M	M	M	-	N	Y	N	Y(AV)	N
Bakach et al. (2021)	N	E	M	1	1	-	N	Y	N	Y(AV)	N
Alfandari et al. (2022)	N	E	M	1	1	-	N	Y(D)	N	Y(AV)	N
Poeting et al. (2019b)	N	E	M	1	1	-	Y	Y	N	N	N
Poeting et al. (2019a)	N	E	M	1	1	-	N	Y	N	Y(AV)	N
Figliozzi (2020)	N	E	M	M	M	M	N	Y	N	Y(AV)	N
Chen et al. (2021b)	N	E	M	1	M	M	Y	Y	N	N	Y
Chen et al. (2021a)	N	E	M	1	M	M	Y	Y	N	N	Y
Jennings & Figliozzi (2019)	N	E	M	M	1	M	N	N	N	Y(AV)	N
Simoni et al. (2020)	N	E	1	M	1	1	Y	N	N	N	N
Yu et al. (2020b)	N	E	M	M	M	M	N	Y	N	Y(AV)	N
Yu et al. (2022a)	N	E	M	M	M	1	Y	Y	Y	Y	N
Boysen et al. (2018)	N	E	M	1	1	M	N	Y	N	Y(AV)	N
Ostermeier et al. (2022)	N	E	M	1	1	M	N	Y	N	Y(AV)	N

LEGEND: Dest: destination; Y: yes; N: no; E: end-customers; S: service hub; M: multiple; (P): platoon length; (D): restricted deadline; (A): restricted arriving time; (AV): by default only AVs can deliver parcel to the door directly

3.4.2 Comparison of Existing Route Planning Models

The techniques for solving VRPs in terms of ALD scenarios can be categorised into exact methods and heuristic algorithms (as shown in Table 3.7). The exact methods primarily employ commercial solvers (e.g., Gurobi, CPLEX) to explicitly solve linear mathematical models. For those nonlinear mathematical models that cannot be directly addressed using solvers, they are typically transformed into linear models through techniques such as constraint relaxation or convex optimisation, allowing the utilisation of solvers for resolution. Considering most reviewed literature used small customer sizes (e.g., nodes ≤ 50), exact methods are effective for solving the models with acceptable computation time.

However, the NP-hard nature of VRPs makes exact solutions inefficient and impractical for large problem instances. Therefore, heuristic (Duhamel et al., 2010) and meta-heuristic approaches (Vincent et al., 2021) are commonly used to solve intermediate and large models because they can find near-optimal solutions with less computation time. These heuristics methods include heuristic lagrangian relaxation (Yang et al., 2022), adaptive large neighbourhood search (Chen et al., 2021a; Yu et al., 2022a), and simulated annealing algorithm (Poeting et al., 2019a; Vincent et al., 2022). These studies have revealed that the effectiveness of (meta-)heuristics on improving solution quality for VRPs.

Nevertheless, (meta-)heuristic algorithms can be susceptible to encountering local optima. Studies within the realm of RL have showcased its capability to efficiently tackle extensive, time-critical VRPs by utilising environmental cues to uncover superior solutions (Bello et al., 2016; Cappart et al., 2021; Nazari et al., 2018). RL-based approaches that have been used to solve VRPs can be divided into two categories: i) end-to-end methods (e.g., Pointer network, Ptr-Net) and ii) the use of deep RL (DRL) to improve iterative search algorithms. End-to-end methods provide a scalable and efficient framework for optimising deep neural networks, but they tend to produce lower quality solutions for medium to large problem instances compared to other combinatorial optimisation solvers such as LKH3, Google OR Tools, and Gurobi (Wang & Tang, 2021).

DRL-based iterative search frameworks use neural networks to learn and choose heuristic

Table 3.7: Summary table of the existing approaches for solving AV-only and AV-assisted VRPs. Based on Srinivas et al. (2022)

Citations	**Methods**	**Customer Size**
Sonneberg et al. (2019)	MILP	10
Gnegel et al. (2021)	MIP, layer graph	50
Ulmer & Streng (2019)	Policy function approximation	1000 (orders)
Reed et al. (2022b)	IP	36
Fang et al. (2013)	IP, cut & solve algorithm	30 (tasks)
Wu et al. (2017)	IP, two-phase exact algorithm	55 (tasks)
Scherr et al. (2019)	ILP	12 (hubs)
Scherr et al. (2020)	IP, dynamic discretisation discovery	12 (hubs)
Lemardelé et al. (2021)	Continuous approximation	500 (orders)
Bakach et al. (2021)	MIP, p-median problem	300
Alfandari et al. (2022)	MILP, Benders de-composition	100
Poeting et al. (2019b)	Agent-based simulation	15,484 (orders)
Poeting et al. (2019a)	IP, Simulated annealing	194,651 (orders)
Figliozzi (2020)	Continuous approximation	200
Chen et al. (2021b)	MILP, two-stage meta-heuristics	50
Chen et al. (2021a)	MILP, improved ALNS	200
Jennings & Figliozzi (2019)	Continuous approximation	192
Simoni et al. (2020)	IP, weighted interval scheduling	50
Yu et al. (2020b)	MILP, hybrid meta-heuristic	30
Yu et al. (2022a)	MILP, ALNS	28
Boysen et al. (2018)	MIP, multi-start local search	25
Ostermeier et al. (2022)	MIP, local search	125

MILP: mixed integer linear programming; MIP: mixed integer programming; IP: integer programming; ILP: integer linear programming; ALNS: adaptive large neighbourhood search

rules within heuristic algorithms to improve the speed of generating new solutions for classical heuristics in large problem instances. While these frameworks offer better solution quality compared to end-to-end methods, they sacrifice generalisation ability in solving similar type of problems. Hence, the trade-off between optimisation capability and generalisation ability in different problem instances should be considered due to the nature of RL-based methods (Sutton & Barto, 2018).

Table 3.8 summarises the representative studies of each RL technique mentioned above. Although there has been significant progress in utilising RL for addressing VRPs, there exist certain research gaps that necessitate attention. To the best of the author's knowledge, most RL algorithms developed for VRP can only handle single-feature or a very limited number of decision variables, lacking the capability to manage the multi-dimensional decision variables

present in real-world operational scenarios. Therefore, novel methodologies are imperative to train RL algorithms using different decisional features for practical application in VRPs (Yan et al., 2022). Furthermore, the computational instances predominantly involve randomly generated nodes on square-shaped maps. Limited attention has been directed towards maps characterised by significantly skewed coordinates, a scenario frequently encountered in practical applications.

The reviewed studies further reveal that the solution quality of RL techniques will be affected by the performance of sampling, which is determined by the structure of encoder & decoder (James et al., 2019) or the state interpretation (Zhao et al., 2020). Hence, Chapter 6 designs a encoder within the RL model to contain a set of vectors that combine task allocation tuples and task execution sequence tuples of VRP, such that provides better initial solution for the main optimisation procedure/loop. Compared with the classic encoder that only uses the coordinates of nodes in traditional pointer networks (Deudon et al., 2018; Vaswani et al., 2017; Vinyals et al., 2015), the designed encoder contains more decision parameters without transforming the original coordinates, improving computing speed and solution quality while can be extended to formulate multiple decision variables accordingly.

Beyond the above processing in sampling phase, Chapter 6 further develops a novel attention-based mechanism to decode the processed vector with superior solutions during the state interpretation and transition phase. The optimal strategies are obtained through continuous training with a trial-and-error mechanism. The training process of the designed RL algorithm follows an offline manner and does not rely on historical samples.

Compared with other online training DRL methods with on-policy or off-policy sampling, the offline training will significantly improve the sample efficiency since the offline data/sample will be generated based on the behaviour policy. To overcome the weakness of solution quality faced by most offline RL methods, we apply ϵ-greedy, and global and local search mechanisms to balance the exploitation-exploration dilemma.

Table 3.8: Classification of RL methods in solving VRP. Based on Liu et al. (2023a)

Class	Citation	Method	Problem Instance & Performance
DRL-based Iterative Method	Vinyals et al. (2015)	Ptr-Net; supervised training	30-TSP: near optimal and better than heuristics algorithm; 40, 50-TSP: significant gap towards optimality
	Bello et al. (2016)	Ptr-Net; REINFORCE & Critic baseline	50-TSP: solution outperforms Vinyals et al. (2015). 100-TSP: near-optimal solved by Concorde. 200-Knapsack: optimal solution
	Nazari et al. (2018)	Ptr-Net; REINFORCE & Critic baseline	100-TSP: save training time around 60% compared with Vinyals et al. (2015); Bello et al. (2016); 100-CVRP: better than heuristics algorithms
	Deudon et al. (2018)	Transformer attention; REINFORCE & Critic baseline	20, 50-TSP: solution obtained better than Bello et al. (2016). 100-TSP: result similar to Bello et al. (2016)
	Vaswani et al. (2017)	Transformer attention; REINFORCE & Rollout baseline	100-TSP: solution obtained better than Bello et al. (2016); Vinyals et al. (2015); Nazari et al. (2018); Deudon et al. (2018). 100-CVRP, 100-SDVRP, 100-OPVRP, 100-PCTSP, SPCTSP: near optimal solved by Gurobi, and better than heuristics algorithms
	Kool et al. (2018)	Graph pointer network; HRL	20, 50, 250, 500, 1000-TSP: solution obtained better than Bello et al. (2016). 20-TSPTW: solution obtained better than OR-Tools, and ant colony optimisation.
	Ma et al. (1911)	Ptr-Net; REINFORCE & Critic baseline & Decompose policy	100, 150, 200, 500-Multi objective TSP: better than MOEA/D, NSGA-II, and MOGLS
DRL-based Iterative Method	Chen & Tian (2019)	Graph attention & PPO	100-CVPR: solution obtained better than Vaswani et al. (2017). 100-CVPRTW: better than multi-heuristics algorithms.
	Gao et al. (2020)	Transformer attention & REINFORCE	20, 50, 100-CVRP: solution obtained better than Nazari et al. (2018); Vaswani et al. (2017), OR Tools, and LKH3. Computing speed better than LKH3

1) SDVRP=Split delivery VRP; 2) OPVRP=Order picking VRP; 3) PCTSP= Prize collecting TSP; 4) Stochastic prize collecting TSP; 5) MOEA/D=Multi-objective evolutionary algorithm based on decomposition; 6) NSGA-II=Non-dominated sorting genetic algorithm II; 7) MOGLS=Multi-objective genetic local search

3.5 Concluding Remarks

This chapter has reviewed different research aspects and state-of-the-art ALD studies, including macroscopic evaluation, mesoscopic decision, and microscopic optimisation for ALD operations. Considering existing AV products are compactly designed to obtain greater flexibility, as well as mainly operated on pedestrian paths compared to conventional delivery vans, current models are increasingly incorporating stochastic issues that consider uncertain environmental elements, including changing demand patterns, time windows, and road networks.

The review has revealed that delivery cost and efficiency are the most two important factors that promote logistics operators to achieve the successful deployment of ALD, given that the operators follow safety regulations. The majority of existing studies employ the CA approach to evaluate models associated with deploying delivery services in the last-mile market. However, since CA does not directly solve VRPs, its evaluation precision is constrained by the substantial simplification of delivery routes in large problem instances. In light of this, Chapter 4 introduces a hybrid method that combines precise VRP solving with CA to assess ALD. This model accurately solves VRP in a specific region and subsequently maps the solution to other regions with similar road networks, thereby improving evaluation precision while ensuring computational efficiency.

No study in the reviewed literature determines customers' decision patterns in the ALD service selection. Considering that the principle of ALD service selection for customers is based on the compromise between cost and convenience, the competition between different autonomous delivery options can be constructed as a non-cooperative game model with different operational features involved. The reviewed studies further reflect that there is a lack of discussion regarding customers' bounded decision behaviours caused by incomplete information being given to their decision process. Hence, Chapter 5 has proposed a non-cooperative game model to formulate the service competition between MPLs and ADRs and designed an improved algorithm incorporating global information to enhance customers' decision rationality. Customers' fully rationality and bounded rationality decision patterns are formulated as Wardrop and Logit equilibrium models respectively, because the significant fitness of these two

models could capture the proposed operational features compared to other reviewed equilibrium models.

The VRP has been extensively used to optimise parcel distribution operations. Operation models are increasingly incorporating stochastic factors that consider uncertain environmental elements, including changing demand patterns, delivery capacity, battery consumption, and road network. The VRP in terms of AV applications can be classified into AV-only operation scenarios and AV-assisted operation scenarios. Considering that the AV-only delivery scenario requires superior adaptability to urban traffic conditions, and delivery capacity, and is applicable for modelling VRP that utilises medium-sized AVs. Therefore, Chapter 6 applies the AV-only VRP model to formulate route planning for MPLs. A novel algorithm based on RL has been devised to address the challenges posed by the exact approach, which often entails substantial computational time, and heuristics, which are prone to encountering local optima. These two methods are prominently employed in the majority of the literature reviewed. As an alternative option to MPL delivery, the reviewed literature suggests formulating ADR delivery operation as an AV-assisted VRP model, given that ADRs have restricted loading capacity and service radius. Chapter 7 develops such an operation model.

Chapter 4

Capability Evaluation for AV Fleets in Last-Mile Delivery

This chapter serves as the fundamental for Chapter 5 to 7. This chapter substantiated the superiority of the MPL and ADR delivery options over the conventional human-van delivery mode across three evaluation dimensions: operational costs, delivery time consumption, and gas emissions. The validation conducted in this chapter underscores the potential and feasibility of MPL and ADR delivery modes for future large-scale deployments. Specifically, this chapter focuses on assessing the delivery scenario that combines three options: MPLs, ADRs, and conventional human-van delivery across diverse geographic areas, spanning from suburban to urban settings. Additionally, the proposed capability evaluation places particular emphasis on the delivery to three distinct categories of destination: apartments, houses, and workplaces, where the property of each destination type shows different fitness in terms of the three proposed delivery options considering the physical accessibility through ramps and curbs, and the storey of customers located may condition the usage of a specific delivery option. This final reflects on different parcel volume that are assigned to different delivery options. Hence, this chapter further assess the performance of different parcel allocation policy.

An improved CA approach is introduced to evaluate the delivery scenario in terms of the proposed performance indicators. Compared to the conventional CA approach which relies on continuous density functions for estimating parameters and decision variables in both long-haul

and last-mile transportation, the proposed improved CA approach maintains the foundational concept of route estimation for long-haul transportation, while solving the exact route for a specific region followed by mapping the solved route to other regions within the same borough considering the homogeneity and variance of road network. This renders a sound compromise between accuracy and feasibility for the evaluation on a large scale. Additionally, it allows to capture the detailed time consumption for different delivery activities involved in a specific service option. Thus, the contribution of this chapter are as follows:

- This chapter provides a new paradigm to evaluate the deployment of autonomous delivery service in real-world geographic regions on a large scale, which enables logistics operators to configure the service based on the perspective of cost structure and delivery capability.
- The evaluation of autonomous delivery capabilities specifically takes into account the suitability of various service options for different destination facilities. By assigning distinct probabilities to specific delivery options based on the type of destination, this study investigates the influence of destination attributes on overall delivery performance.
- In comparison to the conventional CA approach commonly employed in existing research for macroscopic-level assessment of delivery capabilities, the proposed enhanced CA method allows for a superior balance between evaluation accuracy and implementation feasibility from a macro-to-meso perspective. Additionally, it enables the analysis of time consumption for various activities within the delivery process.

4.1 Scenario Definition for Delivery Options

Three delivery options (human-van, MPL, and ADR delivery) are incorporated into serving demands of a borough that can be divided into multiple regions. The delivery performance of the borough will be analysed in terms of delivery efficiency, total operations costs, and externalities emission. Figure 4.1 shows the operation of different delivery options in a single region within the borough, for the illustration purpose.

- **Human-van (HV) Delivery:** this delivery option delineates the prevailing scenario in

urban logistics. A delivery truck will initially transport customers' parcels of a borough from the carrier's Distribution Center (DC) to an Urban Consolidation Centre (UCC). Delivery van will then being employed to deliver parcels directly from UCC to the end customers. Each delivery van is responsible for serving a specific region within the borough based on the predetermined scheduling.

Upon the van's arrival near a customer's location/stopover i, the delivery staff have to spend time $T_i^{van_parking}$ locating a parking space to ensure compliance with traffic regulations. The time consumption of delivery staff for sorting parcels before handing over the parcel to the customer is represented as $T_i^{van_sorting}$. After completing the sorting process, the delivery staff disembarked from the van and proceeded on foot to hand over the parcels to the customer. Once the customer acknowledges receipt, the delivery staff return to the vehicle. The time expended for this process is denoted as $T_i^{van_handover}$. Upon completion of the service for the present customer, the delivery van proceeds to the subsequent customer's location, ultimately returning to the UCC after fulfilling all customers within the designated region.

- **MPL Delivery:** considering the restricted service range of MPLs, the long-haul transportation for parcels of MPL delivery scenario is the same as HV delivery where a delivery truck will first transport customers' parcels within a borough from the carrier's DC to a UCC. MPLs will then being deployed to deliver parcels from UCC to a set of predefined stopovers such that customers can access MPLs for parcel pickup. The selection of stopovers is based on customers' spatial distribution with the objective of minimising customers' walking distance for parcel pickup, which will be introduced in Appendix A.2. Each MPL is responsible for traversing a set of assigned stopovers for parcel handover of a specific region within the borough.

  Once the MPL's arrival near a stopover i, the MPL will spend time $T_i^{mpl_parking}$ locating a parking spot to ensure compliance with traffic regulations. Let $T_i^{mpl_handover}$ denotes the total time consumption of all customers who indicate this stopover i for parcel pick-up to unlock/operate compartments. Let $T_i^{mpl_walking}$ represents the total time consumption of customers walking from their locations to the MPL for parcel collection. Upon completion

of the service for the present stopover, the MPL proceeds to the subsequent stopovers, ultimately returning to the UCC after traversing all stopovers within the designated region.

- **ADR Delivery:** ADR are capable of delivering parcels to the ground floor of the designated destination, provided that the ramp or pathway leading to the building entrance is conducive to ADR operations. Compared to the HV service, the customers have to visit the ground floor for parcel pick-up, while the walking distance yielded is shorter than that of the MPL service.

  The proposed ADR delivery entails a collaborative effort between the ADR and delivery personnel alongside a van, attributed to the limited battery capacity U of the ADR. This limitation necessitates the ADR to be transported by a van for medium to long-distance journeys within the city while the ADR concludes the final stage of the delivery. On the other hand, the process of loading parcels into the ADR's compartment requires human intervention. Moreover, in instances where the ADR's battery power reaches its maximum limit U, delivery tasks can be seamlessly taken over by the delivery staff to ensure the fulfilment of remaining demands.

  In this mode, a delivery truck will transport customers' parcels within a borough from DC to UCC. The delivery van then departs from UCC to deliver parcels to end customers. Each delivery van is on board an ADR and is responsible for serving a specific region within the borough based on the predetermined scheduling.

  Once the delivery van arrives at the customer's location i, the van will spend time $T_i^{adr_parking}$ locating a parking spot. Subsequently, the delivery staff will load parcels onto the ADR before it commences the final leg of the delivery journey. The duration required for loading parcels is represented as $T_i^{adr_loading}$. Each ADR possesses a capacity limitation q, enabling it to serve a maximum of 3 consecutive customers (considering the compartment size of the prevailing ADR product in the market) before returning to the delivery van to reload parcels for the subsequent delivery round (i.e., $0 \leq q \leq 3$). Upon the ADR's arrival, each customer will dedicate time $T_i^{adr_handover}$ to operate or unlock

the compartment. Once the demand of customer i has been satisfied, the ADR has two options for travelling to the subsequent customers of the same delivery round:

1) **Return to van:** the ADR returns to the van after serving the current customer i, and the van takes the ADR to the next customer j. This option holds if the time consumption t_{ij}^{adr} of ADR's individual travelling from i to j is greater than the time consumption t_{ij}^{van} of ADR onboard the van travelling from i to j plus the parking time $T_j^{adr_parking}$ of the delivery van and the parcel loading time $T_j^{adr_loading}$ at customer location j, i.e., $t_{ij}^{adr} > t_{ij}^{van} + T_j^{adr_parking} + T_j^{adr_loading}$.

2) **Self-navigation:** if ADR is about to serve 2 consecutive customers, the ADR directly navigate to the next customer j after serving the current customer i. Simultaneously, the delivery van will proceed to customer j to rendezvous ADR. This option holds when $t_{ij}^{adr} \leq t_{ij}^{van} + T_j^{adr_parking} + T_j^{adr_loading}$. If ADR is about to serve 3 consecutive customers, the ADR will navigate to the second customer j after serving the current customer i, followed by navigating to the last customer $j+1$ of the same delivery round. Under this circumstance, the accompanied van will directly visit customer $j+1$ to rendezvous ADR after releasing the robot at the current customer i. This option holds when $t_{ij}^{adr} + t_{jj+1}^{adr} \leq t_{ij}^{van} + t_{jj+1}^{van} + T_j^{adr_parking} + T_{j+1}^{adr_parking} + T_j^{adr_loading} + T_{j+1}^{adr_loading}$. Hence, the self-navigation of ADR reduces the round trip of the robot travelling between customers and the van.

We made the following assumptions when formulating the delivery scenario:

- **Assumption 1:** each customer only has one parcel for collection. The size of parcels can fit any compartment of van, MPL, and ADR.
- **Assumption 2:** each delivery truck and van equips a driver/delivery staff to drive the vehicle. Considering the existing vehicle products in the market, it is assumed that the delivery truck and van are fuel-powered while the MPL and ADR are electric-powered.
- **Assumption 3:** the average time consumption of finding available parking space within different boroughs are set different, because the parking space occupations among different boroughs varying.

- **Assumption 4:** considering the operational characteristics among different delivery options and the corresponding service time at different customer/stopover i varied, the value of $T_i^{van_parking}$, $T_i^{mpl_parking}$, $T_i^{adr_parking}$, $T_i^{van_sorting}$, $T_i^{van_handover}$, $T_i^{mpl_walking}$, $T_i^{mpl_handover}$, $T_i^{adr_handover}$, $T_i^{adr_loading}$ are set different, which follows normal distribution (will be specified in Section 4.3).

- **Assumption 5:** for ADR delivery option, the delivery van can release and collect ADR at any customer location.

- **Assumption 6:** for all proposed delivery modes, vehicles (e.g., van, MPLs) have sufficient capacity to deliver packages once they are allocated to meet the demand in a particular region within the borough. Otherwise, unmet demand in the region will be assigned for additional vehicles to fulfil.

- **Assumption 7:** given that the parcel delivery within a borough can be subdivided into multiple regions, and the size and demands of an individual region within a borough are not significantly large, we make the assumption that each delivery option possesses a single fleet allocated for a region to meet all demands, with the aim of cost-saving. The capacity of HV, MPL, and ADR delivery option within a region can be denoted as Cap_{hv}, Cap_{mpl}, and Cap_{adr}, respectively.

4.2 Solution Method

In this section, we adjusted the conventional CA formulation introduced by Daganzo (1984a,b). As illustrated in Section 4.1 and Figure 4.1, logistics operators must access all end customer locations and optimise total travel distance to limit total operating costs which is considered as VRP. Therefore, CA approach is capable to be used for estimating the route length and tour time of the VRP, with a sound compromise between accuracy and feasibility. All distances are in L_1 metric (Manhattan distance).

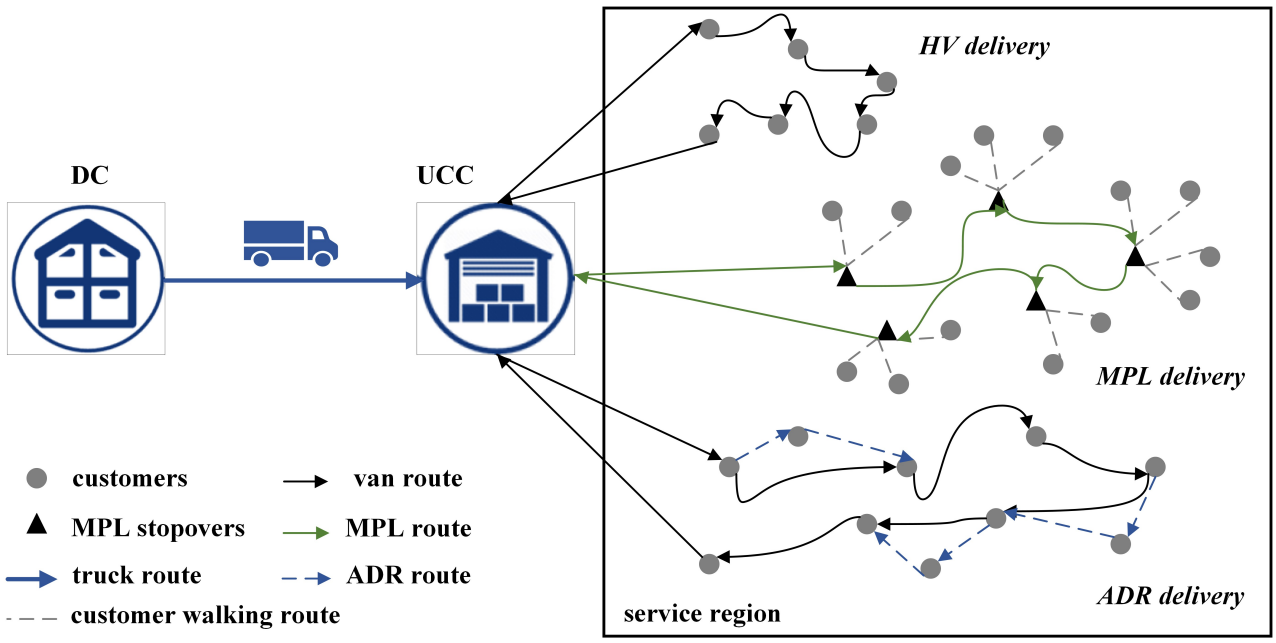


Figure 4.1: Illustration of different delivery options for serving a single region

4.2.1 Conventional Continuous Approximation Framework

A conventional CA-based routing framework entails a single echelon network wherein a carrier offers delivery services for N customers within a rectangular service area of dimensions $A = l_x l_y$. The logistics company deploys a fleet of delivery vehicles, which operates from a DC outside the service region A. The DC is located at x_{DC}, y_{DC} relative to the central point of the service region. A typical delivery route involves the transportation of parcels from the DC to the initial customer-stop, and conversely from the last customer-stop back to the DC, constituting the long-haul journey ρ. Additionally, it encompasses the last-mile journey in which the delivery vehicle traverses between the first and last customer stops. Figure 4.2 presents the above CA-based routing framework.

Conventional CA framework estimates the corresponding tour length based on Equation 4.1 as the sum of the back-and-forth long-haul distance ρ and the last-mile distance (Pahwa & Jaller, 2022). The tour time is estimated based on Equation 4.2 as the sum of the time consumption of loading packages at DC (t_{LP} per package), long-haul transportation time of the vehicle (t_{LH}), last-mile travelling time of the delivery vehicle (t_{LM}), and the time spent related to auxiliary operations when delivering packages (e.g., the total time spent on parking vehicle, sorting parcel, handover parcel) at customer-stops (t_{DP} per customer-stop), represented by each

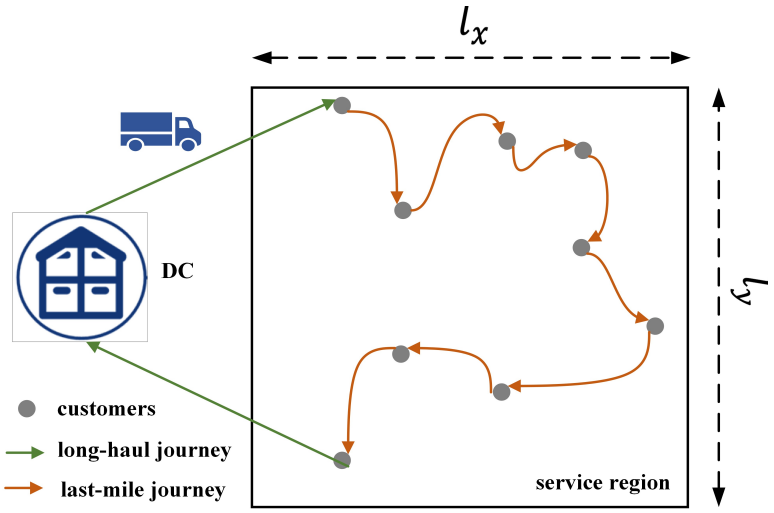


Figure 4.2: Illustration of conventional CA framework

term in the equation, respectively.

$$L = 2\rho + \frac{0.73C}{\sqrt{\delta}} \tag{4.1}$$

$$T = Ct_{LP} + 2t_{LH} + t_{LM} + Ct_{DP} \tag{4.2}$$

$$= Ct_{LP} + 2t_{LH} + \frac{0.73C}{v\phi\sqrt{\delta}} + Ct_{DP} \tag{4.3}$$

where C represents the number of customer-stops of the delivery process, v represent the velocity of the delivery vehicle , ϕ denotes the traffic congestion factor, and δ denotes the customer-stop density ($\delta = N/A$).

Since the distance between DC and the centre of the service region will affect the estimation of the tour length and delivery time of the long-haul journey, the calculation for the tour length of long-haul journey ρ (Equation 4.4) and the corresponding delivery time t_{LH} (Equation 4.5) can be further expanded by considering the different cases of relative locations between DC and the service region, free-flow vehicle speed, and the traffic condition. The proof of the formula can reference Pahwa & Jaller (2022).

$$\rho = \begin{cases} |x_{DC}| + |y_{DC}| & \text{if } |x_{DC}| \text{ and } |y_{DC}| \geq \frac{\sqrt{A}}{2} \\ |x_{DC}| + \frac{y_{DC}^2}{\sqrt{A}} + \frac{\sqrt{A}}{4} & \text{if } |x_{DC}| \geq \frac{\sqrt{A}}{2} \text{ and } |y_{DC}| < \frac{\sqrt{A}}{2} \\ \frac{x_{DC}^2}{\sqrt{A}} + \frac{\sqrt{A}}{4} + |y_{DC}| & \text{if } |x_{DC}| < \frac{\sqrt{A}}{2} \text{ and } |y_{DC}| \geq \frac{\sqrt{A}}{2} \\ \frac{x_{DC}^2 + y_{DC}^2}{\sqrt{A}} + \frac{\sqrt{A}}{2} & \text{if } |x_{DC}| \text{ and } |y_{DC}| < \frac{\sqrt{A}}{2} \end{cases} \tag{4.4}$$

$$t_{LH} = \frac{1}{\phi} \begin{cases} \frac{\rho}{v_{out}} + \sqrt{A}(\frac{1}{v_{in}} - \frac{1}{v_{out}}) & \text{if } |x_{DC}| \text{ and } |y_{DC}| \geq \frac{\sqrt{A}}{2} \\ \frac{|x_{DC}|}{v_{out}} + \frac{(\frac{y_{DC}^2}{\sqrt{A}} + \frac{\sqrt{A}}{4})}{v_{in}} + \frac{\sqrt{A}}{2}(\frac{1}{v_{in}} - \frac{1}{v_{out}}) & \text{if } |x_{DC}| \geq \frac{\sqrt{A}}{2} \text{ and } |y_{DC}| < \frac{\sqrt{A}}{2} \\ \frac{(\frac{x_{DC}^2}{\sqrt{A}} + \frac{\sqrt{A}}{4})}{v_{in}} + \frac{|y_{DC}|}{v_{out}} + \frac{\sqrt{A}}{2}(\frac{1}{v_{in}} - \frac{1}{v_{out}}) & \text{if } |x_{DC}| < \frac{\sqrt{A}}{2} \text{ and } |y_{DC}| \geq \frac{\sqrt{A}}{2} \\ \frac{\rho}{v_{in}} & \text{if } |x_{DC}| \text{ and } |y_{DC}| < \frac{\sqrt{A}}{2} \end{cases} \tag{4.5}$$

where v_{out} and v_{in} are free-flow vehicle speed when the vehicle travelling outside and inside the service region with traffic congestion factor ϕ, since the vehicle will initially depart from the DC where outside the service region to the inner service region.

4.2.2 Modelling Cooperative Deliveries Using Adjusted Continuous Approximation

The model developed above only supports a single echelon structure of last-mile operations, wherein the logistics operators deliver parcels from a single DC to the customers of the service region. Considering the service options defined in Section 4.1 and Figure 4.1 wherein the delivery network is a two-echelon structure, the tour of each delivery option k ($k \in \{HV, MPL, ADR\}$) includes two parts: i) the long-haul journey ρ ($\rho = \rho_1 + \rho_2 + \rho_2^{'}$), which consists of the trip of a truck travelling from DC to the UCC (denoted as ρ_1), the journey from the UCC to the first customer-stop (denoted as ρ_2), and likewise from the last customer-stop back to the UCC (denoted as $\rho_2^{'}$); ii) the last-mile delivery, which involves the journey between the first and the last customer-stops. The DC is located at x_{DC}, y_{DC} relative to the centre

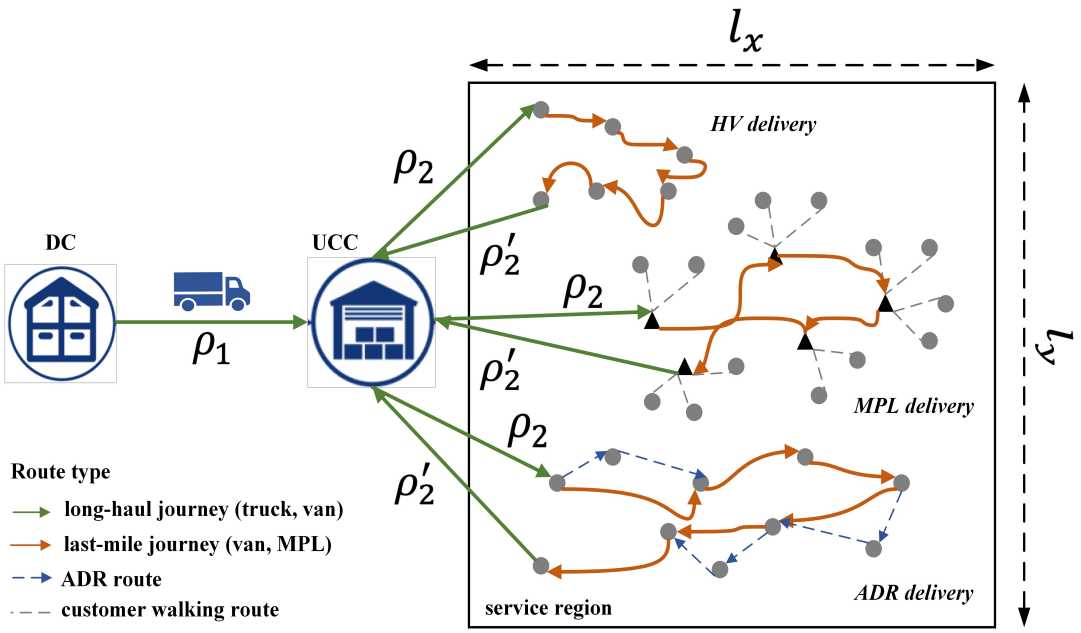


Figure 4.3: Illustration of different route types of a region in CA framework

of the service region, while the UCC is located at x_{UCC}, y_{UCC} relative to the centre of the service region. Figure 4.3 presents the above two types of delivery routes with different delivery options.

Therefore, we adjusted the estimation for the tour length and delivery time of each delivery option k based on the logic of conventional CA framework introduced in Section 4.2.1.

Estimation for the Long-Haul Journey

Equation 4.6 illustrates the calculation for estimating the tour length $L^k_{long-haul}$ pertaining to the long-haul journey with respect to the delivery option k. The first component of the equation signifies the round-trip undertaken by the delivery truck as it travels between the DC and the UCC. The second component of the equation represents the journey of the delivery vehicle (e.g., van, MPL) associated with the specific delivery option k, from the UCC to the first customer-stop. The third component of the equation signifies the journey between the UCC and the final customer-stop.

$$L^k_{long-haul} = 2\rho_1 + \rho_2 + \rho_2^{'} = 2(|x_{UCC} - x_{DC}| + |y_{UCC} - y_{DC}|) + \rho_2 + \rho_2^{'} \tag{4.6}$$

Equation 4.7 can be used to calculate the delivery duration $T^k_{long-haul}$ of the long-haul journey for delivery option k. The first component of the equation represents the sum of the preprocessing time (e.g., loading packages onto the delivery truck) associated with delivery option k. This occurs prior to the truck's journey from the DC to the UCC, given that the delivery truck consolidates all parcels of the same service region to facilitate batched transportation. The second element of the equation denotes the preprocessing time specific to the delivery option k, occurring prior to the departure of the delivery vehicle (e.g., van, MPL) from the UCC performing the last-mile delivery. The third term of the equation symbolises the back-and-forth transit time of the delivery truck for trip ρ_1. The last component of the equation denotes the total time consumption associated with the delivery option k for trips ρ_2 and $\rho_2^{'}$, with the consideration of the locational relationship involving the DC, UCC, and the service region. Additionally, it considers the varying vehicle velocities of the respective delivery option k when traversing areas outside and within the service region.

For Equation 3.8, it first considers the location relationship between DC and the dimension of the service region: if $|x_{DC}|$ or $|y_{DC}| > \frac{\sqrt{A}}{2}$ satisfies, then 0 will be applied to the equation, as denoted by the first brace in Equation 3.8. If $|x_{DC}|$ and $|y_{DC}| \leq \frac{\sqrt{A}}{2}$ satisfies, then the location relationship between UCC and the dimension of the service region has to be considered, as denoted by the second brace in Equation 3.8.

$$T^k_{long-haul} = \sum_k C_k t_{LP} + C_k t_{LP} + \frac{2\rho_1}{v_{truck}\phi} + 2\frac{1}{\phi}(\frac{1}{v^k_{in}} - \frac{1}{v^k_{out}})\times \tag{4.7}$$

$$\begin{cases} 0 \\ \begin{cases} (\frac{\sqrt{A}}{2} - x_{DC}) & \text{if } x_{UCC} > \frac{\sqrt{A}}{2} \text{ and } y_{UCC} \in [-\frac{\sqrt{A}}{2}, \frac{\sqrt{A}}{2}] \\ (\frac{\sqrt{A}}{2} + min(-x_{DC}, y_{DC})) & \text{if } x_{UCC} > \frac{\sqrt{A}}{2} \text{ and } y_{UCC} < -\frac{\sqrt{A}}{2} \\ (\frac{\sqrt{A}}{2} + y_{DC}) & \text{if } x_{UCC} \in [-\frac{\sqrt{A}}{2}, \frac{\sqrt{A}}{2}] \text{ and } y_{UCC} < -\frac{\sqrt{A}}{2} \\ (\frac{\sqrt{A}}{2} + min(x_{DC}, y_{DC})) & \text{if } x_{UCC} < -\frac{\sqrt{A}}{2} \text{ and } y_{UCC} < -\frac{\sqrt{A}}{2} \\ (\frac{\sqrt{A}}{2} + x_{DC}) & \text{if } x_{UCC} < -\frac{\sqrt{A}}{2} \text{ and } y_{UCC} \in [-\frac{\sqrt{A}}{2}, \frac{\sqrt{A}}{2}] \\ (\frac{\sqrt{A}}{2} + min(x_{DC}, -y_{DC})) & \text{if } x_{UCC} < -\frac{\sqrt{A}}{2} \text{ and } y_{UCC} > \frac{\sqrt{A}}{2} \\ (\frac{\sqrt{A}}{2} - y_{DC}) & \text{if } x_{UCC} \in [-\frac{\sqrt{A}}{2}, \frac{\sqrt{A}}{2}] \text{ and } y_{UCC} > \frac{\sqrt{A}}{2} \\ (\frac{\sqrt{A}}{2} + min(-x_{DC}, -y_{DC})) & \text{if } x_{UCC} > \frac{\sqrt{A}}{2} \text{ and } y_{UCC} > \frac{\sqrt{A}}{2} \end{cases} \end{cases} \tag{4.8}$$

where C_k represent the number of customer-stops of the delivery option k. t_{LP} represents the time required to load an individual package either at the DC or the UCC. v_{truck} denotes the velocity of the delivery truck travelling between the DC and the UCC. v^k_{out} and v^k_{in} are free-flow vehicle speed (e.g., van, MPL) of the delivery option k when traversing areas outside and within the service region. ϕ is the traffic congestion factor.

Estimation for the Last-Mile Journey

Though Equation 4.1 and 4.2 provide fundamental and simplified formulas to estimate the tour length and completion time of the delivery without solving the exact VRP model, the existing approximation of the last-mile journey may result in a relatively significant error because it relies on significant assumptions about the coefficient (e.g., 0.73) of the VRP solution and the demand density δ. Additionally, Equation 4.2 assumes that the time consumption of auxiliary operations (e.g., parking vehicle, sorting parcels) at different customer locations/stopovers is the same which does not fit the real-world operation. Further, the road network of different regions

varied, which may increase the variance between estimated results and real-world operation. Such variance will be amplified in a larger network's tour length and time estimation.

Considering this study focuses on the estimation for different boroughs that can be subdivided into multiple regions, we introduce an adjusted estimation approach for the last-mile journey to address the aforementioned constraints, aligning it more closely with the specific characteristics of this study. In contrast to the conventional CA framework, the proposed method initially calculates the tour length and delivery time of the last-mile journey for an individual region within a borough by solving an exact VRP model (denote this region as solved region). The computed tour length and delivery time for this solved single region are then extrapolated to approximate the corresponding values for the remaining regions (denoted as unsolved regions) within the same borough. This extrapolation process considers variations in road network density and demand spatial distribution across different regions within the same borough since the road network density and demand spatial distribution are proven to have significant effects on delivery performances from an urban planning point of view (Cui et al., 2022; Danchuk et al., 2019; Llorca & Moeckel, 2020; Ghelichi et al., 2024).

Figure 4.4 shows the mechanism of the adjusted CA framework to estimate the tour length and delivery time of a borough. The estimation of the tour length and delivery time for the long-haul journey (indicated by the green line) encompassing all regions segmented within a borough adheres to Equations 4.6 and 4.7, respectively. While the estimation of the tour length and delivery time of the last-mile journey (denoted as the orange line) includes two steps: i) obtaining the exact VRP solution of the solved region (e.g., region 1); ii) mapping the exact VRP solution from the solved region to other unsolved regions (e.g., regions 2 to 4) within a borough through mapping functions (Equation 4.14 and 4.15). Therefore, the tour length and the delivery time of the delivery option k for the solved region follow the Equations 4.9 and 4.10.

$$L^k_{total} = L^k_{long-haul} + L^k_{last-mile} \qquad \forall k \in \{HV, MPL, ADR\} \tag{4.9}$$

$$T^k_{total} = T^k_{long-haul} + T^k_{last-mile} + T^k_{DP} \qquad \forall k \in \{HV, MPL, ADR\} \tag{4.10}$$

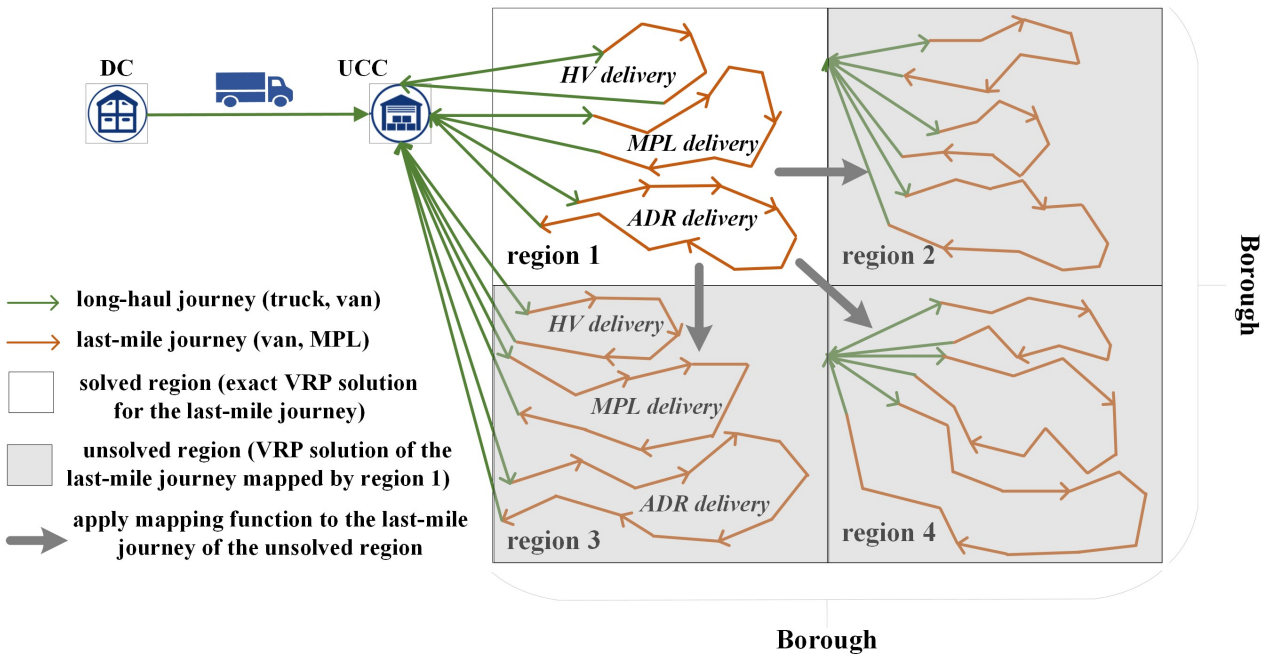


Figure 4.4: Illustration of the adjusted CA framework

$$T_{DP}^{k} = \begin{cases} \sum_{i=1}^{C_k}(T_i^{van_parking} + T_i^{van_sorting} + T_i^{van_handover}) & \forall i \in \{1, \cdots, C_k\} \quad k = \text{HV} \\ \sum_{i=1}^{C_k}(T_i^{mpl_parking} + T_i^{mpl_handover} + T_i^{mpl_walking}) & \forall i \in \{1, \cdots, C_k\} \quad k = \text{MPL} \\ \sum_{i=1}^{C_k}(T_i^{adr_parking} + T_i^{adr_loading} + T_i^{adr_handover}) & \forall i \in \{1, \cdots, C_k\} \quad k = \text{ADR} \end{cases} \tag{4.11}$$

where $L_{last-mile}^{k}$ and $T_{last-mile}^{k}$ are the exact route length and delivery time of the delivery option k solved by the VRP solver. T_{DP}^{k} denotes the total time consumption of the delivery option k associate with the auxiliary operations (e.g., parking vehicle, sorting parcel, handover parcel mentioned in Section 4.1) when delivering packages for all customers/stopovers. Therefore, the total tour length and delivery time of all delivery options within the solved region is represented as Equation 4.12 and 4.13, respectively:

$$L_{region} = \sum_{k} L_{total}^{k} \qquad \forall k \in \{HV, MPL, ADR\} \tag{4.12}$$

$$T_{region} = \sum_{k} T^{k}_{total} \qquad \forall k \in \{HV, MPL, ADR\} \tag{4.13}$$

The approximation for the tour length $\hat{L}^{k}_{last-mile}$, delivery time $\hat{T}^{k}_{last-mile}$, and the auxiliary operation time consumption $\hat{T}^{k}_{DP}$ of the last-mile journey in terms of the delivery option k within an unsolved region (e.g., the region either from 2 to 4) follows the mapping function 4.14, 4.15, and 4.16. The mapping functions consider the joint ratio of the road network density and customer-stop size between the solved and unsolved regions. Relevant literature that considers the effect and significance of demand spatial distribution and road network density within continuous approximation functions includes:Lei & Ouyang (2018); Ansari et al. (2018); Silva et al. (2022); Peng et al. (2014).

$$\hat{L}^{k}_{last-mile} = \frac{\hat{\mu}_k \hat{\psi}}{\mu_k \psi} L^{k}_{last-mile} \qquad \forall k \in \{HV, MPL, ADR\} \tag{4.14}$$

$$\hat{T}^{k}_{last-mile} = \frac{\hat{\mu}_k \hat{\psi}}{\mu_k \psi} T^{k}_{last-mile} \qquad \forall k \in \{HV, MPL, ADR\} \tag{4.15}$$

$$\hat{T}^{k}_{DP} = \frac{\hat{\mu}_k \hat{\psi}}{\mu_k \psi} T^{k}_{DP} \qquad \forall k \in \{HV, MPL, ADR\} \tag{4.16}$$

where $\frac{\hat{\mu}_k}{\mu_k}$ denotes the ratio of customer-stop size between the unsolved and the solved region in terms of delivery option k. The calculation for the customer-stop size follows Equation 4.17 and 4.18, which mimics the estimation of the last-mile journey computed by Equation 4.1. $\frac{\hat{\psi}}{\psi}$ is the ratio of the road network density between the unsolved and the solved region, in which the calculation of the road network density follows Equation 4.19:

$$\mu_k = \frac{C_k}{\sqrt{\delta_k}} \qquad \forall k \in \{HV, MPL, ADR\} \tag{4.17}$$

$$\delta_k = \frac{N_k}{A} \qquad \forall k \in \{HV, MPL, ADR\} \tag{4.18}$$

$$\psi = \frac{R}{A} \tag{4.19}$$

where δ_k is the demand density of the delivery option k, N_k denotes the total number of customers served by the delivery option k, A denotes the size of the service region, and R represents the total number of vehicle roads within the service region.

Correspondingly, the total tour length $\hat{L}^k_{total}$ and the delivery time $\hat{T}^k_{total}$ of the delivery option k within an unsolved region follows the Equations 4.20 and 4.21.

$$\hat{L}^k_{total} = \hat{L}^k_{long-haul} + \hat{L}^k_{last-mile} \qquad \forall k \in \{HV, MPL, ADR\} \tag{4.20}$$

$$\hat{T}^k_{total} = \hat{T}^k_{long-haul} + \hat{T}^k_{last-mile} + \hat{T}^k_{DP} \qquad \forall k \in \{HV, MPL, ADR\} \tag{4.21}$$

where the calculation for $\hat{L}^k_{long-haul}$ and $\hat{T}^k_{long-haul}$ follows Equation 4.6 and 4.7, respectively. By computing $\hat{L}^k_{total}$ and $\hat{T}^k_{total}$, the total rout length and delivery time of the unsolved region could be obtained through Equation 4.22 and 4.23.

$$\hat{L}_{region} = \sum_k \hat{L}^k_{total} \qquad \forall k \in \{HV, MPL, ADR\} \tag{4.22}$$

$$\hat{T}_{region} = \sum_k \hat{T}^k_{total} \qquad \forall k \in \{HV, MPL, ADR\} \tag{4.23}$$

The estimation formulas presented in Section 4.2.2 are the basis of all the delivery service modelling done in the study. The following section will focus on developing evaluation matrix based on cost and efficiency perspectives, in terms of the proposed delivery options.

4.2.3 Evaluation Matrix Modelling

In accordance with the delivery scenario delineated in Section 4.1 and the introduced methodology for estimating tour length and delivery time in Section 4.2.2, Table 4.1 provides a comprehensive compilation of distinct performance indicators. These indicators will be computed for each delivery option by utilising and decomposing the estimation outcomes for tour length and delivery time.

In accordance with Assumptions 2 and 7, the driver cost for each delivery option k within a

Table 4.1: Different performance indicators for evaluation

Performance Indicators	**HV**	**MPL**	**ADR**
Operations Costs (£)			
Driver cost Q^k_{driver}	√	-	√
Vehicle operating cost $Q^k_{operating}$	√	√	√
Fuel cost Q^k_{fuel}	√	-	√
Electricity cost $Q^k_{electricity}$	-	√	√
Operation Efficiency (*min*)			
Total preprocessing time at DC $T^{DC}_{preprocessing}$	√	√	√
Total preprocessing time at UCC $T^{UCC}_{preprocessing}$	√	√	√
Total parking time of the HV delivery $T^{HV}_{parking}$	√	-	-
Total parking time of the MPL delivery $T^{MPL}_{parking}$	-	√	-
Total parking time of the ADR delivery $T^{ADR}_{parking}$	-	-	√
Total walking time of customers $T^{MPL}_{walking}$	-	√	-
Total parcel sorting time of HV delivery $T^{HV}_{sorting}$	√	-	-
Total parcel loading time of ADR delivery $T^{ADR}_{loading}$	-	-	√
Total parcel handover time of HV delivery $T^{HV}_{handover}$	√	-	-
Total parcel handover time of MPL delivery $T^{MPL}_{handover}$	-	√	-
Total parcel handover time of ADR delivery $T^{ADR}_{handover}$	-	-	√
Total time consumption of van travelling T^{HV}_{travel}	√	-	-
Total time consumption of MPL travelling T^{MPL}_{travel}	-	√	-
Total time consumption of ADR travelling T^{ADR}_{travel}	-	-	√
Externalities Emission (*g*)			
CO_2 emission $Q^k_{CO_2}$	√	-	√
CO emission Q^k_{CO}	√	-	√
NO_X emission $Q^k_{NO_X}$	√	-	√
PM emission Q^k_{PM}	√	-	√

Notes: 1) the driver cost comes from the delivery van requiring staff to drive the vehicle. 2) vehicle maintenance costs are applicable for both fuel-powered vehicles (e.g., van, truck) and electric-powered vehicles (e.g., MPL, ADR). 3) the fuel cost is yielded by the delivery truck or van. 4) the electricity cost comes from the electricity consumption of recharging MPL and ADR.

borough is contingent upon the total number of regions that can be subdivided from a singular borough (which equals the total number of drivers invested in the delivery), the hourly salary of the driver, and the total working hours of the delivery, as specified by Equation 4.24.

$$Q^k_{driver} = \begin{cases} \sum_{o=1}^{O} c_{driver} \cdot (\frac{2\rho_1}{v_{truck}\phi} + T^k_{last-mile}) & k = \text{HV} \\ \sum_{o=1}^{O} c_{driver} \cdot \frac{2\rho_1}{v_{truck}\phi} & k = \text{MPL} \\ \sum_{o=1}^{O} c_{driver} \cdot (\frac{2\rho_1}{v_{truck}\phi} + T^k_{last-mile}) & k = \text{ADR} \end{cases} \tag{4.24}$$

where $O(O = \left[\frac{U}{A}\right]^{+})$ represents the total number of regions partitioned by a borough with the size of U (the value of O will be rounded up to guarantee the whole borough has been fully covered). c_{driver} denotes the hourly salary of a single driver.

The vehicle operating cost of delivery option k is contingent upon the type of vehicle utilised by the delivery option k and the respective mileage covered by the vehicle during the delivery process (as shown in Figure 4.3). The operating costs of each vehicle type consider the annualised purchase cost, depreciation, tire replacement, insurance, and maintenance and repair costs. The total vehicle operating cost of delivery option k of the borough is calculated by Equation 4.25.

$$Q^{k}_{operating} = \frac{2\rho_1}{c_{truck}} + \begin{cases} \sum_{o=1}^{O} \left(\frac{\rho_2 + \rho_2^{'}}{c_{van}} + \frac{L^{HV}_{last-mile}}{c_{van}} \right) & k = \text{HV} \\ \sum_{o=1}^{O} \left(\frac{\rho_2 + \rho_2^{'}}{c_{mpl}} + \frac{L^{MPL}_{last-mile}}{c_{mpl}} \right) & k = \text{MPL} \\ \sum_{o=1}^{O} \left(\frac{\rho_2 + \rho_2^{'}}{c_{van}} + \frac{L^{ADR}_{last-mile}}{c_{adr} + c_{van}} \right) & k = \text{ADR} \end{cases} \tag{4.25}$$

where c_{truck}, c_{van} c_{mpl}, and c_{adr} denote the unit operating cost per vehicle mileage of the truck, van, MPL, and ADR, respectively.

The delivery truck and van incur fuel consumption. Hence, the fuel cost calculation for the proposed delivery option follows Equation 4.26, wherein the cost depends upon the vehicle mileage travelled. It should be noted that the last-mile journey of MPL delivery will not yield the fuel cost.

$$Q^{k}_{fuel} = 2\rho_1 f_{truck} + \begin{cases} \sum_{o=1}^{O} \left(\rho_2 + \rho_2^{'} + L^{HV}_{last-mile} \right) f_{van} & k = \text{HV} \\ 0 & k = \text{MPL} \\ \sum_{o=1}^{O} \left(\rho_2 + \rho_2^{'} + L^{ADR}_{last-mile} \right) f_{van} & k = \text{ADR} \end{cases} \tag{4.26}$$

where f_{truck} and f_{van} denote the unit cost of truck and van fuel consumption per mileage, respectively.

The MPL and ADR delivery options resulting in electricity consumption as the vehicles are electrically powered and necessitate battery charging. The calculation for the electricity cost of the proposed delivery options follows Equation 4.27

$$Q^k_{electricity} = \begin{cases} 0 & k = \text{HV} \\ \sum_{o=1}^{O} \left(\rho_2 + \rho_2^{'} + L^{MPL}_{last-mile}\right) e_{mpl} & k = \text{MPL} \\ \sum_{o=1}^{O} T^{ADR}_{travel} v^{adr}_{in} e_{adr} & k = \text{ADR} \end{cases} \tag{4.27}$$

where e_{mpl} and e_{adr} denote the unit cost of MPL and ADR electricity consumption per mileage, respectively. v^{adr}_{in} represents the ADR velocity within the service region.

The indicators within the operation efficiency class in Table 4.1 can be calculated by breaking down Equations 4.7, 4.10, and 4.11, given that the summation of these indicators constitutes the overall delivery time for the respective delivery option. The calculation for different indicators within the operation efficiency class in Table 4.1 are summarised as follows:

- Total preprocessing time at DC: $T^{DC}_{preprocessing} = \sum_k C_k t_{LP} \qquad \forall k \in \{HV, MPL, ADR\}$
- Total preprocessing time at UCC: $T^{UCC}_{preprocessing} = C_k t_{LP} \qquad \forall k \in \{HV, MPL, ADR\}$
- Total parking time of the HV delivery: $T^{HV}_{parking} = \sum_{o=1}^{O} \sum_{i=1}^{C_k} T_i^{van_parking} \qquad \forall i \in \{1, \cdots, C_k\}$
- Total parking time of the MPL delivery: $T^{MPL}_{parking} = \sum_{o=1}^{O} \sum_{i=1}^{C_k} T_i^{mpl_parking} \qquad \forall i \in \{1, \cdots, C_k\}$
- Total parking time of the ADR delivery: $T^{ADR}_{parking} = \sum_{o=1}^{O} \sum_{i=1}^{C_k} T_i^{adr_parking} \qquad \forall i \in \{1, \cdots, C_k\}$
- Total walking time of customer: $T^{MPL}_{walking} = \sum_{o=1}^{O} \sum_{i=1}^{C_k} T_i^{mpl_walking} \qquad \forall i \in \{1, \cdots, C_k\}$
- Total parcel sorting time of HV delivery: $T^{HV}_{sorting} = \sum_{o=1}^{O} \sum_{i=1}^{C_k} T_i^{van_sorting} \qquad \forall i \in \{1, \cdots, C_k\}$
- Total parcel loading time of ADR delivery: $T^{ADR}_{loading} = \sum_{o=1}^{O} \sum_{i=1}^{C_k} T_i^{adr_loading} \qquad \forall i \in \{1, \cdots, C_k\}$

- Total parcel handover time of HV delivery: $T^{HV}_{handover} = \sum_{o=1}^{O} \sum_{i=1}^{C_k} T_i^{van_handover} \quad \forall i \in \{1, \cdots, C_k\}$
- Total parcel handover time of MPL delivery: $T^{MPL}_{handover} = \sum_{o=1}^{O} \sum_{i=1}^{C_k} T_i^{mpl_handover} \quad \forall i \in \{1, \cdots, C_k\}$
- Total parcel handover time of ADR delivery: $T^{ADR}_{handover} = \sum_{o=1}^{O} \sum_{i=1}^{C_k} T_i^{adr_handover} \quad \forall i \in \{1, \cdots, C_k\}$
- Total time consumption of van travelling $T^{HV}_{travel} = \sum_{o=1}^{O} T^{HV}_{last-mile}$
- Total time consumption of MPL travelling $T^{MPL}_{travel} = \sum_{o=1}^{O} T^{MPL}_{last-mile}$
- Total time consumption of ADR travelling $T^{ADR}_{travel} = \sum_{o=1}^{O} T^{ADR}_{last-mile}$

The externalities emissions are yielded by the delivery truck and van, where the calculation of the emission follows Equation 4.28

$$Q^k_{gas} = 2\rho_1 m^{truck}_{gas} + \begin{cases} \sum_{o=1}^{O} \left(\rho_2 + \rho_2^{'} + L^{HV}_{last-mile}\right) m^{van}_{gas} & k = \text{HV} \\ 0 & k = \text{MPL} \\ \sum_{o=1}^{O} \left(\rho_2 + \rho_2^{'} + L^{ADR}_{last-mile}\right) m^{van}_{gas} & k = \text{ADR} \end{cases} \tag{4.28}$$

where m^{truck}_{gas} and m^{van}_{gas} represent the unit gas emission per mileage ($gas \in \{CO_2, CO, NO_X, PM\}$) of the truck and van, respectively.

4.2.4 Route Modelling for a Single Region

As outlined in Section 4.2.2, the estimation of the tour length and delivery time for the last-mile journey within the unsolved region is derived by extrapolating from the precise tour length and delivery time of the solved region. Consequently, we introduce route planning models for different delivery options, enabling the acquisition of an exact solution for the last-mile journey through the resolution of the route planning model of the solved region. Appendix A presents the route model of the proposed delivery options. The route planning models are solved by the commercial solver Gurobi 9.5.2 in the Apple M1 processor with 16GB RAM.

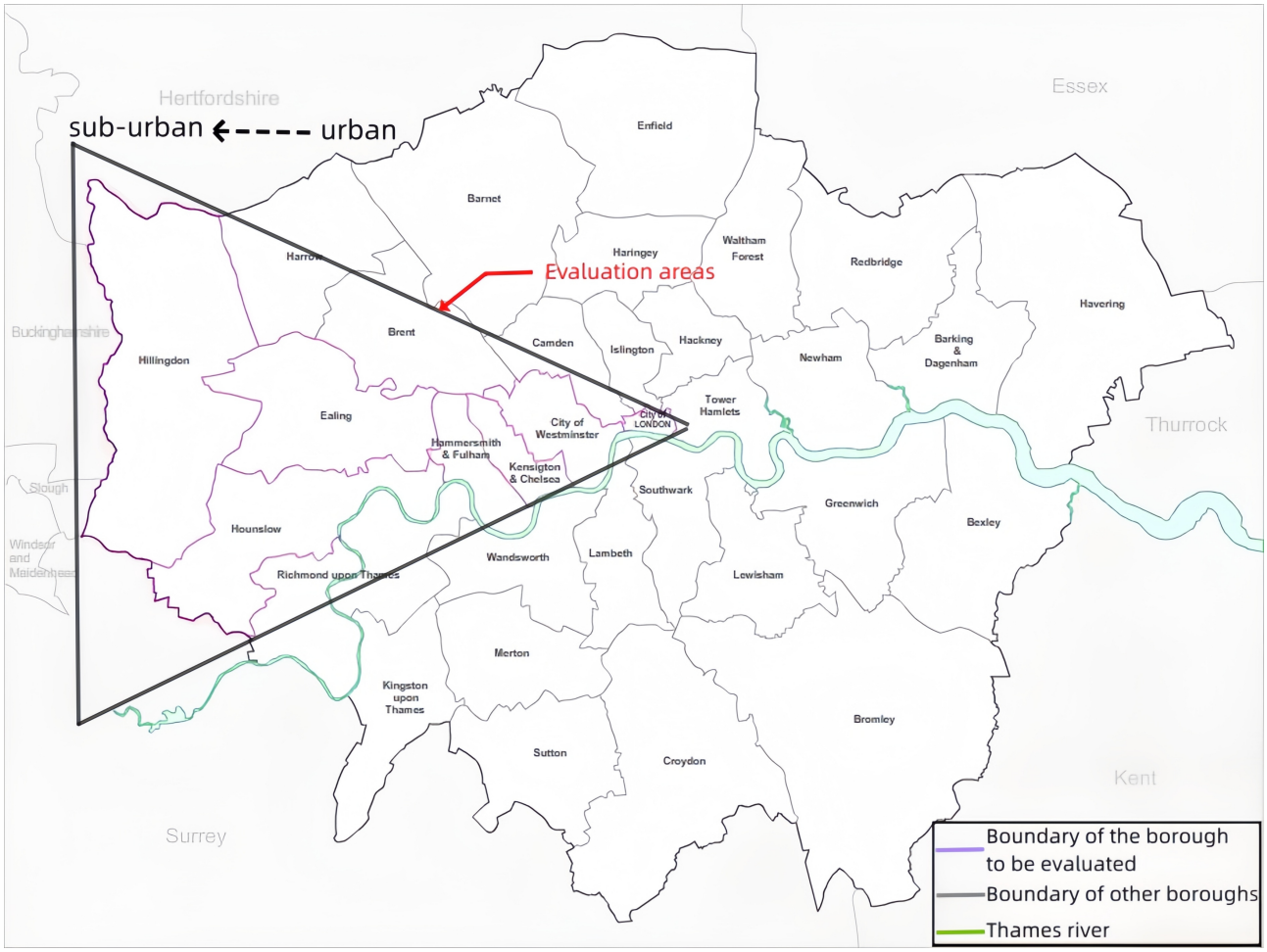


Figure 4.5: Illustration of the case study areas

4.3 Case Study

This section centres on assessing the capability of the suggested autonomous delivery options in various boroughs within London, UK. Precisely, we utilise geographic datasets encompassing diverse boroughs ranging from central London (urban areas) to the western outskirts of London (suburban areas). The selected boroughs consists of (sorted from urban to suburb): 1) City of London; 2) City of Westminster; 3) Kensington and Chelsea; 4) Hammersmith and Fulham; 5) Ealing; 6) Hounslow; and 7) Hillingdon, as illustrated in Figure 4.5.

Table 4.2 provides the size, the population, and the estimated daily e-commerce parcels of different boroughs (Statista, 2023). Given that the daily parcel volumes for various boroughs in London have not been disclosed by logistics operators or relevant industry reports, we opted for the following formula to make an volume estimation for a single borough: e-commerce parcel volume = e-commerce parcels generated per person per year (units) / 365 (days) × population of

Table 4.2: Size, population, and estimated daily e-commerce parcels of different boroughs in London

Boroughs	**Size (km^2)**	**Population (people)**	**Estimated daily volume (parcels)**
City of London	2.90	10,847	2,200
City of Westminster	8.29	211,365	42,852
Hammersmith and Fulham	6.33	185,238	37,555
Kensington and Chelsea	12.13	146,154	29,631
Ealing	21.44	369,937	75,001
Hillingdon	44.67	310,681	62,988
Hounslow	55.98	290,488	58,894

Table 4.3: Location of DC and UCC

DC & Boroughs	**Postcode**
DC	HP2 7XX
UCC in City of London	WC1X 0DG
UCC in City of Westminster	SE1 7HJ
UCC in Hammersmith and Fulham	W6 9NS
UCC in Kensington and Chelsea	W14 0PE
UCC in Ealing	W5 5EH
UCC in Hillingdon	UB8 2GG
UCC in Hounslow	TW3 1AA

the borough, wherein the average parcels generated per person per year in the UK is projected to 74 units according to the industry report (Mordor Intelligence, 2023). Furthermore, considering that this study focuses on assessing the cost and efficiency of deploying autonomous delivery services from the perspective of a single logistics operator, it is necessary to determine the proportion of parcel volumes accounted for by the company within the specified borough. Taking Royal Mail as an example, which holds the highest market share in the UK parcel delivery market at 30% (Statista, 2022), the author selects 30% of the estimated daily parcel volume of the borough as the evaluation benchmark. Table 4.3 shows the location of the central DC, and the UCC of each borough, respectively.

We focus on three destination building characteristics: apartment, house, and office spaces. The primary distinctions among these customer destination lie in varying heights and road conditions (e.g., ramps) in front of the buildings. Therefore, the characteristics of the building and the floor on which the customer resides can influence their choice of delivery service. For

Table 4.4: Proportions of each service mode accounts for the total parcel volume of the borough

Delivery options	**HV-led Scenario**	**ADR-led Scenario**	**MPL-led Scenario**
HV	50%	35%	15%
MPL	35%	15%	50%
ADR	15%	50%	35%
Total percentage		100%	

Notes: the figures in the table denote the percentage of the parcel volume distributed by the respective delivery option relative to the total parcel volume within the borough.

instance, house customers are more inclined to opt for ADR services than apartments. This is because houses typically have only 2 to 3 floors, making it convenient for customers to descend and retrieve their parcels promptly. Moreover, the ramps in front of house courtyards are more conducive to the movement of ADRs. On the other hand, customers residing in apartments, potentially on higher floors, may not promptly descend for parcel retrieval. Consequently, they might opt for MPL services, as MPL services afford customers more flexible pickup times. For HV delivery option, it is universally applicable to all types of buildings, given its broad and flexible scope of application. It is particularly attractive for customers situated in office spaces, as it often involves the delivery of significant commercial packages. Due to security considerations, these customers are more inclined to opt for HV delivery services. Customers' service selection for different delivery options is ultimately reflected in the varying proportions of each service mode accounts for the total parcel volume in the borough. Therefore, we conduct different experiment settings illustrated in Table 4.4, which mimics different scenarios of varying delivery options that dominate the parcel delivery of the borough.

From a geographic perspective, central London boroughs predominantly feature apartment-based housing (accounts for $\approx$ 60% of the total buildings), in contrast to the landed houses more common in the suburbs (accounts for $\approx$ 70% of the total buildings) (Office for National Statistics, 2021). Additionally, demand distribution is more spatially concentrated in urban areas, whereas it is more dispersed in suburban areas (Allen et al., 2018). Given that urban areas have higher traffic volumes and limited road and curb space compared to suburban areas, delivery activities that rely on curbside parking typically spend more time finding available parking spaces in urban areas (Li & Guo, 2014). Moreover, these parking spots are often located at a distance from the customers' locations.

Table 4.5 shows the parameter settings introduced in Section 4.2. The specific numerical values for the parameters within the table are configured on a per-vehicle basis. The estimation of the operating costs per mile for trucks is predicated on the assumption of an annual mileage of 50,000 miles with the fuel consumption rate of 6.5 Miles Per Gallon (MPG) and an ownership period of 10 years (Blakemore, 2018), while the estimation for delivery van is predicated on the assumption of an annual mileage of 30,000 miles with the fuel consumption rate of 21 MPG and an ownership period of 5 years (DriveSmart, 2023). The parameter settings of the table are based on industry and government reports, and research articles. The cost-related parameters are set considering the local rate of the UK market. To mimic the time consumption difference in finding available parking space between urban and suburban areas, we reference the study conducted by Reed et al. (2022a,b) wherein the average parking time consumption for each delivery option from urban to suburban areas is set ranging from 7 to 1 minute, with an interval of 1. Since the relevant government reports and research articles do not disclose the traffic congestion factor ϕ of London city, we assume that the congestion factor $\phi = 0.77$ based on the traffic congestion factor calculation model introduced by Petrik et al. (2014).

Table 4.5: Parameter settings for the evaluation model

Cost-related Parameters	**Value**
Hourly driver cost[1] c_{driver} $(£/hr)$	23
Unit operating cost of truck[2] c_{truck} $(£/mi)$	1.1
Unit operating cost of delivery van[3] c_{van} $(£/mi)$	0.55
Unit operating cost of MPL[4] c_{van} $(£/mi)$	0.1
Unit operating cost of ADR[5] c_{adr} $(£/mi)$	0.05
Unit fuel cost of truck[6] f_{truck} $(£/mi)$	1.21
Unit fuel cost of van[6] f_{van} $(£/mi)$	0.35
Unit electricity cost of MPL[7] e_{mpl} $(£/mi)$	0.03
Unit electricity cost of ADR[7] e_{adr} $(£/mi)$	0.02
Efficiency-related Parameters	
Unit time consumption of loading parcel at DC and UCC[4] t_{LP} $(min/parcel)$	0.003
Unit parking time consumption of HV[8] $T_i^{van_parking}$ $(min/stopover)$	$N([1, \cdots, 7], 1)$
Unit parking time consumption of MPL[8] $T_i^{mpl_parking}$ $(min/stopover)$	$N([1, \cdots, 7], 1)$
Unit parking time consumption of ADR[8] $T_i^{adr_parking}$ $(min/stopover)$	$N([1, \cdots, 7], 1)$
Unit sorting time consumption of HV[9] $T_i^{van_sorting}$ $(min/stopover)$	$N(2.4, 1)$
Unit handover time consumption of MPL $T_i^{mpl_handover}$ $(min/stopover)$	$N(2, 1)$
Unit loading time consumption of ADR[9] $T_i^{adr_loading}$ $(min/stopover)$	$N(2.4, 1)$
Unit handover time consumption of HV $T_i^{van_handover}$ $(min/stopover)$	$N(3, 1)$
Unit walking time of customer $T_i^{mpl_walking}$ $(min/customer)$	$N(5, 2)$
Unit handover time of ADR $T_i^{adr_handover}$ $(min/stopover)$	$N(3, 1)$
Vehicle-related Parameters (parameters in Appendix A included)	
Truck velocity[10] v_{truck} (mph)	55
Delivery van velocity outside & inside the service region[10] v_{out}^{van}, v_{in}^{van} (mph)	60, 25
MPL velocity outside & inside the service region[11] v_{out}^{mpl}, v_{in}^{mpl} (mph)	6, 6
ADR velocity inside the service region[12] v_{in}^{adr} (mph)	6
Maximum capacity of ADR[9] q $(parcels)$	3
Maximum battery power of ADR[13] U (Kw)	120
Power consumption of ADR[13] ζ (Kw/hr)	90
Capacity of HV delivery option (van capacity) Cap_{hv} $(parcels)$	60
Capacity of MPL delivery option (MPL capacity) Cap_{mpl} $(parcels)$	30
Capacity of ADR delivery option (accompanied van capacity) Cap_{adr} $(parcels)$	60
Emission-related Parameters	
Unit CO_2 emission of truck[14] m_{co2}^{truck} (g/mi)	1592
Unit CO emission of truck[14] m_{co}^{truck} (g/mi)	0.81
Unit NO_X emission of truck[14] m_{nox}^{truck} (g/mi)	5.55
Unit PM emission of truck[14] m_{pm}^{truck} (g/mi)	0.09
Unit CO_2 emission of van[14] m_{co2}^{van} (g/mi)	386
Unit CO emission of van[14] m_{co}^{van} (g/mi)	1.77
Unit NO_X emission of van[14] m_{nox}^{van} (g/mi)	0.17
Unit PM emission of van[14] m_{pm}^{van} (g/mi)	0.003

Notes: 1) ERI Economic Research Institute (2023); 2) Blakemore (2018); 3) DriveSmart (2023); 4) Lemardelé et al. (2021); 5) Korus (2018); 6) FleetNews (2023); 7) Li & Kunze (2023); BusinessEnergy.com (2023); 8) the average time consumption of finding parking space is: City of London = 7 min, City of Westminster = 6 min, Kensington and Chelsea = 5 min, Hammersmith and Fulham = 4 min, Ealing = 3 min, Hounslow = 2 min, Hillingdon = 1 min; 9) Reed et al. (2022a,b); 10) Pahwa & Jaller (2023); 11) Li et al. (2021a); 12) Li & Kunze (2023) 13) Liu et al. (2023b); 14) Pahwa & Jaller (2022)

4.4 Empirical Results

This section presents empirical findings from the case study outlined in Section 4.3. Initially, this section assess the projected costs, delivery efficiency, and emissions associated with various delivery options across three demand segmentation scenarios as identified in Table 4.4 (Horizontal comparison). Subsequently, we directs attention to the impact of diverse boroughs, ranging from urban to suburban, on costs and delivery efficiency (Vertical comparison). This analysis aims to ascertain the suitability and effectiveness of different delivery options in distinct borough types and road network density. This section finally focus on the performance comparison among different delivery options given that fulfilling the same parcel volume.

4.4.1 Performance Comparison of Different Demand Segmentation Strategies

This section focuses on quantifying the impact of different demand segmentation strategies on the overall delivery system. The calculation of performance indicators is based on the cumulative costs, time consumption, and gas emissions generated by the proposed three delivery modes across all evaluation boroughs.

Cost of Different Demand Segmentation Strategies

The results in Figure 4.6 indicate that the delivery scenario dominated by MPL service has the lowest total cost, followed by HV-led and ADR-led scenarios. In comparison to HV-led and ADR-led scenarios, the total cost in the MPL-led scenario is reduced by 39.45% and 44.79%, respectively. The primary source of savings originates from a reduced need for drivers, as the majority of customers personally visit MPL stopovers to retrieve their parcels. Across the three delivery scenarios, driver costs constitute the majority of the total costs (average percentage of 84.88%), followed by operational costs (average percentage of 9.26%), and fuel & electricity costs (5.74%). It is noteworthy that ADR-led scenario incurs the highest total costs in all three scenarios. This is attributed to the constrained capacity of ADR, resulting in longer delivery durations and vehicle travel distances during periods of increased delivery demand. Therefore,

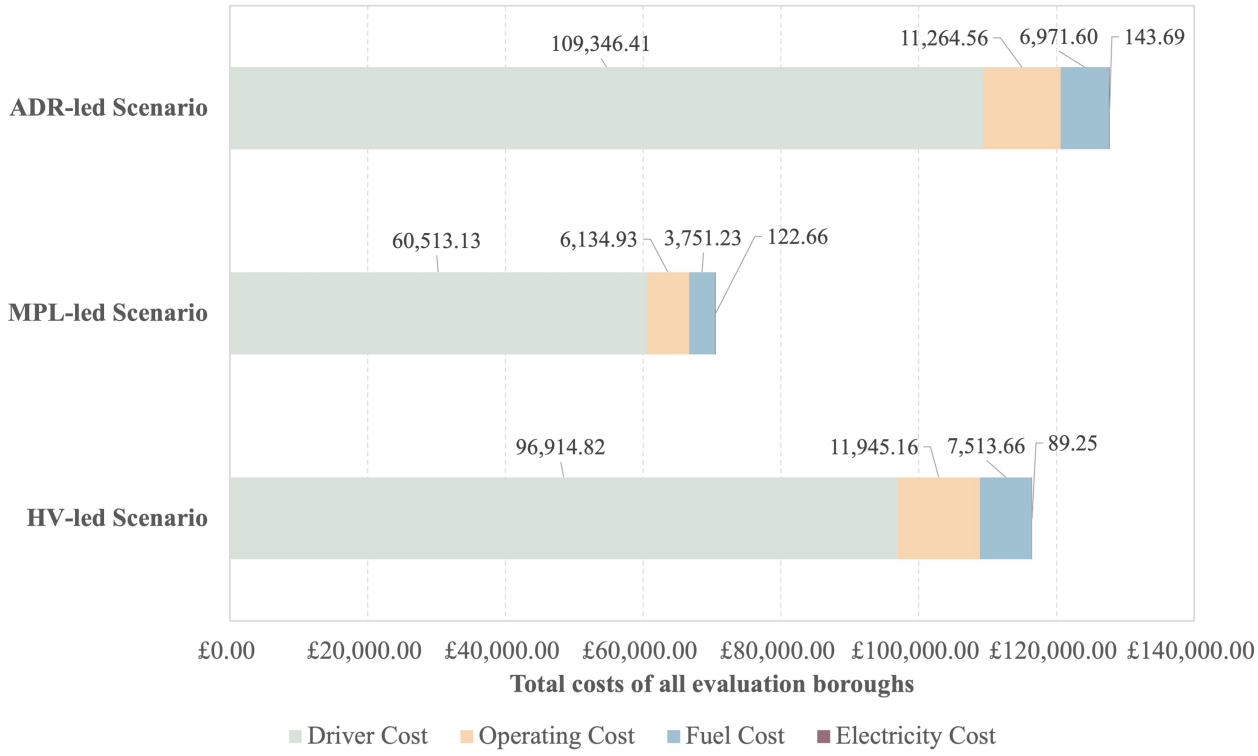
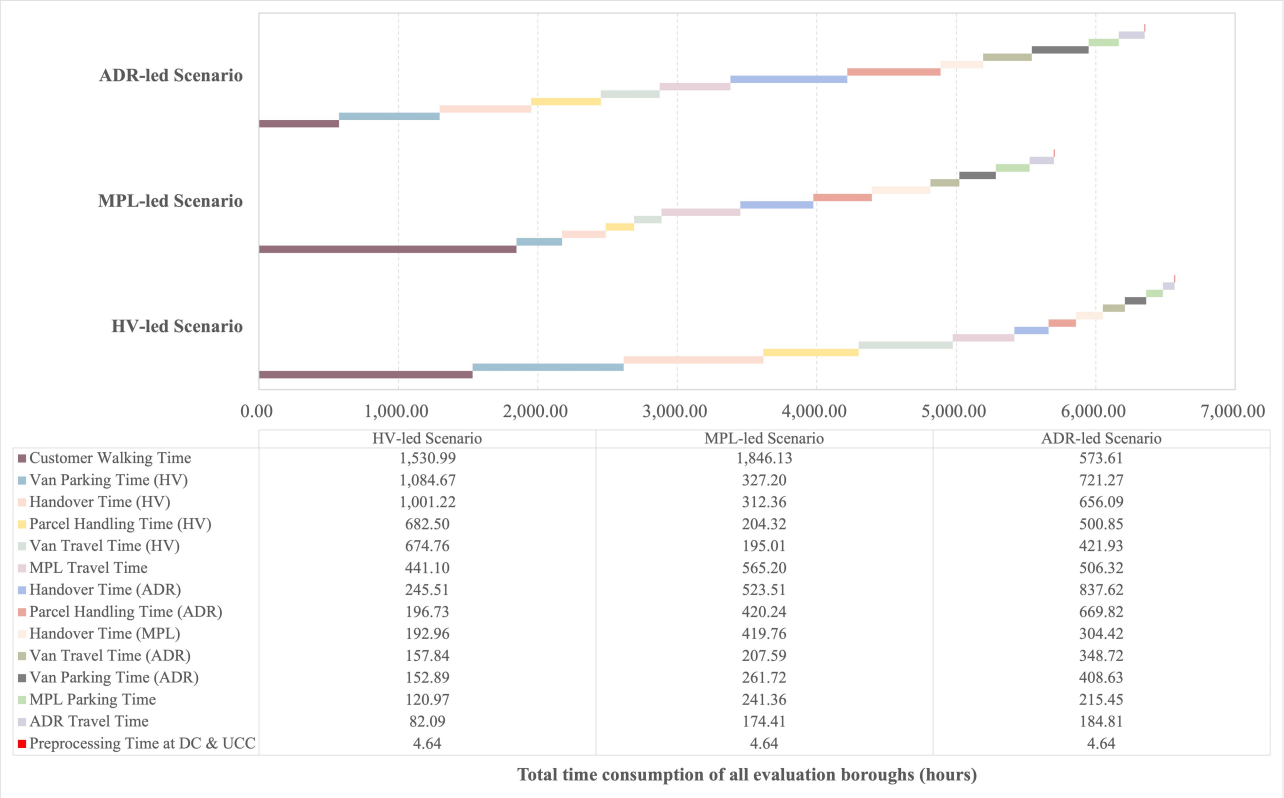


Figure 4.6: Cost comparison of different demand segmentation scenario

ADR is better suited as an auxiliary delivery option to alleviate excess demand.

Time Consumption of Delivery Activities in Different Demand Segmentation Strategies

The results in Table 4.6 reveal that the MPL-led scenario incurs the shortest total delivery time, while the HV-led scenario has the longest duration. For the HV-led scenario, customer walking, van searching for parking spaces under the HV delivery option, and delivery personnel performing the final leg of parcel handover are the three most time-consuming activities among all delivery operations, accounting for 23.31%, 16.51%, and 15.24% of the total delivery duration, respectively. In contrast, for the MPL-led scenario, the proportion of customer walking and MPL travelling in the total delivery time increases, reaching 32.37% and 9.91%, respectively, owing to a higher volume of delivery demands fulfilled by MPL. In the ADR-led scenario, parcel handover and ADR-accompanied van searching for parking spaces account for 13.18% and 10.98% of the total duration, respectively. In summary, parking dependence and parcel handover remain significant bottlenecks in delivery modes involving van participation.



	HV-led Scenario	MPL-led Scenario	ADR-led Scenario
Customer Walking Time	1,530.99	1,846.13	573.61
Van Parking Time (HV)	1,084.67	327.20	721.27
Handover Time (HV)	1,001.22	312.36	656.09
Parcel Handling Time (HV)	682.50	204.32	500.85
Van Travel Time (HV)	674.76	195.01	421.93
MPL Travel Time	441.10	565.20	506.32
Handover Time (ADR)	245.51	523.51	837.62
Parcel Handling Time (ADR)	196.73	420.24	669.82
Handover Time (MPL)	192.96	419.76	304.42
Van Travel Time (ADR)	157.84	207.59	348.72
Van Parking Time (ADR)	152.89	261.72	408.63
MPL Parking Time	120.97	241.36	215.45
ADR Travel Time	82.09	174.41	184.81
Preprocessing Time at DC & UCC	4.64	4.64	4.64

Figure 4.7: Time consumption comparison of different demand segmentation scenario

Table 4.6: Emission comparison of different demand segmentation scenario

Segmentation Scenario	CO_2 (g)	CO (g)	NO_X (g)	PM (g)
HV-led scenario	8,335,121.01	36,995.30	4,586.37	641.44
ADR-led scenario	7,737,298.94	34,253.99	4,323.08	594.98
MPL-led scenario	4,185,698.75	17,968.16	2,758.91	318.95

The MPL delivery, which reduces stopovers and mitigates parking dependence, has the potential to enhance delivery efficiency.

Gas Emissions of Different Demand Segmentation Strategies

The MPL-led scenario is the greenest delivery mode among the proposed three segmentation scenario, as illustrated in Table 4.6. In the HV-led scenario, greenhouse gas emissions are 99.13% higher compared to the MPL-led scenario and 7.73% higher compared to the ADR-led scenario. This is attributed to the predominant use of fuel-driven delivery vans, which perform the majority of parcel deliveries.

4.4.2 Performance Comparison of Different Boroughs

This section focuses on the impact of boroughs situated in urban and suburban areas on delivery performance. To achieve a fair comparison, the experiment sets the demand for each borough at 13,249 parcels, drawn from the average daily parcel volume across all boroughs. Furthermore, the calculation of performance indicators is based on the average value of the proposed three demand segmentation scenario.

Cost of Different Boroughs

The results in Figure 4.8 reveal that the total delivery costs for boroughs in the city centre (e.g., City of London, City of Westminster, and Kensington & Chelsea) are approximately 36.28% higher than those in suburban boroughs (e.g., Ealing, Hounslow, and Hillingdon). In contrast, the borough (e.g., Hammersmith & Fulham) situated in transitional geographical positions between the city centre and suburban areas incur costs that fall between the two. The increasing delivery cost in urban boroughs is primarily attributed to a significant increase in driver costs. Specifically, the driver costs for boroughs in the city centre are approximately £7,873 higher (41.71%) compared to those in suburban boroughs. The main reason for this phenomenon is that the road conditions in the city centre are more complex with narrow vehicle lanes, requiring more time to find available parking spaces. On the other hand, the operating costs for suburban boroughs are 10.06% higher than those for city centre boroughs. This is due to the larger geographical area of suburban boroughs, resulting in increased vehicle travel mileage. The differences between urban and sub-urban boroughs in fuel and electricity costs are insignificant.

Time Consumption of Delivery Activities in Different Boroughs

Table 4.7 presents the time consumption of various delivery activities associated with delivery options in different boroughs. It is noteworthy that the travel time of the accompanying van in the ADR delivery options tends to increase as the borough is located farther away from the city center, while the travel time of the ADR decreases. Specifically, the travel time of ADR

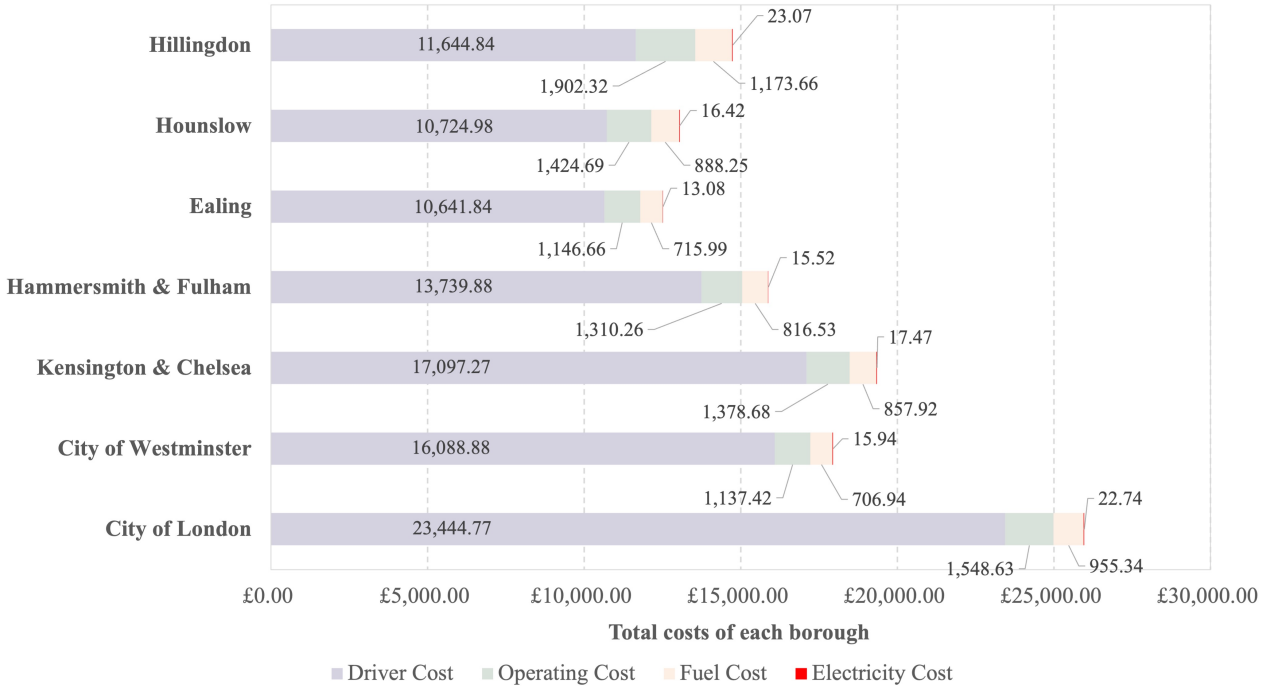


Figure 4.8: Cost comparison of different boroughs

in suburban boroughs is reduced by 84.56% compared to urban boroughs. This implies that in suburban boroughs, due to the dispersed geographical locations of customers, ADR is more frequently loaded onto vans to travel between different customers rather than making solo trips between customers, to save delivery time (since the van travels faster than ADR). Additionally, the time spent searching for parking spaces significantly decreases by approximately 80% from urban to suburban for all three proposed delivery options. There is no explicit relationship between different boroughs for parcel sorting and parcel handover activities, as topographical factors do not influence these activities. Finally, for the customer walking activity associated with MPL service, the walking time in suburban boroughs is reduced by 33.19% compared to urban boroughs. This is because logistics operators need to increase the number of stopovers in suburban areas to cover customer demands, thereby reducing customer walking time.

Table 4.7: Total time consumption (hours) of delivery activities in different boroughs

Boroughs	**Preproceed**	**Van Travel (HV)**	**Van Travel (ADR)**	**MPL Travel**	**ADR Travel**	**Van Parking (HV)**	**MPL Parking**	**Van Parking (ADR)**	**Sorting (HV)**	**Sorting (ADR)**	**Handover (HV)**	**Handover (MPL)**	**Handover (ADR)**	**Customer Walking**
City of London	0.66	80.02	25.63	73.10	79.88	366.61	91.35	133.13	87.94	56.83	117.48	58.15	70.03	227.86
City of Westminster	0.66	54.69	22.32	56.88	47.54	222.94	57.97	80.82	74.35	42.20	98.05	46.51	55.95	175.25
Kensington & Chelsea	0.66	67.88	26.23	64.56	48.74	210.20	53.25	77.63	77.33	53.17	111.39	55.88	70.12	348.30
Hammersmith & Fulham	0.66	59.09	30.42	69.22	25.48	141.31	37.62	58.15	64.61	56.70	120.02	48.66	68.00	184.69
Ealing	0.66	50.99	27.31	61.87	16.16	78.24	20.47	32.31	61.60	51.33	82.13	39.66	62.02	161.65
Hounslow	0.66	58.35	38.81	87.04	6.27	49.59	15.10	16.39	61.65	68.78	79.78	40.57	85.93	165.71
Hillingdon	0.66	79.42	51.53	125.00	4.77	11.36	6.42	7.76	64.75	85.78	93.15	40.20	107.23	174.69

Table 4.8: Gas emission of different boroughs

Boroughs	CO_2 (g)	CO (g)	NO_X (g)	PM (g)
City of London	1060166.46	4695.89	590.56	81.54
City of Westminster	786687.12	3430.17	478.84	60.22
Kensington & Chelsea	953504.51	4187.32	558.13	73.15
Hammersmith & Fulham	907601.31	3983.32	533.06	69.61
Ealing	796203.63	3485.49	474.31	61.02
Hounslow	987728.62	4324.78	587.75	75.70
Hillingdon	1300305.95	5813.27	684.20	100.28

Gas Emissions of Different Boroughs

Table 4.8 illustrates the greenhouse gas emissions generated by the delivery options in different boroughs. The borough at the very centre of the city (City of London) and the borough at the outskirts (Hillingdon) occupy the top two positions in terms of gas emissions. However, the distinction lies in the fact that the City of London experiences more gas emissions due to congested urban road conditions, whereas Hillingdon sees increased vehicle mileage owing to the dispersed locations of customers and, consequently, higher gas emissions.

4.4.3 Performance Comparison of Different Delivery Options

This section evaluates the performance of the proposed three delivery options in meeting the parcel delivery requests across all boroughs. To achieve a fair comparison, we set the demand for each borough at 13,249 parcels, drawn from the average daily parcel volume across all boroughs. The calculation of performance indicators for each delivery option is based on the cumulative costs, time consumption, and gas emissions generated by all vehicle fleets invested.

Cost of Different Delivery Options

The results in Figure 4.9 indicate that the MPL delivery option is the least costly among the three proposed delivery options (excluding vehicle acquisition costs). Specifically, the cost of MPL service accounts for only 0.66% of the conventional HV delivery option's cost and 0.9% of the ADR service option's cost. The primary source of total cost savings stems from a significant reduction in driver costs, where the driver cost for the MPL delivery option only arises from the long-haul journey while the last-mile journey does not. On the other hand, compared to HV and ADR delivery options that incur fuel consumption in the last mile journey, MPL service

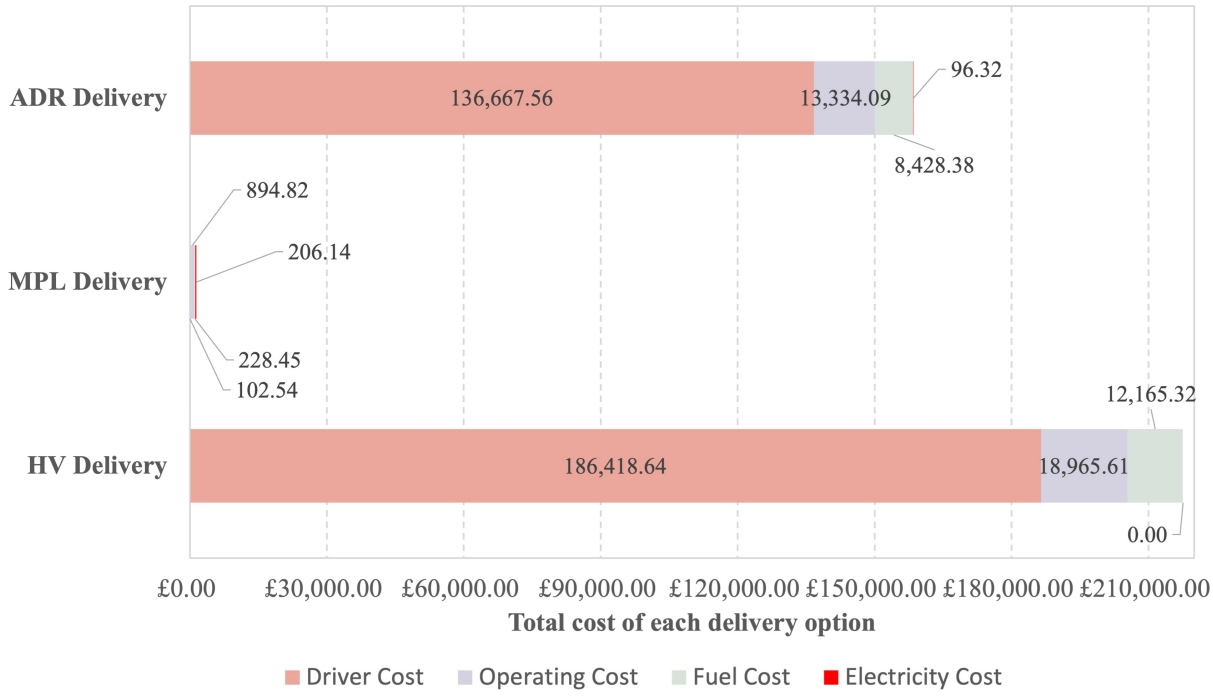


Figure 4.9: Cost comparison of different delivery options

generates no fuel consumption. Additionally, the unit electricity price for MPL is significantly lower than the unit fuel price, resulting in lower operating costs for MPL. For the ADR service option, its cost is reduced by 27.13% compared to the conventional HV delivery. In both HV and ADR delivery options, driver costs still represent the majority of the total costs, accounting for 85.69% and 86.21%, respectively. Therefore, with advancements in autonomous driving and automation technologies, the potential exists to significantly reduce driver costs if ADR can achieve automatic loading and unloading of parcels as well as easy access to the van.

Time Consumption of Delivery Activities in Different Delivery Options

Table 4.9 presents the time consumption for various delivery activities corresponding to the three proposed delivery options. The ADR service option exhibits the shortest time consumption, reducing by 18.2% compared to MPL delivery and 28.14% compared to conventional HV delivery. It is noteworthy that, for the MPL service, customer walking time constitutes 69.25% of the total time, while vehicle (i.e., MPL) travel time and parcel handover time are significantly lower than both HV and ADR delivery options. For HV and ADR service options, the time spent on parking search accounts for 39.89% and 23.75% of the total time consumption, respectively, while parcel handover time represents 25.77% and 26% of the

total time consumption. These two delivery activities are the most time-consuming during the delivery process.

Table 4.9: Total time consumption (hours) of delivery activities in different services

Services	**Preproceed**	**Van Travel (HV)**	**Van Parking (HV)**	**Sorting (HV)**	**Handover (HV)**	**MPL Travel**	**MPL Parking**	**Handover (MPL)**	**Customer Walking**	**Van Travel (ADR)**	**ADR Travel**	**Van Parking (ADR)**	**Sorting (ADR)**	**Handover (ADR)**	**Total**
HV	0.66	1364.21	3296.15	1472.73	2129.83	-	-	-	-	-	-	-	-	-	8263.59
MPL	0.66	-	-	-	-	1145.23	484.22	602.04	5027.71	-	-	-	-	-	7259.86
ADR	0.66	-	-	-	-	-	-	-	-	937.13	802.68	1410.24	1243.37	1544.15	5938.23

Table 4.10: Gas emission of different delivery options

Delivery Options	CO_2 (g)	CO (g)	NO_X (g)	PM (g)
HV service	13465229.24	60519.37	6845.75	1040.15
MPL service	300569.60	152.93	1047.84	16.99
ADR service	9343922.21	41621.15	5030.66	719.84

Gas Emissions of Different Delivery Options

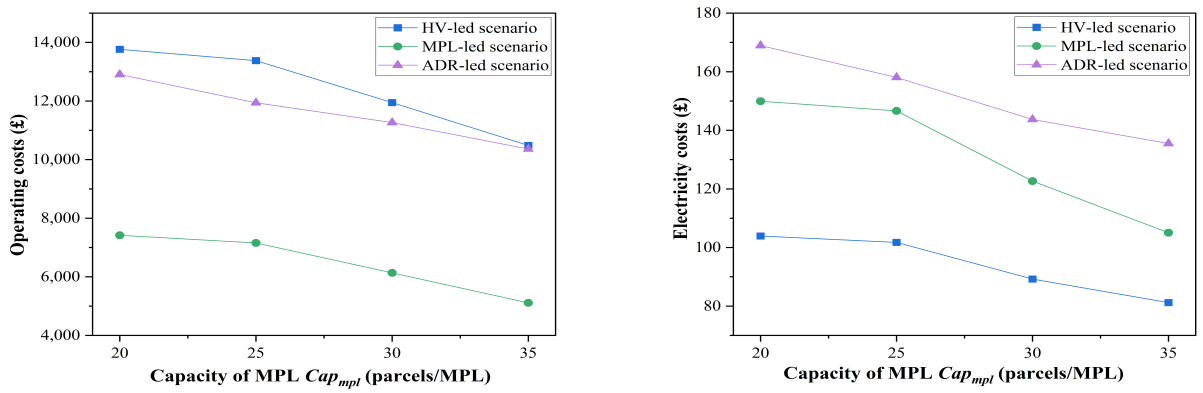
The MPL delivery service exhibits the lowest greenhouse gas emissions among the three proposed delivery options, with ADR delivery following and HV delivery service showing the highest emissions, as delineated in Table 4.10. Emissions from the MPL delivery option are attributed solely to the long-haul journey of parcels, where a truck transports packages from the DC to the local UCC. In comparison to the MPL-led scenario, greenhouse gas emissions for conventional HV delivery service are 97.77% higher, and when compared to the ADR-led scenario, emissions are 96.78% higher. In the case of the ADR delivery option, greenhouse gas emissions are 30.61% lower than those associated with HV delivery. Consequently, the implementation of MPL and ADR holds the potential to enhance urban livability.

4.4.4 Effect of service capacity

This section investigates the effect of different capacity settings of MPL and ADR under varying demand segmentation strategies on the proposed performance indicators. Regarding the impact of MPL capacity on costs, since MPLs are autonomous and electrically powered, changes in MPL capacity will not affect the overall delivery system's driver costs and fuel costs (considering the parcel fulfilment for the same customer set). Thus, Figure 4.10 highlights the effects of capacity on operating and electricity costs.

Larger MPL capacities reduce the number of MPLs needed, leading to savings in operating and electricity costs, especially in MPL-led scenarios (Figure 4.10a, b). In HV-led scenarios, operating costs are lower than in ADR-led scenarios when MPL capacity exceeds 35 parcels, due to a higher proportion of demand being fulfilled by MPLs compared to ADR services.

Differ from MPL service, vans accompanying ADR incur driver costs and fuel expenses (Figure 4.11). The results indicate that greater accompanying van capacity can reduce driver costs since less fleet size is required. In ADR-led scenarios, the cost-saving effect is more



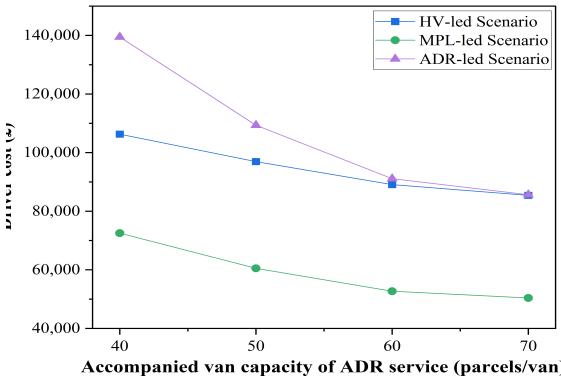
(a) Effect of MPL capacity on operating costs (b) Effect of MPL capacity on electricity costs

Figure 4.10: Effect of different MPL capacity settings on costs

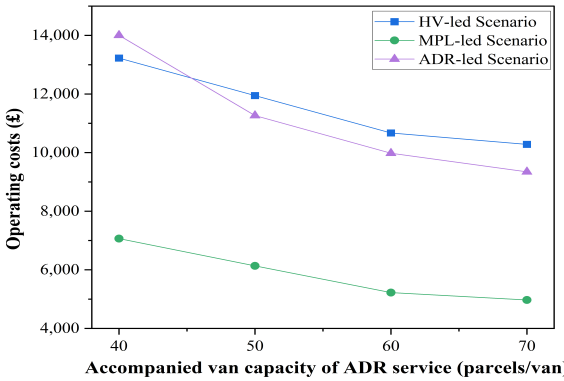
pronounced, especially when the van capacity exceeds 70 parcels/van (Figure 4.11a). In such cases, the driver cost of the ADR-led scenario has the potential to be lower than those in HV-led scenarios. Increasing ADR service capacity also offers potential savings in operating and fuel costs; however, these savings diminish as capacity continues to grow (Figure 4.11b, c). In contrast, electrical costs increase with van capacity because ADR serves more customers per delivery round, leading to longer travel distances and higher electricity consumption (Figure 4.11d). Nevertheless, the increase in electrical costs is less significant.

Figure 4.12 presents the effect of MPL capacity on different delivery activities. The results indicate that increasing MPL capacity allows more customers to be served at a single stopover, thereby reducing the number of stopovers needed to satisfy a specific customer group. This decreases the MPL's travel distance (Figure 4.12a) and the time spent searching for parking at each stopover (Figure 4.12b). However, serving more customers at a single stopover may alter the optimal pre-determined stopover locations (since these locations are determined by customer spatial distribution), potentially increasing customers' walking distance for parcel pick-up (Figure 4.12d). Changes in MPL capacity have no significant impact on the time customers spend operating the MPL compartments for parcel handovers since the parcel handover was only affected by customers (Figure 4.12c).

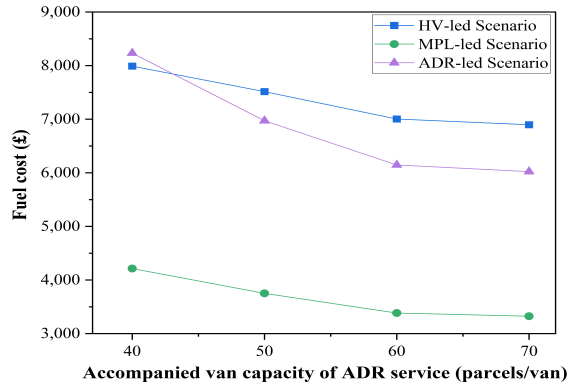
Figure 4.13 illustrates the impact of ADR service capacity on delivery activities. Increasing ADR capacity reduces the travel time of accompanying vans, as fewer vans are required to meet demand. However, these time savings are less significant compared to MPL services (Figure 4.13a). This difference arises because MPL consolidates multiple customers at a single stopover for parcel pickup, significantly reducing vehicle travel time due to spatial aggregation, while



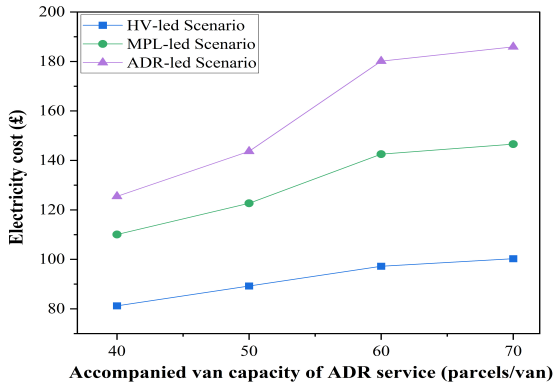
(a) Effect of ADR capacity on driver cost



(b) Effect of ADR capacity on operating costs



(c) Effect of ADR capacity on fuel costs



(d) Effect of ADR capacity on electricity cost

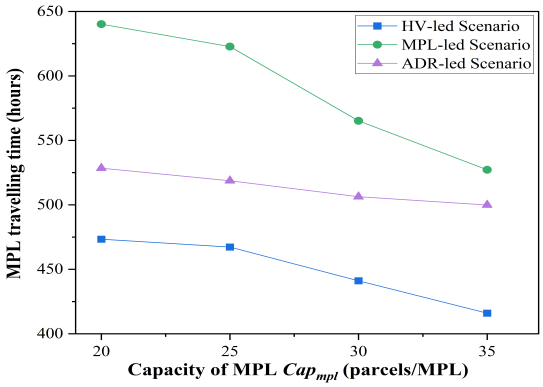
Figure 4.11: Effect of different ADR service capacity settings on costs

ADR must still visit most customers individually. As ADR capacity grows, the number of customers served per delivery round increases, resulting in longer total ADR travel time (Figure 4.13b). Additionally, higher ADR capacity leads to more intermediate stops for parking and parcel handling, increasing time consumption (Figure 4.13c,d).

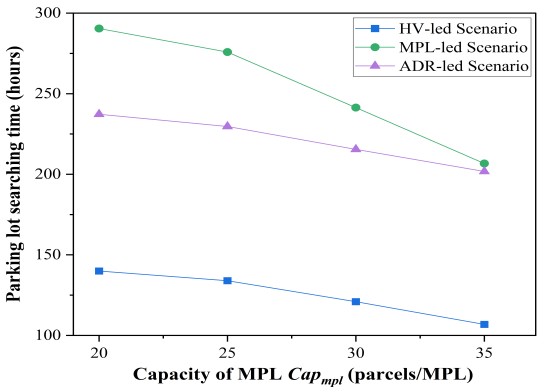
Regarding the impact of service capacity on emissions, MPLs, as electrically powered vehicles, do not produce emissions. Consequently, Figure 4.14 focuses solely on the effect of ADR service capacity on total emissions. The result shows that the total emission reduces with the increasing ADR service capacity as less fleet size is required to satisfy demand.

4.4.5 Implications and discussion

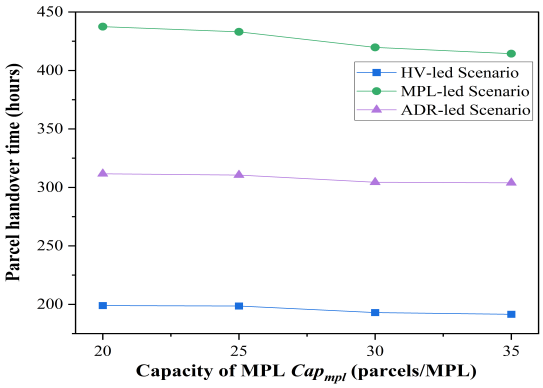
Based on the experimental results in Section 4.4, MPL and ADR delivery modes demonstrate superiority over the conventional HV delivery method in terms of cost, delivery efficiency, and sustainability. Specifically, MPL and ADR delivery modes excel in reducing labour costs and dependency on parking spaces incurred from delivery activities. This advantage



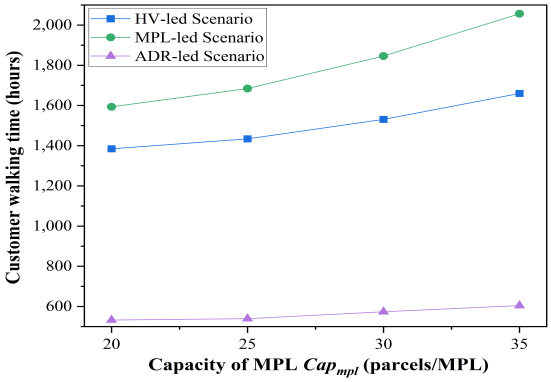
(a) Effect on travelling



(b) Effect on parking lot searching



(c) Effect on parcel handover



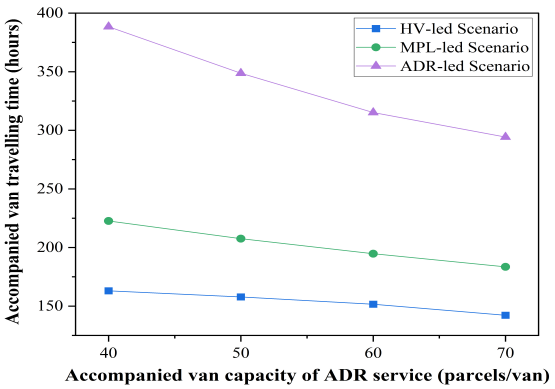
(d) Effect on customer walking

Figure 4.12: Effect of different MPL capacity on delivery activities

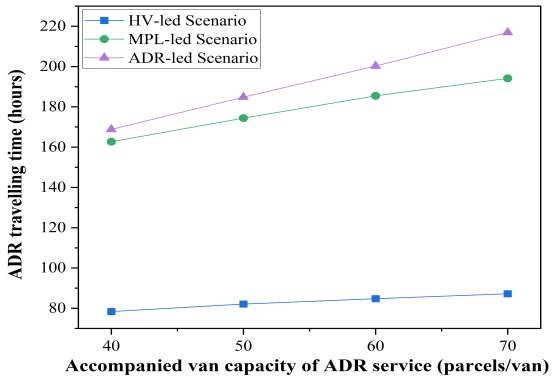
is particularly notable in urban areas where curbside space is more constrained. Unlike the prolonged occupation of parking spaces by the HV model, MPL and ADR services utilise sidewalks more efficiently, thereby minimising parking space occupation.

Table 4.11 presents a fitness matrix comparing various delivery options across different geographic regions and destination types. The delivery options in parentheses are ranked by the percentage of demand they meet, from highest to lowest (i.e., the first service listed is the predominant mode). In suburban and high-rise building-dense areas, HV is predominant due to dispersed demand and easier parking availability, as customers are less willing to walk longer distances for MPL parcel pick-up. In urban areas and non-high-rise building-dense regions, MPL is preferred because of denser demand and greater parking challenges, making customers more receptive to MPL's shorter walking distances. In transitional areas with better road conditions and minor parking shortages, ADR is predominant, offering potential enhancements in delivery efficiency.

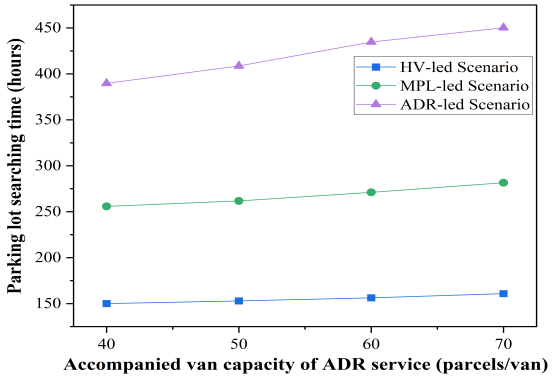
For logistics operators, the large-scale effective deployment of MPL and ADR depends on



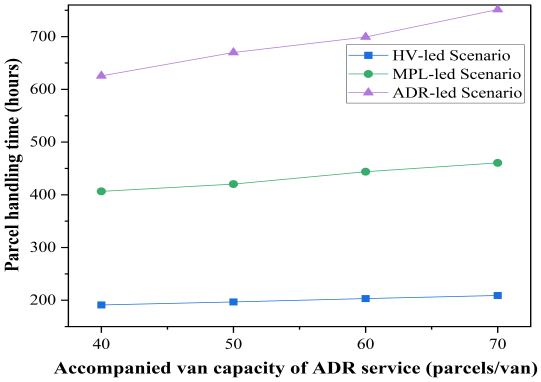
(a) Effect on van travelling



(b) Effect on ADR travelling



(c) Effect on parking lot searching



(d) Effect on parcel handling

Figure 4.13: Effect of different ADR service capacity on delivery activities

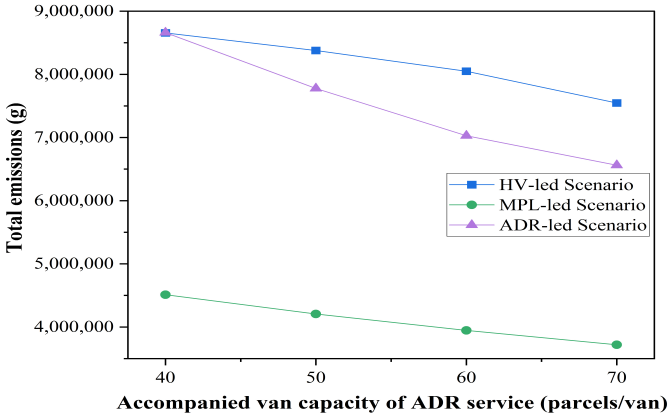


Figure 4.14: Effect of ADR service capacity on total emission

several factors. The first is segmenting demand according to the capacity of different delivery modes. As evidenced by the results in Section 4.4.1, different demand segmentation strategies lead to significant differences in total cost and delivery efficiency. For example, due to ADR's limited compartment size and the manual operation required for accompanying vans, heavy reliance on ADR service does not yield cost savings in high-demand-density areas with dense

Table 4.11: Fitness matrix of different autonomous delivery options against geographic regions and destination types

Proportion of high-rise building/Regions	**Urban**	**Transition area**	**Sub-urban**
<30%	(MPL, ADR, HV)	(ADR, MPL, HV)	(HV, ADR, MPL)
30%-50%	(MPL, HV, ADR)	(ADR, HV, MPL)	(HV, ADR, MPL)
>50%	(HV, MPL, ADR)	(HV, ADR, MPL)	(HV, ADR, MPL)

: MPL service dominant. : ADR service dominant. : HV service dominant.

high-rise buildings. Therefore, ADR is better served as a supplementary option to HV and MPL services to meet demand. The MPL service, being the most economical and environmentally friendly delivery method, is convenient for customers who cannot receive packages in person or live in low-rise buildings/landed houses. However, it is less attractive to those customers who reside on high floors and can sign for their packages in person, due to the need to navigate stairs and visit assigned MPL's parking lot. The actionable insights for logistics operators include: i) implement advanced demand forecasting tools to determine the optimal mix of MPL, ADR, and HV services. ii) develop algorithms to allocate fleets based on real-time demand and capacity constraints dynamically.

The second factor is the predominant destination types within the service area. In areas with mostly high-rise buildings, the HV delivery method may remain dominant, as most customers are unwilling to travel extra distances or navigate stairs to retrieve packages. In areas with primarily low- to mid-rise buildings, MPL and ADR can attract more customers through lower service prices, offsetting the perceived inconvenience. Therefore, operators are suggested to conduct detailed geographic and demographic analysis tailoring delivery strategies to specific areas. Additionally, offering tiered pricing strategies to incentivise customers in different building types to choose the most efficient delivery mode.

The third factor involves considering the trade-offs between fleet capacity, cost, and efficiency. For example, while a larger MPL capacity can serve more customers per delivery round and reduce the number of vehicles needed, it may increase the average walking time for customers to pick up their packages, thereby decreasing customer satisfaction. For ADR services, increasing the van capacity can also serve more customers per delivery round, reducing the number of vans and ADR units required. However, it may also increase the number of intermediate stops for the van.

4.5 Concluding Remarks

This chapter evaluates the feasibility and effectiveness of autonomous delivery modes across various boroughs in London, ranging from the city centre to suburban areas. The assessment primarily focuses on three dimensions: cost, delivery efficiency, and emissions pollution. To achieve this objective, we first define a collaborative delivery scenario comprising traditional HV delivery, MPL delivery, and ADR delivery options. Subsequently, an adjusted CA method is applied to estimate the fleet's travel distances and delivery time consumption within this collaborative delivery scenario. The adjusted CA method builds upon the conventional CA approach for estimating delivery distances and times in long-haul parcel transportation. While the estimation for the last-mile journey of parcels is achieved by precisely solving the delivery route length and time consumption for a specific region within a borough and subsequently mapping these values to the remaining regions within the same borough using a mapping function. This approach ensures computational efficiency while maintaining prediction accuracy and considering the impact of varying road conditions on the delivery process within different regions.

We sequentially selected geographic and demand data from 7 boroughs extending from the city centre to the western suburban areas of London as a case study. The evaluation encompasses four perspectives: i) the performance of different demand segmentation strategies under the collaborative delivery modes of the proposed three delivery options; ii) the performance of the collaborative delivery mode in each borough; iii) the performance of each delivery option independently in meeting the demand of all boroughs; and iv) the effect of different service capacity on the proposed performance indicators. The results indicate that the scenario dominated by MPL delivery option (fulfilling 50% of total demand), with HV (fulfilling 35% of total demand) and ADR as auxiliary delivery options (fulfilling 15% of total demand), achieves optimal cost savings and delivery efficiency improvements among the proposed three order segmentation strategies. On the other hand, the scenario dominated by ADR delivery (fulfilling 50% of total demand), with HV (fulfilling 35% of total demand) and MPL as auxiliary delivery options (fulfilling 15% of total demand), outperforms the HV-dominated delivery scenario in terms of delivery efficiency and gas emissions but comes at the expense of higher cost investment.

In terms of the performance of different boroughs, the central city boroughs incur higher

costs and greater time consumption compared to the suburban boroughs. This is attributed to the more intricate road conditions in urban areas. Driver costs constitute the core of the total costs, approximately 80%, and the activity most time-consuming during the delivery process is the search for parking spaces, particularly pronounced in the central city boroughs. Therefore, optimising curbside parking spaces emerges as a crucial factor for enhancing delivery efficiency.

Regarding the performance of each delivery option, the MPL delivery service emerges as the most cost-effective among the proposed three delivery options, while the ADR delivery option exhibits minimal time consumption. Conversely, the conventional HV delivery scheme proves to be the highest in both cost and delivery time among the three proposed delivery approaches. For the MPL delivery service, the walking time incurred by customers for parcel retrieval constitutes the most time-consuming activity in the delivery process. Logistics operators can alleviate this drawback by increasing the stopovers during the delivery process. However, for HV and ADR delivery options, the search for parking spaces and parcel handover remains a bottleneck in last-mile delivery. Therefore, with the continuous development of autonomous driving technology and the increasing autonomy of ADR, the prospect of ADR autonomously loading parcels could potentially enhance delivery efficiency.

Increasing the capacity of MPL and ADR fleets has the potential to reduce costs but may come at the expense of customer convenience and increased intermediate stops. A larger MPL fleet capacity allows more customers to be served at a single stopover, potentially lengthening customer walking distances due to changes in pre-determined stopover locations. For ADR service, increasing the van capacity may result in more intermediate stops, thus extending the overall delivery time.

To conduct an unbiased discussion, we provide the following considerations. The first pertains to the impact of capital and insurance costs on deploying the proposed autonomous delivery services. Deploying emerging autonomous vehicles requires substantial upfront investment, making it difficult for logistics operators to break even in the short term. However, in the long term, especially considering rising labour costs and shortages, the significantly lower operational costs of autonomous delivery service will offset the initial investment, leading to profitability. Therefore, for logistics operators, devising appropriate purchasing cycles and strategies will help ensure financial liquidity while deploying new delivery services. On the other hand, given that regulations are still evolving and considering the need to protect pedestrians, buildings, and surrounding vehicles in the operating environment, future operators may need

to allocate more funds for insurance.

From the methodological perspective employed in this study, although the improved CA integrates discrete exact route solution techniques to achieve a sound balance between computational feasibility and evaluation effectiveness, there are limitations in quantifying evaluation accuracy. These limitations stem from two factors: i) the lack of standardised and recognised evaluation datasets and performance metrics, which prevents fair pairwise comparisons of different evaluation techniques on the same dimensions. For the same case dataset (such as the London dataset used in this study), no other research has applied diverse evaluation techniques (such as multi-agent simulation) as a benchmark for comparison. ii) given that MPL and ADR services have not yet been broadly adopted and actual operational performance data has not been disclosed by operators, it is challenging to compare the theoretical results derived by evaluation techniques with the truth values. This limitation can also be observed in the relevant literature that employs the CA method for evaluating last-mile delivery services (Ansari et al., 2018; Zhang et al., 2024).

Future research directions could encompass the interaction between autonomous delivery vehicles and passenger vehicles, as well as the interaction between sidewalk delivery robots and pedestrians to the evaluation scope. Additionally, more autonomous delivery modes, such as drone delivery, can be incorporated into urban logistics scenarios to comprehensively assess their impact on cities, subject to large evaluation areas and real-world operating data are given.

Chapter 5

Customer Decision Behaviour Modelling for Autonomous Delivery Service Selection

The work of this chapter is adapted from the published conference paper at the **102nd Transportation Research Board Annual Meeting**, and is currently under consideration by **Transportation Research Part C: Emerging Technologies** [1].

Chapter 4 has already substantiated the feasibility and effectiveness of the MPL and ADR delivery modes for future large-scale deployments, examined from the perspectives of logistics operators and urban sustainable goal. Building upon this foundation, equally crucial is the comprehension of MPL and ADR delivery modes from the consumers' standpoint. Therefore, this chapter investigates the impact of customers' decision behaviour regarding selecting MPLs and ADR services on overall delivery performance. The MPL operation follows the corresponding scenario definition described in Section 4.1 in Chapter 4 while the ADR transports parcels from UCC to customers. This chapter answer the following research questions:

- **Q1:** Under what circumstances do customers choose the ADR or MPL service?

[1] Liu, Y., Escribano-Macias, J., Zhang, K., Ye, Q & Angeloudis, P. (2023), Kernelised-Support-Vector-Machine-Based Congestion-Game Model for Coordinated Delivery Service Selection Using Mobile Parcel Lockers and Delivery Robots. Available at SSRN 4383034.

- **Q2:** How will the pricing of ADRs affect customer preference?
- **Q3:** How can operators configure the operating parameters of MPLs and ADRs service to optimise system performance, such as the investment in MPL fleets and capacity?

Drawing upon the study conducted by Gibbons et al. (1992) and Hayel et al. (2016), this chapter modelled the service selection between MPLs and ADRs as a distinctive congestion game involving non-cooperative customers. The model incorporates knowledge of the theoretical game theory and queue theory while considering the probability of congestion occurring in the service queue when customers are waiting to be served by a specific delivery option. The customer's payoff depends upon the service it selects and the queue length waiting for service. The literature review in Section 3.3.2 reveals three main limitations of the existing game theory studies:

(1) Most non-cooperative game models assume that players are always rational (Muggy & Stamm, 2014; Heidari Gharehbolagh et al., 2016), indicating that players only perform the policy which maximises their interests in response to the actions of other players, with the corresponding equilibrium yielded defined as Nash equilibrium. However, Nash equilibria normally present sub-optimal characteristics compared with the optimal allocation of the system (Wang et al., 2015), this is known as the Price of Anarchy (POA) in which POA can be arbitrarily large (Roughgarden, 2005).

(2) Players might exhibit bounded rational decision behaviours due to inaccurate or incomplete information, which could obstruct an optimum equilibrium of the system. The impact of customers' bounded rationality on the system can be measured by the Price of Rationality (POR), which was defined as the ratio between the system's utility of bounded rationality and the utility of full rationality (Hayel et al., 2016). To the best of the author's knowledge, none of the related studies has discussed the dynamic of POR where both the level of bounded rationality and the global information are considered in customers' decision patterns simultaneously.

(3) No relevant study has considered the inverse process of game models, where the expected equilibrium is predetermined to derive the optimal operational parameters

accordingly. This process enables logistics operators to have an agile adjustment for the capacity configurations in response to the market change, thus obtaining the expected market share.

To bridge the above research gap, this chapter proposes a KSVM-based non-cooperative congestion game model to investigate the delivery service selection between MPLs and ADRs and considers varying levels of customer rationality. On the basis of this model, logistics operators offer customers complete information during their decision-making process, aiming to minimise the impact of incomplete information on customers' payoff estimation. This manner promotes mutual benefits for both the system and the customers through synchronised actions. In achieving this aim, the objectives are:

- **Congestion-Game model establishment:** establishing payoff functions for MPLs and ADRs based on their operational characteristics. A non-cooperative queuing system is established to model the game process of service selection between MPLs and ADRs, in which the congestion situation of the service queue caused by peak demand is considered. Further, we investigate the impact of critical operating parameters on the system's performance.

- **Global information introduction:** developing a novel algorithm that employs global information in customer's decision-making behaviour, where the algorithm induces the player's choice to ensure both system and players benefit from the same action (Branzei et al., 2008). In practice, logistics operators can achieve this mechanism by providing customers with access to global information, such as the current length of the waiting queue for a specific delivery service, through the service portal. The algorithm's performance will be compared with a benchmark without global information applied (addressing limitation 1).

- **Bounded rationality characterisation for customers:** applying a multi-nomial logit model capturing customers' bounded rational behaviour in delivery service selection. Various levels of bounded rationality models are compared with their fully rational counterparts to assess systems' efficiency (addressing limitation 2).

- **Regression prediction:** developing a KSVM model to determine the appropriate price

for ADR service, such that logistics operators can allocate resources efficiently while achieving the expected market share. KSVM incorporates players' decisions via the game mechanism: the operating parameters and the equilibrium are utilised as input for the KSVM. The resulting output of the KSVM corresponds to the service charge of ADR services (addressing limitation 3).

This study adds to the body of knowledge in the following aspects by attaining the above objectives:

- **Modelling contribution:** a novel congestion-game model has been introduced to address the autonomous delivery service selection problem by incorporating customers' bounded rational behaviour and leveraging global knowledge. Additionally, to the author's knowledge, studies on last-mile service selection have not yet introduced the inverse game process where KSVM is applied to achieve key operating parameter prediction.

- **Theoretical contribution:** we introduced a novel notion of Price of Global Information (POGI) that better represents the efficiency of the system when both individual payoff and the global utility are considered. The POGI is, therefore, used for measuring the value of the global information that contributes to improving the system performance, which provides a new measuring metric for future game-related research.

- **Managerial contribution:** we proved the existence and uniqueness of the equilibrium for fully and bounded rationality models regarding the service selection between MPLs and ADRs. Additionally, this study identified the key parameters in the context of MPLs and ADRs operations and investigated their effects on the system performance. Managerial implications are provided to support logistics professionals in improving last-mile operations based on the parametric analysis.

- **Algorithmic contribution:** the computational experiment showed that the proposed algorithm with global information applied increases the system efficiency by 28% to 68%, effectively enhancing customers' decision behaviour.

The rest of the paper is structured as follows. Section 5.1 introduces a congestion-game model for customer service selection behaviour. Section 5.2 proves the existence and uniqueness of equilibrium models and demonstrates the inverse game process. Section 5.3 illustrates the solution algorithms for different rationality models. A novel algorithm that applies global information in customer decision-making behaviour is introduced. Section 5.4.7 discusses the results of experiments and identifies the effects of critical factors on system performance. Section 5.5 provides a summary of the findings.

5.1 Problem Statement and Congestion-game Model

This study investigates the decision-making behaviour of a large customer group in the context of autonomous last-mile delivery service selection that considers varying levels of customer rationality. In this problem, logistics service operators are assumed to provide two types of delivery services for a large customer group: MPLs and ADRs.

We considers the on-demand delivery scenario where each individual within a large customer group has the authority to select either MPL or ADR service before parcels are out for delivery. Additionally, customers could apply mixed strategies to respond to the service choices of other customers. This emulates the flexibility afforded to customers when using the service selection portal, enabling them to opt for the delivery option with fewer people selected, thereby reducing the waiting time for parcel reception.

Table 5.1 summarises the notation and terminology of the model. Customers possess prior knowledge regarding the characteristics of these services, allowing them to make decisions: MPLs will depart from UCC and serve customers with greater flexibility for parcel pickup, providing a buffer time for customers to visit MPLs at each stopover where the MPL remains for a while. Each MPL can serve multiple customers at each stopover. MPLs will return to the UCC once all assigned customers on the route are served. In contrast, the ADR service requires customers to be at designed destinations to unlock the compartment on time, ensuring that subsequent deliveries are not disrupted. Additionally, customers are not required to pay extra fees if the storage duration $1/\mu_s$ of their package occupies the MPL compartment without

exceeding the predefined threshold Ω_s. Customers will be charged $C^{MPL}_{storage}$ (depends on the duration that exceeds the threshold Ω_s) only if the storage duration $1/\mu_s$ exceeds Ω_s. The existing pilot studies have introduced such an operation mechanism for stationary or mobile parcel locker network (Lemke et al., 2016; Kämäräinen et al., 2001). However, customers have to access MPL in person for parcel collection when MPL arrives at the predefined parking location, which yields a travel distance D. Section 5.1.3 provides a detailed description of the MPL service.

In contrast to MPL service, ADR offers door-to-door delivery for an additional service charge q. Customers must be present at their residences to ensure prompt collection of their parcels. Considering the time T consumed by ADR to transport the package from the UCC to customers' locations varied, we consider that customers have different tolerance levels Ω_w for different levels of waiting time. Section 5.1.3 provides a detailed description of the ADR service.

As this study centres on the system cost influenced by the cumulative customer's decision behaviours in the on-demand delivery service selection scenario, it is imperative to employ efficient models to estimate the service cost. Continuous approximation and queue models are two widely used methods for system evaluation (Ansari et al., 2018; Huang et al., 2013; Campbell, 2013; Lei & Ouyang, 2018). The continuous approximation is mainly used for assessing the overall delivery performance of the system, yet it falls short in examining the cumulative effect of individual decision-making behaviour. On the other hand, the queue model is designed to capture the costs experienced from the customer's perspective, because the queue model can map detailed operational configuration to each customer. Thus, we integrate a non-cooperative congestion-game model with the queue model to characterise the service selection of each customer and the interactions among customers' strategies, enabling the modelling of customers' service selection based on system observations. A fruitful discussion regarding the features and the use case of queue models and continuous approximation approach is provided in Section 5.4.7 to have an unbiased standpoint.

The proposed game model considers the congestion probability of the service queue that mimics the peak hour of parcel arrivals. Additionally, we consider customers' bounded

rational behaviours since they cannot observe the complete information during the decision-making process. The following sections sequentially define the model assumptions, agents & action space, payoff function, optimal response policy, and evaluation metrics for the system performance (e.g., POA, POR, POGI) based on game formulation.

5.1.1 Assumptions and Notations

We made the following assumptions and clarifications for the operation scenario and key parameter settings when formulating the model.

Assumptions and Clarifications for Operation Scenario

This section describes the assumptions related to the operation of the two delivery options.

- **Assumption 1:** this study only consider customers' decision behaviour regarding the service selection between MPL and ADR, which is a decision problem from a macroscopic perspective. The operation-related issues such as the stopover and route planning for MPL and ADR are out of the scope of the study since it involves applying optimisation technique to solve vehicle routing problems. It is assumed that the delivery fleet will follow the optimal scheme for parcel deliveries once customers have decided on the delivery option.

- **Assumption 2:** As parcel dimensions and weight are standardised by logistics providers, we assume that any parcel will fit the compartments of both MPLs and ADRs.

- **Assumption 3:** In terms of ADR service, customers are assumed to be able to collect their parcels on time. Each ADR only can deliver one parcel at any time from UCC to the customer given that most existing ADR products in the market have only one compartment (Jennings & Figliozzi, 2019).

- **Assumption 4:**It is assumed that customers have varying tolerance levels for different waiting time when using ADRs, the waiting cost C_{wait}^{ADR} will be formulated as a piece-wise function (Equation 5.9) considering the current mechanism employed by the practice.

Table 5.1: Notation and terminology of the model

Notation	Description	Metric
	Input Parameters of Parcel Arrivals	
X	The amount of arrival parcel	parcels/hr
λ	The time interval of the two adjacent arrival parcel	Constant[-]
	Input Parameters of Customers	
D	Customers travel distance	meter/customer
μ_t, θ_t^2	Mean value and variance of customer travel distance	Constant[-]
γ	Rationality level of customers	[0,1]
	Input Parameters of MPL Service	
μ_s	Number of parcels that MPLs serve per unit time	Constant[-]
K	Capacity of MPLs	units/MPL
Y	The expectation of parcel storage duration for MPL service	hrs/parcel
C_b^{MPL}	Unit cost of a customer selecting MPLs service when blocking occurs	unit cost/block occurs
Ω_s	Threshold of storage duration for MPL service	hrs/parcel
C_s^{MPL}	Unit storage cost if the storage duration of a parcel exceeds the threshold Ω_s	unit cost/hr
C_{travel}^{MPL}	Travel cost of a customer when selecting MPL service	units/customer
	Input Parameters of ADR Service	
q	Service charge of ADRs	unit cost/parcel
T	Time consumption of an ADR delivering parcels from depot to the customer	hr/parcel
μ_d, θ_d^2	Mean value and variance of delivery time consumption	Constant[-]
C_{w1} C_{w2}	Unit waiting cost within different ranges of customers tolerance for ADR service	unit cost/hr
C_b^{ADR}	Unit cost of customers selecting ADRs service when blocking occurs	unit cost/block occurs
	Auxiliary Variables	
p	The possibility of customers selecting MPLs service	[0, 1]
p'	The proportion of customers selecting MPLs service	[0, 1]
$C_{congestion}^{MPL}$	Total blocking cost of a customer selecting MPLs service	units/block occurs
$C_{storage}^{MPL}$	Total storage cost of a parcel	units/parcel
Φ_{MPL}	Congestion probability of MPL service	[0, 1]
Ω_w	The average tolerance of customers regarding the waiting time for ADR service	minutes
C_{wait}^{ADR}	Total waiting cost of a customer when selecting ADR service	units/customer
$C_{congestion}^{ADR}$	Total blocking cost of a customer selecting ADRs service	units/block occurs
Φ_{ADR}	Congestion probability of ADR service	[0, 1]
$\bar{U}_{MPL}(p)$	Payoff of MPL service when MPL selection probability p applied	costs
$\bar{U}_{ADR}(p)$	Payoff of ADR service when MPL selection probability p applied	costs
$\bar{U}(p, p')$	The average system payoff when a customer select MPL service with probability p	costs
	Decision Variables	
$p*$	The optimal probability that guarantees the optimal system payoff	[0,1]

Clarifications for Key Parameter Settings

This section describes the probability distribution used for determining key parameters.

- **Assumption 1:**Since urban consolidation centres receive highly frequent requests involving standard and same-day deliveries, we construct both MPL and ADR services as queue systems to mimic the peak arrival of parcels within a given time period, which follows a Poisson distribution $X \sim P(\lambda)$ that is commonly used in service queue problems.

- **Assumption 2:**Each parcel is served either by MPLs or ADRs. The congestion possibility Φ_{MPL} and Φ_{ADR} of the service queue is captured by Erlang-B model, which is a classic pattern used for modelling service queue and communication channel blocking (Hayel et al., 2016).

- **Assumption 3:**For the parcel served by MPLs, the storage duration that occupies a single compartment follows an exponential distribution $Y \sim exp(\mu_s)$ since a customer may not collect the package immediately and will occupy a compartment for a certain duration that satisfies the characteristics of exponential distribution.

- **Assumption 4:**To better fit the real-world operations, we assume that customers have varying travel distances for parcel pick-up in which the distance follows a normal distribution $D \sim N(\mu_t, \theta_t^2)$, because most customers' walking distance is concentrated in the medium level, neither too long nor too short (Lemke et al., 2016; Chen et al., 2016).

5.1.2 Agents and Action Space

We define the agents as customers with binary action spaces where they could either select MPLs or ADRs to perform parcel delivery tasks.

For the MPL service, each MPL can store a maximum of K parcels at any time. In the event that the MPL compartments are fully occupied, the upcoming customers will incur a congestion cost if they still opt to select the MPL service. In practical terms, the congestion cost can arise from the expenses associated with arranging additional MPLs to accommodate the excess demand.

Regarding the ADR service, the delivery efficiency is represented as the time consumption T of delivering parcels from depot to customers (i.e., each ADR can deliver a maximum number of parcels $1/T$ per unit of time). In real-world operations, the delivery time consumption T varies when serving different customers due to the distance between customers' locations and the depot. Therefore, we set the delivery time consumption T of different customers to follow a normal distribution $T \sim N(\mu_d, \theta_d^2)$ considering the customers that are located in the same region have similar average delivery time consumption μ_d, albeit with a certain degree of variance θ_d due to the influence of different road networks. Additionally, since this study investigates the impact of varying levels of delivery efficiency on the overall system performance, we manipulate the average delivery time consumption μ_d with a set of values to simulate deliveries in regions situated at different distances from the depot. This is done for the purpose of conducting a sensitivity analysis. (will be discussed in Section 5.4.2, 5.4.3, and 5.4.5).

In formulating the competition between the above delivery options as a game-theoretical model, one must first validate the following hypothesis: i) whether the equilibrium achieved by the two delivery services exists and is unique in both fully rationality and bounded rationality models; ii) how to improve the system's performance through applying global information in customers' decision. Hypothesis I will be proven in Section 5.2.1 and 5.2.2 (corresponding to fully rationality model and bounded rationality model respectively). Hypothesis II will be demonstrated in Section 5.3.3.

5.1.3 Payoff Function of Delivery Services

The payoff function of the two delivery services is defined, which serves as a guideline for customer decisions.

Payoff Function for MPL Service

The MPLs service is modelled as $M/M/1/K$ system that includes a single service queue with capacity K (Hayel et al., 2016). It should be noted that the parcel storage duration is independent and identically distributed. Let p denote the possibility of customers selecting

MPLs service, thus the parcel incoming rate of the MPL service queue is $\lambda\ p$.

According to the Erlang-B model, the congestion probability $\Phi_{MPL}(\lambda,\ \mu_s, p,\ K)$ of MPLs queue is represented as follows (Equation 5.1 & 5.2):

$$\Phi_{MPL}\left(\lambda,\ \mu_s, p, K\right) = \frac{\frac{(\rho p)}{K!}^K}{\sum_{i=0}^{K} \frac{(\rho p)^K}{i!}} \tag{5.1}$$

$$\rho = \frac{\lambda}{\mu_s} \tag{5.2}$$

Equation (5.3) represents the calculation for the total congestion cost $C_{congestion}^{MPL}$, which depends on the unit blocking cost C_b^{MPL} and the congestion probability determined in Equation (5.1). Equation (5.4) denotes the storage cost $C_{storage}^{MPL}$ of a parcel, which depends on the unit storage expense C_s^{MPL} and the storage duration $(1/\mu_s)$. Equation (5.5) represents the travel cost C_{travel}^{MPL} of a customer when visiting MPL for parcel pick-up, which relates to the travel distance of customers D and the unit travel cost $(\ln D/10)$.

$$C_{congestion}^{MPL} = \begin{cases} C_b^{MPL} \cdot \Phi_{MPL}\left(\lambda,\ \mu_s, p, K\right) & \text{if queue is blocking} \\ 0 & \text{otherwise} \end{cases} \tag{5.3}$$

$$C_{storage}^{MPL} = \begin{cases} C_s^{MPL} \cdot (1/\mu_s - \Omega_s) & \text{if } 1/\mu_s > \Omega_s \\ 0 & \text{otherwise} \end{cases} \tag{5.4}$$

$$C_{travel}^{MPL} = D\frac{\ln D}{10} \quad (D \sim N(\mu_t, \sigma^2)) \tag{5.5}$$

By determining the three cost units of MPLs above, the payoff function of MPLs can be formulated as follows:

$$\bar{U}_{MPL}(p) = C_{congestion}^{MPL} + C_{storage}^{MPL} + C_{travel}^{MPL} \tag{5.6}$$

Payoff Function for ADR Service

The ADR service is modelled as an M/D/1 queue and First-In-First-Out (FIFO) policy is applied because this study aims to discuss the effect of the deterministic delivery efficiency controlled by the logistics operators. The effect of different delivery efficiency will be discussed in Section 5.4.7. Considering that customers are only permitted to select either ADRs or MPLs, the probability of choosing ADR service is the inverse of the probability of selecting MPLs $(1-p)$. Therefore, the waiting time $\bar{W}(p)$ for customers receiving parcels when the queue is stable (i.e., $\lambda(1-p)T < 1$) can be defined as follows (Hayel et al., 2016):

$$\bar{W}(p) = T + \frac{\lambda(1-p)T^2}{2(1-\lambda(1-p)T)} \tag{5.7}$$

Given that waiting time $\bar{W}(p)$ is a monotonic increasing function of p, the average tolerance Ω_w of customers regarding the waiting time for receive ADR service is determined as Equation (5.8):

$$\Omega_w = \frac{\bar{W}_{min}(p) + \bar{W}_{max}(p)}{2} = T + \frac{\lambda(1-p)T^2}{4(1-\lambda T(1-p)T)} \tag{5.8}$$

Therefore, we introduce a piece-wise function to determine the waiting cost C_{wait}^{ADR} that accounts for different levels of customer tolerance, in which the value of C_{wait}^{ADR} depends on the relationship between the benchmark tolerance Ω_w and customers' actual waiting time $\bar{W}$:

$$C_{wait}^{ADR} = \begin{cases} C_{w1} \cdot \bar{W}(p) & \text{if } \bar{W}(p) \leq \Omega_w \\ C_{w2} \cdot \bar{W}(p) & \text{if } \bar{W}(p) > \Omega_w \end{cases} \tag{5.9}$$

Similar to MPL service, we assume the probability of congestion $\Phi_{\text{ADR}}(\lambda, p, T)$ follows the Erlang-B model. The congestion cost $C_{cogestion}^{ADR}$ of a parcel is presented as Equation (5.10):

$$C_{congestion}^{ADR} = \begin{cases} C_b^{ADR} \cdot \Phi_{\text{ADR}}(\lambda, p, T) & \text{if queue is blocking} \\ 0 & \text{otherwise} \end{cases} \tag{5.10}$$

$$\Phi_{\text{ADR}}(\lambda, p, T) = \frac{\frac{(\beta(1-p))^{\frac{1}{T}}}{\left(\frac{1}{T}\right)!}}{\sum_{i=0}^{\frac{1}{T}} \frac{(\beta(1-p))^{\frac{1}{T}}}{i!}} \tag{5.11}$$

$$\beta = \frac{\lambda}{T} \tag{5.12}$$

We formulate the payoff function of ADRs, which consists of service charge, congestion cost, and waiting cost:

$$\bar{U}_{ADR}(p) = q + C_{congestion}^{ADR} + C_{wait}^{ADR} \tag{5.13}$$

Average Payoff and Optimal Response Policy

The customer decision is assumed to be mutually independent. Customers can apply a mixed strategy when choosing MPLs and ADRs. The average payoff of a customer that selects MPL service with probability p over the population profile p' is defined as follows, where p' is the proportion of customers choosing MPLs:

$$\bar{U}(p, p') = p\bar{U}_{MPL}(p') + (1-p)\bar{U}_{ADR}(p') \tag{5.14}$$

The optimal response policy p^* is defined as the probability that guarantees the optimal response function $r(p, p^*)$ can achieve the equilibrium of the system:

$$p^* = r(p, p^*) = argmin_{p,p' \in [0,1]} \bar{U}(p, p') \tag{5.15}$$

5.1.4 Performance Evaluation Metrics

We introduce the notion of Price of Anarchy (POA), Prince of Rationality (POR), and Price of Global Information (POGI) to measure the performance of the proposed congestion-game

models and the effectiveness of the improved algorithm with global information applied.

Following the definition of POA, if each individual makes a decision based on minimising their own average cost (i.e., the system is considered a decentralised system), the equilibrium reached will result in higher costs than a centralised system. Therefore, the centralised system assumes that the controller possesses the complete authority to determine the proportion of individuals opting for MPLs or ADRs services to minimise the cost of the system. Under this circumstance, the customer's freedom of service selection would reduce the system's performance. Such a phenomenon is counter-intuitive in general game models but can be observed in strategic queuing problems (Hayel et al., 2016). The measurement of POA can be defined as follows:

$$POA_{full} = \frac{\max_{p \in p^*_{full}} \bar{U}(p)}{\min_{p \in [0,1]} \bar{U}(p)} \tag{5.16}$$

$$POA_{bounded} = \frac{\max_{p \in p^*_{bounded}} \bar{U}(p)}{\min_{p \in [0,1]} \bar{U}(p)} \tag{5.17}$$

where POA_{full} denotes the POA of the full rationality model while $POA_{bounded}$ denotes the POA of the bounded rationality model. A higher POA ratio indicates a less efficient system. The system achieves optimal efficiency when POA is close to 1. The implication of POA in practice is that logistics operators would expect POA approaches to 1 because it implies that the performance of the system is more stable (i.e., less likely to be influenced by the selfish decision-making behaviour of customers). Additionally, operators can decide whether incentives are needed to hedge against the selfish decision-making behaviour of customers based on POA, to achieve more cost-efficient resource allocations.

Inspired by the notion of POA, Hayel et al. (2016) introduced POR to measure the average cost ratio between fully rationality and bounded rationality model. In contrast to POA, POR does not compel customers to select their optimal service to measure the system's performance. Equation (5.18) represents the measurement of POR:

$$POR = \frac{\bar{U}(p^*_{bounded})}{\bar{U}(p^*_{full})} \tag{5.18}$$

POR enables operators to determine the gap regarding the system's performance between fully rationality and bounded rationality of customers. If such a gap is significant, logistics operators could consider disclosing more accurate service information to mitigate bounded rational decisions due to incomplete information.

We introduce a new metric, namely 'Price of Global Information', an indicator that has never been introduced to similar equilibrium models. POGI is defined as a ratio of costs resulting from applying global information in customers' decision-making process compared to those without. In other words, POGI represents the value of global information in decision-making patterns. Therefore, logistics operators can assess the enhancement in system performance resulting from the provision of global information by utilising the value of POGI. This measure allows them to determine the relative importance and effectiveness of different types of information in improving system performance. Equation (5.19) & (5.20) represents the measurement of POGI under different rationality models:

$$POGI_{full} = \frac{\min_{p \in p^*_{full}} \bar{U}^{opt}(p)}{\min_{p \in [0,1]} \bar{U}(p)} \tag{5.19}$$

$$POGI_{bounded} = \frac{\min_{p \in p^*_{bounded}} \bar{U}^{opt}(p)}{\min_{p \in [0,1]} \bar{U}(p)} \tag{5.20}$$

where $\bar{U}^{opt}(p)$ represents the average cost of the two equilibriums with global information applied. $\bar{U}(p)$ represents the average cost of the two equilibriums without global information applied.

5.2 Existence and Uniqueness of Equilibrium Models and Inverse Game Process

This section proves the existence and uniqueness of the equilibrium for both the fully and bounded rationality models. Furthermore, we propose the corresponding inverse game process for the two rationality models.

5.2.1 Fully Rationality Equilibrium

The fully rationality model is formulated as Wardrop equilibrium as discussed in the literature review section. Following the first principle of Wardrop equilibrium, the optimal equilibrium p^*_{full} exists and is unique if and only if the Equation (5.21) is satisfied.

$$\bar{U}_{MPL}(p^*_{full}) = \bar{U}_{ADR}(p^*_{full}) \tag{5.21}$$

Theorem 5.1 *For any* $\lambda,\ p, \mu_s, \Omega_s, q, C_b^{ADR}, C_b^{MPL}, C_s^{MPL}, C_w, D, T, K$, *the optimal strategy* p^*_{full} *exists and is unique to achieve the equilibrium of the system if and only if:*

(a) $C_b^{ADR} < C_w \sum_{i=0}^{\frac{1}{T}} \frac{\lambda T^2 i! T!}{2(1+\lambda T)}$

(b) $q \leq C_b^{MPL} \sum_{i=0}^{K} \frac{i!}{K!} + C_s^{MPL} \cdot \left(\frac{1}{\mu_s} - \Omega_s\right) + D\frac{lg(D)}{10}$

Proof 5.1 *To prove the existence and uniqueness of the equilibrium, Equation (5.21) should hold to guarantee that Equation (5.6) and Equation (5.13) are strictly monotonous and only have a single intersection. The proof first start with a particular value of* $P^*_{full} = 0$*:*

$$\bar{U}_{MPL}(0) = 0 \tag{5.22}$$

$$\bar{U}_{ADR}(0) = q + C_b^{ADR} \cdot \frac{\frac{(\frac{\lambda}{T})^{1/T}}{(\frac{1}{T})!}}{\sum_{i=0}^{\frac{1}{T}} \frac{(\frac{\lambda}{T})^{1/T}}{i!}} + C_w \cdot (T + \frac{\lambda T^2}{2(1-\lambda T)}) > 0 \tag{5.23}$$

The derivatives of the two payoff functions are as follows:

$$\bar{U}'_{MPL}(p) = \frac{\partial \bar{U}_{MPL}(p)}{\partial p} = C_b^{MPL} \cdot \frac{\frac{Kp^{K-1}\left(\frac{\lambda}{\mu_s}\right)^K}{K!}}{\sum_{i=0}^{K} \frac{Kp^{K-1}\left(\frac{\lambda}{\mu_s}\right)^K}{i!}} > 0 \tag{5.24}$$

$$\bar{U}'_{ADR}(p) = \frac{\partial \bar{U}_{ADR}(p)}{\partial p} = C_b^{ADR} \cdot \frac{\frac{\left(\frac{-\lambda}{T}\right)^{1/T}}{\left(\frac{-\lambda}{T}\right)!}}{\sum_{i=0}^{\frac{1}{T}} \frac{\left(\frac{-\lambda}{T}\right)^{1/T}}{i!}} - \frac{\lambda C_w T^2}{2(1+\lambda T)} \tag{5.25}$$

Therefore, $\bar{U}'_{ADR}(p) < 0$ *should hold to guarantee that Equation (5.24) & (5.25) are strictly monotone and intersect. The interrelationship between* C_{bADR} *and* C_w *should satisfy the Equation (5.26) to ensure* p^*_{full} *exists.*

$$C_b^{ADR} < C_w \sum_{i=0}^{\frac{1}{T}} \frac{\lambda T^2 i! T!}{2\,(1+\lambda T)} \tag{5.26}$$

To prove the uniqueness of p^*_{full}*, it should be ensured that* $\bar{U}'_{MPL}(p)$ *and* $\bar{U}'_{ADR}(p)$ *have at most one intersection. Considering* $\bar{U}_{MPL}(p)$ *is a monotone increasing function while* $\bar{U}_{ADR}(p)$ *is a monotone decreasing function (if Equation 5.26 is satisfied), Equation (5.27) should be satisfied to guarantee the uniqueness of* p^*_{full}*, given the condition* $p \in [0, 1]$*.*

$$\bar{U}_{MPL}(1) \geq \bar{U}_{ADR}(1) \tag{5.27}$$

Which is equivalent to:

$$q \leq C_b^{MPL} \sum_{i=0}^{K} \frac{i!}{K!} + C_s^{MPL} \cdot \left(\frac{1}{\mu_s} - \Omega_{\mathrm{s}}\right) + D\frac{lg\,(D)}{10} \tag{5.28}$$

Hence, Theorem 5.1 has been proven. Following the process derived above, we proposes the following remarks:

Remark 5.1 *For* $p^*_{full} = 1$*, all customers will select the MPLs service because the MPL service is free, and the queue is not blocked.* $p^*_{full} = 1$ *can be equilibrium as per Equations (5.22) to (5.26), which can be proven by the intersected curve of the two payoff functions. In contrast,*

$p^*_{full} = 0$ *reflects that all customers will select ADRs service. However, the system will never reach equilibrium because there always exists at least one new customer who will select MPL service due to MPL queue is not blocked, and its cost is lower than the ADRs counterpart.*

Remark 5.2 *If Equation (5.28) is not satisfied (i.e.,* $\bar{U}_{MPL}(1) < \bar{U}_{ADR}(1)$*), the equilibrium is* $p^*_{full} = 1$*, meaning that all customers select MPLs service. Otherwise,* p^*_{full} *is the unique equilibrium of the system where Equation (5.21) is satisfied.*

We introduce the inverse game process of the fully rationality model, where the decision-maker can derive the critical operating parameters (e.g., service charge) according to the expected/predefined equilibrium p^{expt}_{full} and the parameter of the system. This is practical for logistics operators since it provides a reference for a more accurate price setting with efficient capacity allocation.

Proposition 5.1 *For any* $\lambda, \mu_s, \Omega_{\mathrm{s}}, C_b^{ADR}, C_b^{MPL}, C_s^{MPL}, C_w, D, T, K$*, and* p^{expt}_{full}*, the service charge* q_{full} *of ADR service should satisfy the following equation such that equilibrium* p^{expt}_{full} *is achieved and unique:*

$$q_{full} = C_b^{MPL} \sum_{i=0}^{K} \frac{i!}{K!} + C_s^{MPL} \cdot \left(\frac{1}{\mu_s} - \Omega_{\mathrm{s}}\right) + D\frac{lg(D)}{10} - C_b^{ADR} \cdot \sum_{i=0}^{K} i!T! - C_w \cdot [T + \frac{\lambda(1-p^{expt}_{full})T^2}{2-2\lambda T(1-p^{expt}_{full})}]$$

Subject to:

$$C_b^{ADR} < C_w \cdot \sum_{i=0}^{1/T} \frac{\lambda T^2 i!T!}{2(1+\lambda T)}$$

Proof 5.2 *Since the* p^{expt}_{full} *is the expected equilibrium of the system, we obtain:*

$$\bar{U}_{MPL}(p^{expt}_{full}) = \bar{U}_{ADR}(p^{expt}_{full}) \tag{5.29}$$

By processing and simplifying Equation (5.29), we obtain Equation (5.30):

$$\begin{aligned} q_{full} = {} & C_b^{MPL} \sum_{i=0}^{K} \frac{i!}{K!} + C_s^{MPL} \left(\frac{1}{\mu_s} - \Omega_{\mathrm{s}}\right) + D\frac{lg\,(D)}{10} \\ & - C_b^{ADR} \sum_{i=0}^{K} i!T! - C_w[T + \frac{\lambda(1-p^{expt}_{full})T^2}{2-2\lambda T(1-p^{expt}_{full})}] \end{aligned} \tag{5.30}$$

Proposition 5.1 has been proven. Analogously, we can derive the corresponding value of other operating parameters (e.g., capacity), given an expected equilibrium p_{full}^{expt} and other parameters of the system. Derivation follows a similar logic as described above with the exploitation of Equation (5.29) and the p_{full}^{expt} has been proven to be unique for the system's equilibrium.

5.2.2 Bounded Rationality Equilibrium

This section exploits the Multi-Nomial Logit Model (MNL) to formulate customers' bounded rational decision behaviour as discussed in the literature review section. Consumers exhibit bounded rational behaviour when they cannot observe complete or perfect information released by the system, or when there are errors in estimating utility functions caused by perception biases. Such imperfect information includes the current length of the service queue and the overall proportion of customers selecting a particular service. Following the MNL introduced by McFadden et al. (1973) and Hayel et al. (2016), the probability p of the customer selecting MPL service can be denoted by Equation (5.31):

$$p = \frac{e^{-\gamma \bar{U}_{MPL}(p)}}{e^{-\gamma \bar{U}_{MPL}(p)} + e^{-\gamma \bar{U}_{ADR}(p)}} \tag{5.31}$$

where γ denotes the rationality level of customers ($\gamma \in [0, \infty)$). If γ tends to infinity, all customers will react rationally to the actions of others, which can be considered as a Wardrop equilibrium. If γ tends to 0, customers will behave bounded rationality. If $\gamma = 0$, customers apply the uniform mixed strategy without considering their payoff functions and $p = 0.5$.

Definition 5.1 *$r_{logit}(p)$ is the best response function of the MNL model. $p_{bounded}^{*}$ is the equilibrium of MNL for MPLs and ADRs if and only if the Equation (5.32) is satisfied:*

$$p_{bounded}^{*} = r_{logit}(p_{bounded}^{*}) = \frac{e^{-\gamma \bar{U}_{MPL}(p_{bounded}^{*})}}{e^{-\gamma \bar{U}_{MPL}(p_{bounded}^{*})} + e^{-\gamma \bar{U}_{ADR}(p_{bounded}^{*})}} \tag{5.32}$$

Theorem 5.2 *For any λ, $p, \mu_s, \Omega_s, q, C_b^{ADR}, C_b^{MPL}, C_s^{MPL}, C_w, D, T, K$, the optimal strategy $p_{bounded}^{*}$ exist and unique to achieve the equilibrium of the bounded rationality system.*

Proof 5.3 *To prove the uniqueness and existence of the MNL equilibrium, one must find a fixed point of the best-response function as Equation (5.32) defined. Let* $I(p^*_{bounded}) = e^{-\gamma \bar{U}_{MPL}(p^*_{bounded})}$ *and* $L(p^*_{bounded}) = p^*_{bounded} \cdot (e^{-\gamma \bar{U}_{MPL}(p^*_{bounded})} + e^{-\gamma \bar{U}_{ADR}(p^*_{bounded})})$*. It is noticed that:*

$$L(0) = 0 < I(0) \tag{5.33}$$

$$L(1) = e^{-\gamma \bar{U}_{MPL}(1)} + e^{-\gamma \bar{U}_{ADR}(1)} > I(1) = e^{-\gamma \bar{U}_{MPL}(1)} \tag{5.34}$$

Given that $\bar{U}_{MPL}(p)$ *and* $\bar{U}_{ADR}(p)$ *are continuous and monotonous with opposite operators,* $L(p^*_{bounded})$ *and* $I(p^*_{bounded})$ *have at least one intersection. Therefore, the existence of* $p^*_{bounded}$ *has been proven. To prove the uniqueness of* $p^*_{bounded}$*, we manipulate the Equation (5.32):*

$$\bar{U}_{MPL}(p^*_{bounded}) + \frac{1}{\gamma} ln(\frac{p^*_{bounded}}{1 - p^*_{bounded}}) = \bar{U}_{ADR}(p^*_{bounded}) \tag{5.35}$$

Since $\bar{U}_{MPL}(p)$ *is a monotone increasing function while* $\bar{U}_{ADR}(p)$ *is a monotone decreasing function, we notice that:*

$$\frac{\mathrm{d}\, ln(\frac{p^*_{bounded}}{1-p^*_{bounded}})}{\mathrm{d}p^*_{bounded}} = \frac{1}{p^*_{bounded}(1 - p^*_{bounded})} > 0 \tag{5.36}$$

Therefore, the left-hand side of Equation (5.35) is monotone increasing while the right-hand side is monotone decreasing. Thus, $L(p^*_{bounded})$ *and* $I(p^*_{bounded})$ *will only intersect at a single point. Theorem 5.2 is proven.*

Accordingly, we propose the inverse of the game model under the bounded rational setting, which gives an expected equilibrium $p^{expt}_{bounded}$ and the system's parameters, to derive the critical operational parameters (e.g., service charge) of logistics operators.

Proposition 5.2 *Given the parameter* $\lambda, \mu_s, \Omega_s, C^{ADR}_b, C^{MPL}_b, C^{MPL}_s, C_w, D, T, K, and p^{expt}_{bounded}$*, the service charge* $q_{bounded}$ *of ADR service should satisfy the following equation such that equilibrium* $p^{expt}_{bounded}$ *is achieved and unique:*

$$q_{bounded} = C_b^{MPL} \sum_{i=0}^{K} \frac{i!}{K!} + C_s^{MPL} \left(\frac{1}{\mu_s} - \Omega_s \right) + D \frac{lg\,(D)}{10} + \frac{1}{\gamma} ln(\frac{p_{bounded}^{expt}}{1 - p_{bounded}^{expt}}) \tag{5.37}$$

$$- C_b^{ADR} \sum_{i=0}^{K} i!T!C_w[T + \frac{\lambda(1 - p_{bounded}^{expt})T^2}{2 - 2\lambda T(1 - p_{bounded}^{expt})}] \tag{5.38}$$

Proof 5.4 *We utilise the interrelationship derived in Equation (5.35), which is equivalent to:*

$$\bar{U}_{MPL}(p_{bounded}^{expt}) + \frac{1}{\gamma} ln(\frac{p_{bounded}^{expt}}{1 - p_{bounded}^{expt}}) = q_{bounded} + C_b^{ADR} \sum_{i=0}^{K} i!T! \tag{5.39}$$

$$+ C_w[T + \frac{\lambda(1 - p_{bounded}^{expt})T^2}{2 - 2\lambda T(1 - p_{bounded}^{expt})}] \tag{5.40}$$

By processing and expanding Equation (5.39), we obtain Equation (5.41):

$$q_{bounded} = C_b^{MPL} \sum_{i=0}^{K} \frac{i!}{K!} + C_s^{MPL} \left(\frac{1}{\mu_s} - \Omega_s \right) + D \frac{lg\,(D)}{10} + \frac{1}{\gamma} ln(\frac{p_{bounded}^{expt}}{1 - p_{bounded}^{expt}}) \tag{5.41}$$

$$- C_b^{ADR} \sum_{i=0}^{K} i!T! - C_w[T + \frac{\lambda(1 - p_{bounded}^{expt})T^2}{2 - 2\lambda T(1 - p_{bounded}^{expt})}] \tag{5.42}$$

Hence, Proposition 5.2 is proven. Moreover, it can be concluded that the service charge $q_{bounded} > q_{full}$*, if* $p_{bounded}^{expt} = p_{full}^{expt}$*. Such extra charge can be regarded as the 'cost of bounded rationality'.*

We further discusses the relationship between the payoff function of the two equilibrium models with the exploitation of Equation (5.35):

Proposition 5.3 *Considering the full and bounded rationality of customers, we define the following proportion of consumers that select MPL service under a bounded rationality setting, compared with the case where the consumers have fully rationality:*

a) $\bar{U}_{MPL}(p_{bounded}^{*}) = \bar{U}_{ADR}(p_{bounded}^{*})$ *achieved if and only if* $p_{bounded}^{*} = 1/2$

b) $\bar{U}_{MPL}(p_{bounded}^{*}) < \bar{U}_{ADR}(p_{bounded}^{*})$ *achieved if and only if* $p_{bounded}^{*} < 1/2$

c) $\bar{U}_{MPL}(p_{bounded}^{*}) > \bar{U}_{ADR}(p_{bounded}^{*})$ *achieved if and only if* $p_{bounded}^{*} > 1/2$

Proof 5.5 *We first utilise the interrelationship derived in Equation (5.35) to prove Proposition*

5.3(a). If $\bar{U}_{MPL}(p^*_{bounded}) = \bar{U}_{ADR}(p^*_{bounded})$ *is achieved, then* $\frac{1}{\gamma} ln(\frac{p^*_{bounded}}{1-p^*_{bounded}})$ *should be equal to 0. Hence, the following equation should hold:*

$$\frac{p^*_{bounded}}{1 - p^*_{bounded}} = 1 \tag{5.43}$$

Therefore, Equation (5.43) is valid only when $p^*_{bounded} = 1/2$ *is satisfied. If* $\bar{U}_{MPL}(p^*_{bounded}) < \bar{U}_{ADR}(p^*_{bounded})$ *is achieved, then the following equation should hold:*

$$ln(\frac{p^*_{bounded}}{1 - p^*_{bounded}}) < 0 \tag{5.44}$$

Hence, $\frac{p^*_{bounded}}{1-p^*_{bounded}} < 1$ *should hold in which* $p^*_{bounded} < 1/2$ *is required. Proposition 5.3(b) has been proven. With the same analysis, we can prove Proposition 5.3(c).*

5.3 Solution Algorithms and Inverse Game Process

Figure 5.1 illustrates the implementation framework of the proposed system, which contains two parts: a congestion-game model for the service selection and a KSVM-based inverse game mechanism that incorporates the equilibrium derived from the previous service selection model.

In terms of the game model for service selection (as illustrated in Section 5.1), we first proposes two solution algorithms (see Section 5.3.1 and 5.3.2) for the fully and bounded rationality models. The algorithms consider the selfish behaviour of individuals where customers select services by minimising their cost (as described in Section 5.2.1 & 5.2.2). We then proposes a novel algorithm that applies global information to support customers' decisions (Section 5.3.3). The input of the service selection game model is the operating parameters and the predefined game mechanisms, while the output is the equilibrium of different rationality models with or without global information applied.

The KSVM-based inverse game mechanism exploits the operating parameters (except the service charge of ADR) and the expected equilibrium as inputs of KSVM, and outputs the service charge $q_{bounded}$ and q_{full} of ADRs. By implementing KSVM, the logistics operators can

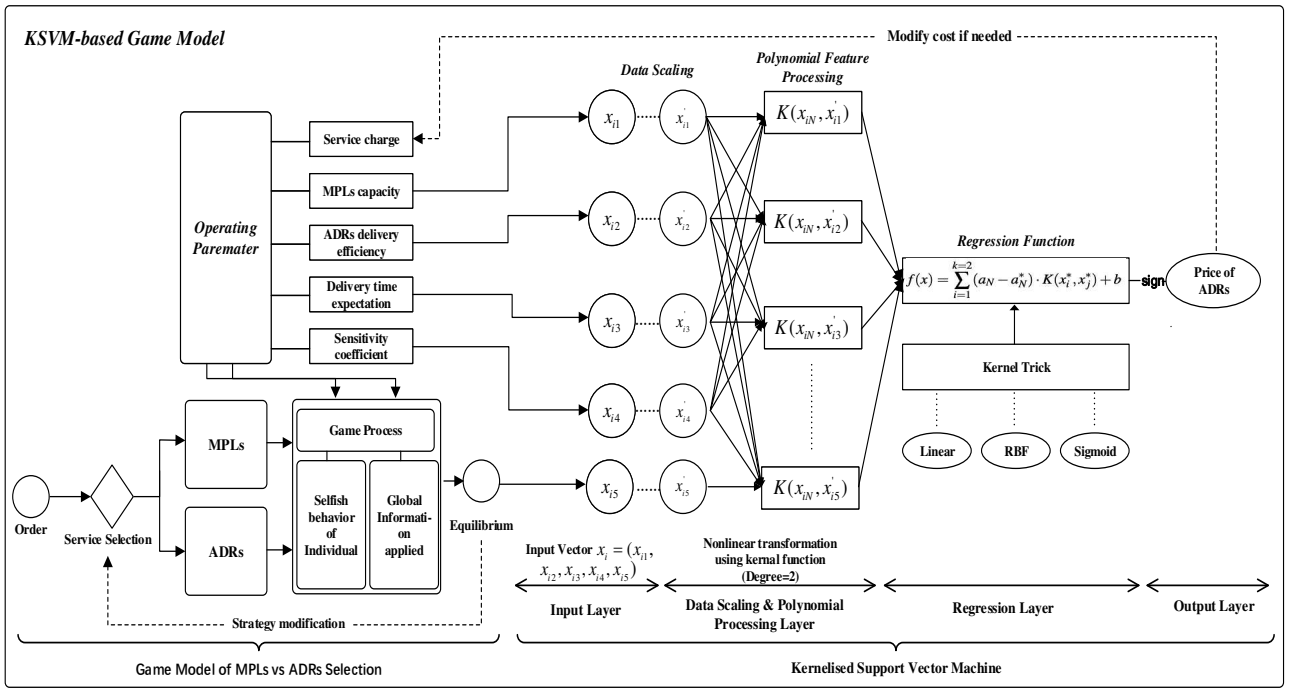


Figure 5.1: Implementation framework

predict the optimal price-setting to obtain the expected market share. The detailed procedure will be illustrated in Section 5.3.4.

5.3.1 Solution Algorithm for Fully Rationality Model

Algorithm 1 outlines the implementation logic of the fully rationality equilibrium model. We first establish the payoff functions of MPLs (Equation 5.6) and ADRs (Equation 5.13). The decision principle of customers is minimising their own costs. The probability p of customers selecting MPLs is then calculated and updated by the cumulative percentage of customers who have chosen MPLs so far. The algorithm then record and update the maximum and minimum payoff at each timestep along with the corresponding action profile to evaluate system performance. Finally, the equilibrium is derived based on Equation (5.15).

5.3.2 Solution Algorithm for Bounded Rationality Model

In contrast to the fully rationality model where customers make decisions strictly based on minimising their payoff function, customers in the bounded rationality model make decisions

Algorithm 1: Equilibrium algorithm of fully rationality model

Input: $p, \mu_s, \Omega_s, q, C_b^{ADR}, C_b^{MPL}, C_s^{MPL}, C_w, D, T, K, max_iteration, threshold$
Output: The optimal equilibrium p^*_{full}

$select_MPLs$, $select_ADRs \leftarrow$ **False**; S_{MPL}, S_{ADR}, S_p, $S_U \leftarrow \emptyset$
$p \leftarrow p \sim U(0,1)$ `// Initialise probability` p
while $i \leq max_iteration$ **do**
  Calculate $\bar{U}_{MPL}(p)$, $\bar{U}_{ADR}(p)$ according to **Equation** (5.6) & (5.13)
  if $\bar{U}_{MPL}(p) \leq \bar{U}_{ADR}(p)$ **then**
    $select_MPLs \leftarrow$ **True**
    Add $select_MPLs$ to S_{MPL} `// Record MPL selection`
    $p \leftarrow S_{MPL}.size()$ / $(S_{MPL}.size() + S_{ADR}.size())$ `// Update probability` p
    $i = i + 1$
  else
    $select_ADRs \leftarrow$ **True**
    Add $select_ADRs$ to S_{ADR} `// Record ADR selection`
    $p \leftarrow S_{MPL}.size()$ / $(S_{MPL}.size() + S_{ADR}.size())$ `// Update probability` p
    $i = i + 1$
  end if
  if $i > threshold$ **then**
    Calculate $\bar{U}(p)$ based on **Equation** (5.14) `// Record the` p `and` $\bar{U}(p)$ `when the system is stable`
    Add p to S_p `// Record the possibility`
    Add $\bar{U}(p)$ to S_U `// Record the payoff`
    if $\bar{U}(p) \leq min(S_U)$ **then**
      $p^*_{full} \leftarrow argmin\ (S_p)$ using **Equation** (5.15)
      $\bar{U}_{min} \leftarrow \bar{U}\,(p)$
    end if
  end if
end while
return p^*_{full}

based on Equation (5.31), where inaccurate observations of the payoff function are captured by the customer rationality level γ. Algorithm 2 illustrates the implementation logic of the bounded rationality model.

5.3.3 Improved Algorithm with Global Information

Both of the above cases assume that customers are self-interested without considering global information, which leads to system inefficiency (e.g., resource redundancy and under-utilisation even if the system has reached equilibrium).

To resolve this limitation, we apply global information to customer decisions. In practice,

Algorithm 2: Equilibrium algorithm of bounded rationality model

Input: $\gamma,\ p, \mu_s, \Omega_s, q, C_b^{ADR}, C_b^{MPL}, C_s^{MPL}, C_w, D, T, K, max_iteration, threshold$
Output: The optimal equilibrium $p^*_{bounded}$

$select_MPLs$, $select_ADRs \leftarrow$ **False**; S_p, $S_U \leftarrow \emptyset$
$p \leftarrow p \sim U(0,1)$ `// Initialise probability` p
while $i \leq max_iteration$ **do**
  $p_{temp} \leftarrow p_{temp} \sim U(0,1)$ `// Probability generator`
  Calculate $\bar{U}_{MPL}(p), \bar{U}_{ADR}(p)$ according to **Equation** (5.6) & (5.13)
  if $p_{temp} \geq p$ **then**
    $select_MPLs \leftarrow$ **True**
    $p \leftarrow$ **Equation** (5.31) `// Update probability` p
    $i = i + 1$
  else
    $select_ADRs \leftarrow$ **True**
    $p \leftarrow$ **Equation** (5.31) `// Update probability` p
    $i = i + 1$
  end if
  if $i > threshold$ **then**
    Calculate $\bar{U}(p)$ based on **Equation** (5.14) `// Record the` p `and` $\bar{U}(p)$ `when the system is stable`
    Add p to S_p `// Record the possibility`
    Add $\bar{U}(p)$ to S_U `// Record the payoff`
    if $\bar{U}(p) \leq min(S_U)$ **then**
      $p^*_{bounded} \leftarrow argmin\ (S_p)$ using **Equation** (5.15)
      $\bar{U}_{min} \leftarrow \bar{U}\,(p)$
    end if
  end if
end while
return $p^*_{bounded}$

customers can estimate the payoff of different services based on the global information provided by logistics operators, such as the current waiting queue length of the service. For this purpose, the principle underlying customer decision-making entails considering the complete utility of the system across the historical record and evaluating the current payoff associated with the chosen service.

Specifically, the customer's decision for both equilibrium models should satisfy two conditions: a) the current cost of the selected service is the lowest among the two delivery options, and b) the current cost of the selected service should not exceed the system's minimum payoff $\bar{U}(p)$ in the prior experience, given a threshold φ ($\varphi \geq 1$). Otherwise, the customer select service follows the probability p that contributes to the minimum payoff $\bar{U}(p)$ in the prior

Algorithm 3: Improved algorithm with global information applied

Input: $\lambda, \gamma, p, \mu_s, \Omega_s, q, C_b^{ADR}, C_b^{MPL}, C_s^{MPL}, C_w, D, T, K, max_iteration, \varphi$
Output: The optimal equilibrium p^*

1: min_value, $min_index \leftarrow 0$; S_p, $S_U \leftarrow \emptyset$
2: $p \leftarrow p \sim U(0,1)$ `// Initialise probability` p
3: **while** $i \leq max_iteration$ **do**
4: **if** fully rationality model **then**
5: Compute $\bar{U}_{MPL}(p)$, $\bar{U}_{ADR}(p)$, and $\bar{U}(p)$ based on **Equation** (5.6, 5.13, and 5.14)
6: Add p to S_p `// Record the possibility`
7: Add $\bar{U}(p)$ to S_U `// Record the payoff`
8: **if** $S_U[min_index] \leq S_U[i]$ **then**
9: $min_index \leftarrow min_index$; $min_value \leftarrow S_U[min_index]$
10: **else**
11: $min_index \leftarrow i$; $min_value \leftarrow S_U[i]$
12: **end if**`// Start to apply global information`
13: **if** $\bar{U}_{MPL}(p) \leq \bar{U}_{ADR}(p)$ **and** $\bar{U}(p) \leq min_value$ **then**
14: $p \leftarrow 1$ `// Select MPL`
15: **else if** $\bar{U}_{MPL}(p) > \bar{U}_{ADR}(p)$ **and** $\bar{U}(p) \leq min_value$ **then**
16: $p \leftarrow 0$ `// Select ADR`
17: **else**
18: $p \leftarrow S_p[min_index]$ `// Select` p `of the minimum cost`
19: **end if**
20: Execute line 16 to line 24 of **Algorithm** (1)
21: **end if**
22: **if** bounded rationality model **then**
23: Execute line 4 to line 12
24: **if** $\bar{U}(p) \leq min_value * \varphi$ **then**
25: $p \leftarrow$ **Equation** (5.31)
26: **else**
27: $p \leftarrow S_p[min_index]$
28: **end if**
29: Execute line 15 to line 23 of **Algorithm** (2)
30: **end if**
31: **end while**
32: **return** p^*

experience as defined by Equation (5.14). Algorithm 3 demonstrates an improved algorithm with global information applied.

5.3.4 Regression Prediction through KSVM

In the game models presented in Section 5.2, it is assumed that the system's operating parameters are predetermined, allowing for the derivation of equilibrium using the

corresponding game mechanism. In the inverse game model based on KSVM, the inputs consist of the system's operating parameter and the expected equilibrium (Figure 5.1). The output of the KSVM is the optimal service charge q for ADRs, which ensures the attainment of the expected equilibrium. To train the KSVM model, we executes the game model with different permutations of operating parameter settings to obtain the corresponding equilibrium value of the game process, which serves as the training sample.

Since the prediction accuracy of KSVM depends on the consistency of input data, (i.e., the metric between different input data will affect the prediction accuracy), we first normalises the input data through $Z - Score$ normalisation. Let x_{iN} denotes the i^{th} input data of feature N ($N = \{\mu_s, \Omega_s, q, C_b^{ADR}, C_b^{MPL}, C_s^{MPL}, C_w, D, T, K\}, i = 1, 2, \ldots, |N|$). Therefore, the normalised i^{th} input data of feature N can be represented as follows, where $\bar{x}_N$ denotes the mean value of all input data for feature N:

$$x'_{iN} = \frac{x_{iN} - \bar{x}_N}{\sqrt{\frac{\sum_{i=1}^{|N|}(x_{iN} - \bar{x}_N)^2}{|N|}}} \tag{5.45}$$

Let x_N^* presents a set of all normalised data for feature N (i.e., $x'_{iN} \in x_N^*$). To enhance the accuracy of predictions, polynomial feature processing is employed to generate higher-dimensional features that amplify the significant interrelationship between two features. We proposes to transfer the normalised input data to higher dimensional terms with the degree of 2, because the model may suffer from over-fitting if the higher-dimensional features (degree>2) are applied. The input data will expand to:

$$N^2 = \{1, x'_{i1}, \ldots, x'_{iN}, (x'_{i1})^2, x'_{i1}x'_{i2}, (x'_{i2})^2, x'_{i1}x'_{i3}, (x'_{i3})^2, \ldots, (x'_{iN})^2\} \tag{5.46}$$

The principle of applying KSVM for regression prediction is to determine an optimal partition hyperplane $\omega x_N^* + b = 0$ (ω, b represent slope and intercept respectively) for a data set, such that minimising the gap between the two data points which are farthest to the hyperplane (as illustrated in Figure 5.2 and Formula 5.47). The mathematical model of KSVM is formulated as follows:

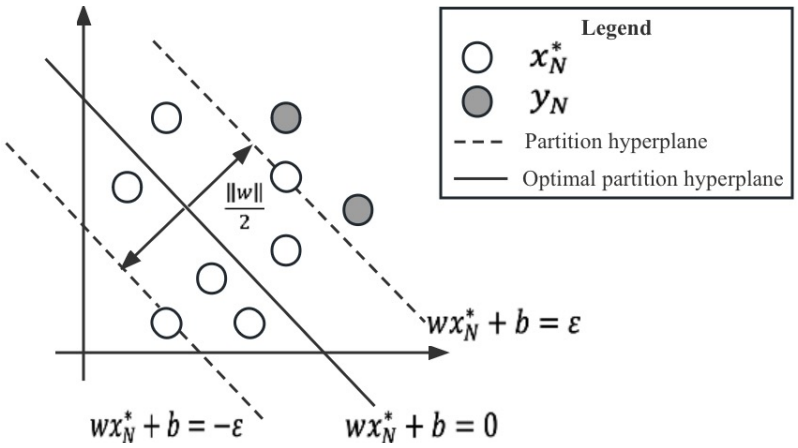


Figure 5.2: Illustration of KSVM principle

$$\min_{w,b} \frac{1}{2} \parallel w \parallel^2 \tag{5.47}$$

s.t.

$$|y_N - (wx_N^* + b)| \leq \epsilon \tag{5.48}$$

where y_N denotes the data points of feature N that within the boundary of hyperplane. Constraint (5.48) ensures that the gap between y_N and the optimal partition hyperplane is less than the error threshold ϵ.

To implement KSVM, kernel functions are used to map the normalised input data of different features to higher dimensional feature space through non-linear transformation. Three typical kernel functions are applied for regression prediction:

$$K(x_i^*, x_j^*) = \begin{cases} [(x_i^{*T} \cdot x_j^*)]^k & \text{Polynomial} \\ exp(-\frac{\|x_i^{*T} - x_j^*\|^2}{2\sigma^2}) & \text{RBF} \\ tanh(vx_i^{*T} \cdot x_j^* + c) & \text{Sigmoid} \end{cases} \tag{5.49}$$

where x_i^*, x_j^* represent the normalised input data of feature i and j $(i \neq j, i, j \in N)$. k, σ, v, c is the built-in parameter of the corresponding kernel function that affects the generalisation ability. We input the higher dimensional feature to the regression layer to derive the regression

value based on the following equation:

$$f(x) = \sum_{i=1}^{k=2}(a_N - a_N^*) \cdot K(x_i^*, x_j^*) + b \tag{5.50}$$

where a_N and a_N^* represent general and the optimal variable of the optimal partition hyperplane, respectively. b denotes the intercept of the optimal partition hyperplane. Hence, $(a_N - a_N^*)$ represents the weights of each feature. Finally, Equation (5.51) is applied to activate Equation (5.50).

$$s(x) = \sum_{N=1}^{N} a_N \cdot y_N \cdot K(x_i^*, x_j^*) + b \tag{5.51}$$

5.4 Computational Results and Discussions

We conduct event-based simulations for the two rationality models, which are evaluated based on six criteria: a) the result of the game process in terms of MPLs and ADRs service selection; b) the effect of the single operating parameter on system performance; c) the dynamic of equilibrium models considering the interaction of varying operating parameters; d) the effectiveness of the improved algorithm on performance improvement; e) the POGI under different rationality models, and f) regression prediction for the service charge through KSVM. The models are implemented in AnyLogic 8 professional version on an Apple M1 (3.20 GHz) processor with 16GB RAM.

Table 5.2 presents the benchmark parameter setting of the experiment scenario. The rationality level γ was set to capture the uniform variation of the evaluation metric. The maximum capacity K of MPLs was determined based on the study conducted by Perboli et al. (2018), which indicated that small to medium-sized MPLs are more suitable for sidewalk operations. The average storage duration $1/\mu_s$ of parcels in MPLs was set to reflect the scenario where customers collect their packages at the end of the day while the parcels are stored in MPLs during working hours. This value was informed by the findings of Kedia et al. (2017) and Iwan et al. (2016). The storage threshold Ω_s was determined considering that MPLs primarily

Table 5.2: Benchmark parameter setting of equilibrium models

Notation	Value/Metric	Notation	Value/Metric
X	$X \sim P(\lambda = 600)$ parcels/hr	γ	[1, 0.9, 0.7, 0.5, 0.3]
K	18 units/MPL	Y	$Y \sim exp(\mu_s = 0.15)$ hrs/parcel
C_b^{MPL}	10 unit cost/block occurs	C_s^{MPL}	5 unit cost/hr
Ω_s	8 hrs/parcel	D	$D \sim N(\mu_t = 250, \theta^2 = 50)$ m/customer
q	5 unit cost/ADR	T	$T \sim N(\mu_d = 0.4, \theta_d^2 = 0.2)$ hr/parcel
C_b^{ADR}	10 unit cost/block occurs	C_{w1},C_{w2}	10, 15 unit cost/hr
$max_iteration$	1100 (-)	$threshold$	1000 (-)

Note: each iteration simulates 1 hour in real-world operation

serve the purpose of storing parcels during working hours (9:00 to 17:00). The walking distance D for customers to collect their parcels were determined based on empirical research conducted by Kedia et al. (2017). The delivery efficiency of ADR that departs from urban consolidation centres to customers lies between 0.5 to 0.2 hr/parcel (Jennings & Figliozzi, 2019; Kapser & Abdelrahman, 2020). The following sections will investigate the effect of different efficiency on the system. Each iteration simulates 1 hour in real operation.

5.4.1 Benchmark Result of the Game Process

To present the game process, queue-based visualisation is used since we formulate the service selection between MPL and ADR as a congestion-game model adapted from queue theory. The computation time of the simulation process is 65 seconds.

Figure 5.3 illustrates the waiting queue length in real-time for two equilibrium models. As the simulation duration extends indefinitely, the waiting queue length approaches zero, as all customers are eventually served once equilibrium is reached. However, customers demonstrating bounded rationality exhibit slower convergence rates compared to those displaying full rationality, attributed to inaccurate cost estimations or limited service information disclosed. Consequently, an increase in the ADR queue length occurs as parcel volumes rise, as indicated by the divergence in the latter portion of the red curve ($\gamma = 0.9$). Moreover, the substantial disparity between the dark red curve (representing the full rationality model for ADRs) and the light red curve (representing the bounded rationality model for ADRs) underscores the considerable influence of customers' bounded rationality on ADR service selection. Conversely,

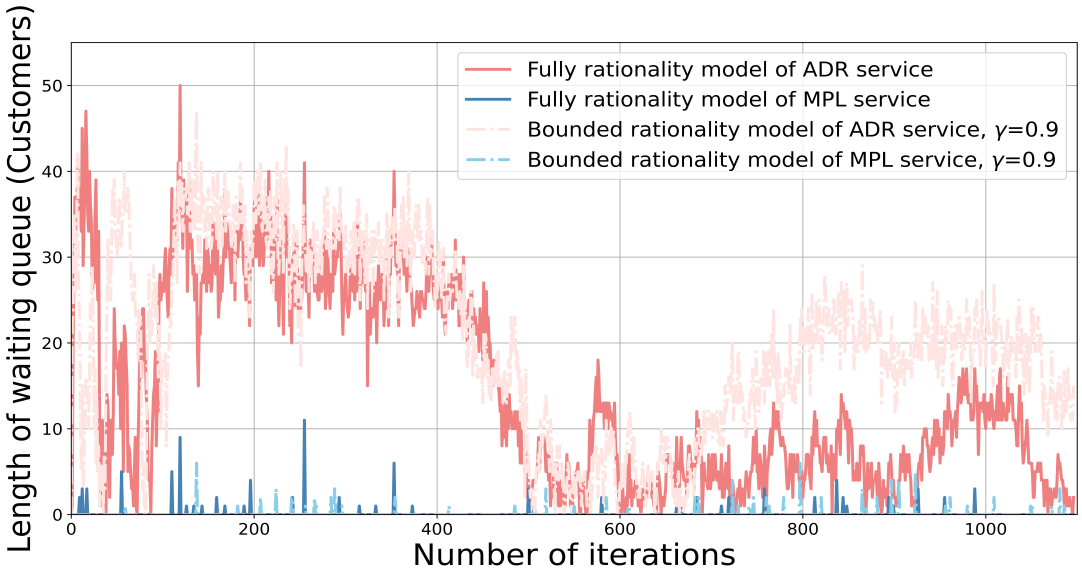


Figure 5.3: Real-time queue length of MPL and ADR services

the waiting queue for MPL service is notably shorter than that for ADR service, owing to its enhanced service capacity and cost efficiency.

5.4.2 Effect of Single Operating Parameter on System Performance

Examining the influence of operating parameters on system performance is useful for logistics operators as it provides insight into how customers make delivery choices under varying operating conditions, thus revealing the dynamics of system performance. The key operating parameters can be grouped into the following categories: (a) MPL capacity; (b) customers' parcel storage duration in MPL services; (c) the average distance travelled by customers when using MPL services; (d) the service charge of ADRs; and (e) the delivery efficiency of ADRs. The results of the experiments are presented in Appendix A.

Figure B.1a indicates that customers are more likely to select MPLs as MPL capacity increases. Specifically, the equilibrium p^* of different rationality models equals 1 when the capacity $K \geq 30$ units/MPL. The higher irrationality results in a lower equilibrium than the fully rational model, meaning that MPL operators must invest more capacity to capture a larger market share. Figure B.1b reveals that customers' bounded rationality results in a smoother trend in POA. Bounded rationality models, however, have a greater POA than the fully rationality model when $K \geq 17(\pm 1)$ units/MPL. Furthermore, the POA increases with

the irrational level of customers. This suggests that when MPL capacity is sufficient, logistics operators typically expect customers to rationally use MPLs to optimise resource allocation. Finally, customers' irrational behaviour leads to a broader POR gap when $K \leq 20$ units/MPL, indicating that the difference between the current utility and the optima is significant. However, the gap will be improved or eliminated when capacity $K > 20(\pm 2)$ units/MPL.

The shorter occupancy duration of the compartments will improve MPL utilisation because more available compartments can serve customers (Figure B.1c). The equilibrium p^* of different rationality models equal 1 when the expectation of storage duration $\mu_s \geq 1/exp(0.3)$ hrs/parcel (storage duration $\leq exp(0.3)$ hrs/parcel). Figure B.1d represents that reducing the average storage duration will improve system efficiency. Specifically, $POA_{bounded} \geq POA_{full}$ when the expectation of storage duration $\mu_s \geq 1/exp(0.18)$ hrs/parcel. On the other hand, the lower level of customer rationality results in a larger POA for the system with the average storage duration increases. Given these concerns, the MPLs will be more efficiently utilised if customers refrain from over-occupying the compartments. Higher irrationality levels and longer storage duration cause a significant utility gap (POR). However, POR will be improved when storage duration $\leq exp(0.2)$ hrs/parcel.

The increase in customer travel distance generates a lower model equilibrium since customers expect a higher degree of convenience from MPL (Figure B.1e). MPL service has less attraction to customers when travel distance $D \geq 450(\pm 20)$ meters/customer. Our result suggests that the travel distance $D \in [200, 300]$ meters/customer has the potential to improve system equilibrium. Figure B.1f indicates that the increase of customer maximum travel distance will significantly improve the system efficiency when travel distance $D \leq 230$ meters/customer. Regarding the tendency of POR, the system with lower levels of customer rationality improves POR over shorter travel distances (e.g., $D \leq 500$ meters/customer). Further, we observe that bounded rationality models generate lower average costs than the fully rationality counterpart (since POR≤ 1). A similar paradox can also be observed in some routing games (Braess, 1968). In this case, providing redundant information to the customer is inefficient and will not improve their decision-making pattern since the accuracy of that information cannot be guaranteed (error-free), thus leading to higher costs (Hassin & Haviv,

2003).

Figure B.1g represents that the model equilibrium increases with ADR service charge. There is a gradual increase in the slope of the curve towards the end compared to other operating parameters. The reason for this is that customers are more price sensitive when compared to MPL services. Therefore, customers have higher expectations for the delivery efficiency of paid/higher charged ADR services. While Figure B.1h reveals that the increase in the service charge of ADRs leads to the worse performance of the system (since POA becomes greater). There is a significant fluctuation of the curves for the fully and bounded rationality models with $\gamma = 0.9$ than for other bounded rationality models. This is due to the fact that the extra charge for ADR service will compromise its attractiveness for customers since most customers will choose charge-free MPLs instead. In terms of the full rationality model and bounded rationality models with $\gamma = 0.5,\ 0.7$ & 0.9, POR=1 is achieved. This shows that the service charge of ADRs has less effect on the system utility in the case of bounded rationality models.

Figure B.1i illustrates that the equilibrium decreases with the increasing delivery efficiency of ADRs. Customers with lower levels of rationality have pronounced negative impacts on the system's equilibrium, given that higher delivery efficiency is attractive to customers selecting ADRs. Specifically, the system's equilibrium decreases when delivery efficiency exceeds 6 units/hr. Figure B.1j reveals that the increase in ADRs' delivery efficiency will improve the system's POA once delivery efficiency exceeds 6 units/hr. Additionally, the system with lower levels of customer rationality improves POR significantly with less delivery efficiency invested in. The threshold of POR improvement for $\gamma = 0.9$, 0.7, 0.5, and 0.3 are 10, 9, 8, and 7 units/hr, respectively. We observe that all bounded rationality models generate a lower average cost than the fully rationality model (since POR ≤ 1).

Lastly, it is concluded that the storage cost of MPLs only has a minor impact on the equilibrium of the system (Figure B.1k). As an explanation for this phenomenon, the storage charge does not apply to all customers who select MPL's service but only to those whose parcels are kept beyond the free storage threshold. Hence, the storage cost of MPLs has no significant impact on customer service selections. We also noticed that POA_{full} maintains at 211.16 (Figure B.1l), while the POA of the system deteriorates as storage costs increase in bounded

rationality models. However, the $POA_{bounded} \leq POA_{full}$ if the storage cost $C_s^{MPL} \leq 3$ unit cost/day.

The above results indicate that customer travel distance and the average storage time of parcels are important factors for improving the utilisation of MPLs. In contrast, service charge and delivery efficiency are the main factors that customers are willing to compromise when choosing ADR services.

5.4.3 Effect on System Considering the Interaction of Homogeneous Operating Parameters

To gain a deeper understanding of how the system is influenced by the interaction between various operating parameters, we examine the relationships between different parameters. We identify two parameter pairings for MPL and ADR services: (i) the service charge of ADRs compared to the storage cost of MPLs (referred to as service charge properties), and (ii) the delivery efficiency of ADRs compared to the capacity of MPLs (referred to as capacity properties). This analysis will help operators determine which homogeneous operating parameters have a more significant impact on the system.

To verify the homogeneity of MPL and ADR attributes under each parameter-based pairing, we first performed correlation tests on POA and POR for each of the two pairings (see Appendix C). The results indicated that the corresponding attributes of MPL and ADR services under the same parameter-based pairings have a significant correlation in POA. In the correlation test regarding POR, the correlation of attributes under different parameter pairings was shown to be weakly correlated, influenced by the smaller differential band in POR (POR $\in [0.75, 1.4]$). However, based on the results shown in Figure B.1b, B.1h, B.1j, and B.1l, the attributes under different parameters pairings are considered as homogeneous because their corresponding POR values are close.

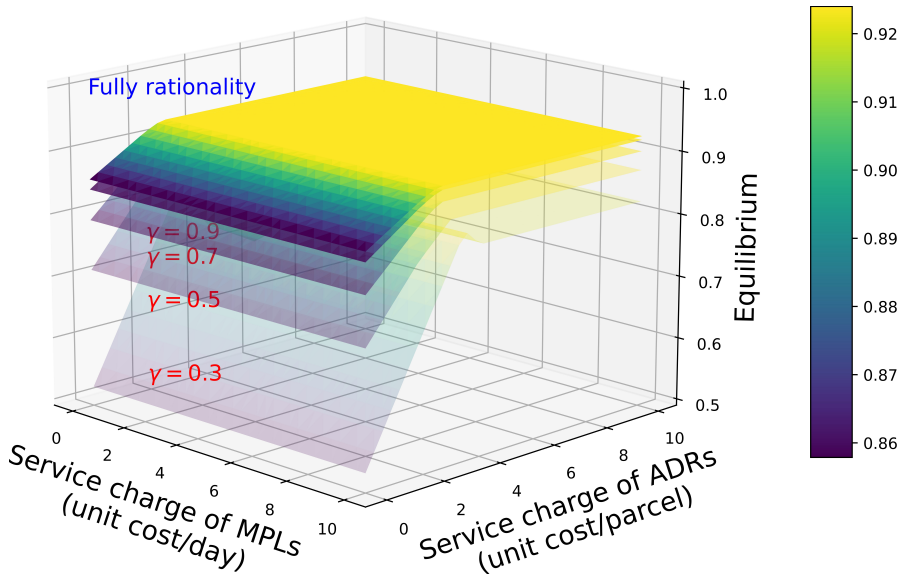


Figure 5.4: Equilibrium of varying customers' rationality under service charge properties

Effects of Service Charges Against System Performance

The service charge of ADRs has a more significant effect on the equilibrium compared with MPL service (Figure 5.4). In contrast, the storage cost of MPLs has almost no influence on the equilibrium, which is consistent with the results shown in Figure B.1i and Figure B.1k. Additionally, the higher level of rationality of customers leads to a narrower range of equilibrium attainable, meaning that customers' choice is becoming more concentrated on MPL service. The interval range of equilibrium is as follows (with descending order of customers' rationality): 0.86-0.92 (full rationality), 0.84-0.92 ($\gamma = 0.9$), 0.79-0.90 ($\gamma = 0.7$), 0.71-0.87 ($\gamma = 0.5$), and 0.52-0.82 ($\gamma = 0.3$).

Figure 5.5 represents the average POA under different levels of customers' rationality. The results indicate that as customer rationality decreases, the system performance declines and the performance difference within the same group becomes more pronounced. Concretely, Figure D.1a to D.1e in Appendix D represents the heatmap of POA for different levels of customers' rationality. The more pronounced the colour disparity, the less significant difference in the POA within the rationality model. We find that the changes in service charge of ADRs q have a mild impact on POA when customers behave more rationally (e.g., full rationality and bounded rationality $\gamma = 0.9$), and the extra storage charge of MPLs C_s^{MPL} has no significant effect on POA. However, for the low rationality models (e.g., $\gamma = 0.5, 0.3$), the POA increases

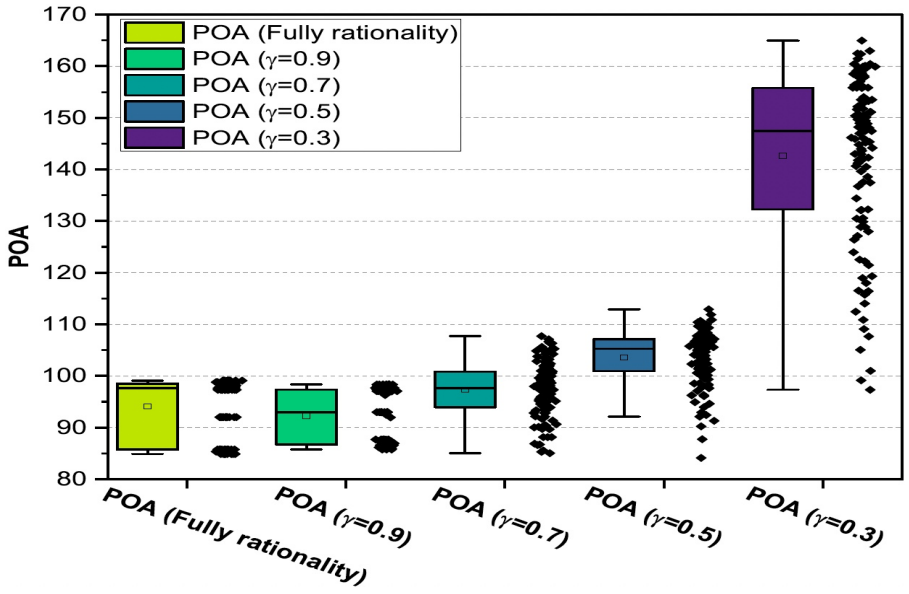


Figure 5.5: POA under different levels of customers' rationality

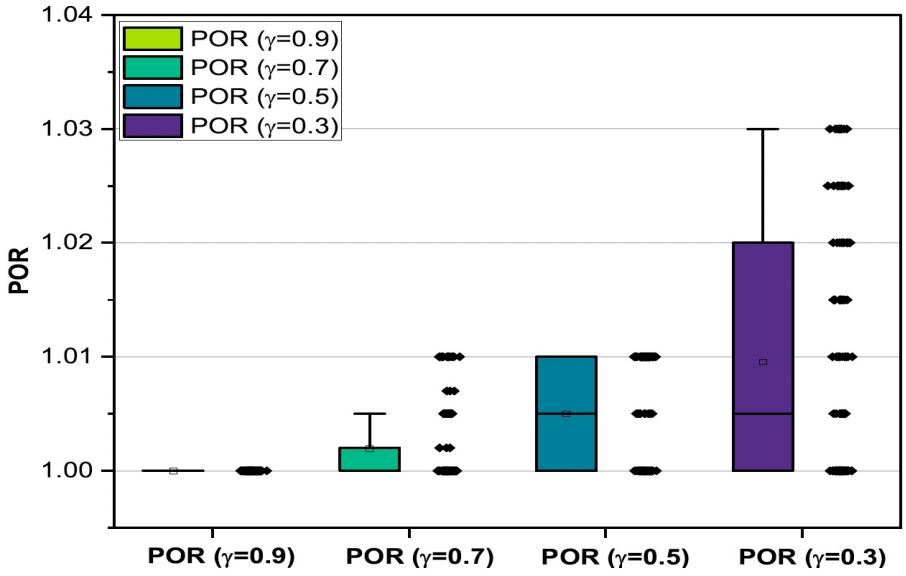


Figure 5.6: POR under different levels of customers' rationality

with C_s^{MPL}. When a certain level of bounded rationality exists ($\gamma = 0.9$), the system becomes more centralised to improve its performance, which confirms that customers' greater freedom of service selection would result in decreased efficiency.

The result in Figure 5.6 shows that the average POR increases as customer rationality decreases but the overall POR of all samples is close to 1. This suggests that the interaction of MPLs and ADRs service charges do not result in a noticeable difference in the optimal average cost between bounded rationality and fully rationality models. Specifically, Figure D.2a to D.2d in Appendix D indicates that the increasing service charge q of ADRs leads to larger POR while the service charge of MPLs C_s^{MPL} has not significant influence on POR.

In summary, from a practical standpoint, logistics operators can enhance the efficiency of the system by setting relatively lower service charges for MPLs and ADRs compared to conventional delivery options. The results of our research indicate that a satisfactory POA can be achieved by setting ADRs' price within the range of 1 to 4 unit costs/parcel and 0 to 2 unit costs/day for MPLs, as evidenced by the small POA in the upper left region of the heatmaps presented in Figures D.1a to D.1e. To increase the appeal of MPLs and alleviate the workload of door-to-door deliveries, it is recommended that the storage cost of MPLs be lower than that of ADRs.

Effects of Capacities Against System Performance

The results indicate that the delivery efficiency of ADR service has a more pronounced effect on the equilibrium compared to the capacity of MPL service. The equilibrium is seen to increase with the delivery efficiency of ADRs (Figure 5.7). When the delivery efficiency of ADRs is below 6 units/hr, the equilibrium is found to increase with the capacity of MPLs, suggesting that MPLs are used as an alternative when ADR service are high in volume. Customers prefer ADR service due to its convenience and lower costs (because higher delivery efficiency would result in lower blocking costs and operating costs) if the capacity of the two delivery options is only slightly different. Therefore, when the delivery efficiency of ADRs exceeds a threshold (e.g., 6 units/hr), the equilibrium cannot be significantly affected by MPLs even though MPLs' capacity increases. Furthermore, a higher level of customer rationality results in a greater equilibrium. The average equilibrium is 0.56 (full rationality), 0.52 ($\gamma = 0.9$), 0.45 ($\gamma = 0.7$), 0.36 ($\gamma = 0.9$), and 0.22 ($\gamma = 0.3$) respectively.

Both the capacity of MPL service and the delivery efficiency of ADRs have a significant impact on the POA within each level of customer rationality (Figure 5.8). The heatmap of POA for various levels of customer rationality can be seen in Figures E.1a to E.1e in Appendix E. It was found that for rationality models with higher levels of customer rationality (e.g., $\gamma \geq 0.7$), the POA improves as the rationality level decreases. However, an excessively low level of rationality ($\gamma \leq 0.5$) has a negative effect on the POA. This can be observed from the average POA under different rationality levels: 42.42 (full rationality), 36.96 ($\gamma = 0.9$), 36.99 (γ

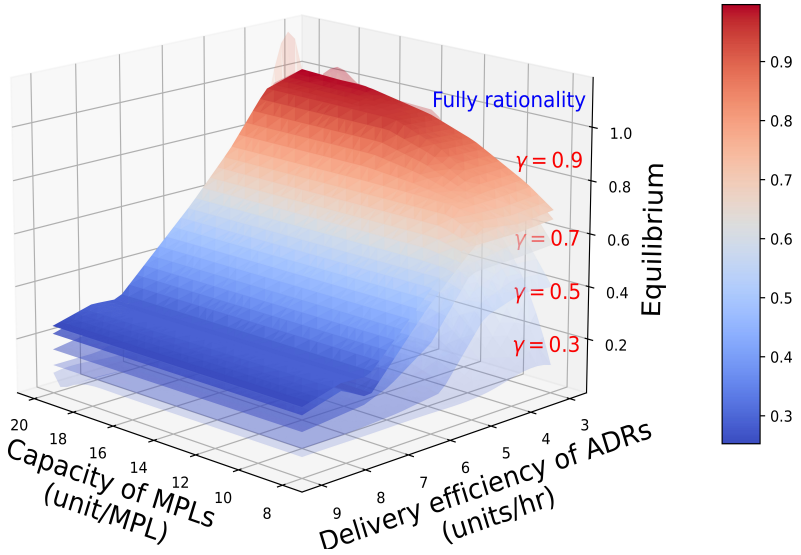


Figure 5.7: Equilibrium of varying customers' rationality under capacity properties

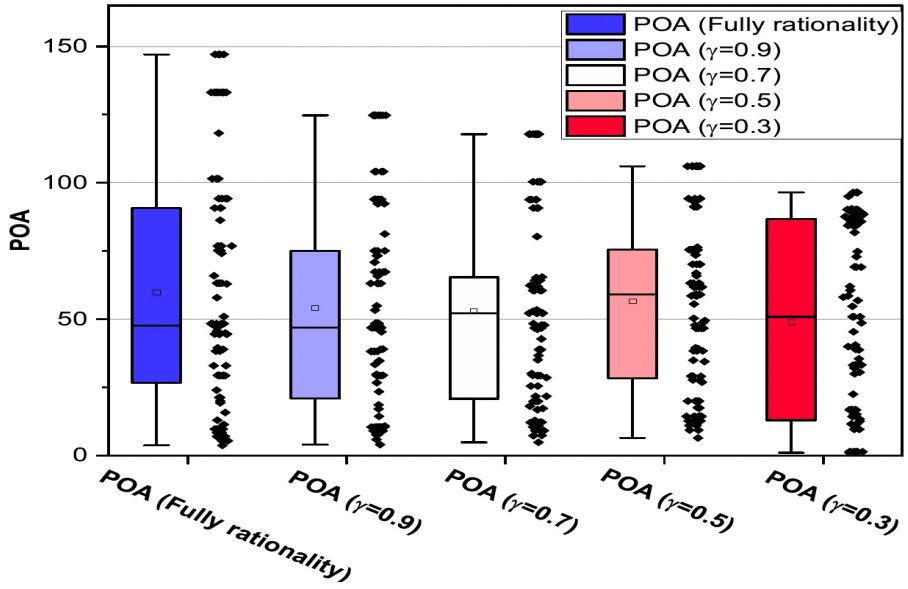


Figure 5.8: POA under different levels of customers' rationality

$= 0.7$), 41.17 ($\gamma = 0.5$), and 49.63 ($\gamma = 0.3$) respectively. On the other hand, the POA decreases with an increase in the delivery efficiency of ADRs if the capacity of MPLs $K < 10$ units/MPL. This is because the lower delivery efficiency of ADRs will cause more customers to choose MPL service, even if its capacity is relatively limited, and such extreme selection will have a negative impact on the performance of the system. As the customer rationality decreases, the difference in POA under the two parameters becomes less significant, as the boundaries of the heatmap's colours become more distinguishable.

As the level of customer rationality decreases, the POR exhibits a downward trend (Figure

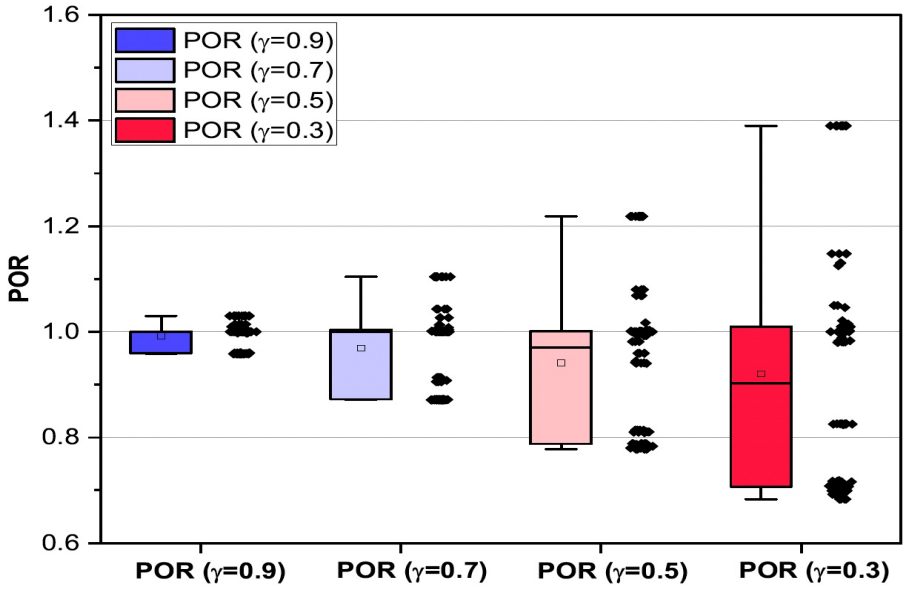


Figure 5.9: POR under different levels of customers' rationality

5.9). This implies that the performance of the system can be improved if customers make decisions that are not solely based on self-interest maximisation. However, the difference in POR within the same rationality model becomes more pronounced. Figure E.2a to E.2d in Appendix E reflect that the POR decreases as ADRs' delivery efficiency improves, meaning that the delivery efficiency of ADRs directly impacts the optimal average cost of the system. Specifically, the POR ≤ 1 when ADRs' delivery efficiency $1/T > 4$ units/hr. On the other hand, POR > 1 when both MPLs' capacity and ADRs' delivery efficiency are at a low level (MPL capacity $K \leq 8$ units/MPL, ADR delivery efficiency $1/T \leq 4$ units/hr). The average POR of different customers' rationality is as follows: 0.99 ($\gamma = 0.9$), 0.96 ($\gamma = 0.7$), 0.91 ($\gamma = 0.5$), and 0.86 ($\gamma = 0.3$).

In summary, the analysis indicates that the service charge and delivery efficiency of ADR service have a greater influence on the system's equilibrium and efficiency compared to the characteristics of MPLs. Extreme customer selection behaviour resulting from excessively high or low service charges for ADR services can lead to an imbalanced allocation of resources and sub-optimal system performance. Additionally, the results demonstrate that customers are more inclined towards higher delivery efficiency of ADRs, even when the capacity of MPLs is increased. In a decentralised system where customers aim to minimise their average cost, the equilibrium reached will yield higher costs and lower POA compared to a centralised system.

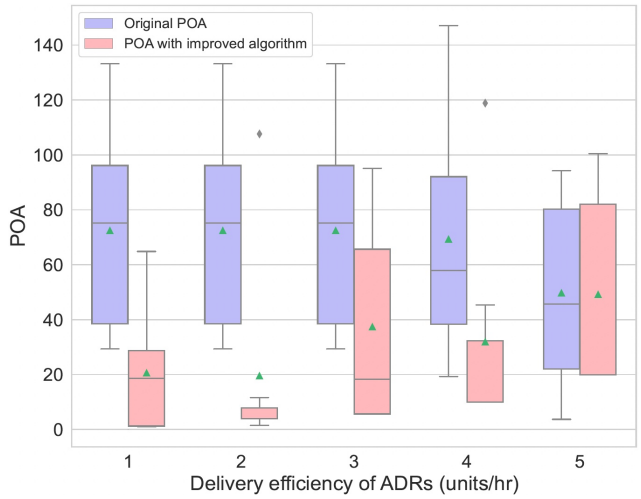
5.4.4 Performance of Improved Algorithm with Global Information

This section investigates the benefits of incorporating global information into customer decision-making. We use POA_{full} and $POA_{bounded}$ under different ADRs delivery efficiency obtained in the previous experiment as the comparison benchmark since the POA is significantly responsive to variations in delivery efficiency, as highlighted in the previous analysis in Section 5.4.2. We set the threshold $\varepsilon = \sum_{i=1}^{max_iteration} \bar{U}(p)/min_value * max_iteration$ because it is expected that the overall cost $\bar{U}(p)$ should not exceed the average cost obtained in the prior experience (Algorithm 3). However, ε can be adjusted in practice accordingly, depending on the requirement towards the system performance.

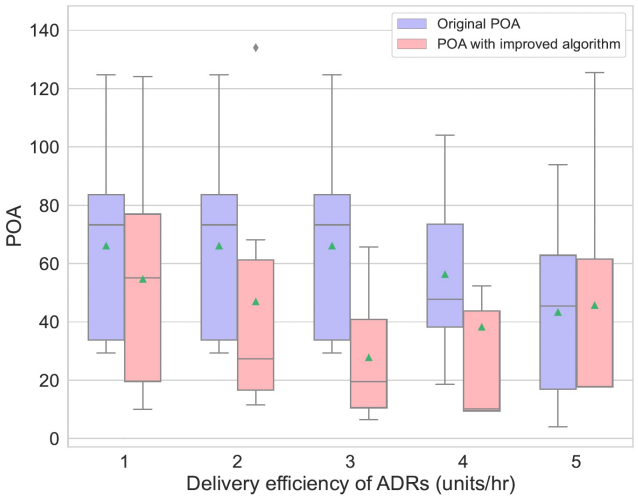
Figure 5.10a to 5.10e illustrates the POA improvement rate of different rationality models. In comparison with the scenario where customers are assumed to make decisions based solely on self-interest and without taking into consideration global information, the average improvement of POA when incorporating global knowledge into customer decisions is observed to be 48.66% (full rationality), 28.40% ($\gamma = 0.9$), 45.70% ($\gamma = 0.7$), 34.09% ($\gamma = 0.5$), and 68.70% ($\gamma = 0.3$), respectively. It is noted that the mean value of POA decreases with the increase in ADRs' delivery efficiency $1/T$, if $1/T \leq 4$ units/hr, followed by an upward trend. This finding underscores the significant enhancement in system performance achieved by integrating global information into customer decision patterns, particularly when the disparity in system performance is more pronounced in the absence of global information.

5.4.5 Price of Global Information in Different Rationality Models

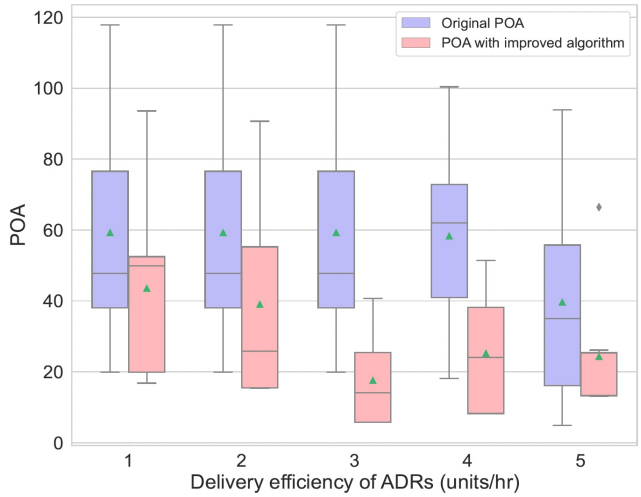
Figure 5.11a to 5.11e represent the POGI of different rationality models. Specifically, we analyse the impact of MPL capacity K and ADR delivery efficiency $1/T$ on POGI. The results show that POGI increases as ADRs delivery efficiency declines in the case of fully rationality model (Figure 5.11a) and the bounded rationality $\gamma = 0.7\&0.9$ (Figure 5.11c & 5.11b). This indicates that global information improves system efficiency when delivery capabilities are in short supply. However, there was an opposite trend of POGI in the bounded rationality model $\gamma = 0.5$ (Figure 5.11d). For the bounded rationality model $\gamma = 0.3$, POGI decreases further with the increasing



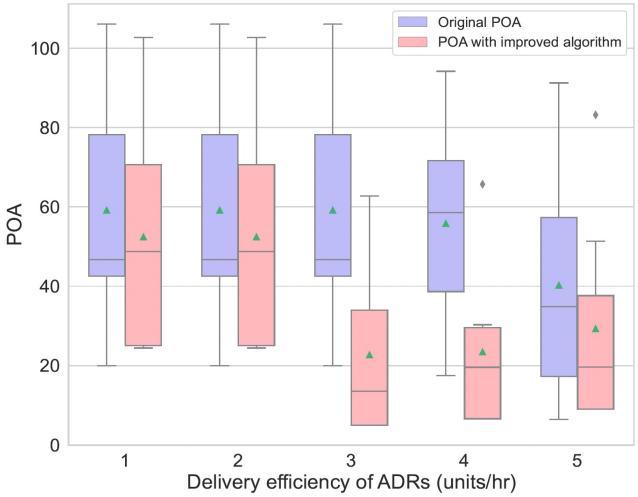
(a) POA improvement (fully rationality model)



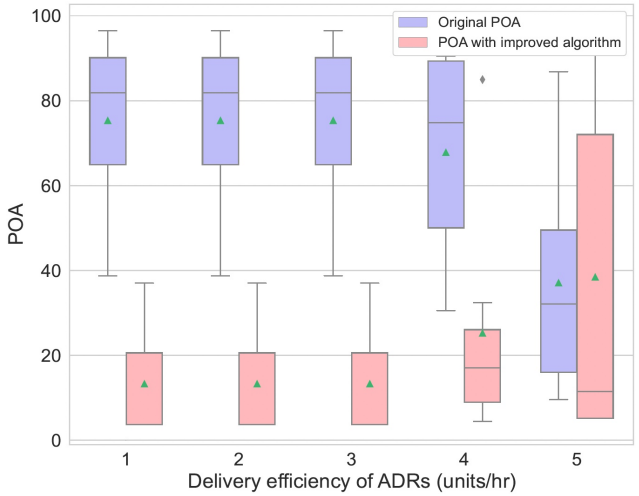
(b) POA improvement ($\gamma = 0.9$)



(c) POA improvement ($\gamma = 0.7$)



(d) POA improvement ($\gamma = 0.5$)



(e) POA improvement ($\gamma = 0.3$)

Figure 5.10: Performance improvement of different rationality models

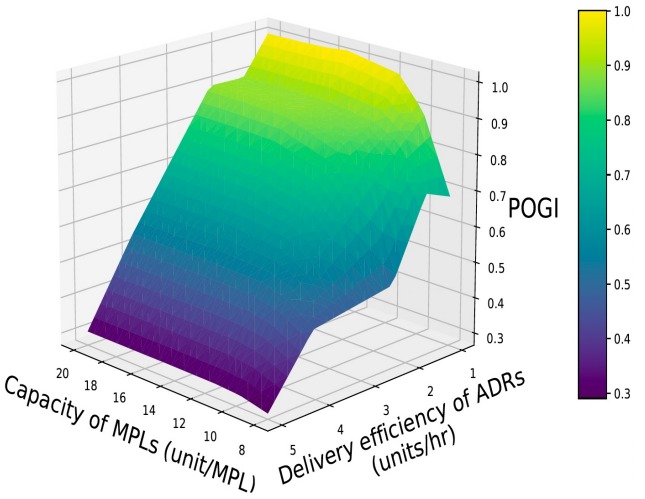
ADRs delivery efficiency and MPLs capacity (Figure 5.11e). Compared to the models with higher rationality levels ($\gamma \geq 0.7$), global information plays a more significant role in system

performance in low rationality models. The average POGI of different rationality models is 0.66 (full rationality), 0.63 ($\gamma = 0.9$), 0.63 ($\gamma = 0.7$), 0.55 ($\gamma = 0.5$), and 0.71 ($\gamma = 0.3$), respectively. These results suggest that incorporating the proposed improved algorithm with global information significantly improves the average cost of the system.

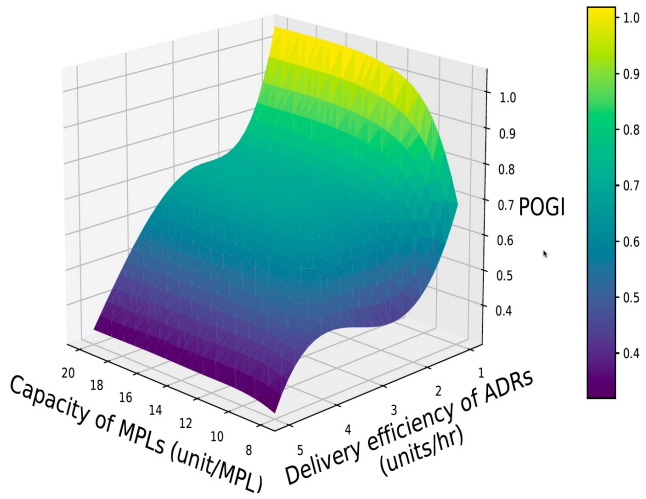
5.4.6 Performance of KSVM for Service Charge Prediction

This section evaluates the performance of the proposed KSVM in predicting the service charge q of ADR service based on expected equilibrium and other operating parameters (as described in 5.3.4). The KSVM model can be employed to forecast other operational parameters such as the capacity or delivery fleet that expects investing in the market, which enables logistics operators to optimise the allocation of their delivery facilities. Since the study intends to evaluate the regression performance of KSVM with a relatively small sample size, we use 400 samples to serve as training data and 80 samples for testing data, determined through cross-validation. To determine the optimal kernel for the employed dataset, we assess the performance of KSVM with three kernel functions (Linear, RBF, and Sigmoid). The regression parameters of KSVM were dynamically determined by predefined regularisation parameter C and kernel coefficient γ_{ksvm} that serve as geometric sequences ranging from 10^{-2} to 10^{2} with an interval length of 5. The performance of the KSVM was measured by the Mean Square Error (MSE), the Explained Variance Regression Score (EVRS), and the Coefficient of Determination (R^2).

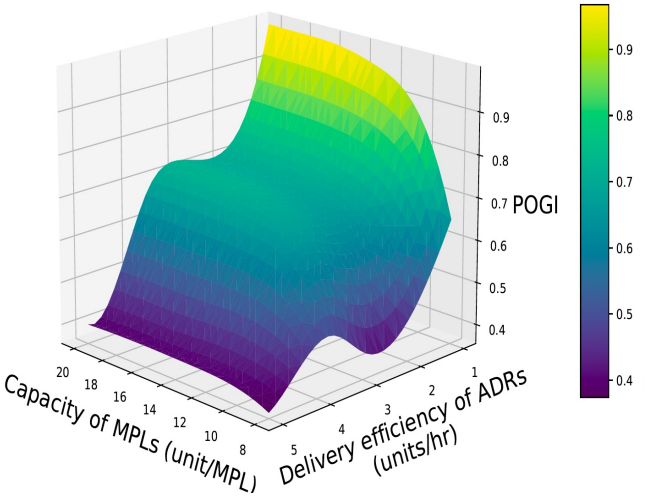
Appendix F presents the performance of different KVSM kernels. The result shows that the RBF kernel with $C = 1$, $\gamma_{ksvm} = 100$ achieves the best regression performance since the MSE value is 0.306. It means that the square of the average difference between the predicted price of ADRs and the true value is 0.306 unit cost/parcel. The obtained lower MSE score proves a good performance of using RBF-KSVM for price prediction since the price range of the training sample is from 0 to 10. In terms of EVRS and R^2, the RBF kernel is also the best fitting our model with the value of 0.912 and 0.893, respectively, since the closer the EVRS and R^2 approaches 1, the better the independent variable (predicted value) is able to explain the variance change of the dependent variable (truth value) thus better fitting the training sample. Neither Linear nor Sigmoid kernel achieves explainable performance since their R^2 is



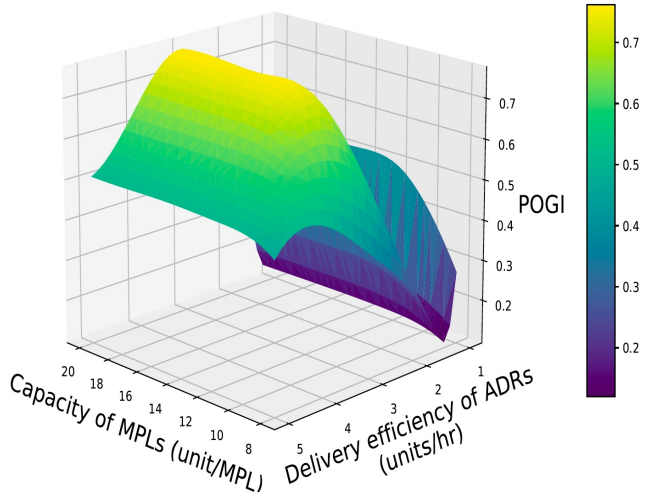
(a) POGI of fully rationality model



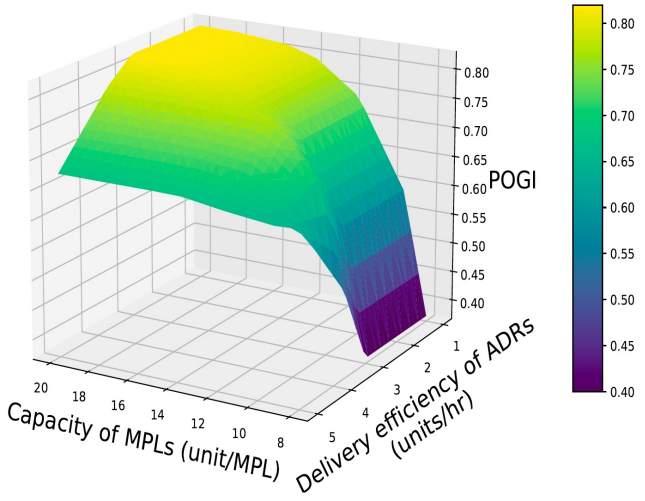
(b) POGI of bounded rationality model ($\gamma = 0.9$)



(c) POGI of bounded rationality model ($\gamma = 0.7$)



(d) POGI of bounded rationality model ($\gamma = 0.5$)



(e) POGI of bounded rationality model ($\gamma = 0.3$)

Figure 5.11: POGI of different rationality models

all negative.

5.4.7 Managerial Implications and Discussions

As the sophistication of autonomous technology increases, a growing number of consumers, particularly young omnichannel consumers and online shopping enthusiasts, are inclined to use MPLs and ADRs as last-mile delivery solutions. These users are willing to pay for more convenient and efficient delivery services (Pani et al., 2020). Logistics operators can capitalise on this trend and increase market share in autonomous deliveries by aligning their operating parameters with the decision-making behaviour of these target consumers. By offering a balance of delivery convenience and efficiency along with cost-effective service charges, logistics operators can position MPLs and ADRs as complementary options to achieve optimal allocation for different demand segments.

Our experiment highlights that customer adoption of MPL service is primarily influenced by the travel distance for parcel collection and the duration of parcel storage. It is recommended that the maximum walking distance for customers when using MPL service should be less than 500 ($\pm$200) meters, which aligns with previous studies as reported by Kedia et al. (2017). Therefore, optimising the placement of MPLs to enhance customer convenience and accessibility is crucial. One potential strategy is to position MPLs in residential areas during non-working hours and relocate them near workplaces during peak commuting times. Additionally, for a cost-effective expansion of MPL capacity, it is advisable to initiate the service with a moderate capacity range, such as 20-30 units per MPL. Furthermore, to distinguish MPL services from ADR offerings and enhance their attractiveness to customers, logistics operators may consider providing complimentary storage for parcels within a specified duration during the initial phase, while implementing tiered fees for parcels exceeding this time frame to discourage inefficient compartment occupancy.

In terms of ADR service, the acceptance of customers primarily depends on the service charge and delivery efficiency. Our analysis reveals that urban customers exhibit a higher inclination to pay for ADR services compared to those residing in suburban or rural areas, particularly if they are located more than 0.5 miles from shopping malls (Bansal et al., 2016). The optimal service charge for ADR service ranges from 1 to 3 cost units per parcel, aligning with previous research by Pani et al. (2020), where approximately 42% of consumers expressed

a willingness to pay between \$2 and \$3. ADR services are recommended specifically for grocery deliveries of low-value yet time-sensitive parcels, as these customers are typically available at the designated location during the scheduled collection time and are less likely to miss the delivery. Conversely, due to security concerns, customers may be reluctant to utilise ADR services for high-value product deliveries, as the ADR remains unattended during the delivery process. To enhance the delivery efficiency of ADR services, it is advisable to load the ADR onto a van and transport it from a suburban distribution centre to a predetermined drop-off area or deploy it in proximity to the end customer at a local service point.

To improve distribution resource allocation, it would be beneficial for logistics operators to provide consumers with more complete service information. One practical way to achieve this goal is to offer more accurate predictions of delivery time or service queue lengths on the service portal when customers are making their service choices. This would mitigate the negative impact of customers' bounded rational behaviour. The pricing and deployment strategies of service providers also play a crucial role in capturing market share. In addition to the proposed non-cooperative game model and predictive technique of KSVM in capturing customers' service selection behaviour, conducting a survey of customer service preferences through a questionnaire can further provide valuable insights for logistics operators to make operational improvements and offer more precise and efficient delivery services.

While this study offers managerial insights for the future deployment of MPL and ADR delivery services based on investigating customer service selection behaviour, it is important to acknowledge certain limitations to maintain an unbiased discussion. Firstly, this paper aims to explore the decision behaviour exhibited by each customer within a large group based on observed states, emphasising the cumulative impact of each individual decision on global efficiency within the on-demand delivery scenario. This requires detailed service information for each customer, such as the expected delivery time for a specific customer, to form the payoff function. For investigating the influence on system performance from an economic-scaled perspective (such as multi-batch delivery and multi-fleet delivery), the proposed game-theoretic approach and queuing mechanism may no longer be applicable. Continuous approximation could be a potential alternative as it allows for a significant simplification of operations in

large-scale, multi-fleet delivery scenarios, thus better capturing the overall efficiency of the system, especially when individual behaviour is not a primary focus of the study. This can be proven by the formulation of continuous approximation introduced by Daganzo (1980), where the approximation of the delivery route's length relies on demand density and the dimensions of the service areas.

Secondly, given the absence of large-scale deployment of MPLs and ADRs in this stage, along with the lack of disclosure of operational data by logistics operators, this study offers theoretical and forward-looking managerial insights based on simulation results. Subsequent research endeavours will employ case studies grounded in actual operational data.

5.5 Concluding Remarks

This chapter provides a thorough investigation of the game-theoretic aspects of the service selection between two types of competitive autonomous delivery services: MPLs and ADRs. When formulating the service selection between MPLs and ADRs, we specifically consider the effects of customer rationality levels on decision-making and the influence of global information on the improvement of system efficiency. To model both fully rational and bounded rational customer decision-making behaviour, we developed Wardrop equilibrium and Logit equilibrium models based on the operating characteristics of MPLs and ADRs, respectively, and proved the existence and uniqueness of these equilibrium models. Subsequently, we defined four performance indicators and analysed the impact of various operating parameters on system performance in the context of different levels of customer rationality. Furthermore, we proposed an improved algorithm that leveraged global information to enhance system performance in models with varying levels of customer rationality and evaluated its improvement rate. Finally, this chapter introduced a regression prediction approach using a KSVM, which is based on the expected equilibrium and other operating parameters, for the pricing of delivery services.

The numerical experiments conducted in this study have provided valuable insights into the factors influencing the efficiency of the delivery system. Specifically, the implementation result has identified the customer's travel distance and the average duration of parcel storage

as significant determinants of equilibrium and efficiency for the MPLs service. In contrast, for the ADRs service, the service charge and delivery efficiency have emerged as key factors impacting system performance. Furthermore, the findings highlight the heightened sensitivity of customers towards high-cost or capacity-limited services like ADRs, underscoring the influence their decisions wield over system performance. Consequently, it becomes imperative for logistics operators to employ effective pricing mechanisms and resource allocation strategies to optimise system utilisation.

Moreover, it is noted that in a decentralised system where customers exhibit fully rationality, there are more significant disparities in system performance. The findings also revealed that by providing customers with global information, such as waiting queue length or estimated delivery time, the system's performance could be substantially enhanced by 28.4% to 68.7%. This improvement can be attributed to the fact that customers are empowered to make well-informed decisions based on more accurate payoff estimations. To facilitate this, delivery service platforms such as websites and mobile apps can disseminate global information, making it readily available to customers during the service selection process. Logistics operators need to address the critical challenges of reasonable pricing and efficient capacity investment. Finally, the numerical experiments demonstrated the effectiveness of the proposed KSVM in predicting service charges with high accuracy, even with a relatively small sample size. The implementation of KSVM enables logistics operators to swiftly adjust their service pricing based on fluctuations in customer demands.

As the scaled implementation of MPLs and ADRs is still in the pilot study phase, this chapter aims to provide theoretical and forward-looking guidance for logistics operators prior to the widespread deployment of these autonomous delivery systems in the last-mile distribution market. One way to utilise MPLs effectively is to deploy them in residential areas during non-working hours and in work areas during peak commuting hours, taking advantage of their mobility to provide customers with convenience and accessibility. ADRs can be utilised in combination with conventional vans or deployed near local service points to increase the delivery radius and efficiency. On the other hand, differentiating MPL and ADR services for latent customer groups enables customers to select their preferred service according to their needs.

This can be achieved through a compromise between the price and delivery efficiency of the two services. A potential strategy is to implement differential pricing, whereby MPLs are offered at no cost or reduced rates compared to ADR service. This pricing scheme serves as an incentive for customers and simultaneously motivates those seeking enhanced convenience to opt for ADRs.

The findings of this chapter offer significant and theoretical implications for logistics operators aiming to optimise the utilisation of ADRs and MPLs in diverse urban delivery contexts. Particularly, the study introduces an innovative and efficient methodology for predicting and assessing the influence of operational factors on the autonomous last-mile distribution market. To have an unbiased discussion, the limitation of the research lies in the direct control mechanism of the improved algorithm with global information applied (i.e., assume that the global information can be utilised by the players straightforwardly). In the research domain of game theoretical control, it will be more interesting if the optimisation of POA in the game system applies an indirect mechanism to constrain the action profile of players (e.g., providing players with rewards as an incentive to promote the utilisation of global information). However, given the main purpose and contribution of this research lies in customer behaviour analysis for delivery service selections and the evaluation of the autonomous last-mile distribution system, this limitation will be resolved in future studies focusing on performance optimisation. Given the above concerns, this chapter provides a first-hand and stepping-stone reference for future performance optimisation research by introducing global information into the decision model of the customer.

Future work could focus on the multi-players or hierarchical game process, which incorporates logistics operators and policymakers as decision-makers. Under these circumstances, logistic operators would consider more complicated scenarios to determine optimal operating parameters cost-efficiently and understand how different stakeholders would react to these scenarios. Additionally, it is valuable to use different simulation techniques (such as multi-agent simulation) to construct the game system and compare them with the discrete event simulation techniques adopted in this study, so as to determine a more effective simulation method. Another research direction is considering optimising the game mechanisms

with equilibrium or POA guarantees, which helps improve the system's POA and POR.

Chapter 6

Route Planning Model for Last-Mile Deliveries using Mobile Parcel Lockers

The work of this chapter has been published at the **Transportation Research Part E: Transportation and Logistics Review** entitled: "*Route Planning for Last-Mile Deliveries using Mobile Parcel Lockers: A Hybrid Q-Learning Network Approach*"[1].

Chapter 5 has examined the service selection behaviour of customers regarding the MPL and ADR delivery modes. The results indicate that significant factors influencing customers' choice towards MPL services include customer walking distance for parcel pick-up and time window considerations. Therefore, judiciously determining the stopover locations for MPL and planning the routes for MPL travel between different stopovers hold the potential to enhance the market share of MPL service. To achieve the goal, this chapter introduces a Mobile Parcel Locker Problem (MPLP), a special case of the Location-Routing Problem (LRP), which determines the optimal stopover location for MPLs throughout the day and plans corresponding delivery routes. As a primary feature of the proposed MPLP, MPLs could dynamically transfer parcels between depots and available parking spaces, with special consideration for parking space availability and the customers' drop-off location/timing requests.

Beyond the novel parcel delivery mode proposed above, this study further considers the

[1]Liu, Y., Ye, Q., Escribano-Macias, J., Feng, Y., Candela, E., & Angeloudis, P. (2023). Route planning for last-mile deliveries using mobile parcel lockers: A hybrid q-learning network approach. *Transportation Research Part E: Logistics and Transportation Review*, 177 , 103234.

dynamic planning technique that can resolve the time window conflict and delivery route adjustment, which has not been widely discussed by the existing MPL-relevant studies. A Hybrid Q-Learning-Network-based Method (HQM) is developed to resolve the computational complexity of the resulting large problem instances while escaping local optima. In addition, the HQM is integrated with global and local search mechanisms to resolve the dilemma of exploration and exploitation faced by classic RL methods. Thus, the contributions of this chapter are as follows:

- **Theoretical contribution:** to the best of the author's knowledge, this research is the first attempt to apply an RL-based approach to solve MPLP. The analysis reveals that HQM with improved sampling efficiency has better optimisation ability than the exact approach and Genetic Algorithm (GA) in large MPLP instances regardless of the network type. We aspire that it will serve as the foundation for future studies in the field of location-routing optimisation or VRPs in last-mile deliveries.

- **Algorithmic contribution:** the proposed novel encoder & attention-based decoder framework of HQM is superior to other existing RL-based frameworks that use coordinates for training input as it can handle more decision parameters (as these parameters can be infinitely extended to the encoder structure) to resolve more complex multi-decision MPLPs and VRPs.

- **Modelling contribution:** we consider the dynamic planning scenario of the proposed model where two route adjustment strategies are developed within the HQM framework to dynamically resolve the time windows conflicts caused by service delay and traffic conditions, which has the potential to be adopted to other RL-based methods in solving time-windows conflicts of VRPs.

- **Managerial contribution:** based on parametric analysis of key factors that influence MPL-based parcel deliveries, this study provides managerial insight from the perspective of logistics operators to stimulate broader adoption of autonomous last-mile delivery modes.

The remainder of the chapter is structured as follows: Section 6.1 describes the definition of MPLP. Section 6.2 formulates mathematical models for MPLP, which includes location

selection model and route planning model. Section 6.3 introduces the solution approach for solving MPLP, which includes task generation algorithm, structure design of the state, and HQM algorithm. Section 6.4 presents the results of experiments and evaluates the performance based on different aspects (e.g., final reward, convergence, computing time), compared with the Gurobi solver and GA-based alternative, used as the baseline. The study further identifies key influence factors (time windows, network sizes, capacity, etc) from both MPLs and customers' perspectives towards different cost units of the objective function. Managerial implications for future MPL operations based on the parametric analysis of different influencing factors are provided in Section 6.5. Section 6.6 provides a fruitful discussion about the advantages and future application of the proposed HQM. Finally, conclusions and recommendations for future work are provided in Section 6.7.

6.1 Mobile Parcel Locker Problem (MPLP)

The MPL operation follows the corresponding scenario definition described in Section 4.1 in Chapter 4, with further consideration for time windows of customers and parking space. A set of customer nodes N serves as the recipient of parcels from a group of MPLs. For each customer $n \in N$, customer n creates a set of locations K_n during the planning horizon. Each location $k \in \{1, 2, \cdots, |K_n|\}$ that a customer visited within the location set K_n is defined by a time interval $[E_{nk},\ L_{nk}](L_{nk} > E_{nk})$ and a corresponding position $[X_{nk},\ Y_{nk}]$. To better fit the real-world operations, we introduce customer's maximum walking range ρ_n to investigate how it affects the delivery system (will be introduced in Section 6.4.4).

A set of MPLs is denoted by M, which MPLs can change their locations to facilitate parcel pick-ups by the customer set N. The service radius of each MPL m ($m \in M$) is denoted as r_m, and the average travel speed is v. We define a set of parking spaces I where locker m can potentially park to serve multiple customers at each parking lot. Each parking space $i \in I$ is represented by a position $[X_i,\ Y_i]$, and the available parking time windows $[E_i,\ L_i](L_i > E_i)$. We use ST_i to denote the service time of parking space i for fulfilling each customer. Therefore, for each parking space i, the parking time windows $[E_i,\ L_i]$ can be

divided into $|U_i|(|U_i| = |L_i - E_i|/ST_i)$ sub-intervals, where U_i represents a set of sub-interval time spans in parking space i. Note that $|U_i|$ needs to be rounded up if it is not an integer such that ensure the length of the original available parking time windows $[E_i,\ L_i]$ is fully considered. We introduce $[e_{ia},\ l_{ia}]$ to represent the $a^{th}(a = 1, \cdots, |U_i|)$ sub-interval within U_i in parking space i (i.e., $[e_{ia},\ l_{ia}] \in [E_i,\ L_i]$). Let $U(U = \{U_1, U_2, \cdots, U_i\})$ to represent a set of all sub-intervals for all parking space i.

For each location $k(k \in K_n)$ of customer n, it can be assigned to the corresponding nearest parking space i to enable customer n visits MPL m for parcel collection (achieved by the location selection model introduced in Section 6.2.3).

We define a task O_{ima} as the movement of MPL m to a specific parking space i, fulfilling the demand q_{ia} within the a^{th} sub-interval $[e_{ia},\ l_{ia}]$ of parking space i (i.e., each sub-interval time span corresponds to one single task). The MPL will depart as soon as the current task O_{ima} is completed. Let O denotes a set of all tasks to be performed $(O_{ima} \in O)$.

We assume that each customer n has $\bar{q_n}$ parcels which can be potentially collected at their any location k within a day. Therefore, for each parking space i, the total demand q_{ia} within the a^{th} sub-interval $[e_{ia},\ l_{ia}]$ can be calculated by the following equations:

$$q_{ia} = \sum_{n=1}^{N} \bar{q_n} \tag{6.1}$$

Subject to:

$$\sqrt{(X_i - X_{nk})^2 + (Y_i - Y_{nk})^2} \leq \rho_n \qquad \forall i \in I, n \in N, k \in K_n \tag{6.2}$$

$$e_{ia} \leq E_{nk} \qquad \forall i \in I, a \in U_i n \in N, k \in K_n \tag{6.3}$$

$$L_{nk} \leq l_{ia} \qquad \forall i \in I, a \in U_i n \in N, k \in K_n \tag{6.4}$$

Equation (6.1) represents the total demand q_{ia} within the a^{th} sub-interval $[e_{ia},\ l_{ia}]$ in parking space i is the aggregated demand of all customers whose stopover k is assigned to parking space i without exceeding customers' maximum walking distance ρ_n (Constraint 6.2). Constraints (6.3) and (6.4) make sure that the time windows of the k^{th} location of customer n should be included within the a^{th} sub-interval of parking space i, such that the demand assigned to the a^{th} sub-interval is valid.

By obtaining a set of tasks to be fulfilled at different sub-intervals of all parking spaces $i \in I$, the tasks are assigned to MPLs to perform while considering the execution sequence. The task execution sequence of MPLs, therefore, formatting the VRP. The proposed MPLP makes the following decisions in sequence:

- The optimal location of a set of parking spaces I, such that customers within the service radius can access MPLs for parcel collection.
- The task assignment scheme for all determined tasks (i.e., by which task to be performed by which MPL), minimising the MPLs deployed.
- The task execution sequence for all MPLs (i.e., the route of MPLs visiting different parking spaces to perform tasks), considering customers' time-windows, available parking time, and shorter route length.

This study does not plan the route between customers and MPLs since each customer will only access the nearest MPL (this only composes a direct shortest path between MPLs and customer locations).

As a result of MPLs' delay, customers might not be served by MPLs on time. In this case, an additional travel distance arises when modifying predefined delivery routes. This route modification is designed to resolve the undelivered parcels within the pre-designed time windows (the corresponding route adjustment algorithms will be introduced in *Route Adjustment Strategy and Route Generation Algorithm* section).

Figure 6.1 illustrates an example of MPLP with three customers and one MPL. The available parking spaces for MPLs are denoted by solid triangles from A to D, and the corresponding available parking time windows are represented by ovals. The three customers are denoted by the solid circle with different numbers from 1 to 3, and their available pickup times at different locations are represented as rectangles. Customers may change location throughout the day as denoted by the dotted arrows. For example, Customer 1 departs from the location covered by Parking Space A to the place covered by Parking Space B, and finally reaches Parking Space D. Customer 2 travels from the location covered by Parking Space C to the destination covered by Parking Space B. Analogously, Customer 3 travels from the location

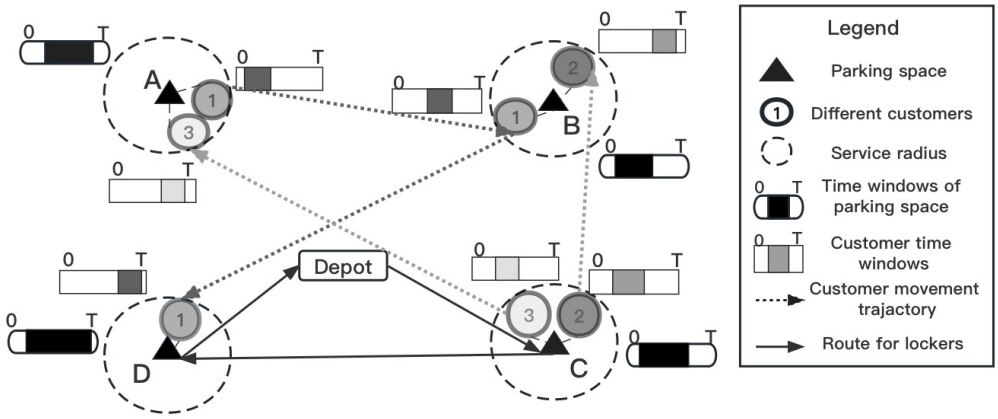


Figure 6.1: Instance of MPLP

covered by Parking Space C to the destination covered by Parking Space A. The optimal solution in this example is to deploy one MPL, which leaves the depot and visits Parking Space C and Parking Space D in sequence before returning to the start when completion (since the available parking time intervals satisfy both customer time windows and service buffer time).

6.2 Mathematical Model for MPLP

This section presents the modelling assumptions, model structure, location selection model of parking spaces, and route planning model for MPLs.

6.2.1 Model Assumptions

When developing the model, we made the following assumptions:

Assumption 1: It is assumed that the logistics operator will ask customers (via the mobile APP or message) if he/she would like to receive MPL service before parcel deliveries. The data usage claim will be made to customers that they are aware of the data (e.g., location, time windows) shared with operators are not too precise. Furthermore, a customer cannot access the data of others. Such an approach would be aligned with practice used by popular food delivery services (e.g., *Deliveroo*): when a rider is fulfilling two consecutive customers at the same time, the location of the previous customer and the rider's route are invisible to the latter customer until the first order has been finished.

Once customer accept the MPL service, they can collect parcels at one of the parking spaces based on their stopovers within a day. The logistics operator will dispatch the MPL to the predefined parking space to serve the customer according to the optimal dispatch strategy. Since the logistics operators only require the approximate location of the customer prior to the service instead of a real-time and precise trajectory, the data privacy of customers can be well-protected.

Other potential incentives of the logistics operators to overcome these concerns, e.g., reduced service fares and government-dominated third-party data supervision, is worth considering in practice. The data supervision is governed by laws and regulations concerning user privacy data issued by relevant government organisations. These measures serve to protect the rights of customers, although the extent of this protection may vary.

Assumption 2: The time windows of a specific customer at different locations shall never overlap to fit real operations.

Assumption 3: The maximum accepted walking ranges of customer ρ_n is assumed to be a constant in which we set different values to evaluate its influence on the delivery system (see Section 6.4.4). However, ρ_n can vary from different customers following a specific probability distribution if required. In addition, the available parking duration $[E_i, L_i]$ varies for different parking spaces due to the different regulations in the region.

Assumption 4: The service charge for parking spaces is not considered. We assume that parking spaces have infinite capacities to accommodate multiple MPLs occupying a parking space. However, the capacity of parking spaces can be adjusted and limited according to real operation.

Assumption 5: Each locker will load parcels at the beginning of the planning horizon and will drive back to the depot once the MPL tasks in the current delivery round are completed.

Assumption 6: Since parcels are standardised in their maximum dimension and weight by logistics providers, we assume that all parcels fit any compartment of the MPL.

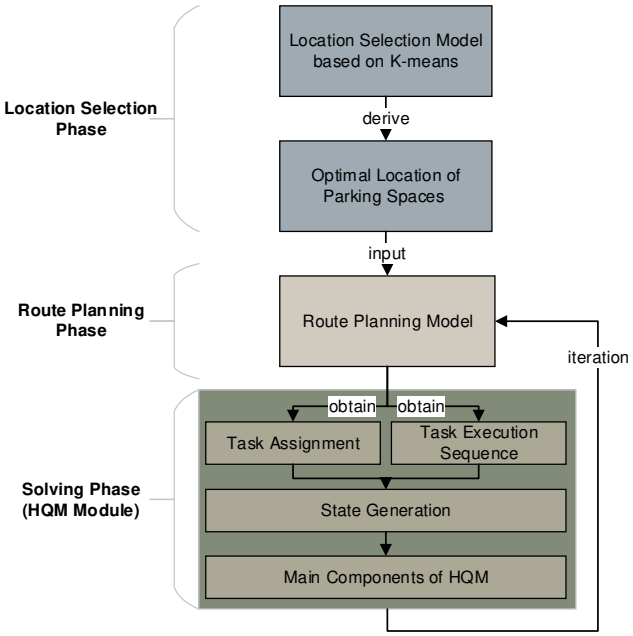


Figure 6.2: Illustration of the model relationship

6.2.2 Model Structure

Following the above problem statement, the proposed MPLP model consists of three parts, as shown in Figure 6.2. The first phase is the location selection, in which the optimal locations for parking spaces are determined based on the spatial distribution of customers. In this phase, all potential stopovers of customers are enumerated and the K-means clustering method is applied to identify the coordinates of optimal parking spaces and the customers' location-parking space pairings, which are used as inputs for the route planning phase. The second phase, route planning, aims to determine delivery routes for MPLs to minimise the MPL fleet size, total distance travelled, and service delays while satisfying all customers and parking time window constraints. To achieve this goal, we first generate the initial feasible solution, which consists of a task assignment scheme and task execution sequence. The generated initial solution will be inputted into the HQM module, where the initial feasible solution will be encapsulated as a state to make it interpretable by the main components of HQM. The main body of HQM includes several crucial optimising components to obtain better state/solution through iteration. The detailed procedure of the HQM module will be introduced in Section 6.3.3.

Table 6.1: Notation table for location selection model

Indices		**Sets**	
n	Customer n ($n \in N$)	N	Set of customer
i	The i^{th} parking space ($i \in \{1, \cdots, P\}$)	K_n	Set of locations for a customer n ($n \in N$)
k	The k^{th} location of a single customer ($k \in K_n$)		
Parameters		**Decision Variables**	
P	The total number of parking spaces (clusters) needs to be determined	g_i	Float: coordinates of the parking space i ($g_i = [X_i, Y_i]$)
ρ_n	Maximum walking range of customer n ($n \in N$)	b_{nk}	Integer: the cluster index of k^{th} location of customer n (serve as true labels)
r_m	Service radius of MPL		
E_{nk}	The earliest time of the k^{th} location of customer n		
L_{nk}	The latest time of the k^{th} location of customer n		
X_{nk}	X coordinate of the k^{th} location of customer n		
Y_{nk}	Y coordinate of the k^{th} location of customer n		
h_{nk}	Coordinates of the k^{th} location of customer n ($h_{nk} = [X_{nk}, Y_{nk}]$)		

6.2.3 Location Selection Model for Parking Spaces

Table 6.1 presents the notations of the location selection phase. All distances are in L_1 metric (Manhattan distance).

Mathematical Model for Parking Space Location Selection

We establish a location selection model for parking spaces (adapted from Ng (2000)).

$$b_{nk} = f(g_i) = argmin_{i \in \{1, \cdots, P\}} \|h_{nk} - g_i\|^2 \qquad \forall n \in N, k \in K_n \tag{6.5}$$

where:

$$g_i = \frac{\sum_{n=1}^{N} \sum_{k}^{K_n} 1\{b_{nk} = i\} \cdot h_{nk}}{\sum_{n=1}^{N} \sum_{k}^{K_n} 1\{b_{nk} = i\}} \qquad \forall i \in \{1, \cdots, P\} \tag{6.6}$$

$$g_i \in \mathbb{R}, b_{nk} \in \mathbb{N} \qquad \forall i \in \{1, \cdots, P\}, n \in N, k \in K_n \tag{6.7}$$

The location model intends to partition all customers' location (data point) into P parking spaces (clusters) in which each customer's location belongs to the parking space with the nearest mean. The objective of K-Means clustering is to determine the cluster index b_{nk} (true label) for each customer's location h_{nk} through minimising squared error (total intra-cluster variance) in Equation (6.5), such that makes this an unsupervised learning problem. Equation (6.6) represents the calculation for all centroids of parking spaces through customer location-parking space matching. Equation (6.7) defines the domains of the decision variables.

Since the K-means requires an explicit number of clustering centres P (parking spaces) before implementation, we determine the minimum value of P by guaranteeing $d_{in}^k \leq \rho_n \leq r_m$ through trial and error.

6.2.4 Route Planning Model for MPLs

By implementing the location model, the optimal location of parking spaces ($g_i = [X_i, Y_i]$) and the customer location-parking space pairing is determined and were used in route planning model. This allows the routing model to calculate the distance between customers and parking spaces. The notations and route planning model for MPLs is demonstrated as follows. Appendix G provides the derivation of auxiliary variables.

Mathematical Model for Route Planning

$$Minimise \quad f(x_m, y_{ijm}, z_{ima}) = W_1 C_1 + W_2 C_2 + W_3 C_3 \tag{6.8}$$

$$C_1 = F_m \sum_{m=1}^{M} x_m \tag{6.9}$$

$$C_2 = \sum_{m=1}^{M} \sum_{i=0}^{V} \sum_{j=0}^{V} y_{ijm} d_{ij} c_{ij} \tag{6.10}$$

$$C_3 = \sum_{m=1}^{M} \sum_{i=1}^{I} \sum_{a}^{U_i} [max(e_{ia} - t_{ima}, 0) + max(t_{ima} - l_{ia}, 0)] \tag{6.11}$$

Table 6.2: Notation table for route planning model

Indices		**Sets**	
m	MPL m ($m \in M$)	M	Set of MPLs
i, j	Two adjacent parking spaces ($i, j \in I$)	I	Set of parking space (excluding depot)
a	The a^{th} task of a parking space i ($a \in U_i$)	V	Set of nodes within the network (including set I and depot)
		U_i	Set of sub-interval/tasks in parking space i
		U	Set of all sub-intervals/tasks for all parking space
Parameters		**Auxiliary Variables**	
$\bar{q_n}$	Number of parcels collected by customer n	d_{ij}	Distance from parking space i to j
Q	The capacity of MPLs	t_{ij}	Actual travel time from parking space i to j
F_m	Fixed cost of MPL m	d_{in}^k	Distance from the k^{th} location of customer n to parking space i
v	Travel speed of MPLs	$[e_{ia}, l_{ia}]$	The sub-interval within $[E_i, L_i]$ of the a^{th} task at parking space i ($a = 1, \cdots, \lvert U_i \rvert$)
Ψ	The maximum number of MPLs that can be used	t_{ima}	The time when locker m arrives at parking space i to perform the a^{th} task of parking space i
c_{ij}	Unit cost travelling from parking space i to j	q_{ia}	The total demand to be fulfilled within the a^{th} sub-interval of parking space i
ST_i	Service time of parking space i	**Decision Variables**	
X_i	X coordinate of parking space i	x_m	Boolean: 1, if locker m is used for service; 0, otherwise
Y_i	Y coordinate of parking space i	y_{ijm}	Boolean: 1, if locker m travel to parking space j from parking space i; 0, otherwise
E_i	The earliest available parking time in parking space i	z_{ima}	Boolean: 1, if locker m fulfils the a^{th} task of parking space i; 0, otherwise
L_i	The latest available parking time in parking space i		
ϕ	A large indefinite number		

Subject to:

$$\sum_{m=1}^{M} x_m \leq \Psi \tag{6.12}$$

$$\sum_{i=1}^{I} \sum_{a=1}^{U_i} z_{ima} \cdot q_{ia} \leq Q \cdot x_m \qquad \forall m \in M \tag{6.13}$$

$$\sum_{m=1}^{M} z_{ima} = 1 \qquad \forall i \in I, a \in U_i \tag{6.14}$$

$$\sum_{i=1}^{I} y_{0im} = \sum_{m=1}^{M} x_m \tag{6.15}$$

$$\sum_{j=1}^{I} y_{j0m} = \sum_{m=1}^{M} x_m \tag{6.16}$$

$$t_{jma} = t_{ima} + ST_i + d_{ij}/v - \phi(1 - y_{ijm}) \qquad \forall i, j \in V, i \neq j, m \in M, a \in U_i \tag{6.17}$$

$$x_m, y_{ijm}, z_{ima} \in \{0, 1\} \qquad \forall i, j \in V, m \in M, a \in U_i \tag{6.18}$$

The objective function (6.8) minimises the fleet size, total travel distance and service delay, where W_1, W_2, and W_3 represent the weights of three cost units. Equation (6.9) to (6.11) denotes the MPLs' deployment cost, travel cost, and delay penalty respectively.

Constraint (6.12) denotes the total quantity of MPLs used should not exceed the predefined value. Constraints (6.13) represents the total demand fulfilled by locker m should not exceed the maximum capacity of MPL. Constraint (6.14) ensures that each task within the parking space i will be executed only once. Constraints (6.15) and (6.16) guarantee MPLs depart from a central depot and will return to the depot when the task is completed, and the number of routes from and back to the depot is equal to the MPL fleet size. Constraint (6.17) represents the time relationship of locker m visit to parking space j from parking space i. Constraint (6.18) defines the domains of the decision variables.

6.3 Solution Methods

The structure of the proposed solution approach is illustrated in Figure 6.3. Firstly, we utilise K-means clustering to determine the optimal location for parking spaces based on customer distribution and assign different customer locations to the nearest parking space, so that a customer can access MPLs at parking spaces at one of their potential locations. Since locating parking spaces through K-means clustering is not considered to be the main problem in this

study, and a number of open-source packages have already been developed for realising K-means clustering, this section will not provide detailed instructions for implementation. Related studies of applying K-means for location selection problems see Kumar & Kumanan (2012) and Pham et al. (2005).

Next, the task generator produces task set O for MPL fleet. MPL m are then allocated to perform each task O_{ima} ($O_{ima} \in O$) within the a^{th} sub-interval $[e_{ia}, l_{ia}]$ in parking space i (as described in Section 6.1).

Then, the HQM algorithm is developed to determine the following set of decisions: i) the task assignment schemes (denoted as x_1) that define the task-locker assignment, and ii) the task execution sequence for each MPL (denoted as x_2). The size of x_1 and x_2 equals the total number of tasks ($|O|$) determined in the task generation phase, respectively. A state $s = (x_1, x_2)$, which consists of a x_1, x_2 pairing, represents a feasible solution of the proposed MPLP. A set of state s forms an agent (denoted as $Agent$). The reward R of HQM is defined as the reciprocal of the objective function (6.8) since the HQM is set to maximise the reward. The action A_t is defined as a decision behaviour in which the Agent selects a policy to obtain a better solution in the next timestep t.

The HQM contains an encoder, the global and local search framework, a route generator, and an attention-based decoder. The encoder first produces a single $Agent$ through a random generator and permutation pool. To speed up the searching process for the optimal solution, the global search mechanism based on the Q-value matrix and the local search mechanism based on single-state adjustments are used concurrently to generate the new state s' for $Agent$.

The updated state s' will then be input into the route generator so that the delivery routes for MPLs can be developed. The decoder is used for calculating the reward R of a state s and decoding the best state s^* such that obtains the optimal task assignment scheme x_1^* and the tasks execution sequence x_2^*. The attention layer within the decoder is used to specify the selecting possibility of each element within x_1 and x_2 that is inherited by the next action A_{t+1}. The current best state s^* will then be recorded, and the Q-value matrix will be updated accordingly. Once the convergence condition has been satisfied, the HQM outputs the best state s^*.

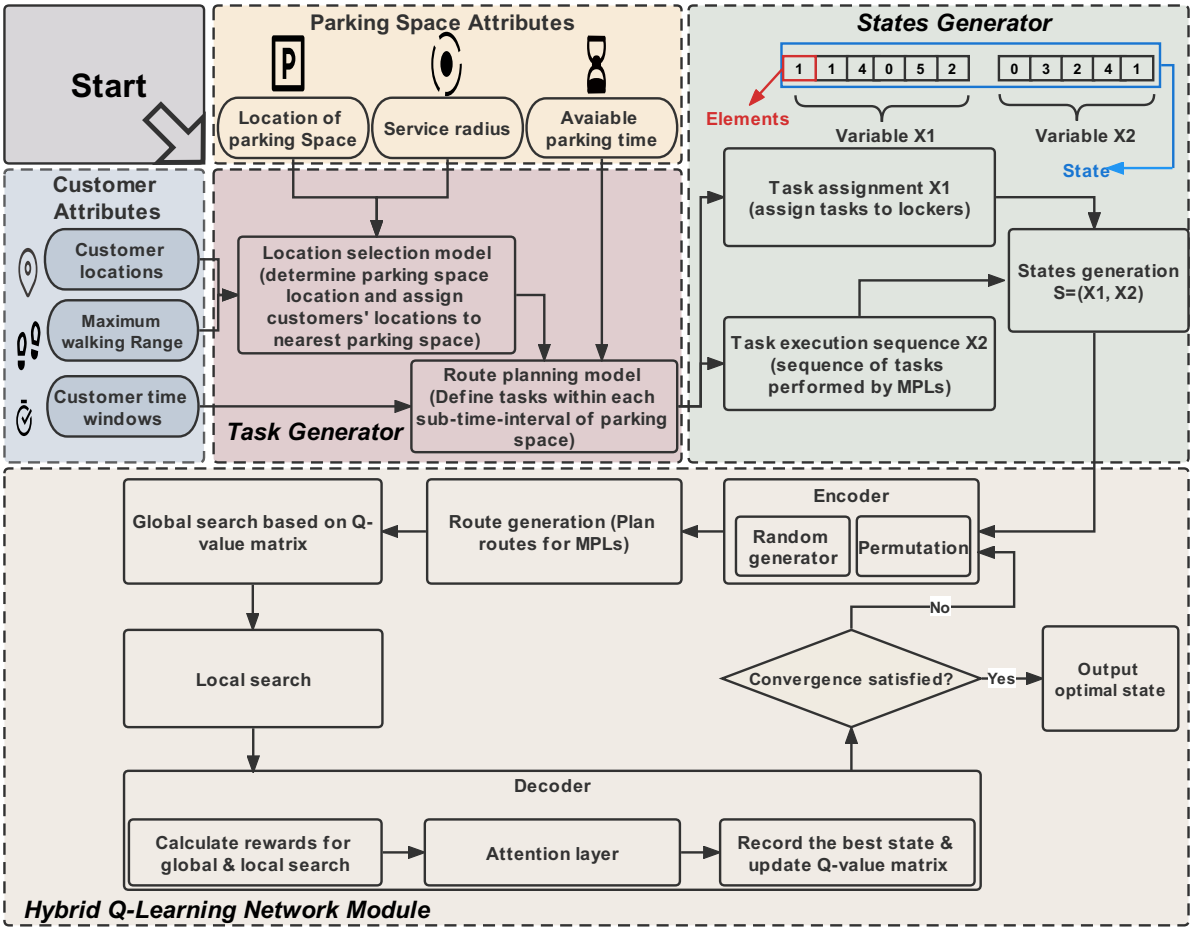


Figure 6.3: Solution framework of the model

The encoder generates *Agent* that is input to the decoder, which then identifies optimal state s^* for each *Agent* according to the current knowledge of the problem. In this manner, the optimal task assignment scheme x_1^* and the task execution sequence x_2^* are determined. A description of the detailed structure and essential elements of the HQM will be provided in Section 6.3.3. The following section will describe the detailed mechanisms of the various components within the solution framework in sequential order.

6.3.1 Task Generation Algorithm

Following the procedure presented in Figure 6.3, we first demonstrate the process of task generation. The objective of this phase is to identify a set of tasks that satisfy both time-window constraints imposed by customers and parking spaces.

The first step is to reduce the redundant length of task set U_i for each parking space i. For instance, if a parking space i is available for an MPL from 10:00 to 16:00

($[E_i, L_i]$=[10:00,16:00]), but no customer time windows within the range between 14:00 and 16:00 ($[E_i, L_i]$=[14:00,16:00]), U_i can be reduced since $[E_i, L_i]$ is now only lasts from 10:00 to 14:00. The available time interval of parking space i can be reduced to $U_i^{'}$:

$$U_i^{'} = \{[E_i, L_i] \cap [E_{nk}, L_{nk}]\} \qquad \forall n \in N, i \in I, k \in \{1, 2, \cdots, |K_n|\} \tag{6.19}$$

To further reduce the range of available time intervals, we determine the earliest time window and the latest time window among all customers within each parking space. The optimal range of available time intervals will be updated as U_i^{opt}. By defining U_i^{opt}, a set of sub-intervals (Equation 6.21) of parking space i can be obtained when executing U_i^{opt} divided by the service time ST_i. The detailed derivation for e_{ia} and l_ia see Appendix G.

$$U_i^{opt} = [min\{E_{1k}, \cdots, E_{nk}\}, max\{L_{1k}, \cdots, L_{nk}\}] \cap U_i^{'} \forall n \in N, i \in I, k \in \{1, 2, \cdots, |K_n|\} \tag{6.20}$$

$$[e_{ia}, l_{ia}] = U_i^{opt}/ST_i \qquad \forall i \in I, a \in \{1, 2, \cdots, |U_i|\}, [e_{ia}, l_{ia}] \in U_i^{opt} \tag{6.21}$$

Thus, for a single task O_{ima} fulfilled by MPL m at parking space i within the a^{th} sub-interval can be represented by Equation (6.22). Following this logic, a task list O that includes all subtasks O_{ima} can be obtained once all parking spaces are enumerated .

$$O_{ima} = \{q_{ia}, m, i, [e_{ia}, l_{ia}]\} \qquad \forall i \in I, m \in M, a \in \{1, 2, \cdots, |U_i|\} \tag{6.22}$$

6.3.2 Structure Design of the State

By obtaining the tasks list for all MPLs, we construct the state of HQM (Figure 6.4). Each column that maps to variable x_1 and x_2 corresponds to a subtask $O_{ima} \in O$, i.e., the length of either x_1 or x_2 denotes the total number of tasks $|O|$ to be performed by MPLs. Each element of variable x_1 represents the unique identification number (ID) of an MPL m that perform the corresponding subtask. For example, the subtask bearing the ID with 0 will be executed by the MPL with ID is 0. The elements within the variable x_2 denote the execution ID for subtask O_{ima} so that forms the task execution sequence (e.g., the subtask with ID is 0 will be executed first followed by the subtask with ID is 1). Once the variable x_1 and x_2 are determined, we

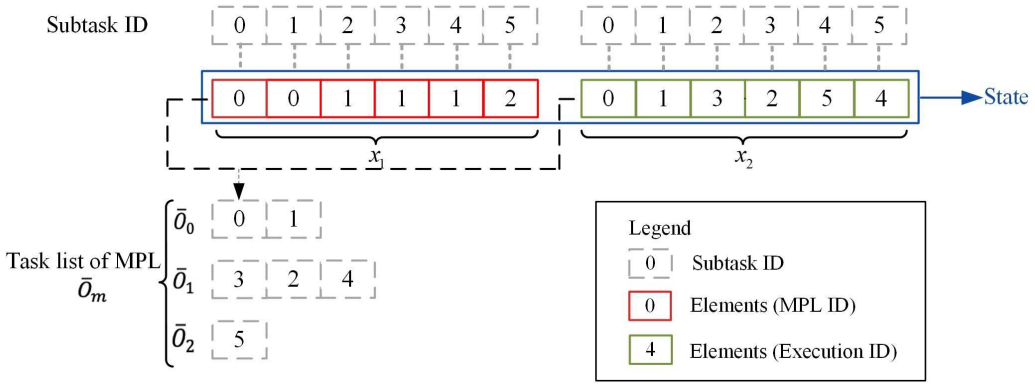


Figure 6.4: Illustration of a state

introduce $\bar{O}_m$ to represent the task list that assigns to a specific MPL m ($m \in M$) where $O_{ima} \in \bar{O}_m$ is satisfied. To add a subtask to the corresponding task list $\bar{O}_m$ of the MPL m in a sequential manner, we combine the inquires for the ID from x_1 and x_2. For instance, given that both subtasks 0 and 1 are assigned to $\bar{O}_0$ (according to x_1), subtask 0 has the priority for execution since the execution ID of subtask 0 is less than that of subtask 1 in x_2.

The element of the variable x_1 is produced by a random integer generator with the range from 0 to $|M|$, while the elements of the variable x_2 is generated within the range from 0 to $|O|$. It is permissible to assign the same MPL to different subtasks O_{ima}, therefore the value of elements can be the same within the variable x_1. However, the task ID should never be repeated, as each subtask O_{ima} is distinct and is executed only once by an MPL m.

A state will map to the actual solution following a sequential manner that decomposes variable x_1 and x_2 as mentioned above. The state will firstly be decomposed by the decoding algorithm (see Appendix H) to obtain the task list $\bar{O}_m$ (with task execution sequence) for each MPL m. Once $\bar{O}_m$ is determined, it serves as an input for route generation algorithm (see Algorithm 4) in which the delivery route of MPL m will be developed. The developed route for MPL m will be finally input to the decoder of HQM (see Algorithm 5) for reward calculation. The above principle is applied to decode and evaluate each state individually, and new states are generated based on reward evaluation until the optimal solution is obtained.

6.3.3 Hybrid Q-Learning Network Algorithm (HQM)

The HQM combines global and local search to generate a set of feasible solutions and obtains the optimal result by updating the Q-value iteratively.

In addition to the state $s = (x_1, x_2)$ defined previously, we define $Agent = \{s_1, \cdots, s_p\}$ as an agent that consists of multiple states s, which provides a set of feasible solution for the problem. The action A_t is defined as a decision behaviour in which *Agent* selects a policy to obtain a better solution/state s in the next timestep. A_t is sampled from the right-hand side of Equation (6.24) during training time and obtained by ε-greedy search of the Q-Learning and global-local search mechanisms. The goal is to maximise the reward R (Equation 6.23) obtained from a set of generated MPL route G:

$$R(Agent|G) = max\{\frac{1}{f_{s_1}(x_m, y_{ijm}, z_{ima})}, \cdots, \frac{1}{f_{s_p}(x_m, y_{ijm}, z_{ima})}\} \qquad \forall s_1, \cdots, s_p \in Agent \tag{6.23}$$

The proposed HQM is parameterised by θ_{x_1} and θ_{x_2}, corresponding to x_1 and x_2 respectively. It updates a strategy $p_{\theta_{x_1},\theta_{x_2}}$ that promotes *Agent* deriving the optimal solution/state s^* by refreshing variable x_1 and x_2. It is a stochastic policy that updates each element within a state and can be factorised as:

$$p_{\theta_{x_1},\theta_{x_2}}(Agent|G) = \prod_{t=1}^{t} p_{\theta_{x_1},\theta_{x_2}}[Agent(t+1)|Agent(t), G] \tag{6.24}$$

To maximise the objective in Equation (6.23), HQM updates every component on the right-hand side of Equation (6.24) based on the Q-value matrix in every timestep t. A new/updated state $s^{'}$ in the next timestep $t+1$ will resemble the current state s with the maximum Q-value in the current timestep t, indicating that the new state $s^{'}$ retains the elements that contributed significantly to the optimal solution in the current timestep t. To accelerate the search for the optimal solution/state s^*, HQM generates multiple *Agent* for the global search.

Encoder

The encoder consists of a random generator and a permutation pool to produce *Agent* as a normalised input of HQM. As shown in the elements in Figure 6.5, each element within x_1 and

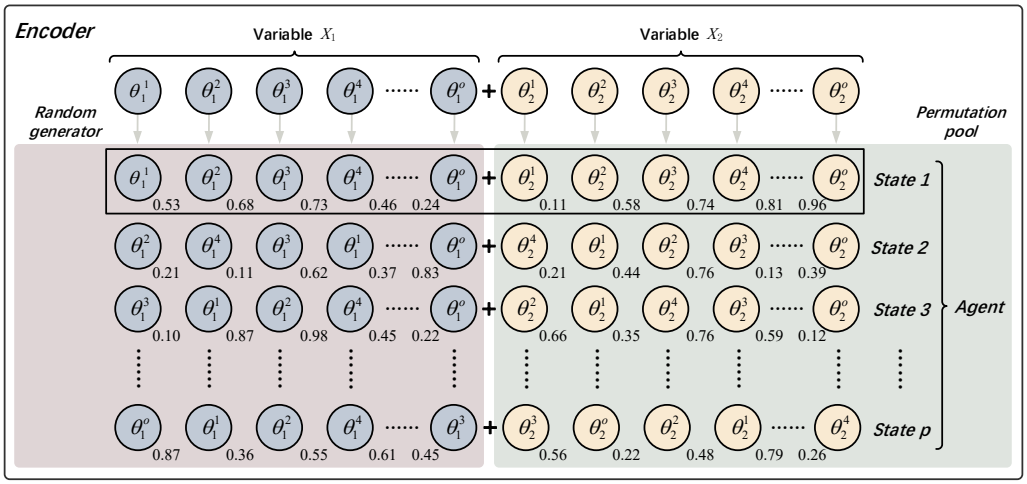


Figure 6.5: Structure of the encoder

x_2 is represented by θ_x^o ($x \in \{x_1, x_2\}$, $o \in \{1, \cdots, |O|\}$). The value next to the elements of x_1 and x_2 represents the probability of each current element that is remain selected by the current subtask o in the next timestep $t+1$, and will be updated by the attention layer embedded within the decoder at each timestep t (will be introduced in *Attention-based Decoder* Section).

Route Generation & Adjustment Strategies

The route generation mechanism is based on utilising *Agent* produced by the encoder. Specifically, each state s within *Agent* corresponds to the route for MPLs. However, since the state s is generated randomly, the resulting route may violate time-window requirements of tasks. We then introduces route adjustment strategies to resolve such time-window conflicts that occur during the route generation phase.

Mechanism of Route Generation

The route generation assigns and sequences tasks for MPLs based on the principle of subtask-MPL paring within the state s (as shown in Figure 6.4). The resulting route will be applied in the global & local search mechanism (will be introduced in *Policy for Action Selection* Section) for adjustment such that obtain better solution (as shown in the global & local search diagram in Figure 6.3). To capture the route generation process for MPLs, we modified the lemma developed by Schwerdfeger & Boysen (2020) and prove that for any given subtask O_{ima}, the task starting/ending time can be adjusted to execute the task as early as possible such that

forms an optimal task O^*_{ima}.

Lemma 6.1 *For any subtask O_{ima} defined in Equation (6.22), there is an optimal task starting time t^*_{ima} and ending time $\hat{t}^*_{ima}$ adjusted from O_{ima} to form an optimal task O^*_{ima}, where t^*_{ima} and $\hat{t}^*_{ima}$ are as follows:*

i) $t^*_{ima} = max\{\hat{t}_{(i-1)ma} + t_{(i-1)i}, e_{ia}\} \qquad \forall m \in M, a \in \{1, \cdots, |U_i|\}, e_{ia} \in [e_{ia}, l_{ia}]$

ii) $\hat{t}^*_{ima} = max\{max\{t_{ima}, E_{nk}\} + ST_i\} \qquad \forall m \in M, n \in N, a \in \{1, \cdots, |U_i|\}, E_{nk} \in [E_{nk}, L_{nk}]$

Proof 6.1 *To simplify the process, we assumes that each customer n only has a single location ($|K_n| = 1$), but the proof is still valid for the multiple-locations situation. Following the nomenclature in Section 6.2, the optimal starting time of MPL m in the first parking space is:*

$$t_{1ma} = max\{e_{1a}, t_{01}\} \qquad \forall m \in M, a \in \{1, \cdots, |U_i|\}, e_{1a} \in [e_{1a}, l_{1a}] \tag{6.25}$$

Since the MPL m departs from the depot to the first parking space, the optimal starting time of MPL m is the maximum value between the travel time and the earliest available task starting time in the first parking space.

We assume that the MPL m serves the customers n as early as possible. Thus, MPL m can leave the first parking space as soon as the task is completed:

$$\hat{t}_{1ma} = \{max(t_{1ma}, E_{nk}) + ST_1\} \qquad \forall m \in M, n \in N, a \in \{1, \cdots, |U_i|\}, E_{nk} \in [E_{nk}, L_{nk}] \tag{6.26}$$

Similarly, the earliest task starting time and ending time in the second parking space are:

$$t_{2ma} = max\{\hat{t}_{1ma} + t_{12}, e_{2a}\} \qquad \forall m \in M, a \in \{1, \cdots, |U_i|\}, e_{2a} \in [e_{2a}, l_{2a}] \tag{6.27}$$

$$\hat{t}_{2ma} = \{max(t_{2ma}, E_{nk}) + ST_2\} \qquad \forall m \in M, n \in N, a \in \{1, \cdots, |U_i|\}, E_{nk} \in [E_{nk}, L_{nk}] \tag{6.28}$$

Following the above logic, the earliest task starting t^*_{ima} and ending time $\hat{t}^*_{ima}$ for each subtask O_{ima} can be represented as:

$$t^*_{ima} = max\{\hat{t}_{(i-1)ma} + t_{(i-1)i}, e_{ia}\} \qquad \forall m \in M, a \in \{1, \cdots, |U_i|\}, e_{ia} \in [e_{ia}, l_{ia}] \tag{6.29}$$

$$\hat{t}^{*}_{ima} = max\{max\{t_{ima}, E_{nk}\} + ST_i\} \qquad \forall m \in M, n \in N, a \in \{1, \cdots, |U_i|\}, E_{nk} \in [E_{nk}, L_{nk}] \tag{6.30}$$

Adjustment Strategy and Route Generation Algorithm

The route generator described in the previous section does not consider the capacity constraint and stochastic events (e.g., traffic congestion, service delay) in real implementation, which violates the constraint outlined in Equation (6.13) and the time-window requirement of a single task, respectively. Thus, a route adjustment strategy is necessary to ensure the feasibility of all MPL routes. Algorithm 4 outlines the procedure of route generations and adjustments.

The route adjustment strategy consists of two algorithms: a back-to-depot (BTD) strategy, and a holding-at-current-parking-space (HCPS) option. The former allows the MPL to reload parcel and postpone deliveries to meet the time-window and capacity constraints. However, this strategy results in an extra distance, impacting the operating costs and delivery efficiency. This process is described in Equation (6.31) where t_{i0} and t_{0j} represent the time consumption of MPL from parking space i to parking space j via the central depot.

The latter enables MPLs to wait at their current parking space, postponing the start time of the next task. Thus, the MPL m will remain at the current parking space i until the latest ending time l_{ia} of the a^{th} sub-interval, as shown in Equation (6.32). In contrast to BTD, HCPS is only used to resolve time-window conflicts.

When the route adjustment algorithm is implemented to solve the aforementioned conflicting scenarios, it also initiates dynamic route planning and adjustment for the upcoming MPL m. This is due to the interdependencies and information interactions among each sub-route of MPL m, which collectively form the final solution. Specifically, as shown in lines 6 to 18 of Algorithm 4, when adding a new subtask to the current route of MPL m, the algorithm will detect whether the formed route violates the demand and time window constraints in each iteration. If violates, the algorithm will apply the corresponding adjustment strategy to resolve the conflict in a dynamic manner.

$$t_{jma} = t_{ima} + ST_i + t_{i0} + t_{0j} \qquad \forall i, j \in I, m \in M, a \in \{1, 2, \cdots, |U_i|\} \tag{6.31}$$

Algorithm 4: Route generation and adjustment algorithm

Input: Task set $\bar{O}_m$ for $m \in M$, Locker speed v, Locker capacity Q
Output: Delivery route of MPLs

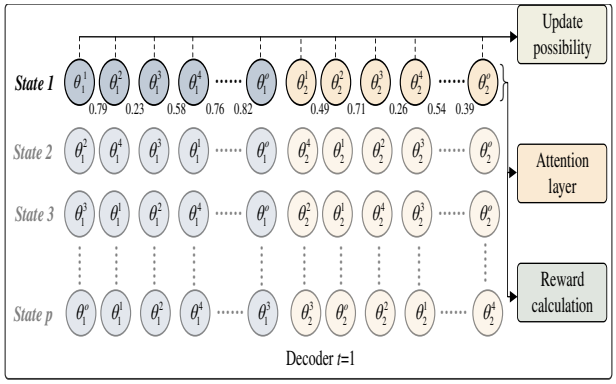
Initialise O_{ima}, q_{ia}, e_{ia}, l_{ia}, E_i, L_i, ST_i
Calculate travel time t_{ij} between parking space i and j for all parking spaces
for each sub-task O_{ima} in $\bar{O}_m$ **do**
  Add depot to current task list $\bar{O}_m$ `// Depart from depot`
  visited_route_list $\leftarrow \emptyset$ `// Record the route of MPLs`
  for *next* in O_{ima} **do**
    current_parking_space $\leftarrow \bar{O}_m[O_{ima}]$ `// Map the current task of MPL` m `to parking space`
    next_parking_space $\leftarrow \bar{O}_m[next]$ `// Map the next task of MPL` m `to the next parking space`
    Add *current_parking_space* to *visited_route_list*
    Update t_{ima}, q_{ia}, e_{ia}, l_{ia}
    travel_time $\leftarrow$ lialg.norm(*current_parking_space* - *next_parking_space*) / v
    if $t_{ima} + ST_i + t_{ij} < e_{ja}$ and $l_{ia} + t_{ij} \geq e_{ja}$ **then**
      Adjust task starting and ending time based on Equation (6.32)`// Apply HCPS`
      Update t_{jma}, q_{ja}, e_{ia}, l_{ja}, E_j, L_j, ST_j
      Add *next_parking_space* to *visited_route_list*
    else if $t_{jma} < e_{ja}$ or $q_{ia} > Q$ **then**
      Adjust task starting and ending time based on Equation (6.31)`// Apply BTD`
      Update t_{jma}, q_{ja}, e_{ia}, l_{ja}, E_j, L_j, ST_j
      Add *next_parking_space* to *visited_route_list*
    else if satisfy constraints **then**
      Add *next_parking_space* to *visited_route_list*
    end if
  end for
end for
return *visited_route_list*

$$t_{ima} = l_{ia} + t_{ij} \qquad \forall i, j \in I, m \in M, a \in \{1, 2, \cdots, |U_i|\} \tag{6.32}$$

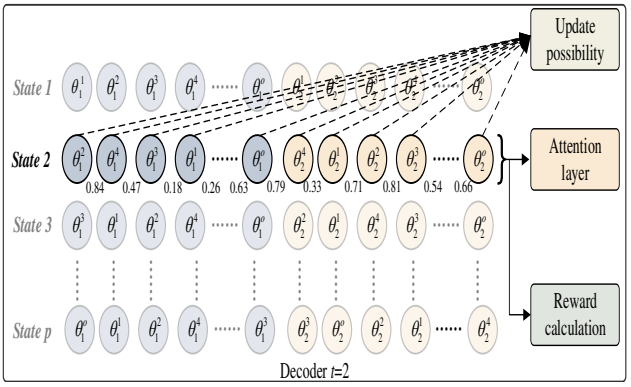
Attention-based Decoder

The attention-based decoder is developed to calculate the reward R of the route generated in *Mechanism of Route Generation* Section and updated the probability of each current element that is remain selected by the current subtask O_ima in the next timestep $t+1$ through the Q-value matrix embedded within the attention layer (see *Mechanism of Attention Layer* Section). The reward calculation is based on Equation (6.23), given the route generated.

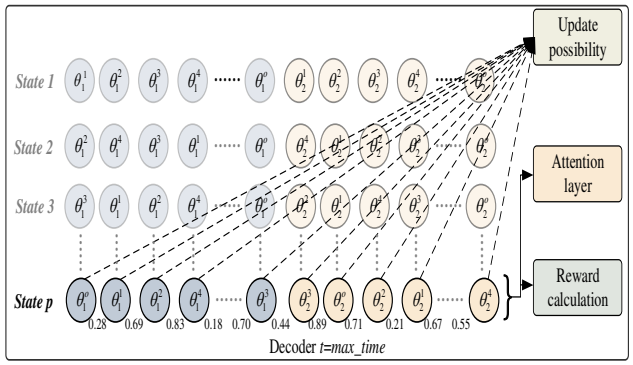
Figure 6.6a to 6.6d demonstrate the decoding procedure for an *Agent*. As shown in Figure 6.6a, the decoder initially decomposes the first state s_1 in the sub-step $t' = 1$ within the iteration



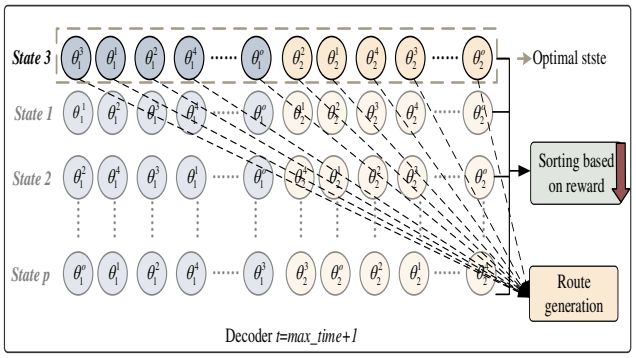
(a) Decoding progress in step $t' = 1$ of iteration t



(b) Decoding progress in $t' = 2$ of iteration t



(c) Decoding progress in $t' = max_time$ of iteration t



(d) Decoding progress in $t' = max_time + 1$ of iteration t

Figure 6.6: Structure of the decoder

t by applying a decoding algorithm (see Appendix H) and calculating the corresponding reward R, followed by updating the probability of each element from state s_2 to s_p within the iteration t. Analogously, the decoding procedure lasts until all states within the *Agent* have been completed (Figure 6.6b and 6.6c). Finally, all states within the decoder are sorted according to the reward obtained, and the optimal state s^* is then extracted to determine the optimal delivery route by executing Algorithm 4 (Figure 6.6d).

Mechanism of Attention Layer

The state-action space of HQM can be represented by a lookup table (Figure 6.7). The size of the lookup table equals $O \times T$, where O equals the length of the task list of a state s, while T equals the total number of stepsize/episode of the whole HQM process. Given this property, the size of the lookup table increases with the problem size exponentially.

The attention mechanism is developed to tackle such a curse of dimensionality, in which

we decompose the state-action space into multiple low-dimensional state-action combination chains linked by the Q-value matrix. We define two Q-value matrices in Equations (6.33) and (6.34), which correspond to x_1 and x_2 within a state s, respectively. Each value within the matrix reflects the contribution of the element to the final reward.

Furthermore, each value within the matrix indicates the interdependence between two adjacent actions, where the larger the value, the closer the two adjacent actions are connected. The new state $s^{'}$ updated in the next action A_{t+1} inherits the elite elements (the elements that contribute to the greater reward) of the current state s, which have greater Q-values in the current action A_t. Therefore, the solution can be enhanced while inheriting the elite elements of the current state s.

$$Q(\theta^o_{x_1}) = \begin{pmatrix} Q_{11}(\theta^1_{x_1}) & \cdots & Q_{1m}(\theta^o_{x_1}) \\ \vdots & \ddots & \vdots \\ Q_{m1}(\theta^1_{x_1}) & \cdots & Q_{mm}(\theta^o_{x_1}) \end{pmatrix} \qquad \forall o \in \{1, \cdots, |O|\}, m \in M \tag{6.33}$$

$$Q(\theta^o_{x_2}) = \begin{pmatrix} Q_{11}(\theta^1_{x_2}) & \cdots & Q_{1o}(\theta^o_{x_2}) \\ \vdots & \ddots & \vdots \\ Q_{o1}(\theta^1_{x_2}) & \cdots & Q_{oo}(\theta^o_{x_2}) \end{pmatrix} \qquad \forall o \in \{1, \cdots, |O|\} \tag{6.34}$$

Based on the Q-value matrices defined above, we determine the possibility of the current element being selected for the current task o in the following timestep $t+1$, denote as $p(\theta^o_{x_1})$ and $p(\theta^o_{x_1})$. These are calculated as follows:

$$p(\theta^o_{x_1}) = \frac{exp(Q(\theta^o_{x_1}))}{\sum_{o=1}^{O} exp(Q(\theta^o_{x_1}))} \qquad \forall o \in \{1, \cdots, |O|\} \tag{6.35}$$

$$p(\theta^o_{x_2}) = \frac{exp(Q(\theta^o_{x_2}))}{\sum_{o=1}^{O} exp(Q(\theta^o_{x_2}))} \qquad \forall o \in \{1, \cdots, |O|\} \tag{6.36}$$

Mechanism of Q-value Update

To update the Q-value matrix in every timestep t, we propose the following Q-value function for each element of a state s within *Agent*:

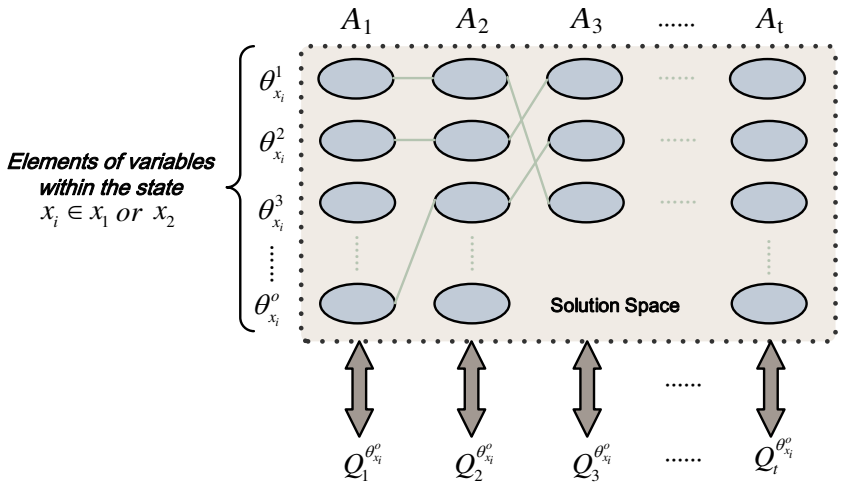


Figure 6.7: Lookup table for state-action space

$$Q_{t+1}^{\theta^o_{x_i}}(s_t^{x_i}, a_t^{x_i}) = Q_t^{\theta^o_{x_i}}(s_t^{x_i}, a_t^{x_i}) + \alpha \left[R_t(s_t^{x_i}, s_{t+1}^{x_i}, a_t^{x_i}) + \gamma \max_{a^{x_i} \in A_t} Q_t^{\theta^o_{x_i}}(s_{t+1}^{x_i}, a_t^{x_i}) - Q_t^{\theta^o_{x_i}}(s_t^{x_i}, a_t^{x_i}) \right] \tag{6.37}$$

where $x_i \in \{x_1, x_2\}$ and $o \in O$, $R_t(s_t^{x_i}, s_{t+1}^{x_i}, a_t^{x_i})$ represents the reward obtained from state $s_t^{x_i}$ by executing action $a_t^{x_i}$ towards a state $s_{t+1}^{x_i}$ in timestep t, α represents the learning rate, and γ represents the discount factor.

The update procedure can be regarded as a Markov decision process that derives satisfactory results through updating the Q-value between two adjacent actions. For each action $a_t^{x_i} \in A_t$ ($t \in \{1, \cdots, max_time\}, x_i \in \{x_1, x_2\}$) of the variable, there is a corresponding Q-value ($Q_t^{\theta^o_{x_1}}$ and $Q_t^{\theta^o_{x_2}}$) associated with action $a_t^{x_i}$. Let A_t denotes the action profile that integrates both actions of variable $x_1(a_t^{x_1})$ and$x_2(a_t^{x_2})$ in timestep t (i.e., $A_t = a_t^{x_1} \cap a_t^{x_2}$).

The action A_t selects the direction of migration that will seek a better solution in the next timestep $t+1$ based on a Q-value such that migrate into the new policy. Once A_t is determined and executed by the *Agent*, and the next action A_{t+1} will be selected according to the estimated Q_{t+1} (Figure 6.7).

Policy for Action Selection

Q-Learning suffers from the trade-off between exploitation and exploration of the information feedback obtained from previous actions (Sutton & Barto, 2018). A strong emphasis on

information exploration may derive the optimal solution but may sacrifice the speed of convergence. While favouring information exploitation can improve convergence speed at the expense of optimality.

As part of the HQM, we propose integrating Q-value-based global search with a local search approach to define the policy for action selection, leading to a near-optimal solution while improving convergence speed. Concretely, the global search mechanism first updates each element within the state s of *Agent* to generate a new state s' given a specific action A_t, in which the resulting reward is denoted as R'. Next, the local search mechanism adjusts the subset of elements within an individual state s' updated by the global mechanism to create another new state s'', in which the resulting reward is represented as R''. The HQM outputs either s' or s'' for the final solution follows $max\{R', R''\}$.

By implementing a combined global and local search approach, the *Agent* and corresponding Q-value matrices will be updated accordingly. The iteration stops once the error of the Q-value matrices generated by two adjacent actions A_t and A_{t+1} is less than 1×10^{-8}. The state s^* with the greatest reward R^* among all *Agent* is defined as the optimal solution of MPLP.

In terms of global search mechanism, the migration between the two adjacent actions A_{t-1}, A_t updates *Agent* based on the Q-value obtained from the previous action A_{t-1}. The policy of action selection is based on the ϵ-greedy search which can be defined as follows:

$$a_{t+1}^{x_i} = \begin{cases} \underset{a_t^{x_i} \in a_t}{argmax}\, Q_t^{\theta_{x_i}^o}(s_{t+1}^{x_i}, a_t^{x_i}) & \tau < \epsilon \\ a_t^{x_i'} & \tau \geq \epsilon \end{cases} \tag{6.38}$$

Where ϵ denotes the greedy factor, τ denotes the random value between 0 and 1, and $a_t^{x_i'}$ denotes the action that is being selected based on the normalised Q-value matrices within the global scope. The greater the Q-value is, the greater probability of the corresponding element inheriting an elite fragment of the current state s.

The local search operation focuses on adjusting the subset of the variable x_1 and x_2 within an individual state s_i to create a new state s_i'. The adjustment for the elements within variable

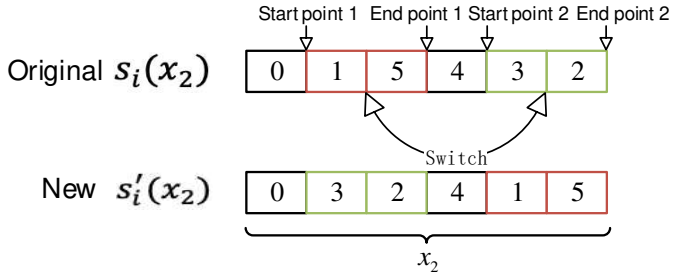


Figure 6.8: Illustration of adjustment for variable x_2

x_1 is represented by the following:

$$s'_i(x_1) = |s_i(x_1) + \vec{\omega} \cdot (s_i(x_1) - s_{neighbour}(x_1))| \tag{6.39}$$

Where $s_{neighbour}$ represent the neighbour state in $\{s_{i+1}, s_{i-1}\}$. $\overrightarrow{\omega}$ is a vector ($|\overrightarrow{\omega}| = |x_1|$) where the i^{th} element $\overrightarrow{\omega}_i$ of $\overrightarrow{\omega}$ depends on the relationship between the corresponding element o and o' within the two neighbour states $s_{neighbour}$, which is captured by the piecewise function (6.40). This allows the modification of only subsets of the elements in x_1. Since all elements within a state s should be non-negative, we take absolute values for all elements within the new state $s'_i(x_1)$.

$$\overrightarrow{\omega}_i = \begin{cases} -1 \quad or \quad 1 \qquad if \quad o < o' \\ 0 \qquad otherwise \end{cases} \qquad \forall i \in \{1, \cdots, |x_1|\}, \overrightarrow{\omega}_i \in \overrightarrow{\omega} \tag{6.40}$$

Since the elements within the variable x_2 is unique and non-repeated, the adjustments for the elements within variable x_2 focus on the subset exchange within the same state to guarantee the uniqueness of elements (Figure 6.8). The start points and end points are generated randomly follows $x_2 \sim U(0, |O|)$. A new state $s'_i(x_1, x_2)$ can be obtained by integrating new variable $s'_i(x_1)$ and $s'_i(x_2)$.

The implementation procedure of HQM is shown in Algorithm 5.

Algorithm 5: Hybrid Q-Learning network method (HQM)

Input: Number of MPLs $|M|$, All task list O, Number of agents $|Agent|$, Timestep T
Output: The optimal state $s = (x_1, x_2)$

Initialise learning rate α, discount factor γ, random value τ, ω, greedy factor ϵ, Q-value matrices $Q_t^{\theta_1^o}$, $Q_t^{\theta_2^o}$
for i in range($|Agent|$) **do**
  $x_1 \leftarrow \mathbb{N} \sim U(0, |M|)$
  $x_2 \leftarrow$ Permutating task list O
  $s = (x_1, x_2)$ `// Generate initial state` s
  Execute Decoding Algorithm, Algorithm 4
end for
Rank and record all state s based on the reward obtained
for t in range(T) **do**
  Record the current best state s^* and the corresponding reward R^* of $Agent$
  for o in range($|O|$) **do**
    Normalise Q-value matrices $Q_t^{\theta_1^o}$, $Q_t^{\theta_2^o}$
  end for
  for i in range($|Agent|$) **do**
    for o in range($|O|$) **do**
      $s'(x_1', x_2') \leftarrow$ Equation (6.38) `// Implement global search and obtain new state` s'
      Execute Decoding Algorithm, Algorithm 4
    end for
    $R' \leftarrow$ Equation (6.23) `// Calculate the reward` R' `for the new state` s'
    if $R' > R^*$ **then**
      $s^*, R^* \leftarrow s', R'$ `// Update the optimal state` s^* `and reward` R^*
    end if
  end for
  for i in range($|Agent|$) **do**
    $s''(x_1'', x_1'') \leftarrow$ Equation (6.39) `// Implement local search and obtain another new state` s''
    Execute Decoding Algorithm, Algorithm 4
    $R'' \leftarrow$ Equation (6.23) `// Calculate the reward` R'' `for the new state` s''
    if $R'' > R^*$ **then**
      $s^*, R^* \leftarrow s'', R''$ `// Update the optimal state` s^* `and reward` R^*
    end if
  end for
  $Q_t^{\theta_1^o}, Q_t^{\theta_2^o} \leftarrow$ Equation (6.37) `// Update Q-value matrices`
  if satisfy convergence constraints **then**
    break
  end if
end for
return s^*

6.4 Results and Discussion

This section evaluates the performance of HQM for the proposed MPLP. To validate the proposed mathematical model and identify the solution gap between HQM and the exact approach counterpart, we first compares HQM's derived solution and computation time with the exact approach. We then apply a Wilcoxon signed-rank to evaluate the result significance between exact approach and HQM. Since the main purpose of this study is to emphasise the advantage of HQM in solving large problem instances compared with meta-heuristic approaches, we next compare the proposed HQM with GA based on the following criteria: a) the performance indicators derived from the objective function (Equation 6.8), b) the demand fulfilment rate, c) the performance of the solution improvement, d) reward acquisition, e) learning convergence speed and f) the computing efficiency. The HQM vs. GA experiments are all executed under both BTD & HCP strategies are applied.

Beyond the performance comparison between HQM and GA, this section further investigates the effect of key factors on influencing different cost units of the objective function. Lastly, this section provides managerial insights and fruitful discussions regarding the implementation of HQM algorithm to benefit academia and the logistics profession.

6.4.1 Generation of Problem Instances

Benchmark instances commonly used for VRP experiments (e.g., Solomon problem instances) are not suitable for the proposed MPLP, since the latter involves customers' trajectory within a day, which is out of the scope of the VRP. On the other hand, since prior MPLP studies (Li et al., 2021a; Schwerdfeger & Boysen, 2020) neither consider the dynamic route adjustment technique as this study nor disclose the data set of problem instances and precise optimal solution of different cost units, we cannot compare the derived results obtained by HQM with the existing studies under the same dimension. Therefore, we create problem instance and analyse the impact of network size on the deployment and operation of MPLs. Problem instance, detailed solutions, and relevant code of experiment are provided in *https://github.com/Yubin-Liu/Hybrid-Q-Learning-Network-Approach-for-MPLP.git*.

In total, 96 problem instances are evaluated based on 24 networks. Each network has a predefined number of MPL parking locations (ranging from $I_i = 5$ to $I_i = 10$), and a set of customer locations N_n close to these I_i ($N_n = \{5, 10, 15, 20\}$).

The HQM and GA code are implemented in Python 3.8.5 on an Apple M1 (3.20 GHz) processor with 16GB RAM, while the exact approach is implemented on the same machine via Gurobi solver. To ensure that the results are comparable, we configured the two algorithms' implementation under the same timestep/iteration and agent/population size, and with similar parameters in terms of learning rate/elite rate and mutation rate to the degree possible. The corresponding HQM and GA parameters are presented in Table 6.3. The setup of the problem instance is as follows:

- **Generation of customer demands:** the number of parcels per customer $\bar{q}_n$ ($n \in N$) is defined as a uniform random distribution $U \sim (1, 4)$.

- **Generation of locations:** for each location of customer $n \in N$, the coordinates are randomly generated within a specified MPL radius r_m of a parking space $i \in I_i$ to easily form different clusters, in other words $d_{ni}^k \leq r_m$ given that $d_{ni}^k = \sqrt{(X_{nk} - X_i)^2 + (Y_{nk} - Y_i)^2}$ ($X_i \sim U(0, X_{max}), Y_i \sim U(0, Y_{max})$) , where d_{ni}^k denotes the distance between the parking space i and the k^{th} stopover of the customer n.

- **Customer stopover generation:** the number of stopovers $|K_n|$ of a customer n within the planning horizontal follows $|K_n| \sim U(1, 3)$. The maximum walking distance ρ_n (km/customer) of customer n is generated by following normal distribution $\rho_n \sim N(0.5, 0.1)$ (Schwerdfeger & Boysen, 2020).

- **Generation of time interval:** generate non-overlapping time intervals for each customer n. The time interval $[E_{nk}, L_{nk}]$ of the k^{th} locations of a customer n is generated by a random generator, in which $E_{nk}, L_{nk} \in T_{list} = \{8 \times 60, \cdots, 18 \times 60\}$ (corresponding to working hours from 8:00 to 18:00) with the length of T_{list} equals 10. Meanwhile, we guarantee that the interval length of time window $|L_{nk} - E_{nk}|$ follows normal distribution $|L_{nk} - E_{nk}| \sim N(60, 5)$ subject to $L_{nk} > E_{nk}$. The generation of parking spaces' time intervals is similar to the above procedure with an interval length $|L_i - E_i|$ following

Table 6.3: Parameters setting in model validation

Parameters in mathematical model		
Parameter	Symbol	Value
Service radius of MPLs	$r_m[km]$	5
Capacity of MPLs	$Q[unit]$	20
Travel speed of MPLs	$v[km/hr]$	40
Fixed cost of locker m	$F_m[£/locker]$	20,000
Unit cost travelling from parking space i to j	$C_{ij}[£/km]$	0.5
A large indefinite number	$\phi[-]$	10^8
Service time of parking space i	$ST_i[min]$	10
Weights of cost units	$W_1 : W_2 : W_3[-]$	10:1:5[a]
Parameters in customer properties		
Maximum walking range	$\rho_n[km]$	$\rho_n \sim N(0.5, 0.1)$
Length of time interval of customer	$T_s = \lvert L_{nk} - E_{nk} \rvert [min]$	$T_s \sim N(60, 5)$
Parameters in parking spaces		
Length of time interval of parking spaces	$T_p = \lvert E_i - L_i \rvert [min]$	$T_p \sim N(50, 5)$
Service buffer time	$\delta[min]$	10
Parameters in HQM		
Number of actions/timestep	$A_t[-]$	1000
Learning rate	$\alpha[-]$	$0.9 \times e^{-(\frac{current_timestep}{total_timestep})}$
Discount factor	$\gamma[-]$	0.9
Greedy factor	$\epsilon[-]$	$\epsilon \sim U(0.1)$
Random variables	$\tau, \omega[-]$	$\tau, \omega \sim U(0.1)$
Total states within an agent	$Agent[-]$	100
Parameters in GA		
Number of iteration/timestep	$Iteration[-]$	1000
Population size	$Pop[-]$	100
Possibility of crossover	$P_{cross}[\%/Pop]$	50
Proportion of elite size	$P_{elite}[\%/Pop]$	5
Mutation rate	$P_{muta}[\%/Pop]$	5

[a]MPLs' fleet size and service delay are given higher weightings because the deployment of MPLs requires high initial costs, while the service delay has a significant effect on the market share for parcel deliveries Iwan et al. (2016). Nevertheless, the weights of different cost units may be adjusted in accordance with real operations.

normal distribution $|L_i - E_i| \sim N(50, 5)$. We ensure that at least one parking space suits each customer.

6.4.2 Performance Comparison between HQM and Exact Approach

We utilise the Gurobi optimiser to solve the problem instances via the exact approach, and the obtained results are compared with the HQM. The solver was terminated after 2 hours or when it has found an optimal solution. The comparison is evaluated based on two aspects: i) the best value of the objective function, cost units, and the computation time; ii) the significance between the results obtained by Gurobi solver and HQM.

Appendix I.1 shows the best reward obtained by Gurobi solver and HQM and the corresponding computation time. Although the Gurobi solver generates superior solutions for small problem sizes (e.g., nodes $\leq$ 100), the computation time is relatively long (average computation time = 3199.74s). While HQM achieve high-quality solutions in large problem instances with significant shorter computation time (average computation time = 609.05s). The average best result between HQM and Gurobi is close with an average gap of 6.21%. Such a gap is influenced by the extreme values caused by HQM obtaining lower rewards in small problem instances (since the solving accuracy of HQM requires larger sample). In real operations, however, we conclude that HQM can achieve better results in shorter computation time than the Gurobi solver based on the results shown in Appendix I.1.

Appendix I.2 shows the best value of different cost units of Gurobi solver and HQM. Result reveals that the Gurobi solver achieves less MPLs deployment in small problem instances while HQM obtains greater distance savings in large problem instances. There is a minor difference between the Gurobi solver and HQM on average service delays (9.07 min of HQM vs. 10.49 min of Gurobi).

We conducted statistical tests to evaluate the performance of HQM statistically. Wilcoxon signed-rank tests were used to determine the significance of the difference between two groups of results obtained by HQM and the Gurobi solver. Wilcoxon signed-rank is a non-parametric statistical test to analyse two related samples (Wolpert, 2013) and was used for a VRP problem (Vincent et al., 2022). Tables 6.4 shows the result for Wilcoxon signed-rank tests in terms of the best reward and computation time, respectively. We used the confidence level of $\alpha = 0.05$ to test the hypothesis. It is concluded that the two paired samples differ significantly if the

Table 6.4: Result of Wilcoxon signed-rank test on best result obtained and computation time

Method	Test on Best Result/Reward Obtained		Test on Computation Time	
	Z	$p-value$(2-tailed)	Z	$p-value$(2-tailed)
HQM vs. Gurobi	1.5	0.145	-4.3	0.000*

*denotes the significant difference exists

significance level ($p-value$) is less than α.

In Table 6.4, the $p-value$ of the best result comparison is larger than α, implying that the performance of the HQM is close to the Gurobi solver in terms of solution quality. Additionally, it is conclude that the computation time of HQM and the Gurobi solver are significantly different since the $p-value$ is less than α. Finally, as a result of Tables 6.4 and Appendix I.1, we infer that the HQM significantly outperforms the the Gurobi solver in terms of solution quality and computation time for large problem instances.

6.4.3 Performance Comparison between HQM and Heuristics Approach

This section compared the performance of HQM with GA regarding the optimal configuration of MPLs (fleet size, travel distance, and service delay) and optimisation performance (reward acquisition, convergence speed, and computation time). We further apply advanced techniques (e.g., roulette wheel selection, elite strategy) to different operations within GA to ensure good computation performance and rigour in comparisons (the implement procedure of GA see Appendix J).

Optimal Configuration of MPLs

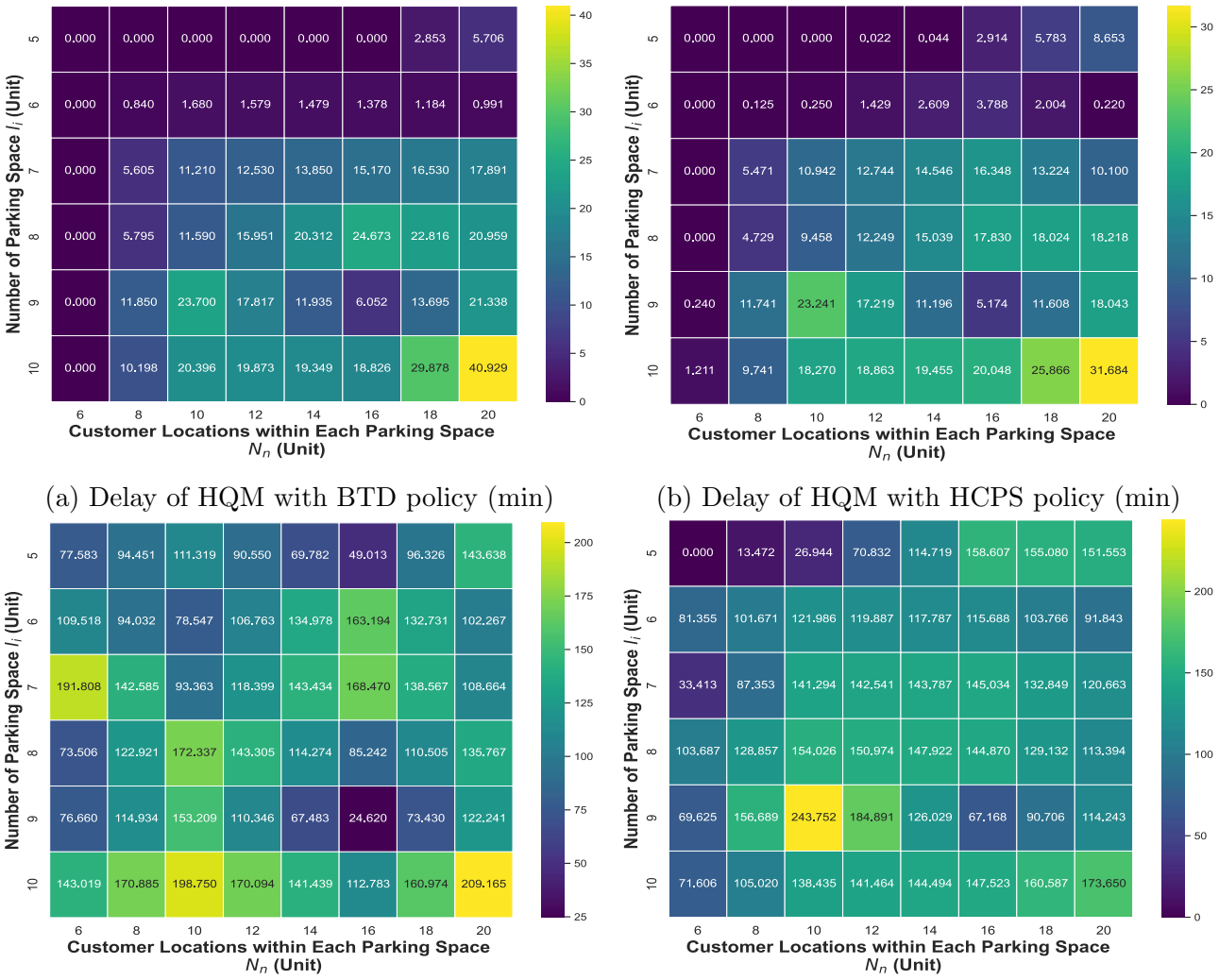
We first investigate the implementation results in fleet size and travel distance under different problem instances (see Appendix K.1). The evaluation is based on two dimensions: i) the performance of algorithms, and ii) the performance of route adjustment strategies. Results indicate that the maximum difference in achieving optimal fleet size between HQM and GA is ± 3 units (the maximum MPL fleet size is 34 units). In fact, the GA achieves a better

performance than the HQM in terms of minimising driving distances but causes higher service delay significantly.

Regarding the performance of BTD and HCPS route adjustment strategies, BTD achieves fewer MPL deployment in small-median size problem instances on average ($I_i \leq 8, N_n \leq 15$). However, such fleet size difference is compromised in large networks ($I_i \geq 9$) because HCPS is regarded as a more cost-efficient approach to resolve time window conflicts compared with deploying dense MPLs caused by BTD. With the increase in parking spaces and customer locations, the HCPS strategy allows for greater driving distance savings than BTD since fewer round trips are required between parking spaces and the central depot.

In terms of average service delay, the results obtained from the two algorithms are 10.10 min (HQM-BTD), 8.91 min (HQM-HCPS), 121.03 min (GA-BTD), and 113.77min (GA-HCPS), respectively. The average service delay obtained by HQM is therefore significantly lower than the GA's, contributing to higher reward (Figure 6.9a to 6.9d). GA, however, shows randomness and non-robust performance in optimising service delay across different network sizes, as GA produces high latency even at small problem instances (Figure 6.9c and 6.9d), while HQM enables on-time delivery in small network sizes (Figure 6.9a and 6.9b). Comparatively to BTD implemented by HQM and GA, HCPS reduces delays by 11.85% (HQM) and 6.00% (BTD), respectively.

We further evaluate the demand fulfilment in the first-delivery round (see Appendix K.2), i.e., the maximum demand that an MPL can fulfil in a single delivery trip. Operators anticipate a higher first-time delivery rate with smaller fleet sizes and shorter travel distances, leading to better utilisation of delivery fleets and reducing repeated deliveries. The result shows that HQM performs better than the GA, as HQM reduces the return to depot frequency. Concretely, the average first-time delivery fulfilment of HQM ranges from 15.5 units to 19.25 units (the maximum capacity is 20 units per MPL), corresponding to the fulfilment rate from 77.50% to 96.30% respectively. In addition, HCPS strategy obtains higher first-time demand fulfilment that BTD when $I_i \geq 8$.



(a) Delay of HQM with BTD policy (min)

(b) Delay of HQM with HCPS policy (min)

(c) Delay of GA with BTD policy (min)

(d) Delay of GA with HCPS policy (min)

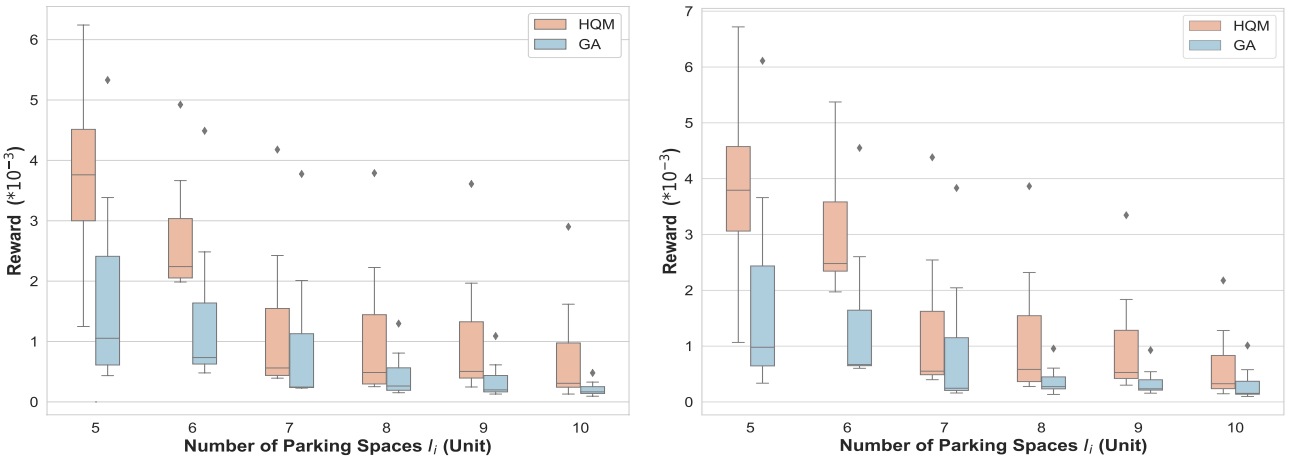
Figure 6.9: Service delay of HQM and GA with different route adjustment polices applied

Optimisation Performance of HQM

The optimisation performance of HQM is evaluated from three dimensions: (i) the final reward obtained, (ii) reward acquisition and convergence, and (iii) computation time. The final reward obtained represents the absolute ability to derive a final optimal solution, while the reward acquisition and convergence speed measure the improvement rate throughout the learning process.

Final Reward Obtained

Figure 6.10a and 6.10b show the final reward obtained from HQM and GA under different problem instances and route adjustment strategies. Results indicate that the proposed HQM



(a) Reward of applying BTD policy (b) Reward of applying HCPS policy

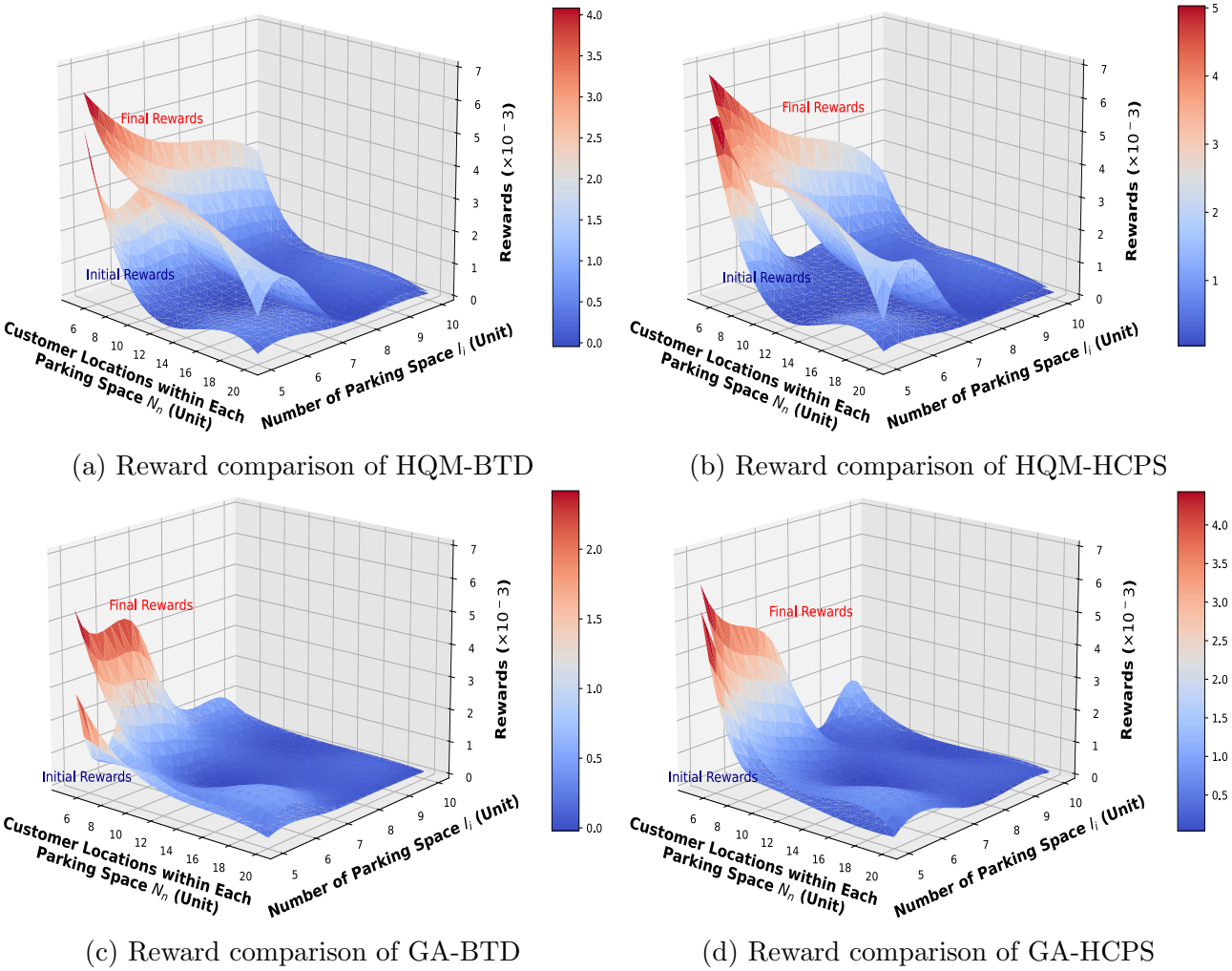
Figure 6.10: Reward obtained under different problem instances

combined with HCPS strategy yields the greatest reward with the value of 1.97×10^{-3}, followed by 1.90×10^{-3} (HQM-BTD), 1×10^{-3} (GA-HCPS), and 0.97×10^{-3} (GA-BTD).

The surface plot from Figure 6.11a to 6.11d illustrates the reward improvement of HQM and GA, where the bottom layer denotes the initial rewards, and the top layer denotes the final rewards optimised by the two algorithms. The calculation of the improvement based on the metric: $R_{t=max} - R_{t=1}/R_{t=1}$. In Figure 6.11b, the larger gap between the two layers within the small and medium problem instances ($I_i \leq 8, N_n \leq 18$) indicates that the HQM-HCPS achieves a greater reward improvement than the HQM-BTD (Figure 6.11a). Contrary to the HQM, the GA is merely good at solving small problem instances ($I_i \leq 7, N_n \leq 10$) and performed poorly when medium and large problem instances are involved (Figure 6.11c and 6.11d). Although HQM shows less reward improvement in large problem instances compared with small-medium networks due to the large gap of axis calibration in the figure, the HQM still perform better than GA, which can be proven by the concrete improvement rate shown in Appendix K.3.

Reward Acquisition and Convergence

Figure 6.12a and 6.12b illustrate the reward acquisition and convergence of HQM and GA. Algorithms are implemented under a large instance setting ($I_i = 10, N_n = 20$). The result concludes that GA tends to fall into a local optimum during the earlier iterations, resulting in a lower reward (Figure 6.12a). The reason for this can be explained by Figure 6.12b, which



(a) Reward comparison of HQM-BTD

(b) Reward comparison of HQM-HCPS

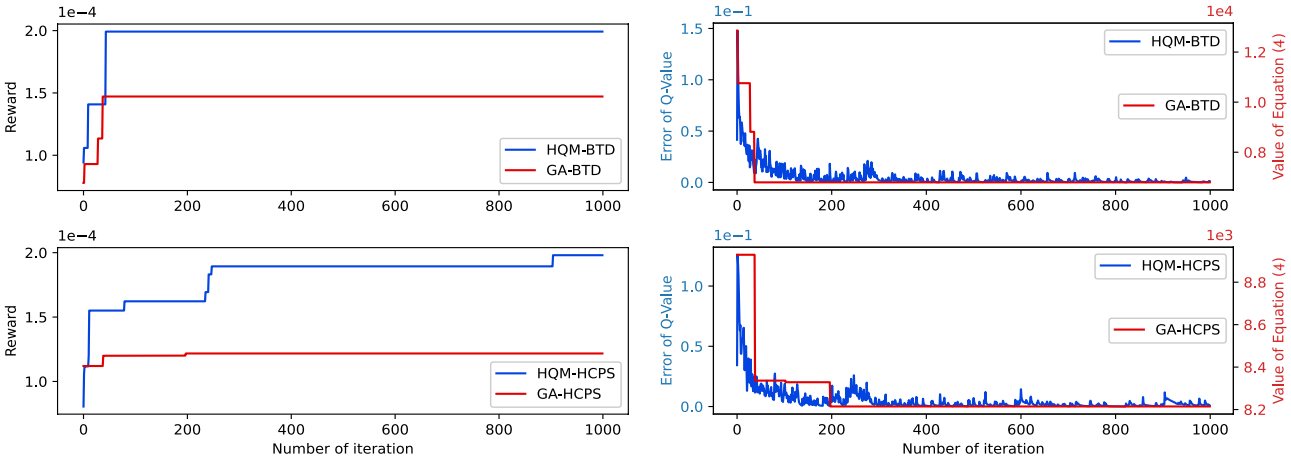
(c) Reward comparison of GA-BTD

(d) Reward comparison of GA-HCPS

Figure 6.11: Optimisation ability of HQM and GA

illustrates the convergence performance of the algorithms: the objective value of GA-BTD and GA-HCPS sharply converges at generation 60 (±20) and generation 200 (±10) respectively, whereas HQM improves solution smoothly and converges at generation 180 (±20), meaning that HQM better exploits the information feedback obtained from the previous action and generate more efficient solution search policy than GA.

On the other hand, considering that the reward quickly converges in the first 100 iterations in the HQM-BTD scenario, we utilise the Gurobi solver (without computing time limitations) to obtain the exact optimal solution such that estimates the gap between the local optimum of HQM and the truth optimum (Table 6.5). The result shows that the optimal gap between HQM-BTD and the optimum is 6.23% while the gap between HQM-HCPS and the optimum is 4.53%. Such gaps are acceptable since the computation time of HQM is around 8 times faster



(a) Convergence comparison for HQM and GA (b) Real-time reward of HQM and GA

Figure 6.12: Reward acquisition and convergence comparison

Table 6.5: Optimal gap between the local optimum of HQM and the truth optimum

Strategy	HQM		Gurobi		Optimal Gap (%)
	Reward ($\times 10^{-4}$)	Computation Time (sec)	Truth Optimum ($\times 10^{-4}$)	Computation Time (sec)	
BTD	1.976	1184.8	1.988	8301	6.23
HCPS	1.979	1265.4			4.53

Notes: Optimal Gap = (Gurobi-HQM) × 100%/Gurobi

than the exact approach.

Computing Time under Different Problem Instances

In Appendix K.4, we compare the computation time between HQM and GA. The results indicate that the average computation time of HQM is 9.64 times that of GA given the same problem instance. This is due to the fact that the heuristic algorithm does not involve data sampling and knowledge mining, thus reducing the computation time required, but sacrifices satisfactory solutions (the final reward of HQM is greater than GA's for all problem instances).

Since there are no MPLP benchmark problem instances introduced and solved by RL-based approaches, we cannot compare HQM's computation time with other RL-based approaches within the same dimension. However, in contrast to other attention-mechanism-based RL approaches for solving VRP with similar network sizes, the proposed HQM provides a shorter computation cycle. We cites the related research conducted by Kool et al. (2018) for reference which has been proven that the proposed REINFORCE approach achieves better optimisation

performance than other RL techniques: the study proposed a model based on attention layers with benefits over the Pointer Network. The REINFORCE is applied to train the model. The computation time under different problem instances with varied sizes is as follows: i) CVRP: 6 min (Node=20), 28 min (Node=50), 2 hr (Node =100); ii) SDVRP: 9 min (Node =20), 42 min (Node =50), 3 hr (Node =100).

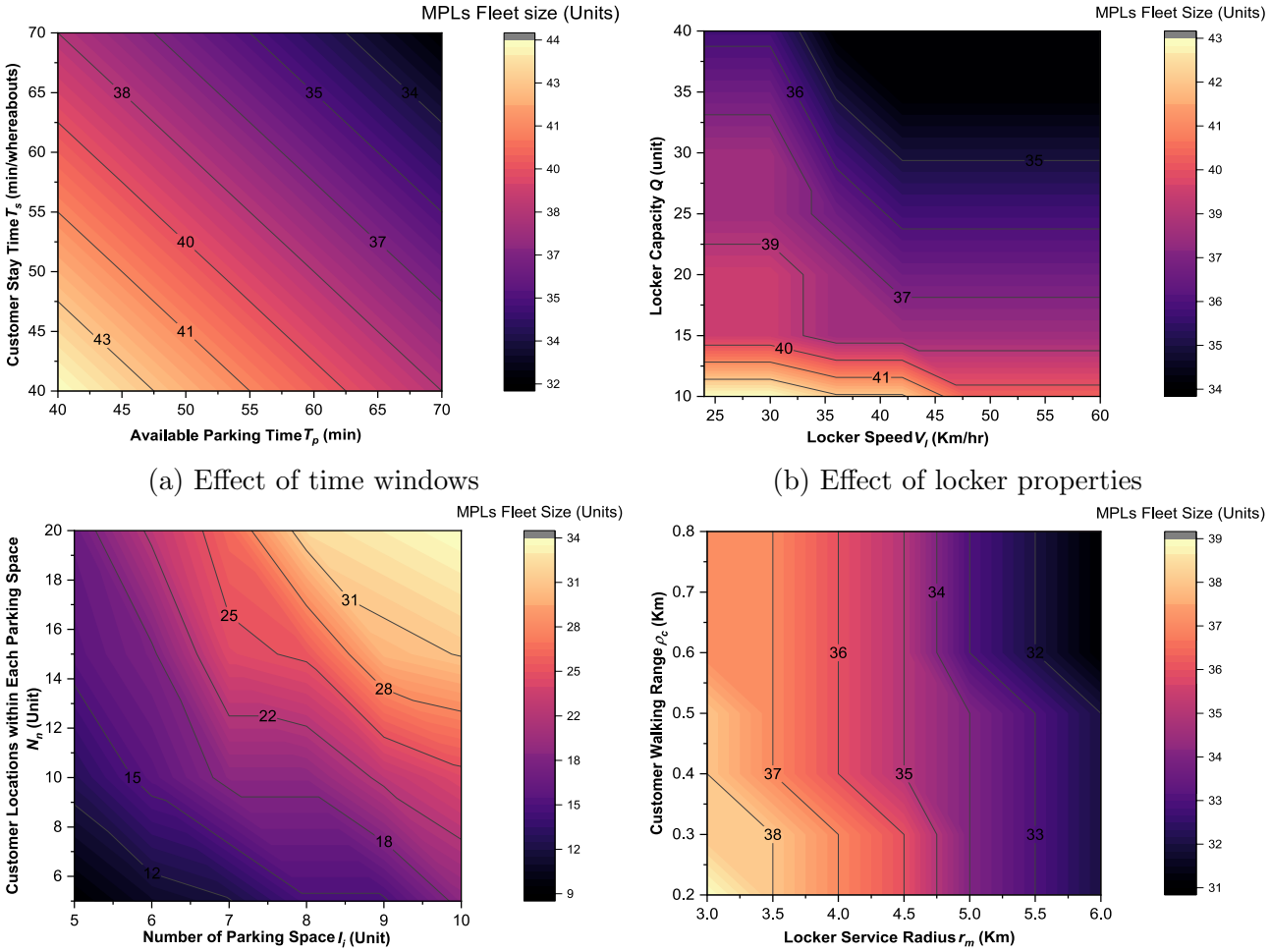
6.4.4 Effect of Critical Factors on Different Cost Units

We eventually investigates the effects of the critical factors on three cost units within the objective function (fleet size, travel distance, and service delay) to provide managerial and operational insight to logistics professionals. Factors are classified into four dimensions, depending on their attributes: i) time windows from customers and parking spaces, ii) locker properties, iii) network density, and iv) locker service radius and customer maximum walking range.

Effects on Fleet Size

Figure 6.13a demonstrates the effect of customer time windows and the available parking time of parking space on fleet size. The length of customer residence time T_s and available parking time T_p are evaluated within the range from 40 to 70 minutes. The result indicates that both customer time windows and the available parking time are negatively correlated to the fleet size. Extending customer residential time provides the potential to reduce time window conflicts, thus resulting in fewer MPLs being required for conflict resolution.

In terms of the effect of MPL properties, MPL capacity Q has a more significant influence on fleet size than speed V_I once $V_I \geq 40$ km/hr (Figure 6.13b). However, the fleet size will increase sharply if $V_I \leq 35$ km/hr and $Q \leq 15$ units because more dense contours are within this area. This can be explained by the fact that lower locker speed may not satisfy the time window requirements of customers and parking spaces, which require assigning extra MPLs to resolve such cases. In addition, a larger MPL fleet is required to meet time window constraints for lower MPL capacities. The result suggests that the MPL capacity Q is expected to be



(a) Effect of time windows

(b) Effect of locker properties

(c) Effect of network size

(d) Effect of service radius and walking range

Figure 6.13: Effects of critical factors against fleet size

greater than 15 units/MPL in the proposed network settings to reduce MPL fleet size.

Figure 6.13c shows that both the number of parking spaces and customer locations have conspicuous impacts on service delays, proven by the positive correlation of the contour plot. Compared with the other three factors, the difference in fleet size with varied network density is the most significant (the range of fleet size from 9 to 34 units).

Finally, customers walking range ρ_c have negligible impact on MPLs fleet size once the walking range $\rho_c \geq$0.4 km (Figure 6.13d), and that MPL service radius has a greater significance on fleet size. In summary, time window constraints, MPL capacity, and network density are the main factors affecting MPLs fleet size.

Effects on Travel Distance

The travel distance is more sensitive with available parking time T_p compared with customer stay duration T_s (Figure 6.14a). The travel distance sharply decreases when available parking time $T_p \geq 55$ min, along with fewer fleet required to meet customer demands, thus leading to shorter routes. On the other hand, the travel distance has increased gradually when customer stay time $T_s \leq 60$ min, because some MPLs are scheduled globally to meet the tight delivery time requirements even though the MPLs are far away from the customer to be fulfilled. This issue is even more significant when time windows constraints are tighter (e.g., 57 min$\leq T_s \leq$ 63 min, and $T_p \leq 45$ min).

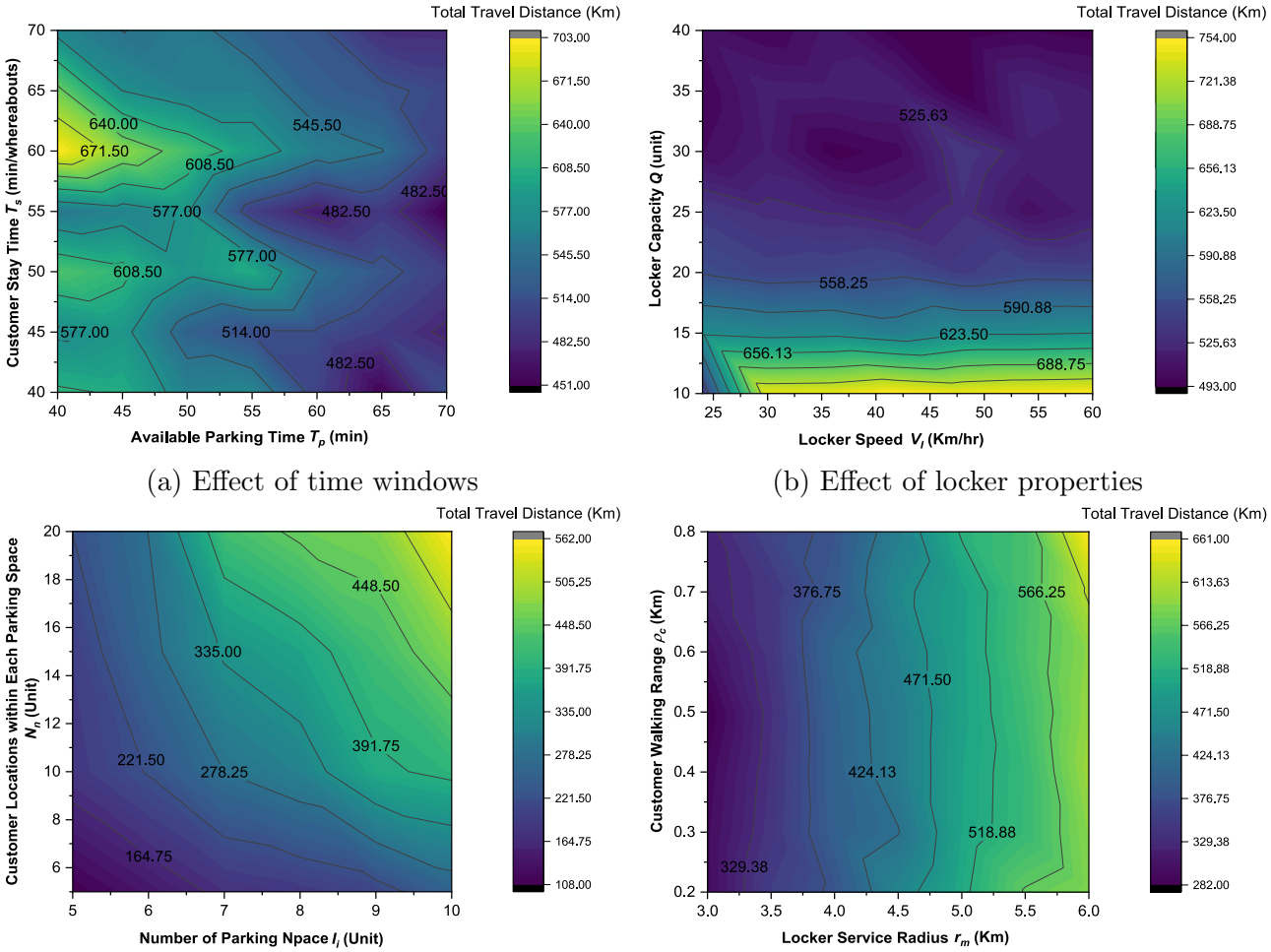
The result indicates that MPL capacity Q is negatively correlated with travel distance, whereas MPL speed V_I does not. The total route length only has a minor change when the locker capacity $Q > 20$ units/MPL (Figure 6.14b).

With the increased number of parking spaces and customer locations, the travel distance of MPLs has increased dramatically, especially the number of locations per parking space $N_n \geq$15 units/parking space and $I_i \geq$8 units (Figure 6.14c). Figure 6.14d demonstrates that the travel distance increased with their service range since more customers are covered by the MPL fleet that results in greater route length.

Effects on Service Delay

Figure 6.15a suggests that the service delay decreases with longer customer stay time T_s, meaning that if customers provide more generous time windows, the MPL can be scheduled more flexibly thus reducing service delays. Specifically, the service delay will be significantly reduced once $T_s \geq$ 50min or $T_p \geq$ 45min. However, the service delay will exceed more than 60 min if both customers and parking spaces time windows are tight (e.g., $T_s <$ 50 min and $T_p <$ 45min). Furthermore, available parking time T_p has a more significant impact on service delay compared with customer stay time T_s.

Figure 6.15b illustrates that the service delay will be improved as locker capacity Q decreases. The reduced MPL capacity means fewer customers can be fulfilled in a single delivery



(a) Effect of time windows

(b) Effect of locker properties

(c) Effect of network size

(d) Effect of service radius and walking range

Figure 6.14: Effects of critical factors against travel distance

round, which saves more travel distance and eliminates the potential delays. However, more MPLs are required to satisfy the remaining demands.

Conversely, the larger capacity Q may promote the algorithm to assign more customers to the same MPL, thus reducing MPL fleet size with the expense of higher service delays. The service delay will be improved as MPL speed increases. However, if the MPL speed is too fast, the travel time between two adjacent parking spaces may be too short to satisfy the available parking time restriction in large problem instances. In this case, the MPL will adopt route adjustment strategies to resolve time window conflicts, which may result in higher delays in the whole system. The capacity $Q \leq 25$ units may achieve satisfactory delay improvement since $Q = 25$ is a distinct boundary of the contour plot. This is of practical significance to logistics operators, as it guides operators to adopt MPLs with appropriate capacity to improve service

delay in an economic fashion.

Figure 6.15c shows that the network sizes have positive conspicuous effects on service delay, especially in the case that $I_i \geq 9$ and $N_n \geq 16$. This is because the increase in parking spaces and customer locations leads to more MPL deployments and longer travel distances, resulting in higher delays. Finally, Figure 6.15d presents that the greater MPL service radius and customer walking distance will cause higher delays since more customers will be served within the service region. Given this concern, $r_m \leq 5$ km and $\rho \leq 0.6$ km achieve lower delays.

In summary, the time window constraints and network density are the crucial factors affecting service delay. This requires collaboration between logistics professionals, customers, and urban planners to identify the most appropriate settings for these parameters that achieve timely service at a reasonable cost. Additionally, carriers can improve customer satisfaction by defining a reasonable service radius and implementing effective scheduling policies based on customer preferences.

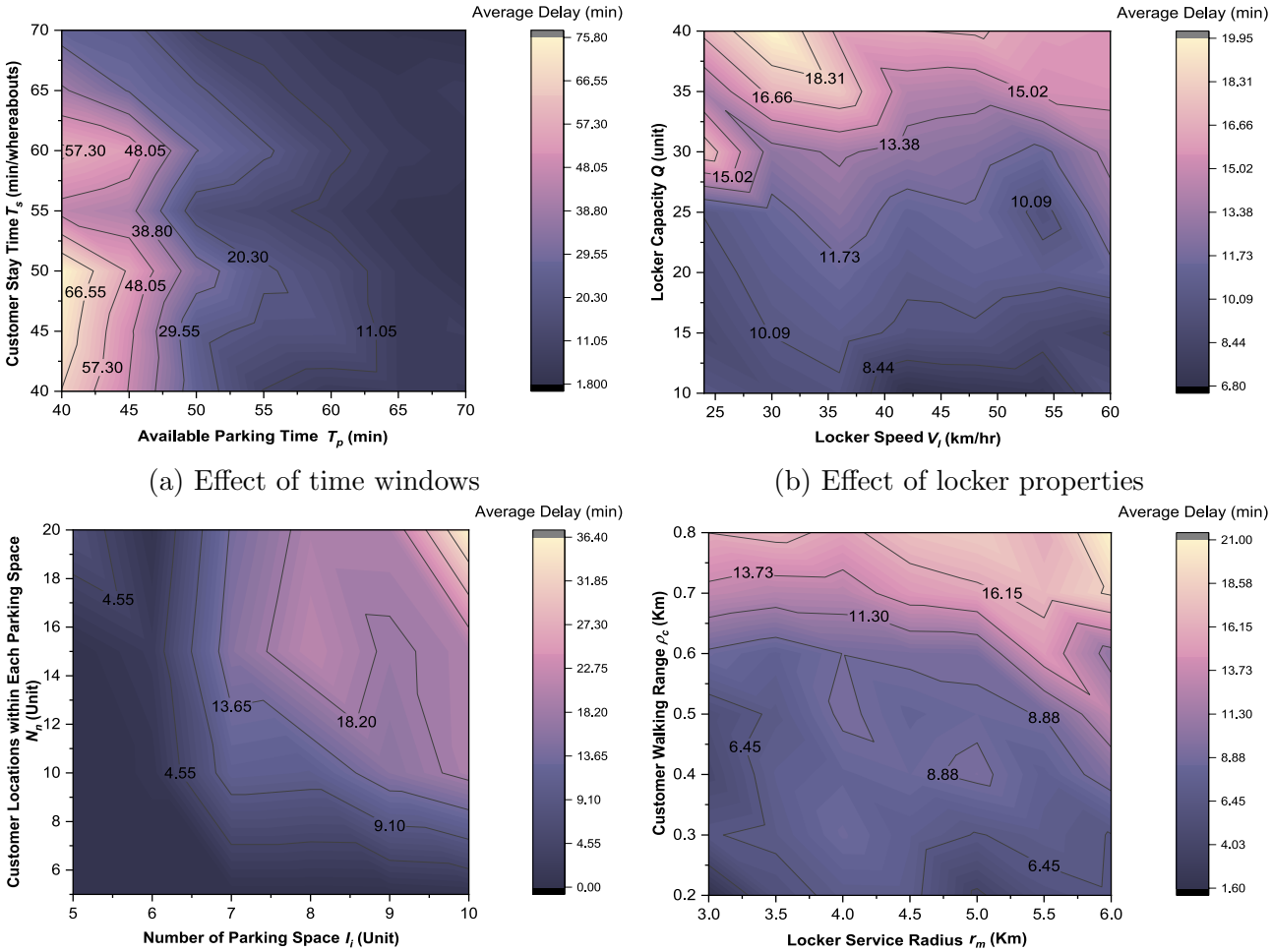
6.5 Managerial Implications

The managerial implications are discussed based on the prior parametric analysis to enable logistics practitioners better deploying MPL operations at scale accordingly in the future.

Customer Time Windows and Available Parking Time

Considering the tendency shown in Figures 6.13a, Figures 6.14a and Figures 6.15a, we conclude the ratio between available parking time T_p and customer stay time T_s (T_p/T_s) is regarded as a crucial indicator for logistics operators since the larger available parking time T_p enable MPLs process more stopovers, thus reach more customers. For $T_p/T_s > 1$, the performance of different cost units has significant improvement subject to $T_s, T_p \geq 60$ min.

Nevertheless, from the perspective of customer stay time T_s, it is difficult for customers to offer a more generous time window in practice given that their schedules are often fixed during the weekdays, making it difficult for them to adapt time windows for parcel collection. Logistics operators will therefore require a more generous available parking time to fulfil orders



(a) Effect of time windows

(b) Effect of locker properties

(c) Effect of network size

(d) Effect of service radius and walking range

Figure 6.15: Effects of critical factors against service delay

more effectively. One possible solution would be to park MPLs in residential areas (there is less traffic on residential streets, so the MPLs can be parked for longer periods without causing significant congestion). As an alternative, the deployment of MPLs near work sites during peak commute hours would allow more customers to be served within a relatively concentrated time interval. Finally, the application of a track & trace system (e.g., smartphone app) allows customers to monitor MPL's position in real-time, thereby reducing parcel overlap time for parcel handover.

Locker Capacity and Locker Velocity

Intuitively, greater MPL velocity contributes to more agility and can reach more customers by visiting more stops along the way. To achieve this, it would be possible to dispatch MPLs prior

to the peak commute period, which would enable them to travel at a faster speed (e.g., $V_i \geq$ 35 km/hr in this study) in the curb-side lane, thereby reducing service delays due to the traffic congestion. However, it is important to note that this policy does not apply to most hours of the day, as there is also an increased demand during off-peak hours, particularly on weekends. Given this concern, increased locker speed on main roads or pavements is not a significant approach to improve delivery efficiency since the average vehicle speed in the urban areas of well-developed cities is limited to 25 km/hr (Kusumaningtyas & Lodewijks, 2008). Therefore, recalling the contour plot in Figure 6.13b, 6.14b, and 6.15b, the results suggest that the medium capacity level of MPLs (e.g., $Q \in [20, 30]$ units/MPL) achieves greater cost savings.

Parking Spaces and Customer Stopovers

The location of the customer stopovers and demand density are important factors that influence the placement of parking spaces. Therefore, the successful adoption of the MPL concept depends on customer willingness to use MPL services and disclose their approximate location. Customers may have concerns about data privacy and the inconvenience of providing this information, which can be addressed by incentives offered by the logistics provider, such as reduced or waived service charges. Alternatively, logistics operators can offer different service options to customers who do not want to provide location information, such as stationary lockers or door-to-door delivery, while still providing MPL service to customers who are willing to use it.

From the perspective of demand density, it becomes easier to reach the maximum capacity of MPLs as demand density increases in an urban area. Hence, combining stationary lockers with MPL is a feasible solution to increase service capacity and convenience in a cost-efficient manner. It is much more likely that MPL will attract customer during the start-up phase due to their novelty in delivery mode. Therefore, having successfully diffused the MPL concept and got more customers on board, more stationary lockers (and thus cheaper) can be added to the existing distribution fleet without jeopardizing customer acceptance.

Service Radius and Customer Maximum Walking Range

Empirical studies have revealed that customers will normally accept a walking distance of no more than 0.7 km for parcel collection (De Oliveira Jr et al., 2016; Kedia et al., 2017). Recall that the result shown in Figure 6.13d, 6.14d, and 6.15d, the results suggest that the service radius of MPL r_m should not exceed 5 km if there is a requirement to guarantee the cost-savings in different cost units over a maximum walking distance of customer ρ_c not exceeding 0.7 km.

On the other hand, as we generate demand within a square area of 5km in length and width (see data generation code provided), it is concluded that by reducing the service radius of each MPL r_m by 1 km, the fleet size and total distance travelled will be reduced by at least 1 unit and 50 km, respectively. Therefore, the break-even point between the service radius of the MPL and the demand density is particularly pertinent for logistics operators. With this in mind, synergistic MPLs and other delivery methods (e.g., traditional van-based delivery, and stationary lockers) remain the high-potential efficient and cost-effective delivery scenarios in the area where demand density is high.

6.6 Analysis and Discussion

In this section, we discuss the proposed method to deal with a special form of LRP, namely MPLP. The main difference between LRP and MPLP is that the movement trajectory of MPL will be determined based on all potential stopovers of the customer set while the classic LRP considers the static characteristics of the demand pattern.

The mathematical model is initially solved using an exact approach solver, Gurobi, to determine the gap between HQM and optimal solutions. The result (Appendix I.1, Table 6.4) shows that HQM has superior ability in deriving high-quality solutions and computational efficiency compared to the Gurobi solver for large problem instances, while it suffers from insufficient training samples for small-scale problems.

However, the exact approach has a weakness in that it takes a considerable amount of time to compute the result. In real-world scenarios, the magnitude of the problem often

makes it impossible to solve the problem accurately within a reasonable timeframe (Yu et al., 2022b). Therefore, the results obtained from the exact approach are only used as a baseline for evaluating the solutions obtained by applying the proposed HQM.

Meta-heuristics, as a typical method used for solving the large problem instances of LRPs or VRPs, we, therefore, focus on the performance comparison between HQM and GA. The comparison covers two dimensions that provide a comprehensive overview of HQM's performance: i) solution quality and convergence/computation efficiency (e.g., Figure 6.10a to 6.12b, Table 6.5, Appendix K.3 and K.4); and ii) cost units and operation indications (e.g., Appendix K.1, Figure 6.9a to 6.9d, Appendix K.2). The results showed that HQM can provide high-quality solutions within a reasonable computational time when compared to GA solutions. However, there is no absolutely efficient algorithm, i.e., algorithms will suffer from the dilemma of compromising solution quality and compromising efficiency (Walpole et al., 2016). Although the proposed HQM performs well in LRPs or VRPs and can provide a novel conceptual RL-based solution framework, further experiments are required to validate its generalisation ability to more complex scenarios.

There are several potential limitations of the proposed HQM model that may arise in larger problem instances. One limitation is that HQM may have differences in optimisation capabilities for different cost units within the objective function, even though the results of different cost units are influenced by the weights assigned to them. Nevertheless, information exchange and exploitation among different variables x within a state s will also have an indirect effect on the optimisation of different cost units since these variables serve as logical abstractions of various cost units. To address this limitation, it may be helpful to design an effective communication mechanism, such as the use of intermediate variables or transformer layers, between different variables to make the information interpretable and usable among the variables.

The second limitation is the difficulty of interpreting the mathematical model into the RL framework, particularly when representing model constraints. One approach to address this is to decompose the original model such that the model constraints are processed and implemented outside the RL framework, while RL is only responsible for training the model.

The last limitation is the computation time. Although HQM's computation efficiency

is significantly superior to the exact approach, it still falls short of heuristic algorithms on large problem instances due to HQM's sampling procedure. The computational time required by HQM can be determined by analysing the procedure defined in the solution framework (Figure 6.3) in terms of complexity. The complexity of the proposed solution is in (1) the task generation procedure, (2) the reward calculation within the attention-based decoder, (3) the Q-value update procedure within the attention-based decoder, and (4) the global search and local search procedure. The other remaining modules can be executed in a constant time.

The procedure of task generation has a complexity of $O(|N \times K_n|)$ because all potential stopovers K_n for customer set N must be enumerated to identify a set of tasks that satisfy time-window constraints imposed by customers and parking spaces. The reward calculation within the attention-based decoder has the complexity of $O(|P|)$ since the decoder will calculate the corresponding reward for all state s ($s \in \{1, \cdots, P\}$) within *Agent*. In the case of Q-value updates, the computation complexity increases dramatically. Since there are two Q-value matrixes $Q(\theta^o_{x_1})$ and $Q(\theta^o_{x_2})$ (see Equation 6.33 and 6.34) in this study which corresponding to variable x_1 and x_2 of a state s respectively, therefore HQM have to update the Q-value matrix simultaneously. Specifically, the complexity of updating $Q(\theta^o_{x_1})$ is $O(M^2)$ where M represents the total number of MPL, while the complexity of updating $Q(\theta^o_{x_2})$ is $O(o^2)$ where o represents the total number of tasks. When it comes to the global search procedure, since all elements within the state s will be updated according to Q-value matrixes in each timestep, the complexity is $O(\Gamma)$,where Γ represents the total number of elements ($\Gamma = |x_1 + x_2|$) within the state s. Finally, the complexity of local search procedure is $O(\Upsilon)$, where Υ is the number of elements that need to be adjusted/shifted between two neighbourhood states. The complexities discussed above are the primary reason that the HQM's computation time increases with the sample size.

As alternatives to the encoder & attention-based decoder and heuristics introduced in this chapter, we provide the following methods to balance solution quality and computation efficiency. The first type of method focuses on generating superior initial solution and improving the efficiency of transcription of input samples. An example of the former method is introducing Negatively Corelated Search (NCS) during the training process to improve the

diversity of better initial solutions (Tang et al., 2016). The latter method typically involves using batch reinforcement learning, which differs from traditional online learning that the learning agent does not interact directly with the environment. It uses a baseline agent that collects data, which is then fed into an algorithm to train a new policy (Laroche et al., 2019). Finally, an emerging technique is '*learning to delegate*' approach, which improves the solution quality and computation efficiency by identifying appropriate subproblems and delegating their improvement to a black box sub-solver (Li et al., 2021b). '*learning to delegate*' framework consists of two components: 1) a sub-solver capable of solving a small problem instance, and 2) a learning model for identifying suitable subproblems within a larger problem to delegate them to an exact solver.

The proposed mathematical model can be extended to solve other variant VRP-related problems. To adapt the mathematical model to a specific scenario, existing constraints must be modified (e.g., considering the parcel pick-up, demand heterogeneity). For the HQM, the encoder and attention-based decoder can be adopted to deal with other VRP variants with some modifications in the proposed solution interpretation, given the excellent generalisation ability of HQM's framework.

6.7 Concluding Remarks

This chapter presents a novel approach for MPLP that accounts for changing customer demand patterns and varied scales, as well as stochastic events. A hybrid Q-Learning network algorithm was developed that combines global and local search mechanisms is presented, which resolve stochastic events that cause delays by employing two route adjusting strategies.

We first evaluated the performance of HQM by using an exact approach solver (Gurobi) as a benchmark to provide a first-hand insight of the solution quality and computation efficiency. The result showed that the proposed HQM achieves higher solution quality with significantly shorter computation time than the Gurobi solver in large problem instances. To highlight the advantage of HQM compared with meta-heuristic algorithms that commonly used for solving LRPs and VRPs, we exploited a customised GA as baseline cases and conducting pairwise

comparison under different network densities (15 to 200 nodes) with different route adjustment strategies applied. The proposed HQM provided solutions in lower computing times than other comparable learning algorithms and obtains a final reward function twice as greater as the benchmark. Additionally, the result revealed that HQM achieves significant improvement in service delay and parcel first-time delivery successful rate compared to GA. Finally, we investigated the influence of critical factors on different cost units within the objective function, providing logistics operators with parameterised and managerial insights for enhancing their delivery service.

In addition to the computational advantages of the HQM model, results suggested that the use of HQM in combination with HCPS leads to reduced service delays and improved MPL configuration compared to using BTD. Although HQM showed a good optimisation ability in the proposed MPLP scenario, further experiments and adaptions are required before migrating it to other LRP and VRP-related problems. To tackle the limitations of sampling efficiency and solution quality that may falls short compared with heuristics, we provided several potential solutions such as applying delegate agent, and batch reinforcement learning approaches.

We finally provided managerial insight based on the parametric analysis for different critical factors against cost units. The most critical factors that affect different cost units are the time window constraints imposed by parking spaces and customers, locker capacity, and network density. Logistics operators face the challenge of balancing deployment costs with improved delivery convenience and punctuality when broadly implementing MPLs in the market. Further, the result highlight that collaborating MPL with other mainstream delivery methods (e.g., traditional van-based delivery and stationary locker) offer the potential for broader application of MPL during start-up phase, especially in cities with high demand density and traffic congestion. As an alternative, dispatching MPLs before peak commuting hours based on customers' travel behaviour allows more demand to be fulfilled at more concentrated sites (e.g., company, school) and shorter time intervals of the day. From the perspective of urban planning and sustainability, deploying MPLs in less congested areas (e.g., residential areas) allows lockers to stay longer without disturbing traffic flow while the characteristic of their electric-powered mobility reduces gas emissions. However, more research is needed to examine

the practicality and validity of integrating customers' daily locations with MPLs for last-mile deliveries, including the collection and protection of customer data.

Future extension of the study could focus on more complex dynamic cases of MPLP that require more flexible route adjustment approaches to respond to changes in demand and delays. It would also be useful to study multi-echelon MPLP, which involves the use of transhipment centres or feeder vehicles to reduce the distance between lockers and the central depot. Another extension of the problem is to consider cooperative deliveries in conjunction with MPLs, such as the use of couriers or delivery robots. An alternative approach is to segment the market so that the method of parcel delivery within a specific area is determined based on predetermined criteria such as demand density, road conditions, and traffic flow. From the perspective of algorithmic innovation, further research on advanced RL-based algorithms for solving MPLP and VRP is valuable. This could include implementing parallel, interoperable interaction between RL methods and the simulator to better match real-world operations. From the perspective of stochastic event modelling, future studies can consider sampling from random variables to build a stochastic event pool, which is different from the triggering mechanism based on mathematical expectations adopted in our study, to reduce bias. Finally, it would be useful to apply the model to real operational case studies in different scenarios, such as urban and rural regions, to validate the model.

Chapter 7

Route Planning Model for Robot-Assisted Deliveries in Last-Mile

The work of this chapter has been published at **26th International Conference on Intelligent Transportation Systems (ITSC 2023)** [1]. Conventional human-vehicle cooperation distribution has always been a challenge for the logistics industry as it involves multiple stops, high costs (up to 53% of the total delivery expense), gas emissions, and unpredictable factors (e.g., traffic congestion) (Van den Berg, 2016). With customers increasingly demanding faster and more convenient delivery options, the complexities of navigation in congested urban areas under unpredictable traffic conditions have caused bottlenecks for efficient distribution.

Beyond the unattended delivery capabilities offered by MPLs introduced in Chapter 6, which enhance the accessibility and convenience of parcel collection, it is pertinent to consider the potential of ADRs in collaborating with delivery staff to streamline delivery operations. This is significant as it allows delivery staff to redirect their efforts towards more productive tasks, such as parcel sorting, while ADRs assume the responsibility of transporting parcels to customers. ADRs may achieve a greater saving in delivery completion time compared to traditional human-dominated delivery by eliminating time-consuming tasks, such as finding

[1]Liu, Y., Zhang, K., Escribano-Macias, J. & Angeloudis, P., (2023). Quantifying the Capability of Autonomous Robot-Assisted Distribution in Last-Mile: Insight from Rural to Urban Configurations. in '*26th IEEE International Conference on Intelligent Transportation Systems ITSC 2023*'. Available at SSRN 4549465

parking spaces and multi-travelling between customers and vehicles (Reed et al., 2022b; Chiara et al., 2020; Abhishek et al., 2021).

Therefore, this chapter i) introduces a route planning model for the ADR-Assisted Deliveries (AAD), ii) evaluate its delivery capability in different geographies (from rural to urban regions), and iii) determine how different operating parameters affect various delivery activities, such that provides managerial insight for future large-scale deployment. To achieve this, this chapter establishes a mixed integer programming model to formulate the ADR-Assisted Delivery Problem (AADP) in the Illinois region via graph network.

In AADP, both ADRs and delivery staff assistance to serve customers. The ADR has a maximum loading capacity and will return to the vehicle to reload parcels once the current delivery is completed. The vehicle can traverse different customers with or without ADR onboard the vehicle. Hence, the vehicle can rendezvous with ADR at the last customer within each service set that ADR performs. The delivery staff only needs to sort packages for ADR and take over delivery tasks once ADR is power exhausted. Therefore, AADP can be regarded as a variant derived from Travelling Salesman Problem with Flying Sidekick (also known as drone-assisted delivery) since both of them involve battery consumption, releasing and rendezvous ADR/drone at customer locations, and planning the accompanied vehicle route within the model (Escribano Macias et al., 2020). The modelling and solving techniques of these two kinds of problems are of reference significance.

Considering the limited capacity of ADR and the size of all possible combinations that determine service sets for ADR delivery, AADP is challenging to solve. To reduce the computing complexity, this chapter introduce a service set & variable reduction technique that decreases the multitude of variables and inefficient sub-route determined in the original problem.

Instead of utilising the dataset from the London area introduced in Chapter 4, this chapter utilised the datasets from various geographic regions in Illinois, US, to fairly validate and compare the delivery performance with the prior study Reed et al. (2022a) in which the same data are used. The travelling time matrix for both vehicle and ADR within each region are generated by the Bureau of Public Road (BPR) impedance function. The vehicle's travel time is configured to replicate real-world traffic conditions such as one-way streets. This captures the

possibility that the ADR may travel faster than the vehicle between customers. This chapter finally compares the performance of ADR-assisted delivery with other delivery modes (e.g., capacitated delivery scenario considering vehicle parking) discussed by Reed et al. (2022b).

Therefore, this study contributes to the knowledge and industries in the following aspects:

- A mixed integer programming model is introduced to formulate AADP while considering battery constraint, from which we evaluate the potential of broad deployment of ADR-assisted delivery in future from activity-operation perspectives.
- We incorporate the BPR impedance functions to capture the travelling time of vehicles on the main roads and ADR on sidewalks, respectively, which considers different characteristics of the working environment.
- The numerical analysis from comprehensive activity-operation aspects provides managerial implications for academia and professions to facilitate efficient configuration of ADR-assisted distribution.

The remainder of the chapter is structured as follows: Section 7.1 introduces the definition of AADP. Section 7.2 formulates the corresponding mathematical model for AADP. Section 7.3 introduce a solution technique that reduces predetermined service set and variables to improve the computation efficiency. Section 7.4 evaluates the performance of the solution technique and analyses the effect of key operating parameters on delivery performance. Section 7.5 provides managerial implications based on experiment results. Section 7.6 concludes our findings and future work.

7.1 ADR-Assisted Delivery Problem

The ADR-Assisted Delivery Problem (AADP) involves fulfilling a customer set C that includes n customers using a van-based ADR assisted by delivery staff. The delivery staff and an ADR are carried by the vehicle when it departs from and returns to the depot. The ADR service is regarded as the preferred service option while delivery staff can sort parcels for ADR and

take over delivery tasks once the maximum battery power b of robot is exhausted. Apart from being on standby for delivery tasks, the delivery staff drives the vehicle travel between different service sets of customers and loading parcels for ADR prior delivery.

Let $\{0\}$ denotes a depot and define $\bar{C} = C \cup \{0\}$. The ADR can deliver maximum q parcels at one time. Let s denote a subset that the ADR serves for customers in a single delivery round ($1 \leq |s| \leq q$) and S represents the set including all possible subsets s. For each $c \in C$, let $O_c = \{s \in S | c \in s\}$ be the service sets that include customer c.

To fulfil a subset of customers s, ADR is released by the vehicle at customer location i (regarded as starting point), loads the parcels, and rendezvous with the vehicle at customer location j that may differ from location i once set s are fulfilled. The duration required for sorting parcels, denoted by f, remains constant irrespective of the number of customers within the subset s. The delivery staff will drive the vehicle to the rendezvous point j at the same time when ADR is out for delivery which achieves time-saving.

Let $V(i, j)$ denote the time of the vehicle travels from customer i to customer j, and $R(i, j)$ represent the time of ADR travels from customer i to customer j. We use tuple (i, s, j) to denote ADR serves subset s starting from customer i and rendezvous with the vehicle at customer j. For each subset s, there exist a customer permutation $(p_1^{isj}, p_2^{isj}, \ldots, p_{|s|}^{isj})$ that achieves the shortest path for ADR travels from customer i to customer j (i, j can or can not be in subset s). Therefore, the time consumption of ADR following the shortest route to serve customer subset s that starts from customer i and ends at customer j can be denoted by $W_{ij}^s = R(i, p_1^{isj}) + R(p_1^{isj}, p_2^{isj}) + \cdots + R(p_{|s|-1}^{isj}, p_{|s|}^{isj}) + R(p_{|s|}^{isj}, j)$.

In one-way streets, the ADR may reach the rendezvous point before the vehicle. Therefore, ADR will wait for the vehicle until it arrives. The waiting time is $max\{V(i, j) - W_{ij}^s, 0\}$.

Figure 7.1 illustrates the operation of conventional human-vehicle delivery scenario (Figure 7.1a) and the proposed AADP scenario (Figure 7.1b). Compared with the conventional human-vehicle delivery scenario, the AADP could significantly eliminate the round trip between the vehicle and customers.

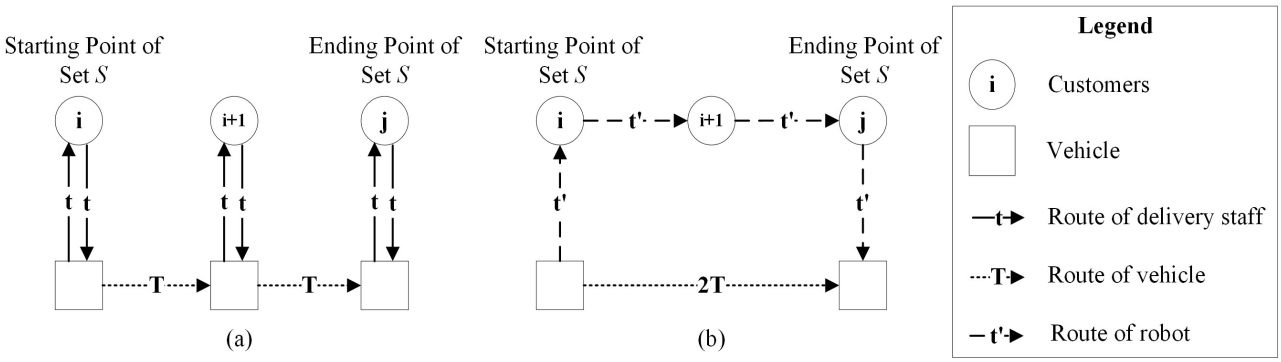


Figure 7.1: Instance of conventional human-vehicle delivery and AADP

7.2 Mathematical Model for AADP

A mixed integer programming model is developed according to the problem defined in Section 7.1. We make the following assumptions when developing model:

Assumption 1: Each customer requests one parcel for delivery in a single delivery round given that most existing ADR products in the market have only one compartment (Jennings & Figliozzi, 2019). As parcel dimensions and weight are standardised by logistics providers, the size of the parcels must fit the compartment of ADR.

Assumption 2: Customers' properties are assumed to satisfy the physical environment that ADRs can access customers through ramps and curbs (Reed et al., 2022b).

Assumption 3: The total demand of set s should be less or equal to the maximum capacity q of ADR, such that all customer demand can be fulfilled.

Assumption 4: The vehicle can release and collect ADR at any customer location (Reed et al., 2022a).

Assumption 5: $V(i,j)$ and $R(i,j)$ satisfy the triangle inequality where the linear distance between i and j is less than any other non-linear distance.

7.2.1 Route Planning Model for AADP

Table 7.1 describes the indices, parameters, sets, and decision variables of the model.

$$minf(x_{ij}, y_{ij}^{s}, z_{ij}, u_i) = F_1 + F_2 + F_3 \tag{7.1}$$

$$F_1 = \sum_{i \in \bar{C}} \sum_{j \in \bar{C} \setminus \{i\}} V(i,j) x_{ij} \tag{7.2}$$

$$F_2 = \sum_{i \in C} \sum_{s \in S} \sum_{j \in C} (W_{ij}^s + max\{V(i,j) - W_{ij}^s, 0\}) y_{ij}^s \tag{7.3}$$

$$F_3 = \sum_{i \in C} \sum_{s \in S} \sum_{j \in C} f \cdot y_{ij}^s \tag{7.4}$$

Subject to:

$$\sum_{i \in C} x_{i0} = 1 \tag{7.5}$$

$$\sum_{i \in C} x_{0i} = 1 \tag{7.6}$$

$$\sum_{i \in C} \sum_{s \in O_c} \sum_{j \in C} y_{ij}^s = 1 \qquad \forall c \in C \tag{7.7}$$

$$\sum_{s \in S} \sum_{j \in C} y_{ij}^s = \sum_{l \in \bar{C}, l \neq i} x_{li} \qquad \forall i \in C \tag{7.8}$$

$$\sum_{i \in C} \sum_{s \in S} y_{ij}^s = \sum_{l \in \bar{C}, l \neq j} x_{jl} \qquad \forall j \in C \tag{7.9}$$

$$\sum_{j \in C} z_{0j} = n \tag{7.10}$$

$$z_{0i} \leq n \cdot x_{0i} \qquad \forall i \in C \tag{7.11}$$

$$z_{ij} \leq n(x_{ij} + \sum_{s \in S} y_{ij}^s) \qquad \forall i,j \in C, i \neq j \tag{7.12}$$

$$\sum_{j \in \bar{C} \setminus \{i\}} z_{ji} - \sum_{j \in C \setminus \{i\}} z_{ij} = \sum_{j \in \bar{C}, i \neq j} x_{ij} + \sum_{s \in S} \sum_{j \in C} (|s| - 1) y_{ij}^s \qquad \forall i \in C \tag{7.13}$$

$$u_j + \rho \cdot W_{ij}^s \leq u_i + M(1 - y_{ij}^s) \qquad \forall i,j \in C, s \in S, i \neq j \tag{7.14}$$

$$W_{ij}^s \cdot y_{ij}^s \cdot \rho \leq u_i \qquad \forall i,j \in \bar{C}, i \neq j \tag{7.15}$$

$$u_0 = b \tag{7.16}$$

$$x_{ij} \in \{0,1\} \qquad \forall i,j \in \bar{C} \tag{7.17}$$

$$y_{ij}^{s} \in \{0,1\} \qquad \forall i,j \in C, s \in S \tag{7.18}$$

$$z_{ij} \in \mathbb{Z}_{+} \qquad \forall i \in \bar{C}, j \in C, i \neq j \tag{7.19}$$

$$u_i \in \mathbb{N}_{+} \qquad \forall i \in \bar{C} \tag{7.20}$$

The objective function (7.1) minimises the total delivery time, which includes the travelling time of the vehicle (Eq.7.2), the travelling and waiting time of ADR (Eq.7.3), and parcel handling time (Eq.7.4). Constraint (7.5) and (7.6) ensures that the ADR is onboard the vehicle when departs and returns to the depot, respectively. Constraint (7.7) guarantees that all customers are fulfilled. Constraint (7.8) ensures that the vehicle will visit customer i to release ADR if ADR serves subset s. Constraint (7.9) guarantees the vehicle will visit the rendezvous point j to pick up ADR travelling to the next customer when subset s has been fulfilled (i.e., $y_{ij}^{s} = 1$).

Constraint (7.10) represents that the total number of parcels to be delivered is n. Constraint (7.11) ensures that the vehicle can visit customers from the depot only when the ADR is onboard (i.e., $x_{0i} = 1$). Constraint (7.12) indicates that the vehicle can travel between customers either with or without ADR onboard. Constraint (7.13) denotes the quantity change of parcels when serving each customer.

Constraints (7.14) and (7.15) ensure that the ADR must have enough power when travelling from customer i to customer j. Constraint (7.16) denotes that the ADR has maximum battery power when departs from the depot. Constraints (7.17) to (7.20) define the domain of the variables.

7.3 Solution Methods

This section demonstrates the method that determines the travelling time for both the vehicle and ADR within the road network in rural and urban regions. We then introduce a technique to reduce the service set of ADR and simplify the solution space of decision variables.

We determine the travelling time between two customers in different geographic regions

Table 7.1: Notation table for route planning model of AADP

Indices		**Sets**	
i	Customer i	C	Set of customers [-]
j	Customer j	$\bar{C}$	Set of customers and depot [-]
s	Service set s	S	Set of all customer service set served by ADR [-]
Parameters		**Decision Variables**	
n	Number of customers [customer]	x_{ij}	Boolean: $x_{ij} = 1$ if the vehicle travels from customer i to customer j with the ADR and delivery staff onboard for $i, j \in \bar{C}$
q	Maximum capacity of ADR [parcels]	y_{ij}^s	Boolean: $y_{ij}^s = 1$ if ADR is released at customer i to service customer set s and rendezvous vehicle at customer j for $i, j \in C, s \in S$
b	Maximum battery power of ADR [kw]	z_{ij}	Integer: the flow of parcels onboard the vehicle when travels from customer i to customer j for $i \in \bar{C}, j \in C, i \neq j$
ρ	Power consumption of ADR [kw/hr]	u_i	Float: the remaining power of ADR in customer i for $i \in C$
f	Time consumption of sorting/loading parcels [minutes]		
$V(i,j)$	Time consumption of the vehicle travels from customer i to customer j [minutes]		
$R(i,j)$	Time consumption of ADR travels from customer i to customer j [minutes]		
W_{ij}^s	Time consumption of ADR following the shortest route to serve customer set s when starts from customer i and ends at customer j [minutes]		
M	A sufficiently large number [-]		

for vehicle $V(i,j)$ and ADR $R(i,j)$, respectively, which is captured by the BPR impedance function considering parking space search.

Since the size of service set S expands exponentially with ADR capacity q and the number of customers n (i.e., $|S| = \sum_{i=1}^{q} \binom{n}{i}$), the size of variable y_{ij}^s is $n^2 \cdot |S|$ which causes computational difficulty. We introduce a solution technique to reduce the size of service set S and variable y_{ij}^s.

7.3.1 Road Impedance

The total travelling time of vehicle $V(i,j)$ from customer i to j consists of the congested travelling time $t_c(i,j)$ of vehicles and the time $t_p(i,j)$ consumed for searching parking spaces. The former is captured by the BPR road impedance function (Eq. 7.21 and Eq. 7.22), which determines congested travel time $t_c(i,j)$ based on traffic flow $g(i,j)$ and link capacity $h(i,j)$. It is obtained from regression analysis based on substantial traffic survey data (US Bureau Public Roads, 1964) and has been widely used in road traffic research. For the value of α and β, we take values in urban streets ($\alpha = 1.28, \beta = 1$) suggested by the literature (Petrik et al., 2014).

In Equation (7.23), we assume that the road impedance for ADR depends on the route length $L(i,j)$ and the operation speed v. v is set as 5.5 km/hr which simulates the general walking speed of young pedestrians (Rastogi et al., 2011).

In experiments, we utilise two parameters, free flow speed t_f and flow capacity ratio g/h, to differentiate the road conditions between rural and urban areas. The park-searching time t_p is set to a constant for a region and varies from 1 to 9 minutes continuously from rural to urban areas.

$$V(i,j) = t_c + t_p \tag{7.21}$$

$$V(i,j) = t_f(i,j)\left[1 + \alpha\left(\frac{g(i,j)}{h(i,j)}\right)^{\beta}\right] + t_p(i,j) \tag{7.22}$$

$$R(i,j) = \frac{L(i,j)}{v} \tag{7.23}$$

7.3.2 Decision Rule for ADR Travelling

ADR may either travel to the next customer autonomously (i.e., $y^s_{ij} = 1$) or on-board the vehicle travelling to the next customer (i.e., $x_{ij} = 1$). The decision principle is to compare the travelling time of the vehicle $V(i,j)$ with ADR travelling time $R(i,j)$ when travelling from customer i to j. If $R(i,j) < V(i,j)$ is satisfied, then ADR has the potential to be travelled from i to j autonomously.

7.3.3 Simplification for ADR Service Set

Based on the above decision rule, we determine the travelling of ADR by compromising the relationship between $V(i,j)$, f, and $R(i,j)$, which contributes to the simplification of the service set S.

Let $G_v(\bar{C}, E_v)$ (E_v denotes edges) represent the driving graph of the vehicle while $G_r(\bar{C}, E_r)$ denotes the travelling graph of ADR. For $i, j \in C$, if there exists an edge $\tilde{E}_r \in E_r$ that satisfies $V(i,j) + f \leq R(i,j)$, it implies that ADR onboard the vehicle takes less time compared with ADR travels from customer i to j independently. Therefore, customer i and j will not be included within the same service set of ADR, and edge $\tilde{E}_r$ can be removed from $G_r(\bar{C}, E_r)$ such that the service set of ADR is reduced.

We further consider the relationship between the rendezvous point and the service set to simplify the ADR service set. Consider set $s \in S$ in which the vehicle travelling is faster than ADR travelling to/from all customers in s and potential loading/rendezvous locations C. Under this circumstance, the vehicle will arrive at customers and rendezvous point prior to ADR. Lemma 7.1 indicates that the ADR does not need to wait for the vehicle at the end of service set s. For all relevant proofs see Appendix L.

Lemma 7.1 *If there exists a service set $s \in S$ satisfies $V(i,j) \leq R(i,j)$ and $V(j,i) \leq R(j,i)$ for all $i \in s$ and $j \in C$, then the vehicle will reach the rendezvous point prior to ADR (i.e., $max\{V(i,j) - W^s_{ij}, 0\} = 0$.*

If ADR does not need to wait for the vehicle, then the modified travelling graph of ADR provides a criterion to judge whether the current set s can be eliminated. Proposition 7.1 indicates that set s will not contribute to the optimal solution if the route determined by set s is excluded from the travelling graph $G_r(\bar{C}, E_r)$ of ADR, set s can be removed from the alternative.

Proposition 7.1 *For set $s \in S$ that satisfies $V(i,j) \leq R(i,j)$ and $V(j,i) \leq R(j,i)$ for $i \in s$ and $j \in C$, the set s will not contribute to the optimal solution (i.e., $y^s_{ij} = 0$) if there is not exist a customer permutation $(p_1^{isj}, p_2^{isj}, \ldots, p_{|s|}^{isj})$ of set s in the travelling graph $G_r(\bar{C}, E_r)$ of ADR.*

Corollary 7.1 outlines the situation that a customer i in graph $G_r(\bar{C}, E_r)$ will be fulfilled individually if i is far away from other customers. Therefore, ADR will be onboard the vehicle that travels from/to customer i.

Corollary 7.1 *For customer $i \in C$ that satisfies $V(i,j) + f \leq R(i,j)$ and $V(j,i) + f \leq R(j,i)$ for $j \in C \backslash \{i\}$, the customer i will be fulfilled individually where ADR loads parcel and rendezvous the vehicle at customer i (i.e., $y_{ij}^s = 1, s = \{i\}$). Additionally, the following equations are satisfied:*

$$\sum_{j \in C} x_{ji} = 1 \tag{7.24}$$

$$\sum_{j \in C} x_{ij} = 1 \tag{7.25}$$

7.3.4 Simplification for Decision Variable

We reduce the size of variable y_{ij}^s by utilising the relationship between starting & rendezvous point and set s. Considering that if both task starting and rendezvous point of ADR are restricted in the customer within a set s while $V(i,j) \leq R(i,j)$, $V(j,i) \leq R(j,i)$ for all $i \in s$ hold, there exists an optimal solution satisfies $y_{ij}^s = 1$ and $j \in s$.

We consider the situation that the task starting point excluded from set s, and ADR has to reach the first customer of set s. If the vehicle with ADR onboard spends less time than ADR travel to the first customer independently, the ADR will on-board the vehicle to the first customer of set s. However, this may cause extra waiting time when ADR serves the final customer of set s and rendezvous the vehicle before travelling to the next set s'. Hence, we compromise the travelling time saved by the vehicle on-board ADR against the waiting time incurred by the ADR travelling alone. Proposition 7.2 reduces the size of y_{ij}^s if the time saved by the vehicle on-board ADR is greater than ADR waiting time.

Proposition 7.2 *For $(i,s,j) \in C \times S \times C$ and $i \notin s$, there exists a customer permutation $(p_1^{isj}, p_2^{isj}, \ldots, p_{|s|}^{isj})$ for set s that achieves the optimal route when ADR loads parcels at customer*

i and reach j. Then $y_{ij}^s = 0$ *and can be eliminated from the solution set if the following equation holds:*

$$R(i, p_1^{isj}) - V(i, p_1^{isj}) \geq max\{V(p_1^{isj}, j) - W^s_{p_1^{isj} j}, 0\} \tag{7.26}$$

Analogously, Proposition 7.3 reduces the size of y_{ij}^s if the time saved by the vehicle onboard ADR when travelling from the last customer of set s to the rendezvous point is greater than the waiting time of ADR incurred in the previous set s'. Then $y_{ij}^s = 0$ and can be eliminated from the solution set if the following equation holds:

Proposition 7.3 *For* $(i, s, j) \in C \times S \times C$ *and* $j \notin s$*, there exists a customer permutation* $(p_1^{isj}, p_2^{isj}, \ldots, p_{|s|}^{isj})$ *for set s that achieves the optimal route when ADR loads parcels at customer i and reach j. Then* $y_{ij}^s = 0$ *and can be eliminated from the solution set if the following equation holds:*

$$R(p_{|s|}^{isj}, j) - V(p_{|s|}^{isj}, j) \geq max\{V(i, p_{|s|}^{isj}) - W^s_{ip_{|s|}^{isj}}, 0\} \tag{7.27}$$

7.3.5 Constraint Tightening

This section introduce tight constraints to capture the relationship between variables x_{ij} and y_{ij}^s. The key principle is that ADR can only visit set s either by itself or onboard the vehicle. Lemma 7.1 implies that each customer will be fulfilled only once. Should the vehicle arrives at customer i, the customer i becomes a loading or rendezvous point. Lemma 7.2 asserts that if the loading and rendezvous points coincide, this particular customer will be the sole recipient of the service.

Lemma 7.2 *For* $i \in C$*,* $|s| = 1$ *and* $s = \{i\}$ *holds if* $y_{ii}^s = 1$ *for some* $s \in S$*.*

Analogously, Lemma 7.3 postulates that if the ADR departs from the vehicle to serve customer i and on-board the vehicle at customer i once the task completed, then the service set is comprised exclusively of customer i.

Lemma 7.3 *For* $i \in C$*, customer i will be served alone (i.e.,* $y_{ii}^s = 1$ *and* $s = \{i\}$*) if* $x_{ci} = 1$ *and* $x_{ik} = 1$ *hold for* $c, k \in C \backslash \{i\}$*,* $c \neq k$*.*

Proposition 7.4 implies that once the ADR is onboard the vehicle travelling from customer i to j ($i \neq j$), customer i and j are within different service sets of ADR.

Proposition 7.4 *For $i, j \in C$ and $i \neq j$, the following equation holds:*

$$\sum_{(a,s,b)\in O} y_{ab}^{s} + x_{ij} \leq 1 \tag{7.28}$$

where $O = \{(a, s, b) \mid a, b \in C, s \in O_i \cap O_j\}$

Proposition 7.5 proposes a mathematical conversion for Proposition 7.4.

Proposition 7.5 *For $i, j \in C$ and $i \neq j$, the following equation holds:*

$$\sum_{s\in S}(y_{ji}^{s} + y_{ij}^{s}) + x_{ij} + x_{ji} \leq 1 \tag{7.29}$$

7.4 Results and Discussion

To validate the model, we use customer datasets from 9 counties in Illinois, which mimics the varying demand distribution based on their population densities (from rural to urban counties) and was introduced by Reed et al. (2022a). For each county, we select 5 random square service regions with 50 customers of each to execute the model. The following rule is applied to code different counties which start from rural to urban: (1) Cumberland, (2) Johnson, (3) Jefferson, (4) Fulton, (5) Adams, (6) La Salle, (7) Champaign, (8) Winnebago, (9) Cook. The model is executed in an Apple M1 processor (3.2GHz) with Python 3.9 using Gurobi 3.10 solver with 16GB RAM. The parameter setting is shown in Table 7.2. Since the service radius of a single ADR is restricted by the battery capacity, in which the robot mainly covers the demands of a small residential area, we set the network as a small size ($n = 50$) for a single delivery round. Considering the compartment size, an ADR can only deliver a few parcels per loading ($q <= 3$). We further conduct a sensitivity analysis for different values of the robot's capacity ($q = 1, 2,$ and 3). All datasets are available at: *https://github.com/Yubin-Liu/Data-Set-for-AAD.git*.

Table 7.2: Parameter Setting for Model

Notation	**Value/Metric**	**Notation**	**Value/Metric**
n	50 [customers]	q	3 [parcels]
b	120 [kw]	ρ	90 [kw/hr]
f	2.8 [minutes]	M	9×10^9 [-]
α	1.28 [-]	β	1 [-]
t_f (Rural)	45 [km/hr]	t_f (Urban)	35 [km/hr]
g/h (Rural)	0.4 [-]	g/h (Urban)	0.8 [-]
t_p	[1,9] [minutes]	v	5.5 [km/hr]

7.4.1 Performance Evaluation for Solution Techniques

Appendix M presents the computational performance of the solution techniques. There are 20,875 service sets of ADR in total when $n = 50$ and $q = 3$, while the maximum number of variable y_{ij}^s is $n^2 \times S = 52,187,500$. Results indicate that the number of edges E_r^* and service set S^* after implementing Proposition 7.1 are positively correlated with the increasing demand density from rural to urban regions. The size of the optimised network degree and service set S^* is reduced by an average of 90.1% and 96.89%, respectively, compared with the maximum.

When applying Corollary 7.1, the number of customers $|y_{ii}^s|$ served alone shows decreasing trend as the areas transition from rural to urban. This is due to the fact that the distance between two customers becomes shorter in urban areas, allowing more customers to be clustered within the same service set. It is noticed that the remaining number of optimised variables $y_{ij}^{s'}$ and $y_{ij}^{s''}$ after executing Proposition 7.2 and 7.3 negatively correlated with the increasing demand density from rural to urban regions, meaning that the autonomous navigation of ADR between two consecutive customers in different service sets shows less possibility in urban areas due to more unpredictable traffic condition compared with rural areas. This significantly increases the percentage of variable y_{ij}^s (increase from 2.43% to 18.48%) that can be precisely determined before the model-solving stage.

7.4.2 Delivery Performance of AAD Scenario

Figure 7.2 represents the time consumption of different delivery activities of AAD scenarios under varying regions. The parcel sorting operation accounts for the highest proportion of the

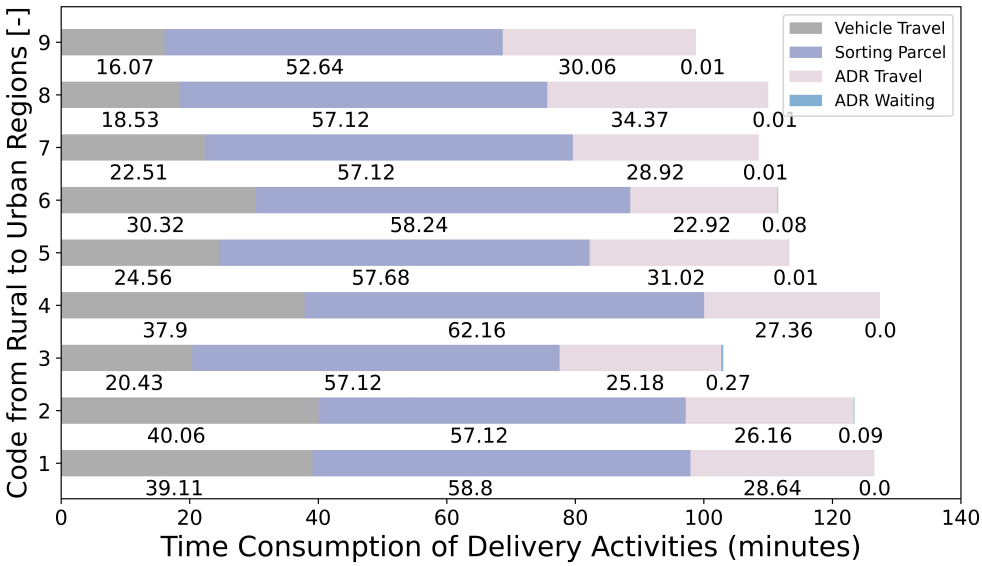
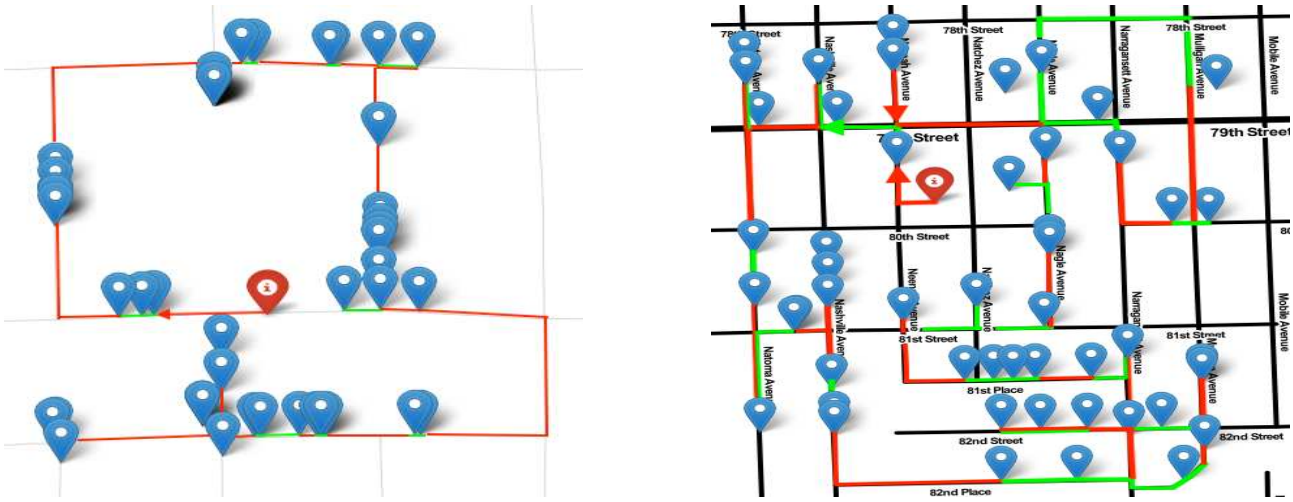


Figure 7.2: Time consumption of different delivery activities in different regions



(a) Rural region (Cumberland) (b) Urban region (Cook)

Figure 7.3: Delivery route of vehicle (——) and ADR (——) in different regions

total delivery time with an average percentage of 50.66%, followed by vehicle travelling (24.40%) and ADR travelling (24.90%). Specifically, the AAD scenario achieves greater delivery time-saving in urban regions compared with rural areas. This can be explained by the fact that the ADR delivery option is used more extensively in urban regions (see Figure 7.3a and 7.3b) that reduces long travel time caused by unpredictable traffic conditions as well as parking space searching.

We further benchmarks the total delivery consumption of the proposed AAD scenario with four classic delivery modes introduced by Reed et al. (2022b): i) Capacitated Autonomous

Vehicle-assisted Delivery Problem (CAVADP) that an AV serving as a mobile depot while the delivery staff departing from AV to distribute parcels before rendezvousing AV. ii) Capacitated Delivery Problem with Parking (CDPP) in which the delivery staff must park the vehicle prior to performing the delivery. The vehicle can park at any customer location with the time p to find parking space is 0. iii) CDPP with Parking-Dependent (CDPP-P) where the time p to find parking space depends on different regions. In the proposed dataset, the value of p from rural to urban regions ranges from 1 to 9 minutes, respectively. iv) FAD scenario where AV traverse different customers and serves them within a restricted time $T = 5.8$ ($T = f + 3$) minutes. The customer must visit the corresponding parking space of AV to collect their parcels because there is no delivery staff.

The result indicates that the proposed AAD scenario outperforms other classic delivery modes in delivery completion time (Figure 7.4). Specifically, the saving percentage of the AAD scenario against the other four delivery modes is 4.8% (CAVADP), 17.46% (CDPP), 47.42% (CDPP-P), and 65.48% (FAD), respectively. The overall savings achieved are attributed to the reduction in vehicle travelling and parcel sorting activities. It is identified that parcel sorting as the bottleneck for achieving efficient deliveries in the CDPP scenario, accounting for 56.91% of the total completion time. In the CDPP-P scenario, delivery staff walking and parking space finding accounted for significant proportions of the total completion time (35.67% and 24.71%, respectively). Finally, considering that customers must visit AV to collect parcels in the FAD scenario, the considerable time consumption of this mode comes from the handover process and unpredictable walking time of customers caused by varying location distributions. The total time of parcel handover and customer walking occupies 88.07% of the total completion.

7.4.3 Effect of Key Operating Parameters on Delivery Performance

We finally investigate the effect of three key operating parameters on delivery performance: i) the delivery capacity of ADR, ii) the parcel handling time, and iii) the maximum battery capacity of ADR.

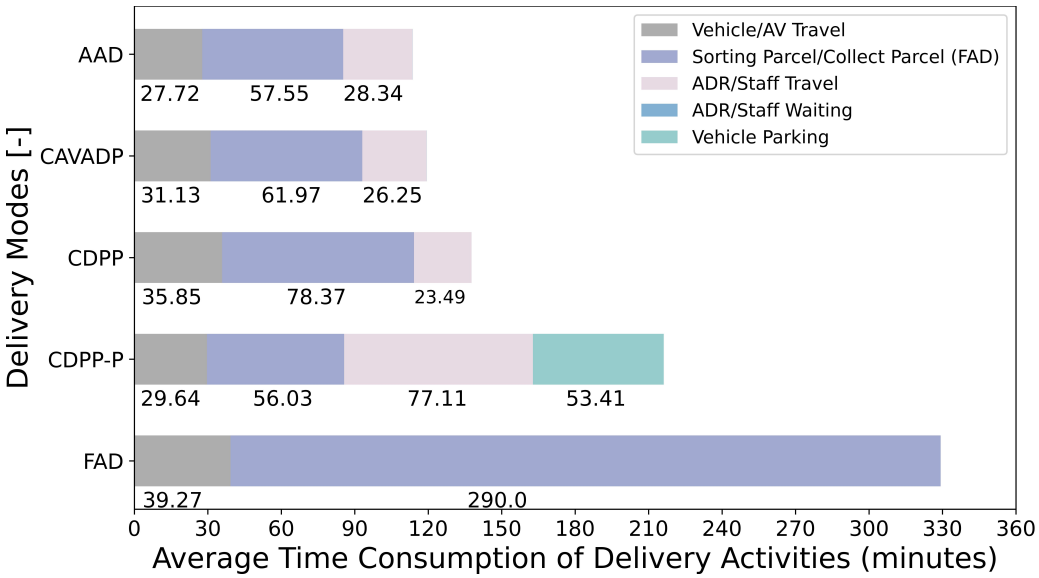
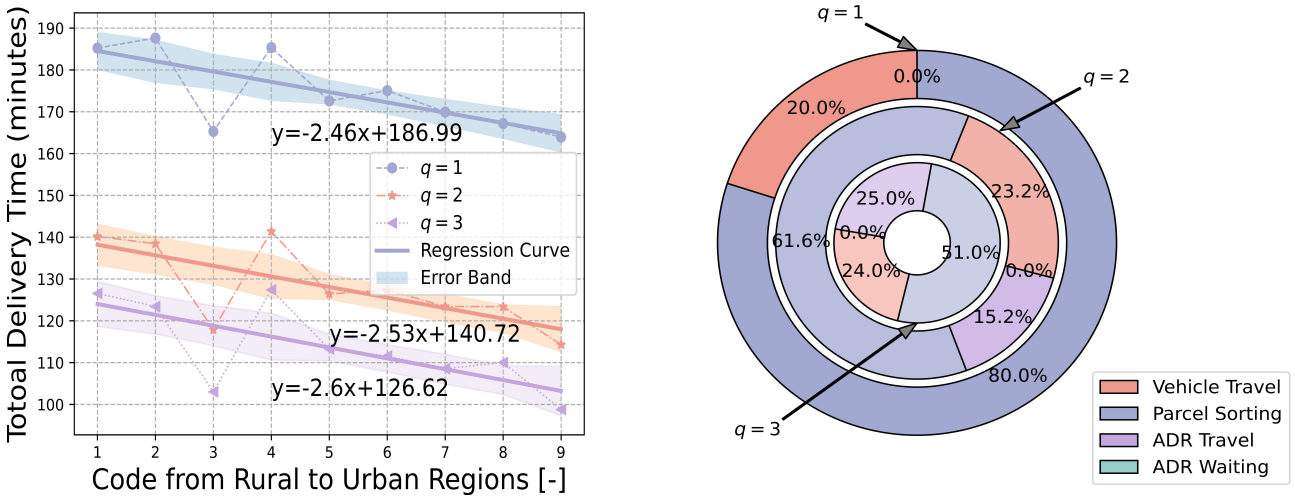


Figure 7.4: Delivery time comparison between different distribution modes



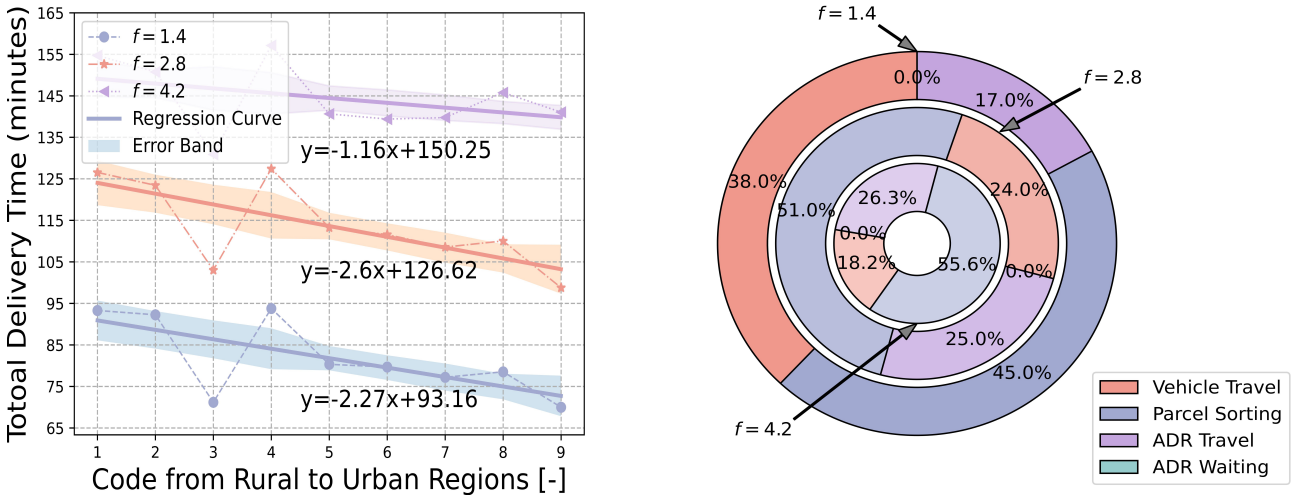
(a) Effect on total completion time

(b) Effect on different activities

Figure 7.5: Effect of ADR capacity on completion time and activities

Effect of Robot Capacity

The results show that the total delivery time increases as the capacity of ADR decrease (Figure 7.5a). This extra delivery time mainly comes from more frequent parcel sorting operations before ADR travels to final customers, which accounts for up to 80% of the total delivery time when ADR capacity $q = 1$ parcels (Figure 7.5b). In contrast, the parcel handling time for $q = 3$ parcels accounts for 51% of the total.



(a) Effect on total completion time

(b) Effect on different activities

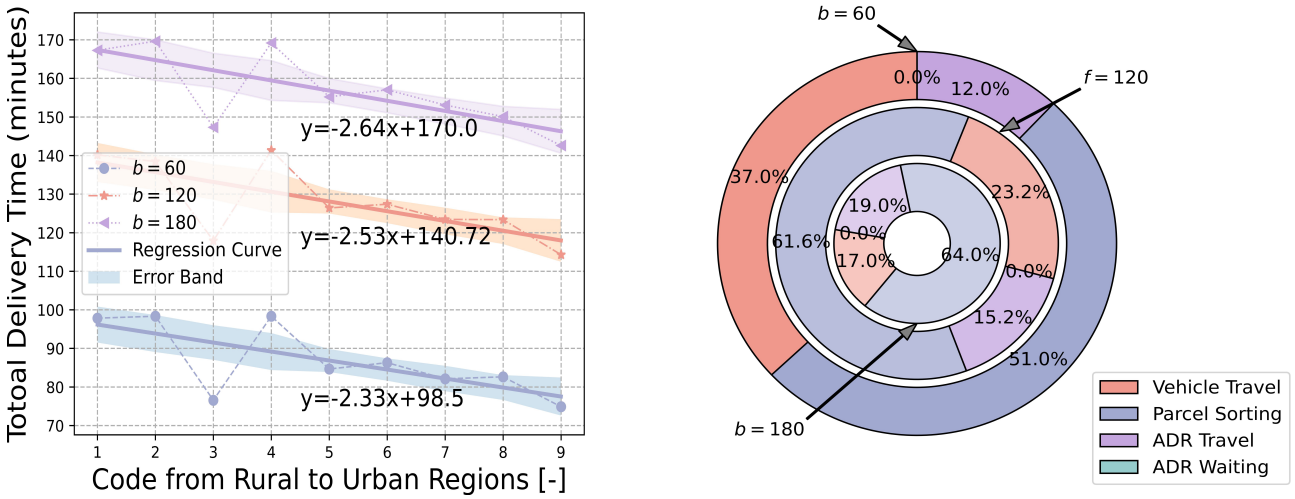
Figure 7.6: Effect of parcel handling on completion time and activities

Effect of Parcel Handling Time

Figure 7.6a indicates that the time reduction in parcel handling will significantly contribute to the total delivery time savings. As can be seen from Figure 7.6b, the increase in parcel handing promotes autonomous navigation for ADR while the proportion of vehicle travelling decreases. Therefore, we infer that deploying AAD within the scenario where the parcel handling operation is a bottleneck may obtain a higher benefit.

Effect of Battery Capacity

For Figure 7.7a and 7.7b, we notice that the total delivery time for battery capacity $b = 60$ is less than that of $b = 80$ and $b = 120$. The greater battery capacity provides the potential for increasing service radius and working duration. However, the rationale behind the decision to utilise an ADR when it has adequate power is primarily determined by the trade-off between the travel time of the vehicle and the travel time of the ADR for two consecutive demands. This decision is not solely based on the remaining power of the ADR.



(a) Effect on total completion time

(b) Effect on different activities

Figure 7.7: Effect of battery capacity on completion time and activities

7.5 Managerial Implications

The capacity and endurance of robots constitute critical indicators of concern for logistics operators in last-mile delivery applications. Given the existing market constraints regarding the size of robots and their manoeuvrability on sidewalks, a robot with a capacity $q = 3$ yields significant time savings in delivery. While larger battery capacities can provide robots with an extended service range, this comes at the cost of higher procurement expenses and increased physical dimensions, potentially impacting their sidewalk operability. One potential solution entails equipping delivery vehicles with spare batteries for convenient replacement when a robot's power is depleted. This approach proves to be more pragmatic and minimizes time consumption compared to recharging batteries during the delivery process. Another factor that logistics operators need to consider is the security of parcels being delivered by robots. For instance, malicious actors may intentionally tamper with the cargo bay to steal packages, necessitating the implementation of enhanced protective measures by logistics operators. These measures could encompass reinforcing the robot's structure, real-time tracking, and instantaneous alerting capabilities.

Parcel handling and loading packages onto robots constitute significant time-consuming activities during the delivery process. One potential solution is to systematically arrange packages in the vehicle according to their delivery sequence before the vehicle departs from

the warehouse. This way, upon reaching the designated parking space, delivery staff can more easily locate the packages for loading onto the robot.

7.6 Concluding Remarks

This chapter evaluates the potentials of the AAD scenario in last-mile distributions specifically considering parking space finding, parcel handling, and battery constraint of a robot. A service set and variable reduction technique are applied to solve the model with lower computational complexity. We evaluate the AAD scenario by using the dataset from nine counties (from rural to urban) in Illinois. For each county, we select five regions for numerical testing.

The results demonstrate that the proposed solution technique could reduce 96.89% infeasible service set on average. Additionally, the proposed AAD scenario outperforms the existing human-dominated modes in delivery efficiency, especially in urban areas. This implies that the significant time consumption in finding available parking spaces is still a bottleneck for efficient deliveries.

This study provides a new dimension to assess the potential of deploying AAD in the last-mile distribution in the future. The solution approach is promising to be applied in solving other AV-assisted delivery problems. To have an unbiased discussion, one limitation of the study is that we only investigates the performance of AAD from the perspective of delivery efficiency improvement.

Future extension of the study could focus on identifying the most suitable autonomous delivery options (e.g., robot or mobile locker delivery etc.) for customers based on the characteristics of the demand channel and customers' properties. It would also be beneficial to investigate the application of using population synthesis techniques and digital twins to present and capture customers' characteristics that better support the decision regarding the delivery service selection. Furthermore, empirical studies that consider future large-scale AAD deployments from the perspectives of cost and benefit are still in short supply, which will provide guidance and implications for the industry. Finally, the AAD operation scenario can incorporate uncertain or stochastic events, such as changes in customer demand patterns, available parking

time variance, and service delays caused by different operating conditions of the ADR, so as to be more in line with the actual operation mode.

Chapter 8

Conclusion

In this chapter, the research findings are assessed in relation to the research objectives outlined in Chapter 1. This involves an examination of the research contributions and limitations, as well as its implications for various stakeholders. Additionally, recommendations for prospective research endeavours are put forth.

8.1 Research Objectives Evaluation

The evaluation encompasses four primary stakeholders in the last-mile distribution sector: governmental bodies, customers, industry collaborators, and academics. The research contributions and discoveries are aligned with the stakeholders and research objectives outlined in Chapter 1, as conveniently outlined below for the reader's reference:

- Obj 1. Evaluate the suitability and effectiveness of implementing AV-based delivery solutions situated from urban to sub-urban of the London city, taking into account the cooperation among the proposed delivery modes (e.g., MPLs, ADRs, and conventional human-van delivery) and the types of customer destinations (e.g, apartments, houses). An improved continuous approximation approach is developed to estimate delivery route length and time consumption of delivery fleets, facilitating the quantification of associated operational costs, delivery efficiency, and gas emissions. Additionally, this study investigates the impact of each delivery option on pre-designed performance

indicators when it occupies different market shares. The fulfilment of this objective resolves Gap 1.

- Obj 2. Develop a non-cooperative gaming framework to delineate the competitive dynamics between different autonomous delivery options (MPLs vs. ADRs). The operation process of delivery options is formulated as a queue system such that investigates the accumulative effect of customer decisions on system efficiency, wherein different settings of key operating parameters are applied for sensitivity analysis. The fulfilment of this objective resolves Gap 2.

- Obj 3. Apply a customised logit equilibrium model to capture different levels of customers' decision rationality for service selection. To improve the overall utility of the delivery system, this study further develops an improved algorithm that incorporates global service information into customers' decision patterns. Finally, this study designs a Machine-learning-based (ML) regression prediction model to determine the optimal service pricing of the delivery options such that obtains the expected market share. The fulfilment of this objective resolves Gap 3.

- Obj 4. Develop a location-route planning model to achieve the optimal configuration of MPL deliveries, wherein customer stopovers and the corresponding time windows within the planning horizontal are considered in the model. This study further introduces dynamic prediction-based route adjustment strategies to reduce the service delay caused by time window conflicts. The detailed fulfilment will be introduced in Chapter 6. As an alternative to the MPL delivery option, we further establish a route planning model for ADR service to optimise the total delivery time consumption of ADR-assisted deliveries, wherein battery consumption constraints and auxiliary activities during the operation are incorporated into the model. The proposed route planning model will be validated through the real-world geographic dataset from rural to urban regions in Illinois, US, as introduced in Chapter 7. The accomplishment of this objective resolves Gap 4.

- Obj 5. Design an optimisation framework that utilises RL-based algorithms, game-theoretical algorithms, heuristics algorithms, and mathematical programming techniques

to solve different sub-problems in accomplishing Obj 1 to 4 above. These solution techniques are specifically tailored to enhance the generalisation ability of the algorithms and solution quality in large problem instances. The proposed problem instances are applied in scale and scope, which allows us to evaluate the performance of the solution techniques effectively. The fulfilment of this objective resolves Gap 5.

Objective 1 is addressed in Chapter 4, which focuses on the macroscopic-level assessment of deploying MPLs and ADRs in the last-mile distribution market. This chapter proposes an evaluation technique hinged on discerning the property of destinations and probability distributions of service selection, to segment the delivery requests performed by different delivery options. The capability of different autonomous delivery options are evaluated through activity-based time consumption, costs, and gas emissions, benchmarked with the conventional human-van delivery. The accomplishment of Chapter 4 reaches Contribution 1 outlined in Section 2.2.

Objective 2 is addressed in Chapter 5, which focuses on investigating customer decision behaviour of the autonomous delivery service selection. The competitive dynamics within the game between MPLs and ADRs are modelled as a queue system, accounting for congestion probabilities in each service queue. To address Objective 3, we devised Wardrop equilibrium and Logit equilibrium models to formulate fully rational and bounded rational customer decision-making behaviours, respectively. The study has proven the existence and uniqueness of these two equilibrium models. Subsequently, four performance metrics have been introduced, and an analysis of the influence of various critical operational parameters on system performance has been conducted, considering varying levels of customer rationality. The fulfilment of the tasks mentioned above reaches Contribution 2 outlined in Section 2.2.

Chapter 5 further incorporates an enhanced algorithm that leverage global information to augment system performance. The efficacy of the algorithm has been assessed across different levels of customer rationality. Lastly, a regression prediction model is developed using KSVM. KSVM is designed to take expected equilibrium and auxiliary operating parameters as inputs, and it produces the optimal pricing strategy for the delivery service. These facilitated the accomplishment of Contribution 3.

Operational-level optimisation for MPL and ADR deliveries are presented in Chapters 6 and 7, respectively. Chapter 6 formulates the location selection model of parking spaces and the route planning model of MPLs as mixed integer and complete integer programming, respectively. The former is tackled using the K-means clustering technique, while the latter is resolved through the proposed HQM. HQM is designed to handle the computational intricacies arising from large problem instances, improving solutions caused by the local optima. Furthermore, it integrates both global and local search mechanisms to grapple with the trade-off between exploration and exploitation encountered by classic RL methods. HQM's performance is assessed across varying problem sizes (up to 200 nodes) and benchmarked against exact approaches and GA in terms of solution quality, computational efficiency, and diverse cost units. Additionally, two route adjustment strategies are devised and assessed across all problem instances. Managerial insights are derived from a parametric analysis of various influencing factors, furnishing guidance for logistics operators in the realm of optimising last-mile distribution operations.

Chapter 7 formulates ADR-assisted delivery as a mixed integer programming model, aiming to minimise the overall delivery time, encompassing vehicle travel, robot navigation, and parcel handling. To refine the solution space, a technique involving service set reduction and decision variable simplification is introduced. The model and its solution approach are evaluated using actual geographic data from all 9 counties of Illinois, USA. Each county consists of 5 regions to simulate robot-assisted deliveries, resulting in a total of 45 scenarios for assessment. The geographic dataset, initially introduced by Reed et al. (2022a), is employed, and Bureau of Public Road (BPR) impedance functions are applied to calibrate the travel time matrix between various nodes for both vehicle and robot, to obtain a closer alignment with real-world practices. Lastly, the study quantifies the impact of different delivery activities on the total time consumption. The accomplishment of Chapter 6 and 7 resolves Objective 4 and reaches Contribution 4.

Achieving Objectives 1 to 4 necessitates addressing different sub-problems encompass macroscopic evaluations and microscopic optimisations, which promotes the accomplishment of Objective 5 and reaches Contribution 5. The proposed RL-based algorithms, game theoretical

methodologies, ML techniques, and traditional optimisation theories contributes to an enhanced knowledge of resolving variant problems that may arise in future research endeavours.

8.2 Research Implications

This thesis provides an effective tool for different stakeholders within the logistics sector to design and evaluate the autonomous last-mile delivery operations, including i) feasibility and capability assessment (Chapter 4) that analyses whether the integration of AV-related delivery solution is beneficial to the current practice, ii) behaviour analysis that understands the decision patterns of customers regarding autonomous delivery service selections and how to improve customers' decision rationality through global information such that increase systems' overall utility (Chapter 5), and iii) location selection and route planning that achieves optimal configuration of the delivery fleets and delivery time savings, considering the changing demand patterns of customers, time window constraints, and battery consumption (Chapter 6 and 7).

8.2.1 Implications for Commerce

The commercial implications encompass three key stakeholders: governmental entities, logistics operators, and customers.

For governmental bodies, the primary considerations regarding autonomous delivery solutions pertain to sustainability and compatibility with the urban geographical environment, as discussed in Chapter 4. These factors significantly impact the potential large-scale deployment of autonomous delivery initiatives. Therefore, utilising case studies, this research presents a realistic customer decision behaviour model towards autonomous delivery systems (as introduced in Chapter 5), providing practical insights into the operational dynamics of the system as outlined in Chapters 4, 6, and 7. This portrayal serves as a foundation for governmental facilitation of business model implementation according to scientific guidelines. Simultaneously, the study's findings contribute to improving streetscape design to support the integration and functionality of autonomous delivery systems in urban environments.

Regarding compatibility with the urban geographical environment, assessments should

encompass factors such as the current pedestrian density, road conditions, traffic flow density, and architectural features within the region to determine their support for autonomous delivery initiatives. Additionally, the introduction of autonomous delivery initiatives raises the issue of potential unfair competition with established delivery modes. This could manifest in a sudden increase in demand for a specific service option, intensifying reliance on curbside spaces. These considerations underscore the need for the formulation of a comprehensive evaluation framework to serve as a benchmark for future assessments.

The significance of data security is on par with the sustainability and compatibility objectives of urban areas, given the incorporation of customers data in autonomous delivery initiatives. It is crucial to engage in discussions regarding the protection and regulation of privacy data. Governmental entities need to develop relevant regulations to ensure safety across the diverse facets mentioned. This involves setting priorities for interactions among pedestrians, vehicles, and AVs, along with incorporating vehicle-to-vehicle information connectivity to minimise conflicts.

For logistics operators, the envisaged autonomous delivery systems possess the potential for deployment in the explored environments, given their self-navigating and decision-making capabilities. This research furnishes a practical model, as elucidated in Chapter 5, for anticipating customer selection behaviour towards these innovative delivery services. This model enables operators to strategically configure resources efficiently by establishing crucial operational parameters (e.g., capacity, service coverage, and pricing) of autonomous delivery services. Additionally, the route planning model proposed in Chapter 6 and 7 exhibits longevity, serving as significant calibrated benchmarks for operating parameters to ensure the profitability and efficiency of autonomous delivery services.

For instance, the findings from Chapter 5 to 7 underscore that critical operational parameters, including capacity, service coverage, and pricing, have a substantial impact on the performance of autonomous delivery systems. Concerning capacity, logistics operators must decide on the sizes of AVs to deploy and the fleet's scale required to efficiently meet demand in a specific area, while considering cost-effectiveness. Regarding service coverage, logistics operators need to determine the maximum distance a fully loaded AV can travel in

a single trip. An excessively extensive service radius may result in delayed parcel deliveries. Service pricing primarily revolves around achieving the desired market share and maintaining supply-demand equilibrium through judicious pricing, thereby enhancing the overall utility of the system. This depends on logistics operators' comprehension of customer decision-making behaviour across different contexts, as introduced in Chapter 5. Additionally, scientifically scheduling the AV fleet significantly impacts delivery timeliness, as discussed in Chapter 6 and 7. This entails determining optimal dispatch times and routes based on an understanding of customer travel patterns. For instance, denser fleet deployment in work areas during commute hours and increased fleet presence in residential areas during off-hours.

For customers, the empirical findings outlined in Chapter 5 indicate a prevailing sensitivity to pricing, where their criteria for service selection primarily revolve around achieving a balance between cost-effectiveness and convenience. Furthermore, the manner in which customers opt for autonomous delivery services based on time windows and daily schedules has implications for the overall efficiency of the system, as exemplified by the outcomes in Chapter 6. This entails customers specifying more accurate or generous time windows for parcel receipt to minimise instances of missed deliveries. Customers with heightened uncertainty regarding time windows may find MPL services more accommodating. Conversely, those who prioritize door-to-door service may discover that robot-assisted delivery mitigates the inconvenience associated with customers having to travel to retrieve their parcels, as demonstrated in the results outlined in Chapter 7.

8.2.2 Implications for Academia

For academia, adhering to a research ideology that progresses from macro-level feasibility evaluation to meso-level planning and decision-making, and ultimately culminates in micro-level operational optimisation holds significant value. This framework ensures that future autonomous delivery solutions align with urban livability and sustainability objectives. Moreover, addressing the future large-scale deployment of autonomous delivery in research requires consideration of the diverse demands of various stakeholders. This results in a notable increase in model complexity, demanding more advanced and diverse solution techniques,

including the integration of RL, heuristics, and classical optimisation methods. Such approaches are essential for efficiently solving large problem instances and models, thereby enhancing solution quality. Another potential strategy involves developing application-oriented or scenario-specific generative artificial intelligence solvers for the logistics domain, leveraging existing generic large models like ChatGPT. These solvers would be constructed based on demand points, vehicle routes, and logistics networks, facilitating the integration of multi-task learning related to various elements, such as facility location selection and route planning. This approach empowers the solver with the capacity for self-learning to generate solving algorithms.

8.3 Future Work

This section provides suggestions for future endeavours that address some constraints of the thesis at hand, extend the findings of this study, and explore intriguing areas relevant to the conducted research.

- While Chapter 4 primarily concentrated on the feasibility assessment of autonomous delivery solutions within specific regions, there is merit in broadening the scope of the evaluation model to encompass city-wide assessments. This expanded model would encompass a more comprehensive set of indicators, entities, and their respective attributes, including customer profiles, interactions with buildings, and vehicular dynamics. This endeavour necessitates the creation of a proficient tool for seamlessly translating real-world scenarios and data into a simulation environment, for instance, through the utilisation of a digital twin.

- The game framework outlined in Chapter 5 can be extended to incorporate a multi-player game model considering more cost units and complex constraints. This extension could involve the inclusion of various logistics operators or governmental entities in the system, resulting in a more intricate game dynamic. It would necessitate a redefinition of the payoff functions and equilibrium models for each player. Furthermore, investigating how to devise an indirect service pricing mechanism to incentivise consumers to make more

rational decisions, thereby achieving a decentralised system and more efficient resource utilisation, is a topic worthy of exploration.

- To the best of the author's knowledge, Chapter 6 presented the first RL-based algorithm to solve the location-routing problems of MPLs. Subsequent research endeavors could involve extending the current model to account for a multi-echelon delivery network. This could entail introducing AVs as the first echelon responsible for transporting parcels between depots and MPLs. The MPLs, constituting the second echelon, would handle parcel storage and accompany delivery staff. These staff staff, representing the third echelon, would then facilitate the handover of parcels to customers. From the perspective of solution algorithm, further work could delve on developing a novel learning-based algorithm that could translate the complex constraints of the mathematical model to the vectors that could recognised by RL, such that improving the solution quality and computational efficiency.

- The presented thesis assumes that the robot-assisted delivery scenario in Chapter 7 do not consider recharging robot's battery during the delivery process. Therefore, incorporating constraints related to robot battery recharge into the model to enhance service coverage represents a potential future research focus. Compared to scenarios commonly addressed in existing literature, which either neglect charging considerations or involve battery replacement, the inclusion of battery recharge introduces more intricate scheduling strategies. Consequently, this leads to path planning solutions distinct from those derived in current studies.

- Subsequent studies could emphasise the operational planning of cooperative delivery scenarios, employing diverse categories of autonomous delivery fleets to partition delivery requests. This requires the design of a rational order allocation mechanism to optimise the overall efficiency of the system.

Bibliography

Abhishek, Legros, B., & Fransoo, J. C. (2021). Performance evaluation of stochastic systems with dedicated delivery bays and general on-street parking. *Transportation Science*, *55*, 1070–1087.

Acemoglu, D., Makhdoumi, A., Malekian, A., & Ozdaglar, A. (2018). Informational braess' paradox: The effect of information on traffic congestion. *Operations Research*, *66*, 893–917.

Aized, T., & Srai, J. S. (2014). Hierarchical modelling of last mile logistic distribution system. *The International Journal of Advanced Manufacturing Technology*, *70*, 1053–1061.

Alfandari, L., Ljubić, I., & da Silva, M. D. M. (2022). A tailored benders decomposition approach for last-mile delivery with autonomous robots. *European Journal of Operational Research*, *299*, 510–525.

Allen, J., Bektas, T., Cherrett, T., Bates, O., Friday, A., McLeod, F., Piecyk, M., Piotrowska, M., Nguyen, T., & Wise, S. (2018). The scope for pavement porters: addressing the challenges of last-mile parcel delivery in london. *Transportation Research Record*, *2672*, 184–193.

Allen, J., Bektaş, T., Cherrett, T., Friday, A., McLeod, F., Piecyk, M., Piotrowska, M., & Austwick, M. Z. (2017). Enabling a freight traffic controller for collaborative multidrop urban logistics: Practical and theoretical challenges. *Transportation Research Record*, *2609*, 77–84.

Alverhed, E., Hellgren, S., Isaksson, H., Olsson, L., Palmqvist, H., & Flodén, J. (2024). Autonomous last-mile delivery robots: a literature review. *European Transport Research Review*, *16*, 4.

Anderson, S. P., Goeree, J. K., & Holt, C. A. (2002). The logit equilibrium: A perspective on intuitive behavioral anomalies. *Southern Economic Journal*, *69*, 21–47.

Ansari, S., Başdere, M., Li, X., Ouyang, Y., & Smilowitz, K. (2018). Advancements in continuous approximation models for logistics and transportation systems: 1996–2016. *Transportation Research Part B: Methodological*, *107*, 229–252.

Arentze, T. A., Dellaert, B. G., & Chorus, C. G. (2015). Incorporating mental representations in discrete choice models of travel behavior: modeling approach and empirical application. *Transportation Science*, *49*, 577–590.

Awasthi, A., Chauhan, S. S., & Goyal, S. K. (2011). A multi-criteria decision making approach for location planning for urban distribution centers under uncertainty. *Mathematical and computer modelling*, *53*, 98–109.

Bakach, I., Campbell, A. M., & Ehmke, J. F. (2021). A two-tier urban delivery network with robot-based deliveries. *Networks*, *78*, 461–483.

Balassa, B. E., Koteczki, R., Lukács, B., & Buics, L. (2023). Sustainability aspects of drone-assisted last-mile delivery systems—a discrete event simulation approach. *Energies*, *16*, 4656.

Ballano, A., Al-Rahamneh, A., Serrano-Hernandez, A., & Faulin, J. (2023). Agent-based modelling and simulation for hub and electric last mile distribution in vienna. *Procedia Computer Science*, *220*, 718–723.

Ban, X. J., Dessouky, M., Pang, J.-S., & Fan, R. (2019). A general equilibrium model for transportation systems with e-hailing services and flow congestion. *Transportation Research Part B: Methodological*, *129*, 273–304.

Bansal, P., Kockelman, K. M., & Singh, A. (2016). Assessing public opinions of and interest in new vehicle technologies: An austin perspective. *Transportation Research Part C: Emerging Technologies*, *67*, 1–14.

Baum, L., Assmann, T., & Strubelt, H. (2019). State of the art-automated micro-vehicles for urban logistics. *IFAC-PapersOnLine*, *52*, 2455–2462.

Bell, M. G. (2022). Future directions in city logistics. In *Global Logistics and Supply Chain Strategies for the 2020s: Vital Skills for the Next Generation* (pp. 175–189). Springer.

Bello, I., Pham, H., Le, Q. V., Norouzi, M., & Bengio, S. (2016). Neural combinatorial optimization with reinforcement learning. *arXiv preprint arXiv:1611.09940*, .

Van den Berg, H. (2016). *Economic growth and development*. World Scientific Publishing Company.

Biba, J. (2022). What is robotics as a service (raas)? URL: `https://builtin.com/robotics/robotics-as-a-service-raas` accessed: 29 May 2024.

Blakemore, T. (2018). Truck operating costs report for 2018. URL: `https://www.thetruckexpert.co.uk/truck-operating-costs-report-for-2018/` accessed: 23 Dec 2023.

Boyer, K. K., Prud'homme, A. M., & Chung, W. (2009). The last mile challenge: evaluating the effects of customer density and delivery window patterns. *Journal of business logistics*, *30*, 185–201.

Boysen, N., Fedtke, S., & Schwerdfeger, S. (2021). Last-mile delivery concepts: a survey from an operational research perspective. *Or Spectrum*, *43*, 1–58.

Boysen, N., Schwerdfeger, S., & Weidinger, F. (2018). Scheduling last-mile deliveries with truck-based autonomous robots. *European Journal of Operational Research*, *271*, 1085–1099.

Braekers, K., Ramaekers, K., & Van Nieuwenhuyse, I. (2016). The vehicle routing problem: State of the art classification and review. *Computers & industrial engineering*, *99*, 300–313.

Braess, D. (1968). Über ein paradoxon aus der verkehrsplanung. *Unternehmensforschung*, *12*, 258–268.

Branzei, R., Dimitrov, D., & Tijs, S. (2008). *Models in cooperative game theory* volume 556. Springer Science & Business Media.

Bray, G., & Cebon, D. (2022). Selection of vehicle size and extent of multi-drop deliveries for autonomous goods vehicles: An assessment of potential for change. *Transportation Research Part E: Logistics and Transportation Review*, *164*, 102806.

BusinessEnergy.com (2023). Compare business electricity rates. URL: `https://www.businessenergy.com/business-electricity/` accessed: 23 Dec 2023.

Cai, J., Luo, J., Wang, S., & Yang, S. (2018). Feature selection in machine learning: A new perspective. *Neurocomputing*, *300*, 70–79.

Cameron, K. R. (2021). Design by learning–increasing the efficacy of last mile services by inferring consumer intentions, . URL: `https://www.pomsmeetings.org/ConfPapers/051/051-0458.pdf`.

Campbell, J. F. (2013). A continuous approximation model for time definite many-to-many transportation. *Transportation Research Part B: Methodological*, *54*, 100–112.

Cappart, Q., Moisan, T., Rousseau, L.-M., Prémont-Schwarz, I., & Cire, A. A. (2021). Combining reinforcement learning and constraint programming for combinatorial optimization. In *Proceedings of the AAAI Conference on Artificial Intelligence* (pp. 3677–3687). volume 35.

Castillo, V. E., Bell, J. E., Rose, W. J., & Rodrigues, A. M. (2018). Crowdsourcing last mile delivery: strategic implications and future research directions. *Journal of Business Logistics*, *39*, 7–25.

Cattaruzza, D., Absi, N., Feillet, D., & González-Feliu, J. (2017). Vehicle routing problems for city logistics. *EURO Journal on Transportation and Logistics*, *6*, 51–79.

Caunhye, A. M., Nie, X., & Pokharel, S. (2012). Optimization models in emergency logistics: A literature review. *Socio-economic planning sciences*, *46*, 4–13.

Chen, C., Demir, E., & Huang, Y. (2021a). An adaptive large neighborhood search heuristic for the vehicle routing problem with time windows and delivery robots. *European journal of operational research*, *294*, 1164–1180.

Chen, C., Demir, E., Huang, Y., & Qiu, R. (2021b). The adoption of self-driving delivery robots in last mile logistics. *Transportation research part E: logistics and transportation review*, *146*, 102214.

Chen, J., Li, R., Ma, J., & An, Q. (2024). Stationary versus mobile parcel lockers: Which self-service technology moves the consumers in the last mile of urban areas? *Travel Behaviour and Society*, *35*, 100742.

Chen, X., & Tian, Y. (2019). Learning to perform local rewriting for combinatorial optimization. *Advances in Neural Information Processing Systems*, *32*.

Chen, Y., Zhang, J., Zeng, Q., & Luo, J. (2016). Equilibrium problem of last mile delivery service choice. *Computer Integrated Manufacturing Systems*, *22*, 2035–2045.

Chiara, G. D., Cheah, L., Azevedo, C. L., & Ben-Akiva, M. E. (2020). A policy-sensitive model of parking choice for commercial vehicles in urban areas. *Transportation Science*, *54*, 606–630.

Clarke, G., & Wright, J. W. (1964). Scheduling of vehicles from a central depot to a number of delivery points. *Operations Research*, *12*, 568–581. URL: `http://www.jstor.org/stable/167703`.

Cregger, J., Machek, E., Epstein, A., Lennertz, T., Shaw, J., Dopart, K., Behan, M. et al. (2020). Emerging automated urban freight delivery concepts: State of the practice scan, .

Cui, Y., Yu, Y., Cai, Z., & Wang, D. (2022). Optimizing road network density considering automobile traffic efficiency: Theoretical approach. *Journal of Urban Planning and Development*, *148*, 04021062.

Daganzo, C. (2005). *Logistics systems analysis*. Springer Science & Business Media.

Daganzo, C. F. (1980). Network representation, continuum approximations and a solution to the spatial aggregation problem of traffic assignment. *Transportation Research Part B: Methodological*, *14*, 229–239.

Daganzo, C. F. (1984a). The distance traveled to visit n points with a maximum of c stops per vehicle: An analytic model and an application. *Transportation science*, *18*, 331–350.

Daganzo, C. F. (1984b). The length of tours in zones of different shapes. *Transportation Research Part B: Methodological*, *18*, 135–145.

Danchuk, V., Ameen, T., Bakulich, O., & Svatko, V. (2019). Simulation processes of construction of optimal routes for the delivery of goods by road transport on urban road network in mountainous cities. In *International Conference TRANSBALTICA: Transportation Science and Technology* (pp. 112–121). Springer.

Daskin, M. S. (1985). Urban transportation networks: Equilibrium analysis with mathematical programming methods.

De Oliveira Jr, G. S., Bialek, J., Marcus, R.-J., & McCarthy, R. (2016). Dose-ranging effect of systemic diphenhydramine on postoperative quality of recovery after ambulatory laparoscopic surgery: a randomized, placebo-controlled, double-blinded, clinical trial. *Journal of Clinical Anesthesia*, *34*, 46–52.

Dekker, R., Bloemhof, J., & Mallidis, I. (2012). Operations research for green logistics–an overview of aspects, issues, contributions and challenges. *European journal of operational research*, *219*, 671–679.

Delfanti, A. (2021). *The warehouse. Workers and robots at Amazon*. Pluto Books.

Deudon, M., Cournut, P., Lacoste, A., Adulyasak, Y., & Rousseau, L.-M. (2018). Learning heuristics for the tsp by policy gradient. In *Integration of Constraint Programming, Artificial Intelligence, and Operations Research: 15th International Conference, CPAIOR 2018, Delft, The Netherlands, June 26–29, 2018, Proceedings 15* (pp. 170–181). Springer.

Deutsch, Y., & Golany, B. (2018). A parcel locker network as a solution to the logistics last mile problem. *International Journal of Production Research*, *56*, 251–261.

Di, X., & Liu, H. X. (2016). Boundedly rational route choice behavior: A review of models and methodologies. *Transportation Research Part B: Methodological*, *85*, 142–179.

Drexl, M., & Schneider, M. (2015). A survey of variants and extensions of the location-routing problem. *European journal of operational research*, *241*, 283–308.

DriveSmart (2023). Van running costs. URL: `https://www.drivesmart.co.uk/runningcostslcv.aspx` accessed: 23 Dec 2023.

Duhamel, C., Lacomme, P., Prins, C., & Prodhon, C. (2010). A grasp× els approach for the capacitated location-routing problem. *Computers & Operations Research*, *37*, 1912–1923.

El Naqa, I., & Murphy, M. J. (2015). *What is machine learning?*. Springer.

Ensafian, H., Andaryan, A. Z., Bell, M. G., Geers, D. G., Kilby, P., & Li, J. (2023). Cost-optimal deployment of autonomous mobile lockers co-operating with couriers for simultaneous pickup and delivery operations. *Transportation Research Part C: Emerging Technologies*, *146*, 103958.

ERI Economic Research Institute (2023). Tractor trailer truck driver salary in london, united kingdom. URL: `https://www.erieri.com/salary/job/tractor-trailer-truck-driver/united-kingdom/london` accessed: 23 Dec 2023.

Escribano Macias, J. (2021). Optimisation of humanitarian response using unmanned aerial vehicles, .

Escribano Macias, J., Goldbeck, N., Hsu, P.-Y., Angeloudis, P., & Ochieng, W. (2020). Endogenous stochastic optimisation for relief distribution assisted with unmanned aerial vehicles. *OR Spectrum*, *42*, 1089–1125.

van Essen, M., Thomas, T., van Berkum, E., & Chorus, C. (2016). From user equilibrium to system optimum: a literature review on the role of travel information, bounded rationality and non-selfish behaviour at the network and individual levels. *Transport reviews*, *36*, 527–548.

European Commission (2022). New rules to improve road safety and enable fully driverless vehicles in the eu. URL: `https://ec.europa.eu/commission/presscorner/detail/en/ip_22_4312`.

Fang, Y., Chu, F., Mammar, S., & Che, A. (2013). An optimal algorithm for automated truck freight transportation via lane reservation strategy. *Transportation Research Part C: Emerging Technologies*, *26*, 170–183.

Fernie, J., & Sparks, L. (2018). *Logistics and retail management: emerging issues and new challenges in the retail supply chain*. Kogan page publishers.

Figliozzi, M., & Jennings, D. (2020). Autonomous delivery robots and their potential impacts on urban freight energy consumption and emissions. *Transportation research procedia*, *46*, 21–28.

Figliozzi, M. A. (2020). Carbon emissions reductions in last mile and grocery deliveries utilizing air and ground autonomous vehicles. *Transportation Research Part D: Transport and Environment*, *85*, 102443.

FleetNews (2023). Car fuel cost calculator. URL: `https://www.fleetnews.co.uk/costs/fuel-cost-calculator/?FuelType=Diesel&PencePerLitre=159.24&Mileage=50000&ManufacturerId=45&Mpg=6.5` accessed: 23 Dec 2023.

Friedman, E., & Mauersberger, F. (2022). Quantal response equilibrium with symmetry: Representation and applications. In *Proceedings of the 23rd ACM Conference on Economics and Computation* (pp. 240–241).

Fudenberg, D., & Tirole, J. (1991). *Game theory*. MIT press.

Gajda, M., Boysen, N., & Gallay, O. (2024). Impact of bot return policies in van-and-bot delivery systems. *International Journal of Production Research*, (pp. 1–20).

Gao, L., Chen, M., Chen, Q., Luo, G., Zhu, N., & Liu, Z. (2020). Learn to design the heuristics for vehicle routing problem. *arXiv preprint arXiv:2002.08539*, .

Gehrke, S. R., Phair, C. D., Russo, B. J., & Smaglik, E. J. (2023). Observed sidewalk autonomous delivery robot interactions with pedestrians and bicyclists. *Transportation research interdisciplinary perspectives*, *18*, 100789.

Gevaers, R., Van de Voorde, E., & Vanelslander, T. (2011). Characteristics and typology of last-mile logistics from an innovation perspective in an urban context. *City distribution and urban freight transport: Multiple perspectives*, (pp. 56–71).

Ghelichi, Z., Gentili, M., & Mirchandani, P. (2024). A simulation-based performance evaluation model for decision support on drone location and delivery scheduling. *Journal of Humanitarian Logistics and Supply Chain Management*, .

Gibbons, R. et al. (1992). A primer in game theory, .

Giuffrida, N., Fajardo-Calderin, J., Masegosa, A. D., Werner, F., Steudter, M., & Pilla, F. (2022). Optimization and machine learning applied to last-mile logistics: A review. *Sustainability*, *14*, 5329.

Gnegel, F., Schaudt, S., Clausen, U., & Fügenschuh, A. (2021). A 2d layered graph approach for scheduling delivery robots, .

Goeree, J. K., Holt, C. A., & Palfrey, T. R. (2016). *Quantal response equilibrium: A stochastic theory of games*. Princeton University Press.

GOV.UK (2023). New laws to safely roll out self-driving vehicles across british roads. URL: `https://www.gov.uk/government/news/new-laws-to-safely-roll-out-self-driving-vehicles-across-british-roads` accessed: 24 Jan 2024.

Gul, F., Natenzon, P., & Pesendorfer, W. (2014). Random choice as behavioral optimization. *Econometrica*, *82*, 1873–1912.

Gunawan, H., Kusuma, A., Abdullah, S. et al. (2020). Parcel locker as the alternative of parcel delivery service of online shopping. *Journal of Computational and Theoretical Nanoscience*, *17*, 1311–1317.

Guo, X. (2013). Toll sequence operation to realize target flow pattern under bounded rationality. *Transportation Research Part B: Methodological*, *56*, 203–216.

Guo, X., Jaramillo, Y. J. L., Bloemhof-Ruwaard, J., & Claassen, G. (2019). On integrating crowdsourced delivery in last-mile logistics: A simulation study to quantify its feasibility. *Journal of Cleaner Production*, *241*, 118365.

Hassin, R., & Haviv, M. (2003). *To queue or not to queue: Equilibrium behavior in queueing systems* volume 59. Springer Science & Business Media.

Hayel, Y., Quadri, D., Jiménez, T., & Brotcorne, L. (2016). Decentralized optimization of last-mile delivery services with non-cooperative bounded rational customers. *Annals of Operations Research*, *239*, 451–469.

Heidari Gharehbolagh, H., Hafezalkotob, A., Makui, A., & Raissi, S. (2016). Cooperative strategies for maximum-flow problem in uncertain decentralized systems using reliability analysis. *Mathematical Problems in Engineering*, *2016*.

Heinen, A. (2021). The impacts of e-commerce growth on the amount of kilometers driven for last-mile delivery in the netherlands: A system dynamic based analysis on the relation between e-commerce and the logistic sector, .

Hinkamp, B., & Ismael, M. (2023). Network optimization for middle mile delivery for an oilfield service company. *Signature*, .

Hoffmann, T., & Prause, G. (2018). On the regulatory framework for last-mile delivery robots. *Machines*, *6*, 33.

Hossain, M. (2022). Self-driving robots: A revolution in the local delivery. URL: `https://cmr.berkeley.edu/2022/04/self-driving-robots-a-revolution-in-the-local-delivery/` accessed: 29 May 2024.

Hu, W., Dong, J., Hwang, B.-g., Ren, R., & Chen, Z. (2019). A scientometrics review on city logistics literature: Research trends, advanced theory and practice. *Sustainability*, *11*, 2724.

Huang, M., Smilowitz, K. R., & Balcik, B. (2013). A continuous approximation approach for

assessment routing in disaster relief. *Transportation Research Part B: Methodological*, *50*, 20–41.

Hughes, S., Moreno, S., Yushimito, W. F., & Huerta-Cánepa, G. (2019). Evaluation of machine learning methodologies to predict stop delivery times from gps data. *Transportation Research Part C: Emerging Technologies*, *109*, 289–304.

Iwan, S., Kijewska, K., & Lemke, J. (2016). Analysis of parcel lockers' efficiency as the last mile delivery solution–the results of the research in poland. *Transportation Research Procedia*, *12*, 644–655.

James, J., Yu, W., & Gu, J. (2019). Online vehicle routing with neural combinatorial optimization and deep reinforcement learning. *IEEE Transactions on Intelligent Transportation Systems*, *20*, 3806–3817.

Jennings, D., & Figliozzi, M. (2019). Study of sidewalk autonomous delivery robots and their potential impacts on freight efficiency and travel. *Transportation Research Record*, *2673*, 317–326.

Ji-Hyland, C., & Allen, D. (2022). What do professional drivers think about their profession? an examination of factors contributing to the driver shortage. *International Journal of Logistics Research and Applications*, *25*, 231–246.

Joerss, M., Neuhaus, F., & Schröder, J. (2016). How customer demands are reshaping last-mile delivery. *The McKinsey Quarterly*, *17*, 1–5.

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, *349*, 255–260.

Kämäräinen, V., Saranen, J., & Holmström, J. (2001). The reception box impact on home delivery efficiency in the e-grocery business. *International Journal of Physical Distribution & Logistics Management*, *31*, 414–426.

Kapser, S., & Abdelrahman, M. (2020). Acceptance of autonomous delivery vehicles for last-mile delivery in germany–extending utaut2 with risk perceptions. *Transportation Research Part C: Emerging Technologies*, *111*, 210–225.

Kapser, S., Abdelrahman, M., & Bernecker, T. (2021). Autonomous delivery vehicles to fight the spread of covid-19–how do men and women differ in their acceptance? *Transportation Research Part A: Policy and Practice*, *148*, 183–198.

Kedia, A., Kusumastuti, D., & Nicholson, A. (2017). Acceptability of collection and delivery points from consumers' perspective: A qualitative case study of christchurch city. *Case Studies on Transport Policy*, *5*, 587–595.

Koh, L. Y., & Yuen, K. F. (2023). Consumer adoption of autonomous delivery robots in cities: Implications on urban planning and design policies. *Cities*, *133*, 104125.

Konstantakopoulos, G. D., Gayialis, S. P., & Kechagias, E. P. (2020). Vehicle routing problem and related algorithms for logistics distribution: A literature review and classification. *Operational research*, (pp. 1–30).

Kool, W., Van Hoof, H., & Welling, M. (2018). Attention, learn to solve routing problems! *arXiv preprint arXiv:1803.08475*, .

Korus, S. (2018). Autonomous delivery robots could lower the cost of last mile delivery by 20-fold. URL: `https://ark-invest.com/articles/analyst-research/autonomous-delivery-robots/` accessed: 23 Dec 2023.

Kötschau, R., Soeffker, N., & Ehmke, J. F. (2023). Mobile parcel lockers with individual customer service. *Networks*, .

Kumar, K., & Kumanan, S. (2012). Decision making in location selection: An integrated approach with clustering and topsis. *IUP Journal of Operations Management*, *11*.

Kusumaningtyas, I., & Lodewijks, G. (2008). Accelerating moving walkway: A review of the characteristics and potential application. *Transportation research part A: policy and practice*, *42*, 591–609.

Lagorio, A., & Pinto, R. (2020). The parcel locker location issues: An overview of factors affecting their location. In *Proceedings of the 8th International Conference on Information Systems, Logistics and Supply Chain: Interconnected Supply Chains in an Era of Innovation, ILS* (pp. 414–421).

Langevin, A., Mbaraga, P., & Campbell, J. F. (1996). Continuous approximation models in freight distribution: An overview. *Transportation Research Part B: Methodological*, *30*, 163–188.

Laroche, R., Trichelair, P., & Des Combes, R. T. (2019). Safe policy improvement with baseline bootstrapping. In *International conference on machine learning* (pp. 3652–3661). PMLR.

Legacy, C., Ashmore, D., Scheurer, J., Stone, J., & Curtis, C. (2019). Planning the driverless city. *Transport reviews*, *39*, 84–102.

Lei, C., & Ouyang, Y. (2018). Continuous approximation for demand balancing in solving large-scale one-commodity pickup and delivery problems. *Transportation Research Part B: Methodological*, *109*, 90–109.

Lemardelé, C., Estrada, M., Pagès, L., & Bachofner, M. (2021). Potentialities of drones and ground autonomous delivery devices for last-mile logistics. *Transportation Research Part E: Logistics and Transportation Review*, *149*, 102325.

Lemke, J., Iwan, S., & Korczak, J. (2016). Usability of the parcel lockers from the customer perspective–the research in polish cities. *Transportation Research Procedia*, *16*, 272–287.

Li, F., & Guo, Z. (2014). Do parking standards matter? evaluating the london parking reform with a matched-pair approach. *Transportation Research Part A: Policy and Practice*, *67*, 352–365.

Li, F., & Kunze, O. (2023). A comparative review of air drones (uavs) and delivery bots (sugvs) for automated last mile home delivery. *Logistics*, *7*, 21.

Li, J., Ensafian, H., Bell, M. G., & Geers, D. G. (2021a). Deploying autonomous mobile lockers

in a two-echelon parcel operation. *Transportation Research Part C: Emerging Technologies*, *128*, 103155.

Li, S., Yan, Z., & Wu, C. (2021b). Learning to delegate for large-scale vehicle routing. *Advances in Neural Information Processing Systems*, *34*, 26198–26211.

Lim, X.-J., Chang, J. Y.-S., Cheah, J.-H., Lim, W. M., Kraus, S., & Dabić, M. (2024). Out of the way, human! understanding post-adoption of last-mile delivery robots. *Technological Forecasting and Social Change*, *201*, 123242.

Lin, Y. H., Wang, Y., He, D., & Lee, L. H. (2020). Last-mile delivery: Optimal locker location under multinomial logit choice model. *Transportation Research Part E: Logistics and Transportation Review*, *142*, 102059.

Lipton, Z. C. (2018). The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue*, *16*, 31–57.

Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. *Neurocomputing*, *234*, 11–26.

Liu, X., & Li, J. (2016). Using support vector machine for online purchase predication. In *2016 International Conference on Logistics, Informatics and Service Sciences (LISS)* (pp. 1–6). IEEE.

Liu, Y., Escribano Macias, J. J., Zhang, K., Ye, Q., & Angeloudis, P. (). Customer behaviour analysis for coordinated delivery service selection using mobile parcel lockers and autonomous robots: A kernelised-support-vector-machine-based congestion-game model. *Available at SSRN 4383034*, .

Liu, Y., Ye, Q., Escribano-Macias, J., Feng, Y., Candela, E., & Angeloudis, P. (2023a). Route planning for last-mile deliveries using mobile parcel lockers: A hybrid q-learning network approach. *Transportation Research Part E: Logistics and Transportation Review*, *177*, 103234.

Liu, Y., Zhang, K., Escribano Macias, J. J., & Angeloudis, P. (2023b). Quantifying the capability of autonomous robot-assisted distribution in last-mile: Insight from rural to urban configurations. *Available at SSRN 4549465*, .

Llorca, C., & Moeckel, R. (2020). Study of cargo bikes for parcel deliveries under different supply, demand and spatial conditions. In *2020 Forum on Integrated and Sustainable Transportation Systems (FISTS)* (pp. 39–44). IEEE.

LMAD (2022). Autonomous delivery legislation in the eu and in the us. URL: `https://www.lmad.eu/news/autonomous-delivery-legislation-eu-us/`.

Lou, Y., Yin, Y., & Lawphongpanich, S. (2010). Robust congestion pricing under boundedly rational user equilibrium. *Transportation Research Part B: Methodological*, *44*, 15–28.

Lunden, I. (2022). Starship technologies picks up 50m from the eu's investment arm to expand its fleet of autonomous delivery robots. URL: `https://techcrunch.com` accessed: 27 Nov 2023.

Lyu, G., & Teo, C.-P. (2022). Last mile innovation: The case of the locker alliance network. *Manufacturing & Service Operations Management*, .

Ma, Q., Ge, S., He, D., Thaker, D., & Drori, I. (1911). Combinatorial optimization by graph pointer networks and hierarchical reinforcement learning. arxiv 2019. *arXiv preprint arXiv:1911.04936*, .

Maas, J., Nitsche, B., & Straube, F. (2023). Systematization of autonomous vehicles in last mile transportation processes–taxonomy development and clustering of existing concepts. *International Journal of Logistics Research and Applications*, (pp. 1–23).

Macharis, C., & Melo, S. (2011). *City distribution and urban freight transport: multiple perspectives*. Edward Elgar Publishing.

Macioszek, E. (2018). First and last mile delivery–problems and issues. In *Advanced Solutions of Transport Systems for Growing Mobility: 14th Scientific and Technical Conference" Transport Systems. Theory & Practice 2017" Selected Papers* (pp. 147–154). Springer.

Malladi, S. S., Christensen, J. M., Ramírez, D., Larsen, A., & Pacino, D. (2022). Stochastic fleet mix optimization: Evaluating electromobility in urban logistics. *Transportation Research Part E: Logistics and Transportation Review*, *158*, 102554.

McFadden, D. et al. (1973). Conditional logit analysis of qualitative choice behavior, .

Melo, S., Macedo, J., & Baptista, P. (2019). Capacity-sharing in logistics solutions: A new pathway towards sustainability. *Transport Policy*, *73*, 143–151.

Merkert, R., Bliemer, M. C., & Fayyaz, M. (2022). Consumer preferences for innovative and traditional last-mile parcel delivery. *International Journal of Physical Distribution & Logistics Management*, *52*, 261–284.

Mishra, S., Sharma, I., & Pani, A. (2023). Analyzing autonomous delivery acceptance in food deserts based on shopping travel patterns. *Transportation Research Part A: Policy and Practice*, *169*, 103589.

Mordor Intelligence (2023). Uk courier express delivery market size and share analysis - growth trends and forecasts (2023 - 2028). URL: `https://www.mordorintelligence.com/industry-reports/united-kingdom-courier-express-and-parcel-market` accessed: 20 Dec 2023.

Muggy, L., & Stamm, J. L. H. (2014). Game theory applications in humanitarian operations: a review. *Journal of Humanitarian Logistics and Supply Chain Management*, .

Nazari, M., Oroojlooy, A., Snyder, L., & Takác, M. (2018). Reinforcement learning for solving the vehicle routing problem. *Advances in neural information processing systems*, *31*.

Ng, A. (2000). Cs229 lecture notes. *CS229 Lecture notes*, *1*, 1–3.

Nguyen, M. H., Pojani, D., Nguyen-Phuoc, D. Q., & Thi, B. N. (2023). What if delivery riders quit? challenges to last-mile logistics during the covid-19 pandemic. *Research in Transportation Business & Management*, *47*, 100941.

Office for National Statistics (2021). Admin-based statistics for property floor space, feasibility research: England and wales. URL: `https://`

www.ons.gov.uk/peoplepopulationandcommunity/housing/methodologies/adminbasedstatisticsforpropertyfloorspacefeasibilityresearchenglandandwales#geographical-analysis-of-median-floor-space accessed: 29 May 2024.

Orenstein, I., Raviv, T., & Sadan, E. (2019). Flexible parcel delivery to automated parcel lockers: models, solution methods and analysis. *EURO Journal on Transportation and Logistics*, *8*, 683–711.

Ostermeier, M., Heimfarth, A., & Hübner, A. (2022). Cost-optimal truck-and-robot routing for last-mile delivery. *Networks*, *79*, 364–389.

Ostermeier, M., Heimfarth, A., & Hübner, A. (2023). The multi-vehicle truck-and-robot routing problem for last-mile delivery. *European Journal of Operational Research*, *310*, 680–697.

Pahwa, A., & Jaller, M. (2022). A cost-based comparative analysis of different last-mile strategies for e-commerce delivery. *Transportation Research Part E: Logistics and Transportation Review*, *164*, 102783.

Pahwa, A., & Jaller, M. (2023). Assessing last-mile distribution resilience under demand disruptions. *Transportation Research Part E: Logistics and Transportation Review*, *172*, 103066.

Pan, S., Giannikas, V., Han, Y., Grover-Silva, E., & Qiao, B. (2017). Using customer-related data to enhance e-grocery home delivery. *Industrial Management & Data Systems*, .

Pani, A., Mishra, S., Golias, M., & Figliozzi, M. (2020). Evaluating public acceptance of autonomous delivery robots during covid-19 pandemic. *Transportation research part D: transport and environment*, *89*, 102600.

Papola, A. (2016). A new random utility model with flexible correlation pattern and closed-form covariance expression: The corum. *Transportation Research Part B: Methodological*, *94*, 80–96.

Pasquesoone, M. (2022). Analysis of the effects from the mobility orientation law on geographical inequalities of mobility and low-carbon mobility in french low-density

inter-municipalities. URL: http://www.diva-portal.org/smash/get/diva2:1660793/FULLTEXT01.pdf.

Pedraza-Martinez, A. J., & Van Wassenhove, L. N. (2012). Transportation and vehicle fleet management in humanitarian logistics: challenges for future research. *EURO Journal on Transportation and Logistics*, *1*, 185–196.

Peng, F., Wang, X., & Ouyang, Y. (2014). Approximation of discrete spatial data for continuous facility location design. *Integrated Computer-Aided Engineering*, *21*, 311–320.

Perboli, G., Rosano, M., Saint-Guillain, M., & Rizzo, P. (2018). Simulation–optimisation framework for city logistics: an application on multimodal last-mile delivery. *IET Intelligent Transport Systems*, *12*, 262–269.

Petrik, O., Moura, F., & de Abreu e Silva, J. (2014). The influence of the volume–delay function on uncertainty assessment for a four-step model. *Computer-based modelling and optimization in transportation*, (pp. 293–306).

Pham, D. T., Dimov, S. S., & Nguyen, C. D. (2005). Selection of k in k-means clustering. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, *219*, 103–119.

Plissonneau, A., Trentesaux, D., Ben-Messaoud, W., & Bekrar, A. (2022). An imitation learning approach for vehicles longitudinal obstacle avoidance in logistics and transportation. In *International Workshop on Service Orientation in Holonic and Multi-Agent Manufacturing* (pp. 527–543). Springer.

Poeting, M., Schaudt, S., & Clausen, U. (2019a). A comprehensive case study in last-mile delivery concepts for parcel robots. In *2019 winter simulation conference (WSC)* (pp. 1779–1788). IEEE.

Poeting, M., Schaudt, S., & Clausen, U. (2019b). Simulation of an optimized last-mile parcel delivery network involving delivery robots. In *Advances in Production, Logistics and Traffic: Proceedings of the 4th Interdisciplinary Conference on Production Logistics and Traffic 2019 4* (pp. 1–19). Springer.

Precedence Research (2023). Autonomous last mile delivery market. URL: https://www.precedenceresearch.com/autonomous-last-mile-delivery-market accessed: 27 Nov 2023.

Protasov, S., Karpyshev, P., Kalinov, I., Kopanev, P., Mikhailovskiy, N., Sedunin, A., & Tsetserukou, D. (2021). Cnn-based omnidirectional object detection for hermesbot autonomous delivery robot with preliminary frame classification. In *2021 20th International Conference on Advanced Robotics (ICAR)* (pp. 517–522). IEEE.

Punakivi, M., & Saranen, J. (2001). Identifying the success factors in e-grocery home delivery. *International Journal of Retail & Distribution Management*, *29*, 156–163.

Quora (2024). What is the difference between macroscopic, mesoscopic, and microscopic models? URL: https://www.quora.com/What-is-the-difference-between-Macroscopic-Mesoscopic-and-Microscopic-models accessed: 24 May 2024.

Raj, G., Roy, D., de Koster, R., & Bansal, V. (2024). Stochastic modeling of integrated order fulfillment processes with delivery time promise: Order picking, batching, and last-mile delivery. *European Journal of Operational Research*, .

Rao, C., Goh, M., Zhao, Y., & Zheng, J. (2015). Location selection of city logistics centers under sustainability. *Transportation Research Part D: Transport and Environment*, *36*, 29–44.

Rastogi, R., Thaniarasu, I., & Chandra, S. (2011). Design implications of walking speed for pedestrian facilities. *Journal of transportation engineering*, *137*, 687–696.

Reed, S., Campbell, A. M., & Thomas, B. W. (2022a). Impact of autonomous vehicle assisted last-mile delivery in urban to rural settings. *Transportation Science*, *56*, 1530–1548.

Reed, S., Campbell, A. M., & Thomas, B. W. (2022b). The value of autonomous vehicles for last-mile deliveries in urban environments. *Management Science*, *68*, 280–299.

Riemann, R., Wang, D. Z., & Busch, F. (2015). Optimal location of wireless charging facilities for electric vehicles: Flow-capturing location model with stochastic user equilibrium. *Transportation Research Part C: Emerging Technologies*, *58*, 1–12.

Rogers, B. W., Palfrey, T. R., & Camerer, C. F. (2009). Heterogeneous quantal response equilibrium and cognitive hierarchies. *Journal of Economic Theory*, *144*, 1440–1467.

Roughgarden, T. (2005). *Selfish routing and the price of anarchy*. MIT press.

Rusmevichientong, P., Shen, Z.-J. M., & Shmoys, D. B. (2010). Dynamic assortment optimization with a multinomial logit choice model and capacity constraint. *Operations research*, *58*, 1666–1680.

Scherr, Y. O., Hewitt, M., Saavedra, B. A. N., & Mattfeld, D. C. (2020). Dynamic discretization discovery for the service network design problem with mixed autonomous fleets. *Transportation Research Part B: Methodological*, *141*, 164–195.

Scherr, Y. O., Saavedra, B. A. N., Hewitt, M., & Mattfeld, D. C. (2019). Service network design with mixed autonomous fleets. *Transportation Research Part E: Logistics and Transportation Review*, *124*, 40–55.

von Schönfeld, K. C., & Ferreira, A. (2021). Urban planning and european innovation policy: Achieving sustainability, social inclusion, and economic growth? *Sustainability*, *13*, 1137.

Schwerdfeger, S., & Boysen, N. (2020). Optimizing the changing locations of mobile parcel lockers in last-mile distribution. *European Journal of Operational Research*, *285*, 1077–1094.

Schwerdfeger, S., & Boysen, N. (2022). Who moves the locker? a benchmark study of alternative mobile parcel locker concepts. *Transportation Research Part C: Emerging Technologies*, *142*, 103780.

Serrano-Hernandez, A., Martinez-Abad, S., Ballano, A., Faulin, J., Rabe, M., & Chicaiza-Vaca, J. (2021). A hybrid modeling approach for automated parcel lockers as a last-mile delivery scheme: a case study in pamplona (spain). In *2021 Winter Simulation Conference (WSC)* (pp. 1–12). IEEE.

Silva, D. F., Vinel, A., & Kirkici, B. (2022). On-demand public transit: A markovian continuous approximation model. *Transportation Science*, *56*, 704–724.

Simoni, M. D., Kutanoglu, E., & Claudel, C. G. (2020). Optimization and analysis of a robot-assisted last mile delivery system. *Transportation Research Part E: Logistics and Transportation Review*, *142*, 102049.

Slabinac, M. (2015). Innovative solutions for a "last-mile" delivery–a european experience. *Business Logistics in Modern Management*, .

Sonneberg, M.-O., Leyerer, M., Kleinschmidt, A., Knigge, F., & Breitner, M. H. (2019). Autonomous unmanned ground vehicles for urban logistics: Optimization of last mile delivery operations, .

Srinivas, S., Ramachandiran, S., & Rajendran, S. (2022). Autonomous robot-driven deliveries: A review of recent developments and future directions. *Transportation research part E: logistics and transportation review*, *165*, 102834.

Starship Technologies (2023). Starship technologies. URL: `https://www.starship.xyz/company/` accessed: 27 Nov 2023.

Statista (2022). Market share of the leading courier service providers in the united kingdom by volume in 2021. URL: `https://www.statista.com/statistics/1245558/courier-service-providers-market-share-united-kingdom/` accessed: 20 Dec 2023.

Statista (2023). Population of london in 2022, by borough. URL: `https://www.statista.com/statistics/381055/london-population-by-borough/` accessed: 20 Dec 2023.

Straits Research (2023). Autonomous last mile delivery market. URL: `https://straitsresearch.com/report/autonomous-last-mile-delivery-market`.

Suguna, M., Shah, B., Raj, S. K., & Suresh, M. (2021). A study on the influential factors of the last mile delivery projects during covid-19 era. *Operations Management Research*, (pp. 1–14).

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Tang, K., Yang, P., & Yao, X. (2016). Negatively correlated search. *IEEE Journal on Selected Areas in Communications*, *34*, 542–550.

Ulmer, M. W., & Streng, S. (2019). Same-day delivery with pickup stations and autonomous vehicles. *Computers & Operations Research*, *108*, 1–19.

US Bureau Public Roads, O. (1964). *Traffic Assignment Manual for Application with a Large, High Speed Computer*. US Department of Commerce.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*.

Vincent, F. Y., Jodiawan, P., Hou, M.-L., & Gunawan, A. (2021). Design of a two-echelon freight distribution system in last-mile logistics considering covering locations and occasional drivers. *Transportation Research Part E: Logistics and Transportation Review*, *154*, 102461.

Vincent, F. Y., Susanto, H., Jodiawan, P., Ho, T.-W., Lin, S.-W., & Huang, Y.-T. (2022). A simulated annealing algorithm for the vehicle routing problem with parcel lockers. *IEEE Access*, *10*, 20764–20782.

Vinyals, O., Fortunato, M., & Jaitly, N. (2015). Pointer networks. *Advances in neural information processing systems*, *28*.

Von Neumann, J., & Morgenstern, O. (2007). *Theory of games and economic behavior (60th Anniversary Commemorative Edition)*. Princeton university press.

Walpole, R. E., Myers, R. H., Myers, S. L., & Keying, E. Y. (2016). Probability & statistics for engineers & scientists, mystatlab.

Wang, B., Li, Z., Wang, S., Li, M., & Ji, A. (2022). Modeling bounded rationality in discretionary lane change with the quantal response equilibrium of game theory. *Transportation research part B: methodological*, *164*, 145–161.

Wang, J. Y., Raith, A., & Ehrgott, M. (2010). Tolling analysis with bi-objective traffic assignment. In *Multiple criteria decision making for sustainable energy and transportation systems* (pp. 117–129). Springer.

Wang, Q., & Tang, C. (2021). Deep reinforcement learning for transportation network combinatorial optimization: A survey. *Knowledge-Based Systems*, *233*, 107526.

Wang, X., & Huang, G. Q. (2021). When and how to share first-mile parcel collection service. *European Journal of Operational Research*, *288*, 153–169.

Wang, X., Xiao, N., Xie, L., Frazzoli, E., & Rus, D. (2015). Analysis of price of anarchy in traffic networks with heterogeneous price-sensitivity populations. *IEEE Transactions on Control Systems Technology*, *23*, 2227–2237.

Wang, Y., Bi, M., & Chen, Y. (2020). A scheduling strategy of mobile parcel lockers for the last mile delivery problem. *Promet-Traffic&Transportation*, *32*, 875–885.

Wang, Y., Bi, M., Lai, J., Wang, C., Chen, Y., & Holguín-Veras, J. (2024). Recourse strategy for the routing problem of mobile parcel lockers with time windows under uncertain demands. *European Journal of Operational Research*, .

Wang, Y., Zhang, D., Liu, Q., Shen, F., & Lee, L. H. (2016). Towards enhancing the last-mile delivery: An effective crowd-tasking model with scalable solutions. *Transportation Research Part E: Logistics and Transportation Review*, *93*, 279–293.

Weltevreden, J. W. (2008). B2c e-commerce logistics: the rise of collection-and-delivery points in the netherlands. *International journal of retail & distribution management*, .

Wolpert, D. H. (2013). Ubiquity symposium: Evolutionary computation and the processes of life: What the no free lunch theorems really mean: How to improve search algorithms. *Ubiquity*, *2013*, 1–15.

Woollacott, E. (2022). Are delivery robots the future of logistics? URL: `https://www.raconteur.net/customer-experience/delivery-robots-retail-logistics` accessed: 29 May 2024.

Wu, P., Chu, F., Che, A., & Fang, Y. (2017). An efficient two-phase exact algorithm for the automated truck freight transportation problem. *Computers & Industrial Engineering*, *110*, 59–66.

Xu, Z., Chen, A., & Liu, X. (2023). Time and toll trade-off with heterogeneous users: A continuous time surplus maximization bi-objective user equilibrium model. *Transportation Research Part B: Methodological*, *173*, 31–58.

Yan, Y., Chow, A. H., Ho, C. P., Kuo, Y.-H., Wu, Q., & Ying, C. (2022). Reinforcement learning for logistics and supply chain management: Methodologies, state of the art, and future opportunities. *Transportation Research Part E: Logistics and Transportation Review*, *162*, 102712.

Yang, S., Ning, L., Tong, L. C., & Shang, P. (2022). Integrated electric logistics vehicle recharging station location–routing problem with mixed backhauls and recharging strategies. *Transportation Research Part C: Emerging Technologies*, *140*, 103695.

Ye, H., & Yang, H. (2017). Rational behavior adjustment process with boundedly rational user equilibrium. *Transportation Science*, *51*, 968–980.

Ying, X. (2019). An overview of overfitting and its solutions. In *Journal of physics: Conference series* (p. 022022). IOP Publishing volume 1168.

Young, P., & Zamir, S. (2014). *Handbook of game theory*. Elsevier.

Yu, N., Xu, W., & Yu, K.-L. (2020a). Research on regional logistics demand forecast based on improved support vector machine: A case study of qingdao city under the new free trade zone strategy. *IEEE Access*, *8*, 9551–9564.

Yu, S., Puchinger, J., & Sun, S. (2020b). Two-echelon urban deliveries using autonomous vehicles. *Transportation Research Part E: Logistics and Transportation Review*, *141*, 102018.

Yu, S., Puchinger, J., & Sun, S. (2022a). Van-based robot hybrid pickup and delivery routing problem. *European Journal of Operational Research*, *298*, 894–914.

Yu, V. F., Aloina, G., Susanto, H., Effendi, M. K., & Lin, S.-W. (2022b). Regional location routing problem for waste collection using hybrid genetic algorithm-simulated annealing. *Mathematics*, *10*, 2131.

Yuen, K. F., Koh, L. Y., Anwar, M. H. D. B., & Wang, X. (2022). Acceptance of autonomous delivery robots in urban cities. *Cities*, *131*, 104056.

Zhang, J., Campbell, J. F., & Sweeney II, D. C. (2024). A continuous approximation approach to integrated truck and drone delivery systems. *Omega*, (p. 103067).

Zhang, J., & Yang, H. (2015). Modeling route choice inertia in network equilibrium with heterogeneous prevailing choice sets. *Transportation Research Part C: Emerging Technologies*, *57*, 42–54.

Zhang, L. (2011). Behavioral foundation of route choice and traffic assignment: Comparison of principles of user equilibrium traffic assignment under different behavioral assumptions. *Transportation Research Record*, *2254*, 1–10.

Zhao, J., Mao, M., Zhao, X., & Zou, J. (2020). A hybrid of deep reinforcement learning and local search for the vehicle routing problems. *IEEE Transactions on Intelligent Transportation Systems*, *22*, 7208–7218.

Zheng, A., & Casari, A. (2018). *Feature engineering for machine learning: principles and techniques for data scientists*. " O'Reilly Media, Inc.".

Zhu, W., Zhang, J., Ye, S., & Xiang, W. (2023). Braess paradox under the bi-objective user equilibrium. *Expert Systems with Applications*, *213*, 118871.

Appendix A

Route Planning Model of Delivery Option

A.1 Route Planning Model of HV Delivery Option

The route planning model of HV delivery option is established based on the scenario definition described in Section 4.1. Table A.1 shows the notation of the model.

Table A.1: Notation table of route planning model of HV delivery option

Indices		**Sets**	
i	Customer stopover i	C	Set of customer stopovers
j	Customer stopover j		
Parameters		**Decision Variables**	
$T_i^{van_parking}$	Time consumption for parking the van at stopover i [minutes]	x_{ij}	Boolean: 1, if stopover i and j are connected; 0, otherwise
$T_i^{van_sorting}$	Time consumption for sorting parcels at stopover i [minutes]	u_i	Integer: to decide by which stopover i is visited (to eliminate the sub-tour)
$T_i^{van_handover}$	Time consumption for parcel handover at stopover i [minutes]		
t_{ij}	Time consumption travelling from i to j [minutes]		

$$minimise \sum_{i \in C} \sum_{j \in C; i \neq j} (t_{ij} + T_i^{van_parking} + T_i^{van_sorting} + T_i^{van_handover}) x_{ij} \tag{A.1}$$

subject to:

$$\sum_{j \in C; i \neq j} x_{ij} = 1 \qquad \forall i \in C \tag{A.2}$$

$$\sum_{i \in C; i \neq j} x_{ij} = 1 \qquad \forall j \in C \tag{A.3}$$

$$u_i - u_j \leq |C| \cdot (1 - x_{ij}) - 1 \qquad \forall i, j \in C; i \neq j; i, j \neq 0 \tag{A.4}$$

$$u_i \leq |C| - 1 \qquad \forall i \in C \tag{A.5}$$

$$x_{ij} \in \{0, 1\}; x_i \in \mathbb{N}_+ \qquad \forall i, j \in C \tag{A.6}$$

The objective function A.1 aims to minimise the total time consumption of the last-mile journey, which consists of the van travelling time, vehicle parking time, parcel sorting time, and parcel handover time. Constraint A.2 and A.3 ensure that all customer stopovers will be visited only once. Constraint A.4 and A.5 eliminate the sub-tours and determine the visiting order among stopovers. Constraint A.6 defines the domain of decision variables.

A.2 Route Planning Model of MPL Delivery Option

The route planning of the MPL delivery option includes two parts: i) determine the location of MPL stopovers based on customer spatial distribution; ii) determine the optimal route traversing among the defined stopovers. Part i will be achieved through the K-means clustering approach. The route model for part ii is similar to the HV delivery option with the adjustment of the objective function.

A.2.1 Location Model of MPL Stopovers

The location model of MPL stopovers is established based on Ng (2000). Table A.2 shows the notation of the model.

Table A.2: Notation table of location model of MPL delivery option

Indices		**Sets**	
i	Customer location i	N	Set of customers
j	MPL stopover j		
Parameters		**Decision Variables**	
P	Total number of stopovers to be determined [-]	g_j	Float: coordinates of MPL stopover j ($g_j = [X_j, Y_j]$)
X_i	X coordinate of the customer i [-]	b_j	Integer: the cluster index of customer location i
Y_i	Y coordinate of the customer i [-]		
h_i	The location of customer i ($h_i = [X_i, Y_i]$) [-]		

$$minimise \quad b_j = f(g_j) = argmin_{j \in \{1,\cdots,P\}} \|h_i - g_j\|^2 \qquad \forall i \in N \tag{A.7}$$

where:

$$g_i = \frac{\sum_{j=1}^{P} \sum_{i=1}^{N} 1\{b_j = j\} \cdot h_i}{\sum_{j=1}^{P} \sum_{i=1}^{N} 1\{b_j = j\}} \qquad \forall j \in \{1, \cdots, P\} \tag{A.8}$$

$$g_i \in \mathbb{R}, b_j \in \mathbb{N} \qquad \forall i \in N, j \in \{1, \cdots, P\} \tag{A.9}$$

The location model intends to partition all customers' location (data point) into P MPL stopovers (clusters) in which each customer's location belongs to the MPL stopover with the nearest mean. The objective of K-Means clustering is to determine the cluster index b_j (true label) for each customer's location h_i through minimising squared error (total intra-cluster variance) in Equation (A.10), such that makes this an unsupervised learning problem. Equation (A.11) represents the calculation for all centroids of parking spaces through customer location-parking space matching. Equation (A.12) defines the domains of the decision variables. The

author determines the minimum value of P through elbow method.

A.2.2 Route Planning Model for MPL Stopovers

By implementing the location model above, the optimal positioning of MPL stopovers is ascertained, facilitating the calculation of travel time consumption between distinct stopovers in route planning model. Table A.3 shows the notation of the model.

Table A.3: Notation table of route planning model of MPL delivery option

Indices		**Sets**	
i	MPL stopover i	M	Set of MPL stopovers
j	MPL stopover j		
Parameters		**Decision Variables**	
$T_i^{mpl_parking}$	Time consumption for parking the MPL at stopover i [minutes]	x_{ij}	Boolean: 1, if stopover i and j are connected; 0, otherwise
$T_i^{mpl_handover}$	Time consumption for handover parcels at stopover i [minutes]	u_i	Integer: to decide by which stopover i is visited (to eliminate the sub-tour)
$T_i^{mpl_walking}$	Time consumption for customer walking from their location to stopover i [minutes]		
t_{ij}	Time consumption travelling from i to j [minutes]		

$$minimise \sum_{i \in M} \sum_{j \in M; i \neq j} (t_{ij} + T_i^{mpl_parking} + T_i^{mpl_handover} + T_i^{mpl_walking}) x_{ij} \tag{A.10}$$

subject to:

$$\sum_{j \in M; i \neq j} x_{ij} = 1 \qquad \forall i \in M \tag{A.11}$$

$$\sum_{i \in M; i \neq j} x_{ij} = 1 \qquad \forall j \in M \tag{A.12}$$

$$u_i - u_j \leq |M| \cdot (1 - x_{ij}) - 1 \qquad \forall i, j \in M; i \neq j; i, j \neq 0 \tag{A.13}$$

$$u_i \leq |M| - 1 \qquad \forall i \in M \tag{A.14}$$

$$x_{ij} \in \{0,1\}; x_i \in \mathbb{N}_+ \qquad \forall i,j \in M \tag{A.15}$$

The objective function A.10 aims to minimise the total time consumption of the last-mile journey, which consists of the MPL travelling time, vehicle parking time, parcel handover time, and customer walking time. Constraint A.11 and A.12 ensure that all MPL stopovers will be visited only once. Constraint A.13 and A.14 eliminate the sub-tours and determine the visiting order among stopovers. Constraint A.15 defines the domain of decision variables.

A.3 Route Planning Model of ADR Delivery Option

The route planning model of the ADR delivery option is established based on the scenario definition described in Section 4.1. Table A.4 shows the notation of the model. The author further elaborates on the delivery process and the corresponding mathematical notations to establish the route planning model of ADR service. Let s denote a subset that the ADR serves for customers in a single delivery round ($1 \leq |s| \leq q$) and S represents the set including all possible subsets s. For each customer $n \in N$, let $O_n = \{s \in S | n \in N\}$ be the service sets that include customer n.

To fulfil a subset of customers s, ADR is released by the vehicle at customer location i (regarded as starting point), loads the parcels, and rendezvous with the vehicle at customer location j that may differ from location i once set s are fulfilled. The delivery staff will drive the vehicle to the rendezvous point j at the same time when ADR is out for delivery which achieves time-saving.

Let $V(i,j)$ denote the time of the van travels from customer i to customer j, and $R(i,j)$ represent the time of ADR travels from customer i to customer j. Let tuple (i,s,j) denotes the ADR serves subset s starting from customer i and rendezvous with the van at customer j. For each subset s, there exist a customer permutation $(p_1^{isj}, p_2^{isj}, \ldots, p_{|s|}^{isj})$ that achieves

the shortest path for ADR travels from customer i to customer j (i, j can or can not be in subset s). Therefore, the time consumption of ADR following the shortest route to serve customer subset s that starts from customer i and ends at customer j can be denoted by $W_{ij}^{s} = R(i, p_1^{isj}) + R(p_1^{isj}, p_2^{isj}) + \cdots + R(p_{|s|-1}^{isj}, p_{|s|}^{isj}) + R(p_{|s|}^{isj}, j)$.

In one-way streets, the ADR may reach the rendezvous point before the vehicle. Therefore, ADR will wait for the van until it arrives. The waiting time is $max\{V(i,j) - W_{ij}^{s}, 0\}$.

$$minimise \quad f(x_{ij}, y_{ij}^{s}, z_{ij}, u_i) = F_1 + F_2 + F_3 \tag{A.16}$$

$$F_1 = \sum_{i \in \bar{N}} \sum_{j \in \bar{N} \setminus \{i\}} (V(i,j) + T_j^{adr_parking}) \cdot x_{ij} \tag{A.17}$$

$$F_2 = \sum_{i \in N} \sum_{s \in S} \sum_{j \in N} (W_{ij}^{s} + max\{V(i,j) - W_{ij}^{s}, 0\}) \cdot y_{ij}^{s} \tag{A.18}$$

$$F_3 = \sum_{i \in N} \sum_{s \in S} \sum_{j \in N} (T_i^{adr_loading} + T_i^{adr_handover}) \cdot y_{ij}^{s} \tag{A.19}$$

Subject to:

$$\sum_{i \in N} x_{i0} = 1 \tag{A.20}$$

$$\sum_{i \in N} x_{0i} = 1 \tag{A.21}$$

$$\sum_{i \in N} \sum_{s \in O_n} \sum_{j \in N} y_{ij}^{s} = 1 \qquad \forall n \in N \tag{A.22}$$

$$\sum_{s \in S} \sum_{j \in N} y_{ij}^{s} = \sum_{l \in \bar{N}, l \neq i} x_{li} \qquad \forall i \in N \tag{A.23}$$

$$\sum_{i \in N} \sum_{s \in S} y_{ij}^{s} = \sum_{l \in \bar{N}, l \neq j} x_{jl} \qquad \forall j \in N \tag{A.24}$$

Table A.4: Notation table for route planning model of ADR delivery option

Indices		**Sets**	
i	Customer i	N	Set of customers
j	Customer j	$\bar{N}$	Set of customers and UCC
s	Service set s	S	Set of all customer service set served by ADR
Parameters		**Decision Variables**	
$\|N\|$	Number of customers [customer]	x_{ij}	Boolean: $x_{ij} = 1$ if the van travels from customer i to customer j with the ADR and delivery staff onboard for $i, j \in \bar{N}$
q	Maximum capacity of ADR [parcels]	y_{ij}^{s}	Boolean: $y_{ij}^{s} = 1$ if ADR is released at customer i to service customer set s and rendezvous van at customer j for $i, j \in N, s \in S$
U	Maximum battery power of ADR [Kw]	z_{ij}	Integer: the flow of parcels onboard the van when travels from customer i to customer j for $i \in \bar{N}, j \in N, i \neq j$
ξ	Power consumption of ADR [kw/hr]	u_i	Float: the remaining power of ADR in customer i for $i \in N$
$T_i^{adr_loading}$	Time consumption of sorting/loading parcels at customer i [minutes]		
$T_i^{adr_parking}$	Time consumption of parking van at customer i [minutes]		
$T_i^{adr_handover}$	Time consumption of handover parcel at customer i [minutes]		
$V(i,j)$	Time consumption of the vehicle travels from customer i to customer j [minutes]		
$R(i,j)$	Time consumption of ADR travels from customer i to customer j [minutes]		
W_{ij}^{s}	Time consumption of ADR following the shortest route to serve customer set s when starts from customer i and ends at customer j [minutes]		
κ	A sufficiently large number [-]		

$$\sum_{j \in N} z_{0j} = |N| \tag{A.25}$$

$$z_{0i} \leq |N| \cdot x_{0i} \qquad \forall i \in N \tag{A.26}$$

$$z_{ij} \leq |N| \cdot (x_{ij} + \sum_{s \in S} y_{ij}^{s}) \qquad \forall i, j \in N, i \neq j \tag{A.27}$$

$$\sum_{j \in \bar{N} \setminus \{i\}} z_{ji} - \sum_{j \in N \setminus \{i\}} z_{ij} = \sum_{j \in \bar{N}, i \neq j} x_{ij} + \sum_{s \in S} \sum_{j \in N} (|s| - 1) y_{ij}^{s} \qquad \forall i \in N \tag{A.28}$$

$$u_j + \xi \cdot W_{ij}^{s} \leq u_i + \kappa(1 - y_{ij}^{s}) \qquad \forall i, j \in N, s \in S, i \neq j \tag{A.29}$$

$$W_{ij}^{s} \cdot y_{ij}^{s} \cdot \xi \leq u_i \qquad \forall i, j \in \bar{N}, i \neq j \tag{A.30}$$

$$u_0 = U \tag{A.31}$$

$$x_{ij} \in \{0, 1\} \qquad \forall i, j \in \bar{N} \tag{A.32}$$

$$y_{ij}^{s} \in \{0, 1\} \qquad \forall i, j \in N, s \in S \tag{A.33}$$

$$z_{ij} \in \mathbb{Z}_{+} \qquad \forall i \in \bar{N}, j \in N, i \neq j \tag{A.34}$$

$$u_i \in \mathbb{N}_{+} \qquad \forall i \in \bar{N} \tag{A.35}$$

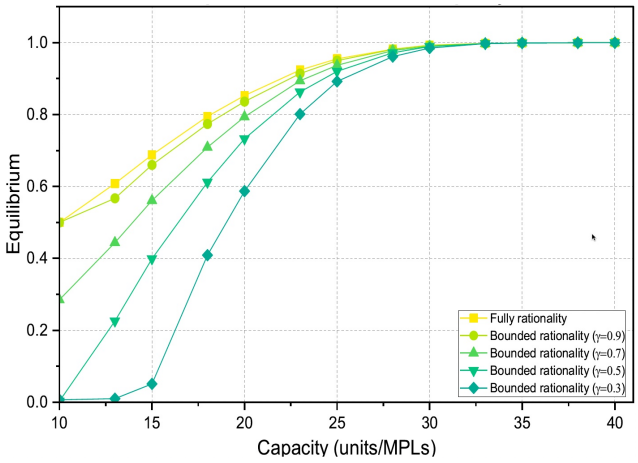
The objective function (A.16) minimises the total delivery time, which includes the travelling time and parking space finding time of the van (Eq.A.17), the travelling and waiting time of ADR (Eq.A.18), and parcel handling and handover time (Eq.A.19). Constraint (A.20) and (A.21) ensures that the ADR is onboard the van when departs and returns to the UCC, respectively. Constraint (A.22) guarantees that all customers are fulfilled. Constraint (A.23) ensures that the van will visit customer i to release ADR if ADR serves subset s. Constraint (A.24) guarantees the van will visit the rendezvous point j to pick up ADR travelling to the next customer when subset s has been fulfilled (i.e., $y_{ij}^{s} = 1$).

Constraint (A.25) represents that the total number of parcels to be delivered is $|N|$. Constraint (A.26) ensures that the van can visit customers from the UCC only when the ADR is onboard (i.e., $x_{0i} = 1$). Constraint (A.27) indicates that the van can travel between customers either with or without ADR onboard. Constraint (A.28) denotes the quantity change of parcels when serving each customer.

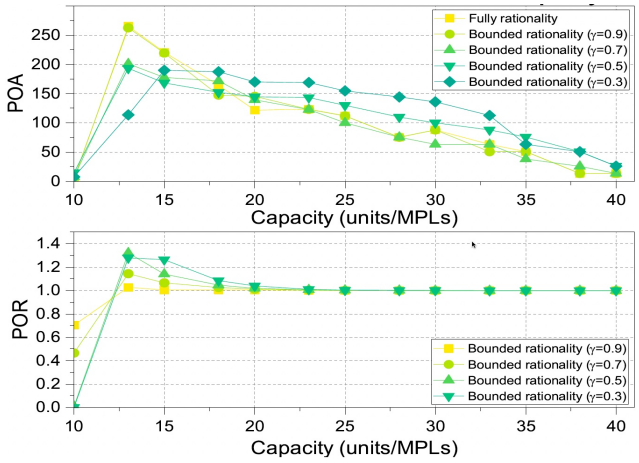
Constraints (A.29) and (A.30) ensure that the ADR must have enough power when travelling from customer i to customer j. Constraint (A.31) denotes that the ADR has maximum battery power when departs from the depot. Constraints (A.32) to (A.35) define the domain of the variables.

Appendix B

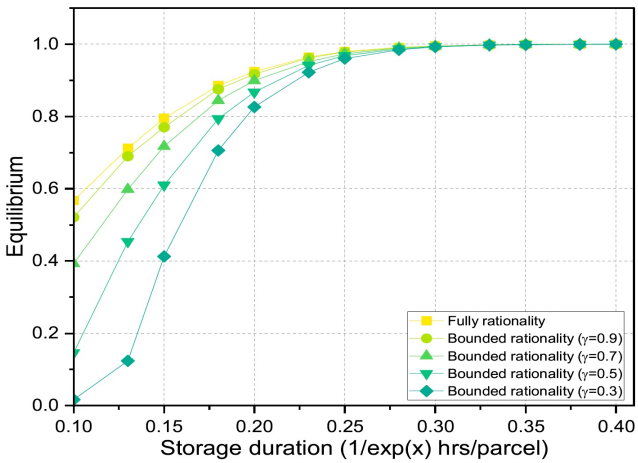
Equilibrium, POA, and POR of Rationality Models with Different Parameters



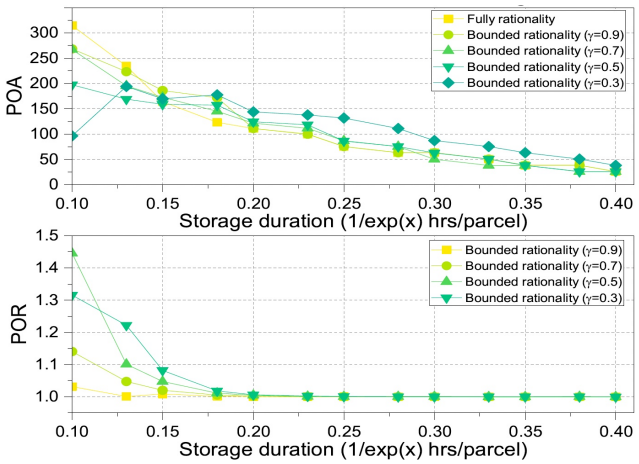
(a) Equilibrium against different capacity



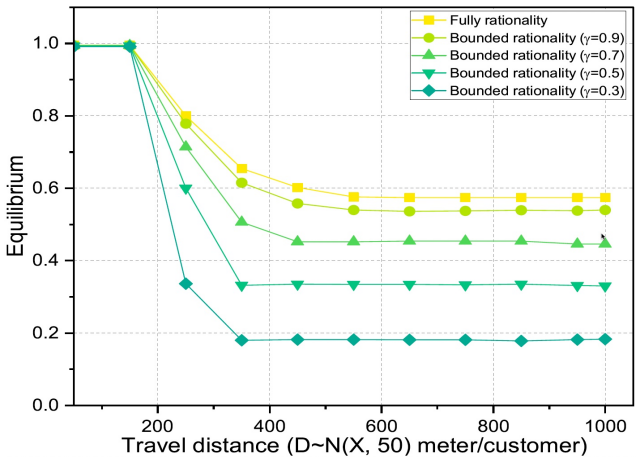
(b) POA & POR against different capacity



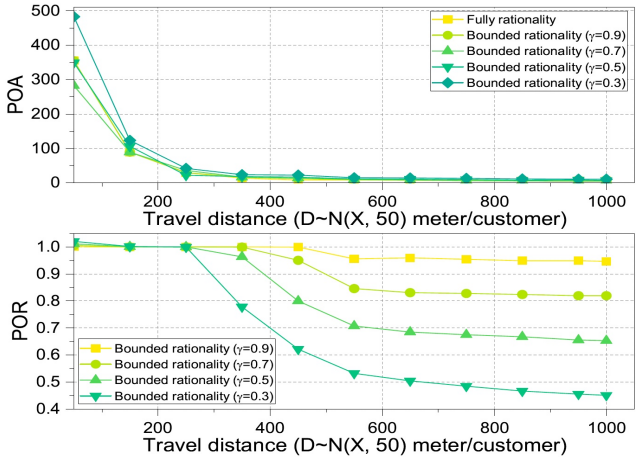
(c) Equilibrium against different storage duration



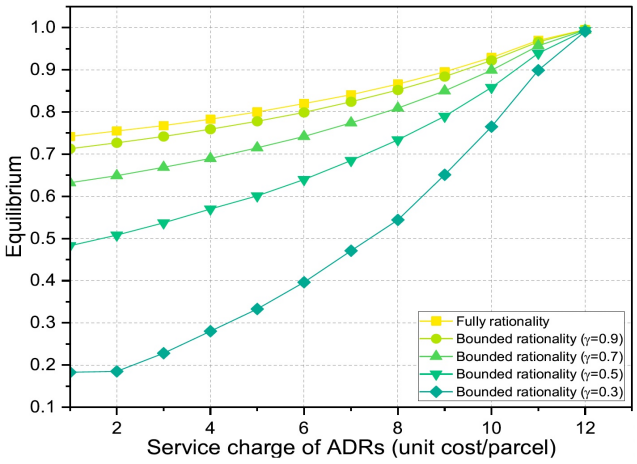
(d) POA & POR against different storage duration



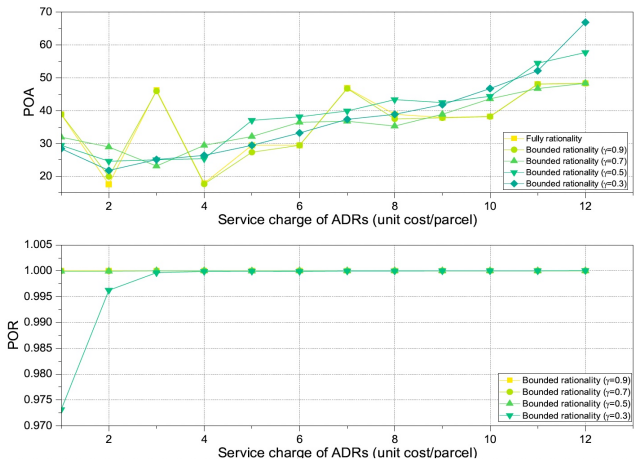
(e) Equilibrium against different travel distance



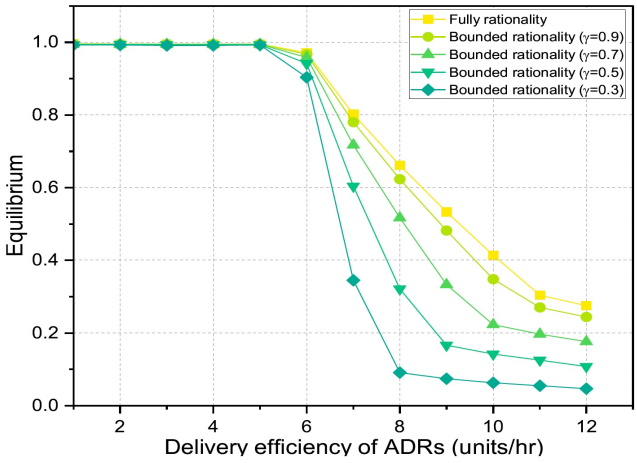
(f) POA & POR against different travel distance



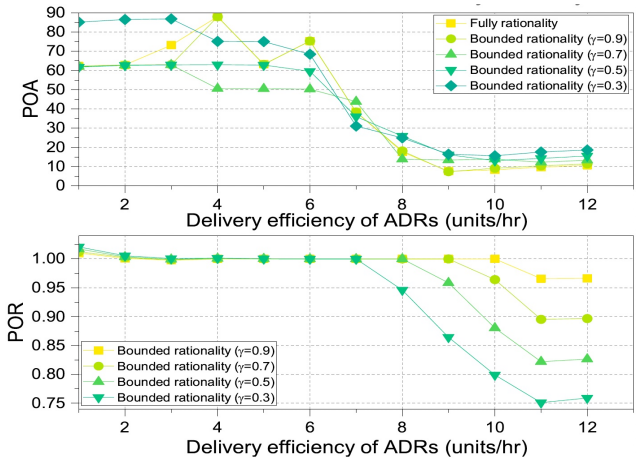
(g) Equilibrium against service charge of ADR



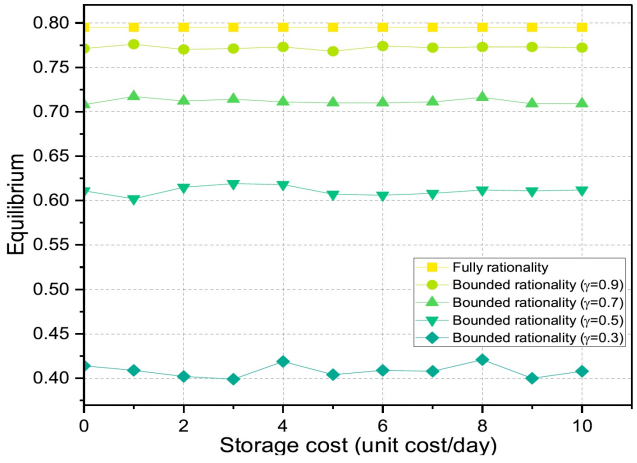
(h) POA & POR against service charge of ADR



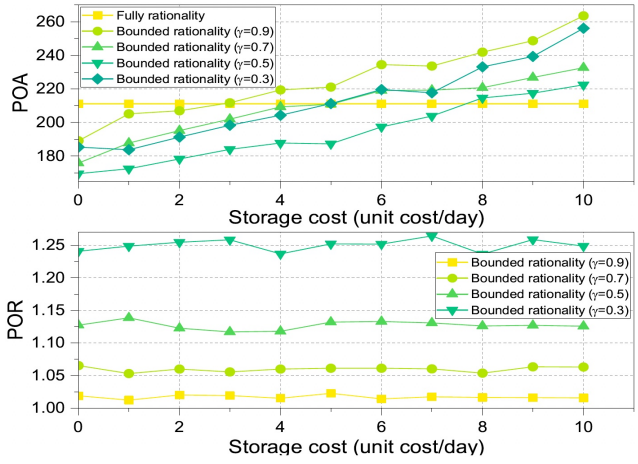
(i) Equilibrium against delivery efficiency of ADR



(j) POA & POR against delivery efficiency of ADR



(k) Equilibrium against storage cost of MPL



(l) POA & POR against storage cost of MPL

Figure B.1: Equilibrium, POA, and POR of rationality models with different parameters

Appendix C

Correlation Analysis between MPLs and ADRs

This appendix shows the POA and POR correlation analysis between MPL and ADR delivery options in terms of service charge and capacity.

Table C.1: POA correlations between MPLs and AGVs in service charge

POA	ADR(rational)	ADR($\gamma = 0.9$)	ADR($\gamma = 0.7$)	ADR($\gamma = 0.5$)	ADR($\gamma = 0.3$)	MPL(rational)	MPL($\gamma = 0.9$)	MPL($\gamma = 0.7$)	MPL($\gamma = 0.5$)	MPL($\gamma = 0.3$)	Avg_ADRs	Avg_MPLs
ADR(rational)	1											
ADR($\gamma = 0.9$)	.996**	1										
ADR($\gamma = 0.7$)	.488	.494	1									
ADR($\gamma = 0.5$)	.626*	.615*	.945**	1								
ADR($\gamma = 0.3$)	.636*	.638*	.934**	.952**	1							
MPL(rational)	-	-	-	-	-	-						
MPL($\gamma = 0.9$)	.540	.537	.896**	.930**	.911**	-	1					
MPL($\gamma = 0.7$)	.484	.466	.826**	.887**	.835**	-	.971**	1				
MPL($\gamma = 0.5$)	.549	.541	.892**	.917**	.907**	-	.978**	.951**	1			
MPL($\gamma = 0.3$)	.588*	.585*	.940**	.954**	.946**	-	.983**	.941**	.980**	1		
Avg_ADRs	.853**	.852**	.861**	.929**	.936**	-	.856**	.785**	.854**	.900**	1	
Avg_MPLs	.551	.544	.904**	.936**	.915**	-	.996**	.975**	.990**	.990**	.864**	1

Notes: 1) Avg_ADRs = Average POA of ADR service in different rationality models; 2) Avg_MPLs = Average POA of MPL service in different rationality models; 3) **: Correlation is significant at the 0.01 level (2-tailed); 4) *: Correlation is significant at the 0.05 level (2-tailed); 5) -: Cannot be computed because at least one of the variables is constant.

Table C.2: POA correlations between MPLs and AGVs in capacity

POA	MPL(rational)	MPL($\gamma = 0.9$)	MPL($\gamma = 0.7$)	MPL($\gamma = 0.5$)	MPL($\gamma = 0.3$)	ADR(rational)	ADR($\gamma = 0.9$)	ADR($\gamma = 0.7$)	ADR($\gamma = 0.5$)	ADR($\gamma = 0.3$)	Avg_MPLs	Avg_ADRs
MPL(rational)	1											
MPL($\gamma = 0.9$)	.993**	1										
MPL($\gamma = 0.7$)	.968**	.966**	1									
MPL($\gamma = 0.5$)	.946**	.948**	.966**	1								
MPL($\gamma = 0.3$)	.706**	.706**	.797**	.860**	1							
ADR(rational)	.635*	.640*	.726**	.591*	.50	1						
ADR($\gamma = 0.9$)	.607*	.612*	.707**	.573*	.478	.996**	1					
ADR($\gamma = 0.7$)	.641*	.659*	.668*	.538	.351	.927**	.15**	1				
ADR($\gamma = 0.5$)	.637*	.660*	.708**	.577*	.437	.974**	.969**	.963**	1			
ADR($\gamma = 0.3$)	.628*	.648*	.669*	.516	.325	.940**	.928**	.959**	.980**	1		
Avg_MPLs	.974**	.975**	.987**	.988**	.840**	.652*	.627*	.609*	.639*	.595*	1	
Avg_ADRs	.640*	.654*	.709**	.569*	.428	.987**	.981**	.967**	.995**	.979**	.636*	1

Notes: 1) Avg_ADRs = Average POA of ADR service in different rationality models; 2) Avg_MPLs = Average POA of MPL service in different rationality models; 3) **: Correlation is significant at the 0.01 level (2-tailed); 4) *: Correlation is significant at the 0.05 level (2-tailed)

Table C.3: POR correlations between MPLs and AGVs in service charge

POR	MPL($\gamma = 0.9$)	MPL($\gamma = 0.7$)	MPL($\gamma = 0.5$)	MPL($\gamma = 0.3$)	ADR($\gamma = 0.9$)	ADR($\gamma = 0.7$)	ADR($\gamma = 0.5$)	ADR($\gamma = 0.3$)	Avg_MPLs	Avg_ADRs
MPL($\gamma = 0.9$)	1									
MPL($\gamma = 0.7$)	.243	1								
MPL($\gamma = 0.5$)	-0.302	-0.077	1							
MPL($\gamma = 0.3$)	.222	.118	.144	1						
ADR($\gamma = 0.9$)	.232	-0.347	-0.51	.309	1					
ADR($\gamma = 0.7$)	.11	.204	-0.379	.184	.051	1				
ADR($\gamma = 0.5$)	.233	.159	-0.27	.385	.118	.928**	1			
ADR($\gamma = 0.3$)	-0.145	-0.355	-0.098	.333	.102	.643*	.677*	1		
Avg_MPLs	.303	.404	.491	.838**	-0.09	.029	.233	.039	1	
Avg_ADRs	-0.143	-0.353	-0.1	.334	.104	.646*	.680*	1**	.04	1

Notes: 1) Avg_ADRs = Average POR of ADR service in different rationality models; 2) Avg_MPLs = Average POR of MPL service in different rationality models; 3) **: Correlation is significant at the 0.01 level (2-tailed); 4) *: Correlation is significant at the 0.05 level (2-tailed)

Table C.4: POR correlations between MPLs and AGVs in capacity

POR	MPL($\gamma = 0.9$)	MPL($\gamma = 0.7$)	MPL($\gamma = 0.5$)	MPL($\gamma = 0.3$)	ADR($\gamma = 0.9$)	ADR($\gamma = 0.7$)	ADR($\gamma = 0.5$)	ADR($\gamma = 0.3$)	Avg_MPLs	Avg_ADRs
MPL($\gamma = 0.9$)	1									
MPL($\gamma = 0.7$)	.984**	1								
MPL($\gamma = 0.5$)	.975**	.999**	1							
MPL($\gamma = 0.3$)	.967**	.992**	.992**	1						
ADR($\gamma = 0.9$)	-0.308	-0.249	-0.235	-0.218	1					
ADR($\gamma = 0.7$)	-0.221	-0.157	-0.142	-0.12	.973**	1				
ADR($\gamma = 0.5$)	-0.228	-0.155	-0.139	-0.112	.904**	.970**	1			
ADR($\gamma = 0.3$)	-0.245	-0.16	-0.142	-0.111	.831**	.905**	.973**	1		
Avg_MPLs	.980**	.999**	.998**	.997**	-0.24	-0.145	-0.141	-0.145	1	
Avg_ADRs	-0.244	-0.167	-0.15	-0.122	.909**	.966**	.997**	.983**	-0.153	1

Notes: 1) Avg_ADRs = Average POR of ADR service in different rationality models; 2) Avg_MPLs = Average POR of MPL service in different rationality models; 3) **: Correlation is significant at the 0.01 level (2-tailed); 4) *: Correlation is significant at the 0.05 level (2-tailed)

Appendix D

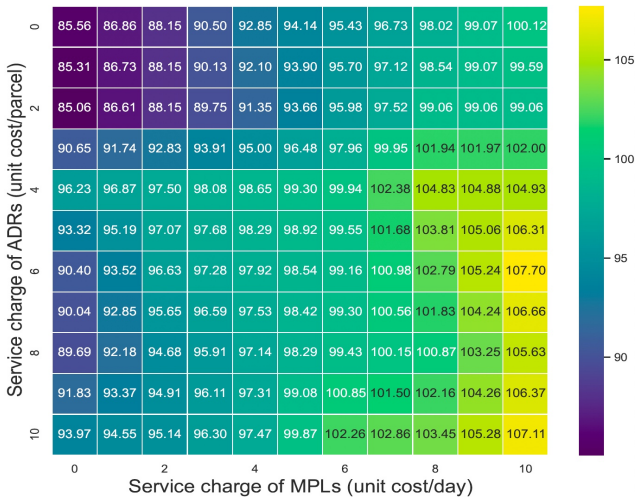
POA & POR of Rationality Models with Different Service Charges



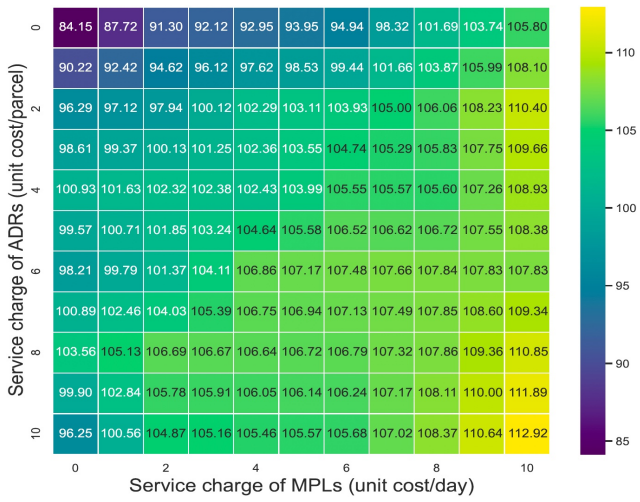
(a) POA of fully rationality model



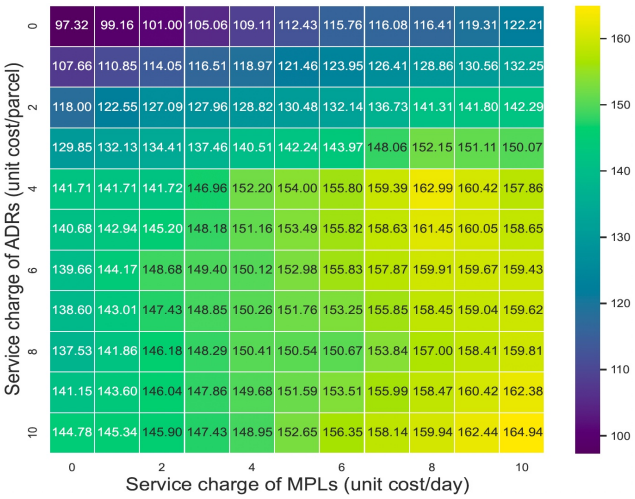
(b) POA of bounded rationality model (γ=0.9)



(c) POA of bounded rationality model (γ=0.7)



(d) POA of bounded rationality model (γ=0.5)

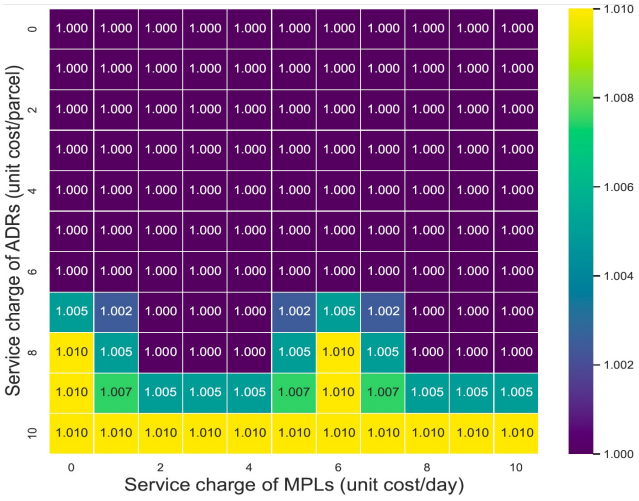


(e) POA of bounded rationality model (γ=0.3)

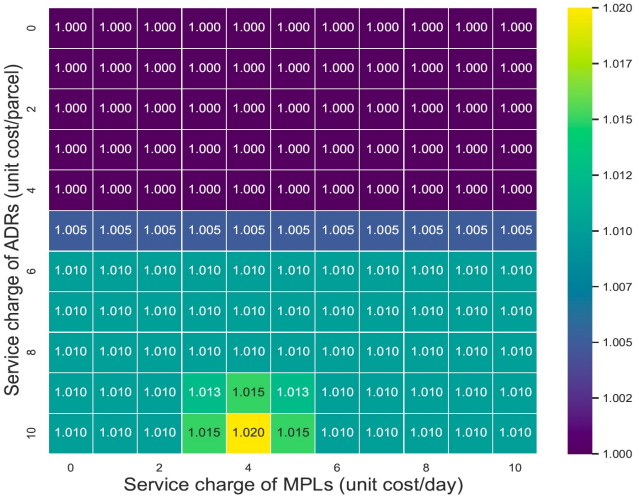
Figure D.1: POA under different service charges of ADRs and MPLs



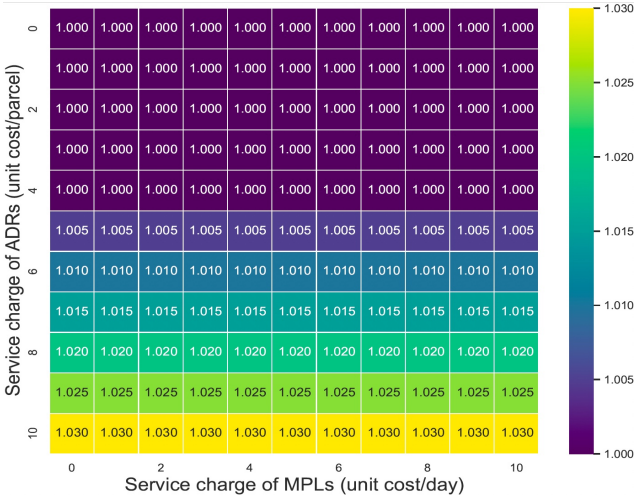
(a) POR of bounded rationality model (γ=0.9)



(b) POR of bounded rationality model (γ=0.7)



(c) POR of bounded rationality model (γ=0.5)

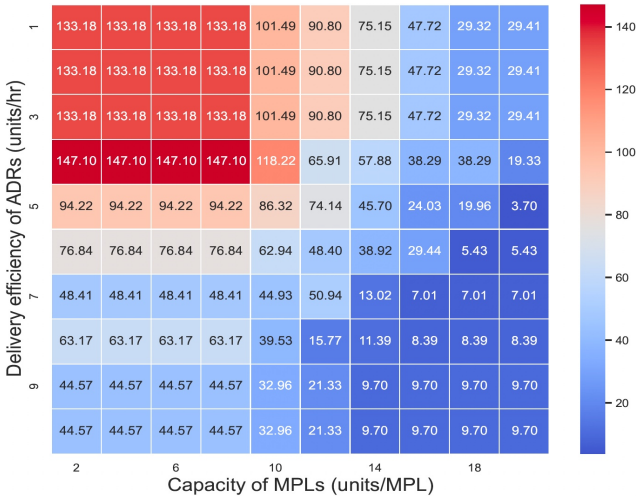


(d) POR of bounded rationality model (γ=0.3)

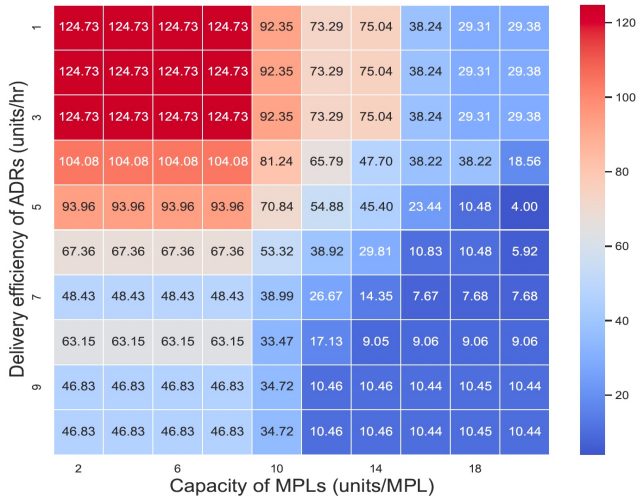
Figure D.2: POR under different service charges of ADRs and MPLs

Appendix E

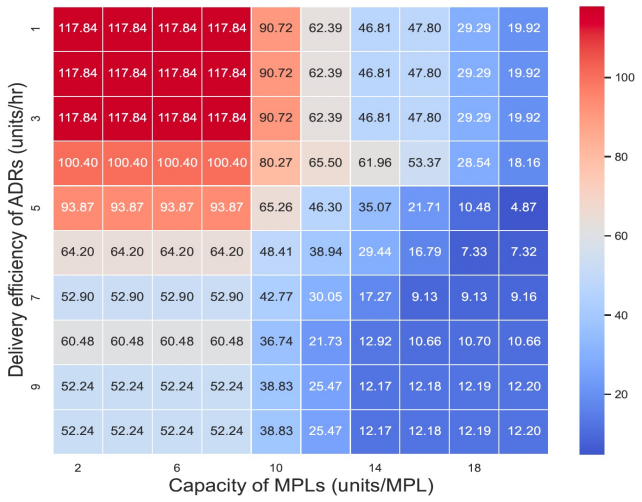
POA & POR of Rationality Models with Different Capacities



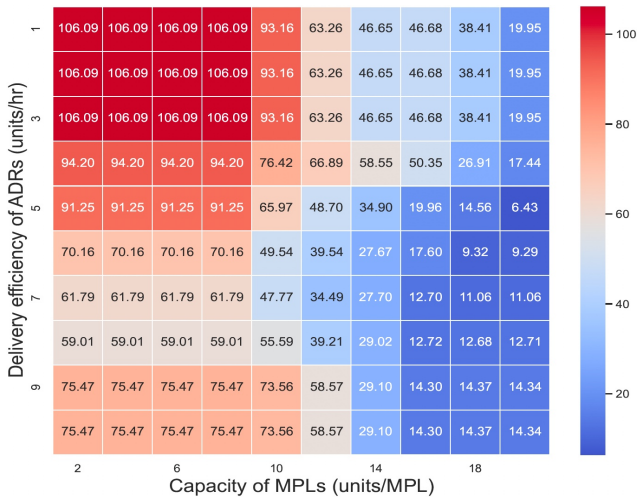
(a) POA of fully rationality model



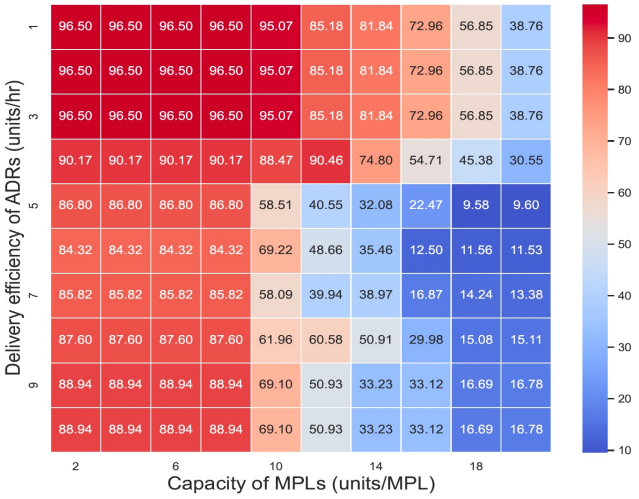
(b) POA of bounded rationality model (γ=0.9)



(c) POA of bounded rationality model (γ=0.7)



(d) POA of bounded rationality model (γ=0.5)

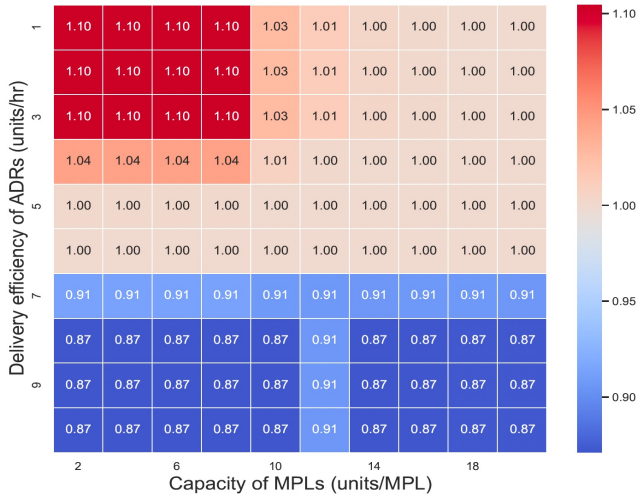


(e) POA of bounded rationality model (γ=0.3)

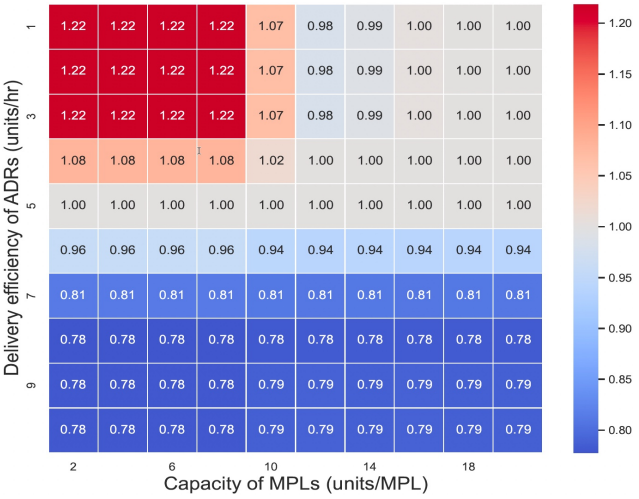
Figure E.1: POA under different capacities of ADRs and MPLs



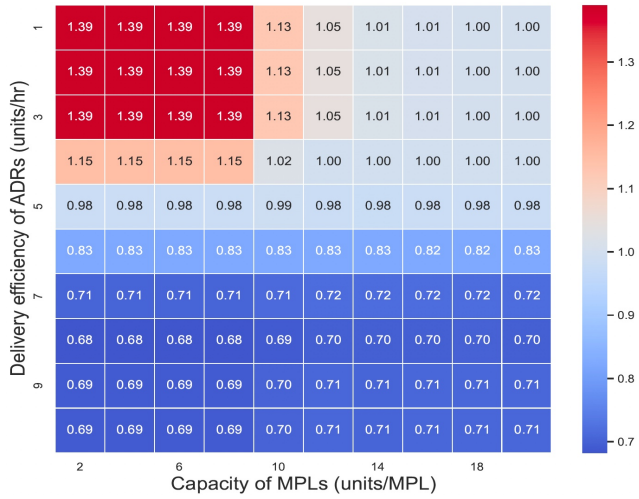
(a) POR of bounded rationality model (γ=0.9)



(b) POR of bounded rationality model (γ=0.7)



(c) POR of bounded rationality model (γ=0.5)



(d) POR of bounded rationality model (γ=0.3)

Figure E.2: POR under different capacities of ADRs and MPLs

Appendix F

Implementation Result of KSVM Regression Prediction

Table F.1: MSE, EVRS, and R Square of different kernel functions of KSVM

Y_{ksvm} \ C	MSE					EVRS					R^2				
	0.01	0.1	1	10	100	0.01	0.1	1	10	100	0.01	0.1	1	10	100
Linear	0.01	1.106	1.106	1.106	1.106	1.106	-0.074	-0.074	-0.074	-0.074	-0.074	-0.095	-0.095	-0.095	-0.095	-0.095
	0.1	1.236	1.236	1.236	1.236	1.236	-0.213	-0.213	-0.213	-0.213	-0.213	-0.223	-0.223	-0.223	-0.223	-0.223
	1	1.377	1.377	1.377	1.377	1.377	-0.355	-0.355	-0.355	-0.355	-0.355	-0.363	-0.363	-0.363	-0.363	-0.363
	10	1.381	1.381	1.381	1.381	1.381	-0.360	-0.360	-0.360	-0.360	-0.360	-0.366	-0.366	-0.366	-0.366	-0.366
	100	1.380	1.380	1.380	1.380	1.380	-0.359	-0.359	-0.359	-0.359	-0.359	-0.366	-0.366	-0.366	-0.366	-0.366
RBF	0.01	1.034	1.060	1.031	1.027	1.025	-0.002	-0.013	0.005	0.001	0.002	-0.024	-0.049	-0.020	-0.016	-0.015
	0.1	1.086	1.113	0.990	1.046	1.029	-0.026	-0.084	0.029	0.003	0.018	-0.075	-0.101	0.020	-0.035	-0.019
	1	1.150	1.235	1.025	1.014	0.972	-0.127	-0.211	-0.014	-0.001	0.041	-0.138	-0.222	-0.014	-0.004	0.038
	10	1.177	1.083	0.805	0.847	0.809	-0.162	-0.063	0.204	0.178	0.200	-0.165	-0.072	0.203	0.162	0.200
	100	1.140	0.938	0.306	0.963	0.809	-0.128	0.080	0.912	0.078	0.200	-0.129	0.072	0.893	0.047	0.200
Sigmoid	0.01	1.030	1.089	1.18e+00	1.34e+00	1.34e+00	0	-0.013	-1.07e-01	-2.29e-01	-2.37e-01	-0.020	-0.078	-1.72e-01	-3.27e-01	-3.27e-01
	0.1	1.092	1.096	3.19e+00	6.51e+00	6.16e+00	-0.012	-0.063	-1.95e+00	-4.8e+00	-4.69e+00	-0.081	-0.084	-2.16e+00	-5.44e+00	-5.10e+00
	1	1.106	1.535	1.97e+02	7.17e+02	7.13e+02	-0.073	-0.465	-1.86e+02	-7.07e+02	-7.04e+02	-0.095	-0.520	-1.94e+02	-7.08e+02	-7.04e+02
	10	1.215	182.976	2.23e+04	7.01e+04	7.02e+04	-0.193	-178.804	-2.20e+04	-6.93e+04	-6.94e+04	-0.202	-180.081	-2.20e+04	-6.93e+04	-6.94e+04
	100	1.288	18957.98	2.76e+06	7.00e+06	7.01e+06	-0.274	-18493.963	-2.73e+06	-6.93e+06	-6.92e+06	-0.275	-18760.687	-2.73e+06	-6.93e+06	-6.93e+06

Appendix G

Derivation for Auxiliary Variables

This appendix provides mathematical derivation for the auxiliary variables according to the parameters and notations introduced in Section 6.2.3 and 6.2.4.

G.1 Derivation for d_{ij}

The author utilised the coordinates of parking space i and j to derive the distance d_{ij} between parking space i and j:

$$d_{ij} = \|g_i - g_j\| = \sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2} \qquad \forall i, j \in I \tag{G.1}$$

G.2 Derivation for t_{ij}

The author utilised the distance d_{ij} from parking space i to j and the velocity v of MPL to derive the time consumption t_{ij} travels from parking space i to j:

$$t_{ij} = d_{ij}/v = \sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2}/v \qquad \forall i, j \in I \tag{G.2}$$

G.3 Derivation for d_{in}^k

To derive the distance from the k^{th} location of customer n to parking space i, we exploit the coordinates of the k^{th} location of customer n and parking space i:

$$d_{in}^k = \sqrt{(X_{nk} - X_i)^2 + (Y_{nk} - Y_i)^2} \qquad \forall i \in I, n \in N, k \in L_k \tag{G.3}$$

G.4 Derivation for e_{ia}, l_{ia}

As mentioned in Section 6.1, for each parking space i, the parking time windows $[E_i, L_i]$ can be divided into $|U_i|$ ($|U_i| = |L_i - E_i|/ST_i$) sub-intervals. Each sub-interval was denoted by $[e_{ia}, l_{ia}]$. We start with the derivation for the first sub-interval of parking space i. Since the starting time of the first sub-interval is the earliest time of parking space i, therefore:

$$e_{i1} = E_i \qquad \forall i \in I \tag{G.4}$$

While the ending time of the first sub-interval l_{ia} is:

$$l_{i1} = e_{i1} + ST_i \qquad \forall i \in I \tag{G.5}$$

Since the ending time of the previous sub-interval and the starting time of the next sub-interval are consecutive. Therefore, the starting time of the second sub-interval of parking space i is:

$$e_{i2} = l_{i1} = e_{i1} + ST_i \qquad \forall i \in I \tag{G.6}$$

While the ending time of the first sub-interval l_{ia} is

$$l_{i2} = e_{i2} + ST_i = e_{i1} + 2 \times ST_i \qquad \forall i \in I \tag{G.7}$$

Following the above logic, the general formulation for deriving e_{ia}, l_{ia} is summarised as follows:

$$e_{ia} = E_i + \varphi \times ST_i \qquad \forall i \in I, a \in \{1, \cdots, |U_i|\}, \varphi \in \{0, 1, \cdots, |U_i| - 1\} \tag{G.8}$$

$$l_{ia} = E_i + \varphi \times ST_i \qquad \forall i \in I, a \in \{1, \cdots, |U_i| - 1\}, \varphi \in \{1, 2, \cdots, |U_i| - 1\} \tag{G.9}$$

Note that Equation (G.9) only applicable for the derivation for the sub-interval from 1 to $|U_i| - 1$. For the last sub-interval where $a = |U_i|$, the derivation of l_{ia} should apply the following

equation such that the ending time of the last sub-interval will not exceed the latest time of parking space i:

$$l_{ia} = L_i \qquad \forall i \in I, a = |U_i| \tag{G.10}$$

G.5 Derivation for t_{ima}

The derivation for t_{ima} is provided in the main body of the manuscript for easier understanding and queries, please see Lemma 6.1 (Equations 6.25, 6.27, and 6.29).

G.6 Derivation for q_{ia}

The derivation for q_{ia} is provided in the main body of the manuscript for easier understanding and queries, please see Equations (6.1) to (6.4).

Appendix H

Decoding Algorithm for State

The input of the decoding algorithm is a state s obtained from the encoder. The output is an interpretable task set of each MPL m. The mechanism of the decoding procedure is to traverse each element of x_1 within a state s, and assign it to the corresponding MPL following the sequence defined by x_2.

Algorithm 6: Decoding algorithm for route generation

Input: State $s = (x_1, x_2)$, Number of MPLs M
Output: Task set $\bar{O}_m$ of MPLs m

$total_dispatch \leftarrow 0$ `// Number of MPLs deployed`
$\bar{O}_m \leftarrow \emptyset$
$x_1 \leftarrow s[0]$ `// Variable of task assignment`
$x_2 \leftarrow s[2]$ `// Variable of task execution sequence`
for m in M **do**
    $\bar{O}_m[m] \leftarrow \emptyset$
    $index_x_1 \leftarrow argwhere(x_1 == m)$ `// Extract index`
    if $index_x_1$ is not $\emptyset$ **then**
        $index_x_2 \leftarrow argwhere(x_2 == index_x_1)$
        $index_x_2 \leftarrow argsort(index_x_2)$ `// Sorting index`
        $task \leftarrow index_x_1[index_x_2]$
        $\bar{O}_m[m] \leftarrow task$ `// Task list of MPL` m `with execution sequence`
        $total_dispatch += 1$
    end if
end for
return $\bar{O}_m$

Appendix I

Performance Comparison between HQM and Gurobi

This appendix provides comparison between HQM and Gurobi in terms of i) reward and computation time, and ii) different cost units.

I.1 Reward and Computation Time for HQM and Gurobi

This section provides the best reward obtained by Gurobi solver and HQM and the corresponding computation time (see Table I.1).

Table I.1: Comparison of average computation time and reward under different problem instances

		Gurobi		HQM		
Number of parking spaces I_i	Customer of each parking space N_n	Computation time (Sec)	Best result[a] ($\times 10^{-3}$)	Computation time (Sec)	Reward[b] ($\times 10^{-3}$)	Gap[c](%)
5	5	896.5	7.247	264.3	6.481	-10.57
	10	1232.4	4.204	384.6	3.899	-7.26

	15	1381.9	3.930	433.5	3.655	-6.99
	20	3497.2	1.096	506.4	1.158	5.66
6	5	898.6	5.774	309	5.149	-10.82
	10	1371.5	3.078	438.9	2.697	-12.38
	15	2491.8	2.162	499.5	2.024	-6.38
	20	3671.6	2.063	580.5	2.227	7.95
7	5	677.1	4.672	367.2	4.281	-8.37
	10	1625.6	0.695	551.7	0.687	-1.15
	15	2884.6	0.383	701.4	0.427	11.49
	20	4965.9	0.408	757.8	0.456	11.76
8	5	948.2	4.062	384.3	3.828	-5.76
	10	2995.3	0.675	562.5	0.717	6.22
	15	4543.7	0.317	734.7	0.354	11.67
	20	5779.9	0.252	951.9	0.264	4.76
9	5	971.5	3.772	411.9	3.478	-7.79
	10	3126.5	0.357	751.2	0.325	-8.96
	15	6079.1	0.639	856.8	0.709	10.96
	20	7200	0.251	978.9	0.274	9.16
10	5	1324.3	2.728	444.6	2.539	-6.93
	10	4796.8	0.339	701.4	0.360	6.19
	15	6233.7	0.240	914.4	0.275	14.58
	20	7200	0.130	1129.8	0.139	6.92
Mean		3199.74	2.061	609.05	1.933	-6.21

[a] To ensure that the results are comparable, the author transpose Best Result = 1 / Equation (6.3); Reward[b] = [R(BTD) + R(HCPS)]/2; Gap[c] = (Final Reward of HQM - Best Result of Gurobi) / Best Result of Gurobi

I.2 Cost Unit Comparisons between Gurobi and HQM

This section focuses on the implementation result comparison of Gurobi and HQM in terms of MPLs deployment size, total travel distance, and service delays (see Table I.2). It should be noticed that the value of different cost units of HQM is the mean value between BTD and HCPS for the corresponding cost unit. The difference between the Gurobi solver and HQM in MPLs deployment is within ± 3 units. Regarding travel distance, HQM achieves a 9.93% improvement compared with that of the Gurobi solver in general. The average service delay of HQM and Gurobi is 9.07 min, 10.49 min, respectively.

Table I.2: MPLs deployed, travel distance, and service delay of Gurobi and HQM

Number of parking spaces I_i	Customer of each parking space N_n	Gurobi			HQM			Gap^a(%)	Gap^b(%)	Gap^c(%)
		MPLs deployed (unit)	Travel distance (km)	Service delay (min)	MPLs deployed (unit)	Travel distance (km)	Service delay (min)			
5	5	8	97.988	0	9	109.519	0	12.50	11.77	0
	10	12	177.869	0	15	183.984	0	20.83	3.44	0
	15	16	174.453	0	16	194.428	0.022	-3.13	11.45	0
	20	16	240.409	0	16	221.891	7.180	0.00	-7.70	-
6	5	9	128.190	0	11	139.608	0	22.22	8.91	0
	10	15	171.281	0.721	16	224.479	0.965	3.33	31.06	33.84
	15	18	254.115	5.684	18	264.484	2.583	-2.78	4.08	-54.56
	20	22	297.026	2.341	22	278.925	0.606	-2.27	-6.09	-74.13
7	5	12	154.041	0	12	173.707	0	0.00	12.77	0
	10	19	244.694	5.076	20	278.558	11.076	2.63	13.84	118.20
	15	25	451.235	20.958	24	341.547	15.759	-4.00	-24.31	-24.81
	20	28	408.578	25.618	26	423.333	13.996	-8.93	3.61	-45.37
8	5	15	171.184	0	15	186.306	0	0.00	8.83	0
	10	20	401.430	15.438	19	309.036	10.524	-5.00	-23.02	-31.83
	15	25	566.556	14.359	23	371.916	21.252	-10.00	-34.35	48.00
	20	33	707.538	20.113	33	456.919	19.589	-1.52	-35.42	-2.61
9	5	14	195.111	0	15	204.009	0.12	7.14	4.56	-
	10	22	348.313	14.293	22	383.603	23.471	0.00	10.13	64.21
	15	33	515.278	20.439	30	421.291	5.613	-9.09	-18.24	-72.54
	20	33	575.891	19.872	31	470.912	19.691	-6.06	-18.23	-0.91
10	5	18	222.569	0	19	252.073	0.606	2.78	13.26	-
	10	26	339.106	31.611	24	402.434	19.333	-9.62	18.67	-38.84
	15	34	519.177	19.243	31	476.686	19.437	-8.82	-8.18	1.01
	20	34	776.938	35.987	33	561.443	25.866	-2.94	-27.74	-28.12

Gap^a=(MPLs deployed of HQM - MPLs deployed of Gurobi)/ MPLs deployed of Gurobi; Gap^b=(Travel distance of HQM - Travel distance of Gurobi)/ Travel distance of Gurobi; Gap^c= (Service delay of HQM – Service delay of Gurobi)/ Service delay of Gurobi

Appendix J

Code Listing of Genetic Algorithm (GA)

The overall implementation process of GA is as follows. It first creates the population, in which the population size equals the number of states within the agent *Agent* in HQM. The population is then ranked based on the fitness function (the value of Equation 6.8). Using an elite strategy and roulette wheel selection, individuals with higher fitness values will be selected to produce the next generation. The next step is to exploit the elite individuals to create a mating pool and breed the population. As a final step, the author applies mutation operation to the new generation to obtain better solutions. Following the notation and parameter setting mentioned in Table 6.3, the code listing of GA is as follows:

Algorithm 7: Decoding algorithm for route generation

Input: Population size Pop; Possibility of crossover P_{cross}; Proportion of elite size P_{elite}; Mutation rate P_{muta}; Number of MPLs $|M|$; All task list O

Output: The optimal individual *individual*

$individual, population, fitness, pop_{index}, pop_{ranked}, cum_{perc}, mating_{pool}, children \leftarrow \emptyset$
for i in range(Pop) **do**
  $x_1 \leftarrow \mathbb{N} \sim U(0, |M|)$, $x_2 \leftarrow$ Permuting task list O
  $individual = (x_1, x_2)$
  Add *individual* to *population*
end for
for *individual* in *population* **do**
  $fitness \leftarrow$ Calculate the fitness value of individual according to Equation (6.8)
  $pop_{index} \leftarrow$ Obtain the index of the sorted fitness score in $fitness$
  $pop_{ranked} = [population[i]$ for i in $pop_{index}]$
end for
$cum_{perc} \leftarrow$ Calculate the cumulative percentage of the fitness value in pop_{ranked}
for i in range(len(pop_{ranked}-$P_{e_{lite}}$, len(pop_{ranked}))) **do**
  Add P_{elite} elite individuals to $mating_{pool}$ `// Select elite individuals within population`
end for
for i in range(0, len(pop_{ranked})-P_{elite}) **do**
  $P_{cross} \leftarrow 100 \times U(0, 1)$ `// Generate random probability`
  for $fitness_{score}$ in cum_{perc} **do**
    if $P_{cross} \leq fitness_{score}$ **then**
      Add the corresponding individual to $mating_{pool}$ `// Create the mating pool`
    end if
  end for
end for
for i in range(0, P_{elite}) **do**
  Add the top P_{elite} elite individual to *children* `// Produce the next generation`
end for
for i in range(0, len($mating_{pool}$)-P_{elite}) **do**
  Exchange some chromosome of the two adjacent individuals `// Regenerate the rest of individual of` $mating_{pool}$
  Obtain the updated population *population′* `// Breed the population`
end for
for *individual* in *population′* **do**
  for *chromosme* in *individual* **do**
    $prob \leftarrow 100 \times U(0, 1)$ `// Generate random probability for mutation operation`
    if $P_{muta} \leq prob$ **then**
      Exchange the current value of the chromosome with the adjacent chromosome
      Obtain the updated population *population′*
    end if
  end for
end for
Recalculate and rank the fitness value of the new population *population″*
Output the best individual $individual^*$ with the highest fitness score if convergence condition satisfies
return $individual^*$

Appendix K

Performance Comparisons between HQM and GA

This appendix focuses on the comparison between HQM and GA in terms of i) deployment size and the total travel distance; ii) parcel first-time fulfilment; iii) reward improvement; and iv) reward and computation time

K.1 Comparison of MPL Deployed and Travel Distance

This section shows the implementation result comparison of HQM and GA in terms of MPLs deployment size and the total travel distance, as illustrated in Table K.1. The comparison is based on two perspectives: i) HQM vs. GA, and 2) BTD vs. HCPS.

Table K.1: MPLs deployed and travel distance of HQM and GA

Parking spaces I_i	Customers per parking space N_n	Route adjustment policy	HQM		GA		Gap[a]	
			MPLs deployed (unit)	Travel distance (km)	MPLs deployed (unit)	Travel distance (km)	MPLs deployed (unit)	Travel distance (km)
5	5	BTD	9	115.196	12	107.922	-3	7.274
		HCPS	9	103.842	15	125.609	-6	-21.767
		Gap[b]	0	11.354	-3	-17.687	-	-
	10	BTD	13	188.929	16	168.477	-3	20.452
		HCPS	16	179.039	16	175.058	0	3.981
		Gap[b]	-3	9.89	0	-6.581	-	-
	15	BTD	15	203.999	16	178.342	-1	25.657
		HCPS	16	184.857	14	162.803	2	22.054
		Gap[b]	-1	19.142	2	15.539	-	-
	20	BTD	17	230.168	16	199.032	1	31.136
		HCPS	15	213.614	15	216.142	0	-2.528
		Gap[b]	2	16.554	1	-17.11	-	-
6	5	BTD	11	148.131	11	135.441	0	12.69
		HCPS	11	131.085	11	127.910	0	3.175
		Gap[b]	0	17.046	0	7.531	-	-
	10	BTD	15	214.336	17	205.476	-2	8.86

		HCPS	16	234.621	14	196.877	2	37.744
		Gap[b]	-1	-20.285	3	8.599	-	-
		BTD	18	268.043	17	227.272	1	40.771
	15	HCPS	17	260.924	17	214.830	0	46.094
		Gap[b]	1	7.119	0	12.442	-	-
		BTD	23	274.511	21	277.209	2	-2.698
	20	HCPS	20	283.338	21	261.223	-1	22.115
		Gap[b]	3	-8.827	0	15.986	-	-
		BTD	12	179.269	12	157.566	0	21.703
	5	HCPS	12	168.145	15	170.794	-3	-2.649
		Gap[b]	0	11.124	-3	-13.228	-	-
		BTD	20	271.938	19	236.719	1	35.219
	10	HCPS	19	285.177	18	224.509	1	60.668
7		Gap[b]	1	-13.239	1	12.21	-	-
		BTD	23	345.110	23	281.598	0	63.512
	15	HCPS	25	337.984	23	279.518	2	58.466
		Gap[b]	-2	7.126	0	2.08	-	-
		BTD	23	371.527	25	346.943	-2	24.584
	20	HCPS	28	475.139	24	373.877	4	101.262
		Gap[b]	-5	-103.612	1	-26.934	-	-
		BTD	15	188.846	16	163.958	-1	24.888
	5	HCPS	15	183.766	15	174.367	0	9.399

		Gap[b]	0	5.08	1	-10.409	-	-
		BTD	19	315.499	19	248.248	0	67.251
	10	HCPS	19	302.573	19	251.514	0	51.059
		Gap[b]	0	12.926	0	-3.266	-	-
		BTD	22	379.610	24	361.005	-2	18.605
	15	HCPS	23	364.222	23	333.354	0	30.868
		Gap[b]	-1	15.388	1	27.651	-	-
		BTD	32	459.328	30	414.339	2	44.989
	20	HCPS	33	454.509	31	406.337	2	48.172
		Gap[b]	-1	4.819	-1	8.002	-	-
		BTD	15	202.043	15	192.197	0	9.846
	5	HCPS	15	205.974	16	194.847	-1	11.127
		Gap[b]	0	-3.931	-1	-2.65	-	-
		BTD	22	378.835	22	360.514	0	18.321
	10	HCPS	22	388.371	21	329.978	1	58.393
9		Gap[b]	0	-9.536	1	30.536	-	-
		BTD	29	432.421	30	411.984	-1	20.437
	15	HCPS	31	410.161	28	387.925	3	22.236
		Gap[b]	-2	22.26	2	24.059	-	-
		BTD	32	480.981	32	450.828	0	30.153
	20	HCPS	30	460.842	30	419.316	0	41.526
		Gap[b]	2	20.139	2	31.512	-	-

		BTD	18	254.706	16	204.258	2	50.448
	5	HCPS	19	249.439	16	217.141	3	32.298
		Gap^b	-1	5.267	0	-12.883	-	-
		BTD	24	420.902	24	345.720	0	75.182
	10	HCPS	23	383.965	23	335.836	0	48.129
10		Gap^b	1	36.937	1	9.884	-	-
		BTD	31	491.675	29	435.426	2	56.249
	15	HCPS	31	461.697	30	446.390	1	15.307
		Gap^b	0	29.978	-1	-10.964	-	-
		BTD	34	580.228	31	507.270	3	72.958
	20	HCPS	32	542.657	32	515.494	0	27.163
		Gap^b	2	37.571	-1	-8.224	-	-

Gap^a= HQM - GA; Gap^b = BTD - HCPS

K.2 Parcel First-time Fulfilment of HQM and GA

This section presents the parcel first-time fulfilment of HQM and GA, as illustrated in Table K.2. The average first-time delivery fulfilment of HQM ranges from 15.5 units to 19.25 units (the maximum capacity is 20 units per MPL), corresponding to the fulfilment rate from 77.50% to 96.30% respectively. For GA, the first-time delivery fulfilment ranges from 12.75 to 17 units, corresponding to the fulfilment rate from 63.75% to 85.00% respectively.

Table K.2: Parcel First-time Fulfilment of HQM and GA

		HQM		GA	
Parking spaces I_i	Customers per parking space N_n	First-time fulfilment-BTD (units)	First-time fulfilment-HCPS (units)	First-time fulfilment-BTD (units)	First-time fulfilment-HCPS (units)
5	5	13	7	7	6
	10	11	15	9	15
	15	18	20	18	18
	20	18	20	17	17
Mean value		15	15.5	12.75	14
6	5	11	12	10	9
	10	17	19	13	14
	15	19	20	18	18
	20	20	16	19	19
Mean value		16.75	16.75	15	15
7	5	11	10	9	9
	10	18	15	14	16
	15	19	14	19	20
	20	18	20	18	18
Mean value		16.5	14.75	15	15.75
8	5	16	17	10	12
	10	14	17	18	18
	15	20	20	19	18
	20	19	20	17	20
Mean value		14.75	18.5	16	17

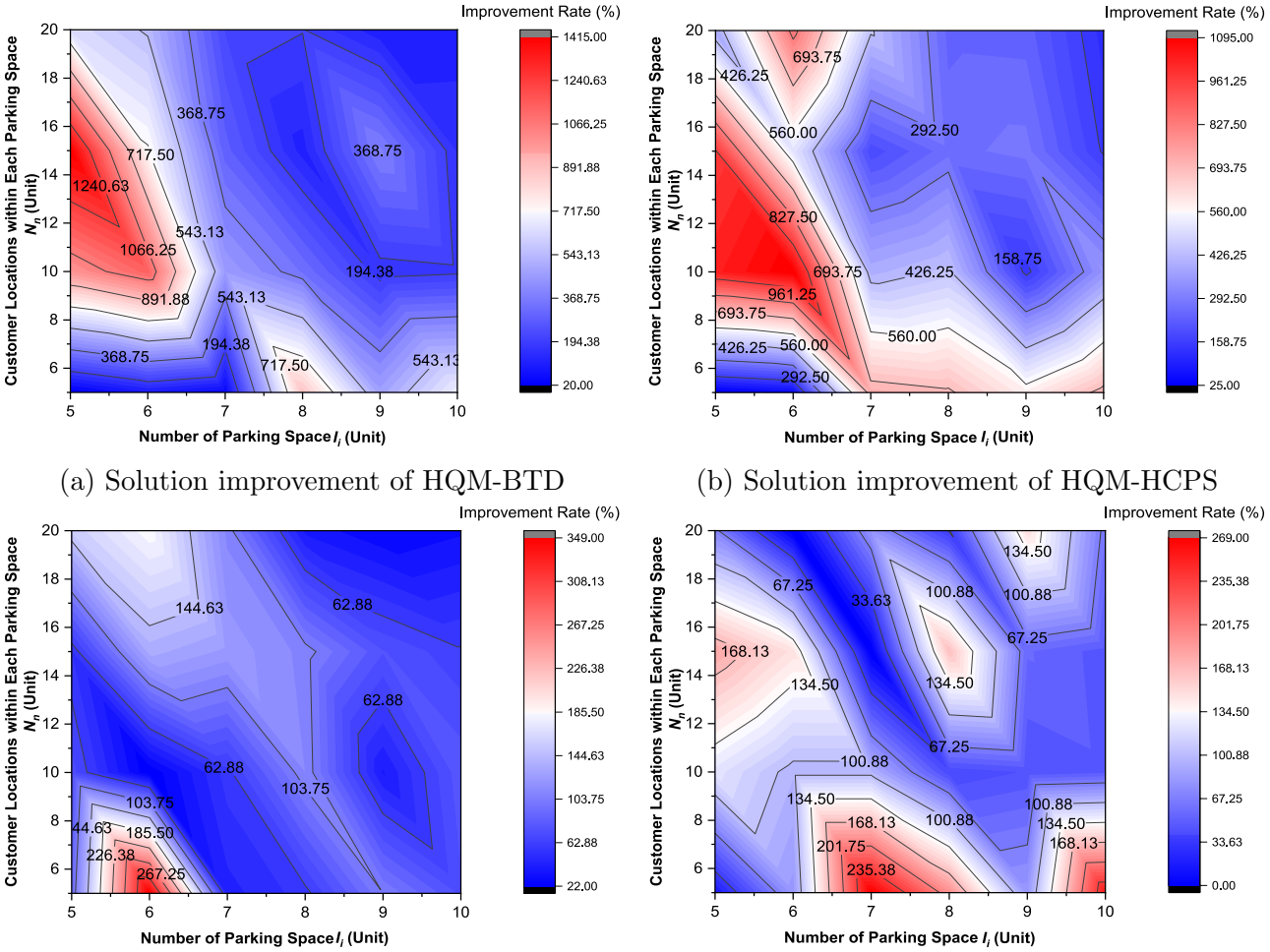
9	5	10	11	9	9
	10	18	20	16	18
	15	20	20	20	15
	20	20	20	20	20
Mean value		17	17.75	16.25	15.5
10	5	20	14	14	15
	10	20	20	15	17
	15	19	20	19	16
	20	18	20	20	20
Mean value		19.25	18.5	17	17

K.3 Reward Improvement of HQM and GA

This section illustrates the reward improvement of HQM and GA, compared with the initial reward (see Figure K.1). Specifically, HQM-HCPS achieves an average increase of 453.58% in final rewards over initial rewards (Figure K.1b), while HQM-BTD achieves an average increase of 433.23% (Figure K.1a). By contrast, GA-HCPS achieves an average improvement of 97.31% in final rewards compared to its initial rewards (Figure K.1d), while GA-BTD achieves an average improvement of 92.50% (Figure K.1c). The optimisation ability of HQM is on average 348.49% greater than that of GA, suggesting that HQM facilitates an efficient search for better solutions and overcomes local optima.

K.4 Reward and Computation Time Comparisons between HQM and GA

This section provides the best reward obtained by Gurobi solver and HQM and the corresponding computation time, as illustrated in Table K.3.



(a) Solution improvement of HQM-BTD

(b) Solution improvement of HQM-HCPS

(c) Solution improvement of GA-BTD

(d) Solution improvement of GA-HCPS

Figure K.1: Improvement rate of HQM and GA

Table K.3: Comparison of average computation time and reward under different problem instances

Number of parking spaces I_i	Customer of each parking space N_n	HQM		GA		Gap^c(%)
		Computation time (Sec)	Reward^a ($\times 10^{-3}$)	Computation time (Sec)	Reward^b ($\times 10^{-3}$)	
5	5	264.3	6.481	23.7	5.721	11.73
	10	384.6	3.899	48.3	1.325	66.02
	15	433.5	3.655	26.7	0.710	80.57
	20	506.4	1.158	30	0.386	66.67
6	5	309	5.149	42	4.520	12.22
	10	438.9	2.697	46.2	0.568	78.94
	15	499.5	2.024	57.9	0.681	66.35

	20	580.5	2.227	48	0.698	68.66
7	5	367.2	4.281	23.4	3.804	11.14
	10	551.7	0.687	46.8	0.252	63.32
	15	701.4	0.427	44.7	0.195	54.33
	20	757.8	0.456	85.8	0.247	45.83
8	5	384.3	3.828	32.4	1.127	70.56
	10	562.5	0.717	48	0.289	59.69
	15	734.7	0.354	61.5	0.249	29.66
	20	951.9	0.264	109.8	0.144	45.45
9	5	411.9	3.478	29.7	1.010	70.96
	10	751.2	0.325	60.9	0.147	54.77
	15	856.8	0.709	126.3	0.258	63.61
	20	978.9	0.274	108	0.174	36.50
10	5	444.6	2.539	47.4	0.746	70.62
	10	701.4	0.360	48.6	0.161	55.28
	15	914.4	0.275	157.8	0.160	41.82
	20	1129.8	0.139	117	0.097	30.22

Reward^a = 1 / Equation (6.8); Reward^b = Equation (6.8); Gap^c = (Reward^a - Reward^b) / Average Final Reward of HQM

Appendix L

Proof for Proposition and Corollary

L.1 Proof for Lemma 7.1

For $s \in S$ and $c \in C$, consider $s \in S$ that satisfies $V(i,c) \leq R(i,c)$ and $V(c,i) \leq R(c,i)$. Assuming that $y_{ij}^s = 1$ for some $i, j \in C$ holds. It is necessary to prove $V(i,j) - W_{ij}^s \leq 0$ such that $max\{V(i,j) - W_{ij}^s\} = 0$. Let $(p_1^{isj}, p_2^{isj}, \cdots, p_{|s|}^{isj})$ is the permutation of customers in set s that achieves the shortest path, with respect to travelling time of ADR, when loading at i and returning to j. According to the definition of the AADP regarding ADR service times in Section 7.1, we got:

$$W_{ij}^s = R(i, p_1^{isj}) + R(p_1^{isj}, p_2^{isj}) + \cdots + R(p_{|s|-1}^{isj}, p_{|s|}^{isj}) + R(p_{|s|-1}^{isj}, j) \tag{L.1}$$

$$\geq V(i, p_1^{isj}) + V(p_1^{isj}, p_2^{isj}) + \cdots + V(p_{|s|-1}^{isj}, p_{|s|}^{isj}) + V(p_{|s|-1}^{isj}, j) \tag{L.2}$$

$$\geq V(i,j) \tag{L.3}$$

Formula (L.3) holds due to the assumption that $V(i,c) \leq R(i,c)$ and $V(c,i) \leq R(c,i)$ for $i \in s$ and $c \in C$. $f + V(p_r^{isj}, p_{r+1}^{isj}) \leq R(p_r^{isj}, p_{r+1}^{isj})$ holds due to the triangle inequality on the vehicle travelling time. Therefore, $V(i,j) - W_{ij}^s \leq 0$.

L.2 Proof for Proposition 7.1

Assuming that $y_{ij}^s = 1$ and customer permutation $(p_1^{isj}, p_2^{isj}, \ldots, p_{|s|}^{isj})$ achieves the shortest path for ADR travels from customer i to customer j in set s. Since $(p_1^{isj}, p_2^{isj}, \ldots, p_{|s|}^{isj})$ do not exist a feasible route in ADR graph

$G_r(\bar{C}, E_r)$, $f + V(p_r^{isj}, p_{r+1}^{isj}) \leq R(p_r^{isj}, p_{r+1}^{isj})(r \in \{1, \ldots, |s|-1\})$ holds. If $y_{ij}^s = 1$, the value of objective function is:

$$W_{ij}^s + max\{V(i,j) - W_{ij}^s, 0\} + f \tag{L.4}$$

$$= W_{ij}^s + f \tag{L.5}$$

$$= R(i, p_1^{isj}) + \cdots + R(p_r^{isj}, p_{r+1}^{isj}) + \cdots + R(p_{|s|}^{isj}, j) + f \tag{L.6}$$

$$\geq R(i, p_1^{isj}) + \cdots + f + V(p_r^{isj}, p_{r+1}^{isj}) + \cdots + V(p_{|s|}^{isj}, j) + f \tag{L.7}$$

$$\geq W_{ip_r^{isj}}^{s'} + f + V(p_r^{isj}, p_{r+1}^{isj}) + W_{p_{r+1}^{isj} s'' j} + f \tag{L.8}$$

$$= W_{ip_r^{isj}}^{s'} + f + max\{V(i, p_r^{isj}) - W_{ip_r^{isj}}^{s'}, 0\} \tag{L.9}$$

$$+ V(p_r^{isj}, p_{r+1}^{isj}) + W_{p_{r+1}^{isj} s'' j} + f \tag{L.10}$$

$$+ max\{V(p_{r+1}^{isj}, j) - W_{p_{r+1}^{isj} j}^{s''}, 0\} \tag{L.11}$$

where $s' = \{p_1^{isj}, \ldots, p_r^{isj}\}$, $s'' = \{p_{r+1}^{isj}, \ldots, p_{|s|}^{isj}\}$. The above process has determined that the solution derived by set s is worse than that derived by s' and s''. Hence, set s will not contribute the optimal solution, which conflicts with $y_{ij}^s = 1$.

L.3 Proof for Corollary 7.1

If an isolated customer i in ADR graph $G_r(\bar{C}, E_r)$ (i.e., no edges from/to customer i in $G_r(\bar{C}, E_r)$), ADR graph $G_r(\bar{C}, E_r)$ does not exist a feasible path such that $|s| > 1$ for $s \in O_i$. Hence, $y_{ij}^s = 0$ for $i, j \in C$ holds based on proposition 7.1. Therefore, $s^* = \{i\}$ is the only set in O_i. Since $y_{lj}^{s^*} = 1$ should hold based on constraint (7.7), we need to prove $l = i = j$. By using refutation, we assume $l \neq i$. For $f \geq 0$, $V(a,i) \leq V(a,i) + f \leq R(a,i)$, $V(i,a) \leq V(i,a) + f \leq R(i,a)$ for $a \in C$ hold. Since $y_{lj}^{s^*} = 1$, the value of objective function is:

$$W_{lj}^{s^*} + max\{V(l,j) - W_{lj}^{s^*}, 0\} = W_{lj}^{s^*} + f \tag{L.12}$$

$$= R(l,i) + R(i,j) + f \geq V(l,i) + R(i,j) + f \tag{L.13}$$

$$= V(l,i) + W_{lj}^{s^*} + f \tag{L.14}$$

$$= V(l,i) + W_{ij}^{s^*} + f + max\{V(i,j) - W_{ij}^{s^*}, 0\} \tag{L.15}$$

The above process has determined that the solution derived by set s^* which loads parcel in customer l is worse than that ADR is onboard the vehicle at customer l to i and loads parcels at i to serve set s^*. Hence, the optimal solution exists when $l = i$. Analogously, we can prove the optimal solution exists when $i = j$ using the above logic.

L.4 Proof for Proposition 7.2 and 7.3

We use refutation to prove proposition 7.2. We assume y_{ij}^s for $i, j \in C$ and $i \notin s$. $R(i, p_1^{isj}) - V(i, p_1^{isj}) \geq 0$ because $max\{V(p_1^{isj}, j) - W^s_{p_1^{isj}j}, 0\} \geq 0$. We consider two situations which includes (i) $max\{V(p_1^{isj}, j) - W^s_{p_1^{isj}j}, 0\} = 0$ and (ii) $max\{V(p_1^{isj}, j) - W^s_{p_1^{isj}j}, 0\} > 0$. For situation (i), the value of objective function is:

$$W_{ij}^s + f + max\{V(i,j) - W_{ij}^s, 0\} \geq W_{ij}^s + f \tag{L.16}$$

$$= f + R(i, p_1^{isj}) + R(p_1^{isj}, p_2^{isj}) + \cdots + R(p_{|s|}^{isj}, j) \tag{L.17}$$

$$\geq f + V(i, p_1^{isj}) + R(p_1^{isj}, p_2^{isj}) + \cdots + R(p_{|s|}^{isj}, j) \tag{L.18}$$

$$\geq f + V(i, p_1^{isj}) + W^s_{p_1^{isj}j} \tag{L.19}$$

$$= f + V(i, p_1^{isj}) + W^s_{p_1^{isj}j} + max\{V(p_1^{isj}, j) - W^s_{p_1^{isj}j}, 0\} \tag{L.20}$$

For situation (ii), since $0 < max\{V(p_1^{isj}, j) - W^s_{p_1^{isj}j}, 0\} = V(p_1^{isj}, j) - W^s_{p_1^{isj}j}$, Equation (7.26) can be conversed as:

$$R(i, p_1^{isj}) \geq V(i, p_1^{isj}) + V(p_1^{isj}, j) - W^s_{p_1^{isj}j} \tag{L.21}$$

Because $y_{ij}^s = 1$, the value of the objective function is:

$$W_{ij}^s + f + max\{V(i,j) - W_{ij}^s, 0\} \geq f + W_{ij}^s \tag{L.22}$$

$$= f + R(i, p_1^{isj}) + R(p_1^{isj}, p_2^{isj}) + \cdots + R(p_{|s|}^{isj}, j) \tag{L.23}$$

$$\geq f + V(i, p_1^{isj}) + R(p_1^{isj}, p_2^{isj}) + \cdots + R(p_{|s|}^{isj}, j) \tag{L.24}$$

$$+ V(p_1^{isj}, j) - W^s_{p_1^{isj}j} \tag{L.25}$$

$$= f + V(i, p_1^{isj}) + R(p_1^{isj}, p_2^{isj}) + \cdots + R(p_{|s|}^{isj}, j) + max\{V(p_1^{isj}, j) - W^s_{p_1^{isj}j}, 0\} \tag{L.26}$$

$$\geq f + V(i, p_1^{isj}) + W^s_{p_1^{isj}j} + max\{V(p_1^{isj}, j) - W^s_{p_1^{isj}j}, 0\} \tag{L.27}$$

Since both of the two situations derive the same result, we infer that there exist a solution that ADR is onboard the vehicle from customer i to p_1^{isj}, then loads parcel to serve set s and rendezvous vehicle at j is bounded by the optimal value of objective function. Therefore, Proposition 7.2 is proven. Analogously, we can prove Proposition 7.3 using the above logic.

L.5 Proof for Lemma 7.2

We make the contrapositive: $y_{ii}^{s} = 0$ if $|s| > 1$. Considering the way of contradiction where $y_{ii}^{s} = 1$ for some $i \in s$. Let $(p_1^{isj}, p_2^{isj}, \cdots, p_{|s|}^{isj})$ is the permutation of customers in set s that achieves the shortest path, with respect to travelling time of ADR, when loading at i and returning to j. Because $y_{ii}^{s} = 1$ while Constraint (7.9) forces $\sum_{j \in \bar{C}} x_{ij} \geq 1$, j' exists such that $x_{ij'} = 1$. Considering $y_{ii}^{s} = 1$ and $x_{ij'} = 1$, the contribution for the objective function is:

$$W_{ii}^{s} + f + max\{V(i,j) - W_{ii}^{s}, 0\} + V(i,j') \tag{L.28}$$

$$= W_{ii}^{s} + f + V(i,j') \tag{L.29}$$

$$= R(i, p_1^{isj}) + \cdots + R(p_{|s|-1}^{isj}, p_{|s|}^{isj}) + R(p_{|s|}^{isj}, i) + f + V(i,j') \tag{L.30}$$

$$\geq R(i, p_1^{isj}) + \cdots + R(p_{|s|-1}^{isj}, p_{|s|}^{isj}) + V(p_{|s|}^{isj}, i) + f + V(i,j') \tag{L.31}$$

$$\geq R(i, p_1^{isj}) + \cdots + R(p_{|s|-1}^{isj}, p_{|s|}^{isj}) + f + V(p_{|s|}^{isj}, j') \tag{L.32}$$

$$\geq W_{ip_{|s|}^{isj}}^{s} + f + V(p_{|s|}^{isj}, j') \tag{L.33}$$

$$= W_{ip_{|s|}^{isj}}^{s} + f + max\{V(i, p_{|s|}^{isj}) - W_{ip_{|s|}^{isj}}^{s}, 0\} + V(p_{|s|}^{isj}, j') \tag{L.34}$$

Formula (L.30) is derived by the conclusion of Lemma 7.1. Formula (L.31) extents the service time based on the definitions in Section 7.1. Formula (L.32) utilises the assumption that $V(i,j) \leq R(i,j)(i,j \in C)$. Formula (L.33) holds because the vehicle driving time satisfies the triangle inequality. Applying the definitions of service time in Section 7.1, Formula (L.35) can be concluded based on Equation (L.34) and the fact that there is no waiting time. It is concluded that the result that includes set s by loading at and back to customer i is bounded below by a scheme that loads at customer i, and returns to customer $p_{|s|}^{isj} \in s$, which makes $y_{ii}^{s} = 0$. This is contradicting to the assumption.

L.6 Proof for Lemma 7.3

Assuming that for $c, k \in C \backslash \{i\}, c \neq k$, $x_{ci} = 1$ and x_{ik} hold. Because $x_{ci} = 1$, Constraint (7.8) forces $\sum_{s \in S} \sum_{a \in C} y_{ia}^{s} \geq 1$. Hence, there exists $y_{ia'}^{s'}$ for $s' \in S$ and $a' \in C$. $i \in s'$ holds. Additionally, because $x_{ik} = 1$, Constraint (7.9) forces $\sum_{a \in C} \sum_{s \in S} y_{ai}^{s} \geq 1$. Hence, there exists $y_{a''i}^{s''} = 1$ for $s'' \in S$ and $a'' \in C$. Also, $i \in s''$. Because Constraint (7.7) forces $\sum_{a \in C} \sum_{s\ in J_i} \sum_{b \in C} y_{ab}^{s} = 1$, tuple $(i, s', a') = (a'', s'', i)$ exists since $y_{ia'}^{s'} = y_{a''i}^{s''} = 1$. Hence, $a' = i$, $a'' = i$, and $s' = s''$ hold. Therefore, $y_{ii}^{s'} = 1$ holds. Based on the conclusion of Lemma 7.2, $s' = \{i\}$ holds and customer i will be fulfilled alone.

L.7 Proof for Proposition 7.4

For $i, j \in C$ and $i \neq j$, we define $\bar{O} = \{(a, s, b) \mid a, b \in C, s \in O_i\}$. Additionally, $O \subset \bar{O}$ holds. If $x_{ij} = 0$, the following equation holds based on definition and Constraint (7.7):

$$\sum_{(a,s,b)\in O} y_{ab}^{s} \leq \sum_{(a,s,b)\in \bar{O}} y_{ab}^{s} = 1 \tag{L.35}$$

We need to prove $y_{ab}^{s} = 0$ for $(a, s, b) \in O$ if $x_{ij} = 1$. $\sum_{s\in S} \sum_{l\in C} y_{jl}^{s} \geq 1$ holds based on Constraint (7.8). Hence, there exists $s' \in S, l^{'} \in C$ such that $y_{jl'}^{s^{'}} = 1$, $j \in s'$. Since Constraint (7.9) requires $\sum_{l\in C} \sum_{s\in S} y_{li}^{s} \geq 1$, there exists $l'' \in C, s'' \in S$ such that $y_{l''i}^{s^{''}} = 1, i \in s''$. Hence, we need to prove $s' \neq s''$. We use refutation and assume that $s' = s'' = s^*$. Therefore, $i, j \in s'$. Based on Constraint (7.7), customer i will only be fulfilled in one set. Hence, $l^{''} = j$ and $l^{'} = i$. We have $y_{ij}^{s*} = 1$. There exists $l \in C$ such that $x_{li} = 1$ based on Constraint (7.8). Because $x_{ij} = 1$ and $x_{li} = 1$, customer i will be fulfilled alone. However, $i, j \in s^*$ and $y_{ij}^{s*} = 1$ hold given a contradiction. Therefore, $s' \neq s''$ and customer i and j will be served in different sets. Hence, $y_{ab}^{s} = 0$ for $(a, s, b) \in O$ holds and Proposition (7.4) has been proven.

L.8 Proof for Proposition 7.5

For $i, j \in C, i \neq j$, since $x_{ij} + x_{ji} \leq 1$ holds. Hence, we consider two situations: (i) $x_{ij} + x_{ji} = 0$ and (ii) $x_{ij} + x_{ji} = 1$. Based on Constraint (7.7), we derive the following inequality that makes Equation (7.29) holds:

$$\sum_{s\in S}(y_{ji}^{s} + y_{ij}^{s}) = \sum_{s\in O_i}(y_{ji}^{s} + y_{ij}^{s}) \leq \sum_{a\in C}\sum_{s\in O_i}\sum_{b\in C} y_{ab}^{s} = 1 \tag{L.36}$$

Assuming that $x_{ij} + x_{ji} = 1$, we need to prove $y_{ij}^{s} = 0$ and $y_{ji}^{s} = 0$ hold for $s \in S$ such that Equation (7.29) holds. We assume that there exists $s \in S$ that makes $y_{ij}^{s} = 1$ satisfied. Since $\sum_{l\in \bar{C}} x_{li} \geq 1$ should be held based on Constraint (7.8), there must exists $l' \in \bar{C}$ such that $x_{l'i} = 1$. Since $\sum_{l\in \bar{C}} x_{il} \geq 1$ should be held based on Constraint (7.9), there must exists $l'' \in \bar{C}$ such that $x_{il''} = 1 (l' \neq l'')$. Because $x_{l'i} = 1$ and $x_{il''} = 1$, customer i will be fulfilled alone. However, considering that $y_{ij}^{s} = 1$ and therefore $i, j \in S$ should hold, which conflicts with i should be served alone. Therefore, we conclude that $y_{ij}^{s} = 0$ holds for $s \in S$. Analogously, we can prove that $y_{ji}^{s} = 0$ holds for $s \in S$.

Appendix M

Implementation Results of the Service Set & Decision Variable Simplifications

This appendix provides implementation results of the service set reduction and variable simplification techniques introduced in Section 7.3.3 and 7.3.4.

Table M.1: Implementation Results of the Solution Techniques Proposed in Section 7.3.3 and 7.3.4

Code	**Degree**	$\lvert E_r^* \rvert$	$\lvert S^* \rvert$	**Imp of** S (%)	t_{p1} (sec)	$\lvert y_{ii}^s \rvert$	t_{c1} (sec)	$\lvert y_{ij}^{s'} \rvert$	t_{p2} (sec)	$\lvert y_{ij}^{s''} \rvert$	t_{p3} (sec)	**Imp of** y_{ij}^s (%)
Maximum	50	2,500	20,875	-	-	-	-	52,187,500	-	-	-	-
1	3.27	366.8	265.6	98.73	1.20	2	0.01	51,554,170	3.33	50,921,891.8	3.30	2.43
2	4.71	240.4	455	97.82	1.20	0.6	0.01	51,107,240	6.59	50,034,324	6.63	4.13
3	4.24	216.4	400.8	98.08	1.25	1.4	0.01	51,235,210	5.69	50,213,081.8	5.62	3.65
4	3.98	202.6	377.2	98.19	1.21	4.2	0.01	51,291,010	5.27	50,394,670.4	5.38	3.44
5	4.37	222.8	431.6	97.93	1.19	2	0.01	51,162,670	6.22	50,138,970.4	6.17	3.93
6	4.20	214.4	394.2	98.11	1.23	1.8	0.01	51,250,770	5.60	50,317,202.2	5.65	3.58
7	5.77	293.6	699.6	96.65	1.24	1.2	0.01	50,531,090	11.24	48,876,883.2	11.05	6.34
8	5.45	277.4	1337.6	93.59	1.19	1.6	0.01	49,047,913.2	23.90	45,916,297.8	23.86	12.02
9	8.57	435	1470.8	92.95	1.44	0.2	0.01	48,715,190	31.44	45,243,657.2	31.73	18.48

1) Imp of $S = (\lvert S \rvert - \lvert S^* \rvert)/\lvert S \rvert \times 100\%$; 2) $t_{p1}, t_{p2}, t_{p3}, t_{c1}$ = Computation time of Proposition 7.1, 7.2, 7.3, and Corollary 7.1; 3) Imp of $y_{ij}^s = (\lvert y_{ij}^s \rvert - \lvert y_{ij}^{s''} \rvert)/\lvert y_{ij}^s \rvert \times 100\%$