

ABSTRACT

Title of Dissertation: ADJUSTMENT FACTORS AND
 APPLICATIONS FOR ANALYTIC
 APPROXIMATIONS OF TOUR LENGTHS

Youngmin Choi, Doctor of Philosophy, 2021

Dissertation directed by: Dr. Paul M. Schonfeld, Professor
 Department of Civil and Environmental
 Engineering

The shortest tour distance for visiting all points exactly once and returning to the origin is computed by solving the well-known Traveling Salesman Problem (TSP). Due to the large computational effort needed for optimizing TSP tours, researchers have developed approximations that relate the average length of TSP tours to the number of points n visited per tour. The most widely used approximation formula has a square root form: \sqrt{n} multiplied by a coefficient β . Although the existing models can effectively approximate the distance for conventional vehicles with large capacities (e.g., delivery trucks) where n is large, approximations that seek to cover large ranges of n , possibly to infinity, tend to yield poorer results for small n values. This dissertation focuses on approximation models for small n values, which are needed for many practical applications, such as for some recent delivery alternatives (e.g., drones). The proposed models show promise in analyzing the real-world problems in which actual

tours serve few customers due to limited vehicle capacity and incorporate realistic constraints, such as the effects of a starting point location, geographical restrictions on movements, demand patterns, and service area shapes. The dissertation may open new research avenues for analyzing the new transportation alternatives and provide guidelines to planners for choosing appropriate models in designing or evaluating transportation problems.

Approximation models are estimated from the following experiments: 1) a total of 60 cases are developed by considering various factors, such as point distributions and shapes of service areas. 2) Solution methods for TSP instances are compared and chosen. 3) After the TSPs are optimized for each n , the TSP tour lengths are averaged. 4) Lastly, models for the averaged TSP tour lengths are fitted with ordinary least squares (OLS) regression.

After the approximations are developed, some possible extensions are explored. First, adjustment factors are designed to integrate the 60 cases within one equation. With those factors, it can be understood how approximation varies with each classification. Next, the approximations considering stochastic customer presence (i.e., probabilistic TSP) are proposed. Third, the approximated tour lengths are compared with the optimal solutions of vehicle routing problem (VRP) in actual rural and urban delivery networks. Here, some additional factors, such as a circuitry factor and service zone shape, are discussed.

Lastly, the proposed methodology is applied to formulate and explore various types of existing and hypothetical delivery alternatives.

ADJUSTMENT FACTORS AND APPLICATIONS
FOR ANALYTIC APPROXIMATIONS OF TOUR LENGTHS

by

Youngmin Choi

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2021

Advisory Committee:

Professor Paul M. Schonfeld, Chair

Professor Ali Haghani

Professor Cinzia Cirillo

Professor Bruce L. Golden

Assistant Professor Myungseob (Edward) Kim

© Copyright by
Youngmin Choi
2021

Acknowledgements

I have received a great deal of support from the following people throughout my journey to Ph.D. I want to thank my esteemed advisor, Dr. Paul Schonfeld, for his academic guidance, encouragement, personality, and patience over the past six years. Any conversation with him keeps me moving and makes me feel incredibly fortunate. I truly appreciate all his effort and warm care.

I am also very grateful to my advisory committee members, Drs. Ali Haghani, Cinzia Cirillo, Bruce Golden, and Edward Kim, for giving me valuable comments and food to thought on my research. I owe special thanks to Dr. Golden, with whom multiple discussions have been challenging and very fruitful.

My gratitude goes to the Appalachian Regional Commission (ARC) and its staff for their support during my studies. Many thanks to Jason Wang at ARC for his generosity and willingness to help me. I cannot thank Jason enough for his kindness.

Finally, I want to express my deepest gratitude to my family for their support and love. Without my family's support, encouragement, and sacrifice, I would not have a chance to complete my Ph.D. journey.

Table of Contents

Chapter 1: Introduction	1
1.1 Background and Motivation	1
1.2 Research Objectives and Scope	3
1.3 Dissertation Overview and Contributions.....	5
Chapter 2: Literature Review.....	8
2.1 Overview of Average Tour Length Approximation	8
2.1.1 Approximations for the TSP Tour Lengths	8
2.1.2 Approximations for TSP Variants and VRP Tour Lengths	11
2.1.3 Special Considerations in Tour Length Approximations.....	13
2.1.4 Guidelines for Using Distance Approximations	15
2.2 Experimental Approach	16
2.2.1 Experiment Procedures: Point Generations, Heuristics, and Sample Size	16
2.2.2 Summary of Literature with Experiments.....	20
2.3 Modeling Deliveries by Small Vehicles	23
2.3.1 Existing Delivery Alternatives.....	23
2.3.2 Hypothetical Delivery Alternatives	24
2.4 Summary	28
Chapter 3: Methodology	31
3.1 Solution Method.....	31
3.1.1 Formulation of Traveling Salesman Problem (TSP)	31
3.1.2 Genetic Algorithm (GA).....	32
3.1.3 Parameter Selection for GA.....	32
3.1.4 Concorde TSP Solver and Comparison of Solution Methods	34
3.2 Simulation Settings	36
3.2.1 Scenario Design	36
3.2.2 Simulation Design: Point Generation, Point Distribution, and Least Squares Method	38
3.3 Results.....	39
3.3.1 Descriptive Statistics of the Optimized TSP Instances.....	39
3.3.2 Standard Deviations of Average TSP Tour Lengths.....	44
3.3.3 Estimated Coefficients β	47
3.4 Summary	54
Chapter 4: Extensions of Tour Length Approximation: Adjustment Factors, Probabilistic Tour Length Approximation, and Comparison on Approximated versus Actual Road Network Distance	55
4.1 Adjustment Factors for Approximations	55
4.1.1 Curve Fitting Methods and Computation Steps.....	56
4.1.2 Validation of Adjustment Factors.....	59
4.1.3 Comparison with Other Tour Length Approximations.....	60
4.2 Tour Length Approximation with Stochastic Customer Presence: Probabilistic Traveling Salesman Problem	62
4.2.1 Simulation Design and Result.....	63
4.2.2 Curve Fitting Result and Validation	65

4.3 Comparison of Approximated Distance versus Actual Road Network Distance	67
4.3.1 Case Study for Rural Area	67
4.3.2 Case Study for Urban Area	77
4.4 Summary	80
Chapter 5: A Comparison of Optimized Deliveries by Robots, Drones, and Trucks	81
5.1 Problem Statement	81
5.2 Methodology	81
5.2.1 Baseline Numerical Values	81
5.2.2 Model Assumptions	84
5.2.3 Model Formulations	85
5.3 Numerical Results and Sensitivity Analyses	86
5.3.1 Numerical Results	86
5.3.2 Comparison the Suggested Model to Other Research work	88
5.3.3 Sensitivity Analyses	88
5.4 Summary	91
Chapter 6: Optimization Approaches for Investigating Various Drone Delivery Alternatives	93
6.1 Problem Statement	93
6.2 Alternative Descriptions	94
6.2.1 Conventional Truck Delivery (CT)	95
6.2.2 Drone Delivery Supported by Truck (DT)	96
6.2.3 One-to-one Delivery by Drone (OD)	97
6.2.4 One-to-many Delivery by Drone (MD)	97
6.3 Methodology	97
6.3.1 Baseline Numerical Values	97
6.3.2 Model Assumptions and Formulations	102
6.3.2.2 Model Assumptions	105
6.3.2.3 Model Formulations	106
6.4 Numerical Results and Sensitivity Analysis	108
6.4.1 Numerical Results	108
6.4.2 Sensitivity Analysis	109
6.5 Summary	113
Chapter 7: Innovative Methods for Delivering Fresh Food to Underserved Populations	116
7.1 Problem Statement	116
7.2 Alternative Descriptions	117
7.2.1 Truck Deliveries	118
7.2.2 E-bikes Deliveries	118
7.2.3 Third-party Delivery by Personal Car (TPC) Deliveries	119
7.2.4 Personalized Ride (PR) Deliveries	119
7.2.5 Parcel Locker Deliveries	119
7.3 Methodology	119
7.3.1 Assumptions for Delivery System	119
7.3.2 Baseline Numerical Values	120
7.3.3 Cost Function	124

7.3.3.3 <i>Third-party Personal Car (TPC) Deliveries Formulation</i>	125
7.3.3.4 <i>Personalized Ride (PR) Formulation</i>	125
7.3.3.5 <i>Parcel Locker Deliveries Formulation</i>	126
7.3.3.6 <i>System Constraints</i>	127
7.3.3.7 <i>Optimization</i>	128
7.4 Numerical Results and Sensitivity Analyses	128
7.4.1 Results.....	128
7.4.2 Sensitivity Analyses.....	130
7.5 Discussion and Summary.....	136
7.5.1 Discussion.....	136
7.5.2 Summary	137
Chapter 8: Conclusions and Future Research	140
8.1 Research Summary and Contributions.....	140
8.2 Future Research	142
References.....	144

List of Tables

<i>Table 1 Summary of Literature with Beardwood’s Formula</i>	9
<i>Table 2 Comparison of Heuristic Algorithms</i>	18
<i>Table 3 Summary of Studies with Experiments for TSP/VRP Tour Approximation</i> ...	20
<i>Table 4 Parameter Section for Genetic Algorithm (GA)</i>	33
<i>Table 5 Optimized TSP Solutions from Heuristic/Solver</i>	34
<i>Table 6 Estimated Average TSP Tour Lengths from Heuristics</i>	35
<i>Table 7 Descriptive Statistics and Normality Test (Case 1)</i>	40
<i>Table 8 Summary of Estimated Coefficients β</i>	49
<i>Table 9 Estimated Coefficients β with Length-to-Width Ratio of 1</i>	50
<i>Table 10 Estimated Coefficients β with Length-to-Width Ratio of 2</i>	51
<i>Table 11 Estimated Coefficients β with Length-to-Width Ratio of 4</i>	52
<i>Table 12 Comparison of Exact and Estimated Tour Lengths</i>	53
<i>Table 13 Curve Fitting for Adjustment Factor Associated n Values (D_0)</i>	58
<i>Table 14 Curve Fitting Results for Equations (10) – (14)</i>	59
<i>Table 15 Percent Adjustment Error (1/2)</i>	59
<i>Table 16 Percent Adjustment Error (2/2)</i>	60
<i>Table 17 Comparison of Percent Differences Using Existing Approximation Models</i>	61
<i>Table 18 Estimation Results for Few/Large n Values Using Adjustment Factors</i>	62
<i>Table 19 Estimated Tour Lengths with Different Range of Uniform Distribution</i>	64
<i>Table 20 Average TSP Tour Lengths for Various Probabilities</i>	64
<i>Table 21 Estimators for Equation (16)</i>	65
<i>Table 22 Percent Errors Using P-TSP Results</i>	66
<i>Table 23 Summary of Service area</i>	70
<i>Table 24 Results for Baseline Scenario</i>	72
<i>Table 25 Results for Alternative Scenario</i>	73
<i>Table 26 Comparison between Approximated and Optimized Tour Lengths (Rural)</i>	74
<i>Table 27 Comparison between Approximated and Optimized Tour Lengths (Rural)</i>	76
<i>Table 28 Optimized Results for Delivery Routes</i>	78
<i>Table 29 Comparison between Approximated and Optimized Tour Lengths</i>	80
<i>Table 30 Variable Definitions and Baseline Values</i>	82
<i>Table 31 Optimization Results for Delivery Alternatives</i>	87
<i>Table 32 Comparison of Results Based on Different Coefficients β</i>	88
<i>Table 33 Results for Sensitivity Analysis</i>	90
<i>Table 34. Variable Definitions and Baseline Values</i>	98
<i>Table 35. Results of Each Delivery Strategy</i>	108
<i>Table 36. Elasticity to Input Parameters</i>	110
<i>Table 37 Variable Definitions and Baseline Values</i>	121
<i>Table 38 Optimization Results of Alternatives</i>	128
<i>Table 39 Results of Combined Delivery Based on Scenarios</i>	134
<i>Table 40 Results of Delivery Service for Washington Village/Pigtown</i>	137

List of Figures

<i>Figure 1 Shape of Areas Developed by Chien (1992)</i>	12
<i>Figure 2 Overall Process for Estimating Beardwood's Coefficient β</i>	16
<i>Figure 3 Sample Variance of the optimized TSP Tour Lengths (Ong and Huang, 1989)</i>	21
<i>Figure 4 Illustration of Point Distributions ($n = 1,000$)</i>	37
<i>Figure 5 Classifications for Distance Approximation</i>	38
<i>Figure 6 PDFs for the TSP Instances for Different n Values</i>	42
<i>Figure 7 Simulation Results</i>	43
<i>Figure 8 Standard Deviations of the TSP Instances</i>	46
<i>Figure 9. Investigation for Estimated β</i>	53
<i>Figure 10 Curve Fitting for Adjustment Factors (D_0)</i>	58
<i>Figure 11 Example of ASD's Delivery Operation</i>	67
<i>Figure 12 Illustration of Service Areas</i>	69
<i>Figure 13 Delivery Scenarios</i>	71
<i>Figure 14 Illustration of Optimized Delivery Routes</i>	73
<i>Figure 15 Some Tours Violating Assumptions in Approximation</i>	75
<i>Figure 16 Optimized Routes for Hypothetical Delivery Points</i>	76
<i>Figure 17 Illustrations of Physical Addresses</i>	77
<i>Figure 18 Illustrations of Optimized Truck Routes</i>	79
<i>Figure 19 Delivery Options Serving Study Area</i>	85
<i>Figure 20 Effects of Inputs on Total Costs</i>	90
<i>Figure 21 Delivery Alternatives</i>	95
<i>Figure 22. Delivery Area A in Response to Consolidation Time h</i>	100
<i>Figure 23. Effects of Large Drones for MD alternative</i>	112
<i>Figure 24. Effects of Location of Distribution Depot</i>	113
<i>Figure 25 Delivery Alternatives Serving Study Area</i>	118
<i>Figure 26 System Constraints on Cost Function</i>	127
<i>Figure 27 System Outputs for Changes in Service Area Z</i>	131
<i>Figure 28 System Outputs for Changes in User Value of Time Spent for Waiting v_u</i>	132
<i>Figure 29 System Outputs for Changes in Demand Density Q</i>	133

Chapter 1: Introduction

1.1 Background and Motivation

The shortest tour distance for visiting all n points exactly once and returning to the origin is computed by solving the well-known Traveling Salesman Problem (TSP) (Applegate et al. 2006). This problem belongs to the class of NP-hard problems in which finding the optimal path requires computation time that increases exponentially with the number of points n (Ansari et al. 2018). Due to this large computational effort, researchers have developed approximations for the relation between the average length of TSP tours and n values. These approximations provide useful estimates to operators who seek to reduce costs and improve system efficiency in large-scale problems or some complex transportation systems. Thus, the approximations have been studied for various transportation planning and system design applications, such as for public transportation services, facility location, and service fleet sizing.

For long-term system planning and design problems in logistics or public transportation, planners and service providers can estimate tour lengths and evaluate routing scenarios before actual demand locations are known (i.e., the exact locations, numbers, and distributions of demand points). The approximations can help in the development of general planning models for large and complex systems, e.g., for optimizing characteristics such as zone sizes and locations, vehicle and fleet characteristics, service quality standards and facility locations. Based on the results of such system planning models, resources can be efficiently allocated.

For short-term planning problems, the approximations can effectively reveal relations among vehicle operating variables, instead of applying the TSP algorithm every time the variables change. Since demands (e.g., package delivery service) vary over days, weeks, and seasons, operators maintain their vehicle fleet based on the peak demand. With a simple

relation among headway, delivery area size, and demand density per hour for the service area, the approximation models can account for the following decision variables: frequency, delivery area size for each vehicle, and the required number of vehicles based on real-time operations. Similarly, each vehicle's optimal loading capacity (e.g., small or large trucks) or delivery area partitioned from the entire service region can be obtained with the approximation. This feature helps to adapt vehicle operation responsively to daily/hourly demands, such as by subdividing large areas into time-varying and possibly overlapping zones served by TSP tours.

The most widely used approximation formula has a square root form: \sqrt{n} multiplied by coefficient β . The existing models can effectively approximate the tour length for vehicles with large capacities (e.g., trucks) where n is large. However, approximations that seek to cover large ranges of n , possibly to infinity, tend to yield poorer results for small n values since the coefficients β decrease as n increases. Therefore, this dissertation focuses on developing approximation models for small n values to analyze the real-world problems in which actual tours visit relatively few points. Note that the “small” n values could be subjective depending on intended applications:

- Flexible-route passenger services (e.g., carpool, dial-a-ride, and airport shuttle)
- Deliveries of large items (e.g., large household appliances)
- Tours by service and repair workers

The approximations proposed in this dissertation may open new research avenues for analyzing recent transportation options, such as deliveries by robotic vehicles with small capacity. Deliveries by robotic vehicles and drones have gained traction in e-commerce due to their potential for reducing labor costs and endeavors to support social-distancing efforts during the pandemic since late 2019. Each shipment by robots and drones costs about \$1.40 and \$0.76, respectively, while the cost per delivery by humans is estimated at \$2.50 (Cuthbertson, 2016;

[Korman, 2019](#)). For these growing needs of autonomous last-mile delivery, the global market size is expected to grow from \$12.0 billion in 2019 to \$91.5 billion by 2030 ([Bloomberg, 2019](#)). Companies, including Amazon, Google, and JD.com, have demonstrated improvements in deliveries by drones and robots. Amazon has shown a few prototype delivery drones since 2013 and announced that its drones could fly up to 30 minutes while carrying a 5 lb (2.23kg) package. Google's Wing drones have completed 3,000 deliveries over an 18-month trial in suburban areas of the U.S and Australia ([Bass et al. 2019](#)). DHL Express ([DHL, 2019](#)) may decrease cost per delivery by up to 80% in urban areas with drones covering a radius of 8-km distance (i.e., a round trip of 16 km). JD.com has developed seven types of delivery drones since 2015 and tested them in rural settings across China and Indonesia, accumulating over five thousand flight hours. The company has been experimenting with autonomous ground robots serving urban populations. Similarly, delivery robots of Starship Technologies can carry items within a 4-mile (or 6-km) radius while cruising at four mph. Besides these efforts by private firms, the Federal Aviation Administration allows UPS and Wing (Google's project) to deliver packages using drones in the U.S as of 2019. Therefore, UPS Flight Forward announces that its drone delivery started in May 2020 for providing prescription medicines in Florida ([UPS Pressroom, 2020](#)).

1.2 Research Objectives and Scope

The overall objective of this dissertation is to develop practical approximations of TSP tour distances for visiting points while considering realistic situations (e.g., salesman's loading capacity and operating conditions). The methodology has the following features:

- Refines the distance approximation models developed by Beardwood et al. ([1959](#)) by focusing on tours with relatively few points.

- Incorporates realistic operating conditions in the methodology, such as the effects of a starting point location, demand patterns, and various service area shapes (e.g., elongation and shape).
- Develops adjustment factors that incorporate the considerations listed above and change the approximation coefficient accordingly.
- Provides guidelines to planners or researchers for choosing appropriate models in designing or evaluating transportation problems.

In seeking to achieve the above features, this dissertation pursues several research goals listed below:

1. Developing a modeling framework that generates random points visited, optimizes TSP tours, and eventually derives the tour length approximation models through statistical estimation
2. Comparing the accuracy of solution methods (i.e., metaheuristics) for optimizing TSP tour instances
3. Identifying the real-world factors which may violate ideal conditions and assumptions for the tour length approximation, such as specific point distributions, elongated service regions, and shapes of regions
4. Comparing model outputs and actual tour distance over real networks in urban and rural areas
5. Providing adjustment factors to conveniently use the approximation methods, considering the abovementioned operating characteristics
6. Applying the proposed methodology to analyze and compare the optimized freight transportation systems, including both existing and hypothetical delivery alternatives

The proposed tour length approximations are designed for a small number of visited points n where the range for n lies between 2 and 100. Typical ranges considered in the literature for n have wider ranges than in this dissertation, i.e., 5 to 100,000 points for n . The difference in accuracy between the two approximations will be explored later.

1.3 Dissertation Overview and Contributions

The organization of this proposed dissertation is as follows. The principal contributions of this dissertation are underlined.

Chapter 2 introduces a comprehensive review of existing studies in 1) approximation methods for the Traveling Salesman Problem (TSP), 2) experiment settings for obtaining the TSP tour length approximation, and 3) planning models that analyze delivery systems by small vehicles (i.e., drones, robots, vans, or bikes). The literature focuses on an overview of the approximation methods and considerations that incorporate real-world constraints. Experiment settings are discussed, including the point generation, solution methods, sample size, and ordinary least squares (OLS) regression analysis. In particular, a total of fourteen metaheuristics and TSP solvers are compared in terms of solution accuracy. Delivery alternatives for existing and hypothetical delivery modes are analyzed with the proposed models. The gaps in the current knowledge and further possible improvements in approximation models are identified from the review.

Chapter 3 develops the TSP approximations through few points. The simulation settings and various factors are introduced for developing the tour length approximation models. This chapter presents the assumptions and evaluation criteria. Then, a solution procedure based on metaheuristics and Concord TSP solver is discussed. The optimized TSP instances are investigated using statistical analysis.

Chapter 4 explores some possible extensions of the TSP tour length approximation. First, adjustment factors are developed for more accurate and convenient use of the model. The factors are designed to integrate six considerations into a single equation. Next, approximations considering stochastic customer presence are developed. Lastly, the approximated tour lengths are compared with actual tour distances using data in urban and rural areas. After urban or rural data are mapped in a GIS platform, data processing (e.g., circuitry factor) and optimized routes by a VRP solver are discussed.

Chapter 5 compares the applicability of various types of autonomous delivery systems. Models are applied to formulate cost functions for deliveries by ground robots, drones, and conventional trucks. The cost function of each alternative is optimized and compared with total costs. Sensitivity analyses are designed to explore how system outputs of such delivery systems vary with changes in baseline inputs.

Chapter 6 discusses the proposed models for analyzing hypothetical delivery alternatives with limited vehicle loading capacity. This chapter identifies the applicability of the drone delivery system in terms of the total cost. In particular, a drone can lift multiple packages within its maximum payload and serve recipients in a service area of a given radius. Battery capacities, the primary energy sources for drone operation, are incorporated as a constraint of the planning model to relate parcel payloads and flight ranges.

Chapter 7 focuses on a last-mile fresh food delivery system for individuals in underserved communities with food deserts. To build self-sustainable and cost-effective alternative in delivering fresh items, a total of five delivery alternatives are proposed and optimized based on total cost.

Chapter 8 summarizes the tasks completed in this dissertation and suggests potential topics for future research.

Therefore, the main contributions are summarized as follows: 1) Beardwood's approximations are refined by incorporating various relevant factors. 2) The exponent for the number of points n is statistically estimated, unlike in the existing studies which assumed that tour lengths should vary with the square root of n . These improvements help estimate accurate TSP tour lengths and solve large system planning and design problems, even when the exact locations of stopping points are not yet known.

Chapter 2: Literature Review

The dissertation selectively reviewed 22 papers approximating TSP tour lengths with low n values and 7 papers chosen for their experimental approaches or solution methods. Excluded from the approximation studies are 1) those dealing solely with many points (e.g., $n > 100$) and 2) those which applied rather than developed approximation methods. The study includes a few research publications that consider large n values if they are pioneering in some way or worth mentioning for their experiments. For experimental approaches, the dissertation focuses comparing solution methods.

2.1 Overview of Average Tour Length Approximation

2.1.1 Approximations for the TSP Tour Lengths

The average distance between two points in both Euclidean and rectilinear space can be mathematically derived (Larson and Odoni, 1981, Phillip, 2007, and Burgstaller et al. 2009). Here, the Euclidean space allows vehicle movements in straight lines between any pair of points, while rectilinear space refers to movements which are restricted to two orthogonal coordinates. Although average TSP distance with three points can still be analytically computed, estimating the tour lengths becomes challenging as the number of points n increases.

In early studies for distance approximation models, Mahalanobis (1940) suggested that average TSP tour lengths for visiting a set of points n in a region served by a single vehicle asymptotically converged to \sqrt{n} with large n , where the points n were scattered at random within the space. Later, Marks (1948) mathematically proved the approximation by providing a lower bound for the expected value of the distance as follows:

$$\text{Average TSP Tour Length } (L) \cong \sqrt{\frac{A}{2} \frac{n-1}{\sqrt{n}}} \quad (1)$$

where A is the zone size.

With a large n , the coefficient β of Equation (1) found by Marks (1948) was roughly 0.7071. Beardwood et al. (1959) later estimated the coefficient β to be 0.749 for \sqrt{nA} (Beardwood's formula) in Euclidean space and numerical experiments by constructing tour instances. After Stein (1977) estimated β at 0.765 through Monte Carlo experiments, many researchers estimated the coefficients using different algorithms. For instance, Ong and Huang (1989) reported that β converged to 0.7425 with normalized TSP tour lengths.

Table 1 Summary of Literature with Beardwood's Formula

<i>Authors</i>	<i>Solution Method</i>	<i>Estimated Coefficient*</i>	<i>Problem Type</i>	<i>Number of Points n</i>	<i>Special Considerations</i>
<i>Marks (1948)</i>	Theoretical Derivation	0.7071	TSP	N/A	N/A
<i>Beardwood et al. (1959)</i>	Theoretical Derivation	0.749	TSP	N/A	N/A
<i>Christofides and Eilon (1969)</i>	N/A	N/A	VRP	10 - 70	N/A
<i>Stein (1977)</i>	Partition Heuristic	0.765	TSP	N/A	N/A
<i>Daganzo (1984)</i>	Theoretical Derivation	0.9	TSP	N/A	Shape of a space *****
<i>Ong and Huang (1989)</i>	3-optimal Heuristic	0.7425	TSP	5 - N/A	N/A
<i>Brunetti et al. (1991)</i>	Cavity Method	0.7251	TSP	50 - 800	N/A
<i>Chien (1992)</i>	Exact Solution	0.88**	TSP	5 - 30	Shape of a space
<i>Fiechter (1994)</i>	Parallel Tabu Search	0.7298	TSP	500 - 100,000	N/A
<i>Lee and Choi (1994)</i>	Multicanonical Annealing	0.7239 ~ 0.8075	TSP	50 - 40,000	N/A
<i>Kwon et al. (1995)</i>	Exact Solution	-**	TSP	10 - 80	Shape of a space
<i>Percus and Martin (1996)</i>	Chained local optimization	0.7120 ± 0.0002	TSP	12 - 100	N/A
<i>Johnson et al. (1996)</i>	Iterated Lin-Kernighan	0.7124 ± 0.0002	TSP	100 - 100,000	N/A
<i>Finch (2003)</i>	N/A	0.75983 ~ 0.98398	TSP	N/A	N/A
<i>Hindle and Worthington (2004)</i>	Cheapest Insertion	-***	TSP	5 - 50	Point distribution
<i>Robusté et al. (2004)</i>	Three Heuristic Algorithms****	-***	TSP, VRP	15 - 139	Shape of a space
<i>Figliozzi (2008)</i>	Monte Carlo Simulation	-***	VRP	N/A	Point distribution, depot location
<i>Applegate et al. (2011)</i>	Cutting-plane method	0.7241373~ 0.7764689	TSP	100 - 2,500	N/A
<i>Cavdar and Sokol (2015)</i>	Exact Solution	-***	TSP	N/A	Point distribution,

<i>Mei (2015)</i>	Cutting-plane method	-***	TSP, VRP	N/A	Shape of a space Point distribution
<i>Lei et al. (2016)</i>	The Concorde TSP Solver	0.7773827~ 0.8584265	TSP	20 - 90	N/A
<i>Nicola et al. (2019)</i>	Pilot Method	-***	TSP, VRP	25 – 1,000	Time window, demands
<i>Madani et al. (2020)</i>	Simulated Annealing	-***	TSP	2 - 15	Shape of a space

* the estimates β in the Euclidean space were listed

** Salesman's origin (e.g., a depot) was positioned at a fixed location

*** The studies considered other decision variables or other terms from Beardwood's formula, such as the spatial distribution and variance of points

**** Clarke and Wright, Fisher and Jaikumar, and Gillet and Miller algorithm

***** The shape of space is the shape of the region in which points are generated (e.g., circular, triangular, or sectorial) to be connected by a tour

Fiechter (1994) estimated the coefficient β at 0.7298 for large values of n ranging from 500 to 100,000. Lee and Choi (1994) showed β to be 0.721, while Percus and Martin (1996) estimated β to be 0.7120 ± 0.0002 in Euclidean space. Johnson et al. (1996) generated large set of points with n up to 100,000 and found the coefficient β to be 0.7124 within the 95% confidence intervals of ± 0.0002 . Note that the estimated β is correlated with the value of n (Franceschetti et al. 2017). Applegate et al. (2011) estimated the coefficient β by running a regression on the optimized TSP solution instances for randomly generated n ranging from 100 to 2000. Lei et al. (2015) used a similar approach to Applegate et al. (2011) where n ranged between 20 and 90. With the two studies combined, the estimated β asymptotically approached an interval ranging from 0.7256264 to 0.8584265 and had a downward trend as n increased, as shown in Table 1. Another loose bound was found between 0.75983 and 0.98398 (Finch, 2003; Arlotto and Steele, 2016). Although most of the coefficients cluster around 0.7, a few studies showed outlying values exceeding 0.8 for the following reasons:

- Experiment settings (e.g., coefficients derived from worst-case TSP tour lengths (Finch, 2003))
- Shapes of area (e.g., elongated (Daganzo, 1984) and sectorial shaped area (Chien, 1992))

This is further explained below.

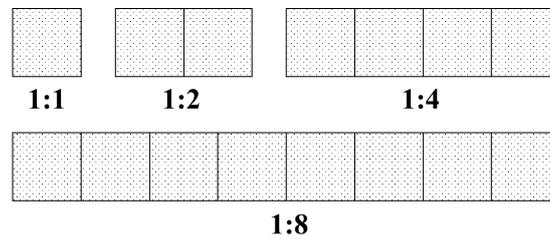
2.1.2 Approximations for TSP Variants and VRP Tour Lengths

For the TSP variants and Vehicle Routing Problem (VRP), many researchers have attempted to estimate the coefficient β through analytical and experimental studies for different operational settings, such as vehicle capacity, zone shape, geometry, or point distributions. The key difference between the TSP and VRP is whether the problem considers vehicle loading capacities, time constraints, or time windows (Kumar and Panneerselvam, 2012). The TSP solution would have a single route served by one vehicle, while the VRP has multiple routes possibly served by multiple vehicles. As such, the number of vehicles should be known a priori for VRP problems. Alternatively, the single TSP route can be split into several equal tours with an optimistic assumption that a penalty in terms of extra travel distance does not exist (Odoni and Larson, 1981).

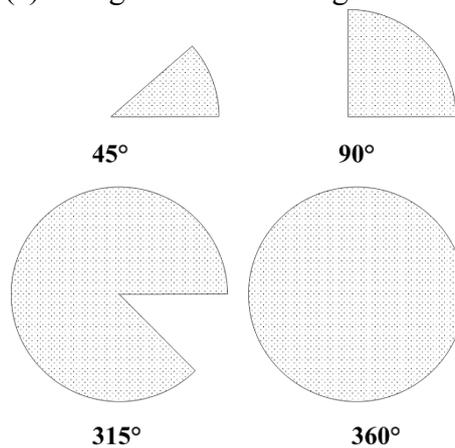
Christofides and Eilon (1969) first incorporated a vehicle capacity per tour in the formula and suggested approximations to the VRP tour length based on the shape and area of a region. Daganzo (1984) proposed an intuitive approximation for a generic irregular service zone, which divided into multiple subareas containing clusters of points. A vehicle route was developed to serve each cluster. In this setting, he estimated β at 0.9 for Euclidean and 1.15 for rectilinear space. Although β for the Euclidean might overpredict the tour distance, it suited spaces with typical shapes.

Chien (1992) derived the coefficient β at 0.88 through empirical simulations and multiple regressions. The paper considered 16 different shapes varying in the 1) elongation and 2) angle of space. Rectangular areas with different length-to-width ratios from 1 to 8 were proposed in Figure 1 (a). Sectorial-shaped areas were developed with eight central angles from 45° to 360° , as illustrated in Figure 1 (b). The starting point (i.e., a depot) was positioned at the lower left

side of the service area. From generated TSP instances, the best-fitted coefficients for Beardwood's formula were derived through OLS regression.



(a) Elongation for Rectangular Areas



(b) Angle of Sectorial-shaped Areas

Figure 1 Shape of Areas Developed by Chien (1992)

Aside from the widely used form of Beardwood, later studies included various terms in the models, such as a length-to-width ratio or area of the smallest rectangle that covered all points. Kwon et al. (1995) carried out both simulations and OLS regressions to test the previous variations (i.e., Beardwood, Daganzo, and Chien).

To the best of our knowledge, most tour length approximations are based on regression methods since the TSP tour lengths associated with n values are non-linear and can be effectively fitted with the square root form with a reasonably good R^2 . However, Kwon et al. (1995) compared results from the regression with a neural network (NN) model for estimating the TSP tour length; the latter model provided slightly better approximations than the former. The NN model was difficult to interpret geometrically due to its characteristic as a so-called “a black box,” where the model would not give any insights. Hindle and Worthington (2004)

developed an alternative expressions for estimating TSP tour lengths, as listed in Equation (2). The authors approximated the average TSP tour length through simulations and regressions.

$$L \cong a \times n + b \times \ln(n) + c \quad (2)$$

where a , b , and c are constants in a 100 x 100 unit square. $a = 3.63$, $b = 85.78$, and $c = 62.67$.

Another formulation for the approximation was considered by Cavdar and Sokol (2015), as presented in Equation (3). The model will be discussed more in detail in the Sections 2.1.3 and 4.1.3.

$$L \cong 2.791 \sqrt{n(cstdev_x \cdot cstdev_y)} + 0.2669 \sqrt{nA(stdev_x \cdot stdev_y)/(\bar{c}_x \cdot \bar{c}_y)} \quad (3)$$

where $cstdev_x$ and $cstdev_y$ are the standard deviations of x (horizontal) and y (vertical) coordinates from center point, $stdev_x$ and $stdev_y$ are the standard deviations of the x and y coordinates, \bar{c}_x and \bar{c}_y are the average distances of points to the central x and y coordinate, and A is a service area size.

Two models were proposed based on demand patterns, namely uniformly random and probabilistic point distribution. The probabilistic demands were designed to simulate point distributions and settlement patterns.

2.1.3 Special Considerations in Tour Length Approximations

Later studies for TSP approximations, considered zone shape, geometry, or point distributions. An extended version of Daganzo's approximation that considered circular and elliptical spaces was proposed by Robusté et al. (2004). Figliozzi (2008) proposed VRP tour length approximations using six different spatial distributions. His models also considered time windows, demands, and depot location. The study showed that time windows negatively affected the accuracy of the models; the time windows increased travel distance not only

because the number of routes was increased but also because the distance between points per route was increased.

In Equation (3), Cavdar and Sokol (2015) developed approximations by incorporating standard deviations of point coordinates. In this way, their approximations can estimate average TSP tour length without knowing the exact point distribution. In Equation (3), the approximation models consisted of a few variables (e.g., the standard deviations of x and y coordinates from center and of distances between the point and center in a region). The models were tested with different spatial distributions, including uniform and triangular distribution. The models performed well for various shapes of a space, such as a triangular or polygonal service area. However, the average TSP tour lengths are underestimated if $n < 1,000$. The use of approximation can be complicated to for the computation of variables, compared to Beardwood's variants (i.e., \sqrt{n}).

Mei (2015) incorporated spatial distributions in approximating the tour lengths. The average nearest neighbor index was introduced for measuring the dispersion of points; the index utilized the distance between centroid and each point. As the point distribution changed from dense (e.g., clustered) to dispersed, the estimates for β increased linearly. Nicola et al. (2019) proposed approximations based on regression models by adding more variables, such as time windows, vehicle capacities, and demands. The proposed model was compared with the previous models from Cavdar and Sokol (2015) and from Hindle and Worthington (2004). Unlike other studies estimating the coefficient β , Madani et al. (2020) investigated the change of the TSP tour length if an additional point is added to the service area. They further considered service area shapes (i.e., square and rectangle).

2.1.4 Guidelines for Using Distance Approximations

Odoni and Larson (1981) pointed out that Beardwood's equation could provide a good approximation if 1) one of the measurements (e.g., width) of space was not much greater than the other measurement (e.g., length) of a region, 2) points n are distributed randomly and uniformly and 3) no obstructions or boundaries existed in the region. Such conditions for a tour's operating zone were generally called "fairly compact and fairly convex." For rigorous definitions of this rule of thumb, numerous measures for both compactness and convexity had been proposed in the literature. Compactness measures were borrowed from geometric concepts, such as perimeters, areas, centroids, and vertices (Kaufman et al, 2017). Some measures are as follows:

- Length-width ratio: the ratio between the length and width of the minimum bounding rectangle
- Convex hull: the ratio of the area between the space and minimum bounding convex hull (i.e., the smallest convex polygon containing all the given points)
- Polsby-Popper: the ratio of the area of the space to the squared perimeter of the space.

Similarly, convexity measures have been based on the area or boundary of a space (Zunic and Rosin, 2004). A boundary-based convexity measure is computed as the ratio of the perimeter of a space and that of convex hull. An area-based convexity measure computes the normalized average visible area of a space, divided by the area of the space (Stern, 1989, and Rote, 2013). The latter method is slightly more challenging to compute.

Most approximation errors here tend to approach zero as n increases: i.e., asymptotically approaching a certain number. The convergence for TSP tour length approximations can be observed between $n = 20$ and $n = 316,228$ (Applegate et al. 2018, Johnson et al. 1996, and Lei et al. 2015). The estimated coefficients β only decrease with increases in n values. Therefore,

when using the approximations for small n , the users must account for discrepancies (e.g., lower and upper confidence intervals (Percus and Martin, 1996 and Johnson et al. 1996), treatments for violating approximation assumptions, or adjustment factors reflecting point distribution).

2.2 Experimental Approach

2.2.1 Experiment Procedures: Point Generations, Heuristics, and Sample Size

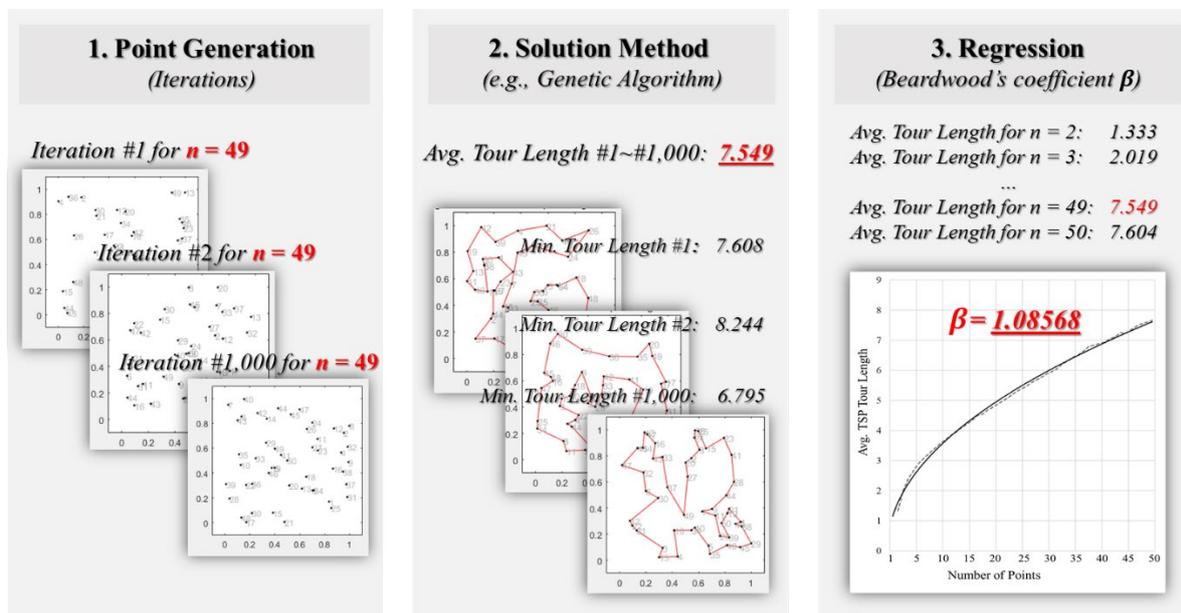


Figure 2 Overall Process for Estimating Beardwood's Coefficient β

Except for the theoretical derivations of Beardwood's coefficients in Table 1, this section shed light on the derivation of the estimates β from experiments. The experimental method is illustrated in Figure 2.

First, n points are generated according to a given distribution (e.g., uniformly and randomly) in a unit space whose area is one. For the point generation, most studies focus on a random and uniform distribution, while the shape of space is limited to a unit square. Random points provided in recent simulation programs are generated with the congruential algorithm, which has been widely used in programming to mimic randomness (Moler, 2008). By generating two

random numbers uniformly distributed in the interval (0,1), the numbers are regarded as a x- and y-coordinate of a point in the space. Each point in the x-y plane with both x and y between 0 and 1 is equally likely to be selected.

Second, a solution method is chosen to compute optimized TSP tour lengths. For every TSP run, the visited points are regenerated after the TSP solution is obtained. From Table 1, no clear preference or explanation is apparent from researchers in choosing the solution method. Furthermore, no consensus exists on the “best” heuristic algorithm for solving the TSP instances as shown in Table 2; ranks imply the shortest TSP solution, while percentage differences show the difference in ratio between the best solution and the solution obtained by the selected heuristic method.

Table 2 Comparison of Heuristic Algorithms

#	Category	Authors	Solution Method													Note
			SA	TS	GA	MA	BCO	ACO	FA	CS	HC	PSO	NN	GH	HS	
10	Rank	Adewole et al (2012)	1	N/A	2	N/A										
	% difference		0.0		1.7											
10	Rank	Gupta (2013)	1	2	7	6	8	3	3	3	N/A					
	% difference		0.0	0.0	-5.6	-5.7	-5.2	-5.7	-5.7	-5.7						
10	Rank	Damghanjiza (2017)	5	N/A	1	N/A	2	N/A	4	3	N/A				Not a random point generation	
	% difference		N/A		N/A		N/A		N/A	N/A						
15	Rank	Adewole et al (2012)	1	N/A	2	N/A										
	% difference		0.0		24.9											
15	Rank	Ansari et al. (2015)	1	N/A				2	N/A							
	% difference		0.0	5.7												
16	Rank	Gupta (2013)	7	8	3	6	5	4	2	1	N/A					
	% difference		-19.4	-19.4	-2.2	-7.3	-4.4	-2.2	-2.2	0.0						
20	Rank	Adewole et al (2012)	1	N/A	2	N/A										
	% difference		0.0		50.2											
20	Rank	Abdulkarim and Alshammari	N/A		1	N/A					2	3	N/A		Elongated Space	
	% difference		0.0	32.8	8.4											
20	Rank	Ansari et al. (2015)	1	N/A				2	N/A							
	% difference		0.0	39.0												
25	Rank	Adewole et al (2012)	1	N/A	2	N/A										
	% difference		0.0		4.4											
25	Rank	Ansari et al. (2015)	1	N/A				2	N/A							
	% difference		0.0	11.9												
25	Rank	Gupta et al (2020)	N/A			3	N/A			1	N/A				2	N/A
	% difference		18.5	0.0	11.2											
29	Rank	Gupta (2013)	6	8	3	4	7	5	2	1	N/A					
	% difference		111.6	159.4	0.1	0.4	128.4	10.3	0.1	0.0						
30	Rank	Gupta (2013)	3	5	3	7	6	8	2	1	N/A					
	% difference		0.5	0.8	0.5	3.8	2.3	7.4	0.5	0.0						
30	Rank	Ansari et al. (2015)	1	N/A				2	N/A							
	% difference		0.0	25.4												
40	Rank	Adewole et al (2012)	1	N/A	2	N/A										
	% difference		0.0		3.3											
42	Rank	Ansari et al. (2015)	1	N/A				2	N/A							
	% difference		0.0	20.3												
50	Rank	Adewole et al (2012)	1	N/A	2	N/A										
	% difference		0.0		32.9											
50	Rank	Ansari et al. (2015)	1	N/A				2	N/A							
	% difference		0.0	23.4												
50	Rank	Gupta et al (2020)	N/A			3	N/A			1	N/A				2	N/A
	% difference		35.0	0.0	21.6											
51	Rank	Gupta (2013)	7	8	4	6	5	2	3	1	N/A					
	% difference		197.6	215.5	4.0	6.9	4.5	0.0	1.6	0.0						
59	Rank	Damghanjiza (2017)	5	N/A	1	N/A	2	N/A	4	3	N/A					
	% difference		N/A		N/A		N/A		N/A	N/A						
60	Rank	Adewole et al (2012)	1	N/A	2	N/A										
	% difference		0.0		24.0											
75	Rank	Gupta et al (2020)	N/A			3	N/A			1	N/A				2	N/A
	% difference		35.0	0.0	39.2											
100	Rank	Abdulkarim and Alshammari	N/A		2	N/A					3	1	N/A			
	% difference		9.3	14.4	0.0											
100	Rank	Gupta et al (2020)	N/A			3	N/A			1	N/A				2	N/A
	% difference		45.7	0.0	36.8											
100	Rank	Crisan et al (2021)	N/A				1	N/A					1			
	% difference		0.0	0.0												

* SA: Simulated Annealing, TS: Tabu Search, GA: Genetic Algorithm, MA: Memetic Algorithm, BCO: Bee Colony Optimization, ACO: Ant Colony Optimization, FA: Firefly, CS: Cuckoo Search, HC: Hill Climbing, PSO: Particle Swarm Optimization, NN: Nearest Neighbor, GH: Greedy Heuristic, HS: Harmony Search, FA: Firefly, and LK: Lin-Kirnighan

For instance, a simulated annealing (SA) algorithm performed better than a genetic algorithm (GA) by 1.7% from Adewole et al. (2012) comparison. This is done mainly because the results sensitively vary with some parameter values of heuristic methods and computation time. In Adewole et al. (2012), a SA procedure for the optimized TSP tour lengths ranging from n of 10 to 60 performed better than a GA. The GA provided a good solution if the time was sufficient, meaning that a large population size was provided. In contrast, Damghanjazi

and Mazidi (2017) showed that the GA performed the best in searching for the TSP solution for 10- and 59-points; the SA and hill climbing method were the worst. More performance comparisons of heuristics were conducted by Gupta. 2013, Ansari et al. 2015, Abdulkarim and Alshammari. 2015, and Gupta. 2020. For a study conducted by Antosiewicz et al. (2013), six well-known metaheuristic algorithms were compared for n values ranging from 20 to 80. The key idea was to find the best solution method when the computation time was restricted (e.g., 100 seconds). The authors presented several criteria for performance (e.g., accuracy, computation time, and standard deviation); however, none of the algorithms outperformed the others for all the suggested criteria. Crisan et al. (2021) examined the quality of the TSP solutions based on a structure of a TSP instance; the instances were classified as semi-structured and unstructured (randomly uniform). Then, the study used a population-based Ant Colony Optimization (ACO) and a local search Lin-Kirnhghan (LK) heuristic for n ranging from 100 to 2,900. At $n = 100$, both heuristics provided the same optimized tour length. In addition to abovementioned metaheuristics, the Concorde TSP solver (Applegate et al. 1998) is currently known as the best-performing TSP solver (Hoos and Stuzle, 2014), and thus widely used for its fast computation and solution accuracy. After an initial solution (and used as an upper bound) is obtained by the chained Lin-Kernighan heuristic, the solver uses a branch-and-bound search for a smaller n or cutting-plane method for a complex large n to narrow the search space. More details on the solver will be discussed in Section 3 of this dissertation.

Third, repeated replications on a given n are produced. After the predefined replications for each n are reached (e.g., 1,000 runs per n values), the TSP tour lengths for each n value are averaged. Then, the repeated runs move the for $n + 1$. Finally, the averaged TSP tour length is fitted with OLS regression to estimate unknown parameter β .

The recommended sample size (i.e., the number of intervals in the 3rd column of Table 3) for running a regression should exceed 23, according to Green (1991). Green compiled a

comprehensive guide for choosing the minimum sample size as a function of the number of independent variables and effect size (e.g., a correlation between two variables); the effect size referred to standardized measures of the size of the mean difference, which generally used in multiple regression analysis. Many metrics could be used for deriving the effect size, such as Cohen’s d (t distribution) or ω (χ^2 distribution). If the effect size was small, a large number of observations were needed. Sample sizes ranged from 23 (large effect size), 53 (medium effect size) and 400 (small effect size). Alternatively, the number of replications N_i is simply derived from the following calculation: $N_i \geq 50 + 8 \cdot N_x$, where N_x is the number of estimates. This guideline for estimating the instance size is simple and easy to use for a parsimonious model.

In brief, this dissertation summarized and compared 15 metaheuristics from the literature as TSP solvers. Although researchers reached no consensus on choosing the best performing algorithm/heuristic for TSP instances, each algorithm has unique features and parameters that may be preferred for a particular research purpose.

2.2.2 Summary of Literature with Experiments

Table 3 summarized the experiment settings for distance approximations from the literature. In Table 3, the number of points n is a range of n considered in estimating the coefficient β . The number of intervals shows how many samples exist within that range (i.e., minimum point to maximum point), while the increment for n is a growth rate from min to max n . Note that irregular means n grows randomly in successive intervals.

Table 3 Summary of Studies with Experiments for TSP/VRP Tour Approximation

<i>Authors</i>	<i>Number of points n</i>	<i>Number of intervals</i>	<i>Increment for n</i>	<i>Replications per n^*</i>	<i>Shape of space</i>	<i>Problem type</i>
<i>Ong and Huang (1989)</i>	5 – N/A	N/A	Irregular	25	Square	TSP
<i>Brunitti et al. (1991)</i>	50 – 800	5	2x**	500 – 20,000	Square	TSP
<i>Fiechter (1994)</i>	500 – 100,000	8	Irregular	10 – 30	Square	TSP
<i>Lee and Choi (1994)</i>	50 – 40,000	14	Irregular	4 – 1,300	Square	TSP
<i>Kwon et al. (1995)</i>	10 – 80	8	10	10	Irregular	TSP

<i>Percus and Martin (1996)</i>	12 – 100	8	Irregular	5 – 20	Square	TSP
<i>Johnson et al. (1996)</i>	100 – 100,000	7	$\sqrt{10}x^{**}$	2 – 2,098	Square	TSP
<i>Hindle and Worthington (2004)</i>	5 – 50	46	1	500	Square	TSP
<i>Applegate et al. (2011)</i>	100 – 2,500	13	Partially Irregular	10,000	Square	TSP
<i>Lei et al. (2016)</i>	20 – 90	8	10	100	Square	TSP
<i>Nicola et al. (2019)</i>	25 – 1,000	N/A	N/A	130 – 400	Square	VRP
<i>Madani et al. (2020)</i>	2 – 15	15	1	100	Square	TSP

* Replications here imply random configurations of point distribution for each n (e.g., Point generation in Figure 2)

** x implies ‘a factor of’

Ong and Huang (1989) used 25 replications for each n value starting from $n = 5$. In their experiments, the sample variable of the optimized TSP tour length was shown to fluctuate, as shown in Figure 3. Although the variance was not discussed in detail for that study, Yang et al. (2020) presented the standard deviations of TSPs to model the travel time reliability (i.e., of tour lengths).

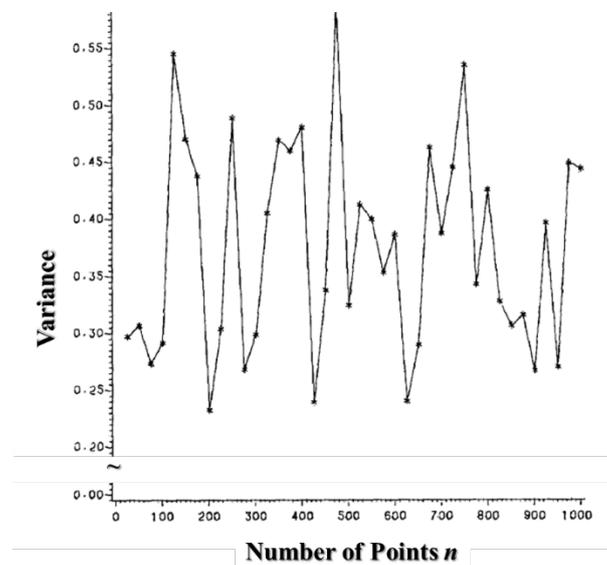


Figure 3 Sample Variance of the optimized TSP Tour Lengths (Ong and Huang, 1989)

Brunetti et al. (1991) found TSP solutions for their selected n values, which were 50, 100, 200, 400, and 800. For each n , replications ranged from 500 to 20,000. Lee and Choi (1994) conducted different replications for the selected 14 intervals of n values, where the values

ranged from 50 to 40,000. As few as four replications were used for large n values (i.e., $n = 40,000$), while 1,300 replications were conducted for small n values (i.e., $n = 50$).

Using the eight intervals of n , Fiechter (1994) ran 10 to 30 replications for each n . Since Kwon et al. (1995) separated training and testing sets for the optimized TSP tour lengths, the number of instances was smaller than in other studies. For Johnson et al. (1996), n ranged from 100 to 100,000 points, increasing by factors of $\sqrt{10}$. The exact TSP tour lengths were obtained for n values between 100 and 316. Then, the number of replications decreased as n increased. Percus and Martin (1996) derived the TSP instances for the eight n values between 12 and 100; replications were conducted between 5 and 12 runs. Unlike other researchers, Hindle and Worthington (2004) and Madani et al. (2020) conducted the replications with the increment of one.

Applegate et al. (2011) ran 10,000 replications for generating the TSP instances visiting each n values. In their experiments, an increment of 100 was chosen for n between 100 and 1,000. Beyond $n = 1000$, the increment of 500 was selected between 1,500 and 2,500 for n values. In Lei et al. (2015) experiments were conducted with 100 replications for each n ranging from 20 to 90. The number of replications for large n increased in Nicola et al. (2019). Since half of the TSP instances were used for test sets, the unused instances were excluded in Table 3. In brief, the number of replications per n was arbitrary. Some researchers have suggested descriptive statistics (e.g., mean or standard deviation) and normality test for the obtained TSP instances (Brunitti et al. 1991; Johnson et al. 1996; Applegate et al. 2011). From this, one can better understand the central tendency and variability of the generated TSP instances. In addition, the instances with few n values can be compared with those for large n values.

2.3 Modeling Deliveries by Small Vehicles

In this section, the vehicles with limited loading capacity are introduced. The existing alternatives include widely available vehicles, including a bike, small van, personal car, or, paratransit (e.g., Lyft or Uber delivery). Two hypothetical modes, namely autonomous ground robotic vehicles (robots) and unmanned aerial vehicles (drones) are discussed.

2.3.1 Existing Delivery Alternatives

Deliveries by bikes could use human-powered or electrically assisted cargo cycles ([Schliwa et al. 2015](#)). An electric cargo bike, called e-bike from here onward, was considered an environment-friendly method for urban parcel deliveries, due to its low emissions, low space requirements in loading zones or curbsides, and relatively low impact on roadway traffic. Sheth et al. ([2019](#)) compared delivery costs for trucks and e-bikes under various operating settings, such as a line-haul distance from a depot, demand density, or delivery volume per stop. They showed that truck delivery was less expensive with a greater line-haul travel or larger volume deliveries per stop. E-bike delivery was cost-effective if the fleet served customers near the depot or covered a dense service area even with low delivery volume per stop. Gruber et al. ([2014](#)) compared the characteristics of e-bikes and passenger cars (or small vans) as package delivery options in urban areas (i.e., third-party delivery by personal car onward). The bikes had a smaller delivery area and tour distance than cars, where the demand for bikes was highly concentrated in inner-city areas. In Berlin, two-thirds of delivery origins and destinations for bikes were located within the inner city, while cars operated extensively throughout the city. Average delivery distance for bikes was 5.1 km versus 11.3 km for cars. Within 10 km, 92% and 56% of deliveries were provided by bikes and automobiles, respectively; the delivery distances for 99% of the bike shipments and 87% of the car shipments were shorter than 20 km. If no constraints were imposed on the weight of deliveries, 42% of the car shipments could

be substituted by bikes with a maximum delivery distance of 10 km. Likewise, a 20-km maximum delivery range could serve 68% of all car shipments and 48% of the resulting mileage. Mean delivery speeds of vehicles turned out to be 15.9 kph and 17.3 kph for bikes and cars, separately, where the bikes could have a speed of up to 25 kph. The team concluded that the e-bikes had great potentials in urban core areas with traffic congestion issues and limited spaces in loading zones.

Service operators might serve demands with temporarily contracted drivers. Such deliveries could be useful in meeting an unexpected surge in demands or be justified when the current demand level was not economically adequate to operate an expensive delivery alternative (e.g., weekend deliveries by trucks). A good example of third-party delivery by passenger car (TPC) was Amazon Flex; Amazon.com has launched Amazon Flex service in 2014 and served more than 50 U.S. cities. The company hired independently contracted drivers. The drivers, with their own cars, usually worked a three-hour time window and delivered an average of 40-50 items within a “small block” of area.

2.3.2 Hypothetical Delivery Alternatives

One promising drone application is parcel delivery, either solely by drones or in collaborative operation with trucks. Recent achievements in the private domain have shown its feasibility ([Kornatowski et al. 2018](#); [UPS Pressroom. 2017](#)). Amazon.com Inc, an electronic commerce company, announced the prototype of its Amazon Prime Air drone in November 2015. The prototype drone could fly up to 15 miles with a maximum speed of 50 miles/hour (= 80 km/hour) and carry packages weighing less than five pounds (= 2.27 kg); about 86 percent of items would be delivered by drones ([Rose. 2013](#)). Joerss et al ([2016](#)) estimated that autonomous drones and robots would deliver 80 percent of all items in the 2020s, while the remaining items would be delivered by conventional transportation. Autonomous delivery

services are expected to be increasingly practical with advanced safety and reliability features, such as automated flight. The drone deliveries are considerably restricted in flight range and parcel payload because most drones are powered by lithium-ion batteries, which currently limit flights to about a half hour (UPS Pressroom. 2017). Due to these key disadvantages, a relatively long tour may be provided by ground transportation (e.g., trucks) while a drone conducts the last-mile delivery. However, Doosan Mobility Innovation Inc., announced in 2019 that a drone's flight time could be extended by over 2 hours with its hydrogen battery, and thus commercial drones delivering multiple items in a single tour could be practical.

Some of the early contributions to delivery-by-drone focus on such delivery supported by trucks (*DT*). The major emphasis was on identifying to what extent resources, such as time, cost or fuel, can be saved with the help of drones. Ferrandez et al (2016) found that *DT* could reduce operating costs. Truck delivery time could be shortened where the speed of drones was 1) about three times faster than truck's or 2) more than two single-package-carrying-drones were assigned to each truck. Wang et al (2016) argued that the maximum delivery completion time could be minimized either by 1) drones which traveled faster than trucks or 2) using more than two drones per truck; the authors found that the delivery time could be reduced by up to 75% with all the above considered. Figliozzi (2017) designed drone deliveries supported by a truck and applied a tour length approximation model to estimate the truck's tour distance. The study also proposed drone energy consumption for level flight at a constant speed. Campbell et al (2017) compared conventional truck delivery (*CT*) and *DT* with operating and delivery stop costs. *DT* offered significant cost savings in suburban areas where demand density was relatively high. The savings were attributed to the fewer tours needed. The authors suggested that assigning multiple drones per truck could reduce operating costs by nearly 40%, depending on the number of drones.

For relatively short delivery ranges, drones are capable of delivering items without truck support. Chowdhury et al (2017) studied a one-to-one delivery by drone (*OD*) for a disaster-relief operation by minimizing the total delivery cost. Increasing the drone flight height could reduce service area and increase the system cost, while increasing drone operating speed could expand the service area and reduce the cost. In addition, unit transportation cost for drones exceeded that for trucks. Some researchers designed services in which delivery drones visited multiple demand points in a single tour (Ham. 2018), while others consider energy storage constraints simultaneously (Rabta et al. 2018; Dorling et al. 2017). Choi and Schonfeld (2018) modeled a delivery service with a one-to-many demand pattern by drones (*MD*) incorporating battery energy storage. They optimized drone fleet size which minimized a total cost function, as well as the costs of operators and users in service area; improvements in battery energy storage could allow drone fleet reductions and increasing drone operating speed could reduce total system cost due to fewer drones and reduced delivery time.

In addition to research on drone deliveries, research on deliveries by robots has been rapidly advancing. Boysen et al. (2018) designed truck-based autonomous robot system (*RT*), where robots launched from trucks. These trucks started from a depot loaded with packages and robots. When the trucks arrived at a customer's location, the robots deliver the single item to the customer's doorstep; the robots essentially conducted "final-mile" deliveries. Next, the study formulated the TSP method for a truck route to establish a launching schedule for the delivery system. The authors explored how the system was affected by 1) the speeds of robots and 2) truck's loading capacity. The speeds varied between 2.8 and 3.7 mph. As speeds increased, the number of delayed deliveries decreased by 75%. Further increases in speeds were less effective due to diminishing returns. In the team's baseline demand density, the optimal number of robots carried by trucks was eight units. Jennings and Figliozzi (2019) investigated the existing regulations in the U.S for delivery robots (e.g., speed, size, or weight

limit) and examined the specifications for currently available robots. They formulated the costs of truck-based robot delivery using a distance approximation model. The results showed that delivery times, Vehicle Miles Traveled (VMT), and costs could be reduced by the proposed system. For the VMT, trucks could reduce travel distance as much as 31% compared to the trucks without robots. Bakach et al. (2020) set up a two-tier delivery network for *RT*. For Tier 1, a truck transports all packages from a depot to a set of micro hubs in the neighborhood of demand points. From there, the robots conduct the last-mile delivery from the hubs (Tier 2). The researchers formulated the system using the TSP for the following objectives: finding the minimum operating cost, minimum number of robot hubs, and minimum number of robots. In the modelling process, various costs were included, such as gas, driver, and electricity. The study found the following: 1) Cost per package for the proposed system was much less than for conventional truck delivery by 67.9% to 92.3% from their baseline. 2) If time windows did not exist (e.g., for unattended delivery), cost per package in suburban areas showed significant economies-of-scale with increasing demand density. The operating cost for robots could be as low as 24-32% of that for conventional trucks. 3) Doubled robot operating speeds from 1.86 to 3.73 mph did not generate meaningful savings. 4) Driver's pay rate had little influence on the operating cost of robots. 5) Although many robots per hub were required in urban areas, fewer micro hubs were needed compared to suburbs.

In contrast with the previous three *RT* literature, Sonneberg et al. (2019) optimized a robot delivery system (*MR*) without aid of trucks where the objective aimed at minimizing the delivery costs. The system was formulated by a variant of the VRP. Moreover, the robots were designed to carry more than one package per tour. They found that increasing shipments per tour could significantly reduce the total daily costs for the system. For instance, the cost for vehicles with two items carried was about 46% lower than with one single item carried.

2.4 Summary

The dissertation reviewed the existing tour length approximations dating back to 1940. Before Beardwood et al. (1959) developed their common formula in the late 1950s, the existing studies focused on theoretical derivation of the TSP tour length approximation. As more advanced computers and efficient solution methods were introduced, researchers explored accurate coefficients for the formula. After Daganzo (1984) considered realistic aspects of tours, recent studies have focused on various conditions (e.g., shape of service area or spatial distribution). The following section discusses some remaining gaps in the literature and opportunities for improvement.

In Section 2.1, most reviewed studies focused on the derivation of asymptotic coefficients of the TSP tour length and on a relatively large number of points visited per tour. In the literature, approximations are found only for five or more visited points, as shown in Table 1 (Chien 1992; Hindle and Worthington, 2004). In addition, the average TSP tour lengths would be inaccurately estimated if the approximations are derived from wide range of n values (Applegate et al. 2011 and Lei et al. 2015). Therefore, such approximations for small number of n points show promise in analyzing new type of vehicles and delivery alternatives could be efficient because actual tours serve relatively few customers, particularly with vehicle loading capacity or working period constraints. Note that each delivery worker may deliver 200-300 packages per working period in an urban area (Sheth et al. 2019; Tipagornwong and Figliozzi. 2014). Holguín-Veras and Patil (2005) showed that more than 50% of truck routes has less than six stops, while 95% of the truck routes had less than 20 stops in Denver, Colorado. In addition, recent transportation alternatives (e.g., dial-a-ride service, paratransit, small vans, deliveries by bikes, drones, and robots) may not be effectively approximated by such models due to limited vehicle loading capacity. Although these types of vehicles may not handle economically many

shipments per tour, new businesses adopting new technologies have grown due to their advantages, which include speed, responsiveness, or freshness for some items.

In estimating Beardwood's coefficients, the following results are found. First, the number of replications for obtaining the optimal TSP tour length significantly varied in the existing studies, as shown in the 5th column of Table 3. Kwon et al. (1995), Applegate et al. (2011), and Lei et al. (2016) used the same runs across all n values, while others did not present any criteria for the number of replications (e.g., fewer replications for large n , and vice versa). Therefore, consistent runs would help in providing descriptive statistics of each n (e.g., mean, median, standard deviation, kurtosis, or skewness); the dataset of the optimum tour lengths can be investigated further, such as by using sample variance provided in Ong and Huang (1989) in Figure 3. In addition, if the computation cost is affordable, large runs (e.g., 1,000 iterations per n value) would provide more reliable estimates of β . Second, except for Hindle and Worthington (2004), researchers have used a discrete interval of n as an independent variable for regression. For instance, most studies used the intervals which increased by some factors or with increments of specified values in Table 3. A larger increment (observable in the 4th column of Table 3) results in a less accurate value of the coefficient β due to the missing samples. Additional (smaller) intervals improve estimates for non-linear relations by reducing interpolation errors.

Furthermore, as Franceschetti et al (2017) pointed out that the estimates β changed with the value of n , other factors (e.g., the point distribution or shape of space) also affected the estimates. Lastly, approximations considering many variables (e.g., distribution-free approximations) may be less applicable than Beardwood's formula if they require variables that are often unavailable or known a priori, such as the number of vehicles, length-to-width ratio of zones, predetermined number of routes, or standard deviations of points. In addition, approximations with many factors and complicated formulas may degrade the usefulness of

the approximation. That is, adding more variables may increase the accuracy of the approximations as well as their complexity. If approximations become too complicated for practical applications, solving the exact solution may become preferable. Therefore, the researchers should consider such trade-offs for approximations.

For practical applications, estimating average tour lengths with relatively small n values becomes important for package delivery services by vehicles with limited carrying capacities (e.g., autonomous ground/aerial vehicles, or bike/passenger car deliveries). Therefore, the approximations providing tour length estimates for few points are valuable for analyzing and planning such systems.

Chapter 3: Methodology

This chapter develops TSP tour length approximation models. First, the simulation settings and factors for the approximation are discussed. Then, the approximation assumptions and evaluation criteria are presented. Lastly, a solution procedure based on metaheuristics and TSP solvers is discussed.

3.1 Solution Method

3.1.1 Formulation of Traveling Salesman Problem (TSP)

The exact algorithm for a TSP tour is formulated as the following integer program:

$$\text{Minimize } \sum_i \sum_j d_{ij} x_{ij}$$

$$\text{Subject to } \sum_{j=1}^n x_{ij} = 1 \quad i = 1, 2, \dots, n \quad (4)$$

$$\sum_{j=1}^n x_{ji} = 1 \quad i = 1, 2, \dots, n \quad (5)$$

$$u_i - u_j + nx_{ij} \leq n - 1 \quad i = 2, 3, \dots, n; \\ j = 2, 3, \dots, n; i \neq j \quad (6)$$

$$x_{ij} = 0 \text{ or } 1 \quad (7)$$

$$u_i \geq 0 \quad (8)$$

where n is the number of n points (i.e., nodes, instances, or vertices), d_{ij} is the travel distance between points i and j (i.e., edges or arcs), x_{ij} are binary decision variables determining whether the sub-route from i to j is chosen in the tour (constraint 7), u_i is the sequence number in which point i is visited (constraint 8), and constraint (6) is designed for sub-tour elimination, which prohibits solutions consisting of several disconnected tours. Thus, the solution must have a single tour covering all points.

3.1.2 Genetic Algorithm (GA)

A genetic algorithm (GA) is a stochastic search method inspired by the process of natural selection to evolve toward a better solution is chosen. In this algorithm, a finite population of candidate solutions to a TSP problem is created; these initial populations are randomly selected from the enumeration of the permuted TSP tours. This population of solutions is represented as a string of encoded genes called a chromosome. Each chromosome is evaluated and selected to produce the next generation based on its fitness. After the evaluation, the selected chromosomes are processed through crossover and mutation operators.

A crossover operator augments the population by selecting some attributes duplicated from one chromosome and the remaining attributes duplicated from the other, while a mutation operator changes the attributes of single chromosomes. The algorithm is terminated in the following cases: 1) when it reaches the pre-specified number of generations (i.e., the number of cycles) or 2) no improvement in the objective function value is found for a certain number of generations. In this dissertation, both cases are considered. Throughout this process, the algorithm leads to an optimal or near-optimal solution ([Potvin, 1996](#)).

3.1.3 Parameter Selection for GA

GA parameters directly impacted the solution quality, and such parameters included crossover rate, mutation rate, population size, and the number of generations ([Hassanat, 2019](#)). Shayanfar ([2015](#)) pointed out a trade-off between population size and computation time. Increasing population size would benefit the solution quality at a decreasing rate. The study also revealed that a crossover value at 0.5 produced a better solution than other parametric values. In finding TSP solutions, Rexhepi et al. ([2013](#)) investigated the impact of population size and mutation rate on GA. Initial populations of 1,000, 5,000, and 10,000 were investigated, while mutation rates were varied between 1% and 10%. Although an increase in initial

population size did not guarantee a good TSP solution, the solution quality was improved by increasing the mutation rate when the initial populations were below 2,000. However, if the number of generations was small (i.e., 51 cycles), it was shown that both large initial population size and high mutation rate were effective in finding a good solution, according to Beed et al. (2017). The population size was investigated for 100, 500, and 1,000, while mutation rates were changed between 1% and 10%.

Table 4 Parameter Section for Genetic Algorithm (GA)

n	Population Size	Crossover Rate	Mutation Rate	# of Generations	Max. Computation Time (sec)
2 ~ 10	100	0.5	0.04	800	0.1
11 ~ 15	100	0.5	0.04	1,000	5
16 ~ 20	100	0.5	0.04	1,500	20
21 ~ 25	110	0.5	0.04	2,500	30
26 ~ 30	110	0.5	0.04	3,000	35
31 ~ 35	120	0.5	0.04	4,000	55
36 ~ 40	130	0.5	0.04	4,500	80
41 ~ 45	140	0.5	0.04	8,000	100
46 ~ 50	150	0.5	0.04	12,000	120

The above parametric modifications are classified as a deterministic parameter control, while adaptive parameter control uses feedback in altering the parameters (Hassanat, 2019). For the quality of TSP solutions, the study adopts the latter approach, which increases population size and the number of generations as n values increase. Note that the parameter values listed in Table 4 may vary with computing performances. For instance, the results in Table 4 are obtained with four computers. The maximum computation time is estimated from the least performing computer. As a result, longer computation times in Table 4 are needed compared to the other types of known TSP solutions and their computation times (i.e., TSPLIB, a library of sample instances for the TSP in Reinelt, 1991). Note that a brute-force method (i.e., exact search) for solving TSP instances with small instances n takes a long time (Lucas, 2018).

3.1.4 Concorde TSP Solver and Comparison of Solution Methods

Before generating optimized TSP instances for tour length approximations, the dissertation evaluates solution methods based on solution quality (i.e., the lowest tour length for TSPs). First, each heuristic/solver is benchmarked against the best-known solutions of TSPLIB. Among 112 optimal solutions in TSPLIB, a total of eleven cases are selected. Second, the solution methods are compared by solving 1,000 randomly and uniformly generated Euclidean instances. TSPs are generated using MATLAB code (Vedenyov, 2011) and package in R-programming (Hahsler and Hornik, 2007).

Table 5 Optimized TSP Solutions from Heuristic/Solver

#	TSPLIB	<i>n</i>	Exact Solution	Heuristic/Solver				
				GA	2-opt	RNN	Concorde	LK
1	eil51	51	426	444.90	474.84	505.28	428.87	428.88
2	berlin52	52	7,542	8,080.43	8,489.47	8,182.19	7,544.37	7,544.37
3	st70	70	675	722.39	754.87	761.69	677.11	677.11
4	eil76	76	538	582.78	609.10	606.77	544.37	544.37
5	pr76	76	108,159	112,496.25	119,364.57	130,921.00	108,159.44	108,159.44
6	rat99	99	1,211	1,315.16	1,405.33	1,369.53	1,219.24	1,219.27
7	kroA100	100	21,282	22,896.66	24,107.01	24,698.50	21,285.44	21,285.44
8	kroB100	100	22,141	23,061.86	24,864.10	25,882.97	22,139.07	22,139.66
9	kroC100	100	20,749	21,816.12	23,740.72	23,566.40	20,750.76	20,750.76
10	kroD100	100	21,294	22,995.16	24,019.85	24,855.80	21,294.29	21,294.29
11	kroE100	100	22,068	23,648.04	25,013.91	24,907.02	22,068.76	22,076.85

* *GA: Genetic Algorithm, RNN: Repetitive Nearest Neighbor, LK: Chained Lin-Kernighan Heuristic*

Table 5 indicates that the Concorde and Lin-Kernighan heuristic find nearly optimal solutions. Their average percent error is about 0.27% and 0.26%, respectively. The gap for Repetitive Nearest Neighbor (RNN) is 14.82%, while the 2-opt search algorithm overestimates the optimal solution by 12.87%. The GA produces intermediate accuracy with an average percent of 6.51%.

Table 6 Estimated Average TSP Tour Lengths from Heuristics

n	Heuristic/Solver					Rank				
	GA	2-opt	RNN	Concorde	LK	GA	2-opt	RNN	Concorde	LK
2	1.0352	1.0352	1.0352	N/A	N/A	1	1	1	N/A	N/A
3	1.5871	1.5871	1.5871	N/A	N/A	1	1	1	N/A	N/A
4	1.8924	1.9545	1.9242	1.9187	2.1184	1	4	3	2	5
5	2.1160	2.1973	2.1335	2.1193	2.6059	1	4	3	2	5
6	2.3021	2.4180	2.3251	2.2977	3.1169	2	4	3	1	5
7	2.4749	2.6511	2.5279	2.4901	3.6625	1	4	3	2	5
8	2.6129	2.8017	2.6570	2.6089	4.1643	2	4	3	1	5
9	2.7545	2.9647	2.8071	2.7531	2.7582	2	5	4	1	3
10	2.8702	3.0896	2.9186	2.8491	2.8491	3	5	4	1	1
11	2.9602	3.2495	3.0643	2.9839	2.9839	1	5	4	2	2
12	3.0880	3.3723	3.1880	3.0977	3.0977	1	5	4	2	2
13	3.2053	3.5115	3.3056	3.1942	3.1943	3	5	4	1	2
14	3.3216	3.6080	3.3964	3.2822	3.2822	3	5	4	1	1
15	3.4297	3.7159	3.5079	3.3797	3.3797	3	5	4	1	1
16	3.5063	3.8317	3.6318	3.4836	3.4836	3	5	4	1	1
17	3.6031	3.9478	3.7420	3.5784	3.5784	3	5	4	1	1
18	3.6863	4.0556	3.8476	3.6667	3.6667	3	5	4	1	1
19	3.7887	4.1549	3.9579	3.7577	3.7577	3	5	4	1	1
20	3.8747	4.2600	4.0503	3.8300	3.8300	3	5	4	1	1
30	4.6747	5.0766	4.9243	4.5518	4.5519	3	5	4	1	2
40	5.4372	5.7846	5.6633	5.1542	5.1542	3	5	4	1	1

* GA: Genetic Algorithm, RNN: Repetitive Nearest Neighbor, LK: Chained Lin-Kernighan Heuristic

Although the Concorde solver and chained Lin-Kernighan provide a good solution, the former outperforms if n is particularly low (e.g., $n < 10$) in Table 6. Table 6 is designed to compare heuristic performances for randomly generated TSPs and shows that GA generally provides good solutions until $n = 12$. Above that n value, the Concorde solver performs better. However, this does not guarantee that GA always performs better than the others across the cases. Neither the Concorde solver nor Lin-Kernighan always provides the optimal solutions below $n = 5$, where the solution methods accept the local optima to save computation times from repetitive computations. (Helsgaun, 2000 and Lin and Kernighan, 1973). Therefore, optimized TSP instances are taken from two algorithms, whichever provides a better solution.

Although computation time is not the major interest of this dissertation due to a focus on TSP with low n values, Lin-Kernighan Heuristic is the fastest among five heuristics (e.g., 0.01 seconds for $n = 50$ and less than 1 second for $n = 100$). The Concord solver generates a solution within an average of 0.08 seconds ($n = 50$), while the GA has the longest computation time.

3.2 Simulation Settings

3.2.1 Scenario Design

This section explains how the dissertation designs various operating conditions in a simulation setting. A depot - distribution center where vehicles start and end their tours - may or may not be in a center of city. Here considers depots located centrally or randomly in a service region. Note that vehicles departing from a depot outside the region would conduct a TSP tour (i.e., a line-haul distance from the depot to the first recipient of the service region is not considered). Although the exact shape of the service area varies with district partitions, two shape categories are considered: square and circle.

The following two categories are essentially relaxing assumptions for approximations discussed in Section 2.1.4; namely, points are scattered randomly and uniformly in the service area. The effects of concentrating the n points toward a particular direction (e.g., non-uniform distribution of the points) will be explored. To do this, the triangular distribution is adopted with different mode (peak) values in Figure 4 (b) and (c), respectively. Then, a bivariate normal distribution is designed to reflect a resident distribution in cities.

Lastly, the elongations of service area focus on changes in service area shape (i.e., reasonably convex but less compact than the square or circular shape region). The length-to-width ratio varies from 1 to 4. Note that the coefficients for changes in area size (A) can be conveniently adjusted by post-processing using the formula $\beta\sqrt{nA}$ of Beardwood et al. (1959).

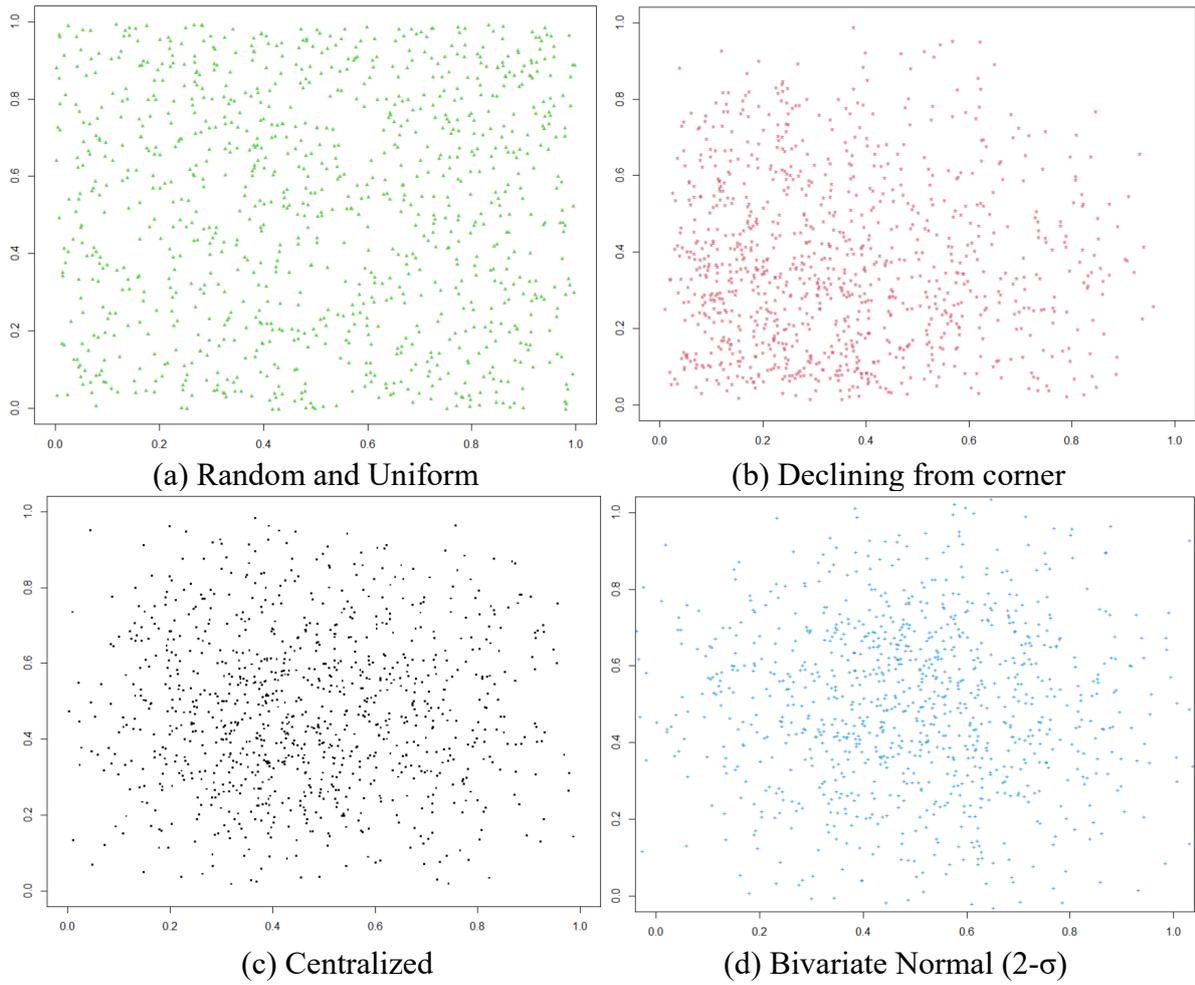
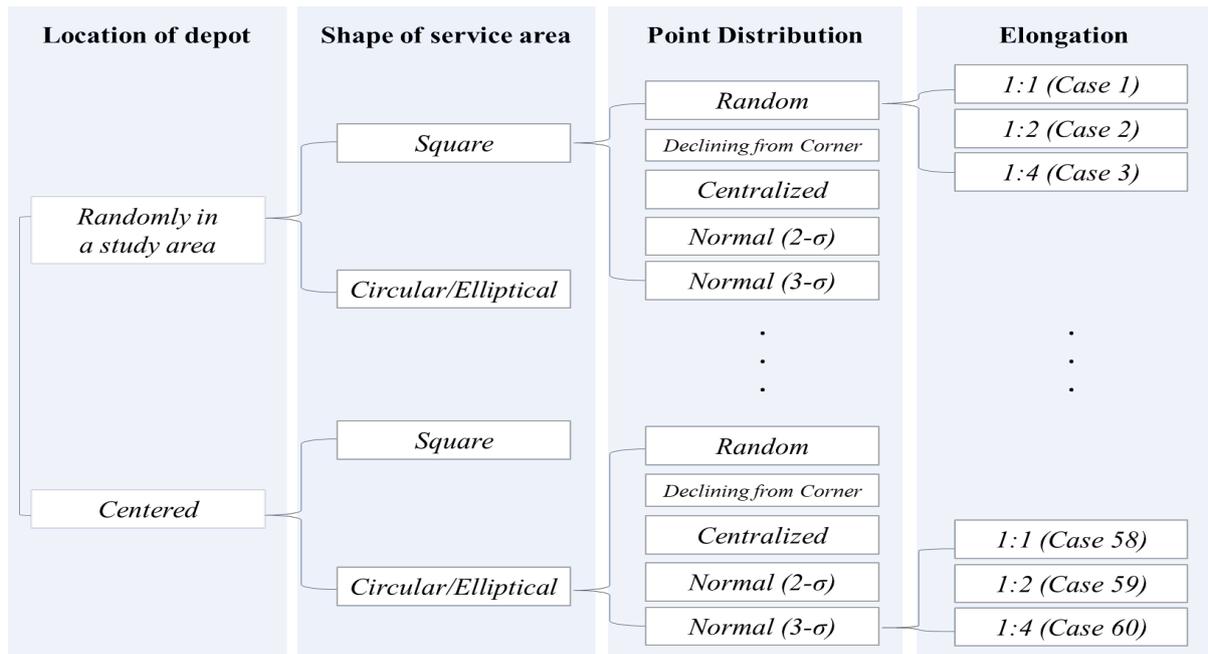


Figure 4 Illustration of Point Distributions ($n = 1,000$)

Overall, Figure 5 summarizes the classifications for the TSP simulation, which are extended from Larson and Odoni (1981).



Case Number	Area Shape	Depot Location	Elongation (L/W ratio)	Point Distribution	Case Number	Area Shape	Depot Location	Elongation (L/W ratio)	Point Distribution	Case Number	Area Shape	Depot Location	Elongation (L/W ratio)	Point Distribution
1	Square	Random	1	Uniform	21	Square	Center	4	Centralized	41	Cir/Ellipt	Random	2	Normal (2σ)
2	Square	Random	2	Uniform	22	Square	Center	1	Declining	42	Cir/Ellipt	Random	4	Normal (2σ)
3	Square	Random	4	Uniform	23	Square	Center	2	Declining	43	Cir/Ellipt	Random	1	Normal (3σ)
4	Square	Random	1	Centralized	24	Square	Center	4	Declining	44	Cir/Ellipt	Random	2	Normal (3σ)
5	Square	Random	2	Centralized	25	Square	Center	1	Normal (2σ)	45	Cir/Ellipt	Random	4	Normal (3σ)
6	Square	Random	4	Centralized	26	Square	Center	2	Normal (2σ)	46	Cir/Ellipt	Center	1	Uniform
7	Square	Random	1	Declining	27	Square	Center	4	Normal (2σ)	47	Cir/Ellipt	Center	2	Uniform
8	Square	Random	2	Declining	28	Square	Center	1	Normal (3σ)	48	Cir/Ellipt	Center	4	Uniform
9	Square	Random	4	Declining	29	Square	Center	2	Normal (3σ)	49	Cir/Ellipt	Center	1	Centralized
10	Square	Random	1	Normal (2σ)	30	Cir/Ellipt	Random	4	Normal (3σ)	50	Cir/Ellipt	Center	2	Centralized
11	Square	Random	2	Normal (2σ)	31	Cir/Ellipt	Random	1	Uniform	51	Cir/Ellipt	Center	4	Centralized
12	Square	Random	4	Normal (2σ)	32	Cir/Ellipt	Random	2	Uniform	52	Cir/Ellipt	Center	1	Declining
13	Square	Random	1	Normal (3σ)	33	Cir/Ellipt	Random	4	Uniform	53	Cir/Ellipt	Center	2	Declining
14	Square	Random	2	Normal (3σ)	34	Cir/Ellipt	Random	1	Centralized	54	Cir/Ellipt	Center	4	Declining
15	Square	Random	4	Normal (3σ)	35	Cir/Ellipt	Random	2	Centralized	55	Cir/Ellipt	Center	1	Normal (2σ)
16	Square	Center	1	Uniform	36	Cir/Ellipt	Random	4	Centralized	56	Cir/Ellipt	Center	2	Normal (2σ)
17	Square	Center	2	Uniform	37	Cir/Ellipt	Random	1	Declining	57	Cir/Ellipt	Center	4	Normal (2σ)
18	Square	Center	4	Uniform	38	Cir/Ellipt	Random	2	Declining	58	Cir/Ellipt	Center	1	Normal (3σ)
19	Square	Center	1	Centralized	39	Cir/Ellipt	Random	4	Declining	59	Cir/Ellipt	Center	2	Normal (3σ)
20	Square	Center	2	Centralized	40	Cir/Ellipt	Random	1	Normal (2σ)	60	Cir/Ellipt	Center	4	Normal (3σ)

Figure 5 Classifications for Distance Approximation

3.2.2 Simulation Design: Point Generation, Point Distribution, and Least Squares

Method

In a simulation setting, n points are generated based on scenarios developed in Section 3.2.1. Random numbers provided in the simulation program (i.e., "rand" function) are used for uniform and random distribution in Figure 5. Using the 'rand' function for producing random

numbers uniformly distributed in the interval (0,1), two sets of random numbers are generated (Moler, 2008). These two are regarded as the x- and y-coordinate of demand location in a service area.

For the cases with non-uniform distributions, the appropriate point generating functions in the program are used to generate points. The parameters of peak value considered are either 0.1 (declining from corner) or 0.4 (centralized) for a triangular distribution, implying that the probability of selecting points is high near the peak value within the interval [0, 1]. Since the randomly generated points for bivariate normal distribution are theoretically unbounded in the interval [0,1], the points generated outside the service area are truncated. Then, new points are added until all points lie within the interval. The mean value is 0.5 (i.e., located at the center coordinate), while the standard deviations (σ) are 0.25 for 2- σ and 0.19 for 3- σ . The 95 or 99 percent of the points are generated near the center point (0.5, 0.5) within the standard deviations of 2- σ or 3- σ , respectively.

Then, sets of 1,000 TSP tour instances are generated by changing the points n from 1 to 100 (i.e., 1,000 runs per n value), in increment of one; each set of averaged tour lengths is fitted using OLS regression to estimate the coefficient β for \sqrt{nA} .

3.3 Results

3.3.1 Descriptive Statistics of the Optimized TSP Instances

Although any of the 60 categories in Figure 5 can be considered, descriptive statistics are provided in Table 7 only for a randomly located depot in Euclidean space with a square service area (Case 1). The case may be practically used to approximate TSP distances in urban road networks, considering that (1) ground vehicles travel on a grid network and (2) distribution depots are typically located away from central business districts in order to reduce costs. The

purpose of providing the statistics is to examine whether a specific pattern or distribution exists for the 1,000 optimized TSPs.

Table 7 Descriptive Statistics and Normality Test (Case 1)

<i>n</i>	Descriptive Statistics							Normality Test	
	Mean	Median	5 th	95 th	STD	Skew	Kurt	W	P value
2	1.0274	1.0195	0.2382	1.8468	0.4920	0.1630	-0.6334	0.9884	0.0000
3	1.5943	1.5825	0.7868	2.3502	0.4735	0.0196	-0.2635	0.9983	<u>0.4530</u>
4	1.8956	1.9049	1.1226	2.5943	0.4327	-0.1502	-0.2009	0.9972	<u>0.0832</u>
5	2.1118	2.1249	1.3975	2.7370	0.4098	-0.1648	-0.2950	0.9958	0.0077
6	2.3023	2.3101	1.6178	2.9185	0.3978	-0.3140	0.2886	0.9939	0.0004
7	2.4731	2.4769	1.8609	3.0506	0.3674	-0.0649	-0.0087	0.9983	<u>0.4499</u>
8	2.6066	2.6292	1.9936	3.2019	0.3528	-0.2096	0.3853	0.9952	0.0030
9	2.7560	2.7583	2.2135	3.2969	0.3360	-0.0692	-0.0287	0.9986	<u>0.6039</u>
10	2.8719	2.8711	2.2942	3.3974	0.3373	-0.1569	0.3420	0.9968	0.0400
11	2.9563	2.9495	2.4276	3.4977	0.3277	0.1346	-0.0022	0.9979	<u>0.2272</u>
12	3.0968	3.1091	2.5433	3.6289	0.3217	-0.0305	-0.0987	0.9986	<u>0.6081</u>
13	3.1948	3.1951	2.6836	3.7117	0.3145	-0.0357	-0.2545	0.9985	<u>0.5547</u>
14	3.2820	3.2911	2.7346	3.8112	0.3225	-0.0162	-0.0153	0.9983	<u>0.4030</u>
15	3.3973	3.3972	2.8733	3.9166	0.3153	-0.0369	0.0269	0.9989	<u>0.8390</u>
16	3.4984	3.5046	2.9546	4.0075	0.3158	-0.0647	-0.1719	0.9979	<u>0.2318</u>
17	3.5912	3.5949	3.0707	4.0914	0.3061	-0.0745	-0.1785	0.9982	<u>0.3547</u>
18	3.6689	3.6700	3.1438	4.1822	0.3085	-0.0800	-0.0765	0.9986	<u>0.6105</u>
19	3.7554	3.7645	3.2228	4.2425	0.3133	-0.2136	-0.0918	0.9963	0.0165
20	3.8308	3.8198	3.3174	4.3258	0.3016	-0.0410	0.0389	0.9989	<u>0.7940</u>
21	3.9220	3.9215	3.4129	4.4192	0.3124	-0.0999	-0.0193	0.9987	<u>0.7035</u>
22	3.9947	3.9970	3.4912	4.4821	0.2994	-0.0984	-0.0387	0.9979	<u>0.2264</u>
23	4.0751	4.0688	3.6187	4.5413	0.2873	-0.0871	-0.2270	0.9977	<u>0.1752</u>
24	4.1492	4.1592	3.6718	4.6046	0.2843	-0.1792	-0.0694	0.9971	<u>0.0708</u>
25	4.2154	4.2348	3.7177	4.6783	0.2876	-0.2153	0.1252	0.9953	0.0037
26	4.2758	4.2945	3.7652	4.7406	0.2950	-0.2193	-0.0771	0.9955	0.0045
27	4.3410	4.3498	3.8432	4.7971	0.2852	-0.2394	0.0091	0.9954	0.0040
28	4.4249	4.4336	3.9377	4.8806	0.2889	-0.2473	0.1734	0.9955	0.0046
29	4.4850	4.4950	4.0297	4.9217	0.2789	-0.1700	0.3394	0.9971	<u>0.0686</u>
30	4.5707	4.5878	4.1135	5.0211	0.2794	-0.1998	-0.0872	0.9964	0.0227
31	4.6213	4.6205	4.1540	5.0803	0.2799	-0.0559	-0.1489	0.9989	<u>0.8280</u>
32	4.6811	4.6767	4.2111	5.1003	0.2671	-0.1415	-0.1379	0.9971	<u>0.0633</u>
33	4.7367	4.7327	4.2849	5.2185	0.2801	-0.0716	0.1786	0.9982	<u>0.3542</u>
34	4.8260	4.8459	4.3263	5.2345	0.2715	-0.3088	0.1108	0.9934	0.0002
35	4.8581	4.8614	4.4104	5.3033	0.2718	-0.1394	-0.0261	0.9982	<u>0.3742</u>
36	4.9258	4.9277	4.4809	5.3664	0.2679	-0.1430	-0.0205	0.9979	<u>0.2357</u>
37	4.9877	4.9926	4.5477	5.4111	0.2659	-0.0668	-0.0587	0.9992	<u>0.9441</u>
38	5.0510	5.0674	4.6217	5.4556	0.2556	-0.2669	0.1511	0.9951	0.0027
39	5.1000	5.1033	4.6697	5.5448	0.2706	-0.1996	0.2302	0.9959	0.0095
40	5.1658	5.1768	4.7270	5.5751	0.2630	-0.2330	0.1815	0.9960	0.0115
41	5.2212	5.2155	4.8053	5.6608	0.2598	0.0260	-0.1390	0.9989	<u>0.8286</u>
42	5.2660	5.2739	4.8205	5.6892	0.2625	-0.1312	0.2328	0.9978	<u>0.2209</u>
43	5.3374	5.3438	4.8960	5.7539	0.2645	-0.2422	-0.0184	0.9957	0.0067
44	5.3916	5.4015	4.9693	5.8275	0.2637	-0.0958	-0.2604	0.9972	<u>0.0827</u>
45	5.4518	5.4510	5.0417	5.8572	0.2532	-0.1501	0.2024	0.9974	<u>0.1031</u>
46	5.5010	5.5120	5.0790	5.9038	0.2524	-0.3308	0.4369	0.9931	0.0001
47	5.5390	5.5424	5.1154	5.9739	0.2532	-0.1241	0.5964	0.9957	0.0068
48	5.5847	5.5865	5.1603	6.0034	0.2567	-0.1807	0.1385	0.9974	<u>0.1161</u>

49	5.6407	5.6456	5.2176	6.0598	0.2566	-0.1329	0.0118	0.9981	0.3337
50	5.6852	5.6949	5.2687	6.0603	0.2507	-0.1535	-0.0022	0.9974	0.1174
51	5.7403	5.7568	5.3242	6.1355	0.2444	-0.2611	0.1071	0.9946	0.0011
52	5.7973	5.8014	5.3751	6.1892	0.2465	-0.1935	0.0685	0.9973	0.0902
53	5.8483	5.8563	5.3944	6.2501	0.2576	-0.2493	-0.0509	0.9957	0.0068
54	5.8903	5.8967	5.4806	6.2934	0.2477	-0.0967	0.1907	0.9977	0.1748
55	5.9429	5.9552	5.5381	6.3334	0.2448	-0.0976	-0.0080	0.9982	0.3815
56	5.9920	5.9904	5.5594	6.4237	0.2579	-0.0009	-0.2316	0.9982	0.3918
57	6.0365	6.0388	5.6262	6.4588	0.2537	-0.1235	0.0937	0.9982	0.3872
58	6.0636	6.0670	5.6738	6.4440	0.2355	-0.0724	-0.1621	0.9987	0.6552
59	6.1318	6.1356	5.7087	6.5384	0.2490	-0.1722	-0.0178	0.9976	0.1487
60	6.1404	6.1531	5.6804	6.5332	0.2554	-0.2005	0.0713	0.9961	0.0121
61	6.2146	6.2211	5.7865	6.6031	0.2473	-0.1266	-0.1240	0.9973	0.0914
62	6.2562	6.2555	5.8434	6.6575	0.2427	-0.0769	-0.0610	0.9988	0.7292
63	6.3096	6.3204	5.8860	6.6859	0.2403	-0.1372	0.2102	0.9971	0.0706
64	6.3506	6.3597	5.9342	6.7374	0.2444	-0.1563	-0.2118	0.9967	0.0324
65	6.3966	6.4027	6.0165	6.7843	0.2333	-0.0689	-0.0921	0.9977	0.1682
66	6.4205	6.4273	6.0108	6.8076	0.2412	-0.2048	-0.0224	0.9961	0.0135
67	6.4759	6.4890	6.0585	6.8620	0.2445	-0.1788	0.0188	0.9971	0.0656
68	6.5192	6.5163	6.1204	6.9154	0.2418	-0.0689	-0.2390	0.9976	0.1599
69	6.5717	6.5753	6.1503	6.9571	0.2408	-0.2112	0.1257	0.9963	0.0169
70	6.6008	6.6095	6.1961	6.9821	0.2368	-0.1802	0.1604	0.9972	0.0820
71	6.6411	6.6511	6.2193	7.0447	0.2465	-0.1311	0.1314	0.9984	0.4624
72	6.6829	6.6894	6.2657	7.0659	0.2417	-0.0830	0.0512	0.9983	0.4348
73	6.7364	6.7482	6.3096	7.1277	0.2474	-0.1988	-0.0241	0.9967	0.0326
74	6.7857	6.7894	6.3716	7.1561	0.2418	-0.1949	0.1696	0.9971	0.0664
75	6.8322	6.8401	6.4326	7.2078	0.2359	-0.1829	0.3072	0.9971	0.0700
76	6.8391	6.8376	6.4233	7.2445	0.2491	-0.0958	-0.1912	0.9979	0.2443
77	6.9125	6.9209	6.5030	7.3224	0.2420	-0.0950	-0.0539	0.9986	0.5848
78	6.9413	6.9475	6.5291	7.3250	0.2397	-0.0990	-0.1370	0.9985	0.5723
79	6.9765	6.9829	6.5841	7.3482	0.2303	-0.1479	-0.0567	0.9980	0.2765
80	7.0184	7.0227	6.6261	7.3856	0.2388	-0.0307	-0.1156	0.9987	0.6820
81	7.0564	7.0639	6.6714	7.4418	0.2358	-0.0182	-0.1617	0.9985	0.5338
82	7.1084	7.1153	6.7163	7.4935	0.2342	-0.0537	-0.1469	0.9985	0.5832
83	7.1331	7.1339	6.7564	7.5147	0.2314	-0.0391	-0.2488	0.9984	0.4917
84	7.1671	7.1736	6.7862	7.5193	0.2306	-0.1728	-0.1016	0.9972	0.0797
85	7.2091	7.2153	6.7784	7.5763	0.2381	-0.2682	-0.0888	0.9944	0.0009
86	7.2449	7.2484	6.8675	7.6044	0.2255	-0.0959	-0.0726	0.9987	0.6889
87	7.2683	7.2759	6.8546	7.6513	0.2357	-0.0598	0.0076	0.9983	0.4539
88	7.3184	7.3323	6.9263	7.6963	0.2336	-0.2242	0.0399	0.9959	0.0093
89	7.3591	7.3679	6.9471	7.7477	0.2390	-0.1479	-0.1703	0.9975	0.1253
90	7.4007	7.4017	7.0068	7.7800	0.2348	-0.0894	-0.0447	0.9985	0.5405
91	7.4416	7.4452	7.0361	7.8269	0.2357	-0.1097	-0.0195	0.9972	0.0773
92	7.4662	7.4654	7.1007	7.8368	0.2313	0.0379	0.0023	0.9989	0.8093
93	7.5110	7.5109	7.1317	7.8946	0.2283	-0.0405	-0.1791	0.9988	0.7758
94	7.5456	7.5518	7.1590	7.9291	0.2346	0.0079	-0.3532	0.9961	0.0135
95	7.5979	7.6037	7.2203	7.9568	0.2300	-0.1244	0.4079	0.9974	0.1094
96	7.6215	7.6234	7.2372	7.9783	0.2232	-0.1126	-0.1641	0.9979	0.2406
97	7.6559	7.6700	7.2633	8.0134	0.2292	-0.1668	0.0154	0.9973	0.0927
98	7.6841	7.6929	7.2670	8.0416	0.2297	-0.2671	0.0361	0.9945	0.0010
99	7.7374	7.7354	7.3556	8.1250	0.2344	-0.1171	0.0491	0.9978	0.2229
100	7.7627	7.7668	7.3844	8.1333	0.2310	-0.0459	-0.0281	0.9988	0.7541

* *Bold numbers infer the p-values exceeding 0.05*

In the 7th and 8th columns of Table 7, measurements for the central tendency and tails of data distributions are listed for the optimized TSP instances. The 5th and 95th percentile lengths are provided to limit the range of values for average TSP tour lengths. To identify whether each set of tour lengths lies in a reasonable range, additional normality tests can be performed, such as Shapiro-Wilk, Chi-Square, Kolmogorov-Smirnov, Cramér-von-Mises, or Anderson-Darling test. These are depending on sample sizes (D'Agostino, 1986). Among the tests, the Shapiro-Wilk test is conducted, which is (1) widely used to test for normality and (2) sensitive for sample sizes up to 2,000 (Yap and Sim, 2011). The p-values exceeding 0.05 in the 10th column of Table 7 indicate that the distribution of the generated TSP instances fits the normal distribution.

From the test outputs, the optimized TSP instances do not have a specific distribution. This finding is aligned with Monte Carlo simulation results from Vinel and Silva (2018), where (1) a consistent deviation from normality exists for $n = 3$ and (2) it is difficult to conclude whether the optimized TSPs follow a normal distribution between $n = 4$ and $n = 10$. The researchers adopted alternative methods (i.e., PP and QQ plots) for examining the normality of TSPs since the Shapiro-Wilk test outputs inconsistent results (i.e., p-value) with the sample size.

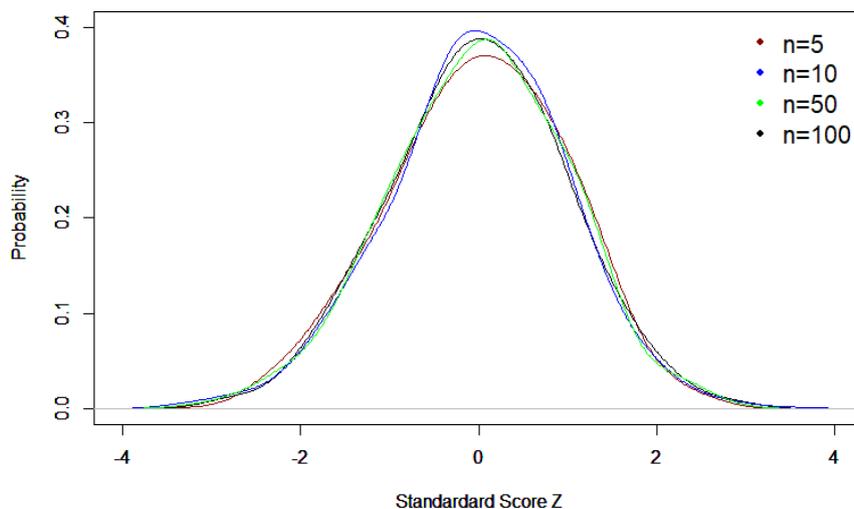


Figure 6 PDFs for the TSP Instances for Different n Values

The probability distribution functions (PDFs) are presented in Figure 6. All the generated instances for each n value are symmetrical based on the Kolmogorov–Smirnov test. Namely, a null hypothesis that the optimized TSPs are symmetric is accepted (e.g., with p-values of 0.143 for $n = 2$ and 0.570 for $n = 100$).

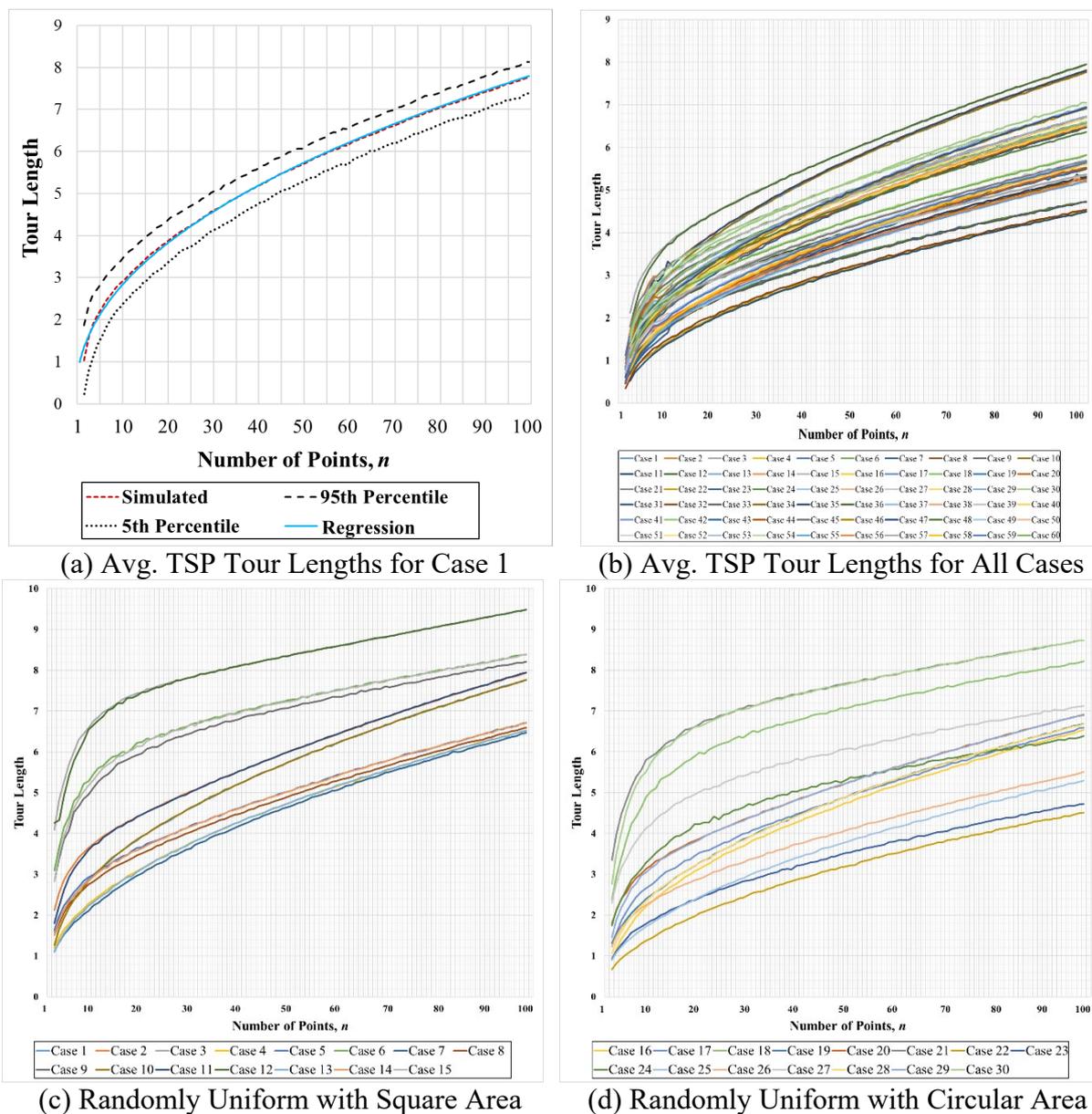


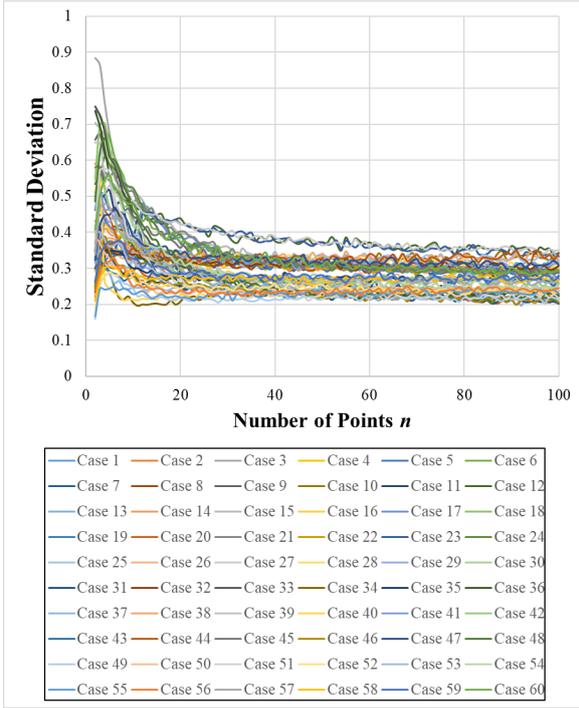
Figure 7 Simulation Results

The curves for average simulated TSP tour lengths, regression results, and 5th percentile, and 95th percentile lengths for Case 1 are presented in Figure 7 (a). Average TSP tour lengths

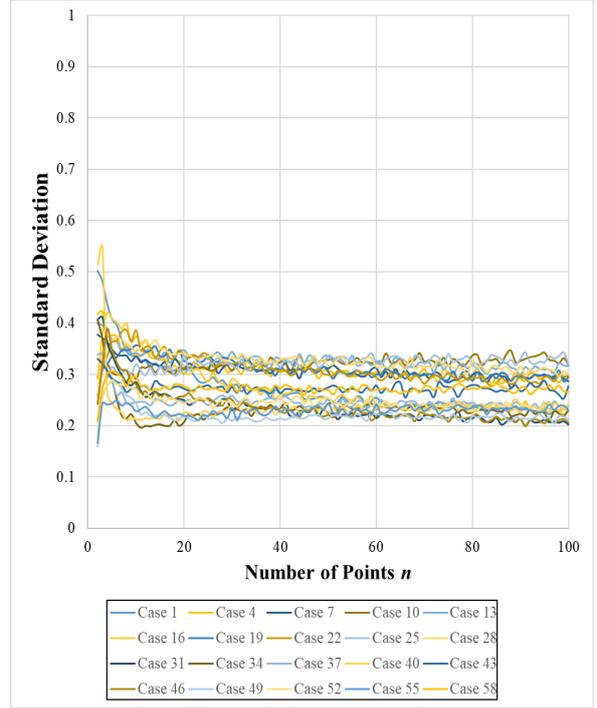
increase rapidly as n increases particularly for the smaller values of n . After that, the TSP tour lengths marginally increase with n .

3.3.2 Standard Deviations of Average TSP Tour Lengths

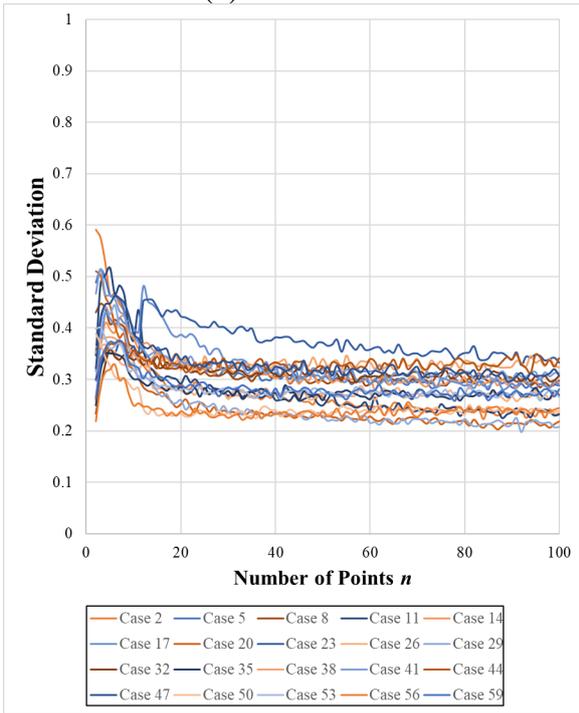
As Ong and Huang (1989) presented the fluctuations in the variance of the optimized TSP tour instances in Figure 3, similar trends in standard deviations (SD) were observed. The SDs decrease as n increases, while the SDs increase as the service area become more elongated in Figure 8 (b), (c), and (d). In comparison to square and circular areas (Figure 8 (e) and (f)), the SDs for circular are smaller than the square since the circle (or ellipse) is more compact than the square.



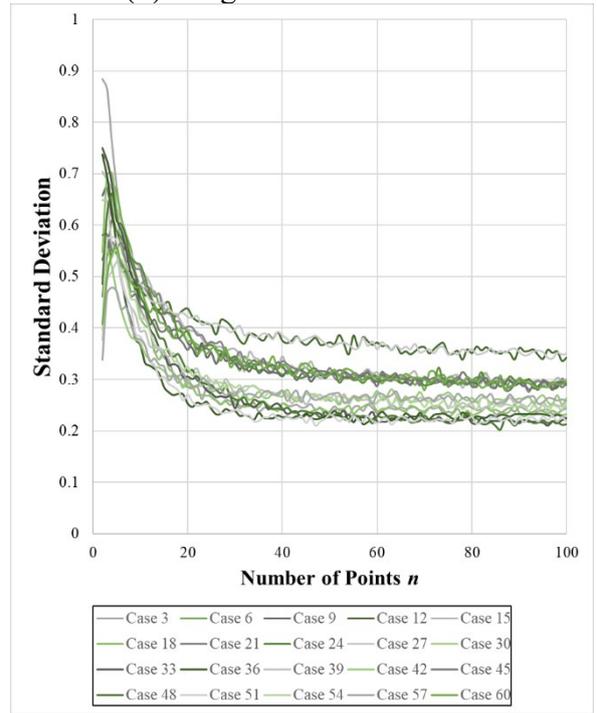
(a) All Cases



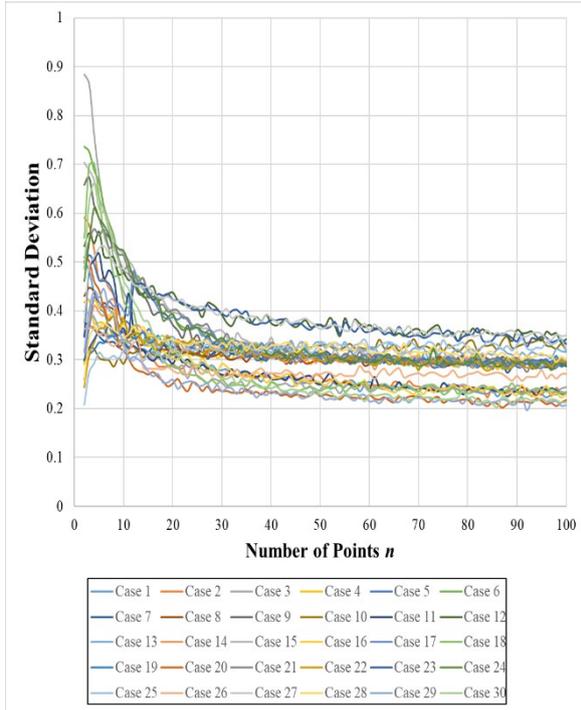
(b) Length-Width Ratio of 1



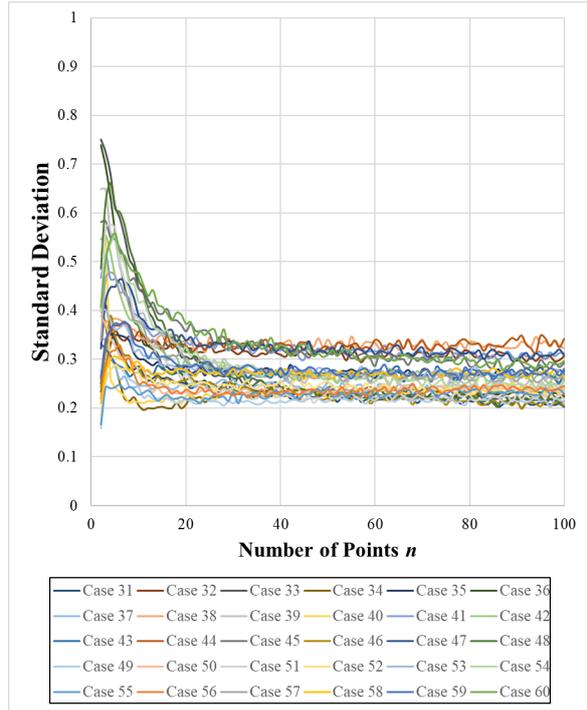
(c) Length-Width Ratio of 2



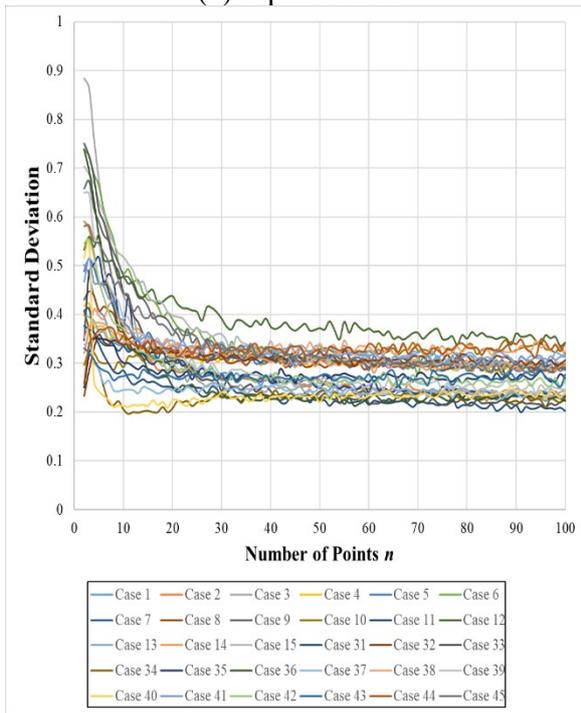
(d) Length-Width Ratio of 4



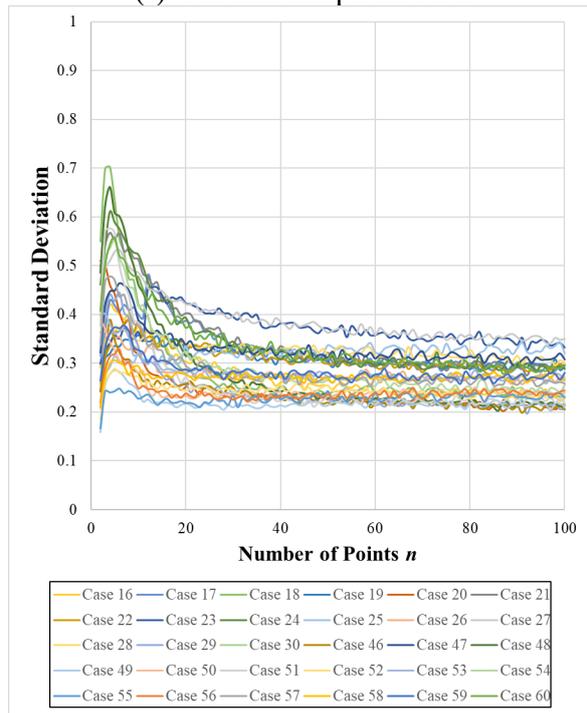
(e) Square Area



(f) Circular/Elliptical Area



(g) Random Depot



(h) Central Depot

Figure 8 Standard Deviations of the TSP Instances

The SDs for the central depot increase until n reaches 4 points. Since one point among n points is positioned at a central location for the latter case, the points are less scattered than in the random depot. It is true for a few n values, but no specific trends can be observed as n

increases. Note that the SDs for point distributions are not presented as the differences are not clearly noticeable.

3.3.3 Estimated Coefficients β

Table 8 shows estimated coefficients β for \sqrt{nA} and adjusted coefficients of determination (R^2) according to changes in points n . Ranges imply the difference in coefficients β estimated by the 5th and 95th percentiles of the TSP tour instance. For a length-width ratio of one, i.e. a square, all estimated coefficients β are consistent with the finding from Finch (2003) that the estimates lie between 0.632499 and 0.91996 for Euclidean space. Note that β estimated by the 5th percentile distances increases as n increases, while β estimated by the 95th distances decreases as n increases. Overall, the gap between the two percentiles decreases.

3.3.3.1 Comparison between Randomly and Centrally Located Depot

Average TSP tours are shorter for central than for randomly located depots since one point is always positioned at the center. While the estimated β in the latter case is larger than the former, the gap diminishes with an increase in n value. In addition, the difference in tour lengths is unnoticeable if n is beyond fifteen (i.e., as small as 0.3% in difference). The coefficients for a central depot are generally smaller than for a randomly located depot.

3.3.3.2 Comparison between Square and Circular/elliptical Service Area

The estimated coefficients are approximately 1.13 times greater for square areas than for circular/elliptical areas since the latter is more compact. The ratios between a squared-shaped and circular/elliptical area decrease as n increases; the difference in tour lengths with few n points is huge. For the same reason, average distances between two random points in a circle are smaller than in a square of equal area (Larson and Odoni, 1981).

3.3.3.3 Comparison among Point Distributions

Among the different point distributions, the coefficients β estimated from random and uniform distribution are the largest since the points are loosely distributed compared to the more concentrated ones in any other distributions in Figure 4. β for centralized and normal distributions is smaller than any other distribution. Since points are more clustered at center for normal distribution with 3- σ than for 2- σ , β for 3- σ is smaller.

3.3.3.4 Comparison among Different Elongated Service Areas

For different length-to-width ratios of the service area, the goodness of fit decreases as the area becomes more elongated. The square root form may not be the best fit if the length-to-width ratios becomes very high.

Table 8 Summary of Estimated Coefficients β

Depot Location	Service Area Shape	n	Length-to-width ratio of 1					Length-to-width ratio of 2					Length-to-width ratio of 4					
			R	D	C	2 σ	3 σ	R	D	C	2 σ	3 σ	R	D	C	2 σ	3 σ	
Random	Square (Cases 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15)	5	0.9094	0.7000	0.6490	0.6900	0.5145	0.9094	0.7000	0.6490	0.6900	0.5145	1.2260	0.9569	0.8738	0.9408	0.7005	
		10	0.9194	0.7179	0.6682	0.7133	0.5363	0.9194	0.7179	0.6682	0.7133	0.5363	1.2050	0.9522	0.8869	0.9486	0.7144	
		20	0.8851	0.7036	0.6661	0.7125	0.5429	0.8851	0.7036	0.6661	0.7125	0.5429	1.0840	0.8739	0.8268	0.8847	0.6779	
		30	0.8612	0.6910	0.6622	0.7095	0.5453	0.8612	0.6910	0.6622	0.7095	0.5453	1.0030	0.8189	0.7819	0.8370	0.6496	
		40	0.8443	0.6829	0.6592	0.7069	0.5461	0.8443	0.6829	0.6592	0.7069	0.5461	0.9506	0.7830	0.7522	0.8049	0.6299	
		50	0.8320	0.6769	0.6571	0.7048	0.5473	0.8320	0.6769	0.6571	0.7048	0.5473	0.9149	0.7583	0.7311	0.7827	0.6166	
	100	0.7979	0.6615	0.6504	0.6981	0.5499	0.7979	0.6615	0.6504	0.6981	0.5499	0.8355	0.7014	0.6840	0.7325	0.5859		
	Cir/Elip (Cases 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30)	5	0.7484	0.5210	0.3952	0.5636	0.6148	0.7484	0.5210	0.3952	0.5636	0.6148	1.0320	0.7215	0.5411	0.7755	0.7716	
		10	0.7601	0.5395	0.4198	0.5799	0.5810	0.7601	0.5395	0.4198	0.5799	0.5810	1.0220	0.7272	0.5665	0.7800	0.7627	
		20	0.7357	0.5361	0.4334	0.5757	0.5490	0.7357	0.5361	0.4334	0.5757	0.5490	0.9317	0.6797	0.5492	0.7271	0.6949	
		30	0.7205	0.5340	0.4397	0.5717	0.5365	0.7205	0.5340	0.4397	0.5717	0.5365	0.8670	0.6446	0.5339	0.6888	0.6483	
		40	0.7102	0.5331	0.4428	0.5695	0.5304	0.7102	0.5331	0.4428	0.5695	0.5304	0.8244	0.6216	0.5210	0.6633	0.6192	
		50	0.7026	0.5323	0.4446	0.5677	0.5273	0.7026	0.5323	0.4446	0.5677	0.5273	0.7952	0.6053	0.5118	0.6454	0.5995	
	100	0.6818	0.5301	0.4486	0.5641	0.5220	0.6818	0.5301	0.4486	0.5641	0.5220	0.7277	0.5686	0.4877	0.6046	0.5570		
	Center	Square (Cases 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45)	5	0.7819	0.6501	0.5619	0.5873	0.4378	0.7819	0.6501	0.5619	0.5873	0.4378	1.0690	0.8909	0.7543	0.7780	0.5948
			10	0.8644	0.6901	0.6539	0.6661	0.5007	0.8644	0.6901	0.6539	0.6661	0.5007	1.1380	0.9133	0.8274	0.7943	0.6634
			20	0.8728	0.6901	0.6811	0.6955	0.5275	0.8728	0.6901	0.6811	0.6955	0.5275	1.0620	0.8573	0.8029	0.7634	0.6572
			30	0.8566	0.6842	0.6804	0.6986	0.5346	0.8566	0.6842	0.6804	0.6986	0.5346	0.9922	0.8100	0.7691	0.7323	0.6375
			40	0.8424	0.6791	0.6768	0.6989	0.5388	0.8424	0.6791	0.6768	0.6989	0.5388	0.9441	0.7775	0.7437	0.7100	0.6218
50			0.8306	0.6745	0.6727	0.6989	0.5410	0.8306	0.6745	0.6727	0.6989	0.5410	0.9108	0.7543	0.7249	0.6940	0.6102	
100		0.7974	0.6610	0.6607	0.6953	0.5470	0.7974	0.6610	0.6607	0.6953	0.5470	0.8351	0.7001	0.6818	0.6567	0.5831		
Cir/Elip (Cases 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60)		5	0.6461	0.4492	0.3308	0.4798	0.4921	0.6461	0.4492	0.3308	0.4798	0.4921	0.8912	0.6125	0.5471	0.6571	0.6706	
		10	0.7181	0.5055	0.3874	0.5405	0.5407	0.7181	0.5055	0.3874	0.5405	0.5407	0.9587	0.6753	0.5725	0.7207	0.7180	
		20	0.7251	0.5235	0.4174	0.5597	0.5430	0.7251	0.5235	0.4174	0.5597	0.5430	0.9088	0.6588	0.5430	0.7046	0.6805	
		30	0.7151	0.5261	0.4287	0.5613	0.5375	0.7151	0.5261	0.4287	0.5613	0.5375	0.8558	0.6335	0.5273	0.6760	0.6430	
		40	0.7067	0.5272	0.4342	0.5619	0.5328	0.7067	0.5272	0.4342	0.5619	0.5328	0.8176	0.6144	0.5158	0.6539	0.6170	
		50	0.7001	0.5275	0.4377	0.5615	0.5298	0.7001	0.5275	0.4377	0.5615	0.5298	0.7908	0.6002	0.5071	0.6384	0.5994	
		100	0.6808	0.5281	0.4452	0.5608	0.5235	0.6808	0.5281	0.4452	0.5608	0.5235	0.7264	0.5668	0.4852	0.6014	0.5585	

* highlights imply the estimates with a low goodness of fit ($R^2 < 0.8$)

* R: Random and uniform, D: declining from corner, C: Centralized, 2 σ and 3 σ : normal distribution within 2 and 3 standard deviations

Table 9 Estimated Coefficients β with Length-to-Width Ratio of 1

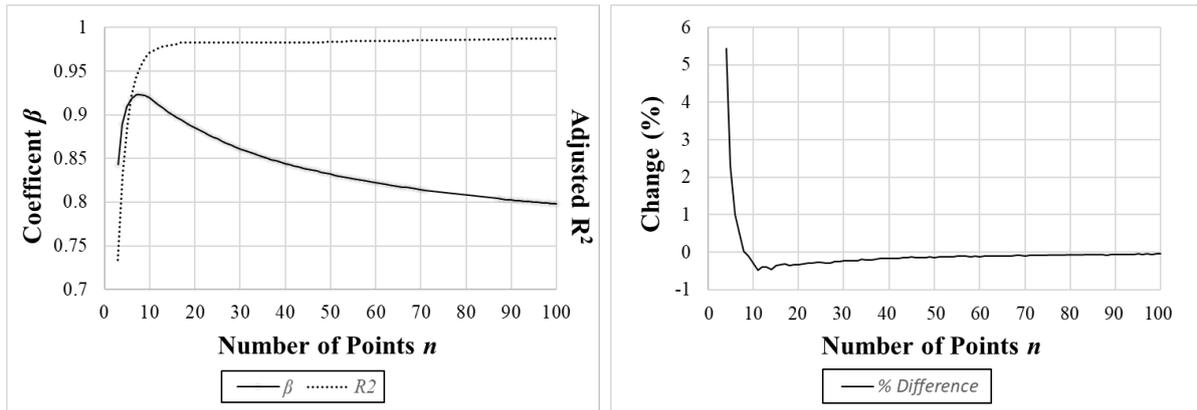
Depot Location	Service Area Shape	n	R			D			C			2 σ			3 σ		
			β	Range	R ²	β	Range	R ²	β	Range	R ²	β	Range	R ²	β	Range	R ²
Random	Square (Cases 1, 4, 7, 10, 13)	5	0.9094	0.505~1.466	0.89	0.7000	0.361~1.043	0.88	0.6490	0.336~0.981	0.88	0.6900	0.378~1.043	0.89	0.5145	0.265~0.805	0.89
		10	0.9194	0.658~1.165	0.97	0.7179	0.474~0.880	0.97	0.6682	0.445~0.900	0.97	0.7133	0.481~0.955	0.97	0.5363	0.347~0.742	0.97
		20	0.8851	0.719~0.818	0.98	0.7036	0.537~0.866	0.99	0.6661	0.510~0.826	0.99	0.7125	0.547~0.879	0.99	0.5429	0.399~0.695	0.99
		30	0.8612	0.734~0.984	0.98	0.6910	0.559~0.886	0.99	0.6622	0.536~0.790	0.99	0.7095	0.574~0.845	0.99	0.5453	0.425~0.670	0.99
		40	0.8443	0.739~0.945	0.98	0.6829	0.571~0.793	0.99	0.6592	0.550~0.769	0.99	0.7069	0.590~0.823	0.99	0.5461	0.440~0.656	0.99
		50	0.8320	0.742~0.919	0.98	0.6769	0.578~0.775	0.99	0.6571	0.559~0.755	0.99	0.7048	0.600~0.808	0.99	0.5473	0.452~0.646	0.99
		100	0.7979	0.741~0.853	0.99	0.6615	0.594~0.728	0.99	0.6504	0.583~0.717	0.99	0.6981	0.627~0.768	0.99	0.5499	0.324~0.622	0.99
	Cir/Elip (Cases 31, 34, 37, 40, 43)	5	0.7484	0.414~1.064	0.89	0.5210	0.277~0.786	0.88	0.3952	0.175~0.658	0.85	0.5636	0.304~0.836	0.90	0.6148	0.324~N/A	0.74
		10	0.7601	0.540~0.962	0.98	0.5395	0.365~0.722	0.97	0.4198	0.245~0.617	0.95	0.5799	0.394~0.768	0.98	0.5810	0.412~N/A	0.91
		20	0.7357	0.597~0.866	0.99	0.5361	0.411~0.666	0.99	0.4334	0.298~0.578	0.99	0.5757	0.443~0.710	0.99	0.5490	0.440~N/A	0.96
		30	0.7205	0.613~0.822	0.99	0.5340	0.430~0.641	0.99	0.4397	0.326~0.559	0.99	0.5717	0.464~0.682	0.99	0.5365	0.447~0.629	0.98
		40	0.7102	0.620~0.795	0.99	0.5331	0.442~0.626	0.99	0.4428	0.343~0.547	0.99	0.5695	0.476~0.664	0.99	0.5304	0.452~0.610	0.99
		50	0.7026	0.624~0.776	0.99	0.5323	0.450~0.616	0.99	0.4446	0.354~0.538	0.99	0.5677	0.484~0.652	0.99	0.5273	0.457~0.599	0.99
		100	0.6818	0.631~0.730	0.99	0.5301	0.472~0.589	0.99	0.4486	0.383~0.515	0.99	0.5641	0.504~0.623	0.99	0.5220	0.471~0.574	0.99
Center	Square (Cases 16, 19, 22, 25, 28)	5	0.7819	0.448~1.113	0.78	0.6501	0.389~0.930	0.85	0.5619	0.301~0.848	0.76	0.5873	0.305~0.900	0.76	0.4378	0.213~0.691	0.77
		10	0.8644	0.613~1.114	0.92	0.6901	0.476~0.916	0.95	0.6539	0.425~0.894	0.89	0.6661	0.438~0.905	0.90	0.5007	0.316~0.702	0.90
		20	0.8728	0.703~1.037	0.98	0.6901	0.529~0.854	0.99	0.6811	0.517~0.850	0.97	0.6955	0.529~0.865	0.97	0.5275	0.386~0.678	0.97
		30	0.8566	0.725~0.981	0.99	0.6842	0.552~0.817	0.99	0.6804	0.547~0.815	0.98	0.6986	0.563~0.835	0.99	0.5346	0.414~0.660	0.98
		40	0.8424	0.735~0.944	0.99	0.6791	0.566~0.791	0.99	0.6768	0.563~0.790	0.99	0.6989	0.582~0.815	0.99	0.5388	0.433~0.649	0.99
		50	0.8306	0.739~0.918	0.98	0.6745	0.574~0.773	0.99	0.6727	0.573~0.772	0.99	0.6989	0.595~0.801	0.99	0.5410	0.445~0.641	0.99
		100	0.7974	0.740~0.853	0.99	0.6610	0.593~0.728	0.99	0.6607	0.593~0.727	0.99	0.6953	0.624~0.765	0.99	0.5470	0.477~0.619	0.99
	Cir/Elip (Cases 46, 49, 52, 55, 58)	5	0.6461	0.366~0.926	0.78	0.4492	0.238~0.683	0.77	0.3308	0.137~0.569	0.75	0.4798	0.247~0.727	0.77	0.4921	0.311~0.697	0.80
		10	0.7181	0.506~0.924	0.98	0.5055	0.337~0.687	0.91	0.3874	0.220~0.578	0.89	0.5405	0.358~0.731	0.91	0.5407	0.387~0.695	0.93
		20	0.7251	0.583~0.857	0.98	0.5235	0.399~0.651	0.97	0.4174	0.283~0.563	0.96	0.5597	0.429~0.694	0.97	0.5430	0.435~0.617	0.98
		30	0.7151	0.605~0.817	0.99	0.5261	0.423~0.632	0.99	0.4287	0.315~0.549	0.98	0.5613	0.453~0.670	0.99	0.5375	0.449~0.626	0.99
		40	0.7067	0.615~0.939	0.99	0.5272	0.437~0.620	0.99	0.4342	0.334~0.539	0.99	0.5619	0.468~0.656	0.99	0.5328	0.456~0.610	0.99
		50	0.7001	0.621~0.774	0.99	0.5275	0.445~0.611	0.99	0.4377	0.347~0.532	0.99	0.5615	0.477~0.646	0.99	0.5298	0.460~0.600	0.99
		100	0.6808	0.629~0.729	0.99	0.5281	0.469~0.587	0.99	0.4452	0.380~0.512	0.99	0.5608	0.501~0.620	0.99	0.5235	0.474~0.574	0.99

Table 10 Estimated Coefficients β with Length-to-Width Ratio of 2

Depot Location	Service Area Shape	n	R			D			C			2 σ			3 σ		
			β	Range	R ²	β	Range	R ²	β	Range	R ²	β	Range	R ²	β	Range	R ²
Random	Square (Cases 2, 5, 8, 11, 14)	5	1.0200	0.597~1.425	0.91	0.7860	0.412~1.196	0.89	0.7260	0.396~1.083	0.90	0.5801	0.307~0.915	0.89	0.7820	0.427~1.189	0.89
		10	1.0140	0.709~1.293	0.95	0.7908	0.501~1.085	0.95	0.7366	0.474~1.003	0.95	0.5917	0.368~0.847	0.95	0.7883	0.513~1.087	0.95
		20	0.9341	0.758~1.096	0.93	0.7458	0.559~0.924	0.96	0.7042	0.532~0.875	0.97	0.5755	0.415~0.749	0.98	0.7538	0.573~0.936	0.97
		30	0.8903	0.759~1.012	0.93	0.7188	0.576~0.855	0.96	0.6864	0.552~0.819	0.98	0.5665	0.437~0.705	0.99	0.7353	0.595~0.876	0.98
		40	0.8637	0.757~0.965	0.94	0.7027	0.584~0.816	0.97	0.6762	0.563~0.787	0.99	0.5622	0.450~0.681	0.99	0.7250	0.607~0.843	0.99
		50	0.8465	0.755~0.934	0.95	0.6923	0.589~0.791	0.98	0.6700	0.571~0.767	0.99	0.5596	0.459~0.665	0.99	0.7181	0.614~0.822	0.99
		100	0.8051	0.747~0.861	0.98	0.6687	0.600~0.734	0.99	0.6559	0.588~0.722	0.99	0.5557	0.483~0.629	0.99	0.7034	0.633~0.773	0.99
	Cir/Elip (Cases 32, 35, 38, 41, 44)	5	0.8524	0.495~1.231	0.90	0.4478	0.210~0.757	0.89	0.5927	0.327~0.909	0.90	0.6401	0.362~0.924	0.91	0.6383	0.359~0.961	0.90
		10	0.8471	0.587~1.105	0.94	0.4737	0.265~0.726	0.95	0.5880	0.376~0.824	0.93	0.6480	0.439~0.859	0.97	0.6437	0.422~0.886	0.95
		20	0.7851	0.630~0.930	0.94	0.5229	0.343~0.727	0.96	0.5680	0.422~0.721	0.97	0.5906	0.464~0.720	0.91	0.6137	0.466~0.767	0.97
		30	0.7523	0.637~0.859	0.95	0.5205	0.368~0.688	0.98	0.5573	0.442~0.675	0.98	0.5639	0.466~0.665	0.93	0.5981	0.481~0.719	0.98
		40	0.7330	0.639~0.821	0.96	0.5134	0.382~0.656	0.98	0.5502	0.451~0.649	0.99	0.5501	0.467~0.636	0.95	0.5892	0.490~0.691	0.98
		50	0.7202	0.639~0.796	0.96	0.5070	0.390~0.632	0.99	0.5458	0.459~0.633	0.99	0.5422	0.468~0.618	0.96	0.5831	0.496~0.673	0.99
		100	0.6902	0.638~0.739	0.98	0.4866	0.408~0.568	0.99	0.5370	0.479~0.597	0.99	0.5287	0.476~0.582	0.99	0.5716	0.511~0.632	0.99
Center	Square (Cases 17, 20, 23, 26, 29)	5	0.9132	0.534~1.297	0.82	0.7408	0.452~1.073	0.87	0.6416	0.347~0.980	0.82	0.5075	0.261~0.818	0.81	0.6862	0.364~1.054	0.80
		10	0.9541	0.653~1.238	0.91	0.7619	0.514~1.026	0.94	0.6990	0.448~0.972	0.91	0.5469	0.334~0.798	0.90	0.7324	0.459~1.020	0.90
		20	0.9184	0.737~1.087	0.95	0.7323	0.557~0.905	0.97	0.7645	0.566~0.961	0.95	0.5569	0.401~0.728	0.97	0.7316	0.548~0.914	0.97
		30	0.8829	0.748~1.008	0.96	0.7120	0.572~0.845	0.98	0.7509	0.595~0.902	0.97	0.5553	0.428~0.692	0.99	0.7237	0.580~0.864	0.99
		40	0.8599	0.751~0.963	0.96	0.6987	0.581~0.810	0.98	0.7328	0.604~0.856	0.97	0.5540	0.442~0.671	0.99	0.7167	0.596~0.834	0.99
		50	0.8438	0.750~0.932	0.97	0.6895	0.587~0.787	0.98	0.7184	0.607~0.824	0.97	0.5533	0.453~0.658	0.99	0.7118	0.606~0.815	0.99
		100	0.8044	0.746~0.860	0.98	0.6680	0.600~0.734	0.99	0.6799	0.610~0.747	0.98	0.5527	0.481~0.626	0.99	0.7003	0.629~0.966	0.99
	Cir/Elip (Cases 47, 50, 53, 56, 59)	5	0.7538	0.441~1.103	0.81	0.3895	0.169~0.682	0.79	0.5200	0.274~0.814	0.80	0.5669	0.342~0.829	0.82	0.5608	0.294~0.858	0.81
		10	0.7931	0.535~1.047	0.91	0.4274	0.231~0.667	0.89	0.5571	0.355~0.783	0.90	0.5934	0.410~0.788	0.91	0.5980	0.373~0.831	0.91
		20	0.7693	0.612~0.916	0.96	0.4455	0.297~0.614	0.97	0.5560	0.416~0.703	0.97	0.5750	0.454~0.701	0.97	0.5956	0.446~0.749	0.97
		30	0.7460	0.629~0.855	0.97	0.4496	0.326~0.583	0.98	0.5498	0.438~0.665	0.98	0.5599	0.465~0.659	0.97	0.5867	0.469~0.707	0.98
		40	0.7294	0.634~0.818	0.97	0.4520	0.346~0.565	0.99	0.5453	0.449~0.644	0.99	0.5502	0.469~0.876	0.98	0.5810	0.481~0.683	0.99
		50	0.7178	0.636~0.794	0.98	0.4524	0.356~0.553	0.99	0.5422	0.456~0.629	0.99	0.5435	0.472~0.618	0.98	0.5768	0.488~0.667	0.99
		100	0.6893	0.637~0.738	0.99	0.4532	0.385~0.522	0.99	0.5356	0.475~0.595	0.99	0.5301	0.478~0.583	0.99	0.5682	0.507~0.630	0.99

Table 11 Estimated Coefficients β with Length-to-Width Ratio of 4

Depot Location	Service Area Shape	n	R			D			C			2 σ			3 σ		
			β	Range	R ²	β	Range	R ²	β	Range	R ²	β	Range	R ²	β	Range	R ²
Random	Square (Cases 2, 5, 8, 11, 14)	5	1.2260	0.552~N/A	0.88	0.9569	0.404~1.610	0.90	0.8738	0.375~1.477	0.89	0.9408	0.413~1.562	0.89	0.7005	0.293~1.243	0.88
		10	1.2050	0.782~N/A	0.95	0.9522	0.561~N/A	0.97	0.8869	0.529~1.266	0.97	0.9486	0.575~1.348	0.97	0.7144	0.405~1.082	0.97
		20	1.0840	0.858~N/A	0.87	0.8739	0.631~N/A	0.92	0.8268	0.599~N/A	0.95	0.8847	0.644~N/A	0.95	0.6779	0.485~0.914	0.97
		30	1.0030	0.851~N/A	0.81	0.8189	0.642~N/A	0.89	0.7819	0.615~N/A	0.93	0.8370	0.658~N/A	0.93	0.6496	0.481~0.830	0.96
		40	0.9506	0.834~N/A	0.80	0.7830	0.641~N/A	0.88	0.7522	0.619~0.879	0.92	0.8049	0.661~0.937	0.92	0.6299	0.490~0.779	0.96
		50	0.9149	0.818~1.001	0.81	0.7583	0.639~0.866	0.89	0.7311	0.618~0.839	0.92	0.7827	0.661~0.894	0.92	0.6166	0.494~0.745	0.96
		100	0.8355	0.776~0.891	0.90	0.7014	0.628~0.770	0.94	0.6840	0.613~0.752	0.96	0.7325	0.658~0.802	0.96	0.5859	0.668~0.505	0.98
	Cir/Elip (Cases 32, 35, 38, 41, 44)	5	1.0320	0.465~1.655	0.90	0.7215	0.320~1.213	0.87	0.5411	0.201~1.004	0.86	0.7755	0.073~1.312	0.87	0.7716	0.360~1.233	0.88
		10	1.0220	0.637~N/A	0.97	0.7272	0.433~1.057	0.96	0.5665	0.291~0.905	0.96	0.7800	0.464~1.129	0.96	0.7627	0.483~1.052	0.96
		20	0.9317	0.717~N/A	0.91	0.6797	0.488~0.886	0.95	0.5492	0.349~0.782	0.98	0.7271	0.527~N/A	0.95	0.6949	0.527~N/A	0.90
		30	0.8670	0.718~N/A	0.85	0.6446	0.497~0.799	0.93	0.5339	0.372~0.714	0.98	0.6888	0.536~0.847	0.93	0.6483	0.524~N/A	0.86
		40	0.8244	0.708~N/A	0.84	0.6216	0.499~0.748	0.93	0.5210	0.384~0.671	0.98	0.6633	0.538~0.792	0.92	0.6192	0.518~0.727	0.86
		50	0.7952	0.700~N/A	0.84	0.6053	0.499~0.738	0.93	0.5118	0.391~0.641	0.98	0.6454	0.537~0.756	0.93	0.5995	0.512~0.693	0.87
		100	0.7277	0.671~0.780	0.91	0.5686	0.500~0.637	0.96	0.4877	0.407~0.570	0.98	0.6046	0.535~0.674	0.96	0.5570	0.498~0.618	0.94
Center	Square (Cases 17, 20, 23, 26, 29)	5	1.0690	0.500~1.698	0.80	0.8909	0.465~1.411	0.85	0.7543	0.330~1.296	0.80	0.7780	0.329~1.365	0.75	0.5948	0.240~1.083	0.77
		10	1.1380	0.726~1.508	0.93	0.9133	0.588~1.280	0.96	0.8274	0.485~1.203	0.93	0.7943	0.428~1.235	0.92	0.6634	0.369~1.020	0.91
		20	1.0620	0.836~1.251	0.93	0.8573	0.630~1.078	0.95	0.8029	0.578~1.028	0.97	0.7634	0.491~1.059	0.96	0.6572	0.451~0.891	0.97
		30	0.9922	0.838~1.122	0.89	0.8100	0.638~0.971	0.93	0.7691	0.601~0.932	0.96	0.7323	0.505~0.964	0.96	0.6375	0.474~0.818	0.97
		40	0.9441	0.959~1.046	0.87	0.7775	0.638~0.906	0.92	0.7437	0.608~0.873	0.95	0.7100	0.510~0.907	0.96	0.6218	0.483~0.771	0.97
		50	0.9108	0.814~0.997	0.86	0.7543	0.636~0.863	0.91	0.7249	0.610~0.834	0.95	0.6940	0.513~0.869	0.96	0.6102	0.488~0.740	0.97
		100	0.8351	0.775~0.890	0.92	0.7001	0.627~0.768	0.95	0.6818	0.610~0.750	0.97	0.6567	0.519~0.786	0.97	0.5831	0.502~0.666	0.98
	Cir/Elip (Cases 47, 50, 53, 56, 59)	5	0.8912	0.409~1.454	0.79	0.6125	0.263~1.077	0.79	0.5471	0.210~0.990	0.88	0.6571	0.266~1.156	0.79	0.6706	0.329~1.096	0.78
		10	0.9587	0.597~1.318	0.93	0.6753	0.392~1.002	0.92	0.5725	0.333~0.853	0.91	0.7207	0.411~1.066	0.93	0.7180	0.455~1.004	0.92
		20	0.9088	0.692~1.105	0.95	0.6588	0.466~0.867	0.97	0.5430	0.354~0.758	0.96	0.7046	0.501~0.923	0.97	0.6805	0.517~0.856	0.95
		30	0.8558	0.704~0.991	0.92	0.6335	0.485~0.789	0.96	0.5273	0.371~0.701	0.97	0.6760	0.522~0.837	0.96	0.6430	0.521~0.775	0.93
		40	0.8176	0.700~0.923	0.90	0.6144	0.491~0.742	0.96	0.5158	0.381~0.662	0.97	0.6539	0.527~0.785	0.95	0.6170	0.516~0.726	0.92
		50	0.7908	0.694~0.878	0.90	0.6002	0.493~0.709	0.96	0.5071	0.388~0.635	0.98	0.6384	0.529~0.750	0.95	0.5994	0.512~0.693	0.92
		100	0.7264	0.670~0.779	0.93	0.5668	0.498~0.636	0.97	0.4852	0.405~0.568	0.98	0.6014	0.532~0.671	0.97	0.5585	0.500~0.620	0.95



(a) Changes in Coefficients and R^2 with n

(b) Percent Change with n

Figure 9. Investigation for Estimated β

Figure 9 (a) shows all coefficients β and goodness of fit values (R^2) estimated with different n values (i.e., with different numbers of intervals) for Case 1. Adjusted R^2 increases as n increases. The coefficients have an uptrend before $n = 8$, and then the estimated β decreases between 8 and 100 points Figure 9 (b) shows details of relative percent changes for this reversal. The relative percent changes in estimated β decrease with n . Beyond $n = 63$, the changes are below 0.01%.

Table 12 Comparison of Exact and Estimated Tour Lengths

β	n	Exact Solution	Estimated Tour Length	MAPE (%)
0.9094	2	1.04	1.2861	23.66
	3	1.56	1.5751	0.97
0.9194	2	1.04	1.3002	25.02
	3	1.56	1.5924	2.08
0.8320	2	1.04	1.1766	13.14
	3	1.56	1.4411	7.62
0.7979	2	1.04	1.1284	8.50
	3	1.56	1.3820	11.41

Since the exact tour lengths for visiting 2 and 3 points can be derived analytically, the estimated distance in Table 12 can be compared with the exact tour lengths presented. The Mean Absolute Percentage Error (MAPE) suggests that the approximation models could

effectively explain the TSP tour lengths when $n \geq 3$. Across the different scenarios developed in Section 3.2.1, the lowest MAPE overall is found at β of 0.7979 in Table 8. For the exact distance of three points ($n = 3$), β of 0.9094 provides the best solutions.

3.4 Summary

Using the optimized TSP instances, approximation models are developed with an OLS regression. The models consider the various scenarios, such as depot location, distance metrics, service area shapes, and point distributions.

Chapter 4: Extensions of Tour Length Approximation: Adjustment Factors, Probabilistic Tour Length Approximation, and Comparison on Approximated versus Actual Road Network Distance

This chapter introduces various extensions for the TSP tour length approximation. First, adjustment factors that integrate various considerations into a single equation are developed for easy use of the proposed approximations. Assuming that the number of visited points is preset and only a subset of the points is visited based on a probabilistic distribution (e.g., uniform distribution), approximations are designed for such probabilistic TSP's (P-TSP). Lastly, the approximation results from Chapter 3 are applied to estimate the tour lengths for rural freight delivery and urban package delivery networks. After the actual and estimated tour lengths are compared, findings and implications are discussed.

4.1 Adjustment Factors for Approximations

To conveniently use the approximation coefficient β , adjustment factors are designed to integrate various considerations from Chapter 3 within one equation. With the factors, it can be understood how sensitively the estimated β varies with each classification. The coefficient β for one classification can be converted to another using Equation (9).

$$\begin{cases} L \cong \beta\sqrt{nA} \\ \beta = D_0 \cdot D_1 \cdot D_2 \cdot D_3 \cdot D_4 \cdot D_5 \end{cases} \quad (9)$$

where D_0 is an adjustment factor associated n values for the coefficient for a random depot in a square service area (Case 1), D_1 is a random-to-center conversion factor, D_2 is, a square-to-

circular conversion factor, D_3 and D_4 are adjacent factors for point distribution, and D_5 is an elongation (i.e., a length-to-width ratio) adjustment factor

In addition to the above adjustment factors, the finding in Krarup and Pruzan (1980) showed that the average ratio of distances between Euclidean and rectilinear space is fixed as 1.26.

4.1.1 Curve Fitting Methods and Computation Steps

Curve fitting methods are used to present the best fit of given data points. First, the relative ratios (1) between coefficients β in Table 8 or (2) between optimized TSP tour lengths are computed for one to another classification. For D_1 , the coefficients β for Case 19 (i.e., central depot) are divided by β of the baseline Case 1 (i.e., random depot). Since the adjustment factors for bivariate normal distributions D_4 are varied according to the standard deviation (σ), the tour lengths with different standard deviations (ranging from 0.01 to 0.5) are generated. Then, the lengths are normalized (i.e., divided by the TSP tour lengths generated on uniform and random distributions). For elongation D_5 , a length-to-width ratio x is introduced as a variable. Note that the range for σ ranges between 0.01 and 0.5, while x is between 1 and 4.

After all fractions are computed, various curve fitting methods (e.g., exponential, polynomial, and power) are applied. The adjustment factors with the highest goodness of fit are chosen, as presented in Equations (10) – (15).

$$D_0 = \begin{cases} -0.0040 * n^2 + 0.0563 * n + 0.7285 & (\text{if } n \leq 8) \\ 1.0580 * n^{-0.0616} & (\text{if } n > 8) \end{cases} \quad (10)$$

$$D_1 = \begin{cases} 0.8337 * n^{0.0462} & (\text{if centered}) \\ 1 & (\text{otherwise}) \end{cases} \quad (11)$$

$$D_2 = \begin{cases} 0.8034 * n^{0.0126} & (\text{if circular/elliptical}) \\ 1 & (\text{otherwise}) \end{cases} \quad (12)$$

$$D_3 = \begin{cases} 0.7381 * n^{0.0249} & \text{(if centralized)} \\ 0.6575 * n^{0.0464} & \text{(if declining from corner)} \\ 1 & \text{(otherwise)} \end{cases} \quad (13)$$

$$D_4 = \begin{cases} 1.6310 * \sigma^{0.5491} & \text{(if normal)} \\ 1 & \text{(otherwise)} \end{cases} \quad (14)$$

$$D_5 = \begin{cases} x^{0.0319} & \text{(if elongated)} \\ 1 & \text{(otherwise)} \end{cases} \quad (15)$$

In Equation (10), the approximated TSP tour length is based on Case 1. In Figure 10 (c), D_0 is divided into two segments at $n = 8$ for the best fit, where an inflection point for coefficients β is found at that n value in Figure 10. From the literature and Table 8, β does not converge to a specific value as n increases. After adding Equation (10), the square root of n no longer holds. This result implies that finding the best exponent of n could improve the approximation accuracy. Equations (10) – (15) are fitted, and the fitting results are described in Table 13 and 14. For D_5 in Equation (15), the tour lengths are not significantly increased within the length-to-width range of 1 and 4, while the unit area is unchanged. The tour lengths increase by 1.1% at $x = 1.4$ and 4.5% at $x = 4$.

Although estimation with the highest R^2 is often regarded as preferable, this study also considers: (1) ease of use, (2) intuitive explanation, (3) overfitting, and (4) reasonable value of the goodness of fit. If the R^2 is not significantly different between the highest and second highest, it would be better to choose a simpler method mainly because of computation and convenience. The same rule applies to other goodness of fit measures, including the sum of square error (SSE) and root mean square error (RMSE). Besides, the estimated adjustment factors should show a clear relation between n and β . Since R^2 can only increase when more variables (e.g., the number of variables or the number of terms) are added, a simpler method would be preferable to avoid overfitting estimations.

Table 13 Curve Fitting for Adjustment Factor Associated n Values (D_0)

Formula		Model Type			
		Exponential	Polynomial	Hybrid	
				Polynomial	Power
Formula		$a \cdot \exp(b \cdot n)$	$a \cdot n + b$	$a \cdot n^2 + b \cdot n + c$	$a \cdot n^b$
Coefficient	a	0.9067	-0.0012	-0.0040	1.0580
	b	-0.0015	0.9042	0.0563	-0.0616
	c	-	-	0.7285	-
Goodness of fit	SSE	0.0104	0.0116	0.0000	0.0002
	RMSE	0.0105	0.0111	0.0020	0.0015
	Adj R^2	0.9185	0.9042	0.9906	0.9980

In Table 13, all the adjusted R^2 exceeding 0.90. Based on the considerations listed earlier, the estimation from the power method is chosen.

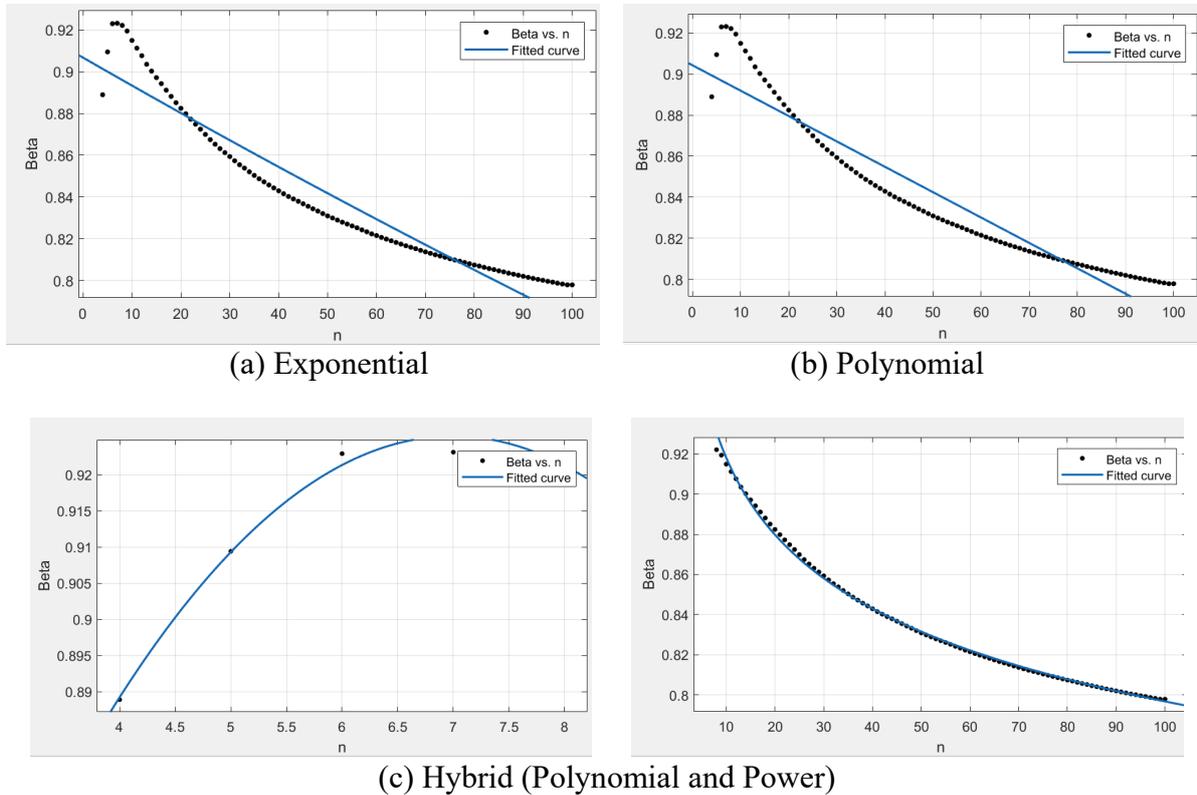


Figure 10 Curve Fitting for Adjustment Factors (D_0)

Table 14 Curve Fitting Results for Equations (10) – (14)

		D_1	D_2	D_3		D_4	D_5
				Clustered	Dispersed	Normal	L/W
Formula		$a \cdot n^b$	$a \cdot n^b$	$a \cdot n^b$	$a \cdot n^b$	$a \cdot \sigma^b$	x^b
Coefficient	a	0.8337	0.8034	0.7381	0.6575	1.6310	-
	b	0.0462	0.0126	0.0249	0.0464	0.5491	0.0319
Goodness of fit	SSE	0.0039	0.0000	0.0000	0.0000	0.0610	0.1617
	RMSE	0.028	0.0026	0.0011	0.0036	0.1105	0.4022
	Adj. R^2	0.71	0.94	0.99	0.99	0.91	0.96

In Table 13 and 14, most relations are well fitted with the power method.

4.1.2 Validation of Adjustment Factors

Table 15 displays absolute percent errors using the proposed β (referring to Table 8), while adjustment errors are percent differences of average tour lengths of 1,000 TSPs and adjustment factors ($D_0, D_1, D_2, D_3, D_4,$ and D_5).

Table 15 Percent Adjustment Error (1/2)

n	n values		Random-Center		Square-Circular		Declining from corner		Centralized		2- σ		3- σ		L/W: 2		L/W: 4	
	Est. β	Eq. (10)	Est. β	Eq. (11)	Est. β	Eq. (12)	Est. β	Eq. (13)	Est. β	Eq. (13)	Est. β	Eq. (15)	Est. β	Eq. (15)	Est. β	Eq. (14)	Est. β	Eq. (14)
2	19.36	11.12	30.21	26.11	18.32	8.29	19.45	8.83	19.78	5.60	18.24	19.26	19.21	17.63	21.57	5.16	19.88	13.90
3	0.76	6.38	6.91	6.99	1.11	8.15	0.86	6.38	0.72	9.58	1.35	6.68	1.49	3.79	3.63	11.62	0.34	35.72
4	4.05	6.35	6.76	2.20	4.21	7.51	3.74	6.99	3.68	9.15	2.83	5.82	3.98	1.68	1.92	14.29	4.48	37.75
5	4.06	3.99	10.83	2.88	3.29	3.87	5.30	5.58	5.74	7.58	5.82	5.24	5.40	2.56	5.30	15.45	4.68	34.94
6	2.22	1.90	3.04	3.64	1.80	2.34	1.68	2.69	1.98	4.41	1.61	7.63	1.27	3.71	4.02	11.86	4.15	30.20
7	1.74	0.95	3.82	3.36	1.64	1.50	2.02	2.15	2.71	3.83	1.93	7.76	1.11	4.31	5.04	12.44	1.88	26.77
8	0.48	0.10	2.55	2.00	0.74	0.83	0.71	0.90	0.71	1.51	1.59	7.70	1.46	3.59	3.76	11.52	0.23	25.22
9	0.13	0.64	3.30	2.18	0.15	0.36	0.03	0.27	1.03	1.09	1.55	7.86	2.19	3.03	3.03	4.08	2.48	21.68
10	1.28	1.14	3.83	2.97	0.53	0.22	0.56	0.70	0.57	0.42	1.44	7.36	1.03	3.50	5.13	2.49	4.35	20.12
20	3.22	2.63	1.24	0.63	2.71	2.34	2.15	1.52	0.37	0.15	0.51	5.18	0.40	0.51	5.58	2.62	8.87	7.44
30	3.10	2.74	2.69	1.69	2.16	1.91	2.04	1.74	0.83	0.51	0.73	3.00	0.91	4.87	5.44	4.42	9.30	1.44
40	3.26	3.10	3.04	2.98	2.49	2.29	1.78	1.62	0.87	0.73	0.93	1.89	0.20	6.15	4.86	4.48	9.08	1.89
50	3.36	3.30	3.29	4.04	2.37	2.17	2.15	2.05	0.66	0.62	0.04	0.53	0.31	7.97	4.25	4.60	7.91	3.05
60	0.65	3.58	0.13	4.44	0.27	2.16	0.22	1.76	0.07	0.74	0.03	0.58	0.57	9.99	0.46	4.66	1.37	4.10
70	1.12	3.12	1.17	5.07	0.70	1.86	0.75	1.73	0.17	0.82	0.74	0.75	0.06	10.48	1.12	4.37	2.66	4.45
80	1.66	2.85	1.57	5.14	1.12	1.64	0.80	1.31	0.48	1.00	0.63	1.70	0.10	11.43	1.97	4.35	3.62	4.61
90	2.23	2.71	2.34	5.60	1.66	1.63	0.84	0.92	0.43	0.84	0.67	2.40	0.15	12.30	2.64	4.39	4.26	4.55
100	2.71	2.56	2.58	5.58	1.91	1.38	1.27	0.97	0.53	0.84	0.82	2.91	0.47	13.40	2.91	4.03	4.90	4.58
Avg.	1.94	2.95	2.14	3.91	1.52	2.10	1.29	1.72	0.75	1.08	0.81	2.75	0.66	7.68	2.81	4.54	4.44	6.65

* L/W implies a length-to-width ratio of service area

The overall absolute percent errors for the proposed β are lower than for the adjustment factors. Although large percent errors are observed for $n = 2$, the percent errors for the proposed β generally decrease. For elongated service areas, large errors are found in proposed models and adjustment factors.

Table 16 Percent Adjustment Error (2/2)

<i>n</i>	Case 27 (Square, Center, L/W: 4, 2 Sigma)		Case 28 (Square, Center, L/W: 1, 3 Sigma)		Case 34 (Circular, Random, L/W: 1, Centralized)		Case 35 (Elliptical, Random, L/W: 2, Centralized)		Case 53 (Elliptical, Center, L/W: 2, Declining)		Case 58 (Circular, Center, L/W: 1, 3 Sigma)		Case 60 (Elliptical, Center, L/W: 4, 3 Sigma)	
	Est. β	Eqs. (10), (11), (14), (15)	Est. β	Eqs. (10), (11), (15)	Est. β	Eqs. (10), (11)	Est. β	Eqs. (10), (12), (13), (14)	Est. β	Eqs. (10), (11), (13), (14)	Est. β	Eqs. (10), (11), (12), (14), (15)	Est. β	Eqs. (10), (11), (12), (14), (15)
2	30.57	4.49	31.59	31.05	23.21	39.26	12.12	13.69	27.38	18.36	30.20	5.87	33.29	24.69
3	1.18	27.80	18.73	11.93	2.02	26.92	37.75	0.04	2.70	5.96	3.04	22.75	1.61	72.02
4	4.52	17.97	11.42	3.38	4.94	25.06	48.09	3.01	15.56	12.00	5.42	27.37	5.07	74.86
5	16.55	39.34	8.71	2.35	6.53	26.26	54.97	4.51	25.17	15.86	9.56	27.98	10.22	77.01
6	0.26	19.11	1.18	0.83	1.01	27.23	46.06	2.54	14.61	15.55	2.68	28.86	2.52	72.08
7	2.30	20.74	2.72	0.50	1.33	27.77	48.83	3.39	16.43	14.96	2.70	27.31	4.29	72.66
8	1.39	19.41	3.77	0.30	0.68	28.32	46.37	1.58	18.16	15.50	3.37	27.79	4.53	72.39
9	0.22	17.24	6.50	2.26	3.33	26.86	38.98	4.09	12.59	8.56	3.45	26.98	1.16	65.48
10	0.13	17.33	6.08	2.02	3.13	26.83	38.40	4.25	12.06	7.56	3.26	26.91	0.03	63.68
20	4.41	8.81	2.55	5.02	1.34	24.59	30.15	5.48	5.05	1.31	1.06	23.11	5.44	46.92
30	5.48	3.84	1.33	5.83	0.88	23.12	26.13	6.15	3.08	1.85	1.51	21.88	6.82	36.96
40	5.19	1.44	1.22	7.01	1.16	21.86	25.09	5.62	1.93	3.95	1.54	21.19	6.90	31.42
50	4.75	0.04	0.89	7.46	0.87	21.36	24.58	5.11	1.35	5.37	1.14	21.30	6.18	28.74
60	0.67	1.08	0.28	8.30	0.49	20.63	23.63	4.57	1.13	6.21	0.02	21.53	0.65	27.09
70	1.66	1.84	0.49	8.79	0.64	20.22	23.44	4.36	1.37	6.62	0.33	21.32	1.92	25.51
80	2.21	2.19	0.58	9.11	0.32	20.24	23.42	4.07	0.60	7.86	0.61	21.17	3.02	24.16
90	2.80	2.60	0.55	9.28	0.87	19.59	23.16	4.00	0.84	8.12	0.36	21.64	3.23	23.92
100	3.08	2.72	0.91	9.84	0.61	19.60	23.30	3.65	0.57	8.79	0.58	21.53	3.86	23.15
Avg	2.84	4.91	1.56	7.25	1.05	22.22	26.72	4.60	3.40	6.01	1.22	22.10	3.62	34.53

Table 16 shows adjustment errors when the adjustment factors are used in combination according to Equation (9). Due to the large errors, it should be recommended that no more than two factors be used at a time, namely D_0 and any other adjustment factor in Equation (9).

The factors for bivariate normal distributions and elongated service areas produce the large absolute average percent errors among the six factors. They apply similarly to the cases if multiple factors are jointly used. Although some limitations exist for their use, these factors are valuable for understanding the relations between the TSP tour lengths and each classification. For instance, planners can roughly estimate how much the actual point distribution may affect vehicle miles traveled (e.g., whether it leads to minor or huge changes).

4.1.3 Comparison with Other Tour Length Approximations

This section computes the approximated TSP tour lengths with different existing approximation models and compares them with average tour lengths of 1,000 TSP instances.

The absolute percent difference is used for the comparison. The approximation models include Daganzo (1984), Hindle and Worthington (2004), and Cavdar and Sokol (2015). Daganzo’s approximation is one of the earliest ones and remains similar to Beardwood’s equation, while other approximations have some variations, as shown in Equations (2) and (3). In particular, Cavdar and Sokol (2015) can estimate average TSP tour length without having the exact distribution of points or service area shape. Note that comparison results in some applications are introduced in Section 5.3.2.

Table 17 Comparison of Percent Differences Using Existing Approximation Models

n	n values			Center			Circular			L/W: 4			3-σ			Other Cases											
	Case 1			Case 16			Case 31			Case 3			Case 13			Case 27			Case 35			Case 60					
	C&S	H&W	D	C&S	H&W	D	C&S	H&W	D	C&S	H&W	D	C&S	H&W	D	C&S	H&W	D	C&S	H&W	D	C&S	H&W	D	C&S	H&W	D
5	28.00	3.44	38.13	18.14	12.96	32.44	7.51	26.63	24.26	64.87	58.54	75.20	15.62	80.51	7.97	77.75	41.61	65.07	17.03	70.17	1.79	71.01	29.04	57.55	21.01	29.04	57.55
10	23.11	3.30	46.88	19.62	4.46	46.28	1.44	24.00	36.23	58.88	54.99	76.85	21.54	73.03	11.01	72.92	42.87	70.62	18.94	66.31	14.47	63.75	28.61	63.29	28.61	63.29	
20	17.48	2.39	54.02	16.88	1.76	54.30	5.65	22.54	44.97	47.72	47.23	76.30	23.13	62.21	27.16	65.67	37.54	71.95	15.52	66.14	25.39	53.27	20.99	64.52	20.99	64.52	
30	15.19	1.37	57.90	14.35	1.48	57.86	7.25	20.00	50.17	38.94	40.62	75.34	21.41	53.73	36.16	60.68	32.32	71.90	13.84	63.70	32.02	45.87	13.39	64.03	13.39	64.03	
40	13.15	1.49	60.40	12.46	1.50	60.40	9.56	19.71	53.30	31.66	35.10	74.68	22.38	51.50	40.89	56.57	27.48	71.71	10.80	67.98	34.46	40.15	7.43	63.89	7.43	63.89	
50	11.74	1.97	62.25	11.15	2.06	62.22	10.59	19.53	55.76	25.96	30.55	74.29	22.24	49.34	44.72	52.98	22.90	71.46	11.01	66.15	38.50	35.45	1.81	63.65	1.81	63.65	
60	10.44	2.87	63.68	10.56	2.40	63.85	11.65	19.93	57.66	20.99	26.34	73.99	21.64	47.47	47.94	50.23	19.11	71.44	10.88	66.83	41.10	32.20	2.42	63.84	2.42	63.84	
70	10.04	3.20	64.97	9.57	3.32	64.93	12.33	20.26	59.18	17.05	22.62	73.73	22.15	47.98	49.77	47.56	15.54	71.33	9.18	68.96	42.65	28.05	7.76	63.42	7.76	63.42	
80	9.52	3.86	66.03	9.22	3.84	66.04	12.75	20.89	60.46	13.23	19.18	73.56	22.36	48.07	51.57	45.52	12.13	71.26	8.96	70.14	44.35	25.19	12.10	63.33	12.10	63.33	
90	8.95	4.77	66.92	8.47	4.95	66.86	13.40	21.90	61.51	10.25	16.16	73.53	22.15	48.41	53.14	43.58	9.07	71.29	8.31	71.76	45.76	23.00	15.76	63.45	15.76	63.45	
100	8.91	5.72	67.70	8.27	5.65	67.72	13.56	22.72	62.50	7.77	13.45	73.56	21.95	48.54	54.61	42.13	6.37	71.39	7.77	73.82	46.89	20.78	19.10	63.61	19.10	63.61	
Avg.	13.69	3.19	59.46	12.22	4.26	59.12	9.77	21.58	52.23	29.57	32.38	74.49	21.80	54.68	40.66	55.34	23.54	70.72	11.89	68.54	35.58	38.89	12.74	62.88	12.74	62.88	

* C&S: Cavdar and Sokol (2015), H&W: Hindle and Worthington (2004), and D: Daganzo (1984)

Table 17 shows the absolute percent differences when applying each approximation model to each case with the 1,000 TSP instances. Cavdar and Sokol (2015) indicate that the model tends to underestimate the tour lengths when $n < 1,000$. For the estimates by Cavdar and Sokol, Table 17 presents the decrease in the percent differences as n increases. It may be noted that the “absolute” percent differences decrease beyond $n = 20$ for the circular area. The approximation from Hindle and Worthington (2004) produces good estimates for Cases 1 and 16, where the cases have similar experiment settings, such as a square area with uniform

distribution. However, the approximation yields poor results for other cases. Lastly, the estimation results from Daganzo’s approximation show the largest differences.

Table 18 Estimation Results for Small/Large n Values Using Adjustment Factors

n	Optimized Length	Approximated Length	% Error	Replication
200	10.7339	11.0754	3.18	100
300	12.9704	13.3101	2.62	100
400	14.8995	15.1640	1.78	100
500	16.5246	16.7781	1.53	100
600	18.0356	18.2237	1.04	100
700	19.5046	19.5427	0.20	100
800	20.7823	20.7621	-0.10	100
900	22.0045	21.9008	-0.47	100
1,000	23.1587	22.9721	-0.81	100
2,000	32.4540	31.4527	-3.09	10
3,000	39.6515	37.7990	-4.67	10
4,000	45.7128	43.0640	-5.79	10
5,000	51.0360	47.6479	-6.64	10
6,000	55.8877	51.7532	-7.40	10
7,000	60.2507	55.4988	-7.89	10
8,000	64.4203	58.9619	-8.47	10
9,000	68.2225	62.1955	-8.83	10
10,000	71.8250	65.2380	-9.17	10

Although the scope of this dissertation covers small n values, Table 18 is designed for the applications of adjustment factors to large n values. With the second column of Table 15 combined, the percent error decreases until $n = 10$. After $n = 100$, the errors keep decreasing. In brief, it is shown that the adjustment factors can estimate the tour lengths for up to 2,000 points within a reasonable percent error range.

4.2 Tour Length Approximation with Stochastic Customer Presence: Probabilistic Traveling Salesman Problem

Jaillet (1985) introduced a probabilistic traveling salesman problem (P-TSP); a probabilistically chosen subset of k points is visited from n known points (i.e., $0 \leq k \leq n$). With this feature, stochastic customer presence (or customer’s acceptance of the service) can be considered in the conventional TSP.

Consider that some points on the optimized TSP tour are absent or unavailable. In Jaillet (1985), the sequence of visiting the points along the optimized tour is preserved instead of re-optimizing the TSP instance with the remaining points. Here, the optimized TSP tour for visiting all points n should be determined before how many points k are chosen. For instance, delivery workers whose daily demands and delivery routes are fixed do not visit some of the preassigned delivery points, possibly due to the absence of the recipient from home for attended delivery or due to the lack of any demand at some points during a particular tour. Those points are removed, and then the route is optimized while maintaining the previous sequence of visits. The P-TSP can be helpful for analyzing such cases with uncertain demands.

In this section, the tour length approximations for P-TSP are developed by introducing the probability p that a pre-located point is actually visited during a tour. However, the detailed steps for computing P-TSPs would be different from Jaillet's original proposal. Preserving the visiting sequence is intended to reduce the computation times, which is no longer an interest of this dissertation. More importantly, the P-TSP solution obtained from the remaining sequence of visits does not guarantee an optimal solution (Wissink 2019). Thus, the TSPs are re-optimized without using the preserved sequence of orders from a prior solution.

4.2.1 Simulation Design and Result

Simulation settings similar to those in Section 3.2 are applied for the P-TSP instances. The instances are in Euclidean space, where the points are uniformly and randomly distributed over a unit square. n value ranges from 10 to 100 with an increment of 10. The uniform distribution is selected to represent the probability of being visited p , where p is a mean varying from 0.1 to 1.0 with an increment of 0.1. If a sample size is large enough, the average TSP tour lengths (i.e., as random variables from the same distribution) depend only on p according to the central

limit theorem, regardless of the range of uniform distribution (i.e., the minimum and maximum values).

Table 19 Estimated Tour Lengths with Different Range of Uniform Distribution

Case	Uniform Distribution			$n = 20$			$n = 40$			$n = 60$			$n = 80$			$n = 100$		
	Min	Mean	Max	k	Tour Length	% Error	k	Tour Length	% Error									
1	0.77	0.88	0.99	17.71	3.6386	0.36	35.10	4.8656	0.27	52.44	5.8994	0.68	70.94	6.6508	0.26	88.20	7.3319	0.08
2	0.80	0.88	0.96	17.53	3.6342	0.49	35.23	4.8883	0.20	52.83	5.8382	0.37	70.32	6.6092	0.37	88.20	7.3330	0.10
3	0.84	0.88	0.92	17.58	3.6287	0.64	35.16	4.8712	0.15	52.78	5.8420	0.30	70.28	6.6031	0.46	88.62	7.3505	0.34
4	0.85	0.88	0.91	17.60	3.6534	0.04	35.19	4.8803	0.04	52.79	5.8198	0.68	70.26	6.6148	0.29	88.09	7.3234	0.03
5	0.87	0.88	0.89	17.74	3.6327	0.53	35.11	4.9001	0.44	52.77	5.8341	0.44	70.42	6.6222	0.17	87.98	7.3438	0.25
6	0.88	0.88	0.88	18.00	3.6519	0.00	35.00	4.8786	0.00	53.00	5.8596	0.00	70.00	6.6338	0.00	88.00	7.3259	0.00

In Table 19, the average TSP tour lengths are estimated from 1,000 TSP instances for each n value (i.e., the same $p = 0.88$ with different ranges of uniform distribution). Compared to the tour length for Case 6 with the other five cases, all the absolute percent errors listed in Table 19 are below 1%.

Table 20 Average TSP Tour Lengths for Various Probabilities

n	p									
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
10	N/A	1.0436	1.5874	1.9196	2.1897	2.3951	2.6781	2.7800	2.9454	3.1149
20	1.0161	1.9428	2.4109	2.7877	2.8635	3.0855	3.2840	3.4750	3.6585	3.8308
30	1.5673	2.4492	2.7554	3.0883	3.3905	3.6818	3.9235	4.1187	4.3697	4.5707
40	1.9390	2.7876	3.0993	3.4839	3.8063	4.1279	4.4246	4.6870	4.9324	5.1658
50	2.2100	2.8821	3.3952	3.8186	4.2230	4.5738	4.8735	5.1692	5.4306	5.6852
60	2.4437	3.0834	3.6521	4.1448	4.5663	4.9372	5.2828	5.5960	5.8866	6.1404
70	2.6351	3.2949	3.9066	4.4281	4.8769	5.2839	5.6351	5.9820	6.3006	6.6008
80	2.7875	3.4950	4.1400	4.6867	5.1723	5.5923	5.9824	6.3628	6.6987	7.0184
90	2.9937	3.6565	4.3477	4.9365	5.4381	5.8916	6.3063	6.6981	7.0546	7.4007
100	3.1182	3.8442	4.5531	5.1576	5.6979	6.1860	6.6039	7.0061	7.3986	7.7627

After a total number of 1,000 replications for each n value is run across all p values, the optimized TSPs are averaged as presented in Table 20.

4.2.2 Curve Fitting Result and Validation

The average tour lengths for unit squares in Table 20 are fitted to derive relation for n and p using Equation (16).

$$L \cong \beta \cdot n^a \cdot p^b \quad (16)$$

where p is a probability of being visited, while a and b are estimated exponents.

Table 21 Estimators for Equation (16)

	Unrestricted Case	Restricted Case
β	1.0213	0.8132
a	0.4386	0.5
b	0.4204	0.4696
R^2	0.9936	0.9763

While estimating the exponents and coefficient β from curve fitting, the two cases are designed. For the unrestricted case, the estimators (i.e., β , a , and b) can have any value. For the restricted case, the exponent a is forced to be 0.5 (i.e., Beardwood's formula). The estimators for both cases are listed in Table 21. The coefficient β is about 26.7% higher for the unrestricted case than for the restricted one, while R^2 is slightly worsened in the restricted case due to the reduced degree of freedom. Note that β and exponent a are interrelated.

Using the estimators, the TSP tour lengths for both cases are estimated in Table 22.

$$L \cong 1.0213 \cdot n^{0.4386} \cdot p^{0.4204} \quad (17)$$

The unrestricted case in Equation (17) also confirms that the tour lengths can be better estimated with statistical estimation for the exponent of n values and relevant coefficient β .

Table 22 Percent Errors Using P-TSP Results

p	n	Unrestricted			Restricted		
		App. Tour Length	Avg. TSPs	% Error	App. Tour Length	Avg. TSPs	% Error
0.2	20	1.9317	1.9428	0.57	1.7080	1.9428	12.09
	40	2.6180	2.7876	6.08	2.4154	2.7876	13.35
	50	2.8872	2.8821	0.18	2.7005	2.8821	6.30
	100	3.9130	3.8442	1.79	3.8191	3.8442	0.65
0.4	10	1.9075	1.9196	0.63	1.6723	1.9196	12.88
	20	2.5852	2.7877	7.26	2.3650	2.7877	15.16
	40	3.5037	3.4839	0.57	3.3447	3.4839	4.00
	50	3.8639	3.8186	1.19	3.7395	3.8186	2.07
0.5	100	5.2367	5.1576	1.53	5.2884	5.1576	2.54
	10	2.0951	2.1897	4.32	1.8571	2.1897	15.19
	20	2.8394	2.8635	0.84	2.6263	2.8635	8.28
	40	3.8483	3.8063	1.10	3.7142	3.8063	2.42
0.8	50	4.2439	4.2230	0.49	4.1526	4.2230	1.67
	100	5.7518	5.6979	0.95	5.8726	5.6979	3.07
	10	2.5528	2.7800	8.17	2.3157	2.7800	16.70
	20	3.4597	3.4750	0.44	3.2749	3.4750	5.76
1	40	4.6890	4.6870	0.04	4.6315	4.6870	1.19
	50	5.1711	5.1692	0.04	5.1781	5.1692	0.17
	100	7.0083	7.0061	0.03	7.3230	7.0061	4.52
	10	2.8038	3.1149	9.98	2.5716	3.1149	17.44
Average % Error	20	3.8000	3.8308	0.80	3.6367	3.8308	5.06
	40	5.1501	5.1658	0.30	5.1431	5.1658	0.44
	50	5.6797	5.6852	0.10	5.7502	5.6852	1.14
	100	7.6976	7.7627	0.84	8.1320	7.7627	4.76
Average % Error		1.52			4.89		

The average absolute percent for the unrestricted case is 1.52%, which is much lower than for the restricted one. The large errors for the restricted case are found at low n values; this implies that the unrestricted model (i.e., estimating the exponent for a) can increase the overall estimation accuracy by reducing the difference between the estimated and optimized tour lengths, particularly at low n . If p is 1.0, β for the restricted case becomes 0.8132, which is about 2% above the previous result of $\beta = 0.7979$ at n with 100 in Table 8.

Although the actual visited points are identical (e.g., $n \cdot p = 20$), the approximated tour lengths in different combinations can be different. For instance, tour lengths are overestimated at small p (e.g., $n = 100$ with $p = 0.2$) and underestimated at large p (e.g., $n = 40$ with $p = 0.5$). The gap in tour lengths between small and large p decreases as p approaches 0.5. One reason for this is rounding errors for the estimators in Table 21. Another possible reason is the estimation errors from the residual (i.e., error terms that represent imperfect goodness of fit).

4.3 Comparison of Approximated Distance versus Actual Road Network Distance

4.3.1 Case Study for Rural Area

4.3.1.1 Network Description

The road network used is agricultural product delivery routes for Appalachian Sustainable Development (ASD). ASD is a non-profit organization focusing on sustainable agriculture development in the central Appalachian region, including Southwest Virginia, Eastern Kentucky, West Virginia, Eastern Tennessee, and Southeast Ohio. ASD's goal is connecting producers to wholesale and retail outlets, searching for local farm products to catalyze economic opportunities in the food and agriculture sector in distressed communities. The organization concentrates on perishable items (e.g., fruits or vegetables) and depends on trucking to reach markets.

With a fleet of two refrigerated trailers, ASD serves South, North, and Kentucky routes biweekly and West Virginia routes weekly. ASD collaborates with several partners for aggregation and food processing. Fresh items are gathered in six aggregation facilities first, and then the items are delivered to five wholesalers/grocers/retailers/food markets in a service region (Figure 11).

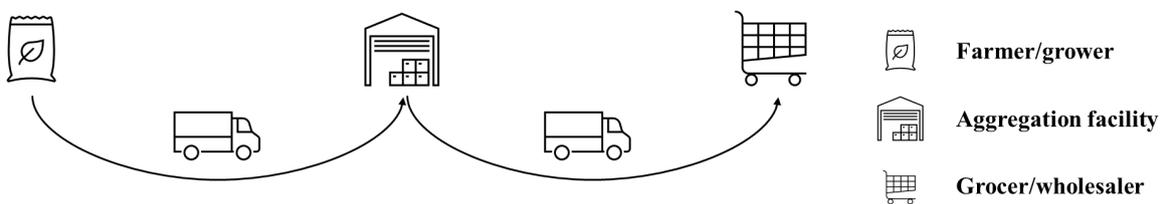


Figure 11 Example of ASD's Delivery Operation

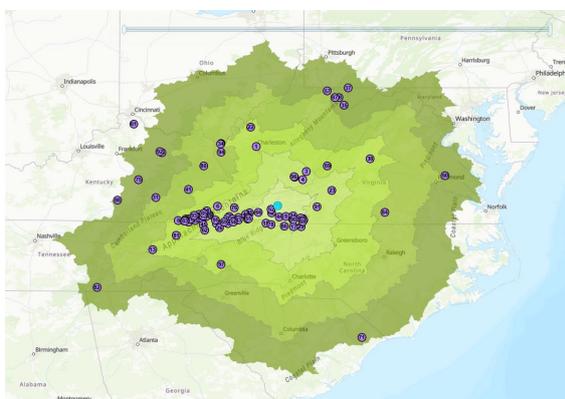
Although the delivery service has been implemented, more profitable transportation routes and various delivery scenarios should be thoroughly explored. Another objective in this chapter is to compare actual and approximated tour lengths. With the comparison results, planners will be informed about some considerations in using approximation models for analyzing transportation system planning problems.

4.3.1.2 Data Processing

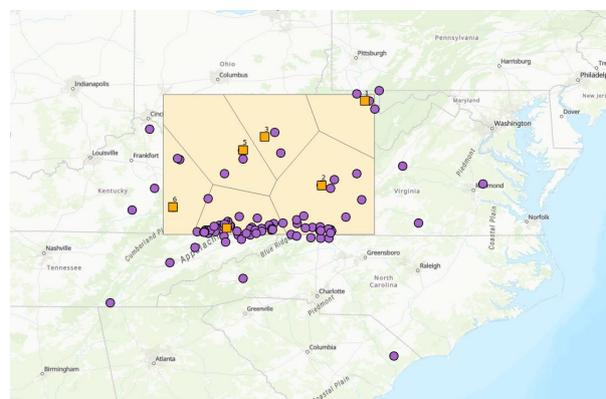
A total of 119 delivery points (e.g., farmer, aggregation, and wholesalers) are mapped using geographic information systems (GIS) software: ArcGIS Pro. This section presents the following: 1) service areas, 2) delivery routes, and 3) a circuitry factor.

4.3.1.2.1 Service Area Z

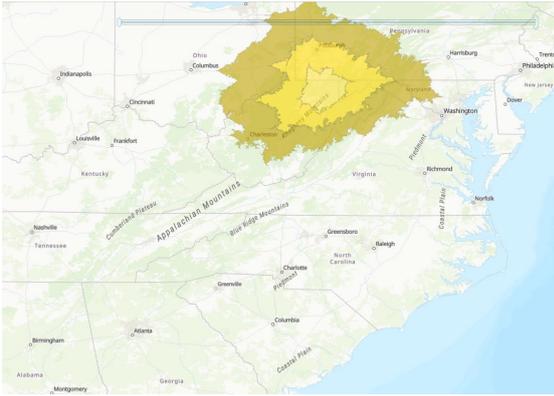
The service area Z is an artificial region that encompasses most of the delivery points and major streets/highways. Z is created using boundaries that can be reached within a 3- and 5-hour driving distance from either aggregation facility (i.e., depot) or centroid. In Figure 12 (a), most of the delivery and pick-up points (colored in purple) are included within a 5-hour driving distance from the centroid, while Figure 12 (c) – (h) are generated based on each depot's location. All isochrones are hourly (e.g., 1-, 2-, 3-, 4-, and 5-hour isochrones from each centroid).



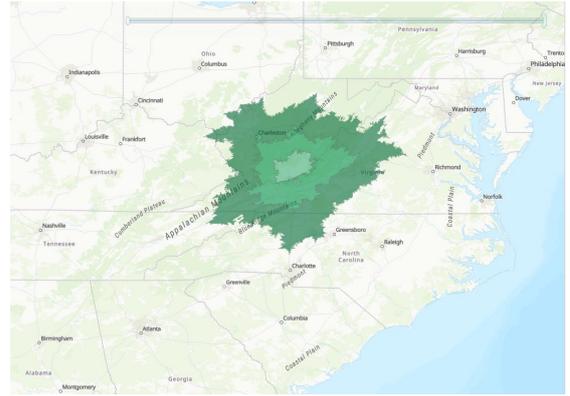
(a) Centroid



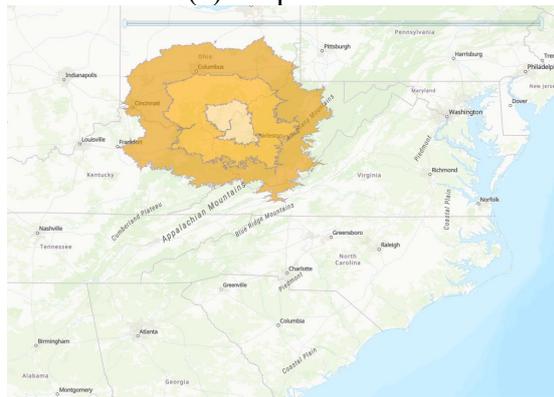
(b) Study Areas Overlapped



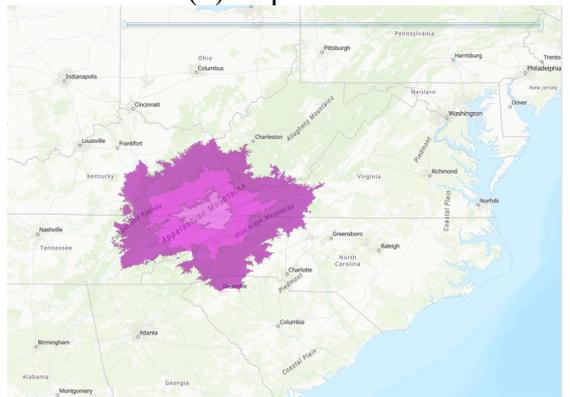
(c) Depot 1



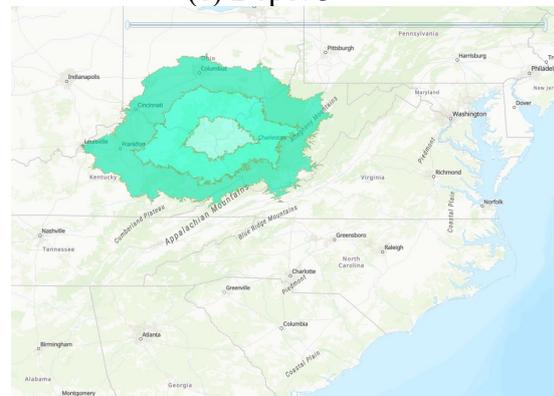
(d) Depot 2



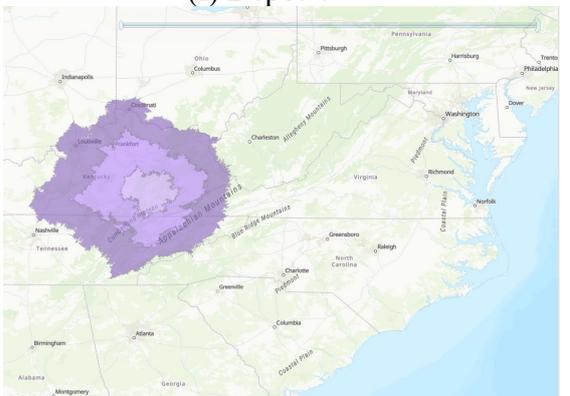
(e) Depot 3



(f) Depot 4



(g) Depot 5



(h) Depot 6

Figure 12 Illustration of Service Areas

Polygons in Figure 12 (b) are partitions of the entire area close to each point (i.e., Voronoi polygons). Although the points are allocated to be served by the nearest depot (orange), some points located outside of the service area (i.e., 3-hr isochrones) may not be served.

Table 23 Summary of Service area

	Z (mi ²)	Driving hour (hr)	n (points)
Centroid	116,961.0	5	119
Depot 1	47,262.2	3	6
Depot 2	38,869.7	3	29
Depot 3	48,481.4	3	2
Depot 4	37,000.9	3	45
Depot 5	48,573.6	3	5
Depot 6	44,965.3	3	9

Table 23 shows the service area size accessible within specified driving hours and the number of delivery points in the region.

4.3.1.2.2 Delivery Scenarios and VRP Solver

The current ASD deliveries serve only wholesalers and retailers. In Figure 13 (a), the baseline scenario consists of two separate deliveries: ASD's truck delivery and farmer's self-delivery. From aggregation facilities (i.e., depots) where growers gather agricultural goods, the ASD trucks start their journeys to the wholesalers/retailers. In an alternative scenario, these trucks visit all farms within the service area for item pick-ups and deliveries to the wholesalers/retailers.

The vehicle capacity is set at 30 items per truck, while the driver's working period is 10 hours/day across scenarios. Each scenario is modeled as the capacitated vehicle routing problem with time windows (CVRPTW): one of the variants of VRP in which vehicles have a homogeneous loading capacity serving customers with a specific visiting hours of delivery points and terminating conditions of delivery. For simplicity, it is further assumed that loading/unloading time per stop or break time for the driver is small enough to be negligible.

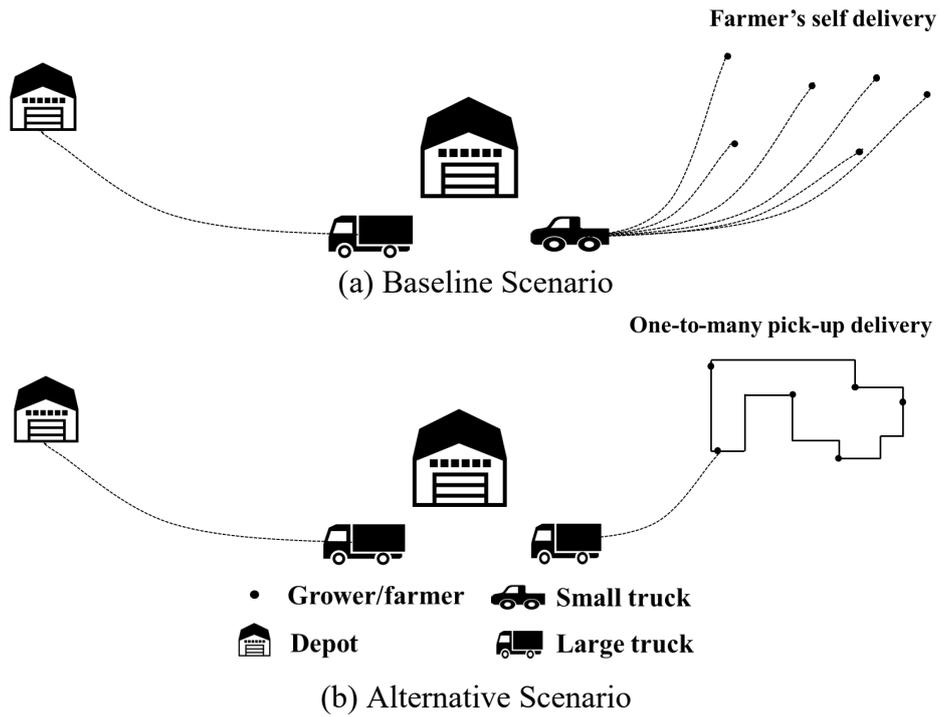


Figure 13 Delivery Scenarios

Each delivery scenario is formulated as a VRP with given vehicle capacity and driver's working period. Then, delivery routes are optimized with a tabu search, which optimizes the sequence for visiting the stops.

4.3.1.2.3 Circuitry Factor c

Since the dead-end or one-way road networks in rural areas increase tour distance, a circuitry factor can be considered in approximating tour lengths. The circuitry factor is the average ratio of actual travel distance to Euclidean distance, as shown in Equation (18). The circuitry factor is greater or equal to 1 (Ballou et al. 2002 and Kweon. 2019):

$$c = \frac{\sum D_n}{\sum D_e} \quad (18)$$

where c is a circuitry factor in the service area, D_n is a summation of network distances of a randomly selected set of two points (i.e., O-D pair) within the service area, and D_e is a summation of Euclidean distances between the two points

At least 30 samples (i.e., distance pairs) are required to estimate the circuitry factor (Ballou et al. 2002). From 87 distance pairs, a circuitry factor of 1.49 is computed with standard deviation of 0.37, according to Equation (18). The estimated circuitry factor is used for approximating tour length in the ASD service region.

4.1.3 Results and Analyses

Delivery routes for baseline and alternative delivery scenarios are optimized, as shown in Table 24 and 25. Note that excluded points are the growers located outside the service area or cannot be visited within the time constraint (i.e., driver’s working period of 10 hours). The delivery distance for all farmers’ self-delivery is reduced from 5,920.2 miles to 3,698.7 miles by ASD’s trucks visiting farmer’s locations.

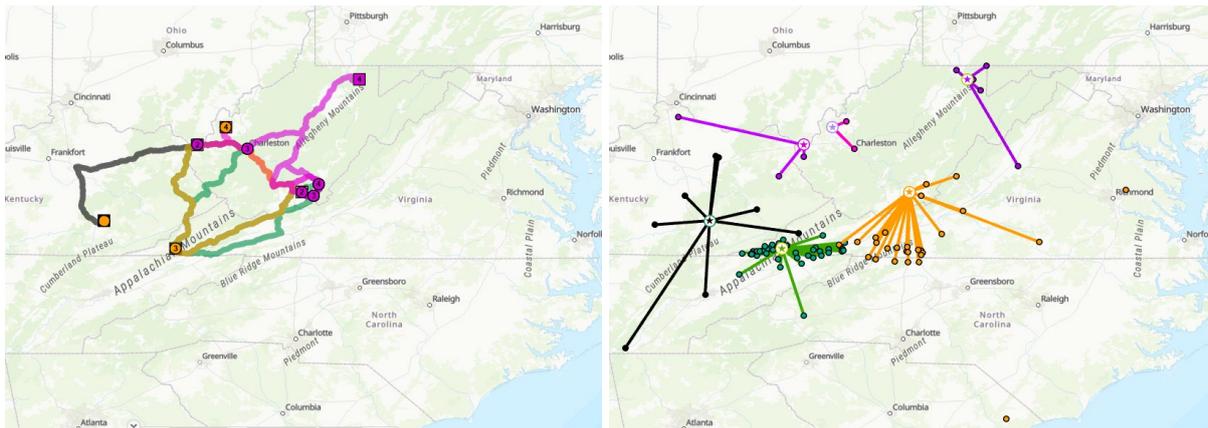
Table 24 Results for Baseline Scenario

Delivery to wholesales/retailers					Farmer's self-delivery				
Depot	Route	D_n	n	Excluded	Depot	# Routes	Avg. D_n	n	Excluded
1	1	452.8	3		1	6	53.5	1	
	2	434.2	2						
2	1	570.7	3	1	2	28	99.7	1	2
	2	491.9	2						
3	1	406.8	5		3	2	42.6	1	
4	1	401.2	3		4	45	35.6	1	
	2	404.8	2						
5	1	363.3	5		5	5	53.4	1	
6	1	503.3	2	3	6	9	93.1	1	
Total	9	4,028.9	27	4	Total	95	5,920.2	6	2

Table 25 Results for Alternative Scenario

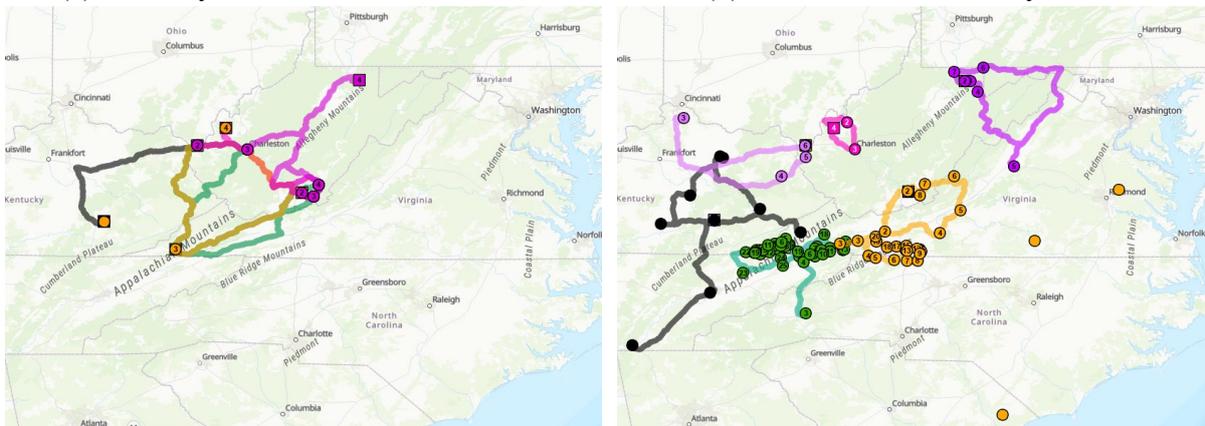
Delivery to wholesales/retailers					Pick-up delivery				
Depot	Route	D_n	n	Excluded	Depot	Route	D_n	n	Excluded
1	1	452.8	3		1	1	463.3	6	
	2	434.2	2		2	1	419.4	8	
2	1	570.7	3	1		2	570.0	19	
	2	491.9	2		3	1	125.2	2	
3	1	406.8	5		4	1	420.1	18	
4	1	401.2	3			2	311.0	27	
	2	404.8	2		5	1	434.9	5	
5	1	363.3	5		6	1	443.1	2	
6	1	503.3	2	3		2	511.7	7	3
Total	9	4,028.9	27	4	Total	9	3,698.7	94	3

Figure 14 shows the optimized tour routes for the baseline and alternative scenarios. Note that the delivery routes in Figure 14 (b) seem straight lines, but the distances for the routes are based on the actual road network.



(a) Delivery to Wholesales/Retailers

(b) Farmer's Self-delivery



(c) Delivery to Wholesales/Retailers

(d) Pick-up Delivery

Figure 14 Illustration of Optimized Delivery Routes

Table 26 compares tour lengths estimated by the proposed approximations with the optimized tour lengths. *App. Tour Lengths* are TSP tour distances estimated from Beardwood's formula (i.e., $\beta\sqrt{nA}$), while *Opt. Tour Lengths* are the optimized tour distances by a Tabu search considering the actual road network in the service area. Note that the Concorde solver discussed in Chapter 3 is originally designed for TSP; the Tabu search is introduced here for a VRP solver. The coefficients β associated with n values are taken from Table 8 (Case 10), while a circuitry factor c for the service area is fixed as 1.49. It is assumed that the service area Z and the number of delivery points served by trucks are divided by the number of routes. For instance, delivery area A for Depot 1 becomes 23,631.1mi².

Table 26 Comparison between Approximated and Optimized Tour Lengths (Rural)

Depot	Z (mi²)	n	β	c	# Routes	App. Tour Length (mi)	Opt. Tour Length (mi)	Difference (%)
1	47,262.2	6	0.7179	1.49	2	284.81	463.3	38.53
2	38,869.7	29	0.6910	1.49	2	555.91	989.4	43.81
3	48,481.4	2	0.7000	1.49	1	324.78	125.2	-159.47
4	37,000.9	45	0.6769	1.49	2	650.72	731.1	10.99
5	48,573.6	5	0.7000	1.49	1	514.01	434.9	-18.19
6	44,965.3	6	0.7179	1.49	1	555.60	954.7	41.80
Total		93				2,885.8	3,698.6	

The average absolute percent difference between actual and approximated tour lengths is 52.1%. Percent differences for service areas (i.e., Depots 1 - 6 in Table 26) vary significantly. Possible reasons for this discrepancy are violations in assumptions (i.e., point distribution and tour characteristics) when applying the approximations. Furthermore, delivery points are clustered rather than uniformly distributed, as shown in Figure 12 (a). For some established routes, tour zones are neither compact nor convex due to low connectivity in the rural road

network, as presented in Figure 15. Note that the approximation assumptions are discussed in Section 2.1.4.

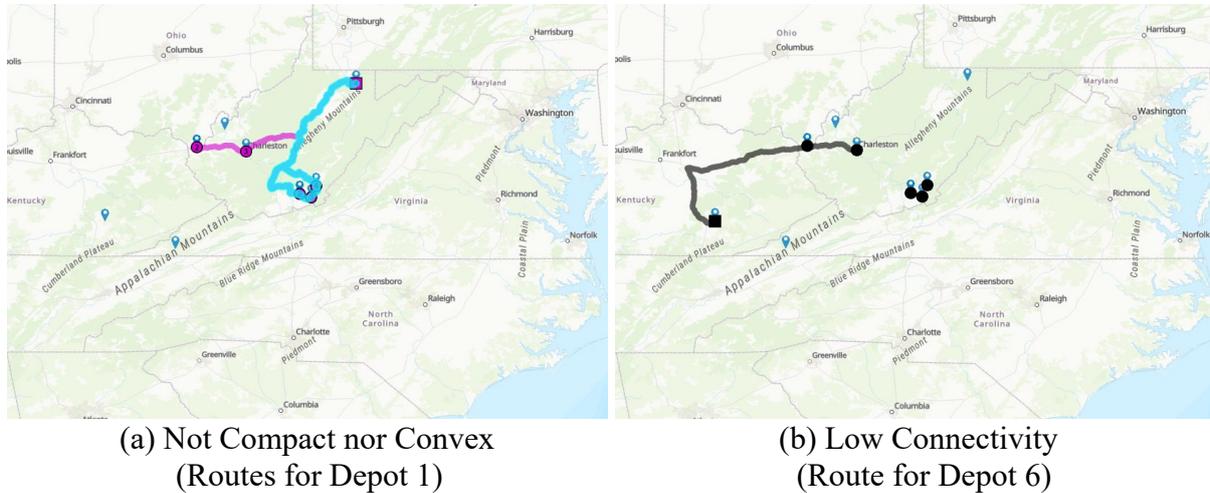


Figure 15 Some Tours Violating Assumptions in Approximation

To overcome that estimation error, lower/upper bounds of coefficients β can be considered, which are presented in Table 8. For instance, the percent error decreases to as low as 0.7% using the upper bound of β (74.5% for absolute percent error).

4.3.1.4. Comparison of Results between Actual and Random Point Distribution

This section is designed for investigating the approximated TSP tour lengths in the previous network with a different point distribution. Therefore, it will be explored how the approximation assumptions (i.e., the point distribution and minimum number of points) affect the tour lengths.

Table 27 Comparison between Approximated and Optimized Tour Lengths (Rural)

Route #	Z (mi ²)	n	β	c	App. Tour Length (mi)	Opt. Tour Length (mi)	Difference (%)
1	47,262.2	5	0.7000	1.49	226.7	371.0	38.9
2	47,262.2	5	0.7000	1.49	226.7	416.7	45.6
3	47,262.2	7	0.7179	1.49	232.5	413.9	43.8
4	47,262.2	6	0.7179	1.49	232.5	418.3	44.4
5	47,262.2	16	0.7036	1.49	227.9	327.1	30.3
6	47,262.2	12	0.7036	1.49	227.9	450.0	49.4
7	47,262.2	12	0.7036	1.49	227.9	217.7	-4.7
8	47,262.2	16	0.7036	1.49	227.9	217.5	-4.8
9	47,262.2	10	0.7179	1.49	232.5	405.0	42.6
10	47,262.2	11	0.7036	1.49	227.9	384.8	40.8

Using the same network in Section 4.1.1, one hundred hypothetical delivery points are randomly generated in the service area. In Table 27, a total of ten routes are then optimized by the tabu heuristic.

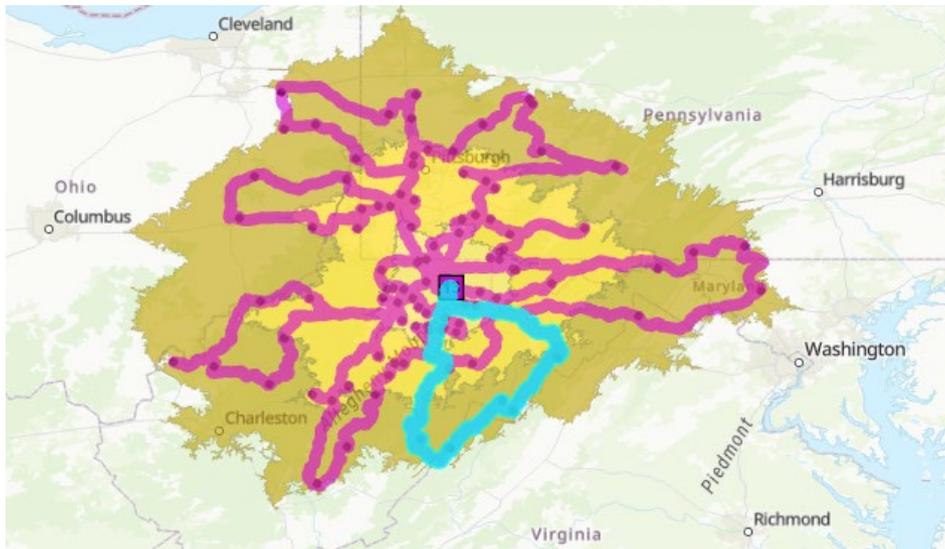


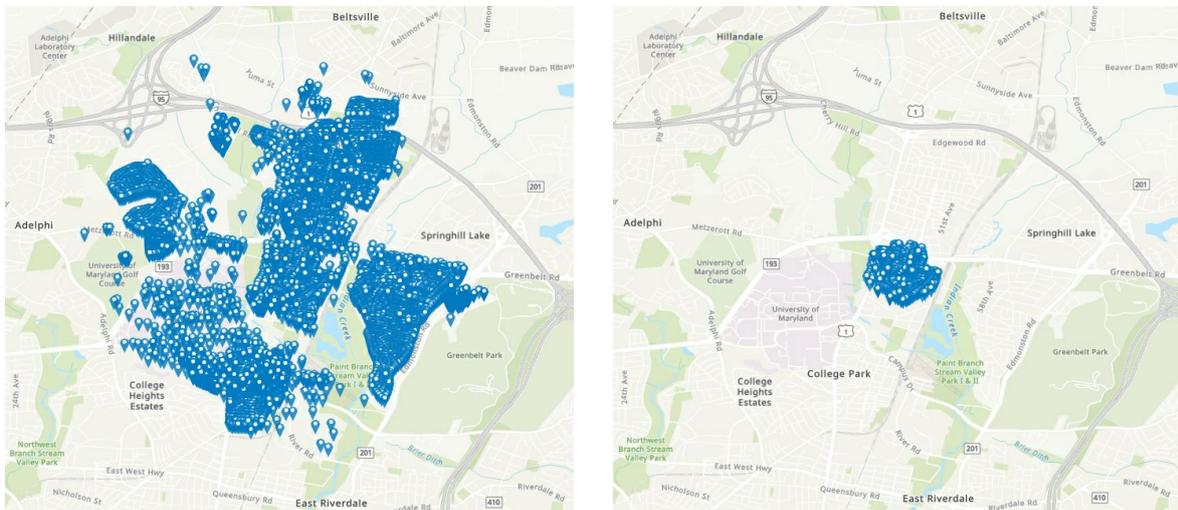
Figure 16 Optimized Routes for Hypothetical Delivery Points

The difference in the average absolute percent error between actual and approximated tour lengths decreases from 32.6% to 13.1% (from the previous result in Table 26). All optimized routes consist of at least five n within a fairly convex delivery area, as illustrated in Figure 16.

4.3.2 Case Study for Urban Area

4.3.2.1 Network Description

To validate the proposed approximations in an urban area, the city of College Park, Maryland, is chosen as a study area. The area of the city is about 8.68 mi², with a population of 32,163 in 2019. Delivery points (i.e., physical addresses of houses/apartments/buildings) are obtained from the OpenAddresses database, an online repository for geocoded addresses. In the dataset 4,288 addresses are available for College Park.



(a) City of College Park

(b) Berwyn Town

Figure 17 Illustrations of Physical Addresses

To reduce the sample size, a subset of the city, Berwyn town, is chosen. In it, 1,025 delivery points are selected. About 54.3 percent of the points are clustered due to many residential apartments in that region. However, the overall delivery points are reasonably distributed uniformly over the service area, as shown in Figure 17 (b). The circular service area surrounding these points is 0.28 mi². It is assumed that all points are served by trucks departing from a single depot located outside of the service region.

4.3.2.2 Route Optimization

Although delivery trucks can carry up to 300 packages in a dense urban area (Figliozzi, 2017), a truck capacity is set here at 200. Trucks spend an average time of three minutes per stop, while the driver’s working shift is 10 hours/day. Note that trucks are not necessarily loaded to their full capacity.

Table 28 Optimized Results for Delivery Routes

Route	n	Opt. Tour Length (mi)	Travel Time (hr)	Unloading Time (hr)
1	133	2.38	0.2	6.7
2	200	1.27	0.1	9.5
3	180	3.86	0.4	9.0
4	131	3.56	0.4	6.6
5	104	3.34	0.2	5.2
6	170	0.88	0.1	8.5
7	107	0.88	0.1	5.9
Average	146.4	2.31	0.2	7.3
Total	1,025	16.17	1.5	51.3

Similar to the rural case study in Section 4.1, each delivery route is optimized by the same solver, and the optimized truck routes are presented in Table 28 and Figure 18. Since the vehicles serve a small block of the delivery area with a large demand density, the delivery time mainly consists of item unloading (e.g., a driver in Route #1 spends 0.2 hours for driving and 6.7 hours for unloading).

Table 29 Comparison between Approximated and Optimized Tour Lengths

n	β	# Routes	Delivery Area (mi ²)	App. Tour Length (mi)	Opt. Tour Length (mi)	Difference (%)
1,025	0.8591	7	0.04	2.09	2.31	9.6

The average number of items carried is about 146 per truck (i.e., n of 1,025 divided by the number of routes). Compared to the case study for a rural area, the average percent error is as low as 9.6% due to the following reasons: 1) uniformly distributed data points, 2) compact and convex circular service area, and 3) high connectivity (e.g., two-way grid road).

4.4 Summary

This chapter explores some possible extensions of the TSP tour length approximations. First, a total of six adjustment factors are developed to integrate the proposed approximations within one equation. Development for adjustment factor associated n values D_0 is a key contribution in this dissertation since β is not a fixed value when n changes. The TSP tour lengths can be more precisely estimated with D_0 than with the existing approximation models. The adjustment factors help planners understand how the tour lengths are sensitively varied by changes in a particular factor. Lastly, researchers can be informed about what was tried and yielded weak results when the factors were combined.

Second, the approximations considering stochastic customer presence are proposed. Third, the tour lengths for rural freight delivery and urban package delivery network are estimated from the dissertation's result. Then, the tour lengths are compared with the approximated distances. The results show that urban areas have favorable conditions for satisfying the imposed approximation assumptions, such as point distribution.

Chapter 5: A Comparison of Optimized Deliveries by Robots, Drones, and Trucks

5.1 Problem Statement

This chapter proposes a distance approximation to estimate the average TSP length for vehicles serving the limited numbers of points n that can be visited per tour and applies it to models for analyzing various types of package deliveries. For approximation models, average TSP tour lengths with different numbers of delivery points n are simulated and then fitted using regression. The models are applied to formulate cost functions for deliveries by ground robot, drone, and conventional truck. Each cost function is optimized and compared with total costs. Sensitivity analyses are designed to explore how system outputs of such delivery systems vary with changes in baseline inputs, including (1) energy cost, (2) user value of time spent waiting for deliveries, (3) service area size, and (4) demand density. For analytic purposes, characteristics of the modes and the baseline for service properties are preset. Several factors that may affect actual applications, such as weather conditions, regulations, and safety issues (e.g., that drones should fly under 400 feet and below 100 mph) are not considered here.

5.2 Methodology

5.2.1 Baseline Numerical Values

Demands (i.e., delivery points) are determined as the product of demand density Q , service area Z , and vehicle departure interval h . The demands are served during working periods W .

Table 30 Variable Definitions and Baseline Values

<i>Symbol</i>	<i>Variable</i>	<i>Units</i>	<i>Value</i>	<i>Range</i>	<i>Note</i>
Decision Variables					
<i>A</i>	Size of Delivery Area	miles ²	-	-	
<i>h</i>	Departure Interval	hr	-	-	
Output Variables					
<i>C_o</i>	Operating Cost	\$ / hr	-	-	
<i>C_t</i>	Total Cost	\$ / hr	-	-	
<i>C_w</i>	User Waiting Cost	\$ / hr	-	-	
<i>L</i>	Average TSP Distance	miles	-	-	
<i>N</i>	Number of Vehicles	vehicles	-	-	
<i>T</i>	Average Delivery Time	hrs	-	-	
Input Variables					
<i>b</i>	Energy Cost	\$/kWh	0.012	0.010-0.014	
<i>B</i>	Battery Capacity	kAh	5.4 (<i>drone</i>) 8 (<i>robot</i>)	-	
<i>D</i>	Driver Pay Rate	\$ / hr	40	-	
<i>H</i>	Handling Cost	\$ / hr	1.5	-	
<i>Q</i>	Demand Density	packages / (mile ² · hr)	20	1 – 40	
<i>S_d</i>	Drone Capacity	packages	5	-	
<i>S_r</i>	Ground Robot Capacity	packages	10	-	
<i>S_t</i>	Truck Capacity	packages	150	-	
<i>T_w</i>	Dwell Time	hrs / stop	0.03	-	
<i>V_d</i>	Drone Speed	mile / hr	50	-	
<i>V_l</i>	Ground Robot Speed	mile / hr	10	-	
<i>v</i>	Truck Speed				
	User Value of Time Spent Waiting for Items	\$ / hr	0.21	0.10 - 0.42	
<i>W</i>	Working Periods	hrs / day	24	-	
<i>Z</i>	Size of Service Area	miles ²	18	1 – 36	

The interval (i.e., headway) h helps concentrate goods for economical loads per vehicle.

Equation (19) indicates how the delivery area A is limited by h and vehicle capacity S_d , S_r , and

S_i . Other conditions being equal, vehicles serve a smaller delivery area A since more demands are generated during longer intervals h . Similarly, A increases as vehicle capacity increases. Most A is smaller of Z , but A may possibly be larger than Z if h is very small.

$$A \leq \frac{S_i}{Qh} \quad (19)$$

Packages are delivered along a last-mile delivery route L at vehicle operating speed V_d or V_r . Average delivery time T is computed using Equation (20); average dwell time per stop T_w is additional time spent per stop for last-mile deliveries.

$$T = \frac{L}{V_l} + T_w(QhA) \quad (20)$$

The number of vehicles serving the service area Z is determined based on vehicle reuse after completing tours. In Equation (21), the number of vehicles N can be found by dividing T by the departure interval h .

$$N = \frac{T}{h} \quad (21)$$

For service alternatives that rely on single vehicles to serve multiple pick-ups or delivery points, the resulting tour lengths are estimated with Equation (22) which approximates the average TSP tour length L :

$$L \cong \beta \sqrt{QhA^2} \quad (22)$$

where β is a constant listed in Table 8 which depends on the shape of service area, location of distribution center, and distance metric.

The energy cost b is proportional to electricity use. The average electricity cost is \$0.012 per kWh ([U.S. Energy Information Administration, 2019](#)). Here proper voltage is assumed to be provided in charging batteries: 18.5 volts for robots and 15.2 volts for drones, for instance.

The handling cost H , for monitoring drone operation, is estimated based on industry rates (Fulfillment by Amazon, 2019). Using the findings from Joerss et al. (2016) that consumers are willing to spend \$5 per shipment in addition to regular delivery prices for same-day delivery, the value of time spent waiting for deliveries v is estimated by converting the additional charge to hourly: \$5 divided by the daily working period W . Since the value is estimated from consumer's willingness to pay for fast delivery, the baseline input for v is not typically regarded as small.

5.2.2 Model Assumptions

For this section, delivery systems for ground robots, drones, and conventional trucks are specified, mathematically formulated and then compared in terms of total cost.

Assumptions for Delivery System

5-1. The demand does not vary with service quality and is served by a single depot located at the center of a circular service area.

5-2. Delivery vehicles carry homogeneous items (e.g., equal package weight and volume) and spend the same dwell time at each last-mile delivery point.

5-3. The demand is uniformly distributed within the service area and over time.

5-4. Ground robots and drones use energy storage completely in every delivery tour (i.e., battery capacity is zero after a completion of each tour).

All the customers' demands are assumed to be non-stochastic and known before a scheduled delivery is initiated. A service area is shaped by a drone's maximum round-trip flight range which encloses a circle. For a fair comparison among delivery options, delivery vehicles are operated in a circular service area, as shown in Figure 19, carrying identical items and spending equal dwell times per stop. Assumption 5-3 is made since the spatial distributions of service

regions are not considered for a general and transferable system design. For Assumption 5-4, the actual amount of energy spent could be computed as in the reference (Choi and Schonfeld, 2018), but estimating it is beyond the scope of this study.

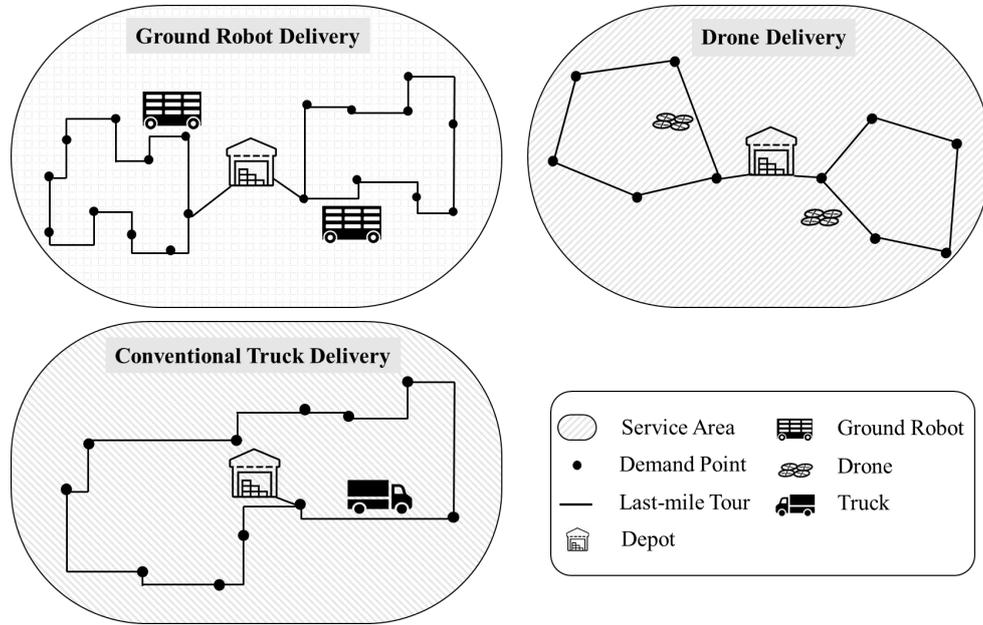


Figure 19 Delivery Options Serving Study Area

For operational settings, deliveries are available throughout a day, i.e., 24-hour operation. The tours of ground robots and conventional trucks are routed in rectilinear space, while drones travel in Euclidean space. Drones maintain a steady level flight at a constant operating speed. Energy consumption associated with other maneuvers, such as acceleration, deceleration, landing and taking off, is not considered.

5.2.3 Model Formulations

The cost function includes operator C_o and user cost C_w as listed in Equation (23). The first term of the equation expresses the costs for system operation associated with item handling H , energy charge cost $b \cdot B$, driver pay rate D , as well as the number of operating vehicles N . The

sum of handling cost and energy cost per average delivery time T is applied to drone and robot delivery, while only D and N are considered for conventional truck. The second term of Equation (23) reflects the waiting time for deliveries which is half the interval h multiplied by total demands and value of time spent waiting for items v .

$$\begin{aligned} \text{Minimize } C_t &= C_o + C_w \\ &= \begin{cases} N \cdot \left(\frac{bB}{T} + H \right) + \frac{(QZh)v_u}{2} & (\text{Drone and Robot}) \\ N \cdot D + \frac{(QZh)v_u}{2} & (\text{Conventional Truck}) \end{cases} \end{aligned} \quad (23)$$

Subject to

$$(QhA) \leq S_i \quad (24)$$

$$T \leq W \quad (25)$$

$$N = \text{integer} \quad (26)$$

where i identifies a delivery option: ground robots, drones, and conventional trucks.

For constraints, the sum of packages carried by vehicles should not exceed the vehicle's maximum capacity. Considering it, Equation (19) is rearranged as in Constraint (24). Each vehicle tour should be completed during the specified working periods W in Constraint (25). Lastly, Constraint (26) requires an integer number of vehicles N for realistic applications.

5.3 Numerical Results and Sensitivity Analyses

5.3.1 Numerical Results

Using the baseline inputs listed in Table 30, the optimization results for deliveries by robots, drones, and trucks are summarized in Table 31. With the imposed Constraints (24) – (26), the

optimal departure interval h^* , which minimizes the total cost function C_t , is found by differentiating the objective function C_t with respect to h .

Table 31 Optimization Results for Delivery Alternatives

	Robots	Drones	Trucks	
Coefficient β	0.9184	0.7336	0.9233	
Departure interval, h^* (hr)	0.162	0.056	1.667	
Delivery area, A^* (mi²)	3.09	4.50	4.50	
Avg. TSP distance, L (mi/vehicle tour)	5.10	3.48	21.45	
Avg. Delivery time, T (hr/vehicle tour)	0.81	0.22	5.75	
Number of vehicles, N^*	6	4	4	
Costs elements (%)	Operating, C_o	72.6	89.5	71.9
	User waiting, C_w	27.4	10.5	28.1
Total cost, C_t (\$/hr)	22.2	19.9	222.5	
Cost per delivery (\$/delivery)	0.38	1.00	0.37	
Critical Constraints	(24)	(24)	(24)	

In ground robot delivery, six robots N with departure interval h of 0.162 hours and delivery area A of 3.09 mile² can optimize the total cost C_t as \$22.2 per hour. Average TSP distance L per vehicle tour is computed as 5.10 miles. For this h^* and A^* combination, the operating and user waiting cost constitute 72.6% and 27.4%, respectively, of the total cost. Cost per delivery of \$0.38 is found by dividing T_c by demands. The optimized intervals h^* for all delivery vehicles are observed away from the cost-minimizing interval due to capacity Constraint (24). In summary, deliveries by robots and drones have lower total cost T_c than by truck for our baseline inputs.

5.3.2 Comparison the Suggested Model to Other Research work

This section compares the previous results using different coefficient β of Equation (22), such as Daganzo (1984). In general, the optimized h decrease as β increases; operating cost C_o is associated with β .

Table 32 Comparison of Results Based on Different Coefficients β

Coefficient β for Equation (22)	Robots		Drones		Trucks		
	Proposed	Daganzo	Proposed	Daganzo	Proposed	Daganzo	
Departure interval, h^* (hr)	0.167	0.167	0.056	0.056	1.667	1.667	
Delivery area, A^* (mi ²)	3.00	3.00	4.50	4.52	4.50	4.50	
Avg. TSP distance, L (mi/vehicle tour)	6.30	6.30	3.55	3.56	23.37	26.73	
Avg. Delivery time, T (hr/vehicle tour)	0.93	0.93	0.22	0.22	5.94	6.27	
Number of vehicles, N^*	6	6	4	4	4	4	
Costs elements (%)	Operating, C_o	70.8	70.8	89.5	89.5	71.9	71.9
	User waiting, C_w	29.2	29.2	10.5	10.5	28.1	28.1
Total cost, C_t (\$/hr)	21.4	21.4	19.8	19.8	222.5	222.5	
Cost per delivery (\$/delivery)	0.36	0.36	0.99	0.99	0.37	0.37	
Critical Constraints	(23)	(23)	(23)	(23)	(23)	(23)	

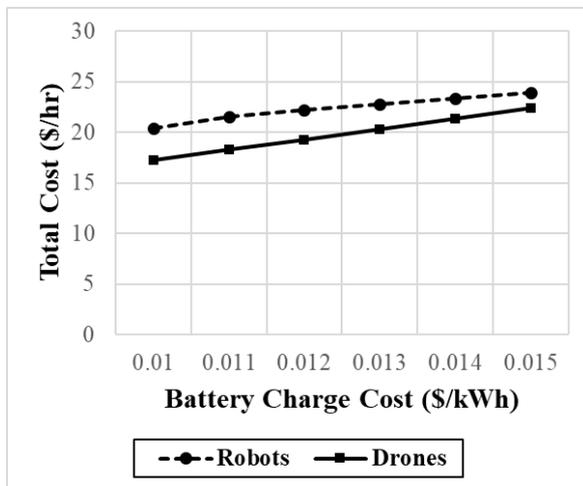
In Table 32, decision variables (i.e., h^* and A^*) for delivery alternatives are nearly unchanged except the total cost C_t . This is due to the integer vehicle Constraint (25). Note that optimized h can be changed for all modes if the constraint is not imposed.

5.3.3 Sensitivity Analyses

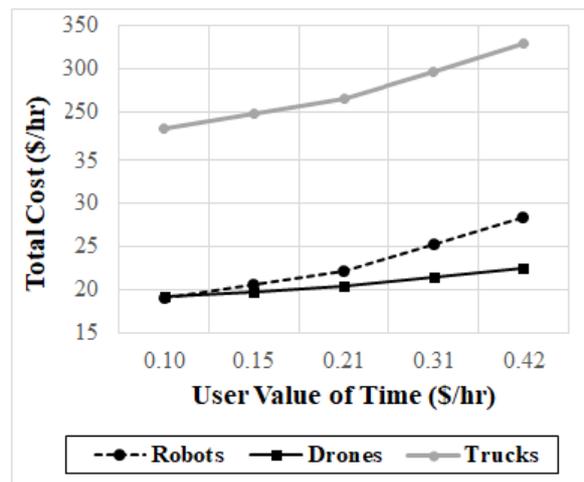
Sensitivity analyses are conducted to explore how outputs vary with changes in baseline inputs: energy cost b , user value of time spent waiting for items v , service area size Z , and demand density Q . These baseline inputs can be flexibly adjusted by the operating conditions for planning purposes.

5.3.3.1 Changes in Energy Cost

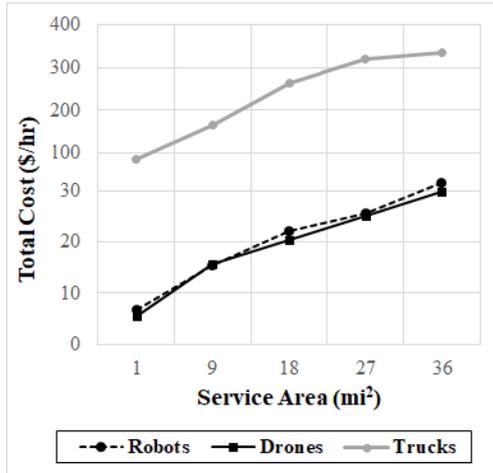
In the United States, the range of electricity rates varies between 0.010\$/kWh and 0.013\$/kWh in recent years (U.S. Energy Information Administration, 2015). The rates change for various reasons, such as seasonal and regional effects, or price changes for raw materials. This section is designed to show how total cost C_t changes with energy cost b . Since C_t for conventional trucks is unaffected by b , that option is omitted. In Figure 20 (a), the difference between total costs decreases as b increases. It is likely that C_t for drone exceeds the cost for robot at large b . Note that the optimal interval h^* is unchanged due to Constraint (23).



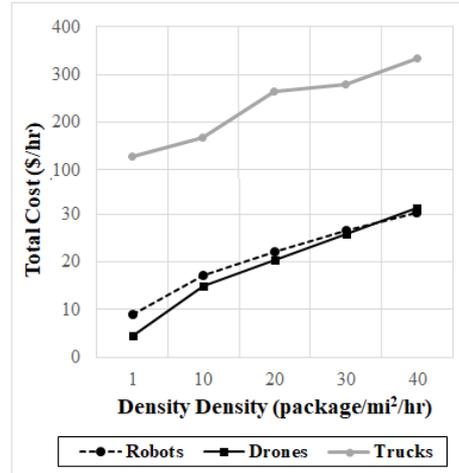
(a) Energy cost b



(b) User value of time spent waiting for goods v



(c) Service area size Z



(d) Demand density Q

Figure 20 Effects of Inputs on Total Costs

5.3.3.2 Changes in User Value of Time Spent Waiting for Items

As noted for the baseline inputs, the baseline value of time spent for waiting items v is an already high value. Willingness to pay for urgent goods, such as blood or medical supplies, could be higher and thus the values are changed within plus or minus 50% of the current baseline. In Table 31, user waiting cost C_w contributes less to the cost function. Thus, total cost C_t for drone delivery is less sensitive to v than the cost for other types of delivery. The gap in C_t between ground robots and drones widens as v increases from Figure 20 (b). Thus, drones could be the most cost-effective delivery option for items at high v .

Table 33 Results for Sensitivity Analysis

b		0.01	0.011	0.012	0.013	0.014	v		0.10	0.15	0.21	0.31	0.42
C_t	Robot	20.4	21.6	22.2	22.8	23.4	C_t	Robot	19.0	20.6	22.2	25.3	28.4
	Drone	17.9	18.9	19.9	20.9	21.9		Drone	18.8	19.3	19.9	20.9	22.0
	Truck	-	-	-	-	-		Truck	230.9	247.6	264.3	297.0	329.7
N^*	Robot	6	6	6	6	6	N^*	Robot	6	6	6	6	6
	Drone	4	4	4	4	4		Drone	4	4	4	4	4
	Truck	-	-	-	-	-		Truck	5	5	5	5	5
h^*	Robot	0.16	0.16	0.16	0.16	0.16	h^*	Robot	0.16	0.16	0.16	0.16	0.16

	Drone	0.06	0.06	0.06	0.06	0.06		Drone	0.06	0.06	0.06	0.06	0.06
	Truck	-	-	-	-	-		Truck	1.72	1.72	1.72	1.72	1.72
	Z	1	9	18	27	36		Q	1	10	20	30	40
C_t	Robot	6.6	15.4	22.2	25.5	31.5	C_t	Robot	9.1	17.0	22.2	26.4	30.3
	Drone	5.3	15.5	19.9	24.4	29.0		Drone	8.1	14.5	19.9	25.4	30.9
	Truck	86.5	166.9	264.3	319.8	333.7		Truck	127.2	166.9	264.3	278.1	333.7
N*	Robot	2	4	6	7	9	N*	Robot	4	5	6	7	8
	Drone	1	3	4	5	6		Drone	2	3	4	5	6
	Truck	2	3	5	6	6		Truck	3	3	5	5	6
h*	Robot	0.33	0.22	0.16	0.13	0.11	h*	Robot	0.43	0.23	0.16	0.13	0.11
	Drone	0.19	0.08	0.06	0.05	0.04		Drone	0.29	0.08	0.06	0.05	0.04
	Truck	3.13	2.50	1.72	1.42	1.25		Truck	3.85	2.50	1.72	1.39	1.25

5.3.3.3 The Effects of Service Area Size & Demand Density

With changes in demand density Q , service area size Z can be used to analyze delivery systems adapted to rural settings. Thus, a rural area may be represented by large Z with low Q . Note that delivery area A is a decision variable determined in the optimization process, while Z is given as an input variable. The rate of increase in total cost C_t for trucks is below the rates for robots and drones. This is mainly due to large capacity for truck S_t with a greater consolidation ability (i.e., large interval h compared to other modes). Delivery by drone has lower C_t than the delivery by robot until demand density Q reaches 34 packages/mi²/hr. In Table 33, the diminishing rate of the optimal interval h^* for drones is greater than for robots, and thus more vehicles are added to the system. In short, vehicles with larger carrying capacity may be favored for Z that are larger or have higher demand densities.

5.4 Summary

This chapter analyzes deliveries by ground robot, drone, and conventional truck. Deliveries by robots and drones have lower total cost than by truck for our baseline inputs. Sensitivity analyses are designed to explore how outputs vary with changes in the inputs for (1) energy

cost, (2) user value of time spent waiting for deliveries, (3) size of service area, and (4) demand density. Drones can be a cheaper delivery option than robots if energy charge cost is near our baseline range, but the difference in total costs diminishes as that cost increases. At high value of time spent waiting for items (e.g., blood or medical supplies), drones may be the most cost-effective option. Changes in the size of service area and demand density can be used to analyze delivery systems in rural settings. According to this analysis, delivery vehicles with larger carrying capacity may be favored for service areas that are larger or have higher demand densities.

Future extensions may consider more evaluation factors for deliveries by ground robots or drones, which include the costs for capital investment. In addition, operating conditions for drone delivery, such as winds, noise, or item drop-off methods for drones may be considered as constraints.

Chapter 6: Optimization Approaches for Investigating Various Drone Delivery Alternatives

6.1 Problem Statement

Drone deliveries are considerably restricted in flight range and parcel payload because most drones are powered by lithium-ion batteries, which currently limit flights to about a half hour (UPS Pressroom. 2017). Although these key disadvantages are likely to be alleviated with improved technology, it is useful to consider these characteristics and examine the operating variables in the overall operation process.

Due to range and payload constraints, some of the early contributions to delivery-by-drone focus on such delivery supported by trucks (*DT*). The major emphasis was on identifying to what extent resources, such as time, cost or fuel, can be saved with the help of drones. Ferrandez et al (2016) found that *DT* could reduce operating costs. Truck delivery time could be shortened where the speed of drones was 1) about three times faster than truck's or 2) more than two single-package-carrying-drones were assigned to each truck.

Wang et al (2016) argued that the maximum delivery completion time could be minimized either by 1) drones which traveled faster than trucks or 2) using more than two drones per truck; the authors found that the delivery time could be reduced by up to 75% with all the above considered. Campbell et al (2017) compared conventional truck delivery (*CT*) and *DT* with operating and delivery stop costs. *DT* offered significant cost savings in suburban areas where demand density was relatively high. The savings were attributed to the fewer tours needed. The authors suggested that assigning multiple drones per truck could reduce operating costs by nearly 40%, depending on the number of drones.

DT and *OD* alternatives may not be optimal for drones delivering a single item per stop, if the drones have an available energy surplus for delivering additional items at different stops. These possibilities motivate the present paper, which assesses alternative delivery approaches and compares their costs.

For analyzing the abovementioned delivery systems, the study adopts distance approximation methods that estimate average tour lengths conducted by vehicles with relatively few visited points. The chapter formulates four alternatives of package delivery services with and without the aid of drones: (1) conventional truck (*CT*), (2) drone supported by truck (*DT*), (3) one-to-one delivery by drone (*OD*), and (4) one-to-many delivery by drone (*MD*). For analytic purposes, characteristics of drones and the baseline for service properties and service area are preset. The specified variables are explored through sensitivity analyses. These tests identify the critical factors contributing to the total costs of a delivery system, which include user's and operator's costs. Several factors that may affect actual applications, such as weather conditions, government regulations, and safety issues (e.g., that drones should fly under 400 feet and below 100 mph), are not yet considered here.

6.2 Alternative Descriptions

This section explains various alternatives of package delivery alternatives with or without the help of drones. Drone delivery is classified by whether it is supported by other types of modes or able to carry multiple packages.

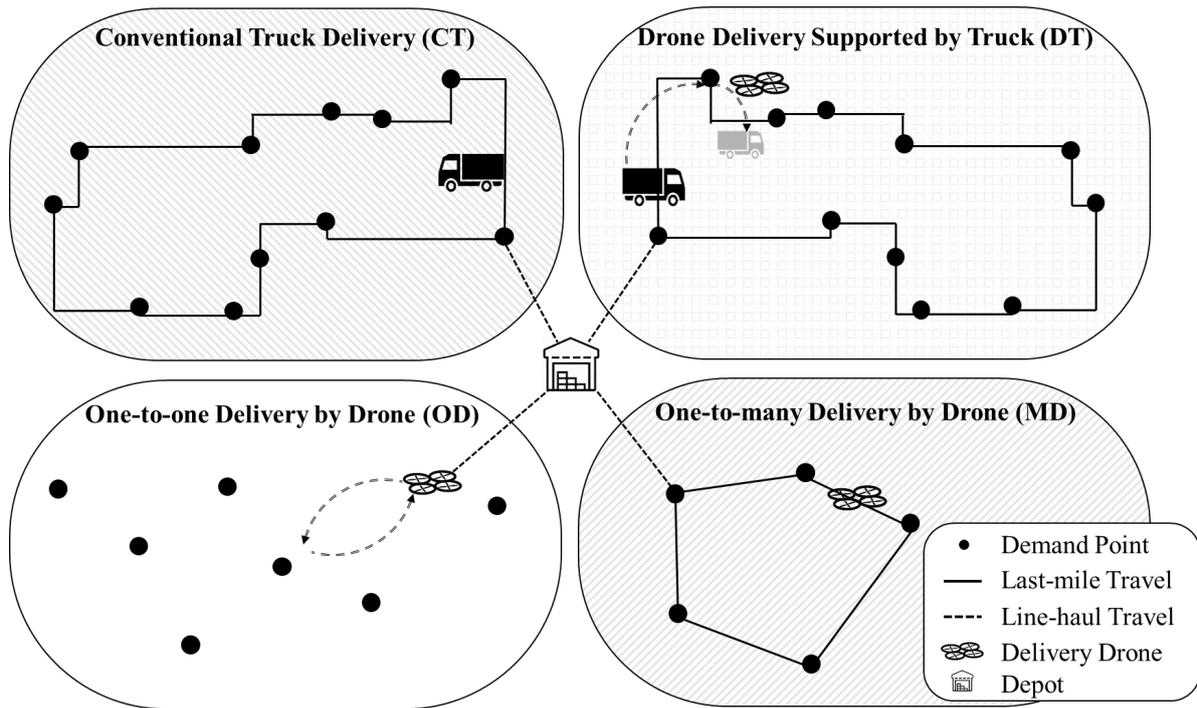


Figure 21 Delivery Alternatives

For fair and consistent analysis among alternatives, no items exceed the drones' maximum allowable payload. Both trucks and drones conduct delivery tasks in a service area of similar size within the drones' maximum delivery range, where a distribution depot is randomly located in the service zone.

6.2.1 Conventional Truck Delivery (CT)

Trucks can carry many more items than drones and thus require longer delivery completion time for each delivery tour due to more stops as well as lower last-mile speeds. A truck's dwell time per delivery point exceeds a drone's since more time is needed for loading/unloading items, searching for a parking spot, parking vehicles, and performing delivery activities. Another characteristic for *CT* is that a last-mile delivery is carried out by human drivers, which implies that the maximum number of deliveries per daily tour may be limited by driver's working period per day rather than by vehicle capacity.

Both the line-haul and last-mile tours by trucks are assumed here to be conducted in a rectilinear space where the vehicle movements are restricted to two orthogonal coordinates. The resulting distance between two visited points is computed as the sum of the absolute differences of their coordinates as shown in Figure 22. Figure 21. For consistent comparison, trucks are assumed to start tours at the center of service area. Both line-haul and local travel speed are assumed to be identical. Although actual delivery time may vary with road traffic conditions and time of day, this study does not consider the conditions.

6.2.2 Drone Delivery Supported by Truck (DT)

Since the delivery range of drones is constrained by battery energy storage, a relatively long tour is provided by ground transportation while a drone serves “last-mile” deliveries only. This drone can be sent to a demand point before a truck arrives there. In Figure 21, by the time the truck arrives at the demand point, the drone has completed its task and is ready to land on the truck (e.g., a grey truck). Note that the truck travels along bold lines while the drone follows dashed lines in Figure 21. By doing so, the trucks move toward demand points without stopping at each point for the last-mile deliveries. The alternative may be subdivided depending on: (1) the number of drones per truck, (2) the drones’ capability of carrying multiple packages, and (3) the possibility of different operating speeds for trucks and drones.

Since information on actual service characteristics for *DT* in private organizations are mostly proprietary, it is challenging to consider all possible cases. If operating speed is higher for drones than trucks, the trucks may skip some delivery points and move directly to the next destination. For simplicity, the study presets that each truck carries one delivery drone. Then, the truck visits multiple customers per tour while the drone conducts a last-mile delivery by carrying a single item per drone stop. The last-mile distance for drones is assumed to be

relatively small since the drones leave from the truck near the delivery point, and the remaining distance is small enough to be negligible. Benefits from employing drones are that some dwell time per stop does not affect the overall delivery time because the two modes move in parallel, and additional deliveries can be made during the saved time. For this case, the maximum number of deliveries per tour is likely to be limited by truck capacity.

6.2.3 One-to-one Delivery by Drone (OD)

In this alternative, delivery drones serve a single destination per tour, i.e. a one-to-one delivery. Once the drones complete their task, they return to the depot and prepare for the next delivery (e.g., battery recharging, maintenance or item fulfillment). One-to-one delivery by drone can be feasible (1) if a service area can be manageably covered by drones, or (2) battery swapping/recharging stations exist in the middle of delivery route to cover a large service area ([Rabta et al. 2018](#); [Hong et al. 2017](#)). The former case is considered here.

6.2.4 One-to-many Delivery by Drone (MD)

For one-to-many delivery drones utilize the energy surplus from the previous *OD* alternative to conduct additional deliveries within their maximum allowable energy storage; the drones serve several customers per tour. The relation between package weight and flight distance will be discussed in the next section.

6.3 Methodology

6.3.1 Baseline Numerical Values

This section discusses the characteristics of delivery modes (i.e., drones and trucks) and their service properties. A cost function is formulated for each delivery alternative and

numerically find an optimal consolidation time which minimizes the system's total cost. A service area is the region where demands are generated and served by delivery modes. Most input variables for delivery drones are adapted from specifications provided by drone manufactures. Other baseline numerical values, such as the service area and drone operating speed, are taken from Amazon.com (Rose. 2013; UPS Pressroom. 2017). The following symbols are used in this paper:

Table 34. Variable Definitions and Baseline Values

<i>Symbol</i>	<i>Variable</i>	<i>Units</i>	<i>Value</i>	<i>Range</i>
<i>Decision Variables</i>				
A_d	Delivery Area for Drone	km ²	-	-
A_t	Delivery Area for Truck	km ²	-	-
C_t	Total System Cost	\$ / hr	-	-
C_c	Capital Cost	\$ / hr	-	-
C_o	Operating Cost	\$ / hr	-	-
C_w	User Waiting Cost	\$ / hr	-	-
h	Consolidation Time	hrs	-	-
L	Last-mile Distance	km	-	-
N_d	Number of Drones	vehicles	-	-
N_t	Number of Trucks	vehicles	-	-
N_p	Number of Packages	package / vehicle	-	-
N_r	Number of Trips	trips / vehicle	-	-
R	Energy Spent per Tour (from full energy charge)	%	-	-
T_{rt}	Delivery Complete Time	hrs	-	-
<i>Input Variables</i>				
α	Average Wait Time Coefficient	-	0.5	-
β	Payload Percentage of Drone (from drone weight)	-	0.5	-
γ, δ	Fractions for Energy Use in Non-Level Flight of Drones	-	0	-
D	Line-haul Distance	km	0	0 - 10
H	Driver Pay Rate	\$ / (truck·hr)	40	36 - 44
i	Interest Rate	%	6.5	-

k	Coefficients for Approximation Equation (30)	-	See Table 8	
M_c	Battery Charging Cost	\$ / trip	0.006	-
M_h	Handling Cost	\$ / drone	1	-
M_i	Indirect Cost	\$ / drone	0.5	-
P_d	Purchase Cost for Drone	\$ / vehicle	3,300	-
P_t	Purchase Cost for Truck	\$ / vehicle	50,000	-
Q	Demand Density	package / (km ² ·hr)	1	0.9 - 1.1
S_d	Max. Allowable Payload for Drone	kg	0.5 * w_d	-
S_t	Vehicle Storage Capacity	kg	200	-
T_d	Dwell Time for Drone	hr / package	0.03	-
T_t	Dwell Time for Truck	hr / package	0.15 (for CT)	-
V_d	Operating Speed for Drone	km / hr	50	30 - 70
V_t	Operating Speed for Truck	km / hr	30	-
v	User Value of Time	\$ / (person·hr)	0.6	0.54 – 0.66
W	Working Periods	hrs / day	12	-
w_d	Drone Weight	kg	11	11 - 17
w_p	Average Package Weight	kg	1	-
Y	Service life for Drone	year	3	-
Z	Service Area	km ²	16 ² π	15.2 ² π - 16.8 ² π

Delivery vehicles travel along a line-haul distance D to the first customer, and the remaining packages are delivered along the shortest last-mile delivery route L at average operating speeds V_d and V_t . The line-haul travel distance D depends on the location of distribution center. For the last-mile delivery distance L , the trucks and drones in alternative MD follow an efficient Traveling Salesman Problem (TSP) tour, while drones in DT and OD travel the shortest direct distance. Details about the last-mile delivery distance will be discussed in the next section. Equation (27) formulates delivery complete time T_r for individual modes in completing a tour, considering the number of packages (i.e., visited points) N_p and the associated dwell time per stop T_d and T_t . The dwell time is estimated by considering a series of delivery processes per delivery point that depend on the mode of transport. For instance, T_d includes take-off, landing, unloading, accelerating, and decelerating, while T_t entails the time spent for parking and last-

mile delivery activity. Note that the dwell time T_d for DT alternative is not added to the entire delivery time, as discussed in the alternative description section.

$$T_{rt} = \begin{cases} \frac{L+2D}{v_t} + T_t(QhA) & \text{(for CT)} \\ \frac{L+2D}{v_t} & \text{(for DT trucks)} \\ T_d & \text{(for DT drones)} \\ \frac{L+2D}{v_d} + T_d \cdot N_p & \text{(for OD and MD)} \end{cases} \quad (27)$$

The demands of a service area are determined by the product of demand density Q , service area Z , and consolidation time h . The demands are served during working period W and assumed to be uniformly generated over time and space. The consolidation time h is a holding time needed to concentrate goods for economical loads per vehicle. Using this relation, Equation (28) shows how delivery area A is associated with vehicle capacity, S_d and S_t , and package weight, w_p . More demands are generated as h increases, and thus vehicles serve compact delivery area A as shown in Figure 22. In most cases, A is smaller (subset) of Z . However, it could be possible that A could be larger than Z where h is extremely small. In addition, A is determined either by considering vehicle storage capacity ($s_t = 2,500$) or the average number of deliveries per hour multiplied by driver's working period ($s_t = 16.7 \cdot W$); each driver can deliver 200-300 packages in an urban area (Figliozzi, 2017). This distinction occurs because the maximum deliveries per tour vary for our alternatives.

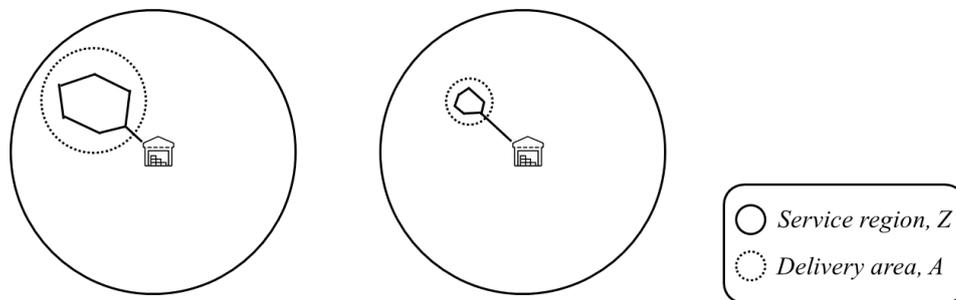


Figure 22. Delivery Area A in Response to Consolidation Time h

The truck fleet is determined based on vehicle reuse after a completion of each tour. In Equation (29), the number of trucks serving the system N_t can be found by dividing round-trip delivery time T_{rt} by the consolidation time h . The numbers of trucks and drones are identical for DT (i.e., a one-to-one paired relation). The drone fleet N_d for OD and MD is determined by vehicle reuse and the number of items carried per tour.

$$A_i = \frac{S_i}{Qh} \quad i = d, t \quad (28)$$

$$\begin{cases} N_t = \frac{T_{rt}}{h} & (CT \text{ and } DT) \\ N_d = N_t & (DT) \\ N_d = \frac{(QZh) T_{rt}}{N_p h} & (OD \text{ and } MD) \end{cases} \quad (29)$$

The number of packages for drones N_p in Equations (27) and (29), ranges from one to the maximum allowable payload S_d . Although payload is related to many factors, including vehicle weight and motor thrust that is a function of air density, rotor diameter, the number of propellers and motor power, the maximum payload $\beta \cdot w_t$ is set by considering the percentage β from the total vehicle mass for drones w_d (Flynt. 2017; Hwang et al. 2018; Lee. 2018). Average package weight per stop w_p can have any value below the maximum allowable payload S_d .

Battery energy storage is set to allow a single package-carrying drone to complete a round trip across a service area; this is the minimum required energy for delivering a parcel to a customer located at the outskirts of the service area. Battery charge cost is proportional to electricity use. Purchasing cost for drones P_d is found by averaging prices of high-end commercial drones. While indirect cost and handling cost, such as monitoring drone operation are considered, the costs related to facility construction or rent are omitted here.

$$N_r = \frac{W}{T_{rt}} \quad (30)$$

In Equation (30), daily trips N_r is the average number of tours made by delivery vehicles during the daily working period. User value of time v is applicable for unattended deliveries, where users would usually wait at their homes, offices, or other convenient places, with little disruption to their other activities. Considering a user's willingness to pay for faster delivery, such as USD 119 for a year subscription to Amazon Prime, the user's expectation for this type of delivery service is reasonably higher than for other types of unattended deliveries. According to a recent survey on the value of time for a same-day delivery, about 9% of consumers are willing to spend USD 5 on top of regular parcel delivery prices. Reflecting this finding, the user value of time is set as \$0.625 per person per hour (Joerss et al. 2016), i.e., USD 5 divided by working period W . It should be noted that the baseline for the value of time is not necessarily set as a small value according to the reference since the value is estimated from consumer's willingness to pay for fast delivery.

6.3.2 Model Assumptions and Formulations

6.3.2.1 Preprocessing Input Variables

Some variables from the baseline are preprocessed for easier computation. First, a distance approximation is introduced for the shortest last-mile travel from a depot to each demand point. Second, battery energy storage is introduced as a constraint for drones considering payload (e.g., package weight) and flight range.

6.3.2.1.1 Approximation of Distance Traveled

For service modes that rely on single vehicles to serve multiple pick-up or delivery points, the resulting tour lengths are estimated with Beardwood's formula in Equation (31). This formula provides good approximations where the shape of the service area is "fairly compact

and fairly convex”, the delivery points are assumed to be uniformly distributed, and the number of delivery points is adequately large, e.g., more than five points. Specifically, this formula approximates the length L of the shortest TSP tour that connects n randomly located delivery points in a delivery zone whose area is A :

$$L = k(nA)^{0.5} \quad (31)$$

where k is a coefficient that depends on the local street pattern, shape of service area, and average number of visited points. To reflect tour length for types of road networks (e.g., rural or urban), the value of k can be multiplied by an appropriate circuitry factor.

6.3.2.1.2 Flight Range and Payload Associated with Battery Capacity

D’Andrea (2014) formulates drone energy consumption considering various factors, such as air resistance, battery cost, and cost of electricity usage. Figliozzi (2018) refines the formula to derive energy E for level flight at a constant speed as shown in Equation (32).

$$E = p \cdot t = \frac{g(m_p + m_v)}{\eta r} d \quad (32)$$

where p is power required for level flight in watts, t is flight duration in seconds, d is flight range in meters, m_p is payload in kg, m_v is drone weight including battery in kg, r is lift-to-drag ratio set as 3, η is power transfer efficiency for motor and propeller set as 0.5, and g is the gravity acceleration constant (9.81 meters/second²).

At least 20% of the full charge energy E should be maintained (i.e., never dip below 20%) for a margin of safety. This is generally known as “the 80% flight rule”, which is commonly used with lithium-ion polymer batteries for the safety, maintenance and protection of drones. Since the exact battery capacity is unknown, that capacity is roughly presumed from drones at Amazon.com, in which a full charge of battery allows a drone to make a round-trip of a 16-kilometer while carrying a single item.

According to Equation (32), the energy consumption of drones increases proportionally with the combined vehicle and parcel weight. Battery capacity is expressed as energy consumption multiplied by duration, and flight range is proportional to the battery energy. Since batteries for delivery drones store a fixed amount of energy, battery capacity can be treated as a constant. It should be noted that drone weight, power transfer efficiency, and lift-to-drag ratio are assumed to be constants in this analysis. Then, the relation among the number of packages N_p (i.e., payload m_p), drone operating speed V_d , and flight duration t can be found. First, the number of packages N_p varies inversely with drone operating speed, d/t . Thus, vehicles can carry more parcels at energy-conserving lower speeds. Second, flight distance d varies inversely with the number of packages. Using Equation (32), it can be computed by the number of packages or average flight distance.

This energy storage relation is used for bounding a drone's maximum flight range associated with the number of parcels carried as well as estimating the drone's operating cost C_o . To derive the operating cost for battery charge, the percentage of remaining energy R can be expressed as follows.

$$R = \frac{E_{base} - E_{used}}{E_{base}} = \begin{cases} \frac{(m_p + \beta \cdot w_d)d - 0.002(N_p + \beta \cdot w_d)V_d \cdot T_d}{(m_p + \beta \cdot w_d)d} * 100 * (1 - \gamma) & \text{(for DT)} \\ \frac{(m_p + \beta \cdot w_d)d - 0.001(N_p + \beta \cdot w_d)(L+D)}{(m_p + \beta \cdot w_d)d} * 100 * (1 - \gamma) & \text{(for OD)} \\ \frac{(m_p + \beta \cdot w_d)d - 0.001(N_p + \beta \cdot w_d)(L+D)}{(m_p + \beta \cdot w_d)d} * 100 * (1 - \delta) & \text{(for MD)} \end{cases} \quad (33)$$

where E_{base} is the energy estimated from drones of Amazon.com allowing a drone to complete a round trip to and from one customer across the service area, E_{used} is the amount of energy spent in a delivery tour, and m_p has a baseline value of 1 kilogram/package in this case. Both γ and δ are fractions of battery energy storage associated with drone landing and takeoff, which reduce the energy usable for delivery distances of drones.

6.3.2.2 Model Assumptions

The assumptions for both truck and drone deliveries are listed here.

6-1. The demand does not vary with service quality and is served by a single depot randomly located in service area.

6-2. The demand is fairly uniformly distributed within the service area and over time (i.e., 12 hours a day).

6-3. All daily demands are served within a predetermined working shift W .

6-4. The tours of each truck are routed in rectilinear space, while drones travel in Euclidean space.

6-5. Drones maintain a steady level flight at a constant operating speed. Energy consumption associated with other maneuvers, such as acceleration, deceleration, landing and taking off, is considered as a fraction (i.e., γ and δ) of the maximum battery energy storage.

6-6. External factors—such as system malfunctions, headwinds, and noise—have no effect on system performance or total cost.

For Assumption 6-1, the circular service area is shaped by a drone's maximum round-trip flight range, which encloses a circle. Assumption 6-2 is proposed since the spatial distributions of service areas are not considered for a general and transferable system design. All the customers' demands (i.e., delivery points) are assumed to be stable and known before a delivery trip is scheduled (Assumption 6-3). The required number of batteries per drone is assumed to be sufficient for all service types supported by drones; the exact figure can be estimated by considering battery energy storage, energy spent per tour, and battery recharge time. For convenience, only the recharging cost is considered. For Assumption 6-5, the specific final-

mile package drop-offs are not considered since various methods are suggested by operators (e.g., Zipline's parachute attached package, Prime Air's drop-off method by landing drone on the ground, and Wing's package dropping by cable). Instead, fractions for energy use of non-level flight of drones γ and δ are introduced for specific drone uses related to energy consumption for package drop-off procedures at a destination and takeoff from depot.

6.3.2.3 Model Formulations

6.3.2.3.1 Cost Function of Conventional Truck Delivery (CT)

The cost function consists of supplier and user cost. The system cost includes the capital which satisfies the peak-period demands and operating cost associated with the number of delivery vehicles, such as battery charge, driver pay rate, management, and maintenance. The user cost reflects the waiting time for deliveries. The total cost function C_t is identical for all the alternatives discussed previously and all costs are hourly.

$$C_t = C_c + C_o + C_w \quad (34)$$

The above cost components are expressed as follows:

$$C_c = \frac{P_t \cdot N_t}{365 \cdot 24} \left\{ \frac{(1+i)^Y - 1}{i(1+i)^Y} \right\} \quad (35)$$

$$C_o = H \cdot N_t \quad (36)$$

$$C_w = \alpha Q Z h v \quad (37)$$

In Equation (35), the capital cost is expressed as the present worth of the investment in truck purchases. Equation (36) includes the costs for system operation, which directly relate to the number of trucks and the tours made within the consolidation time h . Equation (37) specifies the users' cost of waiting to receive packages, which is half (i.e., α is 0.5) the consolidation

time multiplied by total demands and user value of time. The user waiting cost applies similarly for all the delivery strategies.

6.3.2.3.2 Cost Function of Drone Delivery Supported by Truck (DT)

The capital cost includes drone purchases. The operating cost differs from the previous case by considering additional costs related to battery charging and item handling for drones. The user waiting cost remains as in Equation (37).

$$C_c = \frac{(P_t \cdot N_t + P_d \cdot N_d)}{365 \cdot 24} \left\{ \frac{(1+i)^Y - 1}{i(1+i)^Y} \right\} \quad (38)$$

$$C_o = H \cdot N_t + \frac{(M_c \cdot N_r \cdot E_{used} + M_i + M_k)}{h} N_d \quad (39)$$

Another distinction from *CT* is that dwell time does not affect the overall delivery time because drones and trucks do their tasks in parallel.

6.3.2.3.3 Cost Function of Drone-only One-to-one Delivery (OD) / one-to-many Delivery (MD)

The operating cost is adjusted for deliveries solely by drones while user waiting cost remains the same as in Equation (37).

$$C_c = \frac{(P_d \cdot N_d)}{365h} \left\{ \frac{(1+i)^Y - 1}{i(1+i)^Y} \right\} \quad (40)$$

$$C_o = \frac{(M_c \cdot N_r \cdot E_{used} + M_i + M_k)}{h} N_d \quad (41)$$

The difference between *OD* and *MD* is due to the number of vehicles dispatched N_d , and the costs associated with battery recharging $M_c \cdot E_{used}$.

6.3.2.3.4 System Constraints

Constraints apply individually to delivery strategies. The total cost function of *DT* is bounded by constraints (42) and (43). All the listed constraints restrict the other alternatives.

$$N_p \leq \begin{cases} S_t & \text{(for trucks)} \\ \beta \cdot w_t & \text{(for drones)} \end{cases} \quad (42)$$

$$T_{rt} \leq W \quad (43)$$

$$R > 0.2 \quad (44)$$

Constraint (42) specifies that the maximum number of packages per vehicle is less than or equal to its maximum capacity or allowable payload, while constraint (43) requires that a delivery tour should end within one working shift. Constraint (44) binds that the energy spent for each drone tour should not exceed a safety margin. Thus, the drone flight range associated with the number of packages is bounded according to Equation (33).

6.4 Numerical Results and Sensitivity Analysis

6.4.1 Numerical Results

The optimal cost functions in Equations (34) – (37) are found by differentiating the objective function C_t with respect to the consolidation time h . The results must also satisfy the imposed constraints. Using the baseline inputs listed in Table 34, the results for each alternative are summarized in Table 35.

Table 35. Results of Each Delivery Strategy

	CT	DT	OD	MD				
				$N_p=2$	3	4	5	
Total cost, C_t (\$/hr)	387.82	207.8	361.3	312.0	290.0	276.7	267.5	
Cost elements (%)	Operation, C_o	46.1	37.2	44.3	44.3	44.4	44.4	44.4
	Capital, C_c	1.6	1.3	1.6	1.6	1.6	1.5	1.5
	Waiting, C_w	52.3	61.5	54.1	54.1	54.1	54.1	54.1
Deliveries (packages)	676	427	652	563	523	499	483	
Cost / delivery (\$/package)	0.574	0.487	0.555	0.554	0.555	0.555	0.554	
Consolidation time, h (hr)	0.84	0.53	0.81	0.7	0.65	0.62	0.6	
Delivery area, A (km ²)	23.8	37.7	8.2	9.5	10.3	10.8	11.1	
Number of vehicles, N_d or N_t	4.5	1.8	63.3	53.9	49.4	46.5	44.5	
Battery usage (%)	-	16.6	10.2	18.1	25.9	33.8	42.0	
Avg. delivery distance (km)	22.8	54.1	2.4	3.7	4.7	5.6	6.3	
Delivery completion time, T_{rt} (hr)	3.76	0.96	0.08	0.13	0.18	0.23	0.28	

Trucks depart from a depot every 0.84 hours for *CT* and 0.53 hours for *DT*, while all the drones leave the depot every T_{ri} hours. Cost per delivery—total cost over demands generated in consolidation time — ranks *DT*, *MD*, *OD*, and *CT* in ascending order. For drone deliveries, the cost saving from carrying more packages per drone is diminishing. In addition, the fleet size for drones is marginally reduced as more packages can be carried per drone tour.

6.4.2 Sensitivity Analysis

Cases are designed to explore how alternatives are affected by input variables. The results can be considered by operators in planning and managing deliveries.

6.4.2.1 Elasticity to Input Parameters

This section examines how small changes in inputs affect system outputs and thus identifies the critical factors in package delivery systems. The parameters considered here are the driver hourly pay H , demand density Q , value of time v , drone operating speed V_d , and size of service area Z . The driver pay rate is the key cost component in truck deliveries and notably affects the optimized decision variables. For exploring the effects of demand density on system performance, lower demand density may represent rural areas, while higher density may represent urban areas. The user value of time for waiting goods is explored since the value may differ for different customers (e.g., with different incomes), types of items (e.g., fresh products), or places where customers reside. The operating speed of drones not only affects delivery ranges but associated operating cost, based on Equation (41). Lastly, the variation of service area size is examined how that affects the effectiveness and costs of alternatives.

Table 36. Elasticity to Input Parameters

Parameter	Range	Number of Vehicles N_i				Consolidation Time h^*				Total Cost C_i^*						
		CT	DT	OD	MD	CT	DT	OD	MD	CT	DT	OD	MD			
														Np=2	Np=5	Np=2
<i>Driver's Hourly Rate, H</i>	-10%	-0.47	-0.50	-	-	-	0.45	0.35	-	-	-	0.45	0.34	-	-	-
	+10%	-0.58	-0.31	-	-	-	0.50	0.20	-	-	-	0.47	0.36	-	-	-
<i>Demand Density, Q</i>	-10%	0.49	0.27	0.71	0.74	0.79	-0.51	-0.48	-0.11	-0.11	-0.15	0.46	0.40	0.89	0.89	0.90
	+10%	0.47	0.43	0.74	0.73	0.81	-0.55	-0.66	-0.13	-0.13	-0.18	0.49	0.43	0.90	0.90	0.91
<i>Value of Time, V</i>	-10%	0.47	0.51	0.10	0.14	0.10	-0.41	-0.33	-0.31	-0.48	-0.42	0.51	0.60	0.53	0.53	0.53
	+10%	0.43	0.62	0.17	0.13	0.13	-0.41	-0.43	-0.56	-0.49	-0.57	0.54	0.62	0.56	0.56	0.55
<i>Drone Operating Speed, V_d</i>	-10%	-	-	-0.52	-0.46	-0.41	-	-	-0.21	-0.25	-0.15	-	-	-0.21	-0.22	-0.21
	+10%	-	-	-0.58	-0.55	-0.41	-	-	-0.27	-0.16	-0.37	-	-	-0.23	-0.24	-0.22
<i>Service Area, Z</i>	-10%	0.47	0.51	1.00	1.00	1.00	0.41	0.33	0.00	0.00	0.00	0.51	0.60	1.00	1.00	1.00
	+10%	0.43	0.62	1.00	1.00	1.00	0.41	0.43	0.00	0.00	0.00	0.54	0.62	1.00	1.00	1.00

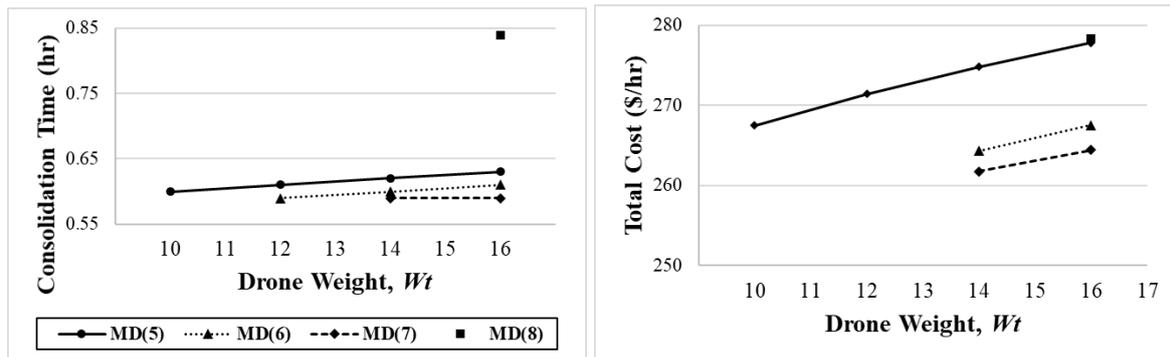
Table 36 summarizes the results of elasticities to inputs. First, changes in **driver pay rate** have greater effects on the service performances of *CT* than *DT* (e.g., either on consolidation time or on user waiting) due to higher operating cost, as shown in Table 35. As **demand density** increases, more trucks and drones are required. Although both truck and drone fleet sizes increase with the density, the increase for drones is much greater than for trucks due to the small drone payload. The optimal consolidation time h^* is reduced as users place a higher **value on waiting time**. Since *DT* has the smallest fleet to serve customers and a relatively large consolidation time, a larger fleet is required for *DT* than for other options. For elasticities to **drone operating speed**, analysis for *DT* is excluded since the delivery completion time is unaffected by that speed; a consolidation time is only determined by the truck capacity S_t regardless of drone speed variations. The optimal consolidation time h^* for drone delivery decreases as the speed increases. This is not attributed to a decrease in user waiting cost but to a decrease in operating cost from energy spent according to equation (8). Overall, the study area can be served with fewer drones as drone speed increases. Lastly, an increase in **service area** changes delivery tour distances L . Consolidation times for all modes decrease with the

size of service area due to the large number of packages generated ($Q \cdot Z \cdot h$). A drone fleet linearly increases with the service area, while truck fleet shows moderate changes according to Equation (28) and vehicle loading capacity (S_d and S_t).

6.4.2.2 Deliveries with Larger Drones

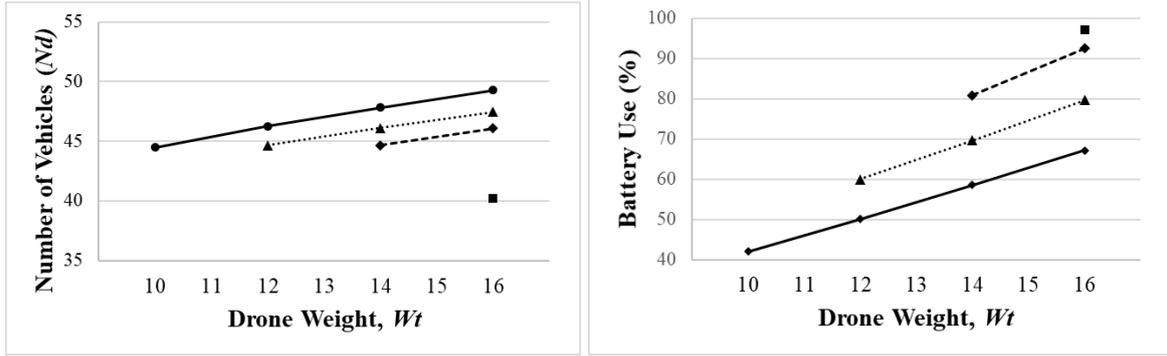
This analysis shows the effectiveness of deploying larger drones in a delivery system for carrying more items per *MD* drone by raising the maximum payload constraint (41). Since the average weight per item is unchanged, this case solely applies to *MD* alternative. As of 2018, the U.S. Federal Aviation Administration defines as “small” unmanned aircrafts which weigh less than 25 kilograms or 55 pounds. This limit allows the baseline drones (i.e., 10 kilograms) to be replaced with larger drones which can carry more packages.

The analysis examines 20%, 40% and 60% heavier drones whose payload is a product of parameter β and drone gross weight w_d . In Figure 23, numbers in parentheses within the legend denote the number of items carried per tour N_p . The associated battery energy storage should increase with the weight of drones, thus affecting the energy used by drones in Equation (33), while drone purchase costs P_d remain as in the baseline.



(a) Consolidation Time, h

(b) Total Cost, C_t



(c) Fleet Size, N_d

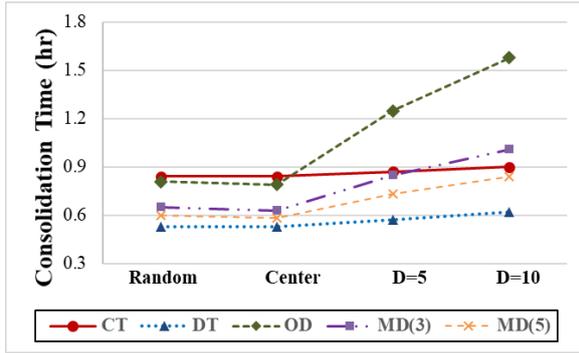
(d) Battery Use, E_{used}

Figure 23. Effects of Large Drones for MD alternative

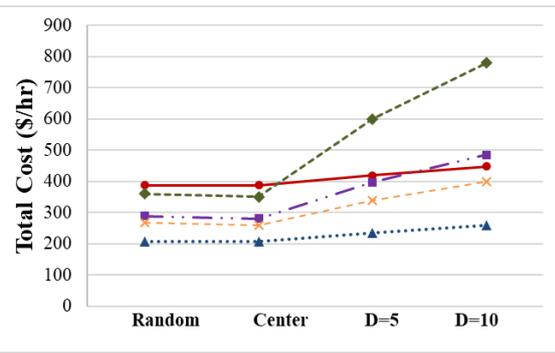
Energy increases with drone weight as in Figure 23 (d). For drones utilizing nearly their maximum energy storage (i.e., 60% heavier drones with 8 items), delivery operation is possible with a high consolidation time. A small delivery area reduces the last-mile tour distance according to Equations (28) and (31). The energy is used inefficiently with heavier drones if all the loaded items N_p can be carried by lighter drones.

6.4.2.3 Special Case: Location of Distribution Depot

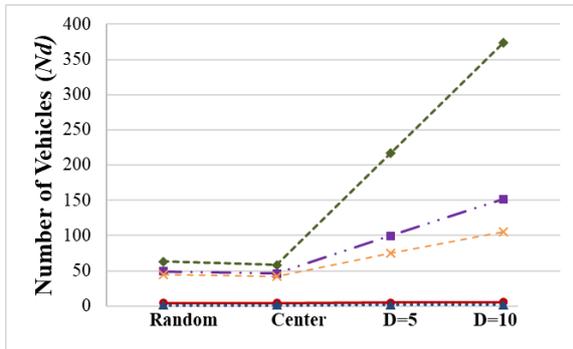
This analysis examines how the location of distribution hub affects the overall system and system performance. This is done by changing the line-haul distance D and coefficient k for last-mile distance L . Including the baseline setting that the hub is randomly placed in a service area, three cases are proposed. For the ‘center’ case, the distribution hub is located at the center of service area. The other two cases change the line-haul distance D .



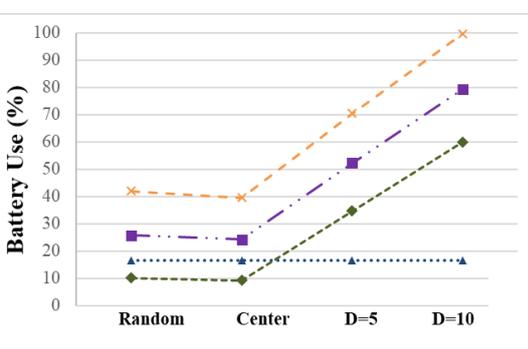
(a) Consolidation Time, h



(b) Total Cost, C_t



(c) Fleet Size, N_d



(d) Battery Use, E_{used}

Figure 24. Effects of Location of Distribution Depot

Comparing the effects of depot locations in a service area (e.g., centered and randomly located), the changes in total costs for drones are approximately 2.6%, while the changes for trucks are unnoticeable, as shown in Figure 24 (b). In external distribution depots, both total cost and fleet size increase with the length of line-haul travel. The drone fleet increases substantially with line-haul distance, where average delivery distance for drones (i.e., a sum of last-mile and line-haul travel) is greater than for trucks.

6.5 Summary

The drone delivery industry is mostly run by private companies, whose achievements indicate that such services are becoming practical. Aside from technical difficulties, many

concerns exist and must be overcome regarding safety, security, regulations, or noise problems. This paper identifies the various alternatives of package delivery services with and without drones. Each delivery method is formulated with a system cost function and compared individually. The authors employ their recently developed distance approximation methods that estimate average tour lengths when only a few points are visited points due to the limited payload of drones. In addition, an energy constraint is incorporated in the model to reflect delivery range associated with payload. It is shown how the optimum delivery area size and consolidation time for minimizing the total cost change as system inputs are varied.

Utilizing drones for package deliveries may be cost-effective compared to conventional trucks. For our baseline values, drones supported by trucks have both the lowest total cost and cost per delivery, while the drone deliveries without trucks become competitive with the cost as more packages are loaded per tour. The study examines sensitivity of alternatives to influential inputs, including driver pay rate, demand density, user value of time, drone operating speed, size of service area, and drone size. Total cost for conventional trucks is more influenced by the driver pay rate than the cost for truck deliveries supported by drones. Although both trucks and drones can conduct frequent delivery tours as the demand density increases, the fleet size for drone-only deliveries increases more than that for trucks because of a payload constraint. Among the four delivery alternatives, a change in user value of time greatly changes total cost for truck delivery supported by drones. The higher operating speeds benefit in reducing costs for both single- and multi-package-carrying drones. Large drones may carry more items per tour but the energy may be used inefficiently with heavier drones if the drone's carrying capacity is underused. As a special case for the effects of distribution hub

location, the study finds that drone fleet size is greatly affected by the locations of the depot compared to fleets for truck delivery if the depot is located outside the service area.

Future extensions may model the possible alternatives for delivery drones supported by trucks (*DT*) and compare them with system cost, such as the number of drones loaded per truck or drones' capability of carrying multiple packages. It is desirable to explore delivery systems while considering additional operating conditions, such as winds and noise, and determine how these affect system cost.

Chapter 7: Innovative Methods for Delivering Fresh Food to Underserved Populations

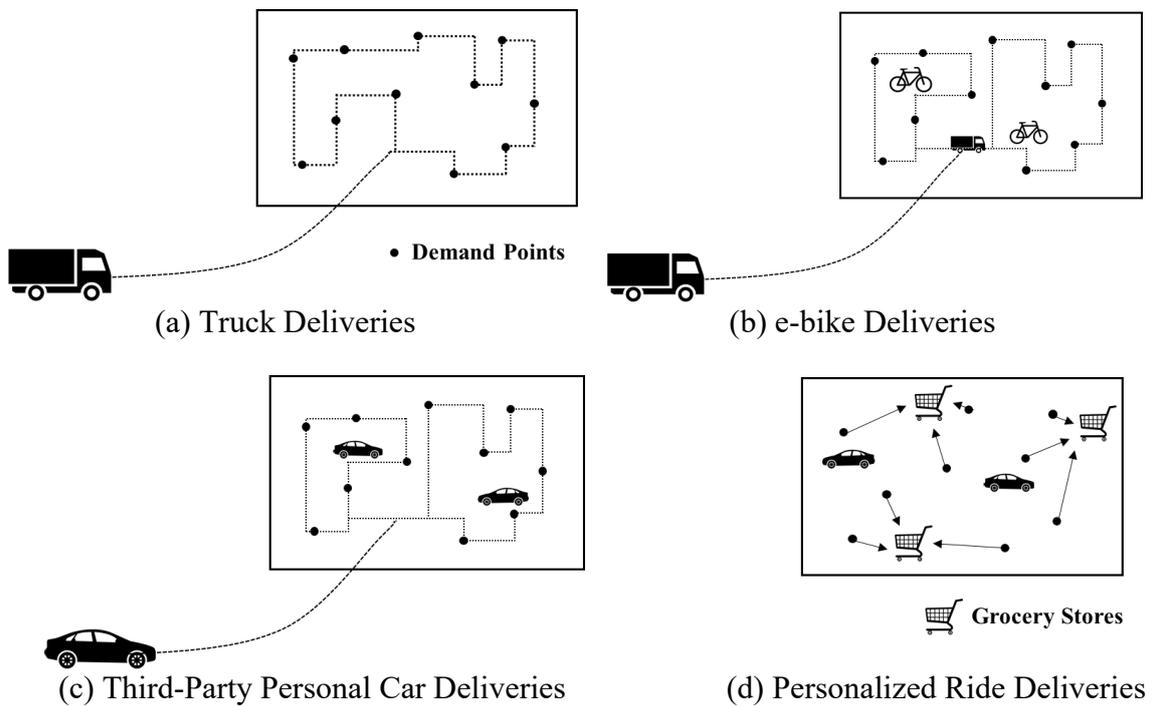
7.1 Problem Statement

The lack of access to fresh foods within reasonable distance and at affordable prices has become a public health concern for individuals living in underserved communities and remote rural areas. Such areas are generally called food deserts. These food deserts are mostly attributed to a scarcity of full-service grocery stores (i.e., selling fresh, canned, dry, and frozen foods), farmers' markets, vehicle availability, or reliable transportation. Thus, residents in food deserts often travel further to access a grocery store, which increases transportation costs and tightens an already limited budget of the household. Furthermore, the lack of access to the foods necessary for a healthy and balanced diet may lead to poor health outcomes. While location decisions for existing grocery stores are based on the profit-maximizing economic principle, system inequity in lower accessibility to fresh foods has emerged as an unintentional by-product. That is, a food desert is an example of market failure that warrants government involvement to improve equity and reduce social costs (e.g., health costs) associated with lower consumption of fresh foods. Therefore, reaching the underserved communities with cost-effective delivery alternatives would be an important service.

The chapter presented here aims to develop a last-mile fresh food delivery system, considering the combinations of transportation modes, for communities with poor access to fresh food. Various fresh food delivery alternatives are identified, including conventional trucks, electric cargo bikes, third-party deliveries by personal car, personalized ride transportation services, and parcel lockers. The corresponding performance and cost functions

for each alternative are formulated. The individual alternatives are separately optimized, and the results are compared. Finally, the study conducts sensitivity analyses in terms of 1) service area size, 2) demand density, 3) user value of time spent waiting for goods, 4) combined deliveries by trucks and estimates 5) mode share for home-deliveries. Using the findings from sensitivity results, the model suggests the optimal mode of transportation for delivering fresh products in the Washington Village/Pigtown section of Baltimore city. The study's main contribution is to evaluate each delivery alternative in terms of total cost, thus enabling local jurisdictions to design the best-suited delivery alternative for the underserved community. Although the delivery alternatives can serve general types of customers and other neighborhoods, the chosen modes are not overly expensive to operate.

7.2 Alternative Descriptions



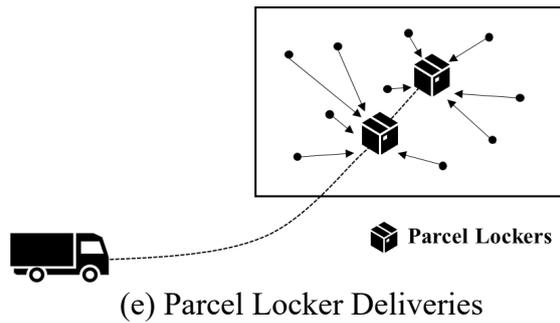


Figure 25 Delivery Alternatives Serving Study Area

The study developed models for the five types of alternatives for fresh food deliveries: trucks, e-bikes, third-party personal cars, personalized ride services, and parcel lockers. Alternative characteristics are discussed below.

7.2.1 Truck Deliveries

In Figure 25 (a), delivery trucks visit all the users (i.e., demand points) in the service area. Trucks travel from the depot a line-haul distance at cruising speed to a corner of the delivery area. From there, drivers drop off groceries at each doorstep by conducting a last-mile delivery tour at average local speed. The study assumes that trucks can load up to 250 packages but may not necessarily travel with a full load.

7.2.2 E-bikes Deliveries

This type of delivery is done by electrically-assisted cargo bikes carrying a small number of items compared to trucks and requires a fulfillment center somewhere inside a service area (Conway et al. 2011). Due to the e-bike's limited loading capacity (150 to 300 Kg), frequent fulfillment trips to the depot are generated. Thus, the depot is replaced with stationed trucks. Therefore, the bike replenishes packages from trucks, while e-bikes serve only the last-mile deliveries. The fulfillment truck as in the truck deliveries is depicted in Figure 25 (b).

7.2.3 Third-party Delivery by Personal Car (TPC) Deliveries

Drivers in third-party delivery by personal car conduct the same delivery process as in truck deliveries (Figure 25 (c)). Aside from a limited loading capacity, delivery characteristics remain the same as for trucks.

7.2.4 Personalized Ride (PR) Deliveries

Instead of delivering items to customers, this alternative considers a vehicle collecting customers in a service area and taking them to the nearest grocery store (Figure 25 (d)). The vehicle is randomly positioned in a service area and travels the shortest distance at an average operating speed to a corner of the customer pick-up locations. The customer's return trip after grocery shopping is considered, as well, possibly with a different driver. For simplicity, the study considers a scheduled-based taxi service, rather than an on-demand service.

7.2.5 Parcel Locker Deliveries

For this case, truck drivers drop off all the items in lockers (Figure 25 (e)). Users then need to access the pick-up locations to receive their items. The costs related to the user's access to the locker are included in the cost function.

7.3 Methodology

7.3.1 Assumptions for Delivery System

7-1. The demand does not vary with service quality (e.g., changes in vehicle operating speed).

7-2. The demand is uniformly distributed within the service area, and deliveries consist of one package per customer (i.e., per delivery point).

7-3. The parcel lockers are evenly distributed within the service area.

7-4. All items in a parcel locker are taken until next scheduled vehicle leaves a depot.

For Assumption 7-1, all the customers' demands are assumed to be non-stochastic and known before a scheduled delivery is initiated. Assumption 7-2 is made since the spatial distributions of service regions are not considered for general and transferable system design. In practice, once packages are delivered, users have some days to pick them up. Since measuring a rate of receiving items is out of the scope of this study, parcel locker users receive their orders before next scheduled vehicles leave a depot (Assumption 7-4) for simplicity.

The study further assumes that parcel lockers are on public property, and the operating cost for the lockers is low enough to be negligible. Some benefits of environmentally friendly modes are not considered, such as gas emissions. Due to difficulties in modeling each produce item with respect to the freshness over time, this study assumes insulated temperature-controlled packaging (e.g., refrigerated bags filled with ice packs) to deliver goods. These items usually stay fresh for up to 24 hours after the expected delivery time. Note that the model users may consider adding a delivery time constraint to the model (which will be discussed in Section 7.3.3.6) to reflect the required freshness.

7.3.2 Baseline Numerical Values

Demands are determined as the product of demand density Q , service area Z , and vehicle departure interval h . The demands are served during regular shift W and assumed to be

uniformly generated over time and space. The headway h is required to concentrate goods for efficient loads per vehicle.

Table 37 Variable Definitions and Baseline Values

<i>Symbol</i>	<i>Variable</i>	<i>Units</i>	<i>Value</i>	<i>Range</i>	<i>Note</i>
<i>Decision Variables</i>					
A	Size of Delivery Area	km ²	-	-	
h	headway	hr	-	-	
<i>Output Variables</i>					
C_o	Operating Cost	\$ / hr	-	-	
C_r	Riding Cost	\$ / hr	-	-	
C_t	Total Cost	\$ / hr	-	-	
C_w	Waiting Cost	\$ / hr	-	-	
C_x	Access Cost	\$ / hr	-	-	
L_t	Average Traveling Salesman Problem (TSP) Distance	km	-	-	
L_s	Expected Shortest Distance	km	-	-	
N	Number of Vehicles	vehicles	-	-	
N_l	Number of Lockers	stations	-	-	
T	Total Delivery Time	hrs	-	-	
<i>Input Variables</i>					
B_p	Driver Pay Rate	\$ / (truck · hr)	40	-	
D	Line-haul Distance	km	16.1	-	
l	Length of Service Area	km	\sqrt{Z}	-	
N_s	Number of Stations for e-Bike Replenishment	stations	1	-	
Q	Demand Density	packages or person/ (km ² · hr)	7.7	0.4 - 15.4	
S_s	Personalized Ride Capacity	person	1	-	
S_b	Bike Capacity	packages	20	-	
S_l	Locker Capacity	packages	50	-	
S_p	TPC Capacity	packages	45	-	
S_t	Truck Capacity	packages	15 · W (truck) 250 (others)	-	
T_m	Dwell Time (Truck, e-bike, and TPC)	hrs / stop	0.05	-	
T_w	Dwell Time (e-Bike replenishment and Locker)	hrs / stop	0.5	-	
T_x	Max. Allowable Access Time	hrs	0.17	-	
V_d	Line-haul Speed	kph	50	-	

V_k	Walking Speed	kph	3	-
V_l	Local Speed	kph	15	-
v_x	User Value of Time Spent for Access	\$ / hr	12	-
v_i	User Value of Time Spent for Riding	\$ / hr	5	-
v_u	User Value of Time Spent for Waiting	\$ / hr	0.625	0.3 - 1.25
W	Working Periods	hrs / day	8	-
w	Width of Study Area	km	\sqrt{Z}	-
Z	Size of Service Area	km ²	46.6	2.6 - 103.6

Equation (45) indicates how delivery area A is associated with h and vehicle capacity S_t , S_b , S_s and S_p . Other conditions being equal, vehicles serve a smaller delivery area A since more demands are generated during longer intervals h according to Equation (45). Based on delivery alternatives, A is determined either by considering vehicle storage capacity ($s_t = 250$) or the average number of deliveries per hour multiplied by driver working period ($s_t = 15 \cdot W$); each driver may deliver 200-300 packages per working period in an urban area (Sheth et al. 2019; Tipagornwong and Figliozzi. 2014). For instance, the capacity for door-to-door services is determined by the driver working hour, while the capacity for fulfillment (e.g., locker or bike replenishment stations) is done by the vehicle storage capacity.

$$A \leq \frac{S_i}{Qh} \quad (45)$$

Delivery vehicles travel along a line-haul distance D to the first customer at average line-haul speed V_d (if applicable to alternatives), and the remaining packages are delivered along a last-mile delivery route L_t or L_s at local speed V_l ; the vehicles return to a depot along the same line-haul route after deliveries are completed. From these, total delivery time T is computed using Equation (46). Average dwell time per delivery point T_m or T_w is the amount of time spent per each stop and depends on alternative types. The dwell time for vehicles conducting last-mile deliveries T_m is shorter than the one for vehicles refilling items T_w since the latter takes

more time for the number of items. T_w for bike replenishment and parcel locker is assumed to be equal, but not obtained from observations; the actual values for T_m and T_w may differ from the time spent for searching parking lot or traffic congestion. The configuration of T is adjustable based on the delivery alternative and will be discussed in a later section.

$$T = \begin{cases} \left(\frac{2D}{V_d} + \frac{L_t}{V_l}\right) + T_m(QAh) & \text{(truck and TPC)} \\ \frac{L_t}{V_l} + T_m(QhA) & \text{(e - bike)} \\ \frac{2D}{V_d} + T_w \cdot N_s & \text{(bike replenishment truck)} \\ \frac{L_s}{V_l} & \text{(PR)} \\ \left(\frac{2D}{V_d} + \frac{L_t}{V_l}\right) + T_w \cdot N_l & \text{(locker)} \end{cases} \quad (46)$$

The number of vehicles serving the area is determined based on vehicle reuse after the completion of each tour. In Equation (47), the number of vehicles N can be found by dividing T by h .

$$N = \frac{T}{h} \quad (47)$$

The operator costs are modeled with cost functions associated with vehicle travel distance at various operating speeds, dwell times at delivery locations, service frequencies, as well as the number and size of vehicles. For service alternatives that rely on single vehicles to serve multiple pick-ups or delivery points, the resulting tour lengths are estimated with Beardwood's formula and its extensions. This formula approximates the length L_t of the shortest Traveling Salesman Problem (TSP) tour that connects any n randomly located points in a zone whose area is A . Beardwood's formula provides good approximations where the shape of the service area is "fairly compact and fairly convex", the delivery points are uniformly distributed, and the number of delivery points is adequately large (Larson and Odoni. 1981). The shortest

expected distance L_s is for a vehicle serving a single destination per tour (i.e., personalized ride service).

$$\begin{cases} L_t \cong k_t \sqrt{QhA^2} & \text{(Truck, e – Bike, TPC, and Locker)} \\ L_s \cong k_s \sqrt{Z} & \text{(PR)} \end{cases} \quad (48)$$

where k_t is a constant that depends on the local street pattern, as discussed in Table 8, while k_s is 0.67 for vehicles randomly positioned in rectilinear space where movements are restricted to two orthogonal coordinates (Larson and Odoni. 1981).

Using the findings from Joerss et al. (2016) that consumers are willing to spend \$5 per shipment in addition to regular delivery prices for same-day delivery, the value of time spent waiting for deliveries v_u is estimated by converting the additional charge to hourly, i.e., \$5 divided by the daily working period W . Since the value is estimated from consumer's willingness to pay for fast delivery, the baseline input for v_u is not necessarily regarded as small. The user value of riding time v_i could be higher than the actual value since it is estimated from commute trips. Note that the value of time spent waiting for deliveries v_u is much smaller than that of time spent for riding v_i and access v_x since the users would usually wait at their homes, offices, or other convenient places, with little disruption to their other activities.

7.3.3 Cost Function

The cost function includes the operator's and user's costs. The operator cost considers the operation costs related to the number of operating trucks and driver pay rate. The user cost can be represented as the cost of the time for waiting C_w , in-vehicle riding C_r , or accessing to service facilities C_x . To sum up, the total cost is expressed in Equation (49).

$$\begin{aligned} C_t &= \text{operator cost} + \text{user cost} \\ &= C_o + C_w + C_r + C_x \end{aligned} \quad (49)$$

It is noted that the elements of the cost function are selectively applicable for each delivery alternative.

7.3.3.1 Truck Deliveries Formulation

Among the user cost components in Equation (50), only user waiting is considered for truck deliveries. Therefore, total cost for truck deliveries consists of C_o and C_w .

$$C_o = B_p \cdot N \quad (50)$$

$$C_w = \frac{(QZh)v_u}{2} \quad (51)$$

Equation (50) expresses the costs for system operation, which directly relate to the number of trucks and driver pay rate. Equation (51) includes the users' cost of waiting to receive packages, which is half the interval h multiplied by total demands and v_u . This user waiting cost applies similarly to all the delivery alternatives.

7.3.3.2 E-bike Deliveries Formulation

E-bike delivery model consists of bikes and trucks in the system, and its total cost follows a similar structure to truck deliveries. Since fulfillment for the bikes (e.g., a stationed truck) is conducted at the center of service region, trucks travel back and forth between a center point and depot as shown in Equation (46). Likewise, a line-haul distance for bikes is omitted.

7.3.3.3 Third-party Personal Car (TPC) Deliveries Formulation

TPC model follows the same structure to the previous alternatives. The key difference from the truck delivery is attributed to vehicle capacity S_p .

7.3.3.4 Personalized Ride (PR) Formulation

Total cost for personalized ride services includes the user riding as a cost. Note that C_w and C_r are doubled due to returning users back to their origins.

$$C_r = \frac{2L(QZh)v_i}{v_l} \quad (52)$$

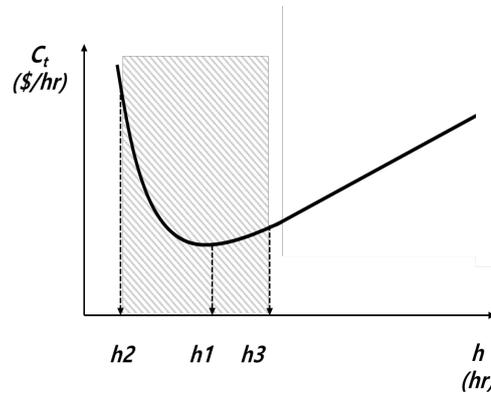
Each demand point represents user pick-up location (e.g., a customer's home). The number of packages QhA is replaced by the passenger. Equation (52) expresses the costs associated with the average in-vehicle travel time spent by users.

7.3.3.5 Parcel Locker Deliveries Formulation

Total cost for parcel locker deliveries consists of an operator, user waiting, and user access cost. Users in locker deliveries need to access their nearest locker, which increases user cost as follows:

$$C_x = \frac{v_x(QZ)(w+l)}{4V_k\sqrt{N_l}} \quad (53)$$

Average dwell time per locker T_m is set as a larger value than to other types of deliveries; a delivery person would place items in bulk to each locker. Note that average TSP distance L_t in Equation (48) is a distance for visiting all the lockers; namely, QhA is replaced by N_l . Average access distance for a service unit (i.e., a parcel locker) is one-fourth of the sum of length l and width w of Z ; the distance is inversely proportional to the square root of the number of lockers N_l under the assumption that these lockers are evenly distributed over Z . Then, user access cost C_x is derived as user access time multiplied by the value of time spent for access v_u .



7.3.3.7 Optimization

The optimal vehicle departure interval h , which minimizes the total cost function as well as meets the imposed constraints, is found by differentiating the objective function C_t with respect to h . Figure 26 shows how these sets of constraints affect total cost. Assuming h_1 to be the cost-minimizing departure interval, h_2 to be the interval bounded by working hour constraint (55), and h_3 to be the one bounded by capacity constraint (54), Case 1 in Figure 26 (a) shows the cost-minimizing h_1 to be optimum.

The optimal C_t is derived at h_3 for Case 2, while none of the intervals satisfy with Case 3. It should be remembered that the number of vehicles N is estimated by considering both vehicle capacity in Constraint (54) and the last-mile distance in Equation (48). On the contrary, Constraints (55) and (57) are the one adding realistic operational considerations in the system without imposing any changes to the system.

7.4 Numerical Results and Sensitivity Analyses

7.4.1 Results

Using the baseline inputs listed in Table 37, the results for each alternative are summarized in Table 38.

Table 38 Optimization Results of Alternatives

Results	Home-delivery			In-store Service	User pick-ups and drop-offs	
	Truck	e-bike		PR	Locker	
Headway, h (hr)	1.66	0.64		0.71	0.05	0.99
Delivery area, A (km ²)	7.8	9.1 (bike)	50.5 (truck)	5.2	-	32.6
Travel distance (km/vehicle tour)	70.8	23.2	32.2	54.2	9.2	45.4
Total Delivery time, T (hr/vehicle tour)	8.0	3.2	0.9	2.1	0.4	2.9

Number of vehicles, N	5	6	2	6	17	3
Number of lockers, N_l	-	-	-	-	-	4
Avg. load per vehicle	119.5	38.5	115.6	42.6	1	118.8
Load factor (%)	99.6	85.6	64.3	94.7	100	47.5
Costs elements (%)	Operating, C_o	52.7	82.2	75.8	49.2	7.8
	Waiting, C_w	47.3	17.8	24.2	0.7	7.0
	Access, C_x	-	-	-	-	85.2
	Riding, C_r	-	-	-	50.1	-
Total cost, C_t (\$/hr)	379	389	317	1,372	1,536	
Cost per delivery (\$/delivery)	0.63	1.68	1.23	40.38	4.31	
Operator cost per delivery (\$/delivery)	0.33	1.38	0.94	20.00	0.34	
Critical constraint equations	(54)	(54)	(54)	(54)	(56)	

* Note: TPC = third-party deliveries by personal car, PR = personalized ride services

In home-delivery services, five trucks N with a headway h of 1.66 hours and delivery area A of 7.8 km² optimize the total cost C_t as \$379 per hour. Travel distance per vehicle tour is computed as 70.8 km, by adding average TSP distance L_t to twice line-haul distance $2D$. Average load per vehicle indicates how many items or passengers are loaded per vehicle, and a load factor shows a percentage of the actual number of items and vehicle capacity. In this h and A combination, the operating and user waiting cost constitute 52.7% and 47.3%, respectively, of the total cost. Cost per delivery of \$0.63 is derived by dividing C_t by total demand $Q \cdot Z \cdot h$. The constraints that bound each alternative's cost function is listed in the last row of Table 38. In e-bike operation, the system outputs for e-bikes and fulfillment trucks are optimized concurrently; two trucks refill the items for six bikes according to bike replenishment schedule. Due to low consolidation h , both TPC and e-bikes show higher costs per delivery than the one for the truck.

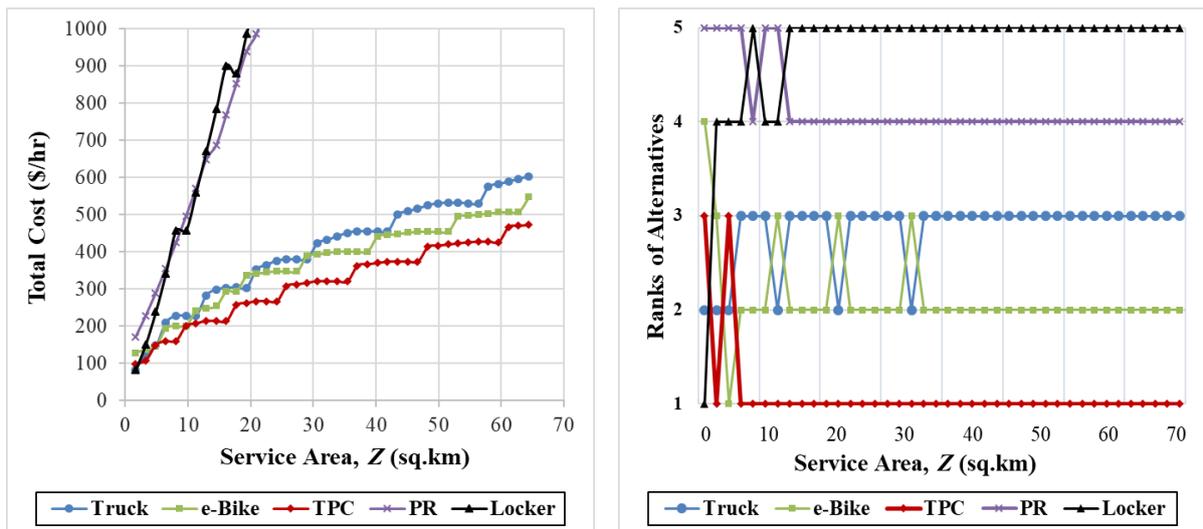
Among home deliveries, TPC shows the lowest total cost C_t due to 1) low user cost C_w resulted from the smaller consolidation time (i.e., the optimal headway h) and 2) the vehicle capacity constraint (54) resulting from storage capacity S_p . PR is an expensive alternative due

to the large N . The cost per delivery for PR is high due to small h and S_s . Note that PR's cost per delivery is one-way ride here. This result may seem surprisingly costlier than the actual operation, possibly for the following reasons: 1) An operator provides access to a few selected stores in the service region, which potentially reduce L_s by decreasing the coefficient in tour distance equation (48). 2) The driver pay rate B_p for PR may be calculated differently from other types, where B_p is decided by various factors, such as service region and surcharges associated with booking, driver supply, and time of day. For simplicity, the above traits are not considered. The majority fraction of total cost in parcel locker consists of user access cost C_x , and the operator cost per package is the lowest among the alternatives.

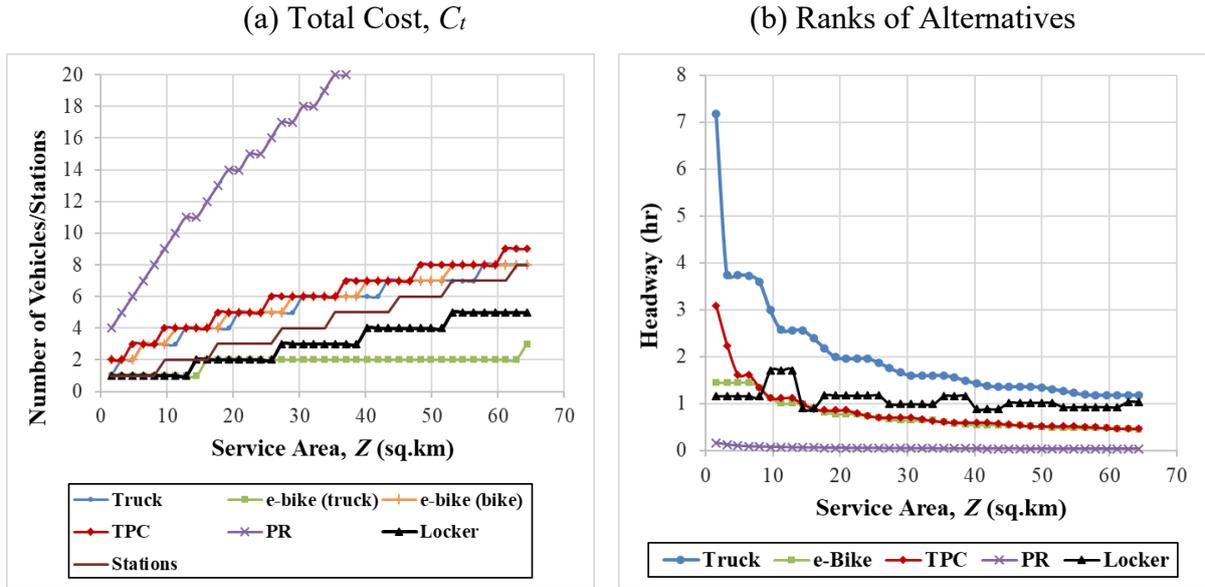
7.4.2 Sensitivity Analyses

7.4.2.1 The Effects of Service Area Size

As more demands are generated with an increase in Z , more delivery vehicles are needed to serve the increased users (the reverse is also true). This analysis finds the effectiveness and costs of alternatives as the size of Z changes. Z is examined from 2.6 km² to 103.6 km².



(Total costs for PR and locker increase up to 3,396\$/hr and 3,516\$/hr, respectively)



(The number of vehicles for PR increases up to 28)

(c) Number of Vehicles, N

(d) Headway, h

Figure 27 System Outputs for Changes in Service Area Z

Figure 27 (a) shows that total cost C_t increases with Z for all alternatives. Home-delivery services have moderate increases compared to PR and parcel lockers. The jumps in C_t correspond to the vehicles incrementally added to the system according to the integer constraint (57). More distinctive jumps are observed for the lockers due to increase in the number of lockers added to the system based on user access time constraint (56); user access cost C_x decreases as more stations are deployed. Figure 27 (b) indicates the ranks of each alternative regarding C_t ; these ranks change as vehicles and stations increase. The locker delivery is the cost-effective service alternative at small Z with a large consolidation h . Figure 27 (c) and (d) show the number of vehicles N and headway h , respectively, which vary with Z .

7.4.2.2 Changes in User Value of Time Spent Waiting for Goods

In Table 38, the user waiting cost C_w shares 0.7% to 47.3% of total cost C_t depending on alternatives. Since C_w is dependent on user value of time v_u , this section explores a case in

which v_u varies by regions or customers who have different values of time spent waiting for goods (Joeress et al. 2016). v_u ranges from about half to twice the baseline.

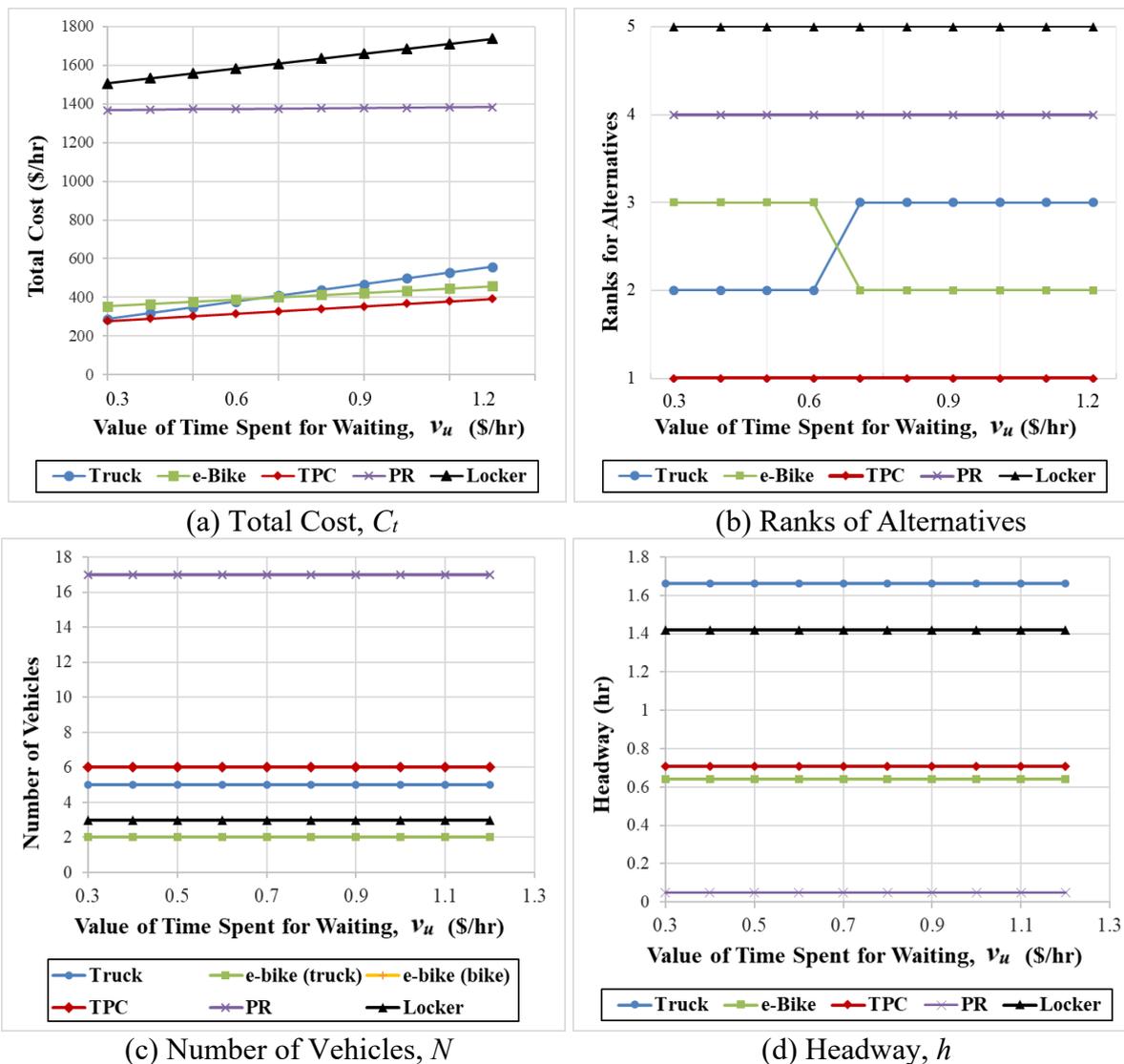


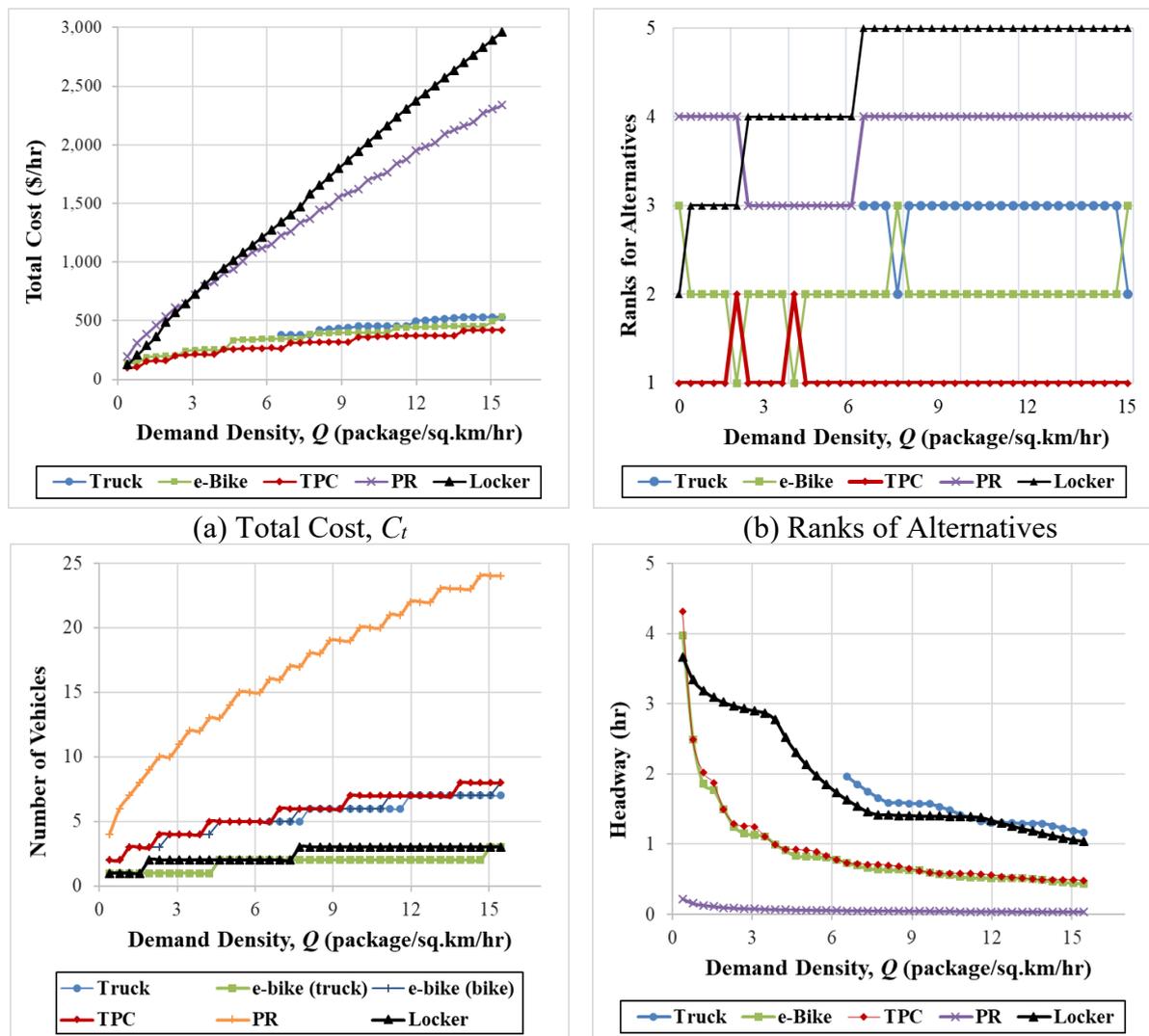
Figure 28 System Outputs for Changes in User Value of Time Spent for Waiting v_u

Although the departure interval h generally decreases with increases in user costs associated with v_u , the intervals stay unchanged in this range of changes due to the vehicle capacity constraint (54) in Figure 28. C_t for PR is nearly unchanged since the alternative is heavily dependent on user riding cost C_r rather than C_w from Table 38. Ranks are inverted between

trucks and e-bikes at v_u of 0.7, and it occurs because C_w constitutes more portions in truck's cost function than the others.

7.4.2.3 Variation in Demand Density

For exploring the effects of demand density Q on system performance, lower demand density may represent suburban areas while higher density may represent urban areas. The range of Q varies from 0.4 to 15.4 packages (or persons) per square kilometer per hour.



(The number of vehicles for PR increases up to 24.3)

(c) Number of Vehicles, N

(d) Headway, h

Figure 29 System Outputs for Changes in Demand Density Q

For truck deliveries below Q of 6.2, the optimal h is unable to be found in Figure 29 since h is unbounded by Constraints (54) and (55): i.e., a Case 3 of Figure 26. Although truck deliveries may be operable in such low Q , the result suggests that the operation is not economical. Either e-bike or PR is the cost-effective mode depending on Q .

7.4.2.4 Combined Deliveries by Trucks

Trucks may be utilized to deliver items to bike or locker fulfillment on the way customers' locations. This analysis is designed to explore trucks performing more than a single task in terms of cost-effectiveness under the assumption that demands are divided by the number of alternatives existed. In this case, total demands for each delivery alternative would be assigned by the given percentages to the alternatives, where the fraction is determined by satisfying working hour constraint (55). Therefore, vehicles serve the equal number of demands while providing more options. Note that the related delivery time T and stops increase; the cost for operating each alternative is added up. Although the baseline inputs remain unchanged, truck capacity S_t is determined by truck loading capacity (i.e., 250 packages) rather than by driver working period.

Two scenarios are designed. For Scenario 1, trucks carry out a door-to-door service while delivering items to an e-bike fulfillment station. Scenario 2 is that trucks serve customers and fulfill items at lockers. In case of 15% allocation for Scenario 1, it implies that 15% of demands are assigned by trucks while the rest is served by e-bikes.

Table 39 Results of Combined Delivery Based on Scenarios

Results	Scenario 1				Scenario 2	
	Truck & e-Bike				Truck & Locker	
<i>Assigned demand</i>	15%	30%			15%	30%
<i>Headway, h (hr)</i>	0.62	0.60			1.44	1.97
<i>Delivery area, A (km²)</i>	9.3	51.8	9.6	53.9	22.5	16.3
	<i>(bike)</i>	<i>(truck)</i>	<i>(bike)</i>	<i>(truck)</i>		

<i>Travel distance (km/vehicle tour)</i>	21.7	82.9	20.1	105.3	67.3	73.5
<i>Total Delivery time, T (hr/vehicle tour)</i>	2.8	4.5	2.4	7.3	5.7	7.9
<i>Number of vehicles, N</i>	5	8	5	13	4	4
<i>Number of lockers, N_l</i>	-	-	-	-	4	4
<i>Avg. load per vehicle</i>	44.9	28.1	43.3	16.6	129.6	177.3
<i>Load factor (%)</i>	99.8	11.2	96.2	6.7	51.8	70.9
<i>Costs Elements (%)</i>	<i>Operating, C_o</i>	88.5	91.7		8.2	11.8
	<i>Waiting, C_w</i>	11.5	8.3		15.4	20.4
	<i>Access, C_x</i>	-	-		76.4	67.8
<i>Total cost, C_t (\$/hr)</i>	587	785		1,428	1,289	
<i>Cost per delivery (\$/delivery)</i>	2.61	3.63		2.75	1.82	
<i>Operator cost per delivery (\$/delivery)</i>	2.31	3.33		0.23	0.22	
<i>Critical constraint equations</i>	(54)	(54)		(56)	(56)	

In Scenario 1 of Table 39, average package load per truck (or load factor for truck) is small since more trucks with underutilized capacity should be deployed. Truck and bike operations are jointly optimized by the same optimized h , where the trucks conducting their own last-mile deliveries are coordinated by the bike fulfillment schedule. Thus, the economic operation for Scenario 1 is not justified, and the combined service might as well serve customers within two separate services. On the other hand, Scenario 2 shows that C_t decreases with consolidation h compared to the results illustrated in Table 39. Therefore, the combined operation would be beneficial only for lockers for our baseline.

7.5 Discussion and Summary

7.5.1 Discussion

The suggested model is applied in analyzing the optimal delivery mode for fresh food delivery of the Washington Village/Pigtown section of Baltimore, Maryland. Washington Village/Pigtown is categorized into a food desert in Baltimore City (Chavis et al. 2018). The size of service area Z is approximately 9.6 km^2 with a population of 5,134. Grocery store density (i.e., the number of corner stores per 10,000 residents) is 38.2 (Baltimore City Health Department. 2017). Since demand density Q for this area is unavailable, the demand density of the area is estimated from the grocery store density divided by working periods W and service area size Z . As a result, Q becomes $0.26 \text{ packages/km}^2/\text{hr}$. Note that all the potential customers are assumed to use the delivery service.

Due to low demand density Q , the required delivery vehicles are much fewer than in the baseline shown in Table 37. The least expensive transportation mode turns out to be the parcel locker delivery. The optimal headways for home deliveries exceed working period W . This indicates that the delivery service operates every h hours; trucks would serve the study area Z every two day, for instance. In Table 40, note that 1) delivery area A is larger than the service area Z , which is feasible, and 2) the critical constraint is changed from vehicle capacity in Table 38 to working hours for truck delivery due to low Q .

Table 40 Results of Delivery Service for Washington Village/Pigtown

Results	Home-delivery			In-store Service	User pick-ups and drop-offs	
	Truck	e-bike	TPC	PR	Locker	
Headway, h (hr)	48.40	18.15	18.15	0.29	1.96	
Delivery area, A (km²)	46.4	9.1 (bike)	53.4 (truck)	37.0	-	483.3
Travel distance (km/vehicle tour)	71.1	24.0	32.2	56.2	4.2	57.5
Total Delivery time, T (hr/vehicle tour)	8.0	3.2	0.9	3.6	0.2	1.9
Number of vehicles, N	1	1	1	1	2	1
Number of lockers, N_l	-	-	-	-	-	1
Avg. load per vehicle	120.0	44.9	44.9	44.9	1	1.9
Load factor (%)	100	99.7	17.9	99.7	35.9	5.0
Costs elements (%)	Operating, C_o	52.6	85.6	74.8	96.9	80.5
	Waiting, C_w	47.4	14.4	25.2	2.6	3
	Access, C_x	-	-	-	-	16.5
	Riding, C_r	-	-	-	0.5	-
Total cost, C_t (\$/hr)	76.0	93.5	53.5	82.6	49.7	49.7
Cost per delivery (\$/delivery)	0.63	2.08	1.19	57.4	10.2	10.2
Operator cost per delivery (\$/delivery)	0.33	1.78	0.89	55.6	8.07	8.07
Critical constraint equations	(55)	(54)	(54)	(54)	(54)	(56)

7.5.2 Summary

An area with limited access to fresh products within reasonable distances and prices was called a food desert. This became a public health concern associated with lower consumption of fresh foods. To mitigate this, the study aimed to develop a cost-effective last-mile fresh food delivery system that addressed the lack of mobility. The chapter identified and optimized five delivery alternatives: conventional trucks, e-bikes, personalized ride transportation services, parcel lockers, and third-party deliveries by personal car. The optimized outputs for alternatives were compared with total cost. Sensitivity analyses were conducted in terms of 1) service area

size, 2) demand density, 3) user value of time spent waiting for goods, and 4) combined deliveries by trucks.

Numerical results showed that third-party deliveries by personal car were the most cost-effective option in delivering fresh items, while the truck delivery ranked second for our baseline values. The personalized ride service and parcel locker delivery were more expensive than home-delivery services. Although more vehicles and frequent trips were needed with an increase in service area size across alternatives, home-delivery services had moderate increases in total cost compared to other types. The personalized ride was less influenced by changes in user value of waiting time. At a low demand density, the truck operation may not be economically operable. The study explored trucks performing more than a single task; trucks carried out a door-to-door service while delivering items to e-bike fulfillment stations or parcel lockers. Only the latter use of trucks was economically justifiable.

Future extensions of this study may include the following. By applying real-world inputs to the suggested model, more specific variables may be considered, such as the effects of roadway network configuration or dividing service areas into several. Sensitivity to changes in public policy variables such as tax incentives to participating grocers may be considered to identify a practical business model that public agencies can manage in collaboration with grocers and carriers. Although the study assumes that all packages are insulated with appropriate temperature-controlled packaging similarly to private meal-kit delivery services, researchers may consider deliveries without the packaging. Then, the mandatory completion time for a delivery tour can be imposed in the model. Finally, the user value of time spent waiting for goods may be explored. Although baseline inputs are intended for food deserts in urban

circumstances, the model can be tailored for rural settings with proper service area size, demand density, and a reasonable coefficient for the distance approximation equation.

This analysis compares alternatives based on their relative costs. Further studies might also compare such alternatives in terms of service quality, capacity, and suitability for various environments.

Chapter 8: Conclusions and Future Research

8.1 Research Summary and Contributions

The dissertation develops the tour length approximations for Traveling Salesman Problem (TSP) when the number of visited points n is relatively small. The principal contributions of this dissertation are underlined and summarized below.

Chapter 2 provides an extensive review of existing research work for the TSP approximations and solution methods. The approximations dealing with small n values are reviewed, while a total of fourteen metaheuristics and TSP solvers are compared in terms of solution accuracy.

Chapter 3 develops the TSP approximations through few points. The approximation models account for various factors, such as area shapes, elongations, point distribution, and depot locations. The optimized TSP instances are further investigated using statistical analysis (e.g., some extreme values, variance, and normality). The effects of those factors on tour lengths are explored.

Chapter 4 introduces some extensions for the approximations. First, a total of 6 adjustment factors are proposed that integrate the above considerations into a single equation. The estimation of the exponent for the number of points n is a key contribution in this dissertation since the previous studies apparently assumed without checking that it should be 0.5, i.e., that tour lengths should vary with the square root of n . When subjecting this exponent to statistical estimation, it is found that its value can be considerably smaller than 0.5. With this estimated exponent, the TSP tour lengths can be more precisely approximated than with the previous models. Second, approximations for probabilistic TSP are developed to reflect stochastic

customer presence (or customer's acceptance of the service). Such approximations are beneficial for analyzing how changes in demand affect tour lengths when n is known. Third, the approximated tour lengths are compared with the actual distances for rural and urban delivery networks. Urban areas have favorable conditions (e.g., point distribution) for satisfying imposed approximation assumptions, and thus the approximated and actual tour lengths differ by as little as 9.6%.

Chapters 5, 6, and 7 present applications of the tour length approximations; most of the vehicles considered in these chapters have a low vehicle carrying capacity. In Chapter 5, a comparison of deliveries by robots, drones, and trucks is presented. The total cost of each alternative is formulated and then optimized for comparison. Some sensitivities are investigated, such as to changes in energy cost, user value of waiting time for delivery, and carrying capacity.

The next application formulates and compares four alternatives of package delivery service with and without the aid of drones. Each delivery alternative is optimized numerically with an objective of total cost minimization. Analyses are conducted with respect to sensitivity to driver pay rate, demand density, user value of waiting time for delivery, drone operating speed, service area size, drone size, and distribution hub location. For reasonable baseline inputs, results indicate that using drones for package deliveries may be cost-effective compared to using conventional trucks.

In Chapter 7, a last-mile fresh food delivery system is proposed for individuals in underserved communities. Five delivery alternatives with various modes are considered. The total cost is formulated and optimized for each alternative. Then, the optimized results for the alternatives are compared. The dissertation examines whether delivery trucks could perform

multiple tasks (i.e., delivering items to customers and fulfillment centers in a single delivery tour). Lastly, mode shares for home deliveries were estimated when multiple delivery alternatives coexisted.

Thus, the key contributions are highlighted as follows. First, Beardwood's approximations have been refined by considering adjustments for various factors. Second, the exponent for n values is statistically estimated rather than assumed to be 0.5. These improvements help estimate accurate TSP tour lengths and solve large system design problems, in which the exact demand points are uncertain the time of planning.

8.2 Future Research

Although this dissertation makes distinct contributions in developing TSP tour length approximations, some of the following potential extensions may be considered in future research.

1. Some considerations for TSP formulations:

- (i) Constraints on the sequence of visits:** some items may need to be dropped earlier than the others, (e.g., heavy items or time-sensitive deliveries) although that tends to increase the tour length. The effect of these conditions on the tour lengths can be analyzed.
- (ii) Changes in the objective function for TSPs and resulting approximations:** the current objective function for TSP instances is to minimize a function related to cost (i.e., the tour length). The function can be changed according to the user's intentions. While minimizing the total distance traveled for the TSPs, the objective function can be modified, e.g., to optimize coverage or maximize profitability.

Some points may be excluded from a tour if the marginal profitability of delivery is considered. For instance, if most of the points to be visited are clustered, some of the remaining points may be skipped and possibly visited on the next tour. The resulting approximations from the optimized TSPs can be compared with the results in Table 8.

2. **Exploration of other forms for adjustment factors:** only a multiplicative form is considered for combining all adjustment factors in Equation (9). The absolute percent errors can be thoroughly investigated when the factors are multiplied (e.g., errors attributed to multicollinearity). Alternatively, other possible forms can be thoroughly explored to reduce the errors.
3. **Representation of the actual road network in the context of approximation:** the actual road network may be more precisely represented in the approximation by considering non-uniform point distributions or circuitry factors, as discussed in Section 4.3. It would be worth investigating which factors should be accounted for and how the approximated tour lengths can better reflect the actual network characteristics.
4. **Various probability distributions for probabilistic TSP:** stochastic customer presence can be modeled with different probabilities, such as Poisson distribution or normal distribution. The estimates (i.e., exponent and coefficient) can be re-computed and compared.

References

- Applegate, D., Bixby, R., Chvatal, V., and Cook, W. On the Solution of Traveling Salesman Problems. *Documenta Mathematica Journal der Deutschen Mathematiker-Vereinigung, International Congress of Mathematicians*, pp. 645–656. 1998 <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.53.3142>
- Applegate, D., Bixby, R., Chvatal, V., and Cook, W. The Traveling Salesman Problem: A Computational Study. *Princeton University Press*. 2006.
- Ansari, S., Basdere, M., Li, X., Ouyang, Y., and Smilowitz, K. Advancements in Continuous Approximation Models for Logistics and Transportation Systems: 1996–2016. *Transportation Research Part B*, 107. 2018. <https://doi.org/10.1016/j.trb.2017.09.019>
- Antosiewicz, M., Koloch, G., and Kamiński, B. Choice of Best Possible Metaheuristic Algorithm for the Travelling Salesman Problem with Limited Computational Time: Quality, Uncertainty and Speed. *Journal of Theoretical and Applied Computer Science*. Vol. 7, No. 1, pp. 46-55. 2013
- Adewole. A., Otubamowo, K., Egunjobi, T. A Comparative Study of Simulated Annealing and Genetic Algorithm for Solving the Travelling Salesman Problem. *International Journal of Applied Information Systems*. Vol. 4. No.4, 2012. <http://doi.org/10.5120/ijais12-450678>
- Arlotto, A., and Steele, M. Beardwood–Halton–Hammersley Theorem for Stationary Ergodic Sequences: a Counterexample. *The Annals of Applied Probability*. Vol. 26, No. 4, pp. 2141–2168. 2016. <https://doi.org/10.1214/15-AAP1142>
- Ansari, A., Ibraheem., and Katiyar, S. Comparison and Analysis of Solving Travelling Salesman Problem using GA, ACO and Hybrid of ACO with GA and CS, *IEEE Workshop on Computational Intelligence: Theories, Applications and Future Directions*. 2015. <http://doi.org/10.1109/WCI.2015.7495512>
- Abdulkarim, H., and Alshammari, I. Comparison of Algorithms for Solving Traveling Salesman Problem, *International Journal of Engineering and Advanced Technology*, Vol. 4. Issue. 6, 2015. <https://www.ijeat.org/wp-content/uploads/papers/v4i6/F4173084615.pdf>
- Ballou, R., Rahardja, H., and Sakai, N. Selected Country Circuitry Factors for Road Travel Distance Estimation. *Transportation Research Part A: Policy and Practice*. Vol. 36, Issue 9, , pp. 843-848. 2002. [https://doi.org/10.1016/S0965-8564\(01\)00044-1](https://doi.org/10.1016/S0965-8564(01)00044-1)
- Baltimore City Health Department. 2017 Neighborhood Health Profile Report. <https://health.baltimorecity.gov/neighborhoods/neighborhood-health-profile-reports>

Bass, J., Leveraging Drones for Delivery: Our Vision to Improve Safety, the Environment, and Customer Service. *Presented at 98th Annual Meeting of the Transportation Research Board*, Washington, D.C., 2019.

Beardwood, J., Halton, J., and Hammersley, J. The Shortest Path through Many Points. *Mathematical Proceedings of the Cambridge Philosophical Society*. vol 55, pp. 299–327, 1959.

Boysen, N., Schwerdfeger, S., and Weidinger, F. Scheduling Last-mile Deliveries with Truck-based Autonomous Robots. *European Journal of Operational Research*. Volume 271, Issue 3, 16 December 2018, Pages 1085-1099. <https://doi.org/10.1016/j.ejor.2018.05.058>

Burgstaller, B and Pillichshammer, F. The Average Distance between Two Points. *Bulletin of the Australian Mathematical Society*. vol. 80, no. 3, pp. 353–359, 2009. <https://doi.org/10.1017/S0004972709000707>

Brunetti, R., Krauth, W., Mézard, M., and Parisi, G. Extensive Numerical Simulations of Weighted Matchings: Total Length and Distribution of Links in the Optimal Solution. *Europhysics Letters*. Vol. 14, No. 4. 1991. <https://iopscience.iop.org/article/10.1209/0295-5075/14/4/002>

Bakach, I., Campbell, A., and Ehmke, J. A Two-Tier Urban Delivery Network with Robot-based Deliveries, Working Paper. *The Otto von Guericke University Magdeburg*. 2020

Bloomberg, Autonomous Last Mile Delivery Market Worth \$91.5 Billion by 2030, *Bloomberg*. 2019. <https://www.bloomberg.com/press-releases/2019-07-15/autonomous-last-mile-delivery-market-worth-91-5-billion-by-2030-exclusive-report-by-marketsandmarkets>

Beed, R., Sarkar, S., Roy, A., and Chatterjee, S. A Study of the Genetic Algorithm Parameters for Solving Multi-objective Travelling Salesman Problem, *2017 International Conference on Information Technology*, Bhubaneswar, 2017, pp. 23-29. <https://www.cs.unh.edu/~sc1242/publications/08423877.pdf>

Campbell, J, F., Sweeney, D, C., Zhang, J., and Pan, D. Strategic Design for Delivery with Linked Transportation Assets: Trucks and Drones. *Midwest Transportation Center*. 2017.

Cavdar, B., and Sokol, J. A Distribution-free TSP Tour Length Estimation Model for Random Graphs, *European Journal of Operational Research*. vol. 243, no. 2, pp. 588-598, 2015. <https://doi.org/10.1016/j.ejor.2014.12.020>

Chowdhury, S., Emelogu, A., Marufuzzaman, M., Nurre, S, G., and Bian, L. Drones for disaster response and relief operations: A continuous approximation model. *International Journal of Production Economics*. 188:167-184. 2017. <https://doi.org/10.1016/j.ijpe.2017.03.024>

- Cuthbertson, A. Self-driving Delivery Robots are Hitting the Streets in the U.K. 2016. <http://www.newsweek.com/self-driving-robot-deliverieslondon-starship-technologies-driverless-435414>,
- Chavis, C., Davis, T., Norris, T., and Jones, A. An Analysis of Food Accessibility Using the USDA Food Acquisition and Purchase Survey. *Transportation Research Board 97th Annual Meeting*, online compendium. 2018.
- Choi, Y., and Schonfeld, P. Drone Deliveries Optimization with Battery Constraints. *Presented at 97th Annual Meeting of the Transportation Research Board*, Washington, D.C., 2018.
- Christofides, N., and Eilon, S. Expected Distances in Distribution Problems. *Operational Research Quarterly*. 20(4), 437–443. 1969. <http://doi.org/10.2307/3008762>
- Chien, W. Operational Estimators for the Length of a Traveling Salesman Tour. *Computers & Operations Research*, 19(6), 469–478. 1992. [https://doi.org/10.1016/0305-0548\(92\)90002-M](https://doi.org/10.1016/0305-0548(92)90002-M)
- Conway, A., Fatisson, P., Eickemeyer, P., Cheng, J., and Peters, D. Urban Micro-Consolidation and Last Mile Goods Delivery by Freight-Tricycle in Manhattan: Opportunities and Challenges. *Transportation Research Board 91th Annual Meeting*, online compendium. 2011.
- Crisan, G., Nechita, E., and Simian, D. On Randomness and Structure in Euclidean TSP Instances: A Study With Heuristic Methods. *IEEE Access*, Vol. 9. pp. 5312 - 5331. 2021. <https://doi.org/10.1109/ACCESS.2020.3048774>
- DHL. DHL Express Launches Its First Regular Fully-Automated and Intelligent Urban Drone Delivery Service, 2019. <https://www.dhl.com/tw-en/home/press/press-archive/2019/dhl-express-launches-its-first-regular-fully-automated-and-intelligent-urban-drone-delivery-service.html>
- D'Agostino, R. Goodness-of-Fit-Techniques. *CRC Press*. 1986.
- Daganzo, C. The Length of Tour in Zones of Different Shapes. *Transportation Research Part B: Methodology*, vol. 18B, no. 2, pp. 135–145, 1984. [https://doi.org/10.1016/0191-2615\(84\)90027-4](https://doi.org/10.1016/0191-2615(84)90027-4)
- Dorling, K., Heinrichs, J., Messier, G., and Magierowski, S. Vehicle Routing Problems for Drone Delivery. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*. 47(1): 70-85. 2017. <https://doi.org/10.1109/TSMC.2016.2582745>
- Damghanijazi, E., and Mazidi, A. Meta-Heuristic Approaches for Solving Travelling Salesman Problem, *International Journal of Advanced Research in Computer Science*, Vol. 8 No. 5. 2017.

Figliozzi, M. Lifecycle Modeling and Assessment of Unmanned Aerial Vehicles (Drones) CO2e Emissions. *Transportation Research Part D: Transport and Environment*. vol. 57, pp. 251-261, 2017. <https://doi.org/10.1016/j.trd.2017.09.011>

Figliozzi, M, Planning Approximations to the Average Length of Vehicle Routing Problems with Varying Customer Demands and Routing Constraints, *Transportation Research Record: Journal of the Transportation Research Board*. No. 2089, 2008, pp. 1–8. <https://doi.org/10.1016/j.trb.2008.01.011>

Fiechter, C. A Parallel Tabu Search Algorithm for Large Traveling Salesman Problems, *Discrete Applied Mathematics*. Vol. 51. pp. 243-267. 1994. [https://doi.org/10.1016/0166-218X\(92\)00033-1](https://doi.org/10.1016/0166-218X(92)00033-1)

Franceschetti, A., Jabali, O. & Laporte, G. Continuous Approximation Models in Freight Distribution Management. *TOP*. Vol. 25, pp. 413–433. 2017. <https://doi.org/10.1007/s11750-017-0456-1>

Ferrandez, S, M., Harbison, T., Weber, T., Sturges, R., and Rich, R. Optimization of a Truck-drone in Tandem Delivery Network Using K-means and Genetic Algorithm. *Journal of Industrial Engineering and Management*. Vol. 9(2). 2016. <http://dx.doi.org/10.3926/jiem.1929>

Fulfillment by Amazon, 2019. <https://services.amazon.com/fulfillment-by-amazon/pricing.html>

Finch, S. *Mathematical Constants*. Cambridge University Press, Cambridge. 2003.

Green, S. How Many Subjects Does It Take To Do A Regression Analysis. *Multivariate Behavioral Research*, Vol. 26, pp. 499–510. 1991. https://doi.org/10.1207/s15327906mbr2603_7

Gruber, J., Kihm, A., and Lenz, B. A New Vehicle for Urban Freight? An Ex-ante Evaluation of Electric Cargo Bikes in Courier Services. *Research in Transportation Business & Management*. Vol. 11. pp. 53-62. 2014. <https://doi.org/10.1016/j.rtbm.2014.03.004>

Gupta, D. Solving TSP Using Various Meta-Heuristic Algorithms, *International Journal of Recent Contributions from Engineering, Science & IT*. Vol 1, No 2. 2013.

Gupta, S., Rana, A., and Kansal, V. Comparison of Heuristic Techniques: A Case of TSP. *10th International Conference on Cloud Computing, Data Science & Engineering*. 2020. <https://doi.org/10.1109/Confluence47617.2020.9058211>

- Hindle, A., and Worthington, D. Models to Estimate Average Route Lengths in Different Geographical Environments, *Journal of the Operational Research Society*, 2004. <https://doi.org/10.1057/palgrave.jors.2601751>
- Holguín-Veras, J., and Patil, G. Observed Trip Chain Behavior of Commercial Vehicles. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1906, pp. 74–80. 2005. <https://doi.org/10.1177/0361198105190600109>
- Hahsler, M., and Hornik, K. TSP – Infrastructure for the Traveling Salesperson Problem. *Journal of Statistical Software*. Vol. 23. Issue 2. 2007. <https://dx.doi.org/10.18637/jss.v023.i02>
- Ham, A, M. Integrated Scheduling of m-truck, m-drone, and m-depot Constrained by Time-window, Drop-pickup, and m-visit Using Constraint Programming. *Transportation Research Part C: Emerging Technology*. Vol. 91. pp. 1-14. 2018 <https://doi.org/10.1016/j.trc.2018.03.025>
- Hassanat, A., Almohammadi, K., Alkafaween, E., Abunawas, E., Hammouri, A., and Prasath, V. Choosing Mutation and Crossover Ratios for Genetic Algorithms—A Review with a New Dynamic Approach. *Information*. 2019. <https://doi:10.3390/info10120390>
- Helsgaun, K. An Effective Implementation of the Lin-Kernighan Traveling Salesman Heuristic. *European Journal of Operational Research*. vol. 126. pp. 106-130. 2000. [https://doi.org/10.1016/S0377-2217\(99\)00284-2](https://doi.org/10.1016/S0377-2217(99)00284-2)
- Hoos, H., and Stuzle, T. On the Empirical Scaling of Run-time for Finding Optimal Solutions to the Travelling Salesman Problem. *European Journal of Operational Research*. Vol. 238, Issue 1, 1. Pages 87-94. 2014. <https://doi.org/10.1016/j.ejor.2014.03.042>
- Jaillet, P. Probabilistic Traveling Salesman Problems, Ph.D. thesis, Massachusetts Institute of Technology. 1985. <https://web.mit.edu/jaillet/www/general/jaillet-phd-mit-orc-85.pdf>
- Jennings, D., and Figliozzi, M. Study of Sidewalk Autonomous Delivery Robots and Their Potential Impacts on Freight Efficiency and Travel, *Transportation Research Record: Journal of the Transportation Research Board*. 2019. <https://doi.org/10.1177/0361198119849398>
- Joerss, M., Schröder, J., Neuhaus, F., Klink, C., and Mann, F. Parcel delivery: The future of Last Mile. *McKinsey&Company*. 2016.
- Johnson, D., McGeoch, L., and Rothberg, E. Asymptotic Experimental Analysis for the Held-Karp Traveling Salesman Bound. *Proceedings of the 7th Annual ACM-SIAM Symposium on Discrete Algorithms*. pp. 341-350. 1996. <https://dl.acm.org/doi/pdf/10.5555/313852.314081>

- Kaufman, A., King, G., Komisarchik, M. How to Measure Legislative District Compactness If You Only Know it When You See it. *American Journal of Political Science*. 2017. <https://j.mp/2Fs3ESc>
- Kumar, S., and Panneerselvam, R. A Survey on the Vehicle Routing Problem and Its Variants. *Intelligent Information Management*, Vol. 4. No. 3. pp. 66-74. 2012. <http://dx.doi.org/10.4236/iim.2012.43010>
- Korman, R. What is 'safe enough' for Drone Deliveries? *The Seattle Times*. 2019. <https://www.seattletimes.com/business/boeing-aerospace/what-is-safe-enough-for-drone-deliveries/>
- Kwon, O., Golden, B., and Wasil, E. Estimating the Length of the Optimal TSP tour: An Empirical Study Using Regression and Neural Networks. *Computers & Operations Research*, Vol. 22. Issue. 10. pp. 1039–1046. 1995. [https://doi.org/10.1016/0305-0548\(94\)00093-N](https://doi.org/10.1016/0305-0548(94)00093-N)
- Kweon, H. Comparisons of Estimated Circuitry Factor of Forest Roads with Different Vertical Heights in Mountainous Areas, Republic of Korea. *Forests*. Vol. 10(12), 1147. 2019. <https://doi.org/10.3390/f10121147>
- Kornatowski, P. M., Bhaskaran, A., Heitz, G., Mintchev, Stefano., and Floreano. D. Last-Centimetre Personal Drone Delivery: Field Deployment and User Interaction. *IEEE Robot. Auto. Lett.* 3(4): pp. 3813-3820. 2018. <https://doi.org/10.1109/LRA.2018.2856282>
- Larson, R., and Odoni, A. Urban Operation Research. *Dynamic Ideas*. 1981.
- Lee J., Choi M.Y. Optimization by Multicanonical Annealing and the Traveling-Salesman Problem. *Computer Simulation Studies in Condensed-Matter Physics VII*. Springer Proceedings in Physics, Vol. 78. Springer, Berlin, Heidelberg. 1994. https://doi.org/10.1007/978-3-642-79293-9_19
- Lei, H., Laporte, G., Liu, Y., and Zhang, T. Dynamic Design of Sales Territories. *Computers & Operations Research*, Vol. 56, pp. 84-92. 2015. <http://dx.doi.org/10.1016/j.cor.2014.11.008>
- Lin, S., and Kernighan, W. An Effective Heuristic Algorithm for the Traveling-Salesman Problem. *Operations Research*, Vol. 21, no. 2, pp. 498-516, 1973. <https://www.jstor.org/stable/169020>
- Lucas, P. R Program to Estimate Time Consumption of Brutal Force in TSP. 2018. <https://github.com/LucasLP/TSP-Brute-Force-Time-Consumption>

- Madani, A., Batta, R., and Karwan, M. The Balancing Traveling Salesman Problem: Application to Warehouse Order Picking. *TOP*. 2020. <https://doi.org/10.1007/s11750-020-00557-y>
- Mahalanobis, P. A Sample Survey of the Acreage under Jute in Bengal. *The Indian Journal of Statistics*. Vol. 4, no. 4, pp. 511-530, 1940. <https://www.jstor.org/stable/40383954>
- Marks, E. A Lower Bound for the Expected Travel Among m Random Points. *The Annals of Mathematical Statistics*, Vol. 19, no. 3, pp. 419-422, 1948. <https://www.jstor.org/stable/2235651>
- Moler, C. Numerical Computing with MATLAB. 2008. *Society for Industrial and Applied Mathematics*. ISBN: 0898716608
- Mei, X. Approximating the Length of Vehicle Routing Problem Solutions Using Complementary Spatial Information. *Ph.D. Dissertation. George Mason University*. 2015. <http://mars.gmu.edu/handle/1920/9623>
- Nicola, D., Vetschera, R., and Dragomir, A. Total Distance Approximations for routing solutions, *Computers and Operations Research*. Vol. 102. pp. 67–74. 2019. <https://doi.org/10.1016/j.cor.2018.10.008>
- Ong, H. L., and Huang, H. C. Asymptotic Expected Performance of Some TSP Heuristics: An Empirical Evaluation. *European Journal of Operational Research*. Vol. 43, 231–238. 1989. [https://doi.org/10.1016/0377-2217\(89\)90217-8](https://doi.org/10.1016/0377-2217(89)90217-8)
- Office of the Secretary. Benefit-Cost Analysis Guidance for Discretionary Grant Programs, *U.S. Department of Transportation*, 2018.
- Phillip, J. The Probability Distribution of the Distance between Two Random Points in a Box. *KTH mathematics*, Royal Institute of Technology, 2007.
- Potvin, J. Genetic Algorithms for the Traveling Salesman Problem, *Annals of Operations Research*, Vol. 63, pp. 339-370, 1996. <https://doi.org/10.1007/BF02125403>
- Percus, A., and Martin, O. Finite Size and Dimensional Dependence in the Euclidean Traveling Salesman Problem, *Physical Review Letter*. 76. pp. 1188–1191. 1996. <https://doi.org/10.1103/PhysRevLett.76.1188>
- Phantom 4 Pro Specs. 2019. <https://www.dji.com/phantom-4-pro/info>
- Rose, C. Amazon's Jeff Bezos Looks to the Future. *CBS News*. 2013.

Rote, G. The Degree of Convexity, *EuroCG 2013*, Braunschweig, Germany, March 17–20, 2013.

Robusté, F., Estrada, M., and López-Pita, A. Formulas for Estimating Average Distance Traveled in Vehicle Routing Problems in Elliptic Zones. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1873, 2004, pp. 64–69. <https://doi.org/10.3141/1873-08>

Rabta, B., Wankmuller, C., and Reiner, G. A Drone Fleet Model for Last-mile Distribution in Disaster Relief Operations. *International Journal of Disaster Risk Reduction*. Vol. 28. Pp. 107-112. 2018. <https://doi.org/10.1016/j.ijdr.2018.02.020>

Reinelt, G. TSPLIB—A Traveling Salesman Problem Library, *ORSA Journal on Computing*. Vol. 3(4). pp. 376-384. 1991. <https://dx.doi.org/10.1287/ijoc.3.4.376>

Rexhepi, A., Maxhuni, A., and Dika, A. Analysis of the impact of parameters values on the Genetic Algorithm for TSP. *IJCSI International Journal of Computer Science Issues*, Vol. 10, Issue 1, No 3, 2013.

Stern, H. Polygonal Entropy: A Convexity Measure. *Pattern Recognition Letters*, Vol. 10. Issue. 4. pp. 229-235, 1989. [https://doi.org/10.1016/0167-8655\(89\)90093-7](https://doi.org/10.1016/0167-8655(89)90093-7)

Stein, D. Scheduling Dial-a-Ride Transportation Systems: An Asymptotic Approach, *Ph.D. Dissertation, Harvard University*, Cambridge, MA, 1977.

Sheth, M, Butrina, P, Goodchild, A, and McCormack, E. Measuring Delivery Route Cost Trade-offs between Electric-assist Cargo Bicycles and Delivery Trucks in Dense Urban Areas. *European Transport Research Review*. 2019. <https://doi.org/10.1186/s12544-019-0349-5>

Sonneberg, M., Leyerer, M., Kleinschmidt, A., Knigge, F., and Breitner, M. Autonomous Unmanned Ground Vehicles for Urban Logistics: Optimization of Last Mile Delivery Operations. *Proceedings of the 52nd Hawaii International Conference on System Sciences*. 2019. <http://hdl.handle.net/10125/59594>

[Shayanfar, E. Prioritizing and Scheduling Interrelated Road Projects Using Metaheuristic Algorithms. Master Thesis. University of Maryland. 2015](#)

Schliwa, G., Armitage, R., Aziz, S., Evans, J., and Rhoades, J. Sustainable City Logistics — Making Cargo Cycles Viable for Urban Freight Transport. *Research in Transportation Business & Management*. Vol. 15. pp. 50-57. 2015. <https://doi.org/10.1016/j.rtbm.2015.02.001>

Tipagornwong, C., and Figliozzi, M. Analysis of Competitiveness of Freight Tricycle Delivery Services in Urban Areas. *Transportation Research Record: Journal of the Transportation Research Board*, 2410, 76–84. 2014. <https://doi.org/10.3141/2410-09>

UPS Pressroom. UPS Drone Delivery Takes Operational Efficiency to New Heights. 2017.

UPS Pressroom, UPS Subsidiary UPS Flight Forward will Use Drones to Deliver Prescription Medicines from a CVS Pharmacy to The Villages, Florida for the Largest U.S. Retirement Community Home, to more than 135,000 Residents. 2020. <https://www.pressroom.ups.com/pressroom/ContentDetailsViewer.page?ConceptType=PressReleases&id=1587995241555-272>

U.S Energy Information Administration, Electric Power Monthly Data. 2019. https://www.eia.gov/electricity/monthly/current_month/epm.pdf

U.S. Energy Information Administration, Independent Statics & Analysis. 2015. <https://www.eia.gov/todayinenergy/detail.php?id=20372>

Vedenyov, M. Travelling Salesman Problem with Genetic algorithm, *MATLAB Central File Exchange*. 2020. <https://www.mathworks.com/matlabcentral/fileexchange/31818-travelling-salesman-problem-with-genetic-algorithm>

Vinel and Silve. Probability Distribution of the Length of the Shortest Tour between a Few Random Points: a Simulation Study, Proceedings of the 2018 Winter Simulation Conference. 2018. <http://doi.org/10.1109/WSC.2018.8632175>

Wang, X., Poikonen, S., and Golden, B. The Vehicle Routing Problem with Drones: Several Worst-case Results. *Optimization Letters*. Vol. 11. pp. 679-697. 2017. <https://doi.org/10.1007/s11590-016-1035-3>

Wissink, P. The Probabilistic Travelling Salesman Off the Beaten Track. Ph.D. thesis, University of Edinburgh. 2019. <https://era.ed.ac.uk/handle/1842/35764>

Yang, H., Liang, X., Zhang, Z., Liu, Y., and Abid, M. Statistical Modeling of Quartiles, Standard Deviation, and Buffer Time Index of Optimal Tour in Traveling Salesman Problem and Implications for Travel Time Reliability. *Transportation Research Record: Journal of the Transportation Research Board*. 2020. <https://doi.org/10.1177/0361198120954867>

Yap, B., Sim, C. Comparisons of Various Types of Normality Tests, *Journal of Statistical Computation and Simulation*. Vol. 81, No. 12, 2011, 2141–2155. <https://doi.org/10.1080/00949655.2010.520163>

Zunic, J., and Rosin, P. A New Convexity Measure for Polygons, *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol. 26, No. 7, 2004.
<https://doi.org/10.1109/TPAMI.2004.19>