2019-01-01

Applications Of A New Genetic Algorithm To Solve The Centralized Carrier Collaboration And Multihub Location Problem Considering Environmental Impacts

Eduardo Jose Castillo Fatule
University of Texas at El Paso, ejcastillof@gmaill.com

APPLICATIONS OF A NEW GENETIC ALGORITHM TO SOLVE THE CENTRALIZED CARRIER COLLABORATION AND MULTIHUB LOCATION PROBLEM CONSIDERING ENVIRONMENTAL IMPACTS

EDUARDO JOSE CASTILLO FATULE

Master's Program in Industrial Engineering

APPROVED:

Jose F. Espiritu, Ph.D., Chair

Heidi A. Taboada, Ph.D.

Carlos Chang, Ph.D.

Stephen Crites, Ph.D.
Dean of the Graduate School

Dedication

To my roommates, Albie and Joel, who have tolerated me and I them for the past few years, and without whom this thesis would've been completed much sooner.
To my friends, Katya, Peter, Alain, Rene, Oscar, Moises, Victor, Colby, Jordan, etc., for your company and moral support.
To my coworkers, Ana, Juan, Oswaldo, Luis, Trini, Yasser, Pedro, Gina, Ileana, Claudia etc., who've stood by me as I complained and procrastinated every step of the way.
To my family, who managed to not care to ask about this thesis despite constantly bothering me to find a job, and believing that that's a possibility in this economy.
To my parents specifically, for all the nagging and love.
To my siblings, for some reason.
To my grandparents, for all their love and caring.
In memory of my grandmother Josefina. I wish you would have been able to see me walk in the ceremony.
Lastly to my aunt Cassie, for receiving me in her home every Thanksgiving as well as her support and understanding.
To my professors, who taught me a lot of what I know.
To my advisors, Dr. Espiritu and Dr. Taboada, thank you for your guidance in the past five years.

APPLICATIONS OF A NEW GENETIC ALGORITHM TO SOLVE THE CENTRALIZED CARRIER COLLABORATION AND MULTIHUB LOCATION PROBLEM CONSIDERING ENVIRONMENTAL IMPACTS.

by

EDUARDO JOSE CASTILLO FATULE, B.S.

THESIS

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

Department of Industrial, Manufacturing, and Systems Engineering

THE UNIVERSITY OF TEXAS AT EL PASO

August 2019

Acknowledgements

This work has been partially supported by the National Institute of Food and Agriculture, under the Hispanic Serving Institutions Program, under award numbers 2015-38422-24112 and S17-502008. Any opinions, findings, conclusions, or recommendations expressed are those of the author(s) and do not necessarily reflect the view of the U.S. Department of Agriculture.

Abstract

The Centralized Carrier Collaboration and Multi-hub Location Problem (CCCMLP) represents a strategy that small-to-medium sized less-than-truckload (LTL) carrier companies can use in order to improve their profit margins. It is a strategy that is being explored in order to make these companies more sustainable as they are forced to reinvent their processes and supply chains. In this work, I will present a metaheuristic approach to optimizing their hub establishment and routing policies in order to better their expected profit margins and reduce their environmental impacts. The study considers the costs of transportation, loading and unloading, maintenance, operations, and inventory holding as part of the costs incurred by the carriers. Additionally, there is a cost incurred by the establishment of a hub for shipment consolidation. Similarly, there are environmental impacts to be considered. These are to be represented and accounted for in terms of the Global Warming Potential (GWP) they represent. Optimality in the results will be represented by a Pareto front, which will be compared with some single-objective solutions for reference. The objective of the CCCMLP is to seek a set of hybrid collaborative consolidation transshipment hubs with the purpose of establishing a collaborative hybrid hub-and-spoke network system to minimize the aforementioned parameters. The problem is modeled using Universal Generating Functions (UGFs) for the stochastic variables and solved using a Multi-Objective Evolutionary Algorithm (MOEA).

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1. Background Information

According to Hernandez et al (2011), there has been a rapid increase in the demand for the less-than-truckload (LTL) industry. This can have widespread negative impacts due to activities such as deadheading, which are required for operations, but which cause the companies to incur losses, as the increased costs carry over into other industries and fields, such as prices for food and other commodities. However, it can be possible to take advantage of information communication technologies (ICT) to create new paradigms for the LTL industry. One such innovation would be the concept of LTL carrier collaboration, which provides opportunities to exploit synergies in operations such as excess capacity, reduce operational costs, increase asset utilization, and improve service levels (Esper and Williams, 2003)

As LTL carrier collaboration becomes a new model for the reduction of supply chain costs and increase of resource utilization, carriers have the opportunity to improve their market position. However, the collaboration relies on seasonal land locational factors, which must be taken into consideration as these companies tend to operate over a point-to-point network structure of warehouses and distribution centers (Hernandez et al, 2011). This kind of network moves shipments from an origin to a destination without middle points for cargo consolidation. An objective of our investigation has been to validate previous research which seeks to identify and locate possible transfer and consolidation hubs to promote carrier collaboration. Additionally, we have studied the environmental impacts associated with these operations as part of a multi-objective study. By introducing transfer points (centralized collaboration transshipment hubs), the transportation costs and environmental impacts are reduced, as we transform the point-to-point networks into hybrid hub-and-spoke networks. Hybrid hub-and-spoke networks, unlike pure hub-and-spoke, allows for direct routes.

LTL carrier-carrier collaboration is not a very widely explored concept within the freight domain. Collaboration within the truckload carrier, liner shipping, and rail industries has been studied by Agarwal and Ergun in 2008, Song and Regan in 2004, Figliozzi in 2006 and Kuo et al in 2008. More recently, Hernandez and Peeta (2010) introduced and examined the viability of LTL carrier collaboration for a single carrier, exploring potential benefits of this type of paradigm shift. The study was furthered by Hernandez et al (2011), changing from a static to a dynamic planning perspective. In both of these, the authors' experiments suggest that collaboration can decrease deadheading for the carriers involved, alleviate costs compared to a non-collaborative alternative. While Hernandez had used exact methods to calculate the savings that could be obtained from varying degrees of collaboration, we used metaheuristic approaches for hub allocation, allowing for shorter computational times.

Additionally, Kamakate and Schipper (2009) saw growing trends in truck freight industry energy use and carbon emissions between 1973 and 2005 for the United States and other countries such as the United Kingdom, Japan, and Australia. While trucks and other transportation methods have become more fuel-efficient, the increased use of these still brings an increase in Carbon Dioxide (CO2) and other gas emissions. As shown in Figure 1.1, Transportation accounted for 28% of total Greenhouse Gas (GHG) emissions in 2015, up from 27% in 2013, and a 17% increase from 1990. Additionally, 23% of the GHGs from the transportation sector can be attributed to Medium-and-Heavy-Duty Trucks, shown in Figure 1.2 (EPA, 2018), (EPA, 2018). Previous studies have been made, showing a relation between the economic value added and the environmental impacts of a product in different stages of the supply chain. In the transportation stage, where our research is focused, the environmental impact is much greater than the added value (Cliff & Wright, 2000). It is therefore important for us to lower the environmental impacts incurred in these stages, as they have little added value for the product compared to high environmental impacts.

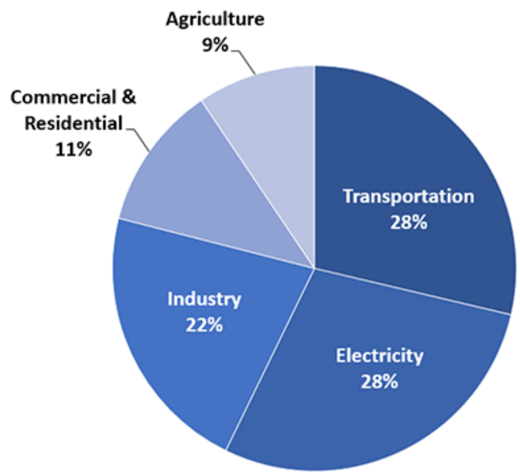

Figure 1.1: Total Greenhouse Gas Emissions by Economic Sector in 2016 (EPA, 2018)

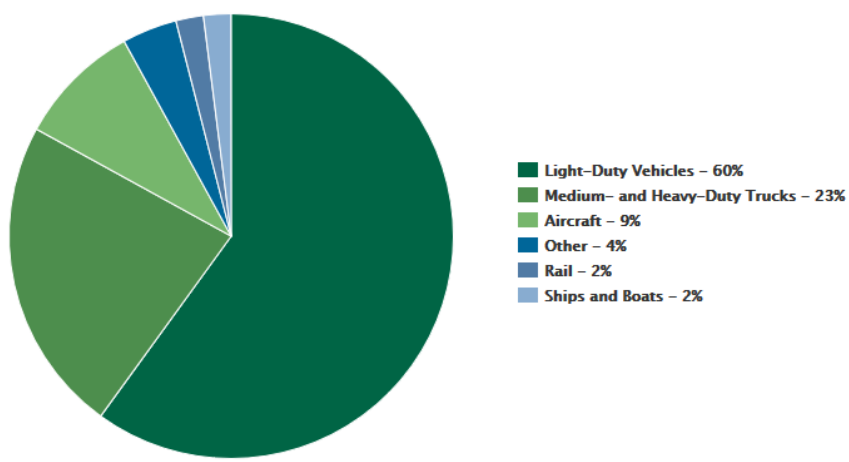

Figure 1.2: 2016 U.S. Transportation Sector GHG Emissions by Source (EPA, 2018)

As part of a greater effort to make industry more sustainable, we believe that the possibility of carrier collaboration can help improve current energy consumption and GHG emission trends. The concept of collaboration is necessary to improve the state of the LTL industry, as truck engines are not expected to noticeably increase their efficiency in the near future due to the heavy weight they carry. As such, we aim to help this industry better their profit margins, reduce their environmental impacts, and plan for uncertainties in their routes. By doing this, we hope to make

the LTL industry more sustainable not only in terms of their environmental impacts, but also in terms of revenue.

1.2. Thesis Objective

Therefore, one of the objectives of this thesis is to address a framework for the Centralized Carrier Collaboration and Multi-hub Location Problem (CCCMLP) among a set of small to medium-sized LTL carriers. Here, a central entity is meant to organize the collaboration between the carriers in order to minimize the total (collaborative) system costs, variability, and environmental impacts subject to each individual carrier's behavior, as well as maximize profits for these same carriers. Thus, it becomes necessary to identify tradeoffs between the objectives in order to find a configuration as close as possible to the ideal. This thesis aims to demonstrate the application of metaheuristics in order to improve the transportation networks used by LTL carriers. This will be done by implementing the approaches suggested by the US Department of Transportation, which are the use of low-carbon fuels, reducing the number of miles traveled, and improving the design of transportation networks to reduce trip frequencies.

Our study will focus more on the last two approaches, which are reducing the total number of miles traveled and improving network design. This will be done by changing the current point-to-point networks into hybrid hub-and-spoke networks. These types of network will be further explained in section 2.1. The implementation of this model will allow to vastly reduce the number of miles traveled, as well as provide a decrease in number of trips taken.

Still, the main objective is to create a multi-objective evolutionary algorithm that can be used to solve the CCCMLP under a non-deterministic scenario while taking into account the environmental impacts of the proposed network configuration.

The use of a metaheuristic method to find a suitable solution has been shown to be efficient in literature. We also believe it to be the best approach since it is capable of solving a problem with more than one objective. This is called multi-objective optimization. There are multiple

approaches to multi-objective optimization, such as the use of utility functions or the NSGA algorithms. We will use a Non-Dominated Pareto-Optimal approach through a Multi-Objective Evolutionary Algorithm (MOEA).

Similar problems to our own have previously been solved through the use of exact mathematical approaches such as Linear Programming, Mixed Integer Programming and Lagrangian Relaxation. While these approaches provide more robust and exact solutions, they have the disadvantage of requiring heavy use of resources as well as an exponential computational time to find a solution. Metaheuristics, on the other hand, may not guarantee finding the most optimal solution; finding instead a number of very good solutions within the search space. However, they have a great advantage in their computational time, requiring mere fractions of what an exact method would require.

1.3. Scope and Limitations

The present work analyzes the scenario of a Centralized Carrier Collaboration and Multi-hub Location Problem (CCCMLP). This scenario is limited to the information and data obtained from other studies. More specifically, most of the data for the CCCMLP was previously collected by Hernandez et al. (2011). The study considered a transportation network of ten nodes, with two carriers that are candidates for collaboration. Tentative data models were created by extrapolation of this data to obtain environmental impacts and probability functions to be used in the multi-objective optimization models.

This thesis aims to minimize transportation costs related to the carriers' operations, as well as the variability of these costs, thus maximizing profit. The other objective to be optimized in the scenario is the minimization of environmental impacts, given as Global Warming Potential. For the optimization process, this work will focus on the use of evolutionary algorithms. The scope of this study is limited to the generation of the Pareto-front of non-dominated solutions. The selection

of an ideal solution is best left to an experienced decision maker. However, methods for pruning the Pareto set will be implemented.

The problem is studied under the assumption that information about the carriers' costs and demands is known beforehand. Costs associated with events that happen in the hub locations and may involve loading and unloading as well as delays caused by various reasons are considered in the holding costs. Another parameter that we will attempt to optimize is the environmental impacts associated with the operation of the carriers involved in the study. These impacts, as well as the costs, are subject to probability functions known as Universal Generating Functions. These will be further explained in section 2.4 of the paper. We plan on attempting to minimize the variability of these costs and impacts as well as their expected value. Lastly, we plan on analyzing the profitability of these operations by the use of constant revenue associated with the carriers' demand and maximizing it.

1.4. Thesis Outline

The remainder of this thesis will be structures as follows:

Chapter 2 will consist of a comprehensive literature review, to include topics such as hub-and-spoke networks, optimization methods, carrier collaboration, and the use of pertinent mathematical models.

Chapter 3 will focus mostly on the proposed mathematical model for the centralized carrier collaboration and multi-hub location problem, including the multi-objective evolutionary algorithm and the relevant parameters and criterions.

Chapter 4 will show a numerical example over which the proposed algorithm is applied, as well as post-pareto optimality analysis.

Finally, chapter 5 will present some concluding statements, after which the works cited will be listed.

Chapter 2: Literature Review

2.1. Hub-and-Spoke Networks Review

Transportation networks used to be designed with direct routes in mind. This is also known as a point-to-point design. These networks are not very practical in design, as they are inefficient and caused an economic loss for the shipping and transportation companies. As information communication technologies progressed, researchers were able to develop more efficient and profitable networks. The Hub-and-Spoke structure was the result, providing cheap, reliable, fast, flexible, and more accessible transportation services to air passengers. After being introduced to the aviation market, the hub-and-spoke concept became a primary model for leading logistics companies. Kim and Soh (2012) attribute consolidating and rerouting at hubs to the structure's achievements in economies of scale.

A Hub-and-Spoke network is a centralized and integrated logistics system designed to reduce costs in the network. The distribution centers within the network, also referred to as the hubs, receive products from different origins, consolidate them, and send them on to either a different hub or their final destination. More specifically, the hub-and-spoke model is "a distribution network consisting of a set of fully interconnected facilities called hubs, optimally situated among other set of facilities and directly connected through arcs known as spokes. At the hub, the transport units are transferred from one service to another connecting the hub with the destination terminal. Ideally, hubs are located near the center of transport demand, in this way distances and trip times between origin and destination terminals can be minimized" (Delgado, 2016). Figure 2.1 provides a visual comparison of a point-to-point model against a hub-and-spoke network:

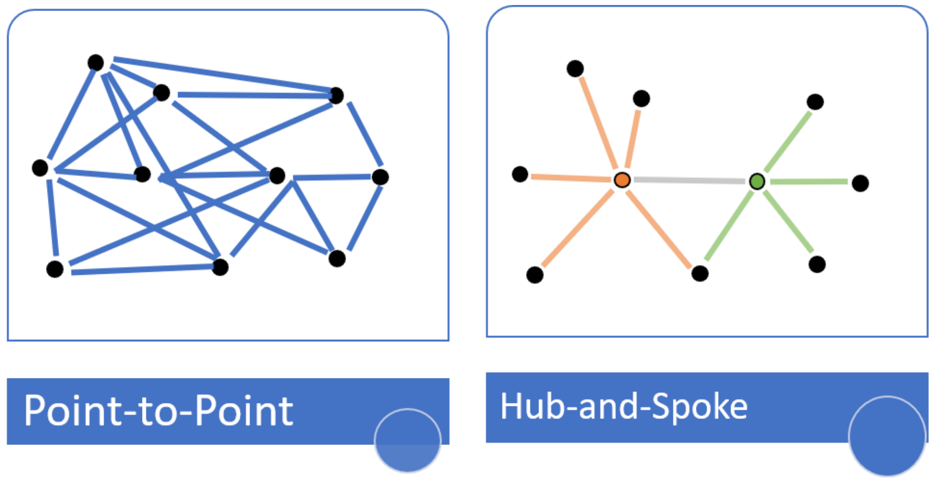

Figure 2.1: Point-to Point vs. Hub-and-Spoke Networks

The hub and spoke structure has been widely applied on various transportation and communication systems after being pioneered within the air travel industry, where the hub and spoke structure was proved to minimize costs. Up to 1979, airlines operated mostly within a point-ezto-point network, where only direct flights were offered from one city to the next. This benefited the passengers, as the total travel times were less, and they did not have to transfer planes. However, the cost for the airlines was higher and the number of routes was limited by demand. The introduction oh hub-and-spoke network structures allowed for more destinations to travel as well as lower costs to the airlines. The network structure has been adopted by most major US passenger airlines. Airlines now use hub and spoke networks to route their flights, operating through centralized hub airports, such as the Dallas-Fort Worth International Airport for American Airlines or the Atlanta-Hartford for Delta. A big advantage of this network structure is that smaller airports can connect to a hub, allowing for routes that would not have enough traffic to be included into a point-to-point network.

Additionally, hub and spoke networks have been implemented into shipping and cargo industries with the objective of speeding up deliveries and reducing costs. For this industry, the hubs are known as central processing facilities, and serve as middle points for cargo on the way from origin to destination. These facilities also serve as a storage location, improving speed for

future deliveries. Companies that have implemented this model include FedEx, UPS, Norfolk Southern, and Yellow Freight., resulting in a competitive logistics advantage and reducing transportation costs and cycle times.

Within the communications industry, the hub-and-spoke model is more commonly known as the star topology. In this structure, the access point is connected to the internet through a wire, and each device is connected wirelessly to the access point, thus the access point serves as a hub between the user's device and the information online. While technologically this may not be the most efficient method, it provides increased security as each device is considered isolated from the others in the network. Other examples in the communications industry are cellular networks, where each phone is connected to a tower, which connects to the call recipient through another tower.

As exemplified above, hub and spoke networks are constantly being implemented for network systems improvement despite the concept being well over 30 years old. Therefore, researchers continue to study and improve specific scenarios by implementing hub and spoke networks. Most of the research to be presented focuses on the improvement of transportation networks with cost minimization objectives. However, the reduction of environmental impacts is scarce in literature. This study intends to address the environmental impact area of the hub and spoke network structure as an objective to be minimized.

2.1.1. Hub Location Problems

Much of the modern hub location research began after Morton O'Kelly's 1987 study on interacting hub facilities. O'Kelly formulated and solved a general hub location problem with a quadratic integer program. He also develops multiple heuristics for this type of problem. Campbell and O'Kelly (2012) define hub location problems as those involving the "location of hub facilities through which flows (e.g., of passengers of freight) are to be routed from origins to destinations." The functions of hubs include, but are not limited to, switching, sorting, or connecting (which

allows the redirection of flows), and consolidation/breakbulk (which allows for the aggregation and disaggregation of the flows).

Hub Location problems have five key distinguishing features:

1. Demand is associated with flows between origin-destination pairs rather than individual points.
2. Flows are allowed to go through hub facilities.
3. Hubs are facilities to be located.
4. There is a benefit of routing flows via hubs.
5. There is an objective that depends on the location of hub facilities and the routing of the flows (Campbell & O'Kelly, 2012).

In pure hub and spoke problems, however, there are two additional requirements defined by Campbell in 1994:

6. Paths between origin-destination pairs visit at most two hubs.
7. Direct origin-destination flows are not allowed.

In recent years, hub location research has strived to solve problems with larger networks and more efficient mathematical formulations and solution algorithms for either the fundamental hub location problem or an extension of it, particularly if these include fully interconnected hubs and flow-independent discounts. Most recently, Contreras, Cordeau, and Laporte (2011) devised an exact algorithm capable of optimally solving a hub location problem on a network of up to 500 nodes using Benders Decomposition, a partitioning method for mixed-integer programs. Their method shows improvements against previous solution methods in both the time it takes to compute and the computational memory requirements. Clustering approaches and simulation-based software have also been implemented for the evaluation of hub-and-spoke networks.

2.1.2. Capacitated and Uncapacitated, Single and Multiple Allocation Hub Location Problems

Hub-and-spoke problems can also be classified as single allocation, where flows go through a single hub en route to their destination (Ernst & Krishnamoorthy, 1999), or multiple allocation, where an origin-destination flow is split among multiple hubs (Ebery et al., 2000). As new technology is developed, recent studies have considered fixed and variable transportation costs on all arcs (Campbell et al., 2015), approaching a hybrid hub-and-spoke network. The concept of a hybrid hub-and-spoke network has been previously introduced by Kuby and Gray (1993) and Aykin (1995). This concept for hub facility location captured the flexibility of flows being sent directly to their destination without passing through hubs.

Other extensions to the hub location problem consider the capacity of the hubs in the decision-making process. Some studies, such as Aykin et al. (1994) and Ernst and Krishnamoorthy (1999), considered capacities un the hubs, limiting the flows that could go through them. On the other hand, researchers like Pirkul and Schilling (1998), Klincewicz (1996) and Topcuoglu et al. (2005) studied the case of uncapacitated hubs. The uncapacitated hub location problem uses fixed costs for the establishment of a hub, but the number of hubs to be established mostly remains as an unconstrained decision variable. This type of problem has been studied as a multi-objective optimization problem, and solved using a multi-objective imperialist competitive algorithm (Mohammadi et al., 2011).

Another variation of the Uncapacitated Hub Location Problem (UHLP) is the Uncapacitates Single Allocation Hub Location Problem (USAHLP). In this version, the number of hubs remains as a decision variable, but the single allocation concept is inserted, forcing all traffic to flow through a single hub for each origin-destination pair. Topcuoglu et al. (2005) solved this problem using genetic algorithms. In comparison to previous literature, their method surpassed previous ones in terms of computational time and CPU requirements, while matching the optimal solution. Similarly, there is an Uncapacitated Multiple Allocation Hub Location Problem, or UMAHLP. For this problem, a 4-index formulation was proposed and solved by an accelerated

Primal (Benders) decomposition and a greedy heuristic by Gelareh and Nickel (2011). This serves to show that there are many different versions of the hub allocation problem. Our problem, the Centralized Carrier Collaboration and Multi-Hub Location Problem, or CCCMLP, falls within the territory of an UMAHLP, although it also incorporates the existence of two separate entities into the transportation network.

2.1.3. P-hub Median Problems

Early instances of the hub allocation problem assumed that the number of hubs must be pre-established rather than left as a decision variable. These were referred to as the p-hub median problem (Kim & O'Kelly, 2009). In such cases, p was used to represent the number of hub facilities that were being located to minimize total flow costs. Some of the solution methodologies for this problem include mixed integer linear programming considering demand uncertainty and congestion effects (de Miranda Jr. et al., 2011). After the initial formulations, researchers have further developed the problem into new directions, adding new features and constraints consistent with various hub-and-spoke applications. After O'Kelly's introduction of the single allocation concept, several studies focused on optimizing this subproblem using tight linear programming formulation and lagrangian relaxation (Pirkul & Schiling, 1998). Other researchers have also used meta-heuristic models for optimizing single and multiple allocation problems (Gomes et al., 2013). Non-linear mixed integer programming functions were also used in conjunction with a generalized Benders decomposition algorithm by de Camargo et al. (2009).

Similarly to the previously mentioned expansions of the single and multiple allocation versions of the hub location problem, there have been studies into the capacitated and uncapacitated directions in which the p-hub constraint has been incorporated. Klincewicz (1996) developed an algorithm based on dual ascent and dual adjustment techniques within a branch and bound scheme to solve a capacitated p-hub location problem. Abdinnour-Helm (1998) developed a new heuristic method which combined aspects of genetic algorithms and tabu search, with

improved results compared to existing methods at the time. For the capacitated single allocation p-hub median problem, stochastic demand and time-based service level hub constraints are incorporated into non-linear programming models (Li & Zhao, 2009; Vidyarthi et al., 2013). Other methods such as heuristic algorithms and linear programming can also be used to solve capacitated and uncapacitated p-hub location problems.

2.1.4. Developed Hub-and-Spoke Models

Regardless of the aforementioned variants of the hub location problem, other authors have proposed new problems related to the hub location problem, such as the minimax hub location problem, which aims to locate a facility that would minimize the maximum weighted interaction cost between origin-destination pairs (O'Kelly, 2009), or the hub arc location model, which attempts to instead locate hub arches with reduced unit flows (Campbell, Ernst, & Krishnamoorthy, 2005). Additionally, other studies consider dynamic (multi-period) hub location problems (Contreras, Cordeau, & Laporte, 2011), networks in continuous Euclidean space (Carlsson & Jia, 2013), generalized models to integrate the operations of pure, stopover, and center-direct hub-and-spoke networks (Lin & Chen, 2008), as well as hierarchical approaches to the network structure (Lin, 2010).

Additional contributions to the hub and spoke model developments include Ishfaq and Sox (2012), who integrated hub operation queuing models into the hub allocation problem. Meng and Wang (2011) developed a mathematical program with equilibrium constraints for the intermodal hub and spoke network design problem with multiple stakeholders and container types. Mohammadi et al (2016) formulated a game-based meta-heuristic for optimizing the configuration of a fuzzy bi-objective reliable hub location problem, in which the authors study the effects of delivery service requirements in the hub-and-spoke problem using a fuzzy queuing approach. Kim and O’Kelly (2009) have also implemented reliable p-hub problems into the Telecommunications field.

2.1.5. Solution Methodologies

As is common with optimization problems, different studies have used different solution methods to reach a solution. A popular alternative for solving hub location problems has been Lagrangian Relaxation, which attempts to solve complex optimization problems by approximating them with simpler ones. Aykin et al. (1994), Pirkul & Schilling (1998), and Elheldi & Wu (2010) have used this type of approach in concerns to hub-and-spoke systems. Additionally, Hernandez et al. (2012) used lagrangian relaxation to optimize a CCCMLP for the less-than-truckload industry over a hybrid hub-and-spoke network.

Klincewicz (1996) attempted the use of heuristic methods for solving the uncapacitated version of the hub location problem. Klincewicz had also attempted to solve the p-hub location problem in 1992 by using tabu search and grasp. Aykin (1994) and Aykin et al. (1995) tried using exact methods such as branch and bound.

Many similarities exist between hub location problems and facility location problems. For both of these, locating the hub facility and designing the network around it is a vital aspect of the problem. The hub-and-spoke network structure has had a great impact on transportation systems, improving service time and quality while reducing operating costs after considering all the different variables. Even with the large number of variants to the problem, the objective function is still the same: Identify the optimal location of the hubs so that the total costs are minimized.

2.1.6. Sustainability and Environmental Impacts in Hub-and-Spoke Networks

Transportation is a key component of human life. It has served as one of the main factors for economic growth, allowing for human mobility and trade. As transportation systems evolve into more efficient structures, however, it becomes more apparent that these changes have great impacts on the economy and on the environment. With the widespread of hub-and-spoke networks, it is possible to better meet sustainability demands according to an empirical study by Liu et al. (2012). Still, it is important to re-evaluate the current concepts of what sustainability and life cycles

mean for the industry, and how these should be incorporated into the economic factors of transportation businesses.

Unfortunately, there is little literature on the environmental impacts of hub-and-spoke networks. O'Kelly (2012) has been one of few who has studied these effects, using city systems as an experimental context for understanding environmental costs and benefits of Hub-and-Spoke network flows, using fuel burn as a measure for these costs. Hub-and-Spoke networks have also been shown to be effective in green supply chains, such as the aforementioned study by Liu et al. (2012), which also proposed a model for integrating marketing and sustainable supply chain management. Other literature focuses on hub-and-spoke network design for biomass supply chains, considering both costs and emissions (Roni, 2013; Roni et al., 2014, 2017). Bio-mass to biorefinery logistics has also been studied by Delgado et al. (2015), performing life cycle assessments to study the environmental impacts of the biofuel production process. These studies show that there is, generally speaking, an improvement in sustainability when using a hub-and-spoke network. Still, there is still much work to do. Hub-and-Spoke networks are still a relatively new concept, regardless of its now widespread use, and though much of the literature shows favorable outcomes in terms of monetary savings, there is still not much research done on its environmental impacts. More research needs to be done to realistically model the variables and constraints that exist within these networks and prove their benefits from an environmental point of view.

2.2. OPTIMIZATION METHODS REVIEW

For decades, many problems were solved by the optimization of a single objective. This objective was usually minimizing cost, maximizing profit, or maximizing reliability. Still, optimizing a single objective may not be a very effective way of improving a system, as other parameters can be negatively affected by this. For most problems, taking into consideration only one objective can be useful in reducing the complexity of the problem to solve, but doing this

renders the solution incapable of fully describing all the nuances of the problem. Almost every real-world optimization problem will involve multiple objectives, often conflicting with each other. This reasoning resulted in the understanding that single-objective optimization is no longer suitable for solving problems, and the framework for multiple-objective optimization was laid down, as it allows for more information to be taken into account (Caramia & Dell'Olmo, 2008).

A basic single-objective optimization problem can be formulated as follows:

$$\begin{gathered} \min\{f(x)\} \\ subject\ to\ x \in S, \end{gathered}$$

Where f is a scalar function and S is the set of constraints.

While a single-objective optimization problem is expressed by a single objective function, a multiple-objective problem is expressed by a number of functions, as well as several equality and inequality constraints. The notation for such a problem can be written as follows:

$$\begin{gathered} min/max\ \ f_i(x), \qquad for\ i = 1,2,\dots,n \\ subject\ to: \\ g(x) \le 0 \\ h(x) = 0 \end{gathered}$$

Where f_i (x) stands for each of the objective functions (i=1…n) that the problem attempts to optimize, and the set of constraints g(x) and h(x) define the feasible region. The decision variables are represented by x.

Several methods exist to solve single and multiple-objective optimization problems. However, we will focus on the solution methodologies for multi-objective problems, which can yield a single solution or a set of non-dominated Pareto-optimal solutions.

2.2.1. Single Solution Approaches

The first set of approaches to solving multi-objective approaches involves combining all objective functions into a single aggregated objective function, thus essentially transforming the problem back into a single-objective problem.

Goal Programming

The Goal Programming approach was first developed by Charnes *et al.* in 1955. The aim was to find specific goal values for each of the objective functions considered in the problem. In their method, each objective function *F(x)* is assigned a goal b_j. The aggregated objective function becomes reducing the total deviation from the goals, d_j. The deviations are separated into positive and negative values, to consider under and overachievement, where achievement or zero values imply reached goals. The optimization problem is then formulated as follows:

$$\min_{x \in Z, d^-, d^+} \sum_{i=1}^{k} (d_i^+, d_i^-)$$

$$Subject\ to\ F_j(x) + d_j^+ - d_j^- = b_j$$

$$d_j^+, d_j^- \geq 0,$$

$$j = 1,2, \dots, k$$

This approach to multi-objective optimization, however, has the drawback of not guaranteeing a Pareto-optimal solution. It can also become increasingly hard to solve given larger problems. Some variations of this method include Weighted Goal Programming, Preemptive Goal Programming, Multi-Goal Programming, and Goal Attainment Method.

Weighted Sum or Scalarization

The Weighted Sum method is one of the most common approaches to multi-objective optimization. This strategy converts a multi-objective problem into a single-objective one by constructing a weighted sum *F(x)* of all objectives in the vector of criteria functions. More specifically, the Weighted Sum method minimizes a positively weighted convex sum of all objectives, and can be represented as follows:

$$F(x) = \sum_{i=1}^{n} w_i f_i(x)$$

Where n represents the total number of objectives I and w_i their respective weights. An advantage to this method is its simplicity, transforming a multi-objective problem into a single-objective one by simple addition, while also allowing the decision maker to assign priority to the objectives through the weights. Its disadvantage, however, lies precisely in assigning the weight coefficients, which can be difficult to choose. Determining the adequate weights for this type of problem has been the subject of much research, with some researchers arguing that weight functions be more appropriate for better representing preferences. Some approaches to weight assignment include ranking, categorization, rating, and eigenvalues.

Lexicographic Method

Yet another way to address multiple objectives is through the lexicographic approach proposed by Fishburn (1974). This method requires the decision maker to establish a priority for each objective. In this method, the objective functions are arranged in order of importance, and solutions are compared in respect to the most important objective. In the event of a tie, the next most important functions are compared, and so on until there are no ties or objectives remaining. This method can be describes with the following formulation:

$$Minimize\ x \in X\ F_i(x)$$

$$Subject\ to\ F_j(x) \leq F_j x * j, \qquad j = 1, 2, \ldots, i-1; i = 2, 3, \ldots, k$$

Where I represents a function's position in the preferred sequence, and $F_j x^*j$ represents the optimum of the jth objective function, found in the jth iteration. After the first iteration (where j=1), $F_j x^*j$ is not necessarily the same as the independent minimum of $F_j(x)$, since new constraints have been introduced from the results of the previous iteration.

Multi-Attribute Utility Theory

In this case, utility refers to the satisfaction that each attribute or objective function provides to the decision maker. This way, the utility theory approach assumes that any decision is

made on the basis of the utility maximization principle. This principle suggests that the best choice is the one that would provide the most satisfaction to the decision maker. In multi-attribute utility analysis, the total utility of a design solution is a scalar on the interval between 0 and 1, where 0 represents no utility and 1 represents the highest possible utility. According to utility theory, if X_i is the measure of satisfaction provided by an attribute *I*, and there are *n* attributes, then the joint utility function for all attributes can be expressed as:

$$U(X_1, X_2, \dots, X_n) = f\big(U_1(X_1), U_2(X_2), \dots, U_n(X_n)\big)$$

In this case, $U_i(X_i)$ is the utility of the i^{th} attribute. The overall utility function is the sum of all individual utilities if the attributes are independent. This function is given as follows:

$$U(X_1, X_2, \dots, X_n) = \sum_{i=1}^{n} U_i(X_i)$$

After assigning weights to each attribute, the function can be expressed as:

$$U(X_1, X_2, \dots, X_n) = \sum_{i=1}^{n} W_i U_i(X_i)$$

While utility optimization is effective and widely used, it requires extensive interviews with the decision maker to determine appropriate utility functions and weights. Once the utility function has been constructed, however, optimization can occur and the design alternative with the maximum utility can be determined.

ε-Constraint Approach

This procedure originally proposed by Chankong and Haimes (1983) overcomes some of the convexity problems that the Weighted Sum technique faces. In this approach, the decision maker chooses one objective to be optimized, and constraints the remaining objective to be within a target range (equal or less than a target in a maximization objective, or equal or larger than a

target in a maximization objective). For example, in a bi-objective minimization problem, the problem is divided into two problems, $P_1(\varepsilon_2)$ and $P_2(\varepsilon_1)$, which are the following:

$$\min f_1\left(\underset{x}{\rightarrow}\right)$$
$$Subject\ to\colon \underset{x}{\rightarrow} \in X$$
$$f_2\left(\underset{x}{\rightarrow}\right) \leq \epsilon_2$$

And

$$\min f_2(\underset{x}{\rightarrow})$$
$$Subject\ to\colon \underset{x}{\rightarrow} \in X$$
$$f_1\left(\underset{x}{\rightarrow}\right) \leq \epsilon_1$$

Respectively.

Contrary to other aggregation methods, the ε-constraints approach is able to identify a number of non-inferior solutions within a nonconvex boundary. However, it has the drawback of its hard constraints not being adequate for representing real design objectives.

2.2.2. Pareto-based Optimization Approaches

In Pareto-optimality based approaches, there is no single optimal solution, but instead a set of non-dominated alternative solutions. These solutions are considered “Pareto optimal”, since none of the other solutions are dominated by other solutions. The concept of dominance in optimization means that there is no other solution in the feasible region which is better than the pareto-optimal solutions in all the objectives considered. This makes it possible for there to be tradeoffs between objectives from the decision-maker’s point of view (Zitzler & Thiele, 1998; Zitzler *et al.*, 2002).

These optimal solutions are called Pareto-optimal solutions and the set of these is denoted as the Pareto-optimal set. Figure 2.2 provides an illustration of the terminologies used in Pareto based optimization:

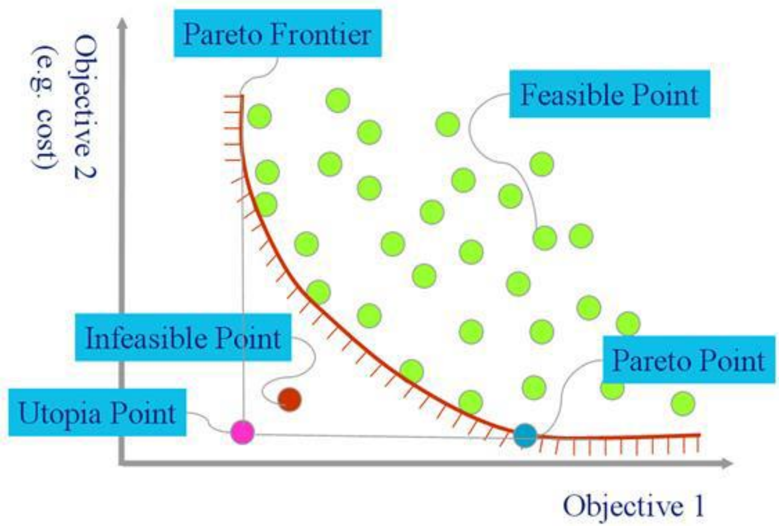

Figure 2.2: Pareto-based optimization Illustration and Terms

As an example of conflicting objectives, such as when maximizing an objective involves increasing another that we want to minimize, we could consider a situation in which cost and efficiency are to be optimized. Cost is an objective that is almost always being minimized, while efficiency is almost always maximized. However, it is also true that increasing efficiency involves increasing costs, creating a conflict in the objectives. The final decision will rely on which of the two objectives is more important to the designer.

Formulating, analyzing, and solving problems with conflicting objectives usually requires a decision maker to express preference relations between alternative solutions. The decision maker must have some expertise in the field for which the problem is solved, as well as knowledge on the resulting set of solutions, in order to make an educated choice as to which solution would be best, as well as to give weights to the objectives being evaluated. The Pareto-optimal set can help reduce the design alternatives from a feasible region into optimal trade-offs (Yancang, Lina, & Shujing, 2010).

Multiple-Objective Evolutionary Algorithms

Optimizing multi-objective problems can be a challenging task, since one of the characteristics of these problems is that the objectives tend to conflict with each other, and the multi-dimensional search space tends to be very complex. As a solution, researchers have proposed several different models to obtain Pareto-Optimal solutions. Much like in single-objective optimization, evolutionary algorithms are some of the more popular models for solving these types of problems due to their capacity to be adapted to different types of problems.

Genetic Algorithms (GAs) optimize a desired objective by altering its encoded variables. Comparing it to biological evolution, the solutions arise from a set of possible "genetic" sequences. Hence, the best solutions result from organisms that were able to survive and reproduce within the environment, which is to say, the solutions with the best objective function values. Genetic Algorithms are a variety of Evolutionary Algorithms (EAs), which apply techniques inspired by evolutionary biology such as inheritance, mutation, and crossover or combination. A set of random solutions represented by a data structure is generated. In technical GA terms, these solutions are considered the chromosomes that constitute the individuals in a population. The chromosomes consist of a sequence of genes, or specific data characteristics which will be used during the evaluation of the objective function or fitness value for the individual, and can be in the form of bits, digits, or letters (Kumar *et al.*, 2010).

A genetic Algorithm simulates the best individuals in successive generations, where a set of individuals composes each generation's population. Each generation's population is evaluated to continually identify the best solutions. The general methodology followed in the formulation of a genetic algorithm is showed below in figure 2.3:

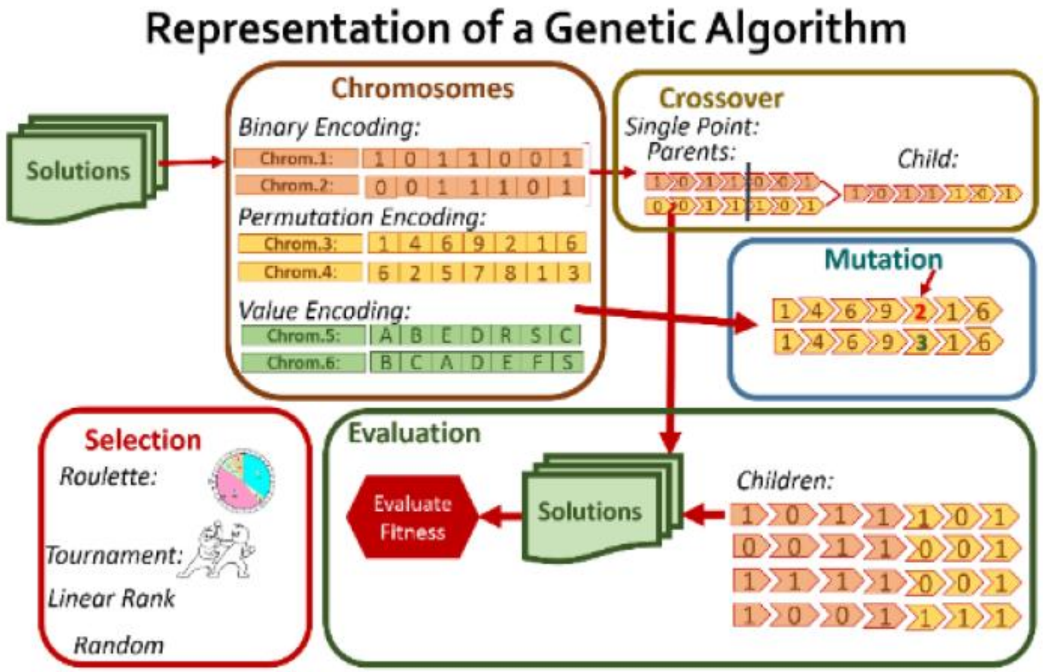

Figure 2.3: Graphic representation of a Genetic Algorithm (Delgado, 2016)

This Process will be explained in more detail in section 3.2.

In multiple objective optimization, Evolutionary Algorithms follow similar processes as Genetic Algorithms. Still, all evolutionary algorithms have different techniques when attempting to achieve diversity in their population set. In multiple-objective evolutionary optimization, it is crucial to achieve diversity in their population, as it means that it's possible to deliver a varied set of alternatives to the decision maker. For this, different fitness assignment methodologies have been explored. Pareto-based approaches can be divided into elitist and non-elitist approaches.

Pareto-based non-elitist Approaches

1. *Multiple Objective Genetic Algorithm (MOGA)* in an approach in which individuals are assigned ranks corresponding to the number of individuals in the current population by

which they are dominated. The non-dominated individuals are ranked and ties are averaged to maintain the sampling rate (Murata & Ishibuchi, 1995).

2. *Niched Pareto Genetic Algorithm (NPGA)* proposed by Horn *et al.* in 1994, has a Pareto dominance-based tournament selection with a sample of the population to determine the winner between two candidate solutions. A subset of individuals is used to determine dominance between the two solutions competing, and the nondominated individual is selected for reproduction. In the event of a tie, the winner is decided through fitness sharing.
3. *Non-Dominated Sorting Genetic Algorithm (NSGA)* in an approach in which all non-dominated individuals are classified into one category, and collectively assigned a fitness value proportional to population size. This group is removed, and the remaining population is repeatedly classified until all the population has been classified (Deb *et al.*, 2002)

Pareto-based Elitist Approaches

1. *Strength Pareto Evolutionary Algorithm (SPEA),* developed by Zitzler and Thiele in 1999, has many similarities with other algorithms in terms of storing previously-obtained Pareto-optimal solutions, as well as the use of the dominance concept and the use of clustering to reduce the number of stored solutions. However, what makes it stand apart from the rest is the use of all three of these concepts in a single algorithm. It determines a fitness function out of the stored solutions, avoiding dominance from the existing population, and uses all the stored Pareto solutions for selection.
2. *Strength Pareto Evolutionary Algorithm 2 (SPEA2),* formulated as an improvement of the original SPEA by Zitzler *et al.* in 2001. This iteration of the algorithm intends to avoid the situations in which individuals in the population can have the same fitness value according to their dominance by stored solutions. This time, the fitness function

is calculated using both the stored solutions and the current population, adopting a new scheme to prevent the loss of boundary solutions during archive updating. Diversity is maintained in this approach with a density-based cluster on the k^{th} nearest neighbor.

3. *Pareto Archived Evolutionary Strategy (PAES)* is a single-parent single-child EA, similar to a (1+1) evolutionary strategy. The PAES method has two main objectives: The first is that the algorithm should be strictly confined to local search, moving from one solution to a nearby neighbor. The second is for the algorithm to be a true Pareto optimizer by treating all nondominated solutions as equal in terms of their value to the decision maker. This can be troublesome when comparing a pair of solutions that do not dominate each other. This is addressed by keeping a record of previous non-dominated solutions, which can be used to estimate a dominance ranking for this pair of solutions. The authors concluded that in a multi-objective routing problem, the PAES method provided competitive results compared to a traditional MOEA (Knowles & Corne, 1999).
4. *Pareto Envelope-based Selection Algorithm (PESA)* is an algorithm that uses a smaller internal (or primary) population and a larger external (secondary) population than that adopted by PAES. PESA uses the same hypergrid division of objective space from PAES to maintain diversity, but its selection mechanism is based on a hypergrid crowding measure. This crowding measure is used to decide which solutions are to be introduced into the external population (Corne *e*t al., 2000).
5. *Pareto Envelope-based Selection Algorithm II (PESA-II),* developed by Corne *et al.* in 2011. This revised version of PESA proposes the use of region-base selection, in which the unit to be selected is a hyperbox rather than an individual. The procedure in this method is to select a hyperbox and then randomly choose an individual within that box.
6. *Elitist Non-dominated Sorting Genetic Algorithm (NSGA-II)* is the revised methosdof the NSGA, proposed by Kalyanmoy *et al.* in 2002. This method was devised to eliminate some of the weaknesses of the original NSGA, particularly its non-elitist

nature and the specification of the sharing parameter. In this version, the individuals still undergo non-dominated sorting as in the first version, with the individuals receiving ranks based on this sorting. However, a new selection technique called crowded tournament selection is proposed. This selection method chooses individuals based on crowing distance (representing neighborhood density). Elitism is implemented by not allowing dominated solutions to pass on to the next generations.

2.3. Carrier Collaboration and Multi-Hub Location Problem Review

2.3.1. Carrier Collaboration

While our main focus of less-than-truckload (LTL) carrier collaboration is still a relatively unexplored concept, extensive literature exists within the domain of ground freight collaboration. In the past, Hernandez and Peeta (2010) studied the possibility of an LTL static single-carrier collaboration, with their findings indicating that a higher degree of collaboration results in increased benefits for the involved carriers as this results in a reduction of dead-heading, as well as a possibility of shared trucks. Voruganti *et al.* (2011) modeled collaboration between truckload carriers, with three different collaboration cases: The first, in which complete collaboration occurs and all carriers work together to maximize total profit; the second, in which only partial collaboration occurs, with each carrier maximizing their individual profits and leasing out their residual capacity to others; and lastly, the case in which no collaboration occurs and the carriers perform in a completely independent manner from each other. Their results indicate that collaboration tends to increase profits, but the ideal type of collaboration varies from case to case.

In other transportation methods, Agarwal and Ergun (2008) studied the effects of collaboration within the liner shipping industry, where cargo scheduling can be difficult due to the cargo-routing process involved. They used a greedy heuristic in order to separate the problem and increase computational efficiency. Their results indicate a high percentage of ship capacity utilization and a significant number of transshipments as part of an optimized shipping solution.

Kuo *et al.* (2008) studied cooperation within the rail-based freight transportation industry. Their results indicated that space leasing and slot cooperation strategies can increase shipments up to 40% in their numerical results.

As can be inferred from these studies, the potential for collaboration within freight industries is high, often leading to an increase in profits and resource utilization. The LTL industry has an advantage of there being more empty space within each shipment when compared to, say, the truckload industry, offering greater potential for resource sharing. This means that the LTL industry has a much bigger potential for reduction of costs and, in turn, an increase in profits, compared to other freight industries.

2.3.2. Road Transportation Collaboration Addressing Environmental Impacts

There are two main directions in which cooperation can happen in transportation flows: horizontally, when carriers collaborate with other carriers or shippers collaborate with other shippers, or vertically, when carriers and shippers collaborate with each other. Ballot and Fontane in 2010 proposed a logistical network concept which achieved a total savings of 25% of CO_2 emissions compared to the traditional setup. They also demonstrated that vertical supply chain collaborations can still be improved by the use of horizontal collaboration. Xu (2013) identified and defined two organizational forms of horizontal collaboration: Centralized and Decentralized. These approaches refer to the decision-making process, where centralized refers to the existence of a central entity which coordinates the collaboration, and a decentralized strategy involves separate entities managing the collaborations within subregions.

Perez-Bernabeu *et al.* (2014) contributed to the subject of environmental impact assessment in these collaborations by analyzing horizontal collaboration for small and medium-sized companies based on a well-known multi-depot vehicle routing problem, assuming an ideal disagreement-free scenario for collaboration. The data collection for emissions consisted of a distance-based method with a fuel conversion factor. Their experimental results provided a

noticeable reduction in both expected costs and GHG emissions using a horizontal cooperation strategy. Also, Espiritu *et al.* (2015) studied the environmental impacts of a small network considering two LTL companies, with results indicating a possibility of up to 67% reduction in Global Warming Potential (GWP) in an ideal collaboration scenario, when compared to a non-collaborative scenario.

Vachon and Klassen (2008) studied the impacts of environment-oriented collaborative activities on manufacturing performance. Their findings indicate broad benefits for companies with collaborative green practices. They specify that these green collaborations can be directed upstream towards suppliers or downstream toward customers. Upstream collaboration practices tend to improve process-based performance, while downstream practices tend to improve product-based performance. The main obstacle for environmental progress in many enterprises was identified to be management commitment, according to Lamming and Hampson (1996). One of the important aspects of green management plans is then to maintain management engagement on the environmental issues by promising economic gains.

From a general environmental impact in supply chains perspective, Srivastava (2007) argues Green Supply Chain Management can reduce environmental impacts without compromising quality, cost, reliability, performance, or efficiency. Still, the change to green supply chain management requires a company-wide paradigm shift in order to achieve its goals. The use of Life Cycle Assessment (LCA) software has been proposed by Hagelaar and van der Vorst (2001), who argue that the use of LCA tools should be used to redesign supply chains in order to meet environmental goals, restructuring the supply chains to be able to use these tools. LCA tools can be very beneficial to supply chains not only in terms of environmental impact reduction, but also for cost reductions, as there is often a correlation between the areas of the supply chain with higher environmental impact production and those with higher incurred costs.

2.3.3. Optimization Models for Carrier Collaboration

Many researchers have used mathematical models to study collaboration in road transportation, as these are commonly used to study real world phenomena that cannot be easily understood using analytic techniques alone. Mathematical models are also useful for understanding the relationships between different parameters and their effects on complex processes.

Pan *et al.* (2013) presented a mixed-integer linear programming model, coded in ILOG's OPL 6.3 software with CPLEX 12.1, which implemented Ballot and Fontane's (2010) concept of supply network pooling. Pan *et al.*'s model explored the effect of pooling the supply chain networks on the reduction of CO_2 emissions from transport with two possible modes: road, referring to Heavy Duty Vehicles, and rail, considering electrically-powered locomotive, considering national distribution networks for two major French retail chains. Their results showed that the merging of the supply chains significantly reduced not only the total CO_2 emissions, but also the number of transport paths thanks to the network pooling. Ozener *et al.* (2009) investigated the potential of collaborative opportunities for the truckload transportation industry, developing models to determine the maximum possible monetary benefit offered by collaboration, modeling the Multi-Carrier Lane Covering Problem as an integer programming model. Lin & NG (2012) transformed a stochastic mixed-integer programming problem of collaboration between carriers into an integer programming model. The results provided evidence that the collaborative method can effectively reduce emissions in a freight network by 3-20%.

Other methods include the use of metaheuristic algorithms. A metaheuristic method is defined as "an iterative generation process which guides a subordinate heuristic by combining intelligently different concepts for exploring and exploiting the search space, in which learning strategies are used to structure the information in order to find efficiently near-optimal solutions (Osman & Laporte, 1996)". These types of algorithmic approaches have been commonly used in optimization methods, including the optimization of transportation systems.

Sadegheih *et al.*, for example, developed a Genetic Algorithm (GA) to study the behavior of global supply chain management and collaborative network design. Their algorithm proved successful for solving transportation problems and their results show alternatives in which the interested parties can lower their total carbon emissions while providing them with competitive advantages. Ozener *et al.* (2009) used heuristic approaches when solving the multi-carrier lane covering problem. Their first heuristic relaxed the precision by specifying a relative optimality criterion, while the second exploited the solution structure by fixing integral flows along cycles before starting a branch-and-bound algorithm.

Other literature includes Jaegler and Burlat's (2012) study, which proposed a new vehicle routing variant to minimize greenhouse gases. One of the most common variants in this routing problem is the vehicles restricted by capacity. In one analysis of the vehicle routing problem with time windows and considering environmental impacts in the search solution, four key focus points were considered for the analysis: energy requirements, mathematical formulation, use of scatter search metaheuristic, and a cooperative game approach option. The analysis results indicated that the environmental impacts are strongly related to the vehicle utilization rate, and their proposed solution enables meeting customer satisfaction while decreasing environmental impacts. Lin & Ng (2012) used a tabu search algorithm with Monte Carlo bounding techniques when studying the possibility of reducing environmental issued within a carbon-constrained context for freight carriers.

Multi-objective approaches to solving routing problems are less common, with some notable examples like Jemai *et al.* (2012), who solved a bi-objective green vehicle routing problem to minimize travel distance and CO_2 emissions. They used the NSGA-II algorithm, which showed good results in emission minimization. This study, however, did not consider collaboration in their analysis. De Mello and Frayret (2014) opted for the use of simulation models, and focused into a complete sustainability assessment rather than only costs or carbon emissions. They used a resource sharing methodology for their freight transportation model, and considered the roas transportation of semi-trailers through a network of hubs in which they wait for their next route

segment. They assess sustainability in terms of logistics performance, working conditions, and environmental impacts. Netlogo software was used to create the simulation model. Results indicated that resource sharing can improve overall performance, including social and environmental metrics.

2.3.4. Collaboration and Hub-and-Spoke Network for Small to Medium-Sized LTL Carriers

Efficient road infrastructure is an important assumption in vehicle routing problems as optimal vehicle routing involved reducing both the number of trucks and their movements within the network. As we delve further into the literature, we find that even within the studies of the transportation industry, Less-than-Truck-Load (LTL) carrier collaboration is still a relatively unexplored concept. Even more so, most of the literature of LTL collaboration focuses mainly on the economic aspect of the problem, ignoring the environmental aspect of the problem. Although not explicitly as collaboration, Zhang *et al.* (2007) introduced the concept of a hybrid hub-and-spoke network for a single LTL carrier, attempting to minimize costs. The term "hybrid," when applied to a hub-and-spoke network, refers to the possibility of incorporating direct routes within the network. Still, this is a fairly new concept for the LTL industry.

More recently, Hernandez and Peeta (2010) and Hernandez, Peeta and Kalafatas (2011) addressed a time-dependent centralized carrier collaboration problem (TD-CCCP), modeled as a binary multi-commodity minimum cost-flow problem, and solved with a branch-and-cut algorithm. Other researchers, such as Cunha and Silva (2007), who focused on configuring a hub-and-spoke network for an LTL company in Brazil, used GAs to determine the optimal network configuration while minimizing costs. In addition, Zhang *et al.* (2007) formulated a hybrid hub-and-spoke model for a single LTL carrier and solved the combinatorial problem using a GA.

Later, Hernandez *et al.* (2012) analyzed a Centralized Carrier Collaboration and Multi-hub Location Problem (CCCMLP) for the small to medium-sized LTL industry. In the CCCMLP, a central entity (e.g., a third-party logistics firm) seeks to find a set of collaborative transshipment

hubs for shipment consolidation in order to establish a hybrid collaborative hub-and-spoke network that minimizes the total costs for all carriers. The CCCMLP was formulated as a variant of the P-hub location problem, which is demonstrably an NP-hard problem, and was solved using Lagrangian relaxation Their results indicate that a larger expected profit margin can be expected when the carriers apply more revenue-oriented behavior, but these effects are less pronounced on larger networks.

2.4. Universal Generating Function Review

In the early days of studying systems reliability, one of the most prevalent methods was binary utility theory. In this method, all components of a system are assumed to have only two states: fully operational and completely failure, denoted by 1 and 0, respectively. The Multi-State Systems approach was introduced later by Barlow and Wu (1978), consisting of a finite number of performance rates, with a much better model for real systems than the binary approach.

Levitin (2005) demonstrated the applicability of the UGF in various reliability analysis problems, while Levitin *et al.* (2000) and Ramirez (2018) applied them in optimization problems.

While the main use for the Universal Generating Function (UGF) is in applications for multi-state systems analysis, it can be used to model any probability function as a discrete one with a polynomial notation. The UGF was first proposed by Ushakov in 1986. It is based on the existence of at least two random variables x_i, each with a corresponding probability function, as shown below:

$$P\{x_i = x(i,j)\} = P(i,j), 1 \leq j \leq k_i$$

Where:

- $x(i,j)$ is the value of the j^{th} state of x_i
- $P(i,j)$ is the probability that x_i is equal to $x(i,j)$
- k_i is the number of states that x_i has

When expressed in the model of a Universal Generating Function, these variables take the following shape:

$$u_i(z) = \sum_{j=1}^{k_i} p(i,j) z^{x(i,j)}$$

Where u_i is the UGF of x_i.

In order to analyze a system in terms of its components, the u functions of each of these components are combined using a composition operator $\otimes\varphi$, where φ denotes the composition function to be applied to the exponents in the UGFs for each component. The composition operators to be used in this problem are further explained in Section 3.2.3.

The composition of two UGFs is illustrated in the equation below:

$$U(z) = u_1(z) \underset{\varphi}{\otimes} u_b(z) = \sum_{i=1}^{k_a} P_{ai} z^{x_{ai}} \underset{\varphi}{\otimes} \sum_{j=1}^{k_b} P_{bj} z^{x_{bj}} = \sum_{i=1}^{k_a} \sum_{j=1}^{k_b} P_{ai} P_{bj} z^{\varphi(x_{ai}, x_{bj})}$$

2.5. Post-Pareto Optimality Review

An aspect of multi-objective optimization is that the selection of the best solution is often left to a decision maker's judgement. There are three approaches to incorporate the decision maker's opinions into the optimization method: Firstly, an *a priori* method, in which the decision-maker's preferences are incorporated into an algorithm before generating solution points. Then, there is the *a posteriori* approach, which first generates the solution points to make up the Pareto-optimal set before a decision is made. Lastly, there is the interactive method, in which the decision maker's preferences are incorporated during the search. The objective of post-pareto analyses is to make the final decision easier for the decision maker, who can often be found struggling to make sense of hundreds and thousands of data points.

As an example of an *a priori* approach, Pinchera *et al.* (2017) introduced a function which they named Quantized Lexicographic Weighted Sum (QLWS), based on the definition of a Global Cost Function, which required the decision maker to define priorities among the targets to

optimized in order to more quickly generate a set of solutions by avoiding the evaluation of the Pareto front. For *a posteriori* methods, Reynoso-Meza *et al.* (2010) used a Differential Evolution algorithm to generate a set of non-dominated, Pareto-optimal solutions, and then introduces spherical pruning, which is less sensitive to the loss of non-dominated solutions, in order to reduce the Pareto set for the decision-maker. Lastly, Gong *et al* (2014) used an interactive method, which they call an Interval Multiple-objective Optimization Problem, which has the goal of finding the decision-maker's preferred solution, by having them input the importance relations between the objectives during the evolution.

The methods that we mostly intend to focus on those which deal with the reduction of the Pareto-optimal set after the Pareto frontier has been established. These methods are intended to alleviate that part of the decision-making stage by delivering a more feasible to analyze amount of solutions. Still, it can be difficult to visualize the Pareto-optimal set when there are more than two objectives, not to mention the challenge of presenting this set to the decision maker. Below are described some *a posteriori* methodologies for Pareto-optimal set analysis.

The non-uniform weight generator with pseudo-ranking scheme, developed by Carrillo and Taboada (2012) uses a weight generating algorithm to generate a set of weights for the solutions. The algorithm then performs a weighted sum of the normalized objectives, where the best solution for each set of weights is marked, and after repeating the process with all different weight sets, the solutions that are marked becomes the new pareto set, while the unmarked solutions are deleted.

Another method for reducing the Pareto set was recently developed by Fernandez (2017), using the concept of Nash Dominance. This approach uses game theory to reduce the number of pareto-optimal solutions. The approach considers each solution a strategy and each objective function a player, and has the objective of maximizing the number of players that benefits from a change in selected solutions, basically choosing to keep solutions that excel in a larger number of objectives compared to others. This generally resulted in the elimination of extreme solutions, which only yield good results in one objective while sacrificing others.

2.6. STATE OF THE ART

As previously mentioned, this is a fairly active research area in the topic of transportation, but the literature is still very limited when concerning environmental impacts. Some studies, such as Delgado (2016) are beginning to consider environmental impacts in their assessments. There is also no literature that considers stochastic effects within the costs and demands for the CCCMLP. Most, if not all research takes a deterministic approach to solving the problem at hand. A lot of the literature also considers a single carrier or a static number of hubs or hub locations. In this problem, a point-to-point network will be transformed into a hybrid hub-and-spoke network using a set of collaborative consolidation transshipment hubs, similar to the approach taken by Hernandez, Unnikrishnan, and Awale (2011), whose CCCMLP approach represented a starting point for studying the effects of rate-setting behaviors by collaborative carriers in a centralized collaborative network.

The present research proposes the development of Single and Multi-Objective Evolutionary Algorithms to solve the Centralized Carrier Collaboration and Multi-hub Location Problem. Moreover, this thesis can provide the LTL industry with initial analysis of a hybrid hub and spoke network design with stochastic cost and demand parameters. In addition, this thesis can be found as a pioneer study in introducing stochasticity into the LTL carrier collaboration.

Chapter 3: Mathematical Model

The Centralized Carrier Collaboration Problem (CCCMLP) is an NP-Complete combinatorial optimization problem. The term NP-Complete refers to the problem's computational complexity and indicates that (a) the solutions can be verified in polynomial time, and (b) the solutions cannot be located in polynomial time. In other words, the time required to solve the problem increases exponentially as the size of the problem grows. For these types of problems in which the exact solution may take extremely long computational times to find, it has become commonplace to utilize heuristic, metaheuristic, and approximation methods in order to obtain a solution. In our case, a Multiple Objective Evolutionary Algorithm (MOEA) is used.

3.1. THE CENTRALIZED CARRIER COLLABORATION AND MULTI-HUB LOCATION PROBLEM

Before explaining the algorithm used to solve and optimize the problem, we will explain the mathematical model behind the Centralized Carrier Collaboration and Multi-Hub Location Problem (CCCMLP). The CCCMLP is traditionally a P-hub median problem which seeks to determine a set of hybrid collaborative consolidation transshipment hubs for a central entity (e.g., third party logistics firm) to help establish a collaborative hybrid hub-and-spoke system that minimizes the total collaborative costs for the set of collaborating carriers. In our study we are considering not only cost as an objective function, but also the variability of this cost, the environmental impacts associated with these network change decisions, and the expected profits that might be gained from the collaboration. Hence, a carrier in this system is classified as either a collaborative carrier (shares the costs to set up hybrid hubs), or non-collaborative (decides to ship directly). The operational networks of the collaborating carriers can be completely identical geographically or overlap in some segments relative to other carriers in the collaboration.

The collaborative rate structure of the collaborative carriers is represented by revenue-oriented behavior. If a collaborative opportunity cannot be identified with regards to hybrid

collaborative consolidation transshipment hubs, a non-collaborative option is considered. It is assumed that the costs of shipping directly fall upon the carrier itself.

The following assumptions are made in the CCCMLP: (i) candidate hybrid collaborative consolidation transshipment hubs are uncapacitated, and (ii) homogenous products are shipped. In addition, the problem is stochastic in the sense that the costs and demands are not known, but instead there is probability-based alternatives in these values, given as UGFs. Also, the available holding times at facilities are time invariant.

In the mathematical formulation of the CCCMLP, we need to understand how the objective functions are being evaluated. As a multi-objective evolutionary algorithm, we need to understand that multiple objectives must be evaluated for each solution before obtaining the fitness metrics for each solution. For obtaining the first objective function, that is, the cost function, we must do the following:

Let the carrier company be denoted by $q \in Q$, the origin of a shipment by $i \in I \subseteq N$, its destination by $j \in J \subseteq N$, and the hubs by which it may travel by $k, l \in N$, where N is the total number of nodes in the network.

Each carrier q has an associated demand denoted by d_{ijq}, the number of shipments that must be made from the origin point i to the destination point j by the carrier q.

The collaborative carrier revenue-oriented cost associated to demand d_{ijq} is given by

$$C_{ijkl} = C_{ik} + \delta C_{kl} + C_{lj}$$

where δ is the collaboration discount associated with transporting from hub k to hub l, and $0 \leq \delta \leq 1$, the collaborative discount of a shipment from hub to hub. The cost associated with a carrier q establishing a hub in node k is denoted as

$$P_{kq} = \vartheta_{kq} + \phi_k$$

where ϑ_{kq} is the holding cost associated with carrier q storing merchandise at the hub in node k, and ϕ_k is the connection cost of the hub, that is, the cost associated with the loading and unloading of merchandise from one truck to another. The costs of shipping directly from node to node by each carrier will be denoted by W_{ijq}.

Additionally, let:

$Y_{ijklq} = 1$, if a shipment is sent from node *i* to node *j* via the hubs *k* and *l* by the carrier q. That is, if the shipment is sent through a collaborative network. Otherwise, it will be equal to 0.

$V_{ijq} = 1$, if a shipment is sent directly from node *i* to node *j* by carrier q, and 0 otherwise.

$X_k = 1$, if the node at point k will become a hub, and 0 otherwise.

Problem Constraints

For the solution to this problem, we have determined the following constraints to be relevant to this problem:

$$\sum_k X_k = p$$

Makes it so that there is an exact number of hubs that can be implemented. In order to limit the number of routes from one point to the next, the constraint below impedes the programming of more than one different route between two points in the system:

$$\sum_k \sum_l Y_{ijklq} + V_{ijq} = 1 \ \forall i, j, q$$

The next constraints state that shipments from origin $i \in I$ to destination $j \in J$ cannot be assigned to a hub at location $k \in K$ or $l \in L$ unless a hybrid collaborative consolidation hub is located in these candidate sites:

$$\sum_l Y_{ijklq} \leq X_k$$
$$\sum_k Y_{ijklq} \leq X_l$$

Also, this next constraint ensures that the shipment will only go through the collaborative network if the cost of going through it is smaller than the cost of direct shipping. Here, γ denotes the profit margin expected by a company in order to participate in the collaboration.

$$C_{ijkl} Y_{ijklq} \leq W_{ijq}\left(1 - V_{ijq}\right)(1 - \gamma) \ \ \forall i, j, k, l, q$$

The last constraint set forces variables X, Y, and V into the binary space:

$$X_k \in \{0,1\}$$

$$Y_{ijklq} \in \{0,1\}$$

$$V_{ijq} \in \{0,1\}$$

The first constraint, however, limits the solution space for multi-objective problems and therefore will not be used in our case studies. It has only been left in the text to better illustrate the format of a hub location problem.

3.2. THE MULTI-OBJECTIVE EVOLUTIONARY ALGORITHM

The proposed MOEA has two main goals in terms of the Pareto front: Proximity and Diversity. Proximity refers to finding solutions that are as close as possible to the Pareto frontier, while Diversity means finding solutions that are spread over it, meaning that it attempts to find solutions that differ in their objective values as much as possible. The algorithm follows the following procedure:

0. START
1. Generate an initial population. This initial population can be either fully random or pseudo-random, and will follow some specific parameters, such as population size.
2. Calculate Objective Function Values for all these solutions.
3. Follow the Pareto Dominance Criterion, explained in section 3.2.7.
4. Calculate the Fitness Metrics, also explained in section 3.2.5.
5. If the stopping criteria have been met, go to step 10. Otherwise, go to step 6.
6. Obtain the Aggregated Fitness Metric from those calculated in step 4.
7. Rank the solutions according to the Aggregated Fitness Metric
8. Select the Parents from the current population to generate the next set of individual solutions. The process for this selection is defined in section 3.2.6

9. Generate a new population as described in section 3.2.6 and return to step 2.

10. Obtain the Pareto-Optimal set

11. END

This process is summarized in Figure 3.1 below. Each step is also described in detail in the following sections.

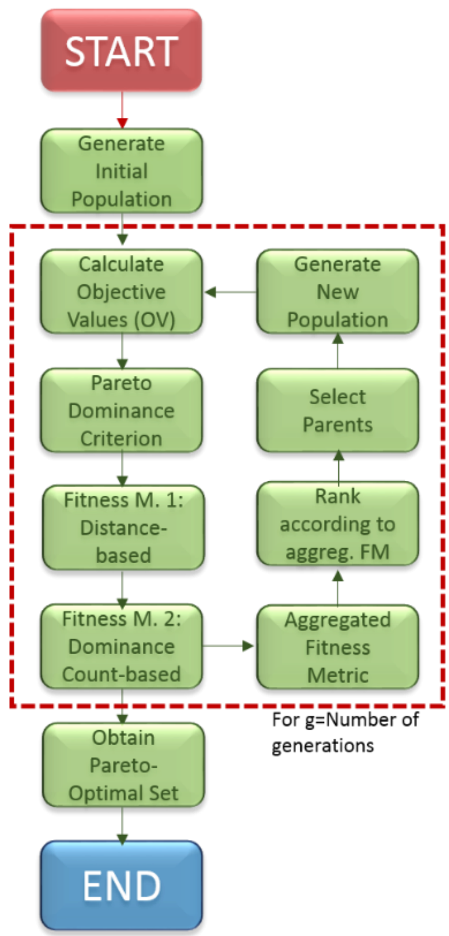

Figure 3.1. MOEA flowchart

3.2.1. The Chromosome

The chromosome is part of the foundations of a Genetic Algorithm. Each chromosome represents a solution inside the search area, and are used to generate and guide new solutions

towards an optimal response. In a multi-objective algorithm, however, rather than being guided to an optimal point, they are guided towards a non-dominated optimality curve.

The MOEA that we're proposing uses a chromosome comprised of several parts: First, a binary vector of length *n*, where *n* represents the number of nodes in the network. Each gene in this part of the chromosome will represent a single specific point in the network, and this node will be selected to be transformed into a transshipment consolidation hub if the value of its corresponding gene is one (1), and will otherwise be denoted with a zero (0). The next part will be used to include the collaboration requirement γ. This requirement symbolizes each carrier's expectations for collaboration and was further explained in section 3.1. Lastly, the objective and fitness functions will be embedded to each chromosome. This final addition to the chromosome is used mainly for ranking and visualization purposes, and is considered a separate part from the previous two, which describe the network configuration. An example of how a chromosome would look is shown below in figure 3.2:

1	0	0	0	1	0	1	0	0	1	γ	OF_1	OF_2	OF_3	OF_4	OF_5	OF_6	AF

Figure 3.2: Example Chromosome

3.2.2. The Initial Generation

This first step in the algorithm will randomly generate a pre-defined number *m* of individuals as an initial set of solutions to explore. This initial population serves to initiate the exploration of the solution space, and the size of this population will influence how quickly the algorithm can converge around an optimum point in the case of single objective optimization, or how quickly it can create a Pareto-optimal set. For the generation of our initial population, the algorithm creates *m* chromosomes, randomly generating zeros and ones for the first part, and choosing a γ value randomly from one of the pre-set values it can take.

3.2.3. The Objective Functions

Cost Functions

This objective function seeks a set of candidate hybrid collaborative consolidation hubs as to minimize the total transportation collaborative costs in a supply chain. It consist of three terms, the first term represents the total transportation costs associated to the carrier collaborative, the second part represents the total costs associated with carriers not collaborating and shipping directly, and third represents the total carrier collaborative costs associated with locating a collaborative candidate hybrid consolidation facilities. The collaborative transportation costs are obtained as the summation of the product of the cost of travel for a shipment C_{ijkl}, the collaborative carrier demand d_{ijq} and Y_{ijklq}(the decision on whether a shipment travels via the collaborative hubs). The non-collaborative costs are obtained as the summation of the cost of shipping directlyW_{ijq}, the collaborative carrier demand d_{ijq}, and the V_{ijq} (the decision on whether to ship directly). The collaborative candidate hybrid consolidation hub location costs are obtained as the summation of the product of the costs of locating a collaborative hubP_{kq}, and the X_k(the decision on whether a collaborative facility is located). The following equation subject to the previously specified constraints represents the mathematical formulation of the centralized carrier collaborative multi-hub location problem (CCCMLP).

$$\min \sum_{i}\sum_{j}\sum_{k}\sum_{l}\sum_{q} C_{ijkl} d_{ijq} Y_{ijklq} + \sum_{i}\sum_{j}\sum_{q} W_{ijq} d_{ijq} V_{ijq} + \sum_{k}\sum_{q} P_{kq} X_k$$

However, it is important to understand that this is a deterministic approach to the formulation. As UGFs are used, however, it is important to specify the operators to be used when aggregating all UGFs in the problem instead of the regular arithmetic operations. Two composition operators will be used. Firstly, an additive operator, and secondly, a multiplicative operator. The functions for these are as follows, respectively:

$$u_1 \underset{+}{\otimes} u_2 = \sum_{i=1}^{k_a}\sum_{j=1}^{k_b} P_{ai} P_{bj} z^{(x_{ai} + x_{bj})}$$

$$u_1 \underset{\times}{\otimes} u_2 = \sum_{i=1}^{k_a} \sum_{j=1}^{k_b} P_{ai} P_{bj} z^{(x_{ai} * x_{bj})}$$

Using these, it will be possible for the algorithm to calculate the aggregated UGF for each solution in order to calculate expected cost. The equation to calculate the expected value for an UGF is as follows:

$$\bar{z} = E(z) = \sum_{i=1}^{k} P_i * x_i$$

As explained in previous sections, the Universal Generating Function works by associating the values that a variable can take with the probability of the variable taking these values in a polynomial notation. In that way, it makes sense that the expected value can be calculated in this way. Similarly, the variance can be calculated with the following formula:

$$\sigma_z^2 = Var(z) = \sum_{i=1}^{k} P_i * (x_i - \bar{z})$$

The previous formula is based on the standard formula for a regular discrete distribution's variance. These two formulas are used to derive the Expected Cost and Cost Variance objective functions from the final aggregated UGF obtained from the procedure explained below:

In obtaining the optimal network configuration, each possible origin-destination pair must select the best route alternative. A greedy approach is used for this step in the algorithm to reduce computational complexity, and in this sense the algorithm will strictly choose the route with the best (least) total expected cost rather than perform a tradeoff or consider other objectives in this specific point in time. The cost of a non-collaborative route is given by W_{ijq} for an i, j origin-destination pair, for a carrier q. For a collaborative route, however, cost is determined by $C_{ik}+\delta C_{kl}+C_{lj}$, where k and l represent the intermediate hubs that the shipment will take, and δ represents the hub-to-hub shipment discount.

The selected routes will have their UGFs stored in what we have denoted as the optimized costs matrix, *S*. These are the UGFs that will be inputted into the aggregating algorithm when calculating the final aggregated UGF for the system.

Environmental Impact Function

A similar method is used with the next objective of minimizing total Global Warming Potential (GWP), an environmental impact related to how much heat the gasses released by the system can trap into the atmosphere. In this case, however, the routes will have previously been chosen by the greedy approach with respect to cost, so the selection of the values that will go into the objective function have already been decided. For the calculation of GWP, the objective function looks as follows:

$$\min \sum_{i}\sum_{j}\sum_{k}\sum_{l}\sum_{q} C_{ijkl} d_{ijq} Y_{ijklq} + \sum_{i}\sum_{j}\sum_{q} W_{ijq} d_{ijq} V_{ijq}$$

The same variables for collaborative cost and non-collaborative cost have been repurposed, this time meaning collaborative and non-collaborative emissions. This function, however, is missing the last term that the cost function had, as it is considered part of the problem that the hubs will be located within already-existing structures, so the environmental impacts associated with these facilities have been left out, as the changes in emissions are thought to be negligible before and after the establishment of the facilities as hubs.

Profit Functions

Lastly, the profit function will be associated with the demand UGFs, as profits will depend entirely on demand serviced. The revenues, however, are fixed, and independent of the resulting costs. The profits differ in collaborative and non-collaborative scenario alternatives and are the result of calculating the expected value of the costs, subtracting these from the revenues, and multiply these values by the expected demand before summing them all. In order to better reflect the stochastic properties of the problem, we'll also include a profit range function, with the objective of minimizing this objective, as certainty in the numbers is important. We will determine

a 95% confidence interval around the expected value of the profit. This can be done by modeling the revenue values as single-state UGFs, with probability of 1, and using a subtractive operator $\underset{-}{\otimes}$ when aggregating the functions:

$$u_1 \underset{-}{\otimes} u_2 = \sum_{i=1}^{k_a} \sum_{j=1}^{k_b} P_{ai} P_{bj} z^{(x_{ai} - x_{bj})}$$

We can then use this to obtain an aggregated profit function. At this point, the following methodology will be used to determine the confidence interval: For the lower level, start at i=1. When

$$\sum_{i=1}^{a} P_i = 0.025,$$

Stop. Determine the value of *a*. At this point, assign x_a to be the lower limit for the confidence interval. For the upper limit, a similar approach is taken: Start at i=1, and when

$$\sum_{i=1}^{b} P_i = 0.975,$$

Stop, and set x_b as the upper interval limit. The range between these two values will be the last objective function value to be considered in this study.

To summarize, the objective functions are:

1. Expected Cost
2. Cost Variance
3. Global Warming Potential
4. Expected Profit
5. Lower Profit Confidence Interval
6. Upper Profit Confidence Interval

3.2.4. The Pareto Dominance Criteria

As part of the optimization process, it is important to clearly delineate the difference between dominated solutions and non-dominated solutions. While it may be easier to define a non-dominated solution as one that is not dominated, we will explain the criteria for domination and non-domination below.

- A dominated solution is one for which another solution exists that is strictly better than it in every one of the objective functions being considered.
- A non-dominated solution is one for which:
    - There is no other solution that is strictly better than it in all objectives, and
    - Is better than other solutions in at least one objective

The set of non-dominated solutions is that which is asymptotically closest to the ideal solution point. This proves valuable to the decision-maker as dominated solutions are worse than non-dominated ones in every single aspect (or objective). Delivering solely non-dominated solutions means delivering those that will have the best possible objective values and reducing the total number of solutions that can be given.

The determination of dominance is also important for the algorithm process, as the non-dominated solutions will take priority when selecting which members of the population will be kept for the next generation and which will be selected for reproduction.

3.2.5. The Fitness Functions

The developed MOEA has two main goals. Proximity which represents the closeness to the Pareto front, and diversity that has the objective of maintaining population diversity as explained in Taboada & Coit, (2012). These two goals are evaluated with two fitness metrics which are distance-based (diversity), and dominance count-based (proximity).

The first fitness metric, which is distance-based, assigns a higher fitness value to those solutions that are farther away from other solutions in the Pareto front. The aim of giving further

solutions a higher fitness value is to give these a better chance of reproducing, thus creating a wider spread of solutions. This way, it is possible to use a fitness metric to increase diversity of the Pareto-optimal solutions. The following steps are followed to assess this fitness metric:

1. **Normalization**: In this step, the objective function values of the solutions are normalized in order to make sure that all values are within the same order of magnitude to prevent unit discrepancies. The normalization of the solutions follows the following equation:

   $$F_{ij}(x) = \frac{f_{ij}(x) - f_j^{min}}{f_j^{max} - f_j^{min}}$$

   Where:

   $F_{ij}(x)$ is the normalized value of objective function j for solution i,

   $f_{ij}(x)$ is the current value of objective function j for solution i,

   f_j^{max} is the maximum value that objective function j has taken in the current solution set, and

   f_j^{min} is the minimum value that the objective function j has taken in the set.

2. **Distance Calculation**: In this second step, the distance metric for each of the solutions is calculated. Euclidean distance is used in this step, and the following formula is used to determine the total distances between one solution and all others:

   $$d_i = \sum_{k=1}^{n} \sqrt{\sum_{j=1}^{m} \left(OF_{ij} - OF_{kj}\right)^2}$$

   Where:

   d_i is the total distance metric for solution i,

   OF_{ij} and OF_{kj} are the objective function values of objective j for solutions i and k, respectively,

   m is the total number of objective functions, and

   n is the total number of solutions in the set.

After obtaining the distances for all points, these values are once again normalized using the previously described normalization formula. This will allow for easier aggregation in the final step of obtaining the Aggregated Fitness Function.

The second distance metric, which is based on dominance count, is based on the previously described concept of Pareto-dominance. The premise for this metric is that solutions which dominate others are closer to the Pareto frontier than those that are dominated, again based on the concept of Pareto dominance, which implies that a non-dominated solution is closer to the Pareto-frontier and the ideal solution than a dominated one, which is to say, it has more proximity. In this way, all solutions are compared against each other, and when a solution i dominates a solution j, the dominance count metric for this solution is increased by one. At the end of this process, the dominance counts for all solutions are normalized in preparation for the final step.

Lastly, the Aggregated Fitness Function is comprised of a simple weighted sum between the distance-based and dominance count-based metrics described above. The weights can be set as equal if no preference exists, or one of the metrics can have higher weights in order to give more priority to one of the objectives of proximity or diversity.

3.2.6. The Crossover Function

Crossover is the process by which a new generation, or set of solutions, is created. The process can have a single crossover point or multiple crossover points. In our multi-objective study, we use two crossover points. The crossover points represents the locations in which the information from both selected parents mixes in the new solution. This is shown in figure 3.3 below. This exemplifies the two-point crossover we are performing. The crossover inserts genes from the first parent up to a randomly selected crossover point p, where $p<n$, n being the number of points in the network we are analyzing. Then, it inserts genes from the other parent up to n. Lastly, the last gene in the chromosome, the γ value, is randomly chosen between one of the two parents.

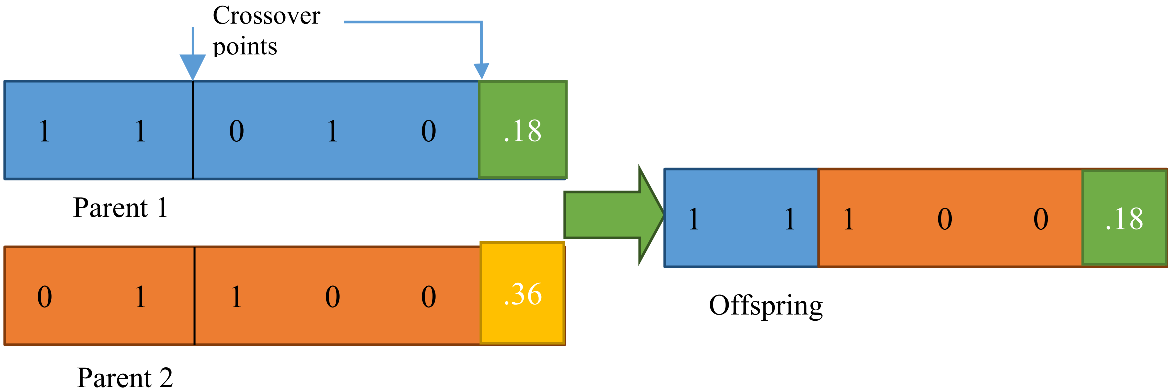

Figure 3.3: Two-Point Crossover Function

The selection of parents for the crossover also has different alternatives. There is a purely random selection process, in which two individuals are chosen from the parent population and used in crossover. This method has the drawback of not choosing better parents, which can cause the algorithm to take longer to find the optimal solutions. Another selection method is the roulette method. In this method, each solution has a probability of being chosen relative to their fitness values. This method can, in this way, have a better chance of selecting more "fit" solutions as the parents, which may create better offspring. Lastly, the method that we are going to use is called the tournament selection. This method randomly chooses two solutions to compete to be the first parent, and the one with the better fitness function gets the right to be the first parent. The same process is used to select the second parent. These two winning individuals are then used to perform the crossover process.

Mutation can also occur during the crossover process. The mutation function helps the GA by inserting some random variation in the evolutionary process, to avoid it falling into local optima. It works by slightly changing the offspring from two parents with a small probability. In this sense, if mutation occurs, and with our mostly binary chromosome, a randomly selected node in the network will change its value from one to zero, or from zero to one.

An example of how mutation can affect a chromosome is shown in figure 3.4 below:

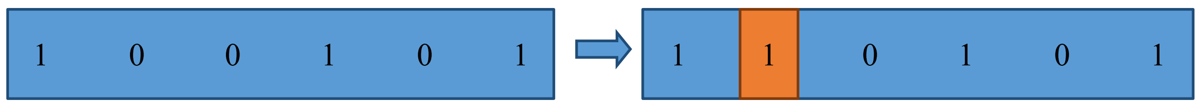

Figure 3.4: Mutation Function Example

On top of the aforementioned process, elitism is also used to fill the new population with individuals. Elitism is a function within the GA to help keep the optimization process moving towards what it may assume at the moment to be. It selects the top percentage of the population in a generation, based on the fitness function, and adds it directly to the next generation of the population.

3.2.7. Obtaining the Pareto-Optimal Front

At the end of each generation and before creating the new one, all non-dominated solutions of a generation are stored separate from the rest of the algorithm. After the stopping criteria (described in the next section) have been met, these solutions are once again checked for dominance, and the dominated solutions are deleted from the set. This final set of solutions is the closest the algorithm has been able to get to the true Pareto frontier. As so, this solution set represents the Pareto-optimal set of solutions. This final set of solutions is what is then analyzed in Post-Pareto optimality procedures or presented directly to the decision-maker.

It is important to reiterate that the basic principle for a solution to be non-dominated is that there is no other single solution which is better than it in all objectives that are being considered. Generally speaking, this becomes more likely as the number of objectives increases. Brockhoff and Zitzler (2006) argue that a high number (≥3) of objectives can cause difficulties not only in approximating the Pareto-optimal set, but also in data visualization, processing time, and decision-making, as the multi-dimensionality of the solutions can present a challenge to both the researchers

and the decision-makers. Dimensionality reduction methods are proposed by Brockhoff and Zitzler. However, these methods are not always suitable to use and are not very commonplace.

3.2.8. Algorithm Termination

The algorithm finishes its optimization process and goes into the Pareto-determination phase when certain criteria has been met. These criteria vary between different problems and the researcher's preference. One such criterion is based on convergence. Here, if the best solution does not change after a set number of generations, the algorithm decides that the best solution has been found and there is no need to continue calculating more iterations. This criterion is mostly applied to single-objective optimization problems but can be slightly adapted to fit multi-objective problems. A suitable change would be to stop if no new non-dominated solutions has been saved after a set number of iterations. Another, simpler stopping criterion is simply a pre-defined number of iterations. With this criterion, there is no need to check the solutions over many iterations, but simply decide how many times the iteration process will be repeated. Lastly, it is possible to use a combination of the two in order to attempt to minimize the total computational time.

In this study, a simple pre-defined number of generations is used as a stopping criterion, as determining convergence can be difficult in multi-objective optimization with a high number of objectives.

Chapter 4: Numerical Example

In this chapter, an example of the previously described algorithm will be presented. In this case, a 10-node network of terrestrial shipment facilities shared by 2 carrier companies will be studied, with the purpose of shifting these carrier's transportation network from a pure point-to-point system to a hybrid hub-and-spoke network, calculating the improvements in the objective functions after such process has been completed.

4.1. Numerical Data

This chapter will serve to show the information that will be used by the algorithm to perform the evaluation and selection of nodes to be transformed into hubs. The data was extrapolated from Hernandez *et al*'s (2011) study.

Due to space and format constraints, the UGFs for each data point will not be presented. Instead, this data will be represented as a summary of the properties of the data (expected value, variance, maximum value, and minimum value). This applies for the costs and demands of the shipments and hub establishment costs, as environmental impacts and revenues are considered as deterministic values.

4.1.1. Non-collaborative Shipping Cost Data

This sub-section presents data summaries regarding costs of non-collaborative shipments.

Table 4.1: Non-collaborative shipping cost means for carrier 1

O/D	1	2	3	4	5	6	7	8	9	10
1	0	134.1	82.8	282.5	182	60	73.7	133.4	212.2	197.3
2	136.6	0	79.8	165.6	85.4	87.3	181.7	171.2	103.7	102.3
3	84.2	78.5	0	187.9	139.2	89.8	84.8	86.7	119.5	197.1
4	283.7	166	213	0	220	276.1	270.4	217.8	81	238.9
5	160.6	69	149.4	211.6	0	108	210	240.4	187	39
6	52	87.8	83.8	265	122.2	0	106.9	185.7	186.8	143.6
7	75.8	153.1	83	244.8	223	130.4	0	75.1	189.2	240

8	150.1	154	103	235	220.3	155.5	76	0	134.05	251.1
9	196.6	103	129.9	79	169.4	181.9	208.8	146.8	0	212
10	185	116	176.3	241	37	129	229.9	286.4	198	0

Table 4.2: Non-collaborative shipping cost means for carrier 2

O/D	1	2	3	4	5	6	7	8	9	10
1	0	134.1	92	299.8	149.9	58	76	155.8	189.8	201.7
2	126	0	78.2	150.8	81	92.7	181	182	101	103.9
3	77.8	90	0	193	131.6	85	90.1	107	118.3	167
4	253.7	153.7	215.4	0	210	273	279.4	193	80.9	262.9
5	172	79	133.7	227	0	113	230	216.8	193.2	35
6	56	106.9	89	256	109.2	0	126	190	176	126
7	71.1	158.8	86	266.7	218	123.1	0	90	187	254
8	156.4	159	91.2	222.4	216.4	176.5	85.2	0	139	268.8
9	206.2	119.1	118	74	179.1	172	194.9	133.9	0	212.3
10	174.05	112	169.2	230	34	139.6	274	253	196	0

Table 4.3: Non-collaborative shipping cost variances for carrier 1

O/D	1	2	3	4	5	6	7	8	9	10
1	0	13.09	3.36	8.25	6	0	2.01	6.64	4.56	3.01
2	7.04	0	2.16	0.84	5.04	3.21	0.81	1.76	4.41	3.81
3	5.76	11.85	0	10.69	11.76	7.96	4.36	19.81	8.85	12.69
4	9.81	17.5	14.4	0	5.4	11.89	6.24	8.76	3	5.49
5	15.24	1.2	7.04	24.64	0	1	30	12.64	4.8	0
6	0.6	10.16	2.16	19.2	2.16	0	4.89	3.01	3.36	4.44
7	2.16	3.09	4	6.36	2.4	3.64	0	3.09	3.36	4
8	4.09	6	4	2.4	4.51	5.25	1	0	4.1475	4.89
9	1.44	4	1.89	0	1.44	1.49	11.76	1.56	0	3
10	1	1	3.01	1.2	2	0	0.09	0.84	1	0

Table 4.4: Non-collaborative shipping cost variances for carrier 2

O/D	1	2	3	4	5	6	7	8	9	10
1	0	1.89	4	5.16	0.09	0	3.2	9.36	11.76	1.81
2	1	0	2.56	7.56	0	0.21	2.4	2.4	1	10.09
3	0.16	3.2	0	4	2.04	1	1.89	1	1.41	2.4
4	4.41	4.41	3.64	0	6	2.4	1.44	1.8	0.09	3.29
5	4	0	4.01	1.6	0	1	1.8	2.76	2.16	0
6	0	3.09	0.8	0.6	11.76	0	3	0.8	0.8	1.8
7	0.09	2.56	0	12.81	7	4.09	0	4	4	5.4

8	3.64	8.4	5.76	7.84	5.04	3.15	5.76	0	4	2.76
9	8.16	1.89	1	0	1.89	4	4.29	8.19	0	5.41
10	3.5475	4	5.76	4	0	10.24	6	6	1	0

Table 4.5: Non-collaborative shipping maximum costs for carrier 1

O/D	1	2	3	4	5	6	7	8	9	10
1	0	139	84	285	184	60	75	136	214	199
2	139	0	81	166	89	89	184	174	107	105
3	86	83	0	191	142	93	87	92	122	203
4	291	171	216	0	226	281	274	222	84	241
5	166	71	151	216	0	109	220	243	189	39
6	53	91	85	271	124	0	109	188	189	146
7	77	155	85	247	225	132	0	77	191	242
8	152	156	105	237	221	157	77	0	137	258
9	198	105	132	79	170	183	212	148	0	213
10	186	117	178	243	39	129	230	287	199	0

Table 4.6: Non-collaborative shipping maximum costs for carrier 2

O/D	1	2	3	4	5	6	7	8	9	10
1	0	135	94	302	150	58	78	159	194	203
2	127	0	79	155	81	93	183	184	102	107
3	78	92	0	195	134	86	91	108	120	168
4	255	155	217	0	213	274	280	194	81	265
5	174	79	135	228	0	114	231	221	195	35
6	56	109	90	257	112	0	129	191	177	127
7	72	162	86	270	222	125	0	92	189	257
8	158	162	96	225	218	178	87	0	141	271
9	209	120	119	74	180	174	197	136	0	215
10	176	114	171	232	34	142	276	256	197	0

Table 4.7: Non-collaborative shipping minimum costs for carrier 1

O/D	1	2	3	4	5	6	7	8	9	10
1	0	125	80	276	179	60	71	129	209	195
2	131	0	78	163	83	84	181	170	101	100
3	80	74	0	183	135	87	81	81	115	191
4	279	161	207	0	217	272	267	213	80	235
5	156	68	143	203	0	107	200	234	183	39
6	51	84	81	259	121	0	104	184	185	141

7	74	151	81	241	221	127	0	73	187	238
8	147	151	101	233	213	151	75	0	131	251
9	195	101	129	79	167	179	203	145	0	209
10	184	115	174	240	35	129	229	285	197	0

Table 4.8:Non-collaborative shipping minimum costs for carrier 2

O/D	1	2	3	4	5	6	7	8	9	10
1	0	132	90	295	149	58	74	151	187	200
2	125	0	75	149	81	92	179	180	100	100
3	77	88	0	191	130	84	88	106	117	164
4	249	149	212	0	208	270	277	191	80	260
5	170	79	130	225	0	112	227	216	192	35
6	56	105	88	255	105	0	125	189	175	124
7	71	158	86	261	214	120	0	88	185	251
8	153	155	90	218	211	174	81	0	137	265
9	203	117	117	74	177	170	192	130	0	210
10	172	110	165	228	34	134	271	251	195	0

4.1.2. Collaborative Shipping Cost Data

This sub-section presents data summaries regarding costs of collaborative shipments.

Table 4.9: Collaborative shipping cost means for carriers 1, 2

O/D	1	2	3	4	5	6	7	8	9	10
1	0	45.5	30.8	101.2	59.8	20	25.95	50	69.9	66.1
2	48	0	28	54.95	27	35	60.1	61.1	38	40.05
3	31.1	26.95	0	66	50.1	29.1	31	33	43	60.1
4	101	57	69.1	0	78.1	87.9	92.1	75	27	89.1
5	57.9	28	48	76.1	0	37.9	78.9	87.15	67.1	14.1
6	19.1	36	31	86	39	0	42	60	65	47
7	27.9	60	29	90	78	41.65	0	28	64.1	86.1
8	54	61	35.1	74	86	61.1	27	0	48	98
9	72.1	40	42	27	63	69	69	50	0	75
10	67	39	63	86	14	46	83	102	77	0

Table 4.10: Collaborative shipping cost variances for carriers 1, 2

O/D	1	2	3	4	5	6	7	8	9	10
1	0	33.25	12.36	26.56	11.76	26.4	19.7475	42	44.39	20.79

2	19	0	19	13.1475	9.6	10.8	20.29	22.29	14.4	20.8475
3	8.89	14.6475	0	8.2	11.49	9.49	14.4	14.4	25.8	11.49
4	15	12	9.69	0	16.29	29.79	14.49	11	28	32.49
5	45.09	27.6	18.4	20.29	0	18.09	32.49	15.7275	27.09	18.09
6	13.29	18.8	36	25	20	0	27.6	16.2	25	8.8
7	24.09	36	16.2	12.8	9.5	1.3275	0	12	24.69	13.29
8	16	5.4	13.29	8.4	26.2	10.89	3.2	0	9	0.9
9	9.49	19	19	28	9	12.8	14.4	25	0	33.6
10	7	13.5	19	18	0.6	27	16	19	18.4	0

Table 4.11: Collaborative shipping maximum costs for carriers 1, 2

O/D	1	2	3	4	5	6	7	8	9	10
1	0	51	35	105	63	26	32	58	77	71
2	52	0	32	58	29	38	65	66	41	44
3	34	31	0	68	55	31	34	36	47	64
4	105	60	72	0	82	93	96	78	32	93
5	66	32	52	81	0	42	85	90	73	19
6	22	40	37	91	44	0	46	66	70	49
7	31	66	32	94	81	43	0	31	68	89
8	58	64	38	78	91	65	29	0	51	99
9	75	44	46	32	66	73	73	55	0	79
10	69	41	69	90	15	52	87	106	81	0

Table 4.12: Collaborative shipping minimum costs for carriers 1, 2

O/D	1	2	3	4	5	6	7	8	9	10
1	0	38	27	92	54	14	21	42	59	61
2	41	0	21	49	21	31	54	53	32	33
3	27	22	0	61	43	22	25	27	35	56
4	95	51	63	0	73	81	87	71	20	79
5	51	19	41	69	0	33	73	81	61	10
6	14	29	25	81	34	0	33	57	60	41
7	19	54	23	86	74	40	0	22	57	81
8	50	58	30	71	80	58	25	0	45	97
9	68	34	36	20	60	65	65	45	0	65
10	63	31	59	81	13	41	79	96	71	0

4.1.3. Demand Data

This sub-section will present demand data summaries for carriers 1, 2.

Table 4.13: Demand means for carrier 1

O/D	1	2	3	4	5	6	7	8	9	10
1	0	482.9	285.4	883	609.2	155	261.4	486.3	689	669.9
2	413.85	0	268.41	540.55	246.35	284.06	498	611.8	401.8	364.6
3	257.62	251.5	0	608.2	437.6	302.6	310.3	336.3	419.3	573.6
4	911.6	588.8	584.2	0	725.5	939	928.75	749.5	273.3	749.5
5	568.75	226.9	489	685	0	330	785	760	540.1	119
6	162	292.4	300.6	810.44	374.2	0	373.7	553.6	595	450
7	254.8	528.2	285	851.2	653	392.2	0	270	700	844.6
8	452	563	357	756.6	831	628	251.2	0	463	847
9	760	397.5	425	235	532.2	619.3	566.8	436.6	0	767.5
10	680.5	348	618.05	764.5	401	764.1	793	944	726	0

Table 4.14: Demand means for carrier 2

O/D	1	2	3	4	5	6	7	8	9	10
1	0	495	293.1	980	498.8	198	240	461.9	718	622.7
2	445.9	0	285	548.6	249	292	535	514.1	337.6	381.7
3	311.4	140.49	0	670.45	531.2	262	300	356.5	367.5	555.8
4	841.5	499	628.2	0	653	874	921	766.4	277	866.75
5	566	271.6	544	653	0	378.5	742	769.4	590.6	124.45
6	162	306.2	289.5	778	375	0	408	523.23	580.2	495
7	282	506.8	278.2	960.3	734.2	429	0	262.2	686	902
8	452	558	303	709	818.1	556.6	285	0	469.4	865
9	623.2	355	416	229	588.9	595	560.2	530.2	0	685
10	608.2	395.1	563.4	768.5	127	409.6	764.1	926.6	718.6	0

Table 4.15: Demand variances for carrier 1

O/D	1	2	3	4	5	6	7	8	9	10
1	0	471	125	4926	61	937	1123	519	819	437
2	103	0	833	166	287	1118	327	150	135	189
3	126	253	0	3552	353	47	66	384	2371	898
4	1045	462	121	0	217	0	181	600	166	600
5	2156	122	54	0	0	25	412	0	686	0
6	0	449	9	142	35	0	48	30	225	1300
7	67	128	0	61	9	187	0	25	0	32
8	9	9	49	22	301	0	138	0	9	0

9	450	681	625	0	219	78	188	180	0	4506
10	110	0	30	32	54	76	9	15204	676	0

Table 4.16: Demand variances for carrier 2

O/D	1	2	3	4	5	6	7	8	9	10
1	0	25	55	225	3	0	361	1273	179	277
2	1073	0	16	529	0	1	540	151	29	101
3	49	19463	0	19	56	0	25	12	30	15
4	212	9	33	0	4	11	2	10	0	120
5	24	1067	49	4	0	7	1	0	13	493
6	0	60	42	49	20	0	28	74	66	25
7	9	29	35	10	25	0	0	13	1	97
8	12	21	0	6	202	4	2	0	112	1
9	19	1	0	4	106	106	39	78	0	25
10	13	0	9	68	0	87	35	20	9	0

Table 4.17: Demand maximums for carrier 1

O/D	1	2	3	4	5	6	7	8	9	10
1	0	502	291	1000	620	210	310	511	710	693
2	425	0	301	550	310	331	509	630	425	397
3	269	298	0	781	447	311	321	356	486	607
4	938	609	601	0	734	939	942	774	400	774
5	631	259	495	685	0	335	800	760	600	119
6	162	303	303	900	379	0	384	562	600	500
7	269	550	285	862	656	411	0	275	700	852
8	455	566	364	763	850	628	261	0	466	847
9	775	425	450	235	550	641	578	472	0	800
10	691	348	623	775	410	772	796	1025	752	0

Table 4.18: Demand maximums for carrier 2

O/D	1	2	3	4	5	6	7	8	9	10
1	0	500	302	993	500	198	259	503	733	644
2	481	0	289	573	249	293	565	551	342	391
3	320	280	0	685	542	262	305	361	373	560
4	861	500	635	0	655	881	923	768	277	882
5	570	300	551	655	0	381	743	770	593	150
6	162	314	296	785	380	0	413	534	594	500
7	285	523	283	969	739	429	0	273	688	931

8	457	561	303	711	833	558	286	0	479	866
9	632	356	416	231	602	607	567	541	0	690
10	610	396	567	781	127	421	773	933	721	0

Table 4.19: Demand minimums for carrier 1

O/D	1	2	3	4	5	6	7	8	9	10
1	0	450	263	800	600	120	220	450	650	651
2	400	0	203	523	230	239	462	600	396	351
3	241	241	0	563	400	290	300	300	350	524
4	850	547	577	0	700	939	900	725	270	725
5	500	220	480	685	0	325	754	760	500	119
6	162	250	297	800	367	0	367	550	550	400
7	240	500	285	842	650	370	0	265	700	840
8	449	560	350	750	800	628	202	0	460	847
9	700	350	400	235	512	615	550	430	0	600
10	670	348	612	760	395	751	790	700	700	0

Table 4.20: Demand minimums for carrier 2

O/D	1	2	3	4	5	6	7	8	9	10
1	0	490	283	952	496	198	221	400	701	601
2	402	0	281	521	249	291	505	510	331	367
3	302	1	0	669	521	262	295	351	362	549
4	821	490	621	0	651	871	919	760	277	851
5	560	210	537	651	0	375	741	769	585	51
6	162	293	283	771	370	0	400	513	573	490
7	279	505	271	957	727	429	0	261	685	897
8	449	551	303	706	801	551	283	0	452	864
9	621	354	416	227	574	583	553	523	0	680
10	601	395	561	759	127	402	759	921	715	0

4.1.4. Environmental Impact Data

In this section, data regarding the environmental impact of the shipments, expressed in Global Warming Potential, 100kg CO_2 equivalent, will be shown. This data is set as deterministic. The data is expressed in thousandths for better visualization of the numbers.

Table 4.21: Global Warming Potential for carrier 1 shipments x10^3

O/D	1	2	3	4	5	6	7	8	9	10
1	0	43.000	26.634	91.134	58.403	19.254	23.746	43.000	68.030	63.216
2	43.963	0	25.672	53.269	27.276	28.239	58.403	54.873	34.015	33.052
3	26.955	25.351	0	60.328	44.604	29.201	27.276	28.239	38.187	63.216
4	91.134	53.269	68.030	0	70.597	88.567	86.642	69.955	25.993	76.373
5	51.664	22.142	47.813	68.351	0	34.978	67.388	77.336	60.007	12.515
6	16.687	28.239	26.313	85.037	39.149	0	34.336	59.687	60.007	46.209
7	24.388	49.097	26.634	78.299	71.560	42.037	0	24.067	60.649	73.806
8	48.134	49.418	33.052	75.410	70.597	50.060	24.388	0	43.000	82.149
9	62.896	33.052	41.716	25.351	54.231	58.403	66.746	47.172	0	67.388
10	59.366	37.224	56.799	77.336	11.873	41.396	73.806	92.097	63.537	0

Table 4.22: Global Warming Potential for carrier 2 shipments x10^3

O/D	1	2	3	4	5	6	7	8	9	10
1	0	41.137	28.243	92.097	46.049	17.805	23.331	48.197	58.328	62.012
2	38.681	0	23.945	46.356	24.866	28.550	55.565	55.872	31.006	31.927
3	23.945	27.629	0	59.249	40.216	26.094	27.629	32.848	36.532	51.267
4	77.668	46.969	66.310	0	64.468	83.808	85.650	59.249	24.866	80.738
5	52.802	24.252	41.137	69.687	0	34.690	70.608	66.617	59.249	10.745
6	17.191	32.848	27.322	78.589	33.462	0	38.681	58.328	54.030	38.681
7	21.796	48.811	26.401	81.966	66.924	37.760	0	27.629	57.407	77.975
8	48.197	48.811	27.322	68.459	66.617	54.337	26.401	0	42.672	82.580
9	63.240	36.532	36.225	22.717	54.951	52.802	59.863	41.137	0	65.389
10	53.416	34.383	51.881	70.608	10.438	42.979	84.115	77.668	60.170	0

Table 4.23: Global Warming Potential for collaborative shipments x10^3

O/D	1	2	3	4	5	6	7	8	9	10
1	0	20.767	13.995	45.597	27.087	9.029	11.738	22.573	31.602	29.796
2	21.670	0	12.641	24.830	12.189	15.801	27.087	27.539	17.155	18.058
3	13.995	12.189	0	29.796	22.573	13.092	13.995	14.898	19.413	27.087
4	45.597	25.733	31.150	0	35.214	39.728	41.534	33.859	12.189	40.180
5	26.184	12.641	21.670	34.311	0	17.155	35.665	39.277	30.248	6.320
6	8.578	16.252	13.995	38.825	17.607	0	18.961	27.087	29.345	21.218
7	12.641	27.087	13.092	40.631	35.214	19.413	0	12.641	28.893	38.825
8	24.379	27.539	15.801	33.408	38.825	27.539	12.189	0	21.670	44.243
9	32.505	18.058	18.961	12.189	28.442	31.150	31.150	22.573	0	33.859
10	30.248	17.607	28.442	38.825	6.320	20.767	37.471	46.049	34.762	0

4.1.5. Revenue Data

In this section, data regarding the revenue acquired by the carriers per each shipment delivered. The data is expressed in dollars.

Table 4.24: Revenue per shipment for carrier 1

O/D	1	2	3	4	5	6	7	8	9	10
1	0	150	90	300	200	75	90	150	225	220
2	150	0	90	175	100	100	200	185	115	115
3	90	90	0	200	150	100	95	100	130	210
4	300	175	200	0	240	295	290	230	95	250
5	200	100	150	240	0	120	230	250	200	50
6	75	100	100	295	120	0	120	200	200	160
7	90	200	95	290	230	120	0	90	200	250
8	150	185	100	230	250	200	90	0	150	275
9	225	115	130	95	200	200	200	150	0	220
10	220	115	210	250	50	160	250	275	220	0

Table 4.25: Revenue per shipment for carrier 2

O/D	1	2	3	4	5	6	7	8	9	10
1	0	150	110	310	160	70	85	170	205	210
2	150	0	95	160	90	100	195	195	110	115
3	110	95	0	210	150	95	100	115	125	180
4	310	160	210	0	220	290	290	205	90	275
5	160	90	150	220	0	120	240	230	205	50
6	70	100	95	290	120	0	140	200	185	135
7	85	195	100	290	240	140	0	100	200	275
8	170	195	115	205	230	200	100	0	150	280
9	205	110	125	90	205	185	200	150	0	225
10	210	115	180	275	50	135	275	280	225	0

4.2. Algorithm Parameters

In each replication, the algorithm will run for 100 generations with 100 individuals in each generation. There will be no constraint for the number of hubs in each individual, to allow the

multi-objective evolutionary algorithm to more freely explore the solution space. As parameters for a genetic algorithm, there is a 5% chance of mutation and a 5% chance of failure to reproduce. Elitism is to be set as the top 10% of the non-dominated solutions at each generation, which may vary from one iteration to the next.

The algorithm will be run multiple times with different criteria to their greedy approach to route selection. These criteria will be minimizing expected costs, minimizing environmental impacts, minimizing variability, and maximizing expected profits. Solutions closest to the ideal points will be saved as “most optimal” solutions and they will later be compared and analyzed.

4.2.3. Cost-Greedy route selection

The first objective to take into account for the greedy approach is cost. Through this approach, the algorithm will choose routes which would minimize the expected cost incurred on each route, ignoring all other objectives during route selection.

The results are summarized in the figure below, where all pairs of objectives are compared, showing the different Pareto fronts that are produced. It is also shown that some objectives are linearly correlated and as so do not produce Pareto fronts between them. The solution marked in red is the solution closest to the “ideal” point by Euclidean distance.

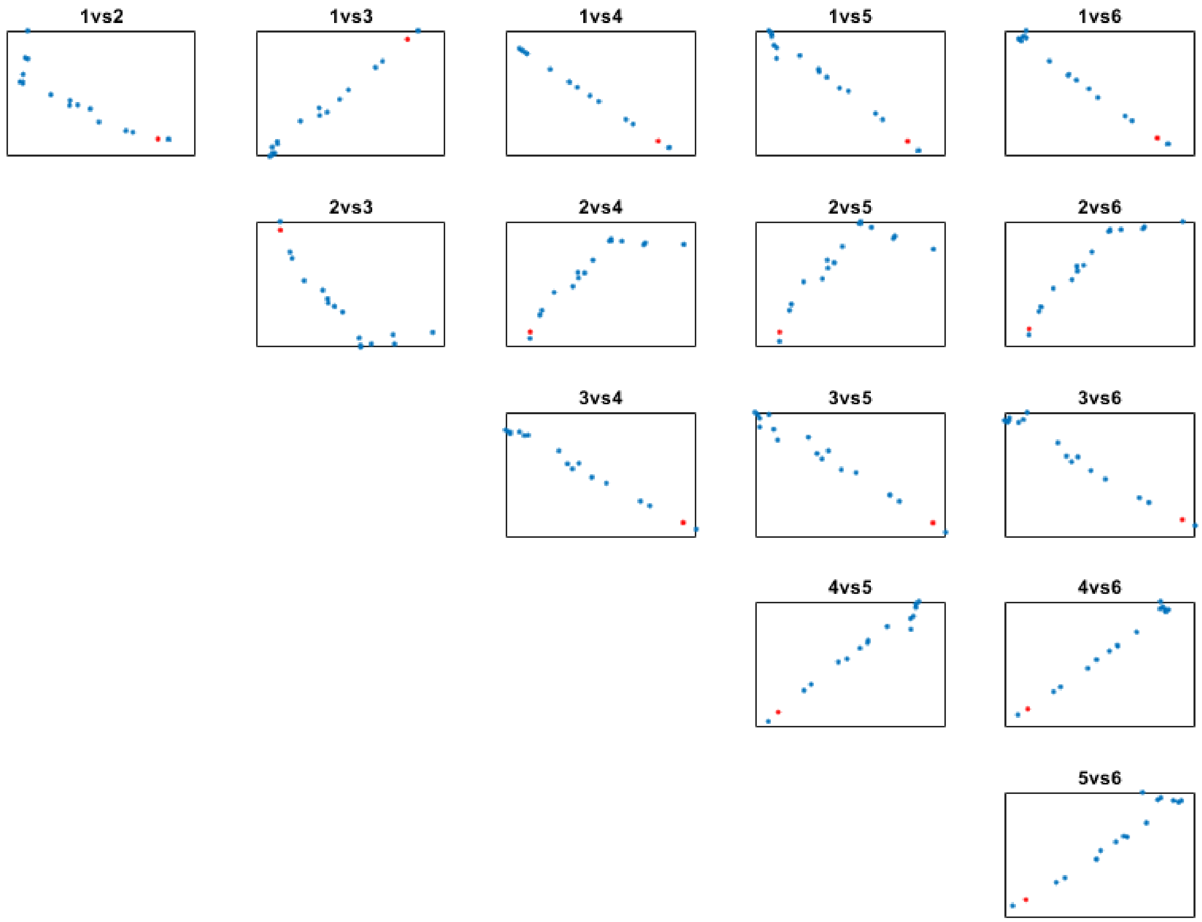

Figure 4.1: Pareto-optimal fronts for algorithm with cost-greedy route selection

The Pareto-optimal plots show strong correlations between costs and profits when cost in the main driver of the route selection process. Similarly, there is strong correlation between expected costs and global warming potential, likely due to the reductions in total transportation distances that occur when shipments are consolidated between hubs instead of deploying straight routes across the network. This was run in on MATLAB 2018a, running on a Dell Inspiron 7779 with an intel i7 7500U CPU and 16GB of RAM, with a total run time of 179 seconds.

Moreover, all solutions are listed below in table 4.26. As previously stated, nodes in the network are selected as hubs when they are denoted by a '1' in the chromosome, while they remain regular nodes if otherwise denoted as '0'. In this instance, the algorithm yielded 41 non-dominated solutions, with the number of hubs ranging from 2 to 10 and Gamma values ranging from 0.09 to

0.96. The minimum values for each objective are{1,463,510; 383,822; 2,712; 1,395,712; 643,422; 2,148,003} and the maximums are {17,200,279; 2,804,519; 5,331; 17,177,101; 13,920,853; 21,843,847} for the objectives of expected cost, cost variance, total GWP, expected profit and 2.5 and 97.5 profit percentiles, respectively.

Table 4.26: Non-dominated solutions for algorithm with cost-greedy route selection

Selected Hubs										Gamma	E(Cost)	Var(Cost)	GWP	E(Profit)	2.5%Prof	97.5%Prof
0	0	0	1	0	0	0	1	0	0	0.84	16014037	386867	5154	2428694	1670435	3186953
0	0	1	1	1	0	1	0	1	0	0.84	9783443	764125	4097	8729892	7232208	10227576
0	1	0	0	1	1	1	1	0	1	0.72	6731663	1244803	3563	11802086	9362273	14241900
0	1	0	1	1	1	1	0	1	0	0.96	17140995	383822	5331	1395712	643422	2148003
0	1	0	1	1	1	1	1	0	0	0.96	17143030	383822	5331	1395712	643422	2148003
0	1	0	1	1	1	1	1	0	1	0.96	17173611	383822	5331	1395712	643422	2148003
0	1	0	1	1	1	1	1	1	0	0.96	17169698	383822	5331	1395712	643422	2148003
0	1	0	1	1	1	1	1	1	1	0.96	17200279	383822	5331	1395712	643422	2148003
0	1	1	0	0	1	1	0	0	1	0.96	17103463	383822	5331	1395712	643422	2148003
0	1	1	1	1	0	0	1	1	0	0.96	17140072	383822	5331	1395712	643422	2148003
0	1	1	1	1	1	1	1	1	1	0.09	1752277	1830570	2776	16863649	13275733	20451566
0	1	1	1	1	1	1	1	1	1	0.18	1752277	1830570	2776	16863649	13275733	20451566
0	1	1	1	1	1	1	1	1	1	0.6	1752277	1830570	2776	16863649	13275733	20451566
1	0	0	0	0	1	0	0	1	1	0.96	17087338	383822	5331	1395712	643422	2148003
1	0	0	1	0	0	1	1	1	1	0.84	6673861	1136314	3719	11877704	9650528	14104879
1	0	0	1	1	0	0	1	1	0	0.96	17124177	383822	5331	1395712	643422	2148003
1	0	0	1	1	0	0	1	1	1	0.72	4693527	1375514	3445	13856944	11160936	16552952
1	0	0	1	1	0	0	1	1	1	0.96	17154758	383822	5331	1395712	643422	2148003
1	0	0	1	1	0	1	0	1	1	0.96	17152954	383822	5331	1395712	643422	2148003
1	0	0	1	1	0	1	1	1	0	0.96	17151076	383822	5331	1395712	643422	2148003
1	0	0	1	1	1	1	1	1	1	0.09	2253042	2804519	3022	16346990	10850132	21843847
1	0	1	0	1	0	1	1	0	0	0.72	8856540	1060169	3898	9647941	7570009	11725873
1	0	1	0	1	1	1	1	0	0	0.72	7526793	1145450	3631	11000350	8755268	13245431
1	0	1	1	0	1	0	1	0	0	0.84	12633255	572097	4564	5876757	4755448	6998067
1	0	1	1	0	1	1	0	0	0	0.84	13376049	541068	4694	5132159	4071665	6192654
1	0	1	1	0	1	1	1	0	0	0.96	17141199	383822	5331	1395712	643422	2148003
1	1	0	1	1	1	1	1	0	0	0.96	17167715	383822	5331	1395712	643422	2148003
1	1	1	0	0	1	1	1	1	1	0.09	2269366	2175852	2970	16309866	12045196	20574536
1	1	1	0	1	1	1	1	1	1	0.09	2000901	2199489	2776	16604136	12293136	20915135
1	1	1	0	1	1	1	1	1	1	0.18	2000901	2199489	2776	16604136	12293136	20915135
1	1	1	0	1	1	1	1	1	1	0.6	2000901	2199489	2776	16604136	12293136	20915135

1	1	1	1	0	1	1	1	0	1	0.96	17192425	383822	5331	1395712	643422	2148003
1	1	1	1	0	1	1	1	1	0	0.96	17188513	383822	5331	1395712	643422	2148003
1	1	1	1	0	1	1	1	1	1	0.09	1731699	1635885	2905	16883107	13676772	20089442
1	1	1	1	0	1	1	1	1	1	0.18	1731699	1635885	2905	16883107	13676772	20089442
1	1	1	1	1	0	1	1	1	0	0.96	17191656	383822	5331	1395712	643422	2148003
1	1	1	1	1	1	1	0	1	0	0.96	17185615	383822	5331	1395712	643422	2148003
1	1	1	1	1	1	1	1	1	0	0.6	1704373	1660086	2750	16905658	13651890	20159425
1	1	1	1	1	1	1	1	1	1	0.09	1463510	1661351	2712	17177101	13920853	20433350
1	1	1	1	1	1	1	1	1	1	0.18	1463510	1661351	2712	17177101	13920853	20433350
1	1	1	1	1	1	1	1	1	1	0.6	1463510	1661351	2712	17177101	13920853	20433350

4.2.4. Variance-greedy route selection

In this case, the algorithm was run with the objective of minimizing variance during the route selection process. The algorithm was run for 100 generations with 100 solutions in each instance. Surprisingly, the algorithm found only one non-dominated solution in this instance, listed below on table 4.27. This was run in on MATLAB 2018a, running on a Dell Inspiron 7779 with an intel i7 7500U CPU and 16GB of RAM, with a total run time of 200 seconds.

Table 4.27: Non-dominated solutions for algorithm with variance-greedy route selection

Selected Hubs										Gamma	E(Cost)	Var(Cost)	GWP	E(Profit)	2.5%Prof	97.5%Prof
0	0	0	1	1	1	1	1	0	1	0.72	15505879	361801	5008	3042798	2333668	3751929

4.2.5. Environmental Impact-greedy route selection

In this case, the algorithm was run with the same parameters of 100 generations of 100 individuals each. However, the route selection parameter was set to prioritize the minimization of the expected Global Warming Potential of the transportation network. We must keep in mind that unlike during the cost-greedy run, there is no collaborative cost-saving measure to be considered due to the different formulation of the objective function. The results are summarized below in figure 4.3, which shows less correlated objective functions, compared with the cost-greedy algorithm results. Some correlation can still be observed between objectives 1 and 4 (expected cost and expected profit), 2 and 6 (cost variance and maximum profit) and 3 and 4 (Global Warming

Potential and minimum profit). The solution marked in red is the solution closest to the “ideal” point by Euclidean distance.

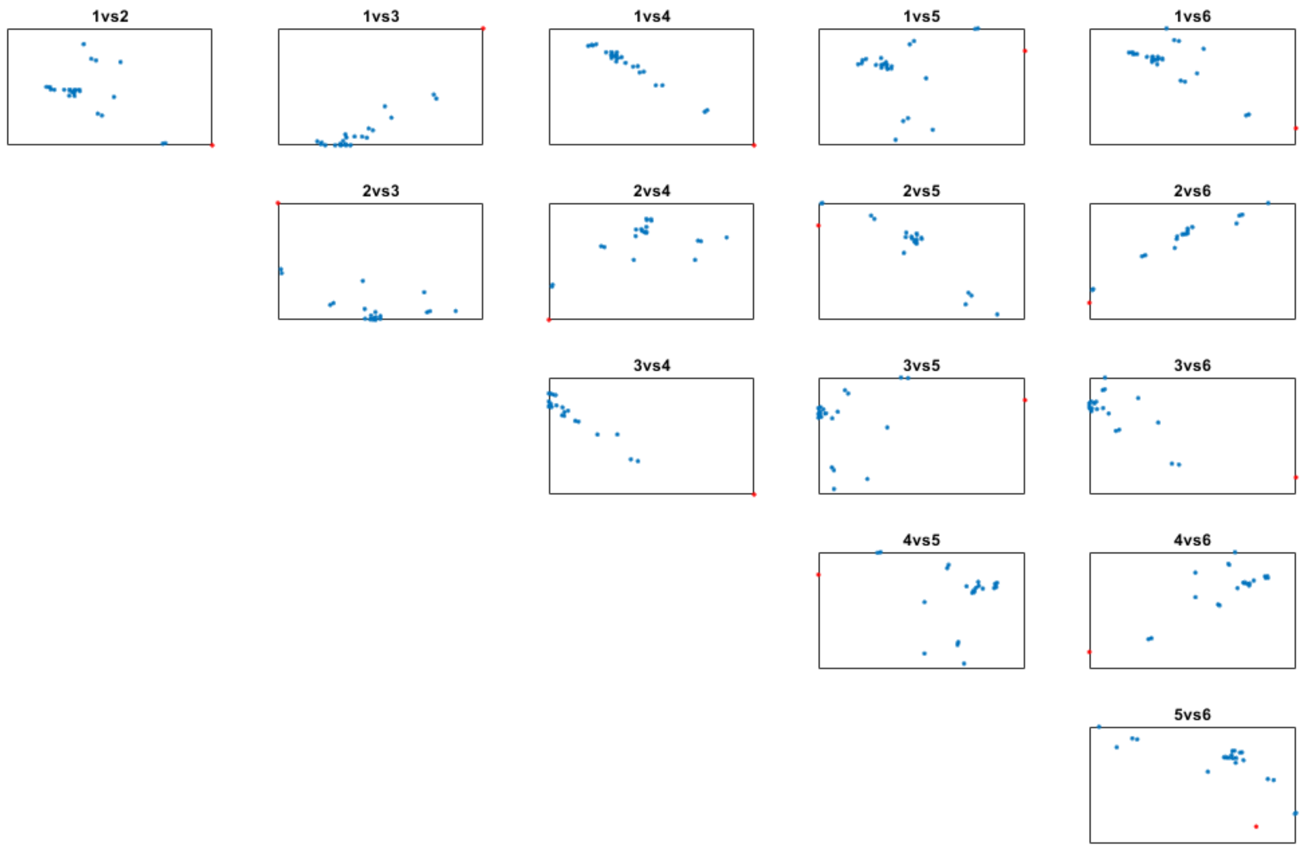

Figure 4.2: Pareto-optimal fronts for algorithm with GWP-greedy route selection

Again, all solutions are listed below in table 4.28. This run had a total yield of 37 non-dominated solutions, with the number of selected hubs ranging from 4 to 10 and a Gamma value ranging between 0.18 and 0.84. In terms of the objective functions, the minimum values were {5,715,399; 1,331,731; 2,688; 11,865,466; 8,072,697; 14,475,659} and the maximums were {6,623,928; 2,343,266; 2,901; 12,846,147; 9,554,462; 17,258,298} for expected cost, cost variance, global warming potential, expected profit, 2.5 percentile for profit and 97.5 profit percentile, respectively. Again, this was run in on MATLAB 2018a, running on a Dell Inspiron 7779 with an intel i7 7500U CPU and 16GB of RAM, with a total run time of 240 seconds.

Table 4.28: Non-dominated solutions for algorithm with GWP-greedy route selection

Selected Hubs										Gamma	E(Cost)	Var(Cost)	GWP	E(Profit)	2.5%Prof	97.5%Prof
0	0	0	1	1	1	1	0	0	0	0.18	6623928	1331731	2901	11865466	9255273	14475659
0	0	0	1	1	1	1	0	0	0	0.6	6623928	1331731	2901	11865466	9255273	14475659
0	1	0	1	1	1	0	1	1	1	0.72	6121576	2164017	2739	12447517	8206043	16688991
0	1	0	1	1	1	0	1	1	1	0.84	6121576	2164017	2739	12447517	8206043	16688991
0	1	1	1	1	1	0	0	0	1	0.72	6085595	1814749	2759	12448060	8891152	16004969
0	1	1	1	1	1	1	0	1	0	0.84	5842772	1879241	2691	12713870	9030559	16397181
0	1	1	1	1	1	1	1	1	1	0.18	5896334	1887084	2688	12719592	9020907	16418277
0	1	1	1	1	1	1	1	1	1	0.36	5896334	1887084	2688	12719592	9020907	16418277
1	0	0	1	1	1	1	1	0	0	0.72	6353118	1348315	2781	12189664	9546966	14832362
1	0	0	1	1	1	1	1	0	1	0.84	6367083	1352969	2773	12206280	9554462	14858099
1	0	0	1	1	1	1	1	1	0	0.84	5996787	1647884	2719	12572663	9342811	15802515
1	0	0	1	1	1	1	1	1	1	0.84	6019725	1629399	2715	12580307	9386685	15773928
1	0	1	1	1	1	0	1	1	1	0.72	5871306	1873498	2689	12721761	9049705	16393818
1	0	1	1	1	1	1	0	1	1	0.18	5870679	1874175	2689	12720584	9047202	16393966
1	0	1	1	1	1	1	0	1	1	0.72	5870679	1874175	2689	12720584	9047202	16393966
1	0	1	1	1	1	1	1	1	0	0.84	5870637	1875482	2690	12718748	9042804	16394692
1	0	1	1	1	1	1	1	1	1	0.84	5898205	1873498	2689	12721761	9049705	16393818
1	1	0	0	1	1	1	1	1	1	0.84	5919605	2343266	2704	12665497	8072697	17258298
1	1	0	1	1	1	1	1	1	0	0.84	5961178	2196934	2704	12628918	8322927	16934908
1	1	0	1	1	1	1	1	1	1	0.84	5987129	2180432	2702	12633548	8359901	16907195
1	1	1	0	1	1	1	0	1	0	0.84	5715399	1915184	2695	12830354	9076594	16584114
1	1	1	0	1	1	1	0	1	1	0.84	5734685	1913532	2690	12841650	9091126	16592173
1	1	1	0	1	1	1	1	0	0	0.84	5868946	1825447	2708	12678842	9100966	16256717
1	1	1	0	1	1	1	1	1	0	0.84	5739605	1888736	2693	12834851	9132930	16536773
1	1	1	0	1	1	1	1	1	1	0.84	5758890	1887084	2688	12846147	9147462	16544832
1	1	1	1	1	1	0	0	1	1	0.84	5874532	1886458	2702	12710477	9013020	16407935
1	1	1	1	1	1	0	1	0	1	0.84	5856642	1861902	2696	12730402	9081074	16379730
1	1	1	1	1	1	0	1	1	1	0.84	5845366	1887084	2688	12768346	9069662	16467031
1	1	1	1	1	1	1	0	1	0	0.84	5812981	1887084	2688	12768346	9069662	16467031
1	1	1	1	1	1	1	0	1	1	0.72	5843562	1887084	2688	12768346	9069662	16467031
1	1	1	1	1	1	1	0	1	1	0.84	5843562	1887084	2688	12768346	9069662	16467031
1	1	1	1	1	1	1	1	0	0	0.18	5840553	1828060	2690	12742809	9159811	16325806
1	1	1	1	1	1	1	1	0	0	0.84	5840553	1828060	2690	12742809	9159811	16325806
1	1	1	1	1	1	1	1	0	1	0.84	5868084	1857558	2689	12745859	9105045	16386674
1	1	1	1	1	1	1	1	1	0	0.18	5841684	1887084	2688	12768346	9069662	16467031
1	1	1	1	1	1	1	1	1	0	0.84	5841684	1887084	2688	12768346	9069662	16467031
1	1	1	1	1	1	1	1	1	1	0.84	5872265	1887084	2688	12768346	9069662	16467031

4.2.6. Profit-greedy route selection

The algorithm was now run with the criterion of maximizing profit during its route selection process. The main MOGA parameters remain the same as in previous runs. The results are summarized below in figure 4.4. Some correlations can be observed between objectives 1 and 4 (expected cost and expected profit), while correlations between the other objectives are harder to see. There may be some correlation between expected cost and GWP (1 vs 3) and expected cost vs 2.5 percentile profit, but any correlations between remaining objectives seem dubious. The solution marked in red is the solution closest to the "ideal" point by Euclidean distance.

Table 4.29 displays all 23 non-dominated solutions obtained by the program with this route selection criterion. In this case, the number of selected hubs ranging from 0 to 10 and a Gamma value ranging between 0.09 and 0.84. In terms of the objective functions, the minimum values were {1,463,510; 383,822; 2,712; 1,395,712; 643,422; 2,148,003} and the maximums were {16,982,742; 4,434,741; 5,331; 17,177,101; 13,920,853; 22,841,551} for expected cost, cost variance, global warming potential, expected profit, 2.5 percentile for profit and 97.5 profit percentile, respectively. Again, this was run in on MATLAB 2018a, running on a Dell Inspiron 7779 with an intel i7 7500U CPU and 16GB of RAM, with a total run time of 357 seconds.

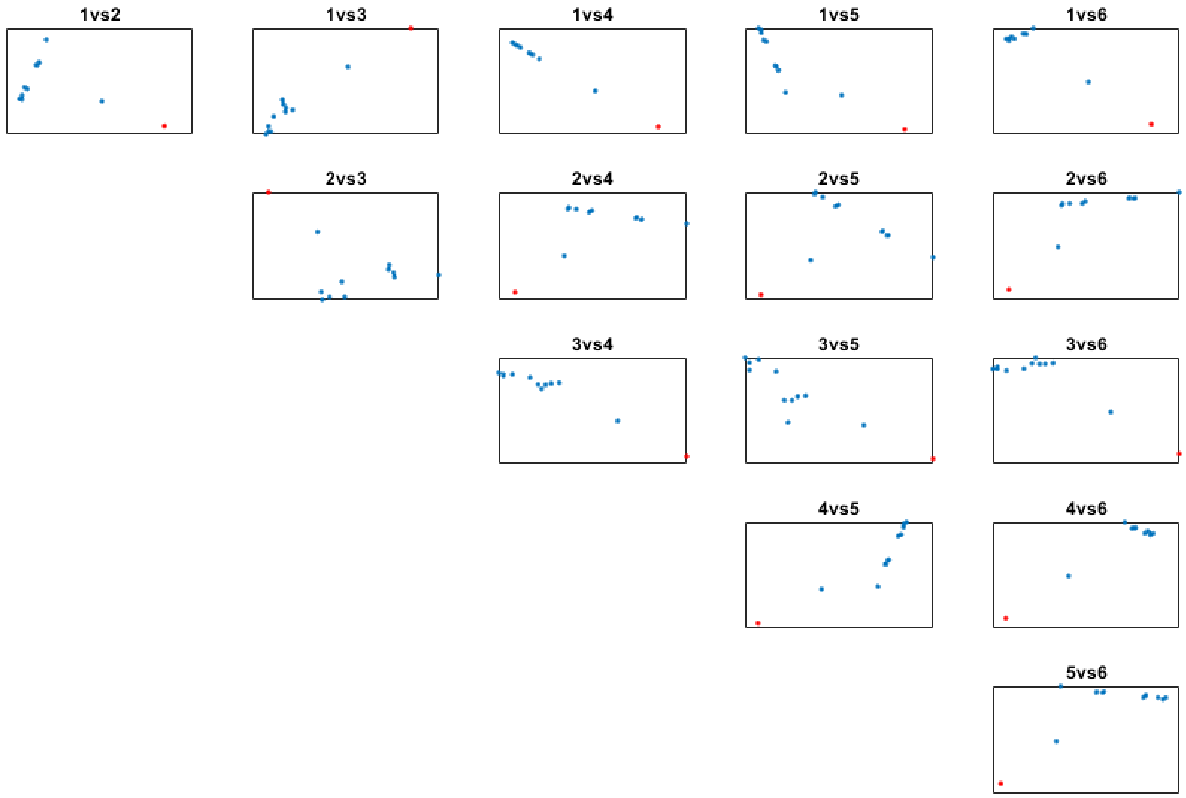

Figure 4.3: Pareto-optimal front for algorithm with profit-greedy route selection

Table 4.29: Non-dominated solutions for algorithm with profit-greedy route selection

Selected Hubs										Gamma	E(Cost)	Var(Cost)	GWP	E(Profit)	2.5%Prof	97.5%Prof
0	0	0	0	0	0	0	0	0	0	0.72	16982742	383822	5331	1395712	643422	2148003
0	0	0	0	0	0	0	0	0	0	0.84	16982742	383822	5331	1395712	643422	2148003
0	0	0	0	0	0	0	0	0	1	0.72	10267249	1549479	4368	8141786	5104807	11178765
0	0	0	0	0	0	0	0	0	1	0.84	10267249	1549479	4368	8141786	5104807	11178765
0	0	0	0	1	1	0	1	1	1	0.84	3594256	3356673	3368	14918617	8339537	21497697
0	1	1	1	1	1	1	1	0	1	0.72	2297843	2131783	3147	16291415	12113121	20469708
0	1	1	1	1	1	1	1	1	1	0.48	1752277	1830570	2776	16863649	13275733	20451566
0	1	1	1	1	1	1	1	1	1	0.72	1752277	1830570	2776	16863649	13275733	20451566
0	1	1	1	1	1	1	1	1	1	0.84	1752277	1830570	2776	16863649	13275733	20451566
1	0	0	0	1	1	0	0	1	0	0.09	4328816	4434741	3311	14149459	5457367	22841551
1	0	0	0	1	1	0	0	1	0	0.84	4328816	4434741	3311	14149459	5457367	22841551
1	0	0	1	1	0	1	0	1	0	0.36	3338593	3233669	3447	15179492	8841502	21517482
1	0	0	1	1	1	0	0	1	1	0.09	3558932	3378796	3261	14985498	8363058	21607937
1	0	0	1	1	1	0	1	0	1	0.48	3236387	3249971	3556	15310078	8940134	21680021
1	1	1	0	1	1	1	1	1	1	0.48	2000901	2199489	2776	16604136	12293136	20915135
1	1	1	0	1	1	1	1	1	1	0.72	2000901	2199489	2776	16604136	12293136	20915135
1	1	1	0	1	1	1	1	1	1	0.84	2000901	2199489	2776	16604136	12293136	20915135
1	1	1	1	0	1	1	1	1	1	0.48	1731699	1635885	2905	16883107	13676772	20089442
1	1	1	1	0	1	1	1	1	1	0.72	1731699	1635885	2905	16883107	13676772	20089442
1	1	1	1	0	1	1	1	1	1	0.84	1731699	1635885	2905	16883107	13676772	20089442

1	1	1	1	1	1	1	1	1	1	0.48	1463510	1661351	2712	17177101	13920853	20433350
1	1	1	1	1	1	1	1	1	1	0.72	1463510	1661351	2712	17177101	13920853	20433350
1	1	1	1	1	1	1	1	1	1	0.84	1463510	1661351	2712	17177101	13920853	20433350

4.2.6. Combined results

After combining all previous solutions and eliminating any newly dominated solutions, we obtained the following pareto-optimal set of solutions. In figure 4.4, we can see all 31 non-dominated solutions displayed. Solutions from different sets have been marked with different colors: blue dots are solutions pulled from the cost-greedy route selection run, magenta from the variance-greedy run, green dots from the Global Warming Potential-greedy run, and red from the profit-greedy run. These runs have 11, 1, 19, and 1 non-dominated solutions respectively in this case, meaning a loss of 31, 0, 18, and 22 solutions to dominance, respectively.

With new aggregates, direct, positive correlations can easily be observed for the following pairs of objectives: 1 and 3, 4 and 5, 4 and 6, and 5 and 6. Direct, negative correlations can be observed for the following pairs of objectives: 1 and 4, 1 and 5, 1 and 6, 3 and 4, 3 and 5, and 3 and 6. On the other hand, the following pairs of objectives can be said to be inversely related: 1 and 2, 2 and 3, 2 and 4, 2 and 5, and 2 and 6. It can be said, in a way, that variance is the biggest factor influencing the formation of the Pareto front, as it is the objective that most causes inverse relations against the remaining objectives. Table 4.30 shows all solutions that were found to be non-dominated after joining all previous sets of solutions, with the set they were originally part of in the last column.

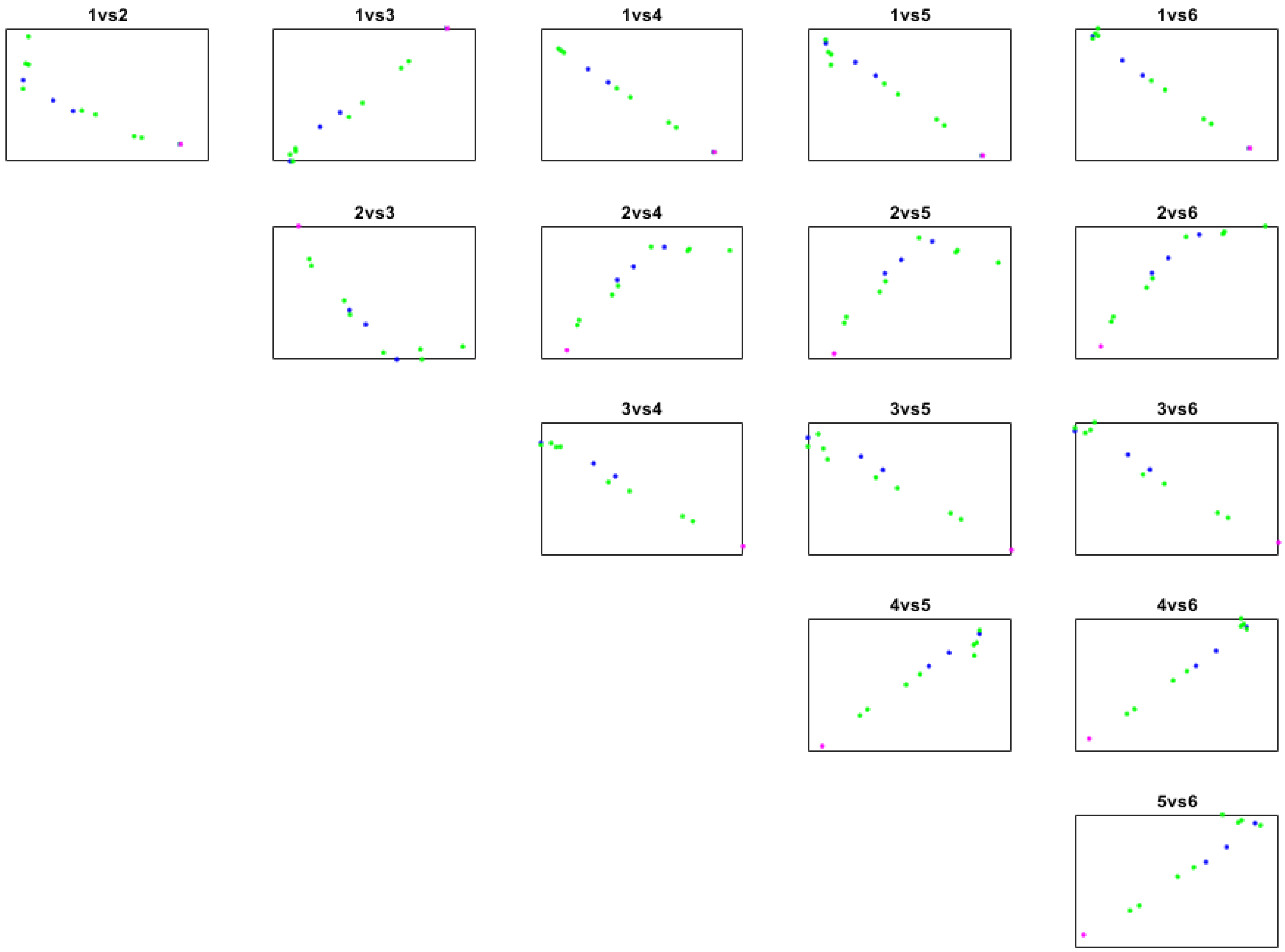

Figure 4.4: Pareto-optimal front for all runs of the algorithm

Table 4.30: Non-dominated solutions from all runs

Selected Hubs										Gamma	E(Cost)	Var(Cost)	GWP	E(Profit)	2.5%Prof	97.5%Prof	Set
0	0	1	1	1	0	1	0	1	0	0.84	9783443	764124.6	4097	8729892	7232208	10227576	cost
0	1	0	0	1	1	1	1	0	1	0.72	6731663	1244803	3563	11802086	9362273	14241900	cost
1	0	0	1	0	0	1	1	1	1	0.84	6673861	1136314	3719	11877704	9650528	14104879	cost
1	0	0	1	1	0	0	1	1	1	0.72	4693527	1375514	3445	13856944	11160936	16552952	cost
1	0	0	1	1	1	1	1	1	1	0.09	2253042	2804519	3022	16346990	10850132	21843847	cost
1	0	1	0	1	0	1	1	0	0	0.72	8856540	1060169	3898	9647941	7570009	11725873	cost
1	0	1	0	1	1	1	1	0	0	0.72	7526793	1145450	3631	11000350	8755268	13245431	cost
1	0	1	1	0	1	0	1	0	0	0.84	12633255	572096.7	4564	5876757	4755448	6998067	cost
1	0	1	1	0	1	1	0	0	0	0.84	13376049	541068.5	4694	5132159	4071665	6192654	cost
1	1	1	0	0	1	1	1	1	1	0.09	2269366	2175852	2970	16309866	12045196	20574536	cost

1	1	1	1	1	1	1	1	1	0	0.6	1704373	1660086	2750	16905658	13651890	20159425	cost
0	1	1	1	1	1	1	0	1	0	0.84	5842772	1879241	2691	12713870	9030559	16397181	GWP
1	0	0	1	1	1	1	1	0	0	0.72	6353118	1348315	2781	12189664	9546966	14832362	GWP
1	0	0	1	1	1	1	1	0	1	0.84	6367083	1352969	2773	12206280	9554462	14858099	GWP
1	0	0	1	1	1	1	1	1	0	0.84	5996787	1647884	2719	12572663	9342811	15802515	GWP
1	0	0	1	1	1	1	1	1	1	0.84	6019725	1629399	2715	12580307	9386685	15773928	GWP
1	0	1	1	1	1	0	1	1	1	0.72	5871306	1873498	2689	12721761	9049705	16393818	GWP
1	0	1	1	1	1	1	1	1	0	0.84	5870637	1875482	2690	12718748	9042804	16394692	GWP
1	1	0	0	1	1	1	1	1	1	0.84	5919605	2343266	2704	12665497	8072697	17258298	GWP
1	1	0	1	1	1	1	1	1	0	0.84	5961178	2196934	2704	12628918	8322927	16934908	GWP
1	1	0	1	1	1	1	1	1	1	0.84	5987129	2180432	2702	12633548	8359901	16907195	GWP
1	1	1	0	1	1	1	0	1	0	0.84	5715399	1915184	2695	12830354	9076594	16584114	GWP
1	1	1	0	1	1	1	0	1	1	0.84	5734685	1913532	2690	12841650	9091126	16592173	GWP
1	1	1	0	1	1	1	1	0	0	0.84	5868946	1825447	2708	1.3E+07	9100966	1.6E+07	GWP
1	1	1	0	1	1	1	1	1	0	0.84	5739605	1888736	2693	1.3E+07	9132930	1.7E+07	GWP
1	1	1	0	1	1	1	1	1	1	0.84	5758890	1887084	2688	1.3E+07	9147462	1.7E+07	GWP
1	1	1	1	1	1	0	0	1	1	0.84	5874532	1886458	2702	1.3E+07	9013020	1.6E+07	GWP
1	1	1	1	1	1	0	1	0	1	0.84	5856642	1861902	2696	1.3E+07	9081074	1.6E+07	GWP
1	1	1	1	1	1	1	1	0	1	0.84	5868084	1857558	2689	1.3E+07	9105045	1.6E+07	GWP
0	1	1	1	1	1	1	1	0	1	0.72	2297843	2131783	3147	1.6E+07	1.2E+07	2E+07	profit
0	0	0	1	1	1	1	1	0	1	0.72	1.6E+07	361801	5008	3042799	2333668	3751929	var

The non-dominated solutions from the aggregated set represent 31 different strategies in which the hypothetical pair of LTL carriers can collaborate in order to optimize the proposed set of objectives, which can be chosen in order to conform to new sets of constraints on hub allocation, collaboration level (Gamma), or expected objective results.

4.3. Post-Pareto Analysis

To finalize our study, we will try to perform a post-pareto operation to further reduce the set of non-dominated solutions obtained by the algorithm. This could be done in a number of ways, such as a simple distance-based fitness metric or some more complicated methods such as the Nash-dominance concept.

For the sake of simplicity and due to the relatively small number of solutions, we will analyze the solutions via a simple distance metric to find which is closest to the "ideal" point of (0,0,0,1,1,1), that is, zero cost, variance, and environmental impacts, with maximum profit. The ideal maximum is set to one because the data is to be normalized from 0 to 1 in each objective in order to be able to properly perform this type of analysis. The results of this analysis are summarized below on table 4.31.

Table 4.31: Non-dominated solutions ranked by Table 4.31: Non-dominated solutions ranked by Euclidean distance to ideal point

E(Cost)	Var(Cost)	GWP	2.5%Prof	E(Prof)	97.5%Prof	#	d
1704373	1660086	2750	16905658	13651890	20159425	11	0.735
4693527	1375514	3445	13856944	11160936	16552952	4	0.883
2269366	2175852	2970	16309866	12045196	20574536	10	0.887
2297843	2131783	3147	16000000	12000000	20000000	30	0.896
6353118	1348315	2781	12189664	9546966	14832362	13	1.002
6367083	1352969	2773	12206280	9554462	14858099	14	1.002
6019725	1629399	2715	12580307	9386685	15773928	16	1.024
5996787	1647884	2719	12572663	9342811	15802515	15	1.029
2253042	2804519	3022	16346990	10850132	21843847	5	1.042
5758890	1887084	2688	13000000	9147462	17000000	26	1.047
5739605	1888736	2693	13000000	9132930	17000000	25	1.047
5868946	1825447	2708	13000000	9100966	16000000	24	1.057
5734685	1913532	2690	12841650	9091126	16592173	23	1.063
5868084	1857558	2689	13000000	9105045	16000000	29	1.063
5715399	1915184	2695	12830354	9076594	16584114	22	1.064
5856642	1861902	2696	13000000	9081074	16000000	28	1.065
5871306	1873498	2689	12721761	9049705	16393818	17	1.068
5870637	1875482	2690	12718748	9042804	16394692	18	1.069
5842772	1879241	2691	12713870	9030559	16397181	12	1.069
5874532	1886458	2702	13000000	9013020	16000000	27	1.073
6673861	1136314	3719	11877704	9650528	14104879	3	1.098
6731663	1244803	3563	11802086	9362273	14241900	2	1.103
5987129	2180432	2702	12633548	8359901	16907195	21	1.154
5961178	2196934	2704	12628918	8322927	16934908	20	1.157
5919605	2343266	2704	12665497	8072697	17258298	19	1.188
7526793	1145450	3631	11000350	8755268	13245431	7	1.205

8856540	1060169	3898	9647941	7570009	11725873	6	1.448
9783443	764124.6	4097	8729892	7232208	10227576	1	1.590
12633255	572096.7	4564	5876757	4755448	6998067	8	2.217
13376049	541068.5	4694	5132159	4071665	6192654	9	2.401
16000000	361801	5008	3042799	2333668	3751929	31	3.000

The original position of the objectives in table 4.30 was stored in the 7th column of table 4.31. The closest point to the ideal solution was found to be original solution number 11. This solution chose all nodes but 10 to convert into hubs. It had a cost-greedy route selection objective and a collaboration level of 60%. However, it is still up to a decision-maker to pick the best solution, as there are still other parameters which may not be ideal in this solution, such as the number of selected hubs, as it is perhaps not unfeasible in terms of resources, but the facilities may not have the capacity to process consolidation and de-consolidation of shipments at nine different facilities simultaneously for such a small network.

Chapter 5: Conclusions and Future Research

The main objective of this thesis was to analyze a Centralized Carrier Collaboration and Multi-hub location problem to show the possibility of reducing costs and environmental impacts for a set of small-to-medium less-than-truckload carriers who are considering collaboration. The problem has been studied and analyzed using a newly designed multi-objective evolutionary algorithm which uses Universal Generating Functions to model the stochastic properties of the data. While this data is purely theoretical, the tool can be implemented for the analysis of a practical scenario in which two or more real companies are considering collaboration. As the case was studied as a multiple objective optimization front, a Pareto-optimal set was created and a tentative solution was chosen via a simple distance-based fitness metric.

In order to demonstrate the effectiveness of the algorithm over hub-and-spoke networks, the CCCMLP was introduces in order to provide a planning framework to analyze the benefits of the network that the algorithm creates. The CCCMLP incorporates the concept of hybrid consolidation hubs based on existing facilities to reduce construction costs, while leveraging current service locations of existing LTL carriers to synergize with new opportunities in e-commerce. The results showed an unsurprising link between the reduction of expected costs and a positive shift in the resulting profit probability distribution. Also, an inverse relation was found between expected costs and corresponding variances. This can greatly be attributed to the network configuration and the uncertainties involved in a collaboration. More interesting, however, is the relation found between the reduction of costs (and positive shift in profits) with a reduction in expected environmental impacts. This is most likely due to the reduction in total travel distances over a collaborative hub-and-spoke network, where shipments between the two carriers can be combined into a single shipment between a number of hubs.

This thesis still leaves work to do in analyzing the implementation of the CCCMLP or other hub-and-spoke network optimization problems onto different scenarios, such as multi-modal transport networks, with or without collaborations between a number of different vendors, as well

as what collaboration could mean between more than two carriers and how it could be measured. Another possible expansion could be into a situation where we must account for a limited number of trucks that need to deliver all shipments, and account for dead-heading when necessary, as well as truck maintenance costs. Other aspects can also be measured such as social impact, which could be measured by expected number of jobs to be created and changes to local economies, as well as public opinion on the transformation of existing infrastructure to accommodate larger facilities for consolidation. Lastly, more thorough post-Pareto optimality studies could be performed given less correlated objectives and a larger non-dominated optimal set.

References

Abdinnour-Helm, S. (1998). A hybrid heuristic for the uncapacitated hub location problem. *European Journal of Operational Research*, 489-499.

Agarwal, R., & Ergun, Ö. (2008). Ship scheduling and network design for cargo routing in liner shipping. *Transportation Science*, 175-196.

Aykin, N. A., Cunningham, J. P., & Rotella, J. P. (1994). Designing Operations Systems Interfaces that are Easy to Use. *AT&T Technical Journal*.

Aykin, T. (1995). Networking Policies for Hub and Spoke Systems with Applications to the Air Transportation System. *Transportation Science*, 201-221.

Ballot, E., & Fontane, F. (2010). Reducing Transportation CO2 Emissions Through Pooling of Supply Networks: Perspectives from a Case Study in French Retail Chains. *Production Planning & Control*, 640-650.

Barlow, R. E., & Wu, A. S. (1978). Coherent Systems with Multi-state Components. *Mathematics of Operations Research*, 275-281.

Brockhoff, D., & Zitzler, E. (2006). Are all objectives necessary? On dimensionality reduction in evolutionary multiobjective optimization. *Parallel Problem Solving from Nature-PPSN IX*, 533-542.

Campbell, J. (1994). Integer Programming formulations of discrete hub location problems. *European Journal of Operations Research*, 387-405.

Campbell, J. F., & O'Kelly, M. E. (2012). Twenty-Five Years of Hub Location Research. *Transportation Science*, 153-169.

Campbell, J. F., Ernst, A. T., & Krishnamoorthy, M. (2005). Hub Arc Location Problems: Part I-Introduction and Results. *Management Science*, 1540-1555.

Campbell, J., Miranda, J. G., Camargo, R., & O'Kelly, M. (2015). Hub Location and Network Design with Fixed and Variable Costs. *Proceedings of the 48th Annual Hawaii International Conference on System Sciences, Computer Society Press*, 1059-1067.

Caramia, M., & Dell'Olmo, P. (2008). *Multi-Objective Optimization.* London: Springer.

Carlsson, J. G., & Jia, F. (2013). Euclidean Hub-and-Spoke Networks. *Operational Research*, 1360-1382.

Carrillo, V. M., & Taboada, H. (2012). A post-pareto approach for multi-objective decision making using a non-uniform weight generator method. *Procedia Computer Science*, 112-121.

Chankong, V., & Haimes, Y. Y. (1983). *Multiobjective Decision Making: Theory and Methodology.* New York: Elsevier Science Publishing.

Charnes, A., Cooper, W. W., & Ferguson, R. O. (1955). Optimal Estimation of Executive Compensation by Linear Programming. *Management Science*, 138-151.

Cliff, R., & Wright, L. (2000). Relationships between environmental impacts and added calue along the supply chain. *Technological forecasting and social change*, 281-295.

Contreras, I., Cordeau, J., & Laporte, G. (2011). The Dynamic Uncapacitated Hub Location Problem. *Transportation Science*, 18-32.

Contreras, I., Cordeau, J.-F., & Laporte, G. (2011). Benders Decomposition for Large-Scale Uncapacitated Hub Location. *Operations Research*, 1477-1490.

Corne, D., Jerram, N., Knowles, J., & Oates, M. (2011). PESA-II: Region-based Selection in Evolutionary Multiobjective Optimization. *Proceedings of the Genetic and Evolutionary Computatoin Conference*, 283-290.

Corne, D., Knowles, J., & Oates, M. (2000). The Pareto Envelope-Based Selection Algorithm for Multiobjective Oprimization. *Proceedings of the 6th Parallel Problem Solving from Nature*.

de Camargo, R. S., de Miranda Jr., G., Ferreira, R., & Luna, H. (2009). Multiple Allocation hub-and-spoke network design under hub congestion. *Computers and Operations Research*, 3097-3106.

de Mello, P. F., & Frayret, J. M. (2014). *Agent-Based Approach of Sustainability Assessment of Resource Sharing in Freight Transportation.* Montreal: CIRRELT, Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport.

de Miranda Jr., G., Saraiva de Camargo, R., Pinto, L. R., Vieira Conceição, S., & Martins Ferreira, R. P. (2011). Hub location under hub congestion and demand uncertainty: the Brazilian case study . *Pesquisa Operacional*, 319-349.

Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. A. (2002). A fasst and elitist multi-objective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 182-197.

Delgado, I. (2016, May). Minimizing Environmental Impacts for Hub and Spoke distribution network problems through the use of multi-objective evolutionary algorithms. The University of Texas at El Paso.

Delgado, I., Fernandez, J., Valles, C., & Taboada, H. A. (2015). Multiple Objective Optimization of a Biomass to Bio-Refinery Logistics System. *IIE Annual Conference Proceedings*.

Ebery, J., Krishnamoorthy, M., Ernst, A., & Boland, N. (2000). The capacitated multiple alocation hub location problem: Formulation and algorithms. *European Journal of Operational Research*, 614-631.

Elhedhli, S., & Wu, H. (2010). A lagrangean heuristic for Hub and Spoke System Design with capacity selection and congestion. *INFORMS Journal on Computing*, 282-296.

EPA. (2018). *1990-2014 Fast Facts.* Washington, D.C.: United States Environmental Protection Agency.

EPA. (2018). *Fast Facts: U.S. Transportation Sector GHG Emissions.* Washington, D.C.: United States Environmental Protection Agency.

Ernst, A., & Krishnamoorthy, M. (1999). Solution algorithms for the capacitated single allocation hub location problem. *Annals of Operations Research*, 262-269.

Esper, T., & Williams, L. (2003). The value of collaborative transportation management (CTM): its relationship to CPFR and information technology. *Transportation Journal, ABI/INFORM*, 55-65.

Espiritu, J. F., Taboada, H. A., Fernandez, J., Castillo, E., Valles, C., & Delgado, I. (2015). Global Warming Potential and Cost Minimization for the Centralized Carrier Collaboration and Multi-Hub Location Problem. *Proceedings of the 2015 International Conference on Operations Excellence and Service Engineering*.

Fernandez, J. V. (2017). *Development of post-pareto optimality methods for multi-objective optimization.* El Paso: The University of Texas at El Paso.

Figliozzi, M. (2006). Analysis and Evaluation of Incentive Compatible Dynamic Mechanism for Carrier Collaboration. *Transportation Research Record: Journal of the Transportation Research Board*, 14-40.

Fishburn, P. C. (1974). Exceptional Paper--Lexicographic Orders, Utilities and Decision Rules: A Survey. *Management Science*, 1442-1471.

Gelareh, S., & Nickel, S. (2011). Hub Location Problems in Transportation Networks. *Transportation Research*, 1092-1111.

Gomes, B. N., Martins, A. X., de Camargo, R. S., & Ramirez, J. A. (2013). An Efficient Genetic Algorithm for the Design of Hub-and-Spoke Networks. *IEEE Communication Letters*, 479-483.

Gong, D., Ji, X., Sun, J., & Sun, X. (2014). Interactive evolutionary algorithms with decisionmaker's preferences for solving interval multi-objective optimization problems. *Neurocomputing*, 241-251.

Hagelaar, G. J., & van der Vorst, J. G. (2001). Environmental Supply Chain Manageent: Using Life Cycle Assessment to Structure Supply Chains. *The International Food and Agrobusiness Management Review*, 399-712.

Hernandez, S., & Peeta, S. (2010). Static single carrier collaboration problem for less-than-truckload carriers. *Proceedings of the 89th Annual Meeting of the Transportation Research Board*.

Hernandez, S., Peeta, S., & Kalafatas, G. (2011, November). A less-than-truckload carrier collaboration planning problem under dynamic capacities. *Transportation Research Part E: Logistics and Transportation Review, 47*(6), 933-946.

Hernandez, S., Unnikrishnan, A., & Awale, S. (2011). A less-than-truckload centralized carrier collaboration multi-hub location problem: a hybrid hub-and-spoke network. *Transportation Research*, 20-28.

Hernandez, S., Unnikrishnan, A., & Awale, S. S. (2012). Centralized Carrier Collaboration Multihub Location Problem for Less-Than-Truckload Industry : Hybrid Hub-and-Spoke Network. *Transportation Research Record*, 20-28.

Horn, J., Nafpliotis, N., & Goldberg, D. E. (1994). A Niched Pareto Genetic Algorithm for Multiobjective Optimization. *Proceedings of the first IEEE conference on evolutionary computation, IEEE world congress on computational intelligence*, 82-87.

Ishfaq, R., & Sox, C. R. (2012). Design of intermodal logistics networks with hub delays. *European Journal of Operations Research*, 629-641.

Jaegler, A., & Burlat, P. (2012). Carbon friendly supply chains: a simulation study of different scenarios. *Production Planning & Control*, 269-278.

Jemai, J., Zekri, M., & Mellouli, K. (2012). An NSGA-II algorithm for the green vehicle routing problem. *Evolutionary computation in combinatorial optimization*, 37-48.

Kamakate, F., & Schipper, L. (2009). Trends in truck freight energy use and carbon emissions in selected OECD cointries from 1973 to 2005. *Energy Policy*, 3743-3751.

Kim, H., & O'Kelly, M. (2009). Reliable P-Hub Location Problems in Telecommunication Networks. *Geographical Analysis*, 283-306.

Kim, J., & Soh, S. (2012). Designing hub-and-spoke school bus transportation network: a case study of Wonkwang University. *Promet - Traffic and Transportation*, 389-394.

Klincewicz, J. (1996). A dual algorithm for the uncapacitated hub location problem. *Location Science*, 173-184.

Klincewicz, J. G. (1992). Avoiding local optima in the p-hub location problem using tabu search. *Annals of Operational Research*, 283-302.

Knowles, J., & Corne, D. (1999). The Pareto achieved evolution strategy: a new baseline algorithm for multiobjective optimisation. *Proceedings of the 1999 Congress on Evolutionary Computation*, 98-105.

Kuby, M., & Gray, R. (1993). Hub network design problem with stopovers and feeders: Case of Federal Express. *Transportation Research*, 1-12.

Kumar, M., Hussian, M., Upreti, N., & Gupta, D. (2010). Genetic Algorithm Review and Application. *International Journal of Information Technology and Knowledge Management*, 451-454.

Kuo, A., Miller-Hooks, E., Zhang, K., & Mahmassani, H. (2008). Train Slot Cooperation in Multicarrier, International Rail-Based Intermodal Freight Transport. *Transportation Research Record: Journal Transportation Research Board*, 31-40.

Lamming, R., & Hampson, J. (1996). The Environment as a Supply Chain Management Issue. *British Journal of Management*, S45-S62.

Levitin, G. (2005). *The universal generating function in reliability analysis and optimization.* London: Springer.

Levitin, G., Lisnianski, A., Hanoch, B. H., & Elmakis, D. (2000). Genetic algorithm and universal generating function technique for solving problems of power system reliability optimization. *DRPT2000. International Conference on Electric Utility Deregulation and Restructuring and Power Technologies. Proceedings*.

Lin, C.-C. (2010). An integral constrained generalized hub-andspoke. *Transportation Research Part E: Logistics and Transportation Review*, 986-1003.

Lin, C.-C., & Chen, S.-H. (2008). An integral constrained generalized hub-andspoke. *Transportation Research Part E: Logistics and Transportation Review*, 986-1003.

Lin, D. Y., & Ng, K. H. (2012). The impact of collaborative backhaul routing on carbon reduction in the freight industry. *Transportation Research Part D: Transport and Environment*, 626-628.

Liu, S., Kasturiratne, D., & Moizer, J. (2012). A hub-and-spoke model for multi-dimensional integration of green marketing and sustainable supply chain management. *Industrial Marketing Management*, 581-588.

Meng, Q., & Wang, X. (2011). Intermodal hub-and-spoke network design: Incorporating multiple stakeholders and multi-type containers. *Transportation Research Part B: Methodological*, 724-742.

Mohammadi, M., Tavakkoli-Moghaddam, H., & Rostami, H. (2011). A Multi-Objective Imperialist Competitive Algorithm for a Capacitated Hub Covering Location Problem. *Journal of Industrial Engineering Computations*, 671-688.

Mohammadi, M., Tavakkoli-Moghaddam, Siadat, A., & Rahimi, Y. (2016). A game-based meta-heuristic for a fuzzy bi-objective reliable hub location problem. *Engineering Applications of Artificial Intelligence*, 1-19.

Murata, T., & Ishibuchi, H. (1995). MOGA: Multi-Objective Genetic Algorithms. *Proceedings of 1995 IEEE International Conference on Evolutionary Computation*, 289-294.

O'Kelly, M. E. (1987). A Quadratic integer program for the location of interacting hub facilities. *European Journal of Operational Research*, 393-404.

O'Kelly, M. E. (2009). Rectilinear minimax Hub Location. *Journal or Geographical Systems*, 227-241.

O'Kelly, M. E. (2012). Fuel Burn and Environmental Implications of Airline Hub Networks. *Transportation Research D*, 555-567.

Osman, I. H., & Laporte, G. (1996). Metaheuristics: A Bibliography. *Annals of Operations Research*, 511-623.

Ozener, O. O., Ergun, O., & Savelbergh, M. (2009). Collaboration for Truckload Carriers. *INFORMS Annual Conference*.

Pan, S., Ballot, E., & Fontane, F. (2013). The reduction of greenhouse gas emissions from freight transport by pooling supply chains. *International Journal of Production Economics*, 86-94.

Pérez-Bernabeu, E., Juan, A. A., & Barrios, B. B. (2014). Horizontal cooperation in road transportation: a case illustrating savings in distances and greenhouse gas emissions. *International Transactions in Operational Research*, 585-606.

Pinchera, D., Perna, S., & Migliore, M. D. (2017). A Lexicographic Approach for Multi-Objective Optimization in Antenna Array Design. *Electromagnetics Research M*, 85-102.

Pirkul, H., & Schilling, D. (1998). An eficient procedure for designing single allocation hub and spoke systems. *Management Science*, 235-242.

Ramirez, L. E. (2018). *Multi-state multi-objective reliability analysis of renewable energy systems.* El Paso: The University of Texas at El Paso.

Reynoso-Meza, G., Sanshis, J., Blasco, X., & Martinez, M. (2010). Design of Continuous Controllers Using a Multiobjective Differential Evolution Algorithm with Spherical Pruning. *European Conference on the Applications of Evolutionary Computation*, 532-541.

Roni, M. S. (2013). *Analyzing the impact of a hub and spoke supply chain design for long-haul, high-volume transportation of densified biomass.* Mississippi State University.

Roni, M. S., Eksioglu, S. D., & Cafferty, K. G. (2014). *A Multi-Objective, Hub-and-Spoke Supply Chain Design Model for Densified Biomass.* Washington, D.C.: U.S. Department of Energy.

Roni, M. S., Eksioglu, S. D., Cafferty, K. G., & Jacobson, J. J. (2017). A multi-multiobjective, hub-and-spoke model to design and manage biofuel supply chains. *Annals of Operations Research*, 351-380.

Sadegheih, A., Drake, P. R., Li, D., & Sribenjachot, S. (2011). Sadegheih, A., et al. "Global supply chain management under the carbon emission trading program using mixed integer programming and genetic algorithm. *International Journal of Engineering Transactions B: Applications*, 37-53.

Song, L., & Regan, A. (2004). An auction based collaborative carrier network. *Proceedings of the 83rd Annual Meeting of the Transportation Research Board*.

Sun, L., & Zhao, L. (2009). Research on optimization of hub-and-spoke logistics network design with impedance effect. *POMS 20th Annual Conference*.

Taboadad, H. A., & Coit, D. W. (2012). A new multiple objective evolutionary algorithm for reliability optimization of series-parallel systems. *International Journal of Applied Evolutionary Computation*, 1-18.

Toppcuoglu, H., F., C., Ermis, M., & Yilmaz, G. (2005). Solving the uncapacitated hub location problem using genetic algorithms. *Computers and Operations Research*, 967-984.

Ushakov, I. (1986). A Universal Generating Function. *Soviet Journal of Computer and Systems Sciences*, 118-129.

Vachon, S., & Klassen, R. D. (2008). Environmental Management an Manufacturing Performance: The Role of Collaboration in the Supply Chain. *International Journal of Production Economics*, 299-315.

Vidyarthi, N., Kumar, R., & Jayaswal, S. (2013). *Hub-and-Spoke System Design for freight transportation with priority consignment classes.* Montreal: Centre interuniversitaire de recherche sur les reseaux d'enterprise.

Voruganti, A., Unnikrishnan, A., & Waller, S. (2011). Modeling Carrier Collaboration in Freight Networks. *Transportation Letters*, 51-61.
Xu, X. (2013). *Collaboration Mechanism in the Horizontal Logistics Collaboration.* Paris: Ecole Nationale Superieure des Mines de Paris.
Yancang, L., Lina, Z., & Shujing, Z. (2010). Review of Genetic Algorithms. *International journalof Information Technology and Knowledge Management*, 451-454.
Zhang, J., Wu, Y., & Liu, P. (2007). Routing Problem for Hybrid Hub-and-Spoke Transportation Network: a Case Study of a LTL Carrier. *2007 IEEE International Conference on Automation and Logisticsc*, 170-175.
Zitzler, E., & Thiele, L. (1998). *An Evolutionay Algorithm for Multiobjective Optimization: The Strength Pareto Approach.* TIK-report.
Zitzler, E., & Thiele, L. (1999). Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach. *IEEE Transactions on Evolutionary Computation*, 257-271.
Zitzler, E., Laumanns, M., & Thiele, L. (2001). SPEA2: Improving the Strength Pareto Evolutionary Algorithm. *TIK Report 103*.
Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C. M., & Grunert da Fonseca, V. (2002). Performance Assessment of Multiobjective Optimizers: An Analysis and Review. *IEEE Transactions on Evolutionary Computation*, 117-132.

Vita

Eduardo Jose Castillo Fatule graduated high school in his home country of the Dominican Republic before beginning his undergraduate studies. In 2015 he transferred to the University of Texas at El Paso, where he completed his bachelor's in Industrial Engineering in 2016 before starting his graduate studies which culminated on this thesis.

During his studies at UT El Paso, Mr. Castillo worked as a Research Assistant under the tutelage of Dr. Jose Espiritu in the department of Industrial, Manufacturing, and Systems Engineering. In 2015 he attended the Institute of Industrial and Systems Engineer's conference where he presented his first studies into metaheuristic optimization. Later the same year he also attended the annual conference of the society of Industrial Engineering and Operations Management, where he received the 2nd place award on the Graduate Student Paper competition, as well as Best Paper on the track of sustainability. More recently, he attended the 2018 Annual IISE conference, presenting multi-objective studies into the CCCMLP which provided part of the framework for this thesis.

Additionally, he worked as a Teaching Assistant under Dr. F. Aguirre and Dr. J. Sanchez, also at the IMSE department of UTEP for the classes of Operations Research, Operations Research II, Design of Experiments, and Statistical Quality Control.

Contact Information:
Eduardo Castillo
ejcastillof@gmail.com