2016-01-01

Minimizing Environmental Impacts For Hub And Spoke Distribution Network Problems Through The Use Of Multi-Objective Evolutionary Algorithms

Ileana Delgado
University of Texas at El Paso, idelgado3@miners.utep.edu

MINIMIZING ENVIRONMENTAL IMPACTS FOR HUB AND SPOKE DISTRIBUTION NETWORK PROBLEMS THROUGH THE USE OF MULTI-OBJECTIVE EVOLUTIONARY ALGORITHMS

ILEANA DELGADO CAMACHO

Master's Program in Manufacturing Engineering

APPROVED:

Heidi A. Taboada, Ph.D., Chair

Jose F. Espiritu, Ph.D., Co-Chair

Vinod Kumar, Ph.D.

Charles Ambler, Ph.D.
Dean of the Graduate School

Dedication

There exist no such words that can profoundly describe how grateful I feel for everything I have been granted. I wish to dedicate this thesis to all of those who believed in me such as my advisors, professors, family and friends. Especially, this achievement is dedicated to my parents, who emphasized the importance of education, taught me to always give the best of me, and the values of commitment and to become a better person day after day. In addition, this work is dedicated to the greatest gift that life has granted me, my family. For their endless love, encouragement and support, for being my motivation to continuously work hard. All effort given while accomplishing this work is in return for all your loving support.

MINIMIZING ENVIRONMENTAL IMPACTS FOR HUB AND SPOKE DISTRIBUTION NETWORK PROBLEMS THROUGH THE USE OF MULTI-OBJECTIVE EVOLUTIONARY ALGORITHMS

by

ILEANA DELGADO CAMACHO, 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

May 2016

Acknowledgements

Now that I am at the end of this journey, I wish to express my sincere gratitude and acknowledge everyone who contributed in my personal and professional development. The completion of this achievement could not have been possible without the assistance and participation of so many people whose names may not be enumerated. Their contributions are sincerely appreciated and gratefully acknowledged. However, I would like to express my deep appreciation and gratitude particularly to the following people:

I firstly acknowledge the continuous support from the Department of Industrial, Manufacturing and Systems Engineering and the Graduate School at The University of Texas at El Paso. In addition, it is my will to acknowledge USDA's support through the BGREEN grant for providing me with the financial support to continue advancing my education while working in the improvement and optimization of complex environmental problems through research tools.

Moreover, I would like to express the deepest appreciation to my committee chair and co-chair, Professor Heidi A. Taboada and Professor Jose F. Espiritu who recognized my potential and trusted me to join their research team. I feel blessed that I had the opportunity to coincide with such extraordinary human beings as they are and become part of their team. I thank them for their continuous support and for having convincingly conveyed that spirit of adventure in regard to research and for that excitement that they always show in regard to teaching. Without their guidance and persistent help, this thesis would not have been possible. My sincere appreciation also goes to all the professors at the IMSE department for their devoted instruction during my graduate years. Furthermore, I would like to thank my dear co-workers, Claudia Valles, Juan Fernandez, Ana Cram, Oswaldo Aguirre, Eduardo Castillo, Karla Gutierrez, and Olivia Moreno for encouraging and accompanying me during this journey. Thank you for those long days and nights of hard work. Thank you for your respective contributions to the completion of this thesis but most of all for your friendship.

Finally my warmest thanks must be to my family. Thanks to my parents for their initial taught, through their guiding example I was granted the value of the desire to learn and continue growing. Thank you father for sharing with me your engineering background and being my motivation of becoming an engineer as well. Thank you mother for always being there at all moments of my academic and personal developments, for providing me no other thing but your unconditional love and sincere support. I appreciate the time you invested in taking care of my children so I could comfortably continue to give my best at school, for having your arms open to every victory or failure I had to affront, and mostly, for motivating me at all moments and reminding me how capable I am of achieving whatever I had envisioned in my life. I thank god for giving me the best children I could have ever had. Thank you for their everyday smiles which have encouraged me to efficiently overcome difficulties encountered during my pursuit of this degree. Lastly, thanks to my dear and beloved husband for his continued and unfailing love, support and understanding.

Abstract

While transportation is crucial to our economy and our personal lives, as a sector it is also a significant source of greenhouse gas (GHG) emissions. The AASHTO report states, "America's transportation system has served us well, but now faces the challenges of congestion, energy supply, environmental impacts, climate change, and threatens the economic, social, and environmental future of the nation". With respect to the environment, transportation is the most visible aspect of supply chains. Transportation emissions can be reduced by implementing approaches that contribute to GHG emissions reduction, for example, the use of low-carbon fuels, new and improved vehicle technologies, improved transportation system efficiency, strategies to reduce the number of vehicle miles traveled and operating vehicles more efficiently. This work is motivated by the urgent need of advancing knowledge and understanding of transportation's current global challenges by contributing to the transition of the current transportation sector into a more sustainable one by implementing a more holistic analysis considering the three approaches stated by the U.S. Department of Transportation, which are, the use of low-carbon fuels, strategies to reduce the number of vehicle miles traveled, and optimizing the design of transportation networks to reduce trip frequencies.

This work provides a design alternative for the centralized carrier collaboration problem and to the biomass-to-biofuels supply chain by simultaneously addressing economic and environmental impacts. A hybrid hub-and-spoke system is used as a set for collaborative consolidation transshipments hubs increasing the efficiency of the operations. The present research proposes the development of Single and Multi-Objective Evolutionary Algorithms to solve the Centralized Carrier Collaboration Multi-hub Location Problem evaluating transportation cost and global warming potential separately and advancing it further into a simultaneous optimization. The second scenario focuses in the design of the supply chain for densified biomass by developing a Multi-Objective Evolutionary Algorithm to determine the Pareto optimal solutions for the operations configuration of a hub and spoke biofuels logistics network. Providing Pareto optimal

yearly configurations for the system's planning operations as well as their corresponding optimal transportation design.

In contrast to previous research, this work not only takes into consideration the GHG emissions, but evaluates the global warming impact category; likewise, it aggregates two other objectives by considering as well acidification and eutrophication potentials creating a more extensive and robust analysis of environmental sustainability of a hub and spoke network. Furthermore, data analytics methodologies were implemented to examine raw data on coordinates searching for optimal locations for facilities in the biofuel supply chain. The approaches used are the K-means and the Silhouette value analyses to select the optimal number of clusters for our data. In addition, a Life Cycle Assessment was conducted to evaluate the environmental potential impacts associated with the problems analyzed providing an adequate instrument for environmental decision support.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Background and Research Motivation

The rapid increase of population, current technology growth and economic globalization are the major factors of the high consumption of non-renewable resources generating detrimental effects on our environment. Currently, world population continues to grow although following a slower growth rate in comparison to past years. The United Nations Department of Economic and Social Affairs states that ten years ago, world population was growing by 1.24 per cent per year. Today, it is growing by 1.18 per cent per year, or approximately an additional 83 million people annually. The world population is projected to increase by more than one billion people within the next 15 years, reaching 8.5 billion in 2030, and to increase further to 9.7 billion in 2050 and 11.2 billion by 2100 (United Nations, 2015). Figure 1.1 provides a graphical representation of the growth pattern of population.

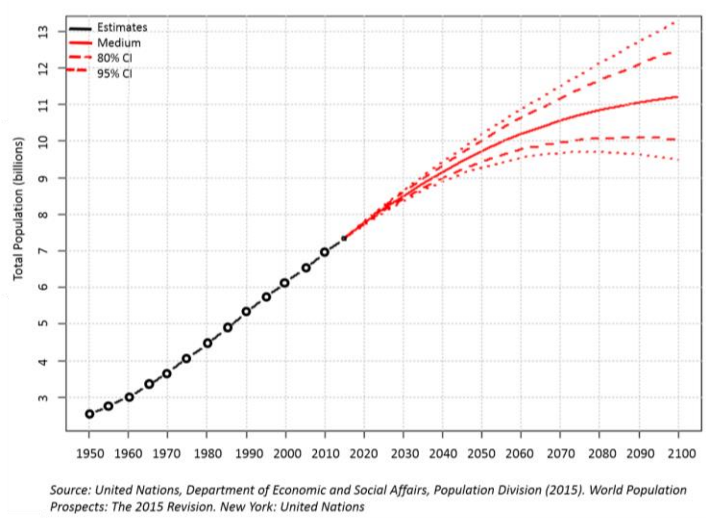

Figure 1.1: Population of the world: 1950-2015, estimates 2015-2100

This growth in population is leading to advancements in technology and demanding a globalization of economy to continue fulfilling population's needs. Nevertheless, these factors are leading to a critical time in human history where environmental sustainability needs to be

addressed when optimizing systems. Today, it is not only profit that matters, but a growing concern about sustainability of our society is emerging among many people, companies, and governments.

The deterioration of the environment is one of the largest issues threatening the existence of human beings today. The United Nations International Strategy for Disaster Reduction characterizes environmental degradation as the lessening of the limit of the earth to meet social and environmental destinations and needs (World Health Organization, 2012)**.** This deterioration is produced through the consumption of valuable resources, such as, air, water, and soil. It occurs when earth's natural resources are depleted, and environment is compromised in the form of extinction of species, pollution in air, water and soil.

Besides the sounded consequences of air pollution, some effects such as high temperatures, rising seas, and severe flooding and droughts have been recorded to be arising from inadequate environmental management. Likewise, not only is the size of population continuously changing throughout history, but earth's climate is as well. The current warming trend is of particular significance since most of the causes are human-induced and proceeding at a rate that is unprecedented in the past 1,300 years (B.D. Santer et al, 1996). In the 1860s physicist John Tyndall recognized the Earth's natural greenhouse effect and later in 1896 the scientist named Svante Arrhenius first speculated that changes in the levels of carbon dioxide in the atmosphere could substantially alter the surface temperature through this effect (B.D. Santer et al., 2003). This environmental impact is known as Global Warming or climate change.

Furthermore, the energy demand has also increased globally towards a situation where the consumption pattern in a particular country depends on the availability of energy resources. Today's economies rely on energy for almost all fundamental needs including food production, heat, transportation, manufacturing and communication. Different energy sources are wood, coal, petroleum products, nuclear power, solar, wind etc. Since the Industrial Revolution, fossil energy resources such as coal, petroleum, and natural gas have become the dominant sources of energy owing to the fact that they are readily accessible and inexpensive. Figure 1.3 represents the energy consumption by source and sector as provided by the U.S. Energy Information Administration,

illustrating the transportation sector as a secondary energy consumer, from which its energy sources are conformed by primarily petroleum with a 93%, natural gas with a 3% and renewable energy with a 4%.

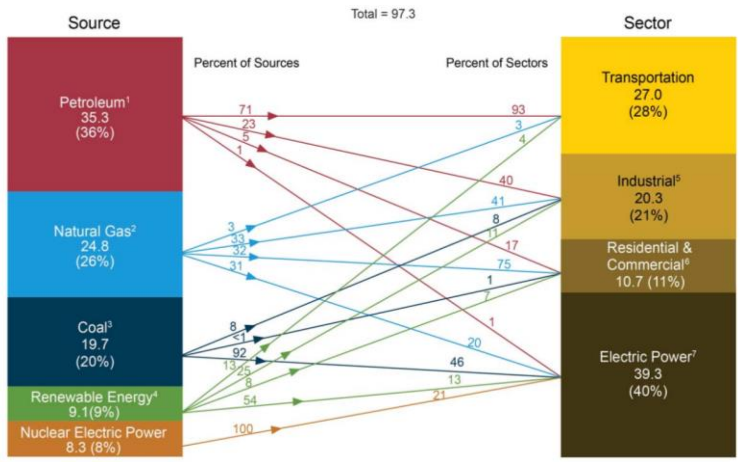

Figure 1.3: Total Energy by source and by sector[1]

Extensive studies have identified CO_2 as the largest source of U.S. greenhouse gas emissions accounting from fossil fuel combustion approximately 78 percent of GWP-weighted emissions since 1990. Emissions of CO_2 from fossil fuel combustion increased at an average annual rate of 0.3 percent from 1990 to 2012 (Environmental Protection Agency, 2014). The U.S. Environmental Protection Agency (EPA) declares that the fundamental factors influencing this trend include (1) a generally growing domestic economy over the last 23 years, (2) an overall growth in emissions from electricity generation and transportation activities, along with (3) a general decline in the carbon intensity of fuels combusted for energy in recent years by most sectors of the economy (Environmental Protection Agency, 2014).

While transportation is crucial to our economy and our personal lives, as a sector it is also a significant source of greenhouse gas (GHG) emissions. By the year of 2012, transportation was responsible for 27 percent of global energy consumption as shown in Figure 1.4 and 22 percent of

[1]US Energy Information Administration (2011). Annual Energy Review. https://www.eia.gov/totalenergy/data/annual/pdf/aer.pdf

fossil fuel burning across the world. In addition, it is attributed a 30 percent of global air pollution and greenhouse gases (Aykin, T, 1995).

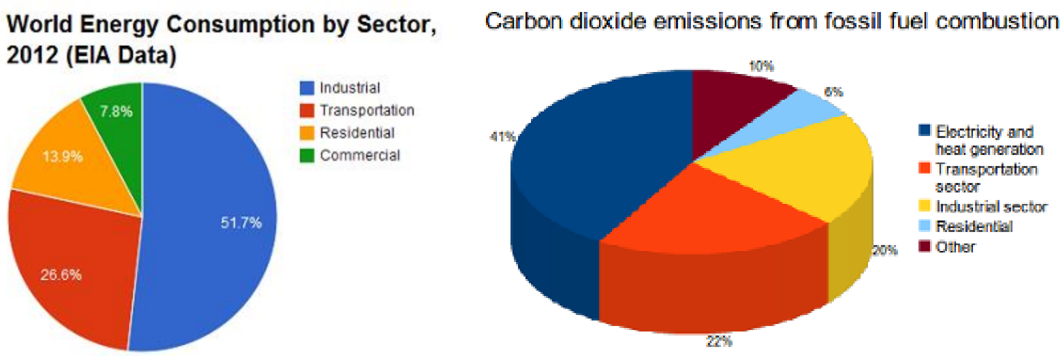

Figure 1.4: Energy Consumption and CO_2 emissions for transportation sector[2]

Fuel combustion emits carbon dioxide (CO_2), methane (CH_4) and nitrous oxide (N_2O). Also, a small amount of hydrofluorocarbon (HFC) is produced by the use of mobile air conditioners and refrigerated transport. Transportation is also a main source for NO_x, SO_2, and PM (particulate matter or fine dust) emissions. The AASHTO report states, “America’s transportation system has served us well, but now faces the challenges of congestion, energy supply, environmental impacts, climate change, and threatens the economic, social, and environmental future of the nation” (American Association of State Highway and Transportation Officials, 2015). With respect to the environment, transportation is the most visible aspect of supply chains.

In 2006, within the transportation sector, the largest sources of GHGs were passenger cars (34%) and light duty trucks (28%). The next largest source were freight trucks (20%) (U.S. Department of Transportation, 2007; AASHTO,2015) as illustrated in Figure 1.5. Emissions from on-road vehicles accounted for 79 percent of transportation GHG emissions.

---

[2]U.S. CO_2 Emissions from Fuel Combustion (2012), International Energy Agency.

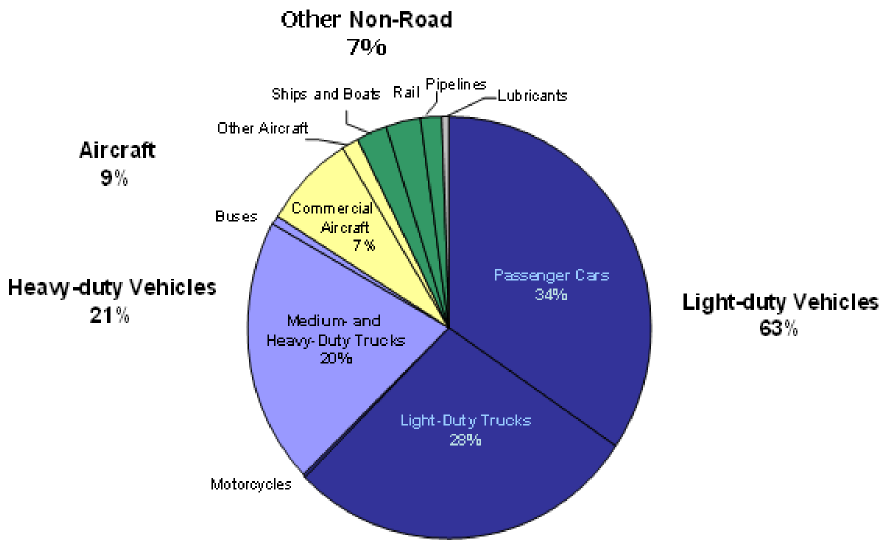

Figure 1.5. U.S. EPA inventory of U.S. greenhouse gas emissions and sinks: 1990-2009[3]

Contrastingly, greenhouse gas emissions in other sectors decreased 15% between 1990 and 2007 but emissions from transportation increased 36% during the same period. This increase happened despite improved vehicle efficiency because the amount of personal and freight transport has increased (European Commission, 2014; De Mello, P. F. B., & Frayret, J. M., 2014). These increases have been induced by the largest demand for travel and limited gains in fuel efficiency. Moreover, The U.S. Energy Information Administration estimates that U.S. gasoline and diesel fuel consumption for transportation in 2014 resulted in the emission of about 1,075 million metric tons and 444 million metric tons of CO_2, respectively, for a total of 1,519 million metric tons of CO_2. This total was equivalent to 83% of total CO_2 emissions by the U.S. transportation sector and equivalent to 28% of total U.S. energy-related CO_2 emissions in 2014 (U.S. Energy Information Administration, 2015). Additionally, between 1990 and 2012, GHG emissions in the transportation sector increased more in absolute terms than any other sector (i.e. electricity generation, industry, agriculture, residential, or commercial) (EPA, 2014). Consequently, it has been estimated that baseline global GHG emissions from human sources will increase between 25

[3]: U.S. EPA, Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2009 (U.S. EPA, 2011)

percent and 90 percent between 2000 and 2030 (Ballot, E., & Fontane, F. 2010).

With regards to the environmental effects, The Intergovernmental Panel on Climate Change (IPCC) projects that global temperatures will raise between 2°F to 11.5°F by 2100, and global sea level will rise between 7 to 23 inches. Nonetheless, the increase of emissions can be reduced by implementing approaches that contribute to GHG emissions reduction, for example, the use of low-carbon fuels, new and improved vehicle technologies, improved transportation system efficiency, strategies to reduce the number of vehicle miles traveled and operating vehicles more efficiently. According to the U.S. Department of Transportation, different system efficiency strategies will help reduce the GHG emissions by optimizing the design, construction, operation, and use of transportation networks to reduce trip frequencies and it has been estimated that the collective impact of these strategies on total U.S. transportation GHG emissions could range from 5-to-17 percent in 2030, or 6-to-21 percent in 2050 (Chen, C. C. F., & Schonfeld, P, 2012).

It is a fact that the transition to a sustainable living involves re-thinking the purpose for which systems are being created and create systems that adapt to the rapid changes and interacts together with the environment without harming it. This work is motivated by the urgent need of advancing knowledge and understanding of transportation's current challenges. It is vital to address environmental impacts within the transportation sector, since not only the world's economy depend on it, but social and environmental threads are being affronted as well. Not only this sector is one of the primary emitters of greenhouse gases, but also is relying in non-renewable resources. An urge for alternative fuel resources for transportation is encountered in addition, efficient systems that reduce the environmental impacts will be necessary to accelerate the transformation to a sustainable transportation.

1.2 THESIS OBJECTIVE

Therefore, the objective of this thesis is to make use of operations research tools to develop a more sustainable and efficient transportation system, identifying trade-offs between

environmental aspects and economy. The main goal of the thesis is to demonstrate the application of metaheuristic approaches in hub and spoke distribution networks. This work will contribute to the transition of transportation sector into a more sustainable by implementing into the analysis three approaches stated by the U.S. Department of Transportation, which are, the use of low-carbon fuels, strategies to reduce the number of vehicle miles traveled, and optimizing the design of transportation networks to reduce trip frequencies.

The metaheuristic approach will be shown to be efficient in reducing the number of miles traveled and consequently the environmental impacts, by testing it under two different scenarios. Both scenarios will implement a hub and spoke network model that will reduce the trip frequencies. This work provides a design alternative for the centralized carrier collaboration problem and to the biomass-to-biofuels supply chain to simultaneously address economic and environmental impacts.

The operations research tool of multiple objective optimization will be employed. In addition, the metaheuristic approach selected is an evolutionary algorithm. Past literature has shown the effectiveness of this algorithm for solving hard combinatorial optimization problems such as the ones presented here.

General approaches exist to solve multiple objective optimization problems, such as methodologies that provide one single solution by combining objectives into an overall aggregated function, and the approaches that obtain a non-dominated Pareto-optimal set. This work will make use of non-dominated Pareto-optimal approach through a Multiple Objective Evolutionary Algorithm (MOEA).

Similar problems as the ones presented in this thesis have been previously solved through mathematical approaches such as linear programming, mixed integer programing, mixed integer linear programming. Although mathematical approaches are exact and able to provide an optimal solution to single objective problems, they may require an exponential computational time to provide a solution of the problem. Metaheuristic approaches does not guarantee obtaining an optimal solution, instead, they provide feasible and robust solutions with a minimum amount of computational time.

1.3 Scope and Limitations

The present work analyzes two scenarios, the first one is a centralized carrier collaboration and multi-hub location problem (CCCMLP), and the second is the biomass-to-biofuel logistic system design problem. Both scenarios are going to be limited by information and data previously given by other studies.

For the CCCMLP, the data used was previously collected by Hernandez *et al.* (2011a). The study considers a transportation network of ten nodes, and two possible carriers collaborating. This problem was previously modeled as a binary (0–1) multi-commodity minimum cost-flow problem formulation for two rate-setting behavioral cases and solved with a branch-and-cut algorithm.

The second scenario focuses in the design of the supply chain for densified biomass. A hub-and-spoke network structure is used to model this supply chain. The model is formulated as a multi-objective, mixed-integer programing problem considering the planning phase and the transportation design. Some studies have addressed either of both aspects such as Zhu & Yao (2011) who addressed the planning phase formulating a Mixed Integer Linear Programming for the logistics system of multiple-feedstock, exploring the possibility of using three feedstock at the same time; and Roni, Md. *et al.* (2014), who addressed the transportation phase formulating a multi-objective mixed integer programming under economic, environmental, and social criteria and approach it with an augmented constraint method. Data from both studies was collected to develop a case study that could provide a design for a yearly planning phase considering a hub and spoke network design for transportation.

This thesis aims to minimize transportation cost in the first scenario and maximize profit in the second in addition to simultaneously minimizing the environmental impacts in both scenarios. As part of the optimization process, this work emphasizes and focus in the use of evolutionary algorithms, however, this thesis may inspire the application of other methodologies as well.

1.4 Thesis Outline

Hereafter, the introduction and motivation of the thesis has been provided in chapter 1. Giving a full in-depth of the need in optimizing environmental impacts in the transportation sector, specifically in hub and spoke networks and how evolutionary algorithms can be a powerful tool to address this problem. The rest of the thesis is structured in the following way.

Chapter 2 provides the definition and history of hub and spoke networks, in addition to an overview of past literature, comprising variations of the problem, methodologies utilized to address it and main applications of hub and spoke networks.

Chapter 3 reviews literature on the impact of Multiple Objective Optimization against single objective and describes the different approaches that are being currently used to address multi-objective problems. As well, a review of the different existing Evolutionary Algorithms is given.

Chapter 4 analyzes the centralized carrier collaboration and multi-hub location problem, providing at first literature in this area to identify the gap in this research. Afterwards, the mathematical formulation, and methodologies are described in this section. Finalizing the chapter with the explanation of the case studies and its results. In this case studies, the problem is going to be analyzed as a single objective and a bi-objective, addressing cost and global warming impact isolated, and then simultaneously.

Chapter 5 describes the optimization of biofuel logistics through the use of hub and spoke networks. As in the past chapter, a brief literature review is provided to understand better the problem and current actions, followed by the mathematical formulation and methodologies. At the end of the chapter, the case study developed will be described and results will be provided.

Furthermore, conclusions on the performance of the algorithm and future work are stated at Chapter 6.

Chapter 2: Hub-and-Spoke Networks

In the past, transportation networks were designed in direct-route operations or what is also known as point-to-point design, nevertheless this was not the most practical design in view of the fact that it resulted in loss of money among shipping and transportation companies. However, as information continues to upgrade and technology keeps advancing, researchers are constantly gaining relevant knowledge to develop competent and profitable networks. Hub-and-Spoke structure was the result of providing cheap, reliable, fast, flexible and higher accessibility transportation services to air passengers. Since the hub-and-spoke concept was introduced to the aviation market after the US airline deregulation, it became a primary distribution model employed by leading international logistics companies. According to J. Kim and S. Soh (2012) “This network structure achieves economies of scale by consolidating and rerouting shipments at hubs”.

A hub and spoke network is centralized, integrated logistics system designed to keep costs down. Hub and spoke distribution centers receive products from many different origins, consolidate the products, and send them directly to destinations. This structural design has been nicknamed as the “wheel” network due to its appearance, such as a bicycle wheel, with the hub as the strategic center of the network, and the spokes coming out of it to connect with remote points. In more technical terms, 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. Figure 2.1 provides an illustration of a hub and spoke model in contrast to a point-to-point network.

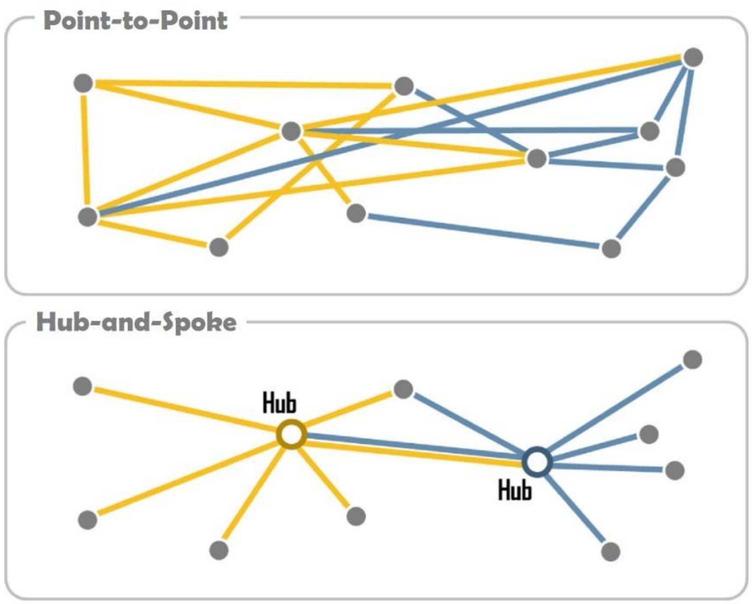

Figure 2.1: Representation of point-to-point and hub and spoke networks

Hub and Spoke structure is being widely applied on various transportation and communication systems. Air transportation systems was the pioneer in successfully implementing hub and spoke networks. In fact, within the aviation industry this structure was proved to be effective in minimizing costs. Going back in history, before the Deregulation on 1979, the Civil Aeronautics Board (CAB) controlled all airline pricing, routes and scheduling. At that time, a single point to point network was used where airlines fly passengers from point A to point B with no stops in between, what is known as direct flights. The benefits to this type of network was mainly to the passenger since it minimized the travel time and there was no need to disembark and transfer to another plane. Although this benefits seemed to be enough to keep direct routes, some disadvantages were brought with this system, mainly for smaller cities, which could not operate on these point to point networks because of the demand, in addition, with this type of network the fares for passengers were higher. On the other hand, hub and spoke airlines provide travelers with hundreds of flight options and locations to travel. Today, most of the major U.S. passenger airlines operation as of 2001 use hub-and-spoke network to route their plane traffic, by operating

centralized hub airports, which are typically larger in size and located in larger cities. In this network, hubs fly to smaller airports that point to point networks may avoid.

In addition to air transport, shipping companies have also adopted the hub-and-spoke model with the objective of speeding up deliveries and reducing its cost. In the shipping industry, hubs are known as central processing facilities where the cargo will be sent from its origin point. The shipment is then either stored or distributed directly from the center of the network. This network has been successfully implemented by shipping companies such as Federal Express, UPS, Norfolk Southern and Yellow Freight resulting in a competitive logistics advantage for them by reducing transportation costs, improving cycle times, and reducing inventory. Wal-Mart and Lowes are some of the many companies that are now realizing that significant cost savings can result from improving their distribution process.

Within the communication industry, hub and spoke network is widely used and in this area it is known as the star topology. In this model, the access point is physically connected to the internet with a wire; like spokes on a wheel, and then, all user devices connect to the wireless router in the center, the hub. All network traffic must go through the hub to reach other spokes in the network or to connect to an outside network. From the IT's point of view, hub and spoke networks does not provide minimum cost or time as they do in transportation, but they offer a high degree of security because each device on the network is isolated from the others through the single connection to the wireless router. Some other examples of communication industries that employ hub and spoke networks are cellular networks, where the cellular tower for a certain region is the hub and all the mobile devices roaming through that area are the spokes. Hub and spoke models can also applied in Social Media, where the structure is about having a central hub (website) and radiating out from that is all your social media and online options that feed back into this hub, driving traffic to it.

As it can be seen from the information presented above, hub and spoke networks are constantly being mentioned for the improvement of systems. Additionally, this network design has been out on the field for approximately 30 years, and is continuing to serve as an efficient design.

Therefore, several researchers continue to study and improve specific research areas through the use of hub and spoke networks. Some of these thorough research and its wide variety of applications are presented in the following sections. The majority of the literature has focused on conducting studies on improving transportation systems through this network considering economic concerns. Nonetheless, there is scarce literature addressing the environmental aspect of hub and spoke network designs. However, this study intends to contribute in initial research for these missing data by focusing on analyzing and minimizing the environmental impacts of hub and spoke distribution network under two different case scenarios.

2.1 Literature Review

2.1.1 Hub location

The studies of hub location problems began with the pioneering studies of a geographer named O'Kelly, (1987) who formulated and solved a hub-and-spoke facility location problem as a quadratic integer program and developed enumeration heuristics to solve the problem. Through time, most of his research has focused on optimizing this network improving the efficiency of the design by developing integer linear programs to represent the mode and route assignment aspects of operational problem (O'Kelly, M. E. and Lao, Yong, 1991), as well as introducing fixed facility costs into a hub location model (M.E. O'Kelly, 1992a). Other research contributed to the advancements by introducing a clustering approach (M.E. O'Kelly, 1992b) and developing a simulation-based software system for evaluating Hub-and-Spoke transportation networks known as HUBNET (Taha, T. *et al*, 1996)

2.1.2 Single and multiple allocation hub location problem

Hub-and-spoke facility locations problems can also be classified as single assignment as stated by Ernst and Krishnamoorthy, (1999) where a specific origin-destination flow is assigned to one hub only and multiple assignment as stipulated by Ebery *et al.*, (2000) where a specific origin-destination flow is split among multiple hubs. With all these new information being

discovered new studies are found considering fixed and variable transportation costs on all arcs (J.F Campbell *et al*, 2015), approaching a hybrid hub-and-spoke network. This hybrid concept was previously introduced by Kuby and Gray, (1993) and Aykin, (1995) who developed models for hybrid hub and spoke facility location which captured the flexibility of flows being sent directly without passing through hubs.

2.1.3 Capacitated and uncapacitated, single and multiple allocation hub location problem

Extensions to this problem were considered in which the capacity of the hubs is part of the decision making process. Some of these variants include capacitated and uncapacitated analysis such as the following studies by Aykin *et al* (1994), Ernst and Krishnamoorthy (1999), and Ebery *et al*. (2000) whom considered capacities on hubs whereas Pirkul H. and Schilling D.A. (1998), Klincewicz, (1996), and Topcuoglu, H. *et al*., (2005) studied the uncapacitated case. In the uncapacitated hub location problem there is a fixed-set up cost for each hub but the number of hubs to locate remains as an unconstrained decision variable. Furthering previous studies extensions addressing this problem as a multi-objective are being studied and solved by a multi-objective imperialist competitive algorithm (M. Mohammadi *et al*, 2011). A variation of the uncapacitated hub location problem (UHLP) is the uncapacitated single allocation hub location problem (USAHLP) where the number of hubs is left as a decision variable and the cost is fixed as in the past problem, but the single allocation concept is added which refers to the allocation of all traffic flow from the spokes to a single hub node. Topcuoglu, H. *et al* (2005) solved this problem using genetic algorithms. Their work was compared to other solutions presented in the literature demonstrating that their methodology surpass in quality and CPU time and matches the optimal solution. Chen J.-F. (2007) uses two approaches to determine the upper bound for the number of hubs along with a hybrid heuristic based on the simulated annealing method, and tabu list. For the uncapacitated multiple allocation hub location problem (UMAHLP) a 4-index formulation was proposed and solved by an accelerated Primal (Benders) decomposition and a greedy heuristic (Gelareh, S. and Nickel, S., 2011)

2.1.4 Single and multiple allocation p-hub median problem

Earliest formulations assumed that the number of hubs to be established is given as an input and not as a decision variable, this problems are referred as the p-hub median problem (Hyun Kim and M.E. O'Kelly, 2009). The p representing the number of hub facilities that are expected to be located to minimize total flow cost. Contributions made to this problem include mixed integer non-linear programming model considering demand uncertainty and congestion effects (Miranda Jr. *et al*, 2011). This initial formulation was driven to multiple directions to add new features consistent with various applications of hub-and-spoke facilities. As O'Kelly introduced the concept of single allocation, several studies focused in optimizing this subdivision of the problem through the use of tight linear programming formulation, langranian relaxation (Pirkul, H. and Schilling, D.A., 1998). Some others proposed meta-heuristic algorithms for the optimization of single allocation problems (Gomes, B. *et al*, 2013). On the other hand, for the multiple allocation problem, non-linear mixed integer programming formulations with a generalized Benders decomposition algorithm was presented (R.S. de Camargo *et al*, 2009) and to better model costs in hub networks, fixed costs were entered into the design (M.E. O'Kelly *et al*, 2014).

2.1.5 Capacitated and uncapacitated p-hub median problems

As the general hub location problem, variants of this problem were also introduced not only as single and multiple allocation but also as capacitated and uncapacitated. On the capacitated case, Klincewicz, J., (1996) developed an algorithm based on dual ascent and dual adjustment techniques within a branch and bound scheme. Abdinnour-Helm, Sue., (1998) developed a new heuristic method based on genetic algorithms and tabu search, providing improved results in comparison to GAs alone and matching the best solutions found in literature at that moment. For the capacitated single allocation p-hub median problem, non-linear programming models (Li, S. and Zhao L., 2009) are presented with stochastic demand and time-based service level constraints at the hubs (Nayneet Vidyarthi *et al*, 2013). Some other studies suggest the use of heuristic

algorithms (Abdinnour-Helm, 2001), and tight linear programming (D. Skorin-Kapov, *et al*, 1996) to solve the uncapacitated single allocation p-hub location problem.

2.1.6 Developed Hub-and-Spoke models

Besides all the variants of these problems, some other studies have proposed different models and problems arising from the hub location problem such as the minimax which sites a facility to minimize the maximum weighted interaction cost between pairs of fixed nodes considering distances represented by a rectilinear norm (M.E. O'kelly, 2009) or the hub arc location model, which instead of locating discrete hub facilities, the model locates hub arcs, which have reduced unit flow (J.F. Campbell *et al*, 2005). In addition to the models previously mentioned, some other studies consider dynamic (or multi-period) hub location problem (Contreras, I. *et al*, 2010), hub and spoke network in continuous Euclidean space (Carlsson, J.G. and Jia, F., 2013), a generalized hub and spoke network that integrates the operation of pure, stopover and center directs hub-and-spoke networks (Ling and Chen, 2008), and hierarchical hub and spoke network (Cheng-Chang Lin, 2010). Additional researchers that contributed to innovation of models are Ishfaq and Sox (2012) who integrated hub operation queuing model in the hub-allocation model. Whereas, Meng and Wang, (2011) developed a mathematical program with equilibrium constraints model for the intermodal hub and spoke network design problem with multiple stakeholders and multi-type containers.

2.1.7 Solution Methodologies

In terms of solution methodologies, Lagrangian relaxation has been widely used to solve different variations of the hub-and-spoke facility location problem as did Aykin *et al*, (1994), Pirkul, H. and Schilling, D.A. (1998), Elhedhli, S., Wu, H., (2010). Other solution techniques which have been explored include heuristic methods as by Klincewicz, (1996), meta-heuristics as did Klincewicz, (1992), and exact methods such as branch and bound as did Aykin(1994, 1995) and Klincewicz, (1996). Hub location problems are in some way similar to facilities location

problems in the sense of locating the hub facilities and designing hub networks. This network structure has had a great impact on systems by improving service quality and reducing operating costs considering all of its different variants. However, in all of these variants, the objective function is to find the location of the hubs and the allocation of the nodes so that the total cost of the network is minimized.

2.2 Hub-and-spoke networks and environmental sustainability

From the beginning of human history transport has actuated as a principal factor for growth, allowing commuting and trade between regions. Through time, our transportation systems have been evolving and upgrading to more cost-efficient structures. Resulting in strategic and organized networks as it is with hub-and-spoke networks. Most of these advances are intended to impact the economies of the different industries. Nonetheless, the actual situation inquire for a re-evaluation of life cycle thinking transitioning to a comprehensive analysis of sustainable systems. Since the concept of sustainability was transformed into business mainstream, hub and spoke networks environmental aspect should be analyzed equally as economic factors.

Unfortunately, until today minor concern has been noticed in considering environment on hub and spoke networks. Scarce studies have been found throughout the literature at this present moment, some of these analyses include M.E. O'Kelly, (2012) who uses the interactions between a system of cities as an experimental context for understanding selected environmental costs and benefits of concentrated flow. Fuel burn is used as an indicator of environmental cost. Moreover, hub-and-spoke networks have also been implemented and proved to be affordable and efficient in green supply chains. Sustainable supply chain has become key topic of academic research and managerial practices. For instance, Shaofeng L. *et al*, (2012) proposed a new hub and spoke integration model to integrate marketing and sustainable supply chain management. In addition, Roni, Md, (2013), Roni, Md *et al* (2014), and Mohammad, R., (2016) analyzed a hub-and-spoke

network design for biomass supply chain taking into consideration costs and emissions. Hub-and-Spoke networks have subsisted for the advancements in their modelling enumerating new variables and constraints to better approach realistic models. Correspondingly, this advancements need to be proven and optimized under environmental scenarios to continue validating this structure as an efficient network and to lessen Hub-and-Spoke networks' environmental impacts.

Chapter 3: Multiple Objective Optimization and Evolutionary Algorithms

In the past, systems were optimized considering one single objective at the time, whether it was minimizing cost, maximizing reliability, or any other aspect to be improved from a system. Nevertheless, optimizing systems per objectives separately was not the most effective way to better a system, when considering that improving a part of the system could mean reducing the quality of another function of the system. Considering only one objective to evaluate a determined problem can be useful for terms of simplicity whereas it is far from a comprehensive description of the real problem. Most real world problems involve simultaneous optimization of several objectives which often have conflict among each other. For instance, having maximization and minimization objective forms, as well as solutions expressed in different units. Therefore, the need to face real applications brought up a new hypothesis of a single-objective function being no longer suitable, and the introduction of multi-objective optimization framework allows the management of more information. (Caramia, M & Dell'Olmo, P., 2008)

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

$$\min f(x)$$

$$x \in S,$$

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

Whereas, a multiple objective optimization problem is expressed by a number of objectives and several inequality and equality constraints. The mathematical expression could be written as follows:

Minimize/Maximize $f_i(\mathbf{x})$ for i=1, 2,… n

Subject to:

$g(\mathbf{x}) \leq 0$

$h(\mathbf{x}) = 0$

Where, $f(\mathbf{x})$ stands for each of the objective functions whether being minimized or maximized and the subscript i representing the actual objective function until n total number of

functions. The constraints which define the feasible region, are represented in the form of $g(\mathbf{x})$ for the inequalities and $h(\mathbf{x})$ for the equalities. Moreover, the decision variables are represented by x.

Several approaches are being currently utilized to solve multi-objective optimization problems. Some methodologies obtain a single solution whereas others obtain a set of non-dominated Pareto set.

3.1 Mathematical Approaches

This first set of approaches, combine the objective functions into an overall aggregated function to find a single optimized solution.

3.1.1 Goal Programming

This approach was developed by Charnes *et al.* (1995), aiming to find specific goal values for objective functions being analyzed. In this methodology goals b_j are specified for each objective function $F_j(x)$. Moreover, the total deviation from the goals is aimed to be minimized, where d_j is the deviation from the goal b_j for the *jth* objective. The deviation is separated into positive and negative values to consider underachievement and overachievement, respectively, where achievement implies that a goal has been reached. The optimization problem is formulated as follows:

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

$$\text{Subject to } F_j(x) + d_j^+ - d_j^- = b_j, j = 1,2 \dots \dots \dots . k$$

$$d_j^+, d_j^- \geq 0, j = 1,2 \dots \dots \dots \dots , k$$

$$d_j^+, d_j^- = 0, j = 1,2 \dots \dots \dots \dots , k$$

However, its major disadvantage is that it does not guarantee providing a Pareto Optimal solution. In addition, its formulation can be problematic when analyzing larger problems. Some variations of this formulation include, Archimedean goal programming (or weighted goal

programming), preemptive (or lexicographic) goal programming approach, multigoal programming, and goal attainment method.

3.1.2 Weighted Sum or Scalarization Technique

The most common approach to multi-objective optimization is the weighted sum method, strategy that converts the multi-objective problem of minimizing the vector of criteria functions, into a scalar problem by constructing a weighted sum *F(x)* of all objectives. In more detail, the weighted-sum method minimizes a positively weighted convex sum of the objectives and can be represented as follows:

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

Where *n* represents the total number of objectives *i* and w_i their respective weights.

An advantage to this methodology is the simplicity of the multi-objective problem transformed into a single through the aggregation of its objectives, in addition to providing the decision maker with the opportunity to assign a level of priority through weights. The problem with this methodology is encountered in choosing the best weighting coefficients to each of the objectives. Highlighted importance is being given to determining the adequate weights some researchers have even proposed weights to be functions instead of constants, in order to better mimic a preference function accurately. In addition many other systematic approaches have been developed to select weights, for instance, ranking, categorization, rating, and eigenvalue methods just to mention some.

3.1.3 Lexicographic method:

Another way to address multiple objectives is through lexicographic approach developed by Fishburn, (1974), a technique that requires the decision-maker to establish a priority for each objective. With the lexicographic method, the objective functions are arranged in order of importance. Then, two solutions are compared with respect to the most important objective. If this

results in a tie, the algorithm proceeds to compare the solutions but now with respect to the second most important objective. This tie-breaking process is repeated until no objectives are left to be examined. This optimization problems solved one at a time can be represented with the following formulation:

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

$$\text{Subject to } F_j(x) \leq F_j x * j\ , j = 1,2, \ldots\ldots i-1, i > 1, i = 1,2 \ldots., k$$

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 ($j = 1$), $F_j x * j$ is not necessarily the same as the independent minimum of $Fj\ (x)$, because new constraints have been introduced.

Some researchers have introduced to this approach constraint relaxation methodologies, other variations include positive tolerances, and a combination of this methodology with the ε-constraint approach discussed in Sect. 3.1.5. The lexicographic approach can be advantageous when a small number of objectives is being considered. In addition, the performance of this methodology is related to the ordering given by the set priorities. Therefore, a disadvantage can be found considering the best set for this priorities.

3.1.4 Multi-Attribute Utility Theory

Utility refers to the satisfaction that each attribute provide to the decision maker. Thus, utility theory assumes that any decision is made on the basis of the utility maximization principle, according to which the best choice is the one that provides the highest satisfaction to the decision maker. In multiattribute utility analysis (MAUA) the total utility of a design solution is a scalar on the interval between 0 (no utility) and 1 (highest utility). According to the utility theory if X_i is the measure of effectiveness of an attribute (or quality characteristics) i and there are n attributes, then the joint utility function can be expressed as:

$$U(X_1, X_2, \ldots\ldots.., X_n) = f(U_1(X_1), U_2(X_2), \ldots\ldots., U_n(X_n))$$

Here, $U_i(X_i)$ is the utility of the *ith* attribute. The overall utility function is the sum of individual utilities if the attributes are independent, and is given as follows:

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

The overall utility function after assigning weights to the attributes can be expressed as:

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

While utility optimization is effective and widely used it requires extensive interviews to determine appropriate utility functions and weights. Once the utility function has been constructed, optimization can occur and the design with maximal utility can be found.

3.1.5 ε-constraint approach

A procedure that overcomes some of the convexity problems of the weighted sum technique is the Epsilon-constraint method. These method was proposed by Chankong and Haimes (1983) for general multi-objective problems. Here, the decision maker chooses one objective out of *n* to be minimized; the remaining objectives are constrained to be less than or equal to a given target value. For instance, when evaluating a bi-objective case, the problem is subdivided into two problems, $P_1(\varepsilon_2)$ and $P_2(\varepsilon_1)$ which are the following:

$$\min f_1(\underset{x}{\rightarrow})$$

$$\text{Subject to: } \underset{x}{\rightarrow} \in X,$$

$$f_2(\underset{x}{\rightarrow}) \leq \epsilon_2$$

And

$$\min f_2(\underset{x}{\rightarrow})$$

$$\text{Subject to: } \underset{x}{\rightarrow} \in X$$

$$f_1(\underset{x}{\rightarrow}) \leq \epsilon_1$$

Respectively.

In contrast to other aggregating approaches, this methodology is able to identify a number of non-inferior solutions on a nonconvex boundary. Nonetheless, a disadvantage is that the use of hard constraints is rarely adequate for expressing true design objectives.

All of these methodologies are found suitable when all objectives get better or worse together, these conventional approaches can effectively find the optimal solution. However if the objectives conflict, then there is not a single optimal solution. In this case, a multi-objective optimization study should be performed that provides multiple solutions representing the tradeoffs among the objectives. This is commonly called Pareto optimization.

3.2 Pareto Optimization

Usually, there is no single optimal solution, but instead a set of alternative solutions. These alternative solutions are considered optimal since no other solutions in the search space are superior to them when considering all objectives. In other words, optimal solutions are integrated by solutions that are not dominated by any other solutions making possible the existence of different tradeoffs between objectives (Eckart Zitzler and Lothar Thiele, 1998; Eckart Zitzler *et al*, 2002).

These optimal solutions are known as Pareto-optimal solutions, and the set of these solutions is denoted as the Pareto-optimal set. Figure 3.1 provides an illustration of the different terminologies used in Pareto optimization.

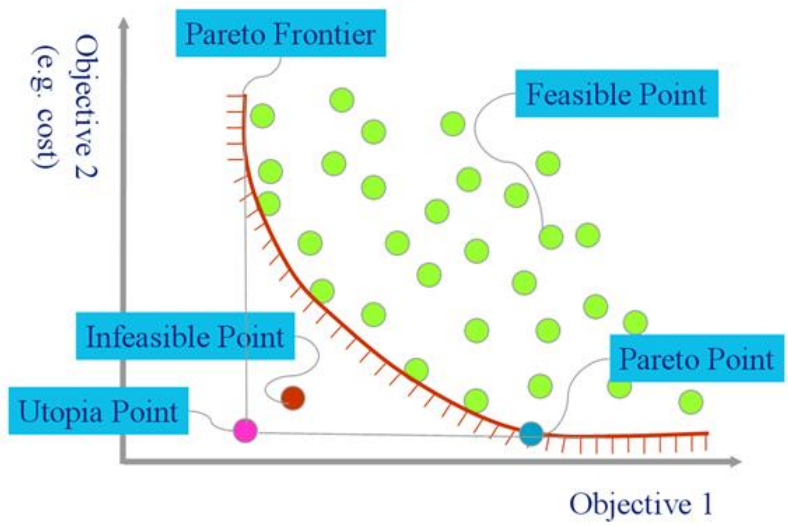

Figure 3.1: Pareto Optimization illustration and terms

Furthermore, an illustration of conflicting objectives could be when the maximization of an objective involves maximizing another objective that it is aimed to be minimized, for instance, when trying to minimize cost and maximize efficiency, it is most probable that the cost will increase as efficiency increases, therefore a conflict will be found and a decision will have to be made based on which of the two objectives has a greater weight of importance to the designer.

Analyzing and solving problems with conflictive objectives usually requires the role of a decision maker who can express preference relations between alternative solutions. Knowledge about this set aids the decision makers in choosing the best solution according to their preference and weights given to each of the objectives being evaluated. The Pareto-optimal set can help in reducing the design space alternatives from a feasible region on to optimal trade-offs (Yancang & Shujing, 2010)

3.4 Multiple Objective Evolutionary Algorithms

Finding the solution of problems involving several objectives that have to be satisfied simultaneously can be quite challenging. The main characteristics in multi-objective problems are that objectives are in conflict and the search space is highly complex. Many researchers have

proposed different models to obtain Pareto Optimal solutions. Such as in single objective optimization evolutionary algorithms are also being employed to solve multi-objective optimization.

Genetic Algorithm optimizes the desired variable through encoding. In comparison to biological terms, solutions arise from a set of possible genetic sequences, hence, the best solutions result from organisms that were able to survive and reproduce within their own environment. Genetic Algorithm is one of a variety of Evolutionary Algorithms (EAs) which applies techniques inspired by evolutionary biology such as inheritance, mutation and crossover. A random possible solution represented as form of data structure is generated, in technical terms of a genetic algorithm, this possible solution is recognized as a chromosome which constitutes an individual in our entire population. The data structure is arranged in a sequence of genes, string form that encodes the data characteristics that will provide the best solution evaluated by the objective function. The chromosome is completed by a sequence of genes which represents information about the individual and can be in form of bits, digits, or letters (Kumar, H. *et al*, 2010). A Genetic Algorithm simulates the best individuals inside successive generations; a set of individuals composes the entire population being evaluated at each generation. A specific methodology is followed when formulating and evaluating a genetic algorithm which is showed in the figure below.

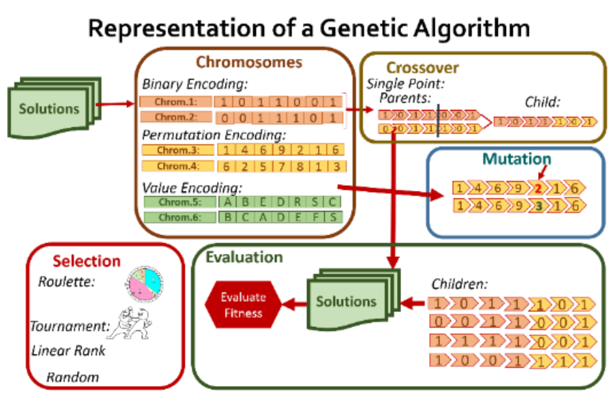

Figure 3.2: Graphic representation of genetic algorithm

For multi-objective applications, Evolutionary Algorithms follow the same process as a genetic algorithm, though several multiple objective evolutionary algorithms differ from each other in their techniques to achieve diversity in the population. For multiple objective evolutionary algorithms achieving diversity is crucial to obtain a good set of solutions. Therefore, different fitness assignment methodologies have been implemented. Pareto-based approaches are categorized into non-elistist approaches and elitist approaches.

3.2.1 Pareto-based non-elitist approaches

1. *Multiple objective GA (MOGA)* (Murata, T.; Ishibuchi, H., 1995): In this approach, an individual is assigned a rank corresponding to the number of individuals in the current population by which it is dominated. All nondominated individuals are ranked and the fitness of individuals with the same rank is averaged so that all of them are sampled at the same rate.

2. Niched Pareto GA (NPGA) (Horn, J.; Nafpliotis, N.; Goldberg, D.E, 1994): Here, Pareto dominance-based tournament selection with a sample of the population is used to determine the winner between two candidate solutions. Around ten individuals are used to determine dominance, and the nondominated individual is selected. If both individuals are either dominated or nondominated, then the result of the tournament is decided through fitness sharing.

3. Nondominated sorting GA (NSGA) (Kalyanmoy, Deb *et al,* 2002): In this approach, all nondominated individuals are classified into one category, with a fitness value proportional to the population size. This group is then removed, and the remaining population is reclassified. The process is repeated until all the individuals in the entire population are classified

3.2.2 Pareto-based elitist approaches

1. *Strength Pareto evolutionary algorithm (SPEA)* (Zitzler, Eckart et al, 2001): This algorithm is similar to others in storing the pareto-optimal solutions found so far, uses the concept of Pareto dominance and performs clustering to reduce the number of nondominated solutions stored. On the other hand it differs from others in that it combines those three techniques in a single algorithm,

the fitness is determined from the solutions stored, avoiding dominance in population, and all solutions in the external pareto set are considered for the selection.

2. *Strength Pareto evolutionary algorithm 2(SPEA2)* (Zitzler, Eckart et al, 2001): In SPEA2, a technique is developed to avoid the situation where individuals dominated by the same archive members have the same fitness values. Here, both the main population and the archive are considered to determine the fitness values of the individuals. A different scheme has been adopted that prevents the loss of the boundary solutions during the updating of the archive. Diversity is maintained by using a density-based approach on the kth nearest neighbor.

3. *Pareto archived evolutionary strategy (PAES) (*Knowles, J., Corne, D., 1999): Knowles and Corne suggested a simple MOEA using a single-parent, single-child evolutionary algorithm which is similar to a (1+1) evolutionary strategy. A Pareto Archive Evolution Strategy (PAES) has two main objectives in mind. One objective is that the algorithm should be strictly confined to local search, for instance it should use a small change operator, allowing it to move from a current solution to a close by neighbor. The second objective is for the algorithm to be a true Pareto optimizer by treating all nondominated solutions as having equal value. In some cases this is problematic because when comparing pair of solutions neither one will be dominated by the other. PAES overcomes this problem by maintaining an archive of previous nondominated solutions, then this archive is used as the form of estimating true dominance ranking of pair of solutions. Knowles concluded that when used to address the multiobjective form of the offline routing problem, (1+1)-PAES provided result extremely competitive with a MOEA in each case.

4. *Pareto envelope-based selection algorithm (PESA) (*Corne, David W. *et al*, 2000): In this approach, a smaller internal (or primary) population and a larger external (or secondary) population are used. PESA uses the same hypergrid division of objective space adopted by PAES to maintain diversity. Its selection mechanism is based on the crowding measure used by the hypergrid. This same crowding measure is used to decide which solutions to introduce into the external population.

5. *Pareto envelope-based selection algorithm-II (PESA-II)* (Corne, David W. et al, 2011): PESA-II is the revised version of PESA in which region-based selection is proposed. In case of region-based selection, the unit of selection is a hyperbox instead of an individual. The procedure of selection is to select (using any of the traditional selection techniques) a hyperbox and then randomly select an individual within such hyperbox.

6. *Elitist non-dominated sorting GA (NSGA-II)* (Kalyanmoy, Deb *et al,* 2002): NSGA-II was proposed to eliminate the weaknesses of NSGA, especially its nonelitist nature and specification of the sharing parameter. Here, the individuals in a population undergo nondominated sorting as in NSGA, and individuals are given ranks based on this. A new selection technique called crowded tournament selection is proposed where selection is done based on crowding distance (representing the neighborhood density of a solution). To implement elitism, the parent and child population are combined and the nondominated individuals from the combined population are propagated to the next generation.

Chapter 4: Centralized Carrier Collaboration and Multi-Hub Location Problem

The highest percentage of greenhouse gas emissions is attributed to road transportation, therefore thorough effort is given in achieving an optimized transportation system. Collaboration among partners in the transportation seems to be the solution to this high GHG emissions, as it can also reduce the number of necessary trips, minimizing environmental impacts and increasing efficiency of the system. Extensive literature can be found addressing the environmental aspect of collaboration in road transportations, although scarce literature is found addressing specifically the LTL industry. Literature vary on the utilization of different methodologies from mathematical approaches such as linear, integer and mixed integer programming. Some other successfully implemented metaheuristic approaches to optimize either one or multiple objectives. The first section will describe approaches taken to address collaboration in transportation in general, whereas on the second section, more specifically will review literature on LTL industry.

4.1 Literature Review

4.1.1 Collaboration in road transportation addressing environmental impacts

There are several ways to cooperate: carriers can collaborate with each other (horizontal cooperation), shippers can collaborate among themselves (horizontal cooperation), and carriers and shippers can also collaborate (vertical cooperation). Ballot E. & Fontane, (2010) proposed a concept of logistical network pooling achieving a saving of at least 25% of CO2 emissions from pooled networks versus current setup. Demonstrated that vertical supply chain optimizations can still be improved by horizontal collaboration. The need of defining strategies for collaboration lead Xu, (2013) to identify two organizational forms of horizontal logistics collaboration; centralized and decentralized. Later, Bernabeu *et al*, (2015) contributed to the subject by analyzing the importance of horizontal transportation for small and medium companies based on well-known Multi-depot vehicle routing problem with the assumption that carriers and shippers were directly

controlled by the same companies, giving an ideal scenario of collaboration where disagreements and own interests are eliminated. Data collection for emissions consisted of a distance-based method making use of a fuel conversion factor. Experimental results provided a noticeable reduction in expected costs as well as in terms of greenhouse gas emissions for the horizontal cooperation strategy.

4.1.1.1 Mathematical Approaches

Mathematical models are frequently used to study real world phenomena that are not susceptible to analytic techniques alone, and to investigate the relationships among parameters that affect the functioning of complex processes. Likewise, mathematical models have been used to study the functioning of collaboration in road transportation.

A mixed integer linear programming model coded in ILOG's OPL 6.3 software with CPLEX 12.1 was presented by Pan, S. *et al* (2013), implementing the concept of network pooling of Ballot E. & Fontane (2010) on freight industry. This model explored the effect of pooling supply chains' networks on reducing CO_2 emissions from transport with two possible modes, i.e. road referred to Heavy Duty Vehicle and rail considering only the electrically powered locomotive, in the context of a national distribution network of two major French retailer chains. The results show that the approach of merging supply chains significantly reduces CO_2 emissions from transport. Besides the reduction of CO_2, it is also important to indicate that the number of transport paths falls because of the network pooling. Okan"Orsan"Ozener *et al*, (2007) investigated the potential of collaborative opportunities in truckload transportation by developing optimization models to determine the maximum benefit that can be derived from collaborating. The Multi-Carrier Lane Covering Problem was modeled into an integer programming problem. Lin, D. Y. & Ng, K. H, (2012) transformed a stochastic mixed integer programming problem of collaboration between carriers into an integer programing model. The results provided evidence that the collaborative method can effectively reduce emissions in a freight network by 3-20%.

4.1.1.2 Heuristic and Meta-Heuristic Approaches

A metaheuristic is formally defined as an iterative generation process which guides a subordinate heuristic by combining intelligently different concepts for exploring and exploiting the search space, learning strategies are used to structure information in order to find efficiently near-optimal solutions (Osman and Laporte, 1996). Researchers have also found suitable optimizing transportation systems through the use of meta-heuristic algorithms.

Addressing collaboration in road transportation, Sadegheigh *et al*, (2011) developed a genetic algorithm to study the behavior of global supply management and collaborative network design. The algorithm showed to be successful for transportation problems obtaining reductions through iterations. His studies showed advantageous results that can lower carbon emissions and still provide parties competitive advantages. Okan”Orsan”Ozener *et al*, (2007) use heuristic approaches to solve a multi-carrier lane covering problem. The first heuristic approach relaxes the precision by specifying a relative optimality criterion, whereas the second exploits the structure of the solution by fixing integral flows along cycles before starting branch-and-bound process.

Some more literature found on transportation collaboration in supply chain to reduce carbon emissions was done by Jaegler, A., & Burlat, P., (2012), proposing a new vehicle routing variant to minimize greenhouse gases. One of the most common variants is vehicle restricted by capacity, and in one analysis the VRP with time windows considers now environmental impacts in the search solution. To complete this four elements were considered: energy requirement, mathematical formulation, use of scatter search metaheuristic, and cooperative game approach as an option. Based on the results found conclusions were that environmental impacts of the cost of vehicle routing are relate to rate of vehicle use, and the method proposed enables customer requirements to be satisfied while also decreasing the impact on the environment. Lin, D. Y. & Ng, K. H, (2012) investigated to see if collaboration between carriers can reduce the environmental issues of freight movement in a carbon constrained business context. This was solved using tabu search algorithm with Monte Carlo bounding techniques.

4.1.1.3 Multiple Objective Optimization Approaches

Jemai J. *et al*, (2012), proved the success of evolutionary algorithms on transportation problems by solving a bi-objective green vehicle routing problem. The objectives were to minimize the total traveled distance and the CO2 emissions. The results of the NSGA-II showed that the algorithm obtain good results and prove the explicit interest grant to emission minimization objective. In contrast to past research, collaboration in transportation was not considered in his analysis.

Besides mathematical models and heuristic approaches, some others prefer simulation models as De Mello, P. F. B., & Frayret, J. M., (2014) who extended the literature from focusing only on costs or emissions into focusing on a complete sustainability assessment. They present a freight transportation model based on the resource sharing methodology. The analysis considers the road transportation of semi-trailers through a network of hubs in which they wait for their next route segment. Sustainability is assessed in terms of logistic performance, drivers working condition, and environmental impact. An agent based simulation model was developed by making use of the software Netlogo. Results indicated that resource sharing can significantly improve all performance, social and environmental.

4.1.2 Collaboration and Hub and Spoke Network for the small to medium-sized LTL

Having efficient road networks becomes very important since having an optimal vehicle routing can reduce the number of trucks and utilize the network better by reducing vehicle movements. Going deeper into the study of collaboration within the transportation industry, LTL carrier collaboration is a relatively unexplored concept within the ground freight domain. Nonetheless, up to date, most, if not all, literature focus mainly on the economic aspect, leaving a side the environmental aspect. This section provides an in view of current research on the LTL carrier collaboration within a hub and spoke transportation network. Although not explicitly as collaboration, Zhang *et al*., (2007) introduced the concept of hybrid hub-and-spoke for a single LTL carrier trying to minimize transportation costs. In their work, hybrid refers to the addition of

direct routes to a pure hub-and-spoke system. The concept of hybrid hub-and-spoke is a relatively new one in the LTL industry.

4.1.2.1 Mathematical and Exact Approaches

Recently, Hernandez and Peeta, (2010, 2011b) addressed a time-dependent, centralized multiple-carrier collaboration problem (TD-MCCP) with a static and dynamic context for the small to medium-sized less-thantruckload (LTL) industry. The TD-MCCP is modeled as a binary (0–1) multicommodity minimum cost-flow problem formulation for two rate-setting behavioral cases and solved with a branch-and-cut algorithm.

4.1.2.2 Heuristic and Metaheuristics Approaches

Other researchers have approached hub-and-spoke LTL problems through the use of genetic algorithms for instance Cunha and Silva, (2007) whom focused on configuring a hub-and-spoke network for an LTL trucking company in Brazil. They sought to determine the number of consolidation terminals (hubs), their locations, and the assignment of the spokes to the hubs, while aiming to minimize the total cost. In addition, Zhang *et al*., (2007) formulated a hybrid hub-and spoke for a single LTL carrier into a combinatorial problem and solved it using a genetic algorithm methodology.

Later, Hernandez *et al.*, (2012), analyzed centralized carrier collaboration multi-hub location problem (CCCMLP) for the small to medium-sized less-than-truckload industry. In the CCCMLP, a central entity (e.g., a third-party logistics firm) seeks a set of collaborative consolidation transshipment hubs to establish a hybrid collaborative hub-and-spoke system that minimizes the total collaborative costs for the set of collaborating carriers. The CCCMLP was formulated as a variant of the P-hub location problem, which is NP-hard and solved with Lagrangian relaxation. The results indicate that larger expected profit margins from collaborative carriers applying revenue-generating behavior would increase the likelihood of collaboration by carriers. As the network size increases, the effect of hybrid hub location costs drops.

4.2 State-of-the-art

As previously mentioned, there is a very active research area in transportation, but there is still limited literature addressing environmental impacts, some studies are beginning to consider CO_2 emissions into their studies but they lack of a complete analysis of how the emissions are impacting the environment and society. The scarcity of research increases when analyzing deeper hub-and-spoke applications. Until today, no research has been done on CCCMLP within LTL carrier collaboration that considers the economic aspect and at the same time the minimization of environmental impacts. In relation to this problem, the majority of the literature addresses collaboration without consideration of multi-hub location or addresses the multi-hub location from the context of a single LTL carrier. In addition, the hybrid hub-and-spoke system will be used as a set of collaborative consolidation transshipment hubs from a current point-to-point network structure, as employed by Hernandez *et al.*, (2011a). Simply speaking, a hub-and-spoke system is formed without costly investments in new facilities. A study similar to this one was conducted by Hernandez *et al.*, (2011a). From a planning perspective, the CCCMLP represents a starting point for studying the effects of rate setting by collaborative carriers in a centralized carrier multi-hub collaborative network.

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

4.3 Model Formulation

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

$$C_{ijkl} = C_{ik} + \delta C_{kl} + C_{lj} \tag{4.1}$$

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

$$P_{kq} = \vartheta_{kq} + \phi_k \tag{4.2}$$

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}.

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.

For the analysis of this problem, we have determined the following constraints to be relevant to the problem. Constraint (4.3) 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, constraint (4.4) impedes the programing of more than one different route between two points in the system. Constraints (4.5, 4.6) 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. Constraint (4.7) 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. γ

denotes the profit margin expected by a company in order to participate in the collaboration. Constraints set (4.8-4.10) constraints variables X, Y, and V into the binary space.

$$\sum_k X_k = p \tag{4.3}$$

$$\sum_k \sum_l Y_{ijklq} + V_{ijq} = 1 \quad \forall i, j, q \tag{4.4}$$

$$\sum_l Y_{ijklq} \leq X_k \tag{4.5}$$

$$\sum_k Y_{ijklq} \leq X_l \tag{4.6}$$

$$C_{ijkl} Y_{ijklq} \leq V_{ijq}\left(1 - V_{ijq}\right)(1 - \gamma) \quad \forall i, j, k, l, q \tag{4.7}$$

$$X_k \in \{0,1\} \tag{4.8}$$

$$Y_{ijklq} \in \{0,1\} \tag{4.9}$$

$$V_{ijq} \in \{0,1\} \tag{4.10}$$

The analysis of this problem will be addressed through single objective optimization in the first case, to facilitate observing the behavior of each objective separately. Then, these same objectives will be optimized by a bi-objective evolutionary algorithm, minimizing both objectives simultaneously to provide a broader view and comprehension of the system.

The first objective function shown in Equation 4.11, seeks a set of candidate hybrid collaborative consolidation hubs as to minimize the total transportation collaborative costs in a supply chain.

$$\min \sum_i \sum_j \sum_k \sum_l \sum_q C_{ijkl} d_{ijq} Y_{ijklq} + \sum_i \sum_j \sum_q W_{ijq} d_{ijq} V_{ijq} + \sum_k \sum_q P_{kq} X_k \tag{4.11}$$

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

shipmentC_{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 directly W_{ijq}, the collaborative carrier demandd_{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 hub P_{kq}, and the X_k (the decision on whether a collaborative facility is located).

The second objective function considered in the present study (Equation 4.12) considers the minimization of the Global Warming Potential.

$$\min \sum_i \sum_j \sum_k \sum_l GWPCC_{ijkl} d_{ijq} Y_{ijklq} + \sum_i \sum_j \sum_q GWPNCCE_{ijq} d_{ijq} V_{ijq} \tag{4.12}$$

Similarly as in the previous cost function, the objective consists of two main terms, the first term represents the Global Warming potential associated to the carrier collaborative, the second part represents the total Global Warming Potential associated with carriers not collaborating and shipping directly.

Equations (4.11 and 4.12) subject to constraints (4.3) through (4.10) represents the mathematical formulation of the centralized carrier collaborative multi-hub location problem.

4.4 Optimization Methodology and Model Development

The optimization methodology used in this analysis is the Genetic Algorithm for the single objective case, and the multiple objective evolutionary algorithm for the bi-objective analysis. In this section, a thorough description of the Genetic Algorithm will be provided, nevertheless, the multi-objective evolutionary algorithm will be described on Chapter 5.

4.4.1 The Genetic Algorithm

As a promising intelligent algorithm, the Genetic Algorithm was developed during the 1970s by an American scientist, John Holland. As a metaheuristic probabilistic search algorithm based on the mechanics of natural selection and natural genetics. Evolution in a genetic algorithm is considered as a process in which possible solutions evolve into a better solution through generations. It is widely used in computer science and operations research categorized as a global search heuristics that will enable finding exact or approximate solutions to optimization and search problems.

For the developments of this genetic algorithm, a random population is first generated, after which it will proceed to evaluate the objective function value for each chromosome in the population and rank them according to the result. Elitism is used to determine the parents of the next population, which will generate the next population through means of a random crossover. The process of evaluating, ranking, and crossover will be repeated a number of iterations until a satisfactory result is found. A flow chart of the procedure is presented on figure 4.1, afterwards, a thorough explanation for each of the steps will be provided.

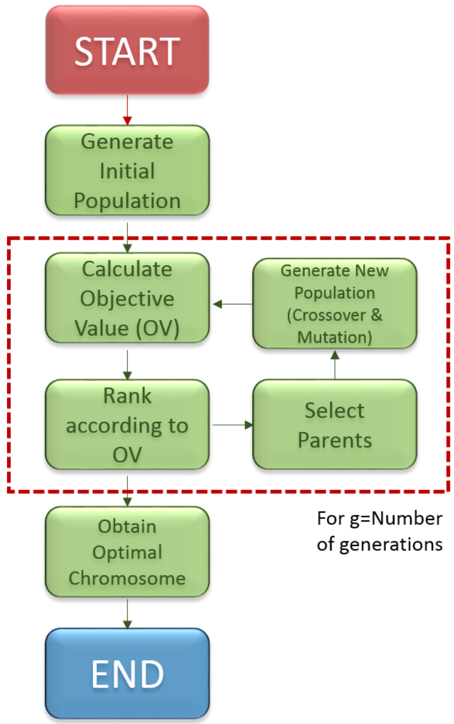

Figure 4.1: Genetic Algorithm Flow Chart

4.4.2 The Chromosome

The first step to initiate the algorithm is to model a chromosome. The chromosome is modeled as a set of parameters which define a possible solution. When searching for the structure of the chromosome, encoding becomes an important factor which decision depends at the most on the problem being solved. Type of encoding include: binary, integer, real and characters. The chromosome for this Genetic Algorithm consists of a binary vector of length n, n being the number of nodes in the network. Each of these will correspond to a node in the network. The value for each gene will be one (1) if such node is to be transformed into a hub, or zero (0) if otherwise. Figure 4.2 illustrates how the chromosome or possible solution was modeled in the algorithm.

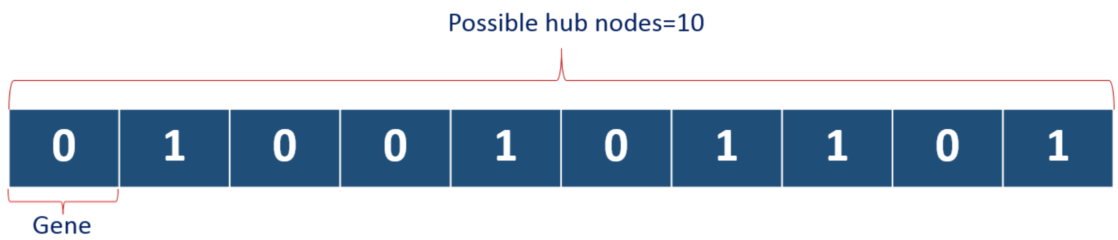

Figure 4.2: Representation of a chromosome (possible solution)

4.4.3 The initial generation

After having modeled our possible solution, a random initial population of individual chromosomes is generated. For this, it will create a zero matrix of size *m x n*, where *m* is the size of the population. Following this, the matrix will randomly select points in the individual chromosomes, until each of them has *p* elements equaling one, thus meeting the constraint for number of hubs allowed.

4.4.4 The Fitness Function

The Fitness Function is then calculated for each individual in the population by the utilizing an optimized cost matrix S. This matrix starts being the same as the cost matrix W given as a superposition of tables 4 and 5, and changes the values in the matrix to those in the Collaborative cost matrix C if $\gamma S_{ijq} > C_{ijkl}$ for all values of l, k. The fitness value, which is equal to the objective function, is determined as a point product of the Optimized Cost Matrix S and the Demand Matrix D, added to the Hubs cost, given as a scalar multiplication of the hub creation costs for each carrier times the first n values of the chromosome.

4.4.5 The Crossover and Mutation Functions

For the crossover function, it was decided that a random process would be more likely to produce results that vary more between each other. A set of parents is selected from the population by elitism, as the top *x* percentage of the population. As so, the algorithm selects a random point in two random chromosomes amongst the parent population, taking the first part of the new

chromosome from the first parent and the second from the second parent, this process is illustrated in figure 4.3. Then, if a random number between zero and one is smaller than the specified mutation factor, a random point in the chromosome is changed from zero to one or from one to zero, which is represented in figure 4.4. An adjuster for the number of chromosomes then randomly selects points in the chromosome to change to one or zero if the number of hubs is different to *p*, the selected number of hubs specified in equation 3.

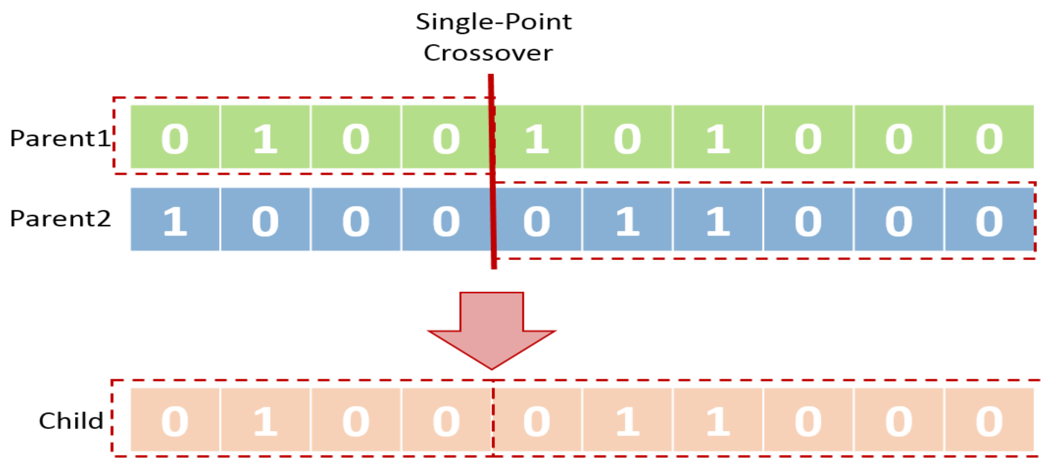

Figure 4.3: Crossover Representation

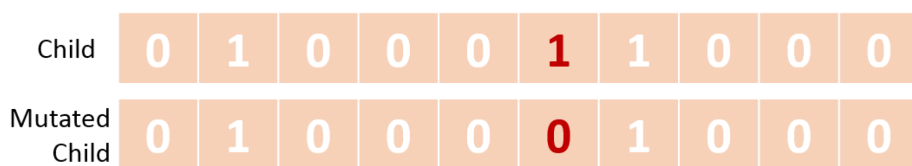

Figure 4.4. Mutation Representation

4.5 Single Objective Case Studies Results

4.5.1 Case Study 1: Cost Minimization

Using data from Tables 1-6, located in APPENDIX A the savings obtained by implementing a 10 node hub-and-spoke network and establishing collaboration between two carriers with varying degrees of expected profit margins γ were determined. The data considered

for the analysis consists on assumed demand for carrier 1(Table A.1) and carrier 2 (TableA.2), individual hub establishing cost for each carrier (Table A.3), non-collaborative shipping costs for carrier 1 (Table A.4) and 2 (Table A.5), and collaborative shipping costs for carrier1 and 2 (Table A.6) For the analysis of collaboration between carriers a δ of 10% and values for γ of 9%, 18%, 36%, 48%, 60%, 72%, 84%, and 96% were considered. On the other hand, for the genetic algorithm parameters, an initial population of 100 is considered, running the algorithm for 200 iterations with a 10% elitism, a mutation factor of 10%.

The results for the case of considering only cost minimization are provided in the following tables for different values of *p* number of hubs in terms of γ, total cost, number of collaborative and non-collaborative routes, percentage of routes picked as collaborative, and relative saving percentage. Furthermore, a summary of the results is given in figure 4.5.

Table 4.1 Cost results with respect to gamma when only 1 hub is considered

p-HUBS=1	**Hub Selected**	**Cost**	**Col. Routes**	**Non-Col. Routes**	**% Col. Routes**	**% Savings**
Total Operations Cost with γ=9%	3	$7,365,653.50	88	92	48.90%	56.50%
Total Operations Cost with γ=18%	3	$7,389,737.50	87	93	48.30%	56.40%
Total Operations Cost with γ=36%	3	$7,579,774.50	81	99	45.00%	55.20%
Total Operations Cost with γ=48%	3	$8,503,193.50	79	101	43.90%	49.80%
Total Operations Cost with γ=60%	3	$10,524,897.00	71	109	39.40%	37.90%
Total Operations Cost with γ=72%	3	$16,936,922.50	27	153	5%	0.00%
Total Operations Cost with γ=84%	3	$16,936,922.50	9	171	5%	0.00%
Total Operations Cost with γ=96%	3	$16,936,922.50	9	171	1.60%	0.00%

Table 4.2: Cost results with respect to gamma when 2 hubs are considered

p-HUBS=2	**Hub Selected**	**Cost**	**Col. Routes**	**Non-Col. Routes**	**% Col. Routes**	**% Savings**
Total Operations Cost with γ=9%	3,5	$5,698,410.10	88	92	48.90%	66.40%
Total Operations Cost with γ=18%	3,5	$5,710,847.10	87	93	48.30%	66.30%
Total Operations Cost with γ=36%	3,5	$5,781,077.10	81	99	45.00%	65.90%

Total Operations Cost with γ=48%	3,5	$6.102.394.10	79	101	43.90%	64.0%
Total Operations Cost with γ=60%	3,5	$7,029,645.10	71	109	39.40%	59.50%
Total Operations Cost with γ=72%	4,6	$12,233,957.40	27	153	15.0%	27.80%
Total Operations Cost with γ=84%	3,4	$14,019,024.20	9	171	5.0%	17.20%
Total Operations Cost with γ=96%	1,4	$16,041,353.90	9	171	5.0%	5.30%

Table 4.3: Cost results with respect to gamma when 3 hubs are considered

p-HUBS=3	**Hub Selected**	**Cost**	**Col. Routes**	**Non-Col. Routes**	**% Col. Routes**	**% Savings**
Total Operations Cost with γ=9%	4,5,7	$4,046,494.30	88	92	48.90%	76.10%
Total Operations Cost with γ=18%	4,5,7	$4,055,681.10	87	93	48.30%	76.10%
Total Operations Cost with γ=36%	3,4,5	$4,112.193.40	81	99	45.00%	75.70%
Total Operations Cost with γ=48%	3,4,5	$4,202,134.00	79	101	43.90%	75.2%
Total Operations Cost with γ=60%	3,5,9	$4,707,665.70	71	109	39.40%	72.20%
Total Operations Cost with γ=72%	3,5,9	$6,997,032.20	27	153	15.0%	58.70%
Total Operations Cost with γ=84%	3,4,10	$10,985,652.40	9	171	5.0%	35.10%
Total Operations Cost with γ=96%	4,7,10	$14,612,113.30	9	171	5.0%	13.7%

Table 4.4: Cost results with respect to gamma when 4 hubs are considered

p-HUBS=4	**Hub Selected**	**Cost**	**Col. Routes**	**Non-Col. Routes**	**% Col. Routes**	**% Savings**
Total Operations Cost with γ=9%	1,4,5,8	$3,299,158.60	88	92	48.90%	80.50%
Total Operations Cost with γ=18%	1,4,5,8	$3,302,644.60	87	93	48.30%	80.50%
Total Operations Cost with γ=36%	1,4,5,8	$3,341,486.00	81	99	43.90%	80.30%
Total Operations Cost with γ=48%	1,5,8,9	$3,469,788.90	79	101	43.90%	79.5%
Total Operations Cost with γ=60%	1,3,5,9	$3,649,780.40	71	109	39.40%	78.50%
Total Operations Cost with γ=72%	4,5,6,8	$5,065,719.40	27	153	15.0%	70.10%
Total Operations Cost with γ=84%	1,4,8,10	$8,971.205.80	9	171	5.0%	47.0%
Total Operations Cost with γ=96%	1,4,7,10	$13,144,337.70	9	171	5.0%	22.4%

Table 4.5: Cost results with respect to gamma when 5 hubs are considered

p-HUBS=5	Hub Selected	Cost	Col. Routes	Non-Col. Routes	% Col. Routes	% Savings
Total Operations Cost with γ=9%	1,4,5,8,9	$2,803,372.30	88	92	48.90%	83.40%
Total Operations Cost with γ=18%	1,4,5,8,9	$2,806,858.30	87	93	48.30%	83.40%
Total Operations Cost with γ=36%	1,4,5,8,9	$2,845,699.70	81	99	45.00%	83.20%
Total Operations Cost with γ=48%	1,3,4,5,8	$2,921,397.90	79	101	43.90%	82.8%
Total Operations Cost with γ=60%	1,3,5,8,9	$2,973,975.00	71	109	39.40%	82.40%
Total Operations Cost with γ=72%	2,5,6,8,9,	$3,633,435.30	27	153	15.0%	78.50%
Total Operations Cost with γ=84%	1,4,8,9,10	$6,755,821.10	9	171	5.0%	60.1%
Total Operations Cost with γ=96%	1,4,7,8,10	$11,479,722.50	9	171	5.0%	32.2%

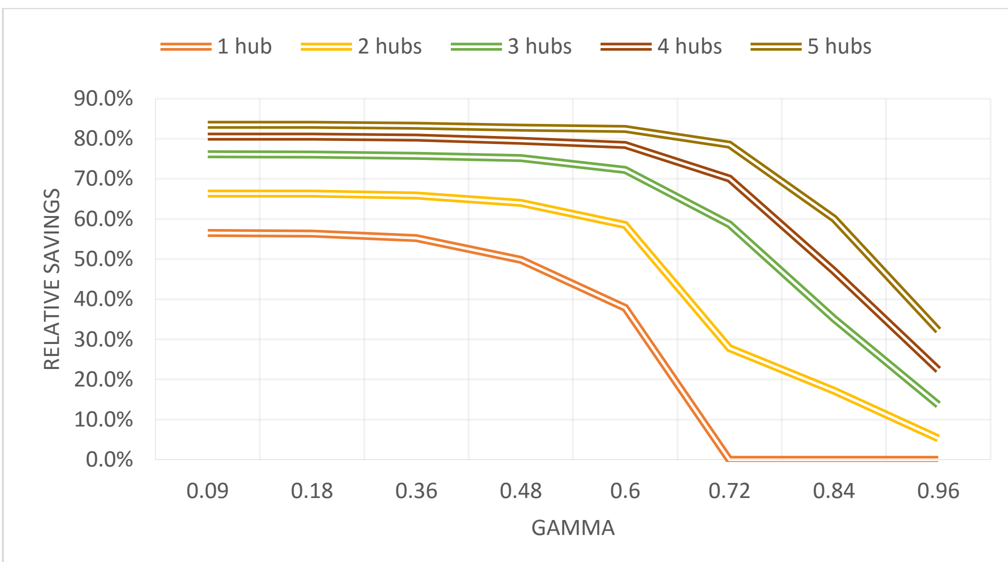

Figure 4.5. Summary of Results from tables

4.5.2 Case Study 2: Global Warming Potential Minimization

As stated earlier the second objective function (equation 4.12) considers the minimization of the Global Warming Potential Impact of the hub and spoke network. Global Warming Potential is defined as a relative measure of how much heat is trapped in the atmosphere by a greenhouse gas, compared to the heat trapped by a similar mass of CO_2. For this case study, a Life Cycle

Analysis (LCA) was performed in order to calculate the impacts that the transportation vehicle of this model emits. The GWP was calculated utilizing GaBi Life Cycle Assessment Software, by using a Diesel truck with a 12.4 ton capacity from its database.

Furthermore, using the demands from tables 1 & 2 on APPENDIX A, and the estimated emissions data from tables B.2 through B.4 on APPENDIX B the total Global Warming Potential (100kg CO_2 equivalent) caused by the trucks traveling through the network was estimated. In the genetic algorithm parameters, an initial population of 100 was considered, running it for 200 iterations with a 10% elitism, a mutation factor of 10%, and considering a δ of 10% and values for γ of 9%, 18%, 36%, 48%, 60%, 72%, 84%, and 96%. The results of these analysis for different values of *p*, number of hubs, will be provided in the following tables in terms of γ, Global Warming Potential, number of collaborative and non-collaborative routes, percentage of routes picked as collaborative, and relative saving percentage. Furthermore, a summary of the results is given in Figure 4.6.

Table 4.6 Global Warming results with respect to gamma when only 1 hub is considered

p-HUBS=1	Hub Selected	Global Warming Potential (100kg CO_2-equivalent)	Col. Routes	Non-Col. Routes	% Col. Routes	% Savings
Total GWP with γ=9%	3	3267.059	78	102	43.30%	38.40%
Total GWP with γ=18%	3	3295.969	78	102	48.30%	37.90%
Total GWP with γ=36%	2	3574.27	70	110	38.90%	32.60%
Total GWP with γ=48%	3	4488.766	64	116	35.60%	15.40%
Total GWP with γ=60%	3	5305.738	24	156	13.30%	0.0%
Total GWP with γ=72%	5	5305.738	12	168	6.7%	0.00%
Total GWP with γ=84%	7	5305.738	6	174	3.3%	0.00%
Total GWP with γ=96%	7	5305.738	0	180	0.0%	0.00%

Table 4.7 Global Warming results with respect to gamma when 2 hubs are considered

p-HUBS=2	Hub Selected	Global Warming Potential (100kg CO_2-equivalent)	Col. Routes	Non-Col. Routes	% Col. Routes	% Savings
Total GWP with γ=9%	3,5	2515.663	78	102	43.30%	52.6%
Total GWP with γ=18%	3,5	2525.737	78	102	43.30%	52.4%
Total GWP with γ=36%	3,5	2612.515	70	110	38.90%	50.8%
Total GWP with γ=48%	5,9	3105.245	64	116	35.60%	41.5%
Total GWP with γ=60%	4,6	3935.305	24	156	13.30%	25.8%
Total GWP with γ=72%	3,4	4307.896	12	168	6.7%	18.8%
Total GWP with γ=84%	4,6	4815.957	6	174	3.3%	9.2%
Total GWP with γ=96%	3,5	5305.738	0	180	0.0%	0.00%

Table 4.8 Global Warming results with respect to gamma when 3 hubs are considered

p-HUBS=3	Hub Selected	Global Warming Potential (100kg CO_2-equivalent)	Col. Routes	Non-Col. Routes	% Col. Routes	% Savings
Total GWP with γ=9%	4,5,7	1735.755	78	102	43.30%	67.3%
Total GWP with γ=18%	4,5,7	1742.015	78	102	43.30%	67.2%
Total GWP with γ=36%	5,7,9	1803.993	70	110	38.90%	66.0%
Total GWP with γ=48%	3,9,10	1994.051	64	116	35.60%	62.4%
Total GWP with γ=60%	4,5,7	2395.949	24	156	13.30%	54.8%
Total GWP with γ=72%	4,5,7	3022.115	12	168	6.7%	43.0%
Total GWP with γ=84%	4,5,7	4212.987	6	174	3.3%	20.6%
Total GWP with γ=96%	1,5,8	5305.738	0	180	0.0%	0.00%

Table 4.9 Global Warming results with respect to gamma when 4 hubs are considered

p-HUBS=4	Hub Selected	Global Warming Potential (100kg CO_2-equivalent)	Col. Routes	Non-Col. Routes	% Col. Routes	% Savings
Total GWP with γ=9%	1,4,5,8	1388.61	78	102	43.30%	73.8%
Total GWP with γ=18%	1,4,5,8	1402.479	78	102	43.30%	73.6%

Total GWP with γ=36%	1,4,5,8	1420.479	70	110	38.90%	73.2%
Total GWP with γ=48%	1,4,5,8	1508.217	64	116	35.60%	71.6%
Total GWP with γ=60%	4,5,6,7	1788.755	24	156	13.3%	66.3%
Total GWP with γ=72%	1,4,8,10	2477.865	12	168	6.7%	53.3%
Total GWP with γ=84%	4,5,7,8	3505.531	6	174	3.3%	33.9%
Total GWP with γ=96%	2,5,6,10	5305.738	0	180	0.0%	0.0%

Table 4.10 Global Warming results with respect to gamma when 5 hubs are considered

p-HUBS=5	**Hub Selected**	**Global Warming Potential (100kg CO_2-equivalent)**	**Col. Routes**	**Non-Col. Routes**	**% Col. Routes**	**% Savings**
Total GWP with γ=9%	1,4,5,8	1388.61	78	102	43.30%	73.8%
Total GWP with γ=18%	1,4,5,8	1402.479	78	102	43.30%	73.6%
Total GWP with γ=36%	1,4,5,8	1420.479	70	110	38.90%	73.2%
Total GWP with γ=48%	1,4,5,8	1508.217	64	116	35.60%	71.6%
Total GWP with γ=60%	4,5,6,7	1788.755	24	156	13.3%	66.3%
Total GWP with γ=72%	1,4,8,10	2477.865	12	168	6.7%	53.3%
Total GWP with γ=84%	4,5,7,8	3505.531	6	174	3.3%	33.9%
Total GWP with γ=96%	2,5,6,10	5305.738	0	180	0.0%	0.0%

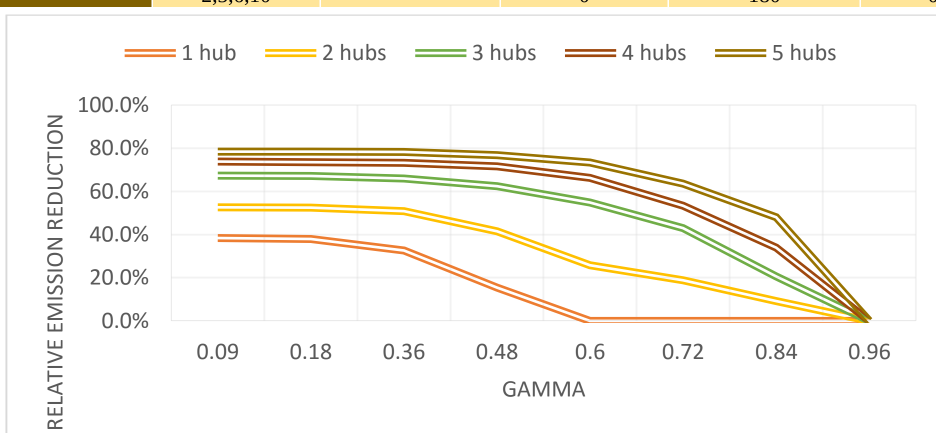

Figure 4.6. Summary of results for GWP

4.5.3 Comparison Analysis

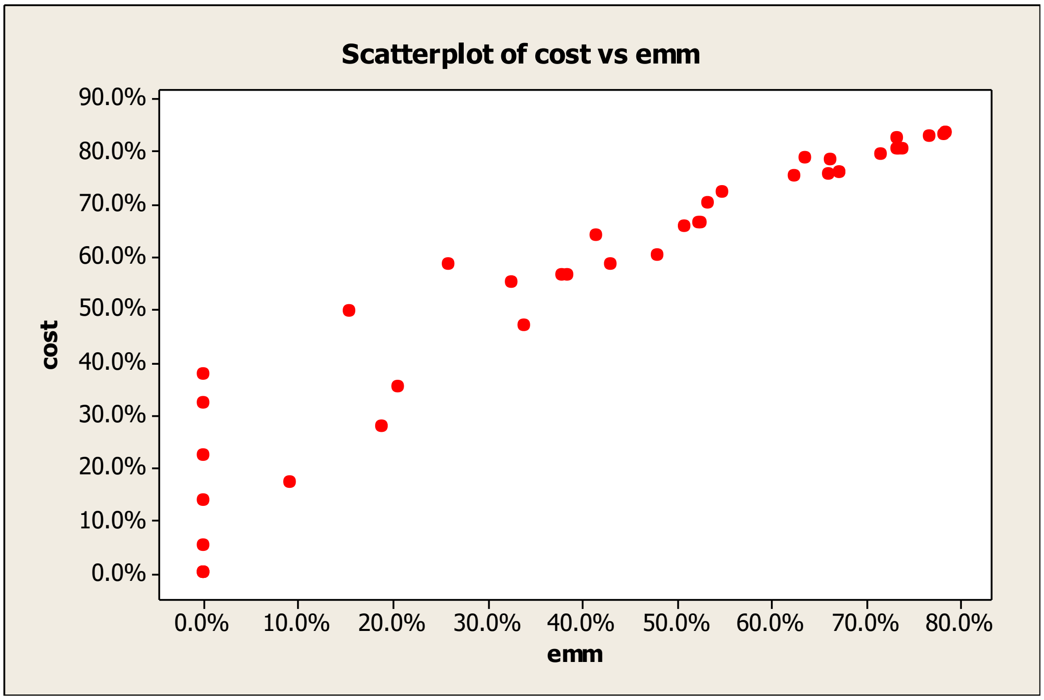

Figure 4.7: Relation between cost savings and emission reduction percentages

It can be stated from Figure 4.7, there is a relationship between the savings in money and the reduction in greenhouse gas emissions. This is due to the decrease in trucks needing to be used in a hub and spoke network with collaboration as opposed to those used in a point-to-point non-collaborative network. Another finding is that, in both studies, as the number of *p* possible hubs in the network increases, so does the decrease in cost and emissions, respectively. Inversely, the increase in the minimum expected savings brought a decrease in the actual savings.

4.6 Bi-Objective Case Study Results

For the bi-objective case study, the same data provided for the two past case studies was utilized. Considering simultaneously the non-collaborative shipping costs for carrier 1 (Table A.4) and 2 (Table A.5), and collaborative shipping costs for carrier1 and 2 (Table A.6.), as well as the non-collaborative GWP for carrier 1(Table B.2) and 2 (Table B.3), and collaborative GWP for

carrier 1 and 2 (Table B.4). For this analysis, the *p* number of hubs was implemented as an upper limit instead of a fixed number of hubs, allowing the algorithm to decide the best number of hubs. The analysis of a centralized carrier collaboration in a 10 node hub and spoke network using a multi-objective evolutionary algorithm reflected the performance of it by effectively obtaining the Pareto Optimal solutions. The algorithm was run for 20 generations with a population size of 250, elitism percentage of 10% and mutation factor of 10%. Different analysis of the network were analyzed evaluating the case where the limit of hubs is specified, in addition to the case in which no limit is considered regarding the hubs to establish, results are given in Table 4.11.

Table 4.11: MOEA's Pareto Optimal set per limit of hubs

	No of Hubs	Nondominated Solutions										Cost	GWP	Time
MOEA	0	0	0	0	0	0	0	0	0	0	0	16900243	5305.738	3.078
	1	0	0	1	0	0	0	0	0	0	0	7365654	3267.059	7.308
	2	0	0	1	0	1	0	0	0	0	0	5698410	2515.663	9.677
	3	0	0	0	1	1	0	1	0	0	0	4046494	1735.755	14.16
	4	1	0	0	1	1	0	0	1	0	0	3299159	1388.61	20.204
	5	1	0	0	1	1	0	0	1	1	0	2803372	1142.27	26.108
	6	0	0	0	1	1	1	1	1	1	0	2342975	913.1721	30.687
	7	1	0	1	1	1	0	1	1	1	0	1911589	701.3547	35.793
	8	1	1	1	1	0	0	1	1	1	1	1602500	535.3647	42.35
		1	1	1	1	1	0	1	1	1	0	1595074	539.2624	
	9	1	1	1	1	0	1	1	1	1	1	1331887	393.5511	45.763
		1	1	1	1	1	1	1	1	1	0	1324272	397.3637	
	10	1	1	1	1	1	1	1	1	1	1	1106881	269.9526	48.891
	No limit	1	1	1	1	1	1	1	1	1	1	1106881	269.9526	47.449

Based on the information presented, when no hubs are utilized in the transportation network, the cost of transportation is \$16,900,243 and the global warming potential impact is 5305.738CO2-equivalent. On the other hand, when hub establishment is considered with a collaboration discount (δ) of 0.1000 and with a 0.0900 profit margin expected (γ), the transportation cost lowers significantly by a 56.4% starting with a transportation cost of 7365654 when considering up to one hub, and continues to reduce when adding number of hubs to be established, ending up with a cost of 1106881 when all nodes in the network are established as hubs. The analysis considered the limit of h number of hubs, to be set as the upper limit of hubs, for instance when analyzing h=4 hubs the study considered the option of having none and up to 4

hubs in the transportation network. For the network being limited from 1-7 hubs, only one optimal solution was found and it considered opening all possible hubs. Although for hubs 8 and 9, the solutions also considered opening the upper limit which was 8 and 9 hubs, although for this studies the Pareto optimal set consisted on two solutions. Moreover, when no limit is given to the algorithm in the sense of number of hubs, the algorithm evolved until reaching a single optimal solution consisting of establishing all 10 nodes as hubs, and its transportation cost can be reduced down to 1106881 and the global warming impact to 269.9526 CO2-equivalent. Conclusions can be drawn from these analyses stating that the costs of establishing the nodes as hubs are significantly low resulting in the algorithm to open all possible hubs which results in a lower transportation cost.

4.6.1 Computational Analysis and Sensitivity Analysis

Exhaustive evaluation was utilized as a comparison indicator to test the algorithm's accuracy. Since the problem being analyzed is considered as a small scale transportation network, it is not considered to be an NP-hard problem, therefore, all possible combinations can be evaluated exhaustively to find the Pareto-optimal set and evaluate the algorithm's performance in reaching this Pareto-optimal set.

Table 4.12: Exhaustive non-dominated solutions

Exhaustively	No of Hubs	Nondominated Solutions										Cost	GWP	Time	No. of Possible Combinations
	0	0	0	0	0	0	0	0	0	0	0	16900243	5305.738	9.675	1
	1	0	0	1	0	0	0	0	0	0	0	7365654	3267.059	14.019	11
	2	0	0	1	0	1	0	0	0	0	0	5698410	2515.663	29.334	56
	3	0	0	0	1	1	0	1	0	0	0	4046494	1735.755	49.33	176
	4	1	0	0	1	1	0	0	1	0	0	3299159	1388.61	71.566	386
	5	1	0	0	1	1	0	0	1	1	0	2803372	1142.27	99.723	638
	6	0	0	0	1	1	1	1	1	1	0	2342975	913.1721	132.919	848
	7	1	0	1	1	1	0	1	1	1	0	1911589	701.3547	163.417	968
	8	1	1	1	1	0	0	1	1	1	1	1602500	535.3647	180.147	1013

		1	1	1	1	1	0	1	1	1	0	1595074	539.2624		
	9	1	1	1	1	0	1	1	1	1	1	1331887	393.5511	196.727	1023
		1	1	1	1	1	1	1	1	1	0	1324272	397.3637		
	10	1	1	1	1	1	1	1	1	1	1	1106881	269.9526	199.994	1024

When comparing this exhaustive solutions to the MOEA's it can be stated that the multiple objective evolutionary algorithm is able to efficiently find the Pareto Optimal set for each instance. Besides, its speed in reaching the solutions is higher than through exhaustive methodology. As represented in Figure 4.8 the computational time for exhaustive evaluation increments in approximately a 6 order polynomial time, whereas the MOEA runs a linear time. This analysis proves the unfeasibility of exhaustive methodologies since their computational time increases at a polynomial rate as the number of combinations increases making it impossible to solve hard combinatorial problems during a lifetime. On the other hand, the MOEA's computational time is shown to be linear with a slight slope making it feasible to solve hard combinatorial problems within a satisfactory timeframe.

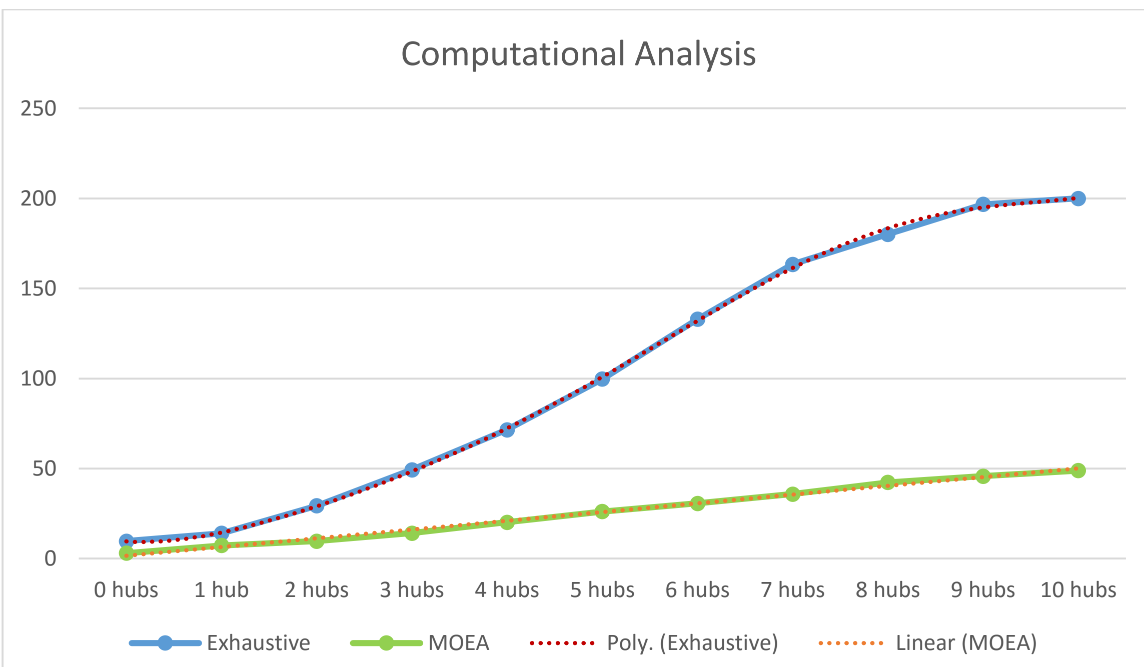

Figure 4.8 Computational time comparison between Exhaustive and MOEA

Therefore, it can be stated that the MOEA developed is suitable in searching for Pareto Optimal Set in transportation networks, and in larger problems, NP-hard, the MOEA will be able to provide robust Pareto-optimal solutions in lesser time that it would take by exhaustively analyzing each possible combination. The optimal data provided through exhaustive methodology was utilized as a base point to analyze the accuracy of the MOEA in reaching the optimal solution when no limit in the number of hubs is given. Table 4.13 provides the results for 20 runs of two selected configuration of parameters, the computational time of each configuration was also recorded, in addition to determining with a 1 whether if the optimal solution was achieved in the run, and a 0 otherwise.

The first configuration of parameters consisted of a population size of 250, with 100 generations, and 0.3 elitism and 0.1 mutation. The second set, considered a population size of 250, with 20 generations, with a 0.1 for elitism and mutation.

Table 4.13: Computational Time and Accuracy Analysis

Top Elitism	Popsize	Gen.	Mut.		Run1	Run2	Run3	Run4	Run5	Run6	Run7	Run8	Run9	Run10		
0.3	250	100	0.1	Time	147.995	152.124	145.453	150.964	151.55	155.057	150.808	152.944	147.993	152.84		
				Accurracy	1	1	1	1	1	1	1	1	1	1		
					Run11	Run12	Run13	Run14	Run15	Run16	Run17	Run18	Run19	Run20	Avg. Time	Accurracy
				Time	142.669	151.514	149.996	151.926	152.509	152.291	150.859	154.051	152.986	150.075	150.8302	
				Accurracy	1	1	1	1	1	1	1	1	1	1		100
0.1	250	20	0.1		Run1	Run2	Run3	Run4	Run5	Run6	Run7	Run8	Run9	Run10		
				Time	50.21	44.597	46.709	49.01	44.448	50.416	48.891	47.909	51.886	45.489		
				Accurracy	1	1	1	1	1	1	1	1	1	1		
					Run11	Run12	Run13	Run14	Run15	Run16	Run17	Run18	Run19	Run20	Avg. Time	Accuracy
				Time	49.142	48.594	47.845	49.768	46.427	47.938	49.714	55.519	52.431	54.214	49.05785	
				Accurracy	1	1	1	1	1	1	1	1	1	1		100

As represented in Table 4.13 the MOEA was able to reach the optimal solution for these two configuration of parameters all of the 20 runs. Making the algorithm accurate and efficient in finding the optimal solution. The average time for the first set of configuration was 150.8302

seconds, and the second set lasted 49.05785 seconds in average. Therefore, since both configurations provided the optimal solution in all runs, both are consider to be suitable for this problem.

Notwithstanding, Multiple Objective Algorithms have gain popularity as a suitable approach in many different fields to solve complex optimization problems; controversy among the optimality of its parameter and their impact has caused an increased research to find the optimal arrangement of these parameters. Literature have shown the existence of a relation among the optimal parameters and the problem being evaluated. If optimality or the nearest value to the optimal is expected from an approach, the optimization of the same should be also taken into consideration. Therefore, a sensitivity analysis together with a design of experiments was performed to evaluate different settings of parameters and their impact on solutions to the scenario where no limit of hubs is specified. The analysis was done following an experimental plan of a 2^k full factorial design, hence, two levels (low and high) were given for each of the factors.

The factors considered for the analysis are: Population size, number of generations, mutation and elitism percentage. Moreover, the total number of parameters analyzed equals to the total number of factors considered in the full factorial design which are four. Thereupon, a 2^4 full factorial design will be analyzed. The high and low levels considered for each of the parameters are presented in the tables below:

Table 4.14: High and Low Levels chosen for MOEA's parameters

Factor A	Elitism
High	Low
0.3	0.1
Factor B	Population
High	Low
250	50

Factor C	Generations
High	Low
20	100
Factor D	Mutation
High	Low
0.1	0.005

Since the 2^4 factorial designs has 16 treatment combinations and two replications were made, a total of 32 runs were performed. The experiment was conducted as a double replicate factorial design in order to better block any noise, to avoid the risk of an unusual response observations affecting the results, to be able to calculate the error and variation, and to better see the interaction between the factors. In the design high and low levels are denoted as +1 and -1.

The data collected in all 32 runs was the following:

Table 4.15: Data collected from 32 runs performed according to the experimental plan

A	B	C	D	Top Elitism	Popsize	Gen.	Mut.	Cost	GWP	Time	Reached Optimal	Cost	GWP	Time	Reached Optimal
1	1	1	1	0.3	250	100	0.1	1106900	269.9526	152.43	YES	1106900	269.9526	145.09	YES
1	1	-1	1	0.3	250	20	0.1	1377574	411.801904	33.614	NO	1377574	411.801904	33.585	NO
-1	1	-1	1	0.1	250	20	0.1	1106900	269.9526	48.906	YES	1106900	269.9526	47.412	YES
-1	1	1	1	0.1	250	100	0.1	1106900	269.9526	225.915	YES	1106900	269.9526	242.12	YES
1	1	1	-1	0.3	250	100	0.005	1106900	26a9.9526	165.606	YES	1602500.4	535.364687	146.822	NO
1	-1	-1	1	0.3	50	20	0.1	1106900	269.9526	10.006	YES	1106900	269.9526	11.36	YES
1	1	-1	-1	0.3	250	20	0.005	1106900	269.9526	38.394	YES	1106900	269.9526	37.064	YES
1	-1	-1	-1	0.3	50	20	0.005	1106900	269.9526	12.904	YES	1424691	432.629555	10.166	NO
-1	-1	-1	1	0.1	50	20	0.1	1106900	269.9526	18.65	YES	1106900	269.9526	18.293	YES
-1	1	1	-1	0.1	250	100	0.005	1106900	269.9526	285.196	YES	1106900	269.9526	285.896	YES
-1	-1	-1	-1	0.1	50	20	0.005	1106900	269.9526	21.641	YES	1377574	411.801904	33.173	NO
1	-1	1	-1	0.3	50	100	0.005	1657395.8	547.875886	44.76	NO	1657395.8	547.875886	43.942	NO
1	-1	1	1	0.3	50	100	0.1	1106881.1	269.952625	46.507	YES	1106881.1	269.952625	44.411	YES
-1	-1	1	-1	0.1	50	100	0.005	1106900	269.9526	198.564	YES	1543476.2	491.586412	81.252	NO
-1	1	-1	-1	0.1	250	20	0.005	1106881.1	269.952625	57.504	YES	1331886.6	393.551118	54.929	NO
-1	-1	1	1	0.1	50	100	0.1	1106881.1	269.952625	92.089	YES	1106881.1	269.952625	93.084	NO

Based simply on the information gathered, some conclusions can be drawn from the best setting of the parameters for this specific problem. The best solution generated by the MOEA was found to be $1106900 for cost, and 269.9526 for global warming impact, reaching the Pareto optimal solution obtained before. This solution is constantly reached when the level of population and mutation is set to its highest. It can be analyzed that the ratio of accuracy increases when the level of population is set to its highest parameter, increasing its accuracy by a 25%. In addition, the mutation parameter also increases the accuracy of the algorithm when it is set to its highest level increasing it by a 50%. Population size and mutation are considered to have a significant effect on the problem. This is due to the relation mentioned before between GA parameters and the problem, since this problem is modeled as binary a high level of mutation aids in introducing variation to the solutions, whereas a large population size allows analyzing more possible solutions per generations. On the other hand, for this problem the number of generations and percentage of elitism do not have a significant effect on the accuracy of the solution. This is probably due to the scale of the problem, since the search space is small there is not a need for too much generations or having a high percentage of elitism for the algorithm to find the optimal solution.

Chapter 5: Biomass-to-Biofuels hub and spoke network logistics design

Regarding the environment, transportation is the most visible aspect of supply chains. Besides improving it by collaboration and optimizing the network with hubs, the increase of GHG emissions produced by transportation can also be diminished by the use of low-carbon fuels. The biofuel supply chain can be benefited from an efficient transportation system, in the meantime, making bioenergy cost efficient can benefit the transportation sector as an effect.

Studies conducted by the Department of Energy (DOE) and United States Department of Agriculture (USDA) have demonstrated that bioenergy can reasonably satisfy 30% of the present petroleum consumption. The Energy Independence and Security Act (EISA) of 2007 mandates an increase in cellulosic biomass being harvested requiring approximately 180 million dry tons by 2022, therefore 16 billion gallons of cellulosic biofuels are expected to be produced yearly. A challenge in supplying that quantity created the necessity for an improvement of the biofuels supply chain, as well as the improvement of the machinery performing each biofuels supply chain operation such as harvesting, storage, preprocessing and transportation. Biomass logistics supply chain optimization requires comprehensive simulation models to optimize both economic and environmental sustainability.

5.1 Literature Review

Extensive literature can be found addressing the problem of optimizing biomass logistics supply chain and is presented below. Most of these approaches consist on formulation of linear equations to achieve a more realistic calculation on the cost of producing biofuel. Recent approaches reformulated linear equations into Mixed-Integer Linear Programs (MILP), while some others continue to use linear optimizers.

5.1.1 Mathematical Approaches

To determine the cost to harvest lignocellulosic biomass to use as a biorefinery feedstock Thorsell *et al*, (2004); conducted a study that determined the potential economies of size that might

result of a coordinated structure. A three-step procedure was used to estimate the harvest cost from the Oklahoma model. Although some studies only take into consideration one type of biomass other studies consider multiple feedstocks. The most common used feedstocks are wheat straw, corn stover, and woody biomass. An analysis on different delivery costs of biomass was developed by Sultana & Kumar, (2011) for this three feedstocks in order to find the optimal configuration and combination at a lower cost demonstrating that the delivery of wheat straw, and corn stover packaged in bales and sending the woody biomass as chips is economically better.

5.1.2 Linear Programming

Cundiff *et al*; (1997) contributed towards the design of a biomass delivery system, considering storage, scheduling and transportation issues. The problem analyzed was to determine optimal schedules for shipping biomass from each producer's field to the central plant, along with a schedule of storage capacity to accommodate the shipping satisfying weather scenarios. A solution methodology called two-stage decision approach is described when dealing with uncertainties in Linear Programming problems. Similar studies are being done not only in the United States, but all around the world such as the case in Italy where Bruglieri & Liberti; (2008) proposed mathematical models for solving problems such as determination of biomasses to produce and/or buy, transportation decisions, and plant site locations. The analysis resulted in a linear programing model describing a biomass-based energy production process.

5.1.3 Integer Programming

Judd *et al*; (2010) proposed a possible solution to the problem encountered in the storage of biomass in-route to bio-energy plants. The study provided storage locations throughout an area for the temporary storage and loading of bales by employing an integer programming formulation.

5.1.4 Mixed Integer Programming

Some have provided thorough work on the economic aspect of the supply chain by conducting research on the costs of different biomass feedstock. For instance, Tembo *et al*, (2003); Mapemba *et al* (2007) conducted several studies in the Southern Plains of Kansas, Oklahoma, and

Texas to identify the key components and potential bottlenecks of conversion of lignocellulosic biomass to ethanol coupled with opportunities for reducing and determining cost to procure, harvest, store, and transport to a biorefinery a flow of lignocellulosic biomass. The studies aided in determining which source of lignocellulosic biomass is found to be the most economical and some restrictions in their harvest seasons. By making use of a multi-region, multi-period mixed integer mathematical programming model the optimal location for a biorefinery was found and the maximization of the total net present worth. Interpretation of results showed a direct proportion among the size of the biorefinery and the transportation distance, meaning that as the biorefinery size increased the transportation distance increased as well; consequently rising transportation cost. Meanwhile, Epplin *et al*; (2007) contributed in analyzing the cost to produce switchgrass for two different alternatives: land-lease alternative and farmers-contract alternative. The main objective of the paper is to determine the cost for each of the alternatives and to determine if NREL's estimated delivered cost of $35 per dry ton is realistic; in addition to identifying challenges of switchgrass as a dedicated bioenergy crop. A multi-region, multi-period, mixed integer mathematical programming was modeled to describe the land lease alternative. The results generated by the comparison of the Oklahoma model and the Tennessee model showed that both costs range among the same numbers, while in comparison to the NREL's estimated delivery cost this range is substantially more than their estimated cost. The analysis also provides an insight on the requirement of increasing switchgrass yield per acre to satisfy a reduction in cost that will satisfy the NREL's estimated cost. In 2012, Marvin *et al*, (2012) formulated and solved a mixed integer programing (MILP) problem of an optimization study of the net present value of a biomass-to-ethanol supply chain in the Midwestern, United States. A single parameter sensitivity analysis and Monte Carlo simulation were approaches performed to determine the robustness of the supply chain. The optimization problem consisted on determining the location of the biorefineries in conjunction with their capacity in addition to harvesting quantity and shipping so as to maximize the net present value.

After showing the improvement made by mathematical methods, two years later Jud *et al*; (2012) utilized a mathematical programming-based method to determine the Satellite Storage Locations and equipment routes that will minimize the total cost of a feedstock logistics system that relies on the use of SSLs for temporary storage and loading of round bales. The paper proposed a Mixed Integer Linear Program to determine optimal number and locations of the SSLs while minimizing transportation, equipment use and labor costs. Shastri *et al*.; (2011) conducted a study of an integrated framework to connect various feedstock production related activities such as pre-harvest crop management, harvesting, transportation and pre-processing. A breadth level mathematical programming model called BioFeed is introduced as a Mixed Integer Linear Program (MILP.) The model has the main objective of determining the optimal configuration to the feedstock production system to maximize the total profit while incorporation long term design decisions as well as management decisions. The study found that the total profit is highly sensitive to factors such as truck idling time and output density.

A multi-commodity network flow model to design the logistics system for a multiple-feedstock biomass-to-bioenergy industry was constructed by Zhu & Yao, (2011) to explore the possibility of using these three same feedstocks. The problem was formulated as a Mixed Integer Linear Programming to determine the locations of warehouses, the size of harvesting team, the types and amounts of biomass being harvested or purchased, quantity of biomass to be stored, and total to be process for each month during a planning horizon of a year. A comparison between using only one type of feedstock against multiple was made. The numerical study showed that the usage of different types of biomass can increase the supply of biomass, alleviate the seasonality caused by the non-harvesting season of switchgrass, smooth the biofuel production, and eventually increase the unit profit of biofuel. This same study was improved by Zhu *et al*, (2011) who also analyzed the challenges in designing the logistics systems for a biomass-to-bioenergy industry, including features such as low bulk density, restrictions on harvesting season and frequency, content variation with time, weather effects, and scattered distribution over a wide geographical area, among others. A mixed integer linear programming model was implemented to serve as a

decision-making tool that will maximize the profit in this multi-echelon, multi-period, and multi commodity supporting system.

5.1.5 Multi-Objective Optimization Approaches

Although one of the major concerns of biofuels is their cost-efficiency, nowadays is not only profit that is important, many people, companies and government itself are concerned about the sustainability of biofuels as well. Therefore, several researchers have addressed the problem by incorporating the environmental aspect of the biofuels supply chain.

Multiobjective optimization models have been designed to examine the optimal supply system taking into consideration greenhouse gas (GHG) and feedstock cost. For example the study of Yu *et al*, (2014) done at Tennessee considered the plant gate cost of switchgrass, including the farm gate cost and transportation cost. The greenhouse gas being analyzed include direct and indirect emissions. This study applied a multiobjective mathematical model to minimize both cost and GHG emissions in a switchgrass supply chain and resulted in that both feedstock cost and GHG emissions were heavily influenced by the type of land converted for switchgrass production.

Other methods employed that also include the environmental aspect of the biofuels supply chain can be found to make use of Arena Simulation Software such as the design of Zhang *et al*, (2012) that evaluates the supply chain based on delivered feedstock cost, energy consumption and GHG emissions. The model is tested under a case study for Michigan's Lower Peninsula. The model demonstrated to be a useful tool to represent a supply chain for biomass feedstock for several biofuel facility locations and it can be extended nationwide. Some studies such as Kumar and Sokhansanj, (2007) focus specifically on one type of biomass feedstock such is the case of this study developed to evaluate cost, energy input and carbon emissions for a number of switchgrass supply options. The study was done with the use of the Integrated Biomass Supply Analysis and Logistics (IBSAL) model developed at Oak Ridge National Laboratory.

Jiuh-Biing Sheu *et al,* (2005) proposed a methodology to address both economic and environment aspects at the same time is a linear multi-objective programming model. Results of

numerical studies indicate that the chain-based aggregate net profits can be improved by 21.1%, compared to the existing operational performance in the particular case studied. Fan Wang *et al*, (2011) proposed a normalized normal constraint method to solve the model by MILP solver CPLEX 9.0 to get a Pareto optimal. You and Wang, (2011) addressed the optimization of cellulosic ethanol supply chains under economic, environmental and social objectives. The model is developed as a multi-objective mixed integer linear program (MO-MILP) while taking into account the characteristics presented by cellulosic ethanol, such as seasonality, as well as supply, demand, regional economic structure and government incentives. The main objectives of the paper are minimizing total annual cost of the project, minimize greenhouse emissions and maximize the number of jobs created yielding a Pareto-optimal curve.

Although a representative number of literature can be found in optimizing the operations of the supply chain. Almost none is found in optimizing the transportation aspect of the biofuels supply chain, only these following.

In his dissertation Roni, Md, (2013) proposes a framework in support of biomass supply chain network design that relies in the use of trucks for short distance biomass transportation, and relies in the use of rail for long-haul, and high-volume transportation of densified biomass. A hub and spoke network design model is proposed for the case when biomass is shipped by rail. Problems analyzed are modeled as Mixed-Integer Linear Programming (MILP). A Benders' decomposition-based algorithm is developed to solve the MILP. This MILP identifies the number, capacity and location of biorefineries needed to make use of the biomass available in the region. A case study is created using data from a number of States in the Midwest USA. The numerical analysis also reveals the tradeoffs that exist among the economics, environmental impact, and social objectives of using densified biomass for production of biofuel. In (2014), Roni, Md. *et al*, formulated this same problem as a multi-objective, mixed integer programming problem under economic, environmental, and social criteria. The problem was solved with an augmented ϵ-constraint method to find the exact Pareto solution. Later, Mohammad R. *et al*, (2016), advanced this same analysis capturing the trade-offs that exist between costs, environmental and social

impacts of delivering biofuels. The model proposed optimizes the CO_2 emissions due to transportation-related activities in the supply chain. The model also optimizes the social impact of biofuels by the number of jobs created. The multi-objective optimization model is solved using an augmented ϵ-constraint method. The method provides a set of Pareto optimal solutions.

5.2 State-of-the-art

At this present moment, society has reached a point where economic profitability is not sufficient to assure the survival of future generations which brought with it the sustainability movement. Environmental challenges are being faced and searching for sustainable systems has become the priority for business owners and researchers. From the literature presented above it can be noticed there are vast of studies addressing biofuels optimization. This topic became of great interest since it was shown to be a potential candidate to replace fossil fuel usage. Some studies have served as a base for the discovery of new sources of energy such as the case of biofuels. Studies addressing the physical and chemical characteristics, most addressing the economic aspect of feedstocks itself, and considering the logistics design process, and some few addressing emissions and sustainability. Nevertheless, there is always area for improvement, especially in emerging topics as it is bioenergy.

As previously stated, there is still limited literature addressing environmental impacts, some studies are beginning to conduct complete assessments of sustainability by introducing CO_2 emissions and social indicators into their studies, but there is a deficiency of studies that analyze how these emissions are impacting the environment. In contrast to past research, this paper not only takes into consideration the GHG emissions, nor evaluates only the global warming impact category; likewise, it aggregates two other objectives by considering as well acidification and eutrophication potentials creating a more extensive and complete analysis of environmental sustainability of biofuels supply chain. Furthermore, non-previous literature provides the stakeholder with trade-offs in the planning process in addition to an optimized transportation system. In order to optimize the transportation network of biofuels logistics system, a hybrid hub-

and-spoke system will be employed considering multiple assignment of hubs. This network structure can be beneficious to the biofuels logistics system by consolidating shipments to increase the efficiency of the operations. A study similar to this one was conducted by Roni, Md. *et al*, (2014) and Zhu and Yao (2011).

The present research proposes the development of a Multi-Objective Evolutionary Algorithm to determine the Pareto optimal solutions for the operations configuration of a hub and spoke biofuels logistics network. Providing robust yearly configurations for the system's planning operations as well as its corresponding transportation design. Moreover, this thesis can provide an initial analysis on environmental assessment of a hybrid hub and spoke network design for the biofuels supply chain, besides being found as a pioneer study in introducing environmental impacts.

5.3 MODEL FORMULATION

The mathematical formulation considers a single feedstock logistics system aggregating different feedstock residues, for instance, corn stover and wheat straw, on to one single referred as crop residue. The analysis considers optimal placement for preprocessing facilities, depots to be established as hubs, and biorefineries, efficiently allocating biomass flow on network nodes to satisfy customer demands. The system is composed of 233 possible nodes for preprocessing facilities, 8 depot nodes, 67 biorefineries available to process biomass, and a total of 113 customers. The operations period is based on a one-year period and only two transportation modes are considered for the analysis. Biorefineries and depots are assumed to have access to train, therefore, rails are available for traveling between depot to depot, and between depots to biorefinery. The transportation network considered in the design of the system follows a structure where the tons of biomass can be transported from any preprocessing facility to a depot (hub) or several depots (hubs) and from there to its nearest biorefinery and finally the biofuel is distributed to the customers. Moreover, since the logistics network considers a hybrid hub-and-spoke network, biomass can also be transported from any of the preprocessing facilities directly to the biorefinery.

The figure below provides a graphic illustration of the supply chain network structure considered for the hub-and-spoke biomass-to-biofuel logistics design network.

Figure 5.1: Representation of biomass & biofuel flow

Let the transportation mode be denoted by $t \in T$, 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. In this system, only customers c have an associated demand which is denoted by d_c, and the shipment begins at the preprocessing of biomass as the origin point i to the first destination point j which is the biorefinery with the allowance of consolidating shipments on a depot (hub) k. After biomass has been processed, biorefinery becomes the second origin point i and customers is the last destination point j.

5.3.1 Model Constraints

This model is subject to constraints that need to be taken into consideration when obtaining solutions. The death penalty method was used in the developed MOEA to ensure that only solutions satisfying the model constraints were considered. First set of constraints are related to the transportation, whereas the second to biomass-to-biofuel production.

Transportation Constraints

- Routes Constraint

$$\sum_k \sum_l Y_{ijklt} + V_{ijt} = 1 \quad \forall i, j, t \tag{5.1}$$

In order to limit the number of routes from one point to the next, constraint (5.1) impedes the programing of more than one different route between two points in the system.

- Shipments assignment

$$\sum_l Y_{ijklt} \leq X_k \tag{5.2}$$

$$\sum_k Y_{ijklt} \leq X_l \tag{5.3}$$

Constraints (5.2, 5.3) 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.

Biomass-Biofuel Production Constraints

- Preprocessing Capacity

$$DT_p \leq PCAP_p K_p \tag{5.4}$$

This constraint (5.4) limits the amount of biomass on a Preprocessing facility *p* to be less or equal than the capacity for each *p*.

- Depots Capacity

$$DTC_d \leq DCAP_d X_k \tag{5.5}$$

The depots capacity constraint (5.5) sets the limit on the biomass feedstock that can be consolidated at the hub, according to its size. However, in this study all depots are the same size and its capacity is assumed to be the same for all.

- Production Capacity

$$DTP_b \leq BCAP_b Z_b \tag{5.6}$$

Production capacity constraint (5.6) provides an upper limit on the feedstock amount that can be processed at each biorefinery per year.

- Demand

$$\sum_{b=1}^{67} BFG_b \geq \sum_c^{113} d_c \tag{5.7}$$

Demand constraint (7) provides a limit on minimum production to meet customers demand by ensuring that the summation of biofuel gallons produced at biorefineries *b* opened is greater than de summation of all customers' demand d_c

- Binary constraints

$X_k \in \{0,1\}$ (5.8)

$Y_{ijklq} \in \{0,1\}$ (5.9)

$V_{ijq} \in \{0,1\}$ (5.10)

$K_p \in \{0,1\}$ (5.11)

$Z_b \in \{0,1\}$ (5.12)

Constraints set (5.8-5.12) constraints variables *X*, *Y*, V, K, and Z into the binary space.

5.3.2 Model Objective Function

Four objective functions are considered such as maximization of the total annual profit as well as minimization of three environmental impacts which are global warming impact, acidification and eutrophication. The four objective functions are described below.

- Maximization of Profit, defined by

$$Max\ R - \sum_{n=1}^{4} C_n \tag{5.13}$$

Where R is the revenue produced by the biofuel generating system

$$R = \sum_{b=1}^{67} \rho BFG_b \tag{5.14}$$

Calculated as the summation of output of gallons of biofuel BFG produced at biorefinery *b*

The total annual cost is represented by

$$\sum_{n=1}^{4} C_n \tag{5.15}$$

Considering the following costs:

C_1-Transportation cost

$$\textstyle\sum_i \sum_j \sum_k \sum_l \sum_t C_{ijkl} d_{ijt} Y_{ijklt} + \sum_i \sum_j \sum_t W_{ijt} d_{ijt} V_{ijt} \tag{5.16}$$

It consist of two terms, the first section represents the total transportation costs associated to the utilization of two modes of transportations, the second part represents the total costs associated with the usage of a single transportation mode and shipping directly. The variable C_{ijkl}

considers the consolidation revenue oriented cost associated to demand d_{ijt} considering the reduction of cost associated with transporting from hub k to hub 1 utilizing both transportation modes. Whereas W_{ijt} represents the cost associated with shipping directly solely utilizing truck.

C_2-Hub establishment cost

$$\sum_k \sum_t P_{kt} X_k \qquad (5.17)$$

Which represent costs associated with locating a collaborative candidate hybrid consolidation facilities. The collaborative candidate hybrid consolidation hub location costs are obtained as the summation of the product of the costs of locating a hubP_{kt}, and the X_k (the decision on whether a collaborative facility is located). Where $P_{kt} = \vartheta_k + \phi_k$ is the investments required to build rail ramps. And it is conformed of ϑ_k is the cost associated with additional tracks required to establish as hub node *k,* and ϕ_k is the turnout cost which allows rail cars to switch tracks.

C_3: Processing Cost

$$\sum_b^{67} PC_b DTP_b \qquad (5.18)$$

Where DTP_b are the dry tons of biomass processed at biorefinery *b*

- C_4-Operation cost of depots and biorefineries

$$\sum_{d=1}^{8} OD_d Y_d + \sum_{b=1}^{67} OB_b Z_b \qquad (5.19)$$

Where Y_d is a binary variable equal to 1 if depot d is open and set as a hub and 0 otherwise, and Z_b is a binary variable equal to 1 if biorefinery *b* is open and 0 otherwise.

- Minimization of Environmental Impacts
- Minimization of Global Warming Potential (measured in kg of $CO2_{\text{equivalents}}$)

$$\min GWP = GWP_p X_b + GWP_e X_b + \sum_i \sum_j \sum_k \sum_l GWPM_{ijkl} d_{ijt} Y_{ijklt} + \sum_i \sum_j \sum_t GWPS_{ijkl} d_{ijt} V_{ijt} \qquad (5.20)$$

- Minimization of Acidification Potential (measured in kg of $SO_{2\,\text{equivalents}}$)

$$\min AP = AP_p X_b + AP_e X_b + \sum_i \sum_j \sum_k \sum_l APM_{ijkl} d_{ijt} Y_{ijklt} + \sum_i \sum_j \sum_t APS_{ijkl} d_{ijt} V_{ijt} \quad (5.21)$$

- Minimization of Eutrophication Potential (measured in kg of phosphate $_{\text{equivalents}}$)

$$\min EP = EP_p X_b + EP_e X_b + \sum_i \sum_j \sum_k \sum_l EPM_{ijkl} d_{ijt} Y_{ijklt} + \sum_i \sum_j \sum_t EPS_{ijkl} d_{ijt} V_{ijt} \quad (5.22)$$

In all three environmental impacts objectives, the impacts resulting from preprocessing biomass is denoted by the first term, and the impact of processing ethanol from crop residues is considered by the second term. Meanwhile, the impacts generated during the transportation from origin point *i* to destination point *j*, traveling thru hubs *k* and *l*, is taken into account in the last two terms of the formulation. The first term represents the impacts associated to a multi-modal network, where both truck and train are utilized for the transportation of biomass/biofuel, and the second part considers the impact resulting from a single mode transportation network.

A table listing all variables employed in the formulation is provided in APPENDIX C to facilitate the comprehension of the model.

5.4 Optimization Methodology

5.4.1 Data Analytics Methodologies

Data analytics was implemented to examine raw data on coordinates within the United States, to search for optimal locations for facilities in the biofuel supply chain. The approaches used are the K-means and the Silhouette to select the optimal number of clusters (k) for our data. A computer program was developed to generate strategic coordinates for preprocessing facilities, and customers. On the other hand, biorefineries coordinates were obtained from Oak Ridge Laboratories database.

5.4.1.1 Silhouette analysis

This analysis was introduced by Peter J. Rosseeuw in (1987) proposed as a graphical display for partitioning techniques based on tightness and separation. It is used to study the separation distance between the resulting clusters by measuring the closeness between a sample point in one cluster to points in neighboring clusters; thus this analysis provides an approach to assess parameters like number of clusters. In order to construct silhouettes, it is necessary to have a partition already obtained through any clustering technique and the collection of all proximities between objects. Then, for each point (i) a silhouette value s(i) will be introduced by the following formula and then a plot will be constructed.

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

Where *a*(i) equals to the average dissimilarity of i to all other objects of cluster A. In other words, this is the average length of all lines within A. Afterwards, another cluster C different from A is considered in the analysis. And a computation of the length from i to all other objects in C is done and then an average of this length is obtained. This average dissimilarity of *i* to all other objects in C is represented as *d*(i, C). After computing *d*(*i*, C) for all clusters different to cluster A, the smallest of those numbers is selected and denote it by *b*(i).

This silhouette coefficient is ranged between -1 and +1, where a positive number indicates that the point is far from the neighboring clusters, a measure of 0 indicates closeness between two neighboring clusters whereas negative value indicates a bad assignment of cluster. When searching for a ‘natural’ value of k, silhouettes values are desired to be as wide as possible. Once silhouette values are computed for each data object, an average silhouette width can be computed for each cluster. This will allow to distinguish strong from weak clusters in the same plot. In addition, an overall average silhouette width for the entire plot can be considered by the average of *s*(*i*) for all objects *i* in the whole data set. Therefore, each value of k will yield a different overall average silhouette width *S*(*k*), and the largest *S*(*k*) value will provide the appropriate number of clusters *k* for the data set.

For this analysis, the silhouette concept was applied to locate optimal centroids for our depots. The raw data on coordinates was tested using k-means for different values of *k* to provide the overall average silhouette width $S(k)$ for each *k*, and find the best k for this data. A table of the top five silhouette values is shown below in Table 5.1.

Table 5.1: Top k's ranked according its silhouette value

Top 5 *k*'s	
k	*Silhouette Value*
2	0.8732
3	0.7127
4	0.6116
5	0.5977
8	0.5879

According to the presented silhouette values the optimal *k* for the data introduced is *k*=2, this signifying that optimal partitioning of this data is two, nevertheless, a greater number that was still in the top *k*'s was chosen to enlarge search space when considering biomass supply chain data as well. Introducing more possible consolidation hubs into the supply chain can result in reduction of transportation costs and lower emissions as proved in Chapter 4. Therefore, a *k* equal to 8 was chosen to allow the MOEA evaluate if when considering data from the supply chain, *k*=2 number of hubs is still the optimal.

5.4.1.2 K-Means

Data clustering techniques are powerful tools for handling large amount of data. Some applications are structuring data information and study correlation among data. *K*-means is a form of unsupervised classification, as the clusters are formed by evaluating similarities and dissimilarities of intrinsic characteristics. A vast quantity of classification methodologies exist now a days, nonetheless, *k*-means continues to be one of the simplest algorithm using unsupervised learning method, easily programmed and is computationally economical, so that it is feasible to process very large samples on a digital computer. Possible applications include methods for

similarity grouping, nonlinear prediction, approximating multivariate distributions, and nonparametric tests for independence among several variables. Therefore for the purpose of this study, k-means was used to determine the location of the depots, partitioning the data in several regions. This technique was introduced by MacQueen (1967), describing it as an iterative relocation of data points between clusters, and a centroid model.

The process which is represented in Figure 5.2, begins by selecting the number of k's we want to explore, k referring to number of clusters. Once a decision has been made, a random selection from the data is made for k number of centroids. This random centroids will initiate the process of encountering the accepted mean for each cluster in the entire population. Once centroids are selected, Euclidean distance is used to calculate the distance from the centroids to each of the other solutions in the population. A matrix of size *m x n* will be formed *m* representing the size of total population, and *n* the size of *k* centroids. Once the distance matrix is constructed, the minimum distance for each individual is located defining to which centroid that individual belongs. Once data have been grouped in k clusters, new centroids are calculated as the average mean of the population in each cluster.

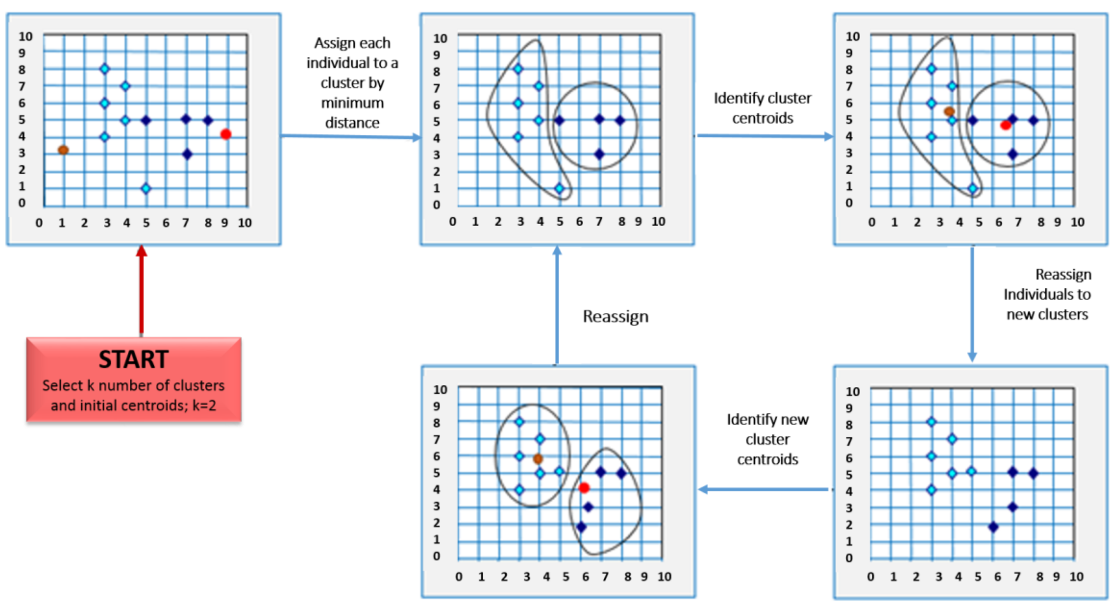

Figure 5.2: K-means representation for k=2

This process is done iteratively until no change is found between centroids in iteration *i-1* and iteration *i*. As mentioned above, for the selection of the optimal number of k the silhouette technique was applied, and from the 5 top silhouette values, 8 was chosen as k for this data. This implies that the data was partitioned in eight regions and the ultimate centroids where used as possible location for depots to be considered in the MOEA. Figure 5.3. Illustrates the 8 clusters formed, and table 5.2 provides their centroids.

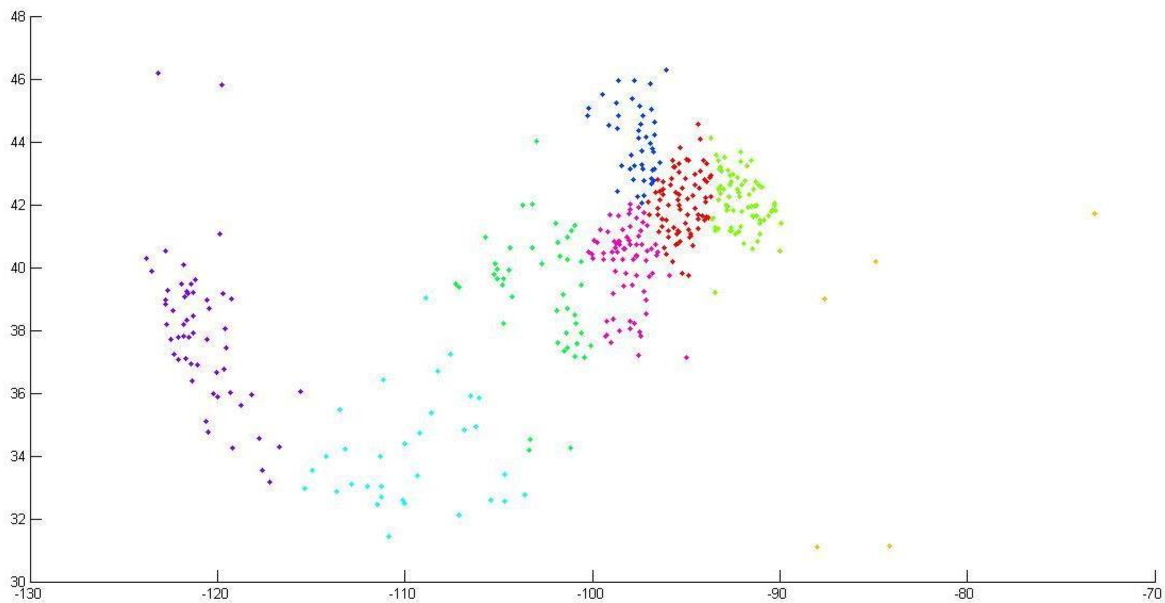

Figure 5.3: Representation of Clusters/regions for data

Table 5.2: Centroids coordinates for depot location

Centroids		
C1	-118.805	35.3794
C2	-104.546	37.80519
C3	-93.8052	42.1166
C4	-97.232	43.22523
C5	-89.718	40.83563
C6	-98.723	39.82165
C7	-111.491	33.53685
C8	-121.607	38.97819

In this study, each observation represents a facility location of the biofuels hub and spoke network to be designed. The data was divided in k=8 regions allowing the location of depots to be used as possible hubs in the network. Nonetheless, the MOEA determines whether a depot is stablished as a hub or not.

5.4.2 Multi-Objective Evolutionary Algorithm Description

Finding the solution to problems involving several objectives which have to be simultaneously satisfied can be difficult to solve. The main characteristics of multi-objective problems are: the objectives are in conflict and the search space is highly complex. Different models have arisen proposing several methodologies to obtain Pareto optimal solutions, differing in most part in their approach to achieve diversity.

A Multiple Objective Evolutionary Algorithm was developed as a decision tool in providing transportation networks and supply chain designs that will minimize environmental impacts of a hub and spoke design. The proposed MOEA has two main evaluating goals: proximity and diversity. Moreover, the algorithm uses a genetic algorithm based on rank selection and elitist reinsertion. A flow-chart representing the general process is provided in Figure 5.4. Furthermore, a thorough description of each stage of the algorithm is given below in order to provide a better understanding of its development.

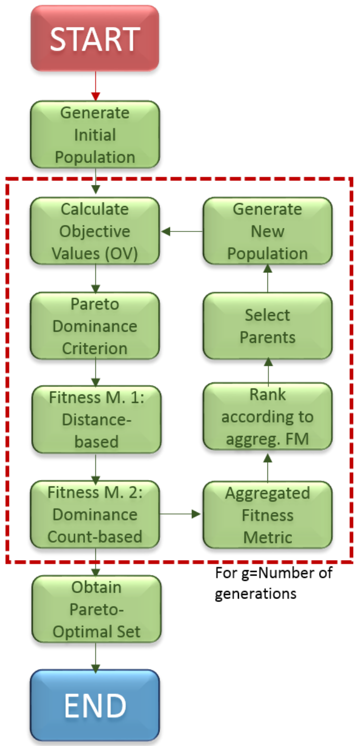

Figure 5.4: MOEA's flow chart

5.4.2.1 The Chromosome

As explained in Chapter 4, modeling the chromosome is the first step to initiate the algorithm, and its structure is based upon the problem and data. Encoding for this problem was selected to be a mixture of binary and integer. Based on the solution intended to be provided which is the planning design of a hub and spoke biomass supply chain, a representation of an actual chromosome is given by figure 5.5 specifying each of the sections.

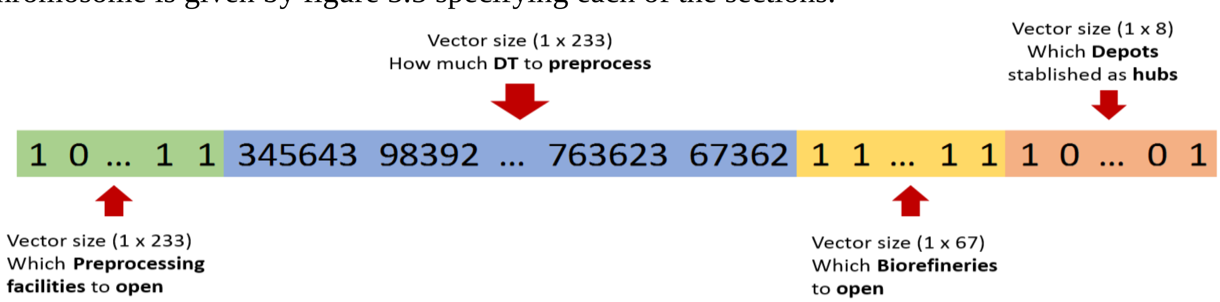

Figure 5.5: Representation of the chromosome for a hub and spoke biomass supply chain

Where the entire chromosome is a vector of length m, m being 541. The first section of the chromosome is a binary vector of length 233, which determines how many and which preprocessing facilities should be open. The following section consists of an integer encoding vector of the same length (233), where the quantity of dry tons to be preprocessed by each facility is determined. Followed by the section represented in yellow in the figure above, this section determines which biorefineries to open in a binary vector of length 67, and the final section, a binary encoded vector of length 8 establishes depots as hubs.

Once the final model structure of a chromosome is selected, an initial population is randomly generated. The size of the population is dependent on the nature problem, but it commonly constitutes several hundred of possible solutions (search space). And each individual or chromosome in the population is evaluated by each of the objective functions. In this analysis, maximization of profit and minimization of environmental impacts is aimed as explained in the model formulation.

5.4.2.2Dominance count

Considering that a simultaneous optimization of all objectives is intended to be achieved, after all individuals in the population are analyzed by the n objective functions, the initial solutions are then evaluated by checking the Pareto dominance criterion. For terms of generality, all objectives have to be written in the minimization form. In case that one objective is in the maximization form, the objective function is multiplied by negative one to achieve a conversion onto a minimization problem as shown in Figure 5.6.

Population	Profit	GWP	Acidif.	Eutrop.
1 0 ... 345643 98392 ... 1 1 ... 1 0 ...	9876635	1265182	20938	24632
0 1 ... 76352 6253 ... 1 0 ... 1 1 ...	6528392	1254645	39879	13287
1 0 ... 123452 12876 ... 1 0 ... 1 0 ...	8372123	1523936	32097	12456
1 0 ... 32345 87653 ... 1 0 ... 1 1 ...	7365234	1432933	20654	23765

(-1)

=

Population	Profit	GWP	Acidif.	Eutrop.
1 0 ... 345643 98392 ... 1 1 ... 1 0 ...	-9876635	1265182	20938	24632
0 1 ... 76352 6253 ... 1 0 ... 1 1 ...	-6528392	1254645	39879	13287
1 0 ... 123452 12876 ... 1 0 ... 1 0 ...	-8372123	1523936	32097	12456
1 0 ... 32345 87653 ... 1 0 ... 1 1 ...	-7365234	1432933	20654	23765

Figure 5.6: Minimization of maximization objectives

Afterwards, a dominance count operation is performed and dominated solutions are eliminated in an attempt to achieve proximity to the True Pareto front. The Pareto criterion is initiated with the dominance count. The dominance count represents the number of solutions dominated by each individual. Figure 5.7 provides a representation on the process for Pareto Dominance Criterion.

Solution	Dominance count	Solutions Dominated
1	**5**	**4,7,10,11,12**
2	**1**	**3**
3	**0**	**0**
4	**0**	**0**
5	**0**	**0**
6	**0**	**0**
7	**0**	**0**
8	**0**	**0**
9	**2**	**3,12**
10	**2**	**7,12**
11	**2**	**4,7**
12	**0**	**0**

Dominated Solutions	Non-dominated Solutions
3,4,7,10,11,12	1,2,5,6,8,9

Non-dominated Set	Profit	GWP	Dominance Count
1	-89425456.83	1553199716	5
2	-71265912.1	1368428579	1
5	-92689433.42	1789450550	0
6	-97708041.98	1950596365	0
8	-96891705.08	1828161389	0
9	-77577377.68	1497750276	2

Figure 5.7: Pareto Dominance Criterion

5.4.2.3 Fitness assignment

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).

- Fitness Metric 1: Distance-based, $f_1(i)$

This fitness metric gives highest fitness to those solutions that are farther away from other solutions in the Pareto front, giving those solutions a greater possibility to be chosen later for reproduction. With this fitness function, we aim to maintain diversity of the Pareto-optimal solutions.

Several steps are followed to assess this fitness metric:

1. *Normalization:* Each objective's solution is normalized according to Equation 5.23. Any discrepancies with the units are aimed to be eliminated by normalizing the values.

$$\frac{f_{i(x)}-f_i^{min}}{f_i^{max}-f_i^{min}} \qquad (5.23)$$

f_i^{min}=minimumvalue in the nondominated set
f_i^{max}=maximum value in the nondominated set
$f_{i(x)}$=value in the nondominated set

2. *Distances:* Euclidean distance between each solution to the rest of the solutions is calculated and a summation of all distances per solution is computed. The maximum and minimum are obtained to determine the intervals to categorize solutions. Father away non-dominated solutions are given the highest fitness to ensure diversity.

- Fitness Metric 2: Dominance count-based, $f_2(i)$

The second fitness metric is based on the dominance count concept. The second fitness aims to select those individuals which are more dominating in an attempt to achieve proximity to the true Pareto front.

- Aggregated Fitness Metric

The two different fitness metrics are then aggregated assigning equal weight to each of the fitness metrics aiming to achieve proximity and diversity; which are two of the most common desirable characteristics in MOEAs.

5.4.2.4 Selection

Through generations a proportion of the population is selected to reproduce and create a new population. During the selection process, the fittest individuals are given a higher probability of reproduction. The fitness of an individual is determined by the overall aggregated fitness metric. Several methods could be chosen when selecting the parents of the next generation. Random selection, roulette wheel, rank selection, and tournament are the popular and widely used selection methods. For this MOEA, a tournament selection was utilized where two individuals are selected randomly and the individual with the best fitness value wins the tournament. In addition, elitism is considered by selecting the parents with the highest fitness value and send them intact to the next generation.

5.4.2.5Crossover

After selection of parents is accomplished, selected parents will crossover to create new generation and population of individuals (solutions). Reproduction can be achieved through different types of crossover, such as, single-point, double-point or uniform crossover or segment crossover. The suitability of the method chosen for the reproduction process depends highly on the type of problem and its encoding. The crossover method chosen for this study was the segment crossover and single point crossover. From each parent the vector is divided into segment, the first segment for this crossover considers the preprocessing facilities and the dry tons preprocessed at each, therefore crossover for this two sections must be considered as a single segment to assure the DT are at opened preprocessing facilities. The two other segments are the binary vectors considering biorefineries to open and depots established as hubs. The crossover is executed as shown in Figure 5.8.

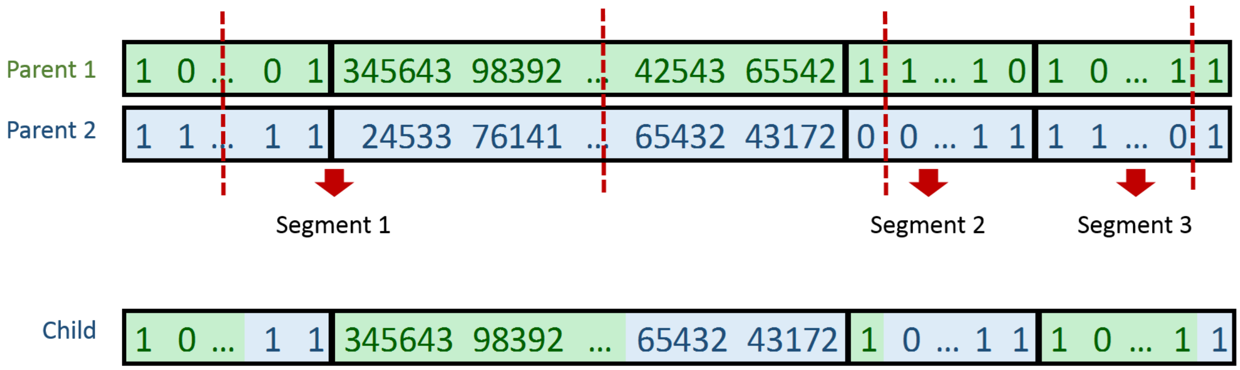

Figure 5.8. Representation of segment crossover in the algorithm

5.4.2.6 Mutation

Within the reproduction process, some individuals of the new population generated undergo a mutation process. This mutation process avoids falling into a local optimum and allows entering of variation in the solutions, usually mutation is taken into consideration with a very small percentage. The mutation process followed is a single point mutation, where randomly a gen (cell) is selected and if it is binary a 1 changes to 0 and 0 to 1, if the random point selects the integer vector, another random number satisfying constraint of preprocessing capacity is selected.

5.4.2.7 Termination

Iteratively, this process of generating new solutions is done until the stopping criterion is met. The termination condition can be either a solution that satisfies a specific criterion such as a predetermined number of generations or using a threshold to detect when the algorithm has reached a steady state and it is no longer evolving.

5.5 Life Cycle Assessment

Nowadays, different methodologies are being analyzed to measure the impact of the different gases being released by our daily activities. Life cycle thinking and life cycle assessment are key elements to better address climate change, and environmental effects of depletion of resources. Impacts considered in a Life Cycle Impact Assessment include climate change, ozone depletion, eutrophication, acidification, and human toxicity, among others. Not only will this

assessment allow a conversion from qualitative degradations into quantitative measures, but as well it will serve as a tool to a better management of the resources.

Life Cycle Assessment is a technique for the systematic evaluation of the environmental aspects and potential impacts that are associated with a process, product or services through all stages of its life cycle. LCA provides an adequate instrument for environmental decision support. Therefore, the International Organization for Standardization (ISO) has standardized this technique, dividing the process into phases such as described in Figure 5.9.

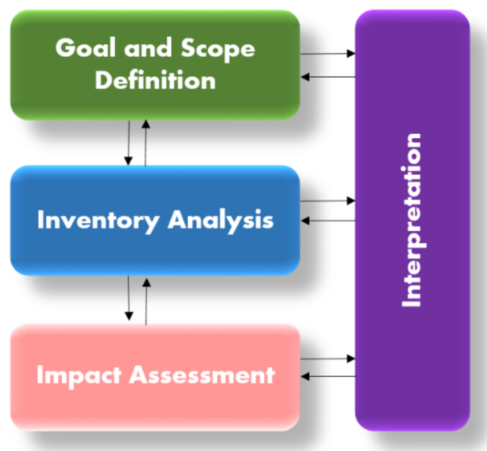

Figure 5.9: Life Cycle Assessment Phases

As shown in the figure above, the process initiates by defining the goal and scope of the project. Next in order to assess the environmental impacts the compilation of inventory material inputs and environmental releases are necessary. Followed by the evaluation of this inputs and releases determine the potential environmental impacts, and finally the interpretation of the results can help make more informed decisions. This assessment has emerged as a valuable decision support tool for both policy makers and industry when assessing impacts of a product or service.

5.5.1 Goal and Scope Definition

In the Goal and Scope definition phase, the product or service to be assessed is defined, a functional basis for comparison is chosen and the required level of detail is selected. For the

purpose of this study, the goal of is to provide a measure of different impact categories of hub and spoke networks utilized in the biomass to biorefinery logistics system.

It is worth noting that the results expressed in this study are relative and not absolute. The present study deals with crop residue as one type of feedstock. In order to provide a good measure of the burdens the production of biofuels are causing in the environment three categories will be taken in to account. The Global Warming Potential (GWP) was calculated over a specific period of a 100 years measured in kg CO2-equivalent, Acidification Potential (AP) is measured in kg SO2- equivalent, and Eutrophication Potential (EP) in kg-phosphate-equivalent. A life cycle assessment (LCA) approach was performed utilizing the life cycle assessment modeling software GaBi. GaBi software will allow to test the different possible scenarios generated by the algorithm thus providing emissions in which the functional unit is the production per ton of biofuel allowing the previous program generated in MATLAB to solve the different objectives of maximizing profit and reducing emissions as explained earlier.

5.5.2 System Boundaries

Since the goal of this study is to find environmental impacts for the production of biofuels, the boundaries are set to be from cradle to grave which is illustrated in Figure 5.10. This includes the processes involved with the preprocessing of crop residues, transportation emissions from origin point (preprocessing facilities) to hubs (depots) and destination (biorefinery), the production of biofuel, and transportation form new origin point (biorefinery) directly sent to destination (customers).

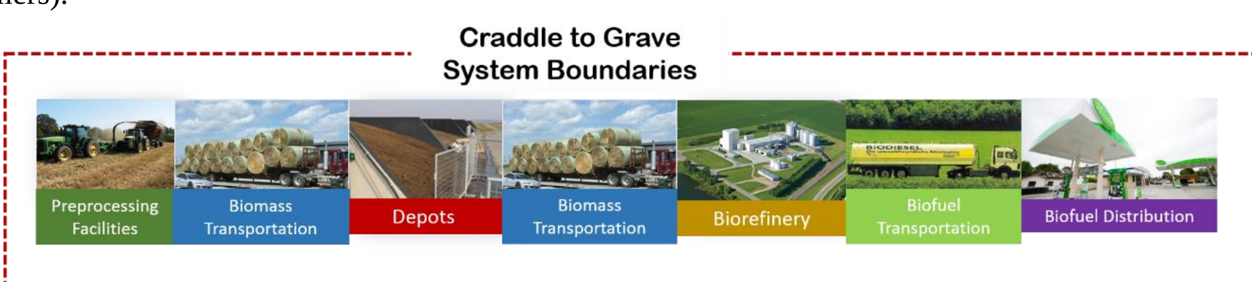

Figure 5.10. System Boundaries for Hub and Spoke Biofuels Logistics

5.5.3 Inventory Analysis

In every Life Cycle Assessment, it is always ideal to capture each and every relevant flow but there are always some discrepancies since not always the information is provided. Inventory Analysis of extractions and emissions, the energy and raw materials used, and emissions to the atmosphere, water and land, are quantified for each process, then combined in the process flow chart and related to the functional basis. Therefore in this study several assumptions had to be made in order to provide a good estimate on the impacts the process is creating.

- GABI software by PE INTERNATIONAL, provides several process and flows and from this database the process of crop residue preprocessing, with its respective flows was utilized. GABI software also provided the source of transportation for this study the truck being utilized is a Truck-trailer, diesel driven, Euro 4, cargo with a payload capacity of 40 tons.

5.6 Multiple Objective Case Study

5.6.1 Numerical Example

The developed multi-objective evolutionary algorithm was evaluated using a numerical example adapted from the work presented in Zhu & Yao; (2011) and Roni, Md. *et al* (2014). The theoretical example considers general crop residue as one feedstock. A total of 421 nodes are considered for the analysis, assigning 233 fields to preprocessing facilities *p*, 8 nodes are selected as depots *d*, 67 biorefineries *b* and 113 customers *c*. Two transportation modes (t) are considered, truck and train. Only facilities located near the rails have access to the train. Biorefineries and depots are assumed to have access to rails. The case study scenario was located in the U.S. Data on the capacity for each preprocessing facility was obtained from USDA BioSISMap, based on biomass per county selecting 8 states, Arizona, California, Iowa, Nebraska, Nevada, New Mexico and South Dakota, data was filtered on having a crop residue production above 100,000 dry tons/Yr. Depots were selected centroids of the study mentioned in section 5.4. Biorefineries data on coordinates and size was acquired from Oak Ridge National Laboratory's biofuels database

utilizing Bioethanol Refineries located at the mentioned states, and filtered by their accessibility to railroad. Since the assumption made in the study is that biorefineries always have access to railroad. Customer's location and demand was based on gasoline consumption 2013 for each state, an approximation of consumption per county was done using total population per state and per county. Data on biofuels operations was gathered from Zhu and Yao (2011) studies, such as storage capacity and cost of operating depots, cost of operating a biorefinery, sale price of biofuel, the conversion equivalence of crop residue dry tons to gallons of biofuel, transportation by truck and train. The capacity of the truck is assumed to be 40 ton. The investment cost of considering a depot as hub was obtained from the studies of Roni, Md *et al* (2014). For further reference, the data collected and utilized for the study is located in Appendix C.

5.6.2 Results

The MOEA for the case study of hub and spoke biofuels supply chain considering maximization of profit and minimization of environmental impacts was run with a population size of 500 for 100 generations. A small percentage of mutation of 1% was introduced, as well an elitism factor of 25%. The algorithm provided Pareto-Optimal solutions for a yearly planning process design indicating which preprocessing facilities to open, as well as the amount of biomass to be preprocessed at each facility, followed by determining in which biorefineries this biomass is going to be converted into biofuel, and which depots will be established as hubs to increase the efficiency of the operations by decreasing the number of trips made to satisfy the demand of customers. Table 5.3 contains the resulting objectives of these solutions, while, Figure 5.11 graphically illustrates the Pareto Front.

Table 5.3: Pareto-Optimal Set

Solution	Profit	GWP	Acidification	Eutrophication
1	4.162E+10	6.81E+11	3458998015	2399375673
2	4.306E+10	5.35E+11	3147203040	2292705483
3	3.821E+10	7.42E+11	4598940203	3388364472
4	4.401E+10	5.71E+11	2889593677	1994942621
5	4.392E+10	6.21E+11	2729044771	1788265165
6	4.021E+10	7.21E+11	4333286217	3164812397
7	5.292E+10	5.57E+11	3014122424	2139069435
8	4.092E+10	6.52E+11	3580074825	2550778425
9	3.998E+10	6.87E+11	4367083207	3239813200
10	4.654E+10	5.79E+11	2716481309	1830628489
11	3.651E+10	6.77E+11	5404166222	4229425890
12	4.081E+10	6.15E+11	4257855632	3228522183

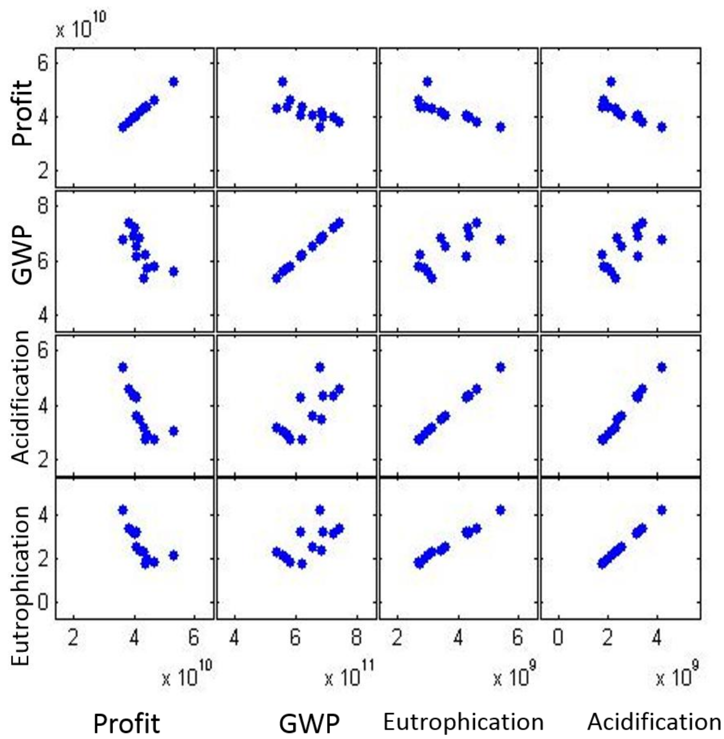

Figure 5.11: Bi-dimensional graphs of each objective with respect to the others

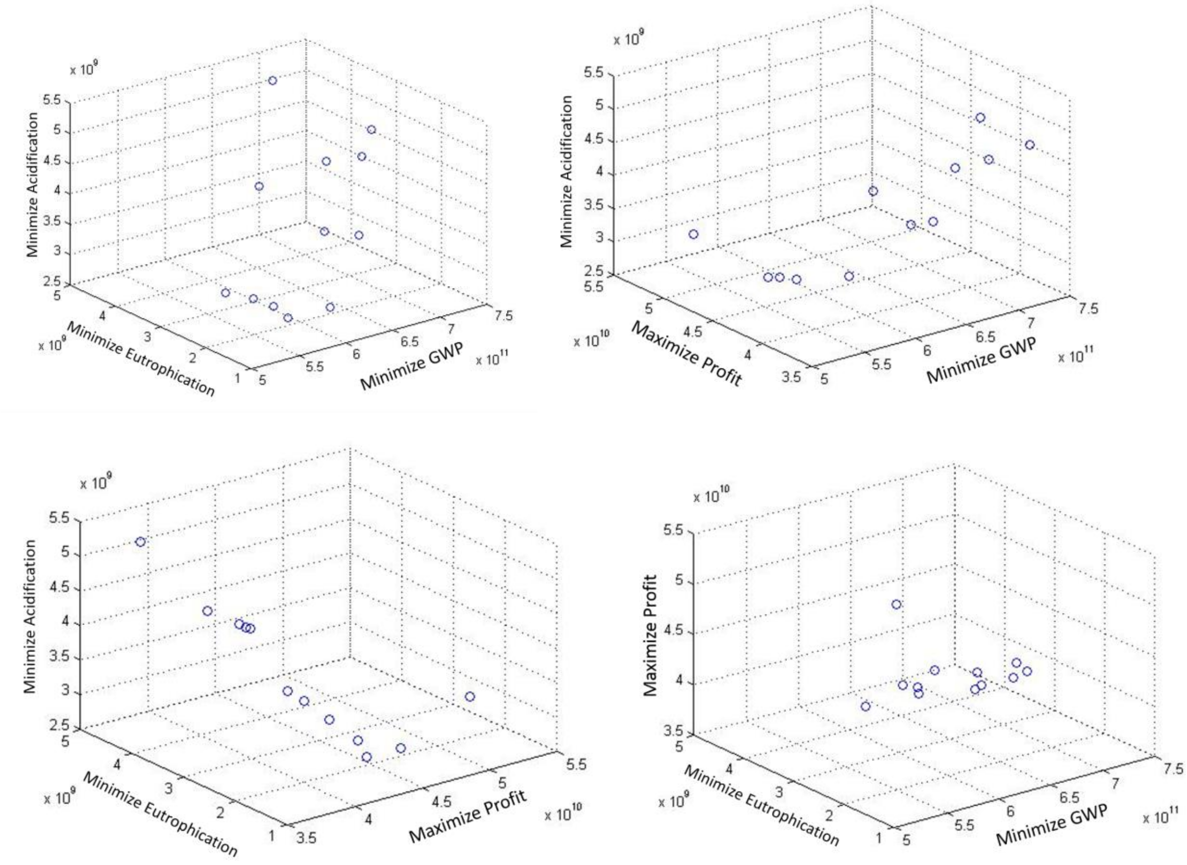

Figure 5.12: 3-dimensional plotting of Pareto-set

From these results it can be noticed a positive correlation among the environmental objective, whereas a negative correlation can be encountered when comparing each to profit. These negative correlations represent the fact that when environmental impacts reduce the profit increases. This increase in profit is the result of an optimized transportation system which transportation costs are reduced, causing as well the minimization of environmental impacts. A stronger correlation is found between Eutrophication and Acidification.

From these 12 Pareto optimal set the closest solution to the ideal point (1,0,0,0) was selected to demonstrate the potential decision tool developed in this work. The solution positioned at the 7th place of the set is the closest to the ideal point, the output of this solution is given in Table 5.4. and its corresponding objective values in Table 5.5.

Table 5.4: Solution 7 Closest to the Ideal Point (1,0,0,0)

Preprocessing Facilites (120 out of 233)														
3	4	5	6	7	11	13	14	15	17	19	22	24	26	27
28	31	32	35	36	37	40	43	44	45	47	48	50	52	53
55	56	57	58	59	62	64	66	68	72	73	75	78	80	82
86	87	89	91	99	101	104	105	106	108	109	110	111	113	114
118	119	120	121	122	123	124	125	126	128	130	131	133	135	137
138	139	142	145	146	148	152	156	158	159	160	162	163	167	168
170	171	172	175	179	180	184	185	186	187	188	189	191	195	196
206	210	211	212	213	215	217	218	219	221	222	226	229	231	233
Dry Tons of Biomass														
142840.3	53390.85	158206.3	166082.5	165500.7	17782.28	79919.15	127042.5	51780.44	111947.8	143564.1	166977.9	119795.7	114961.2	55087.41
190153.7	103490.3	51899.31	211799.3	12435.55	71273.35	169317.6	35845.02	204678.7	233495.9	182628.3	229659.1	66391.19	243266.5	174851.4
85736.41	10233.68	133461.2	35705.86	145959.9	118024.1	79952.73	91669.01	102706.5	260960.4	68607.82	112739.9	177660.5	270299.9	94035.16
16294.92	44024.93	244274.9	262924	39196.75	41660.49	9028.246	58659.32	92583.4	110038	32650.11	80921.39	37873.66	132567.7	50122.02
59433.56	75858.72	14938.3	117840.5	40890.02	143129.2	50855.45	125019	125070.5	90381.41	163014.2	73318.88	125228.1	118146.7	129106
128692.1	81943.05	80774.07	111709.8	14709.72	43483.93	17263.99	34677.23	133746.6	55623.84	154153.7	9107.023	152221.6	83114.83	141574.3
10157.95	189627.4	171896.7	31742.35	61510.51	138041.1	45381.59	118692.1	17335.04	71102.18	137448.4	186099.7	115905.2	121411.1	186313.6
32908.92	96090.38	2751.123	34440.11	98667.73	74202.28	78803.1	163218.5	140686.4	110625.7	57152.2	126084.2	80792.67	68551.69	225865
Biorefinery (33 out of 67)														
2	3	5	6	8	9	10	13	14	16	19	24	27	28	30
34	37	42	45	46	47	50	51	52	54	55	56	58	60	61
62	63	67												
Depots as Hubs (5 out of 8)														
1	3	4	5	7										

Table 5.5: Objective values of Solution 7

Solution
7
Profit
52915576336
GWP
5.56905E+11
Acidification
3014122424
Eutrophication
2139069435

These tables provide an illustration of a possible solution which recommends to open 120 preprocessing facilities out of the 233 possible, specifying which exactly to open, and how much dry tons of biomass to preprocess at each given in the green section of Table 5.4. Selecting next a total of 33 biorefineries necessary to satisfy the customer demand, as stated before each biorefinery has its own capacity and production limit. The shipments of biomass will be consolidated at depots 1, 3, 4, 5 and 7 which will be used as hubs. From this planning design we are expected to obtain a

total profit of 5.292E10, a global warming potential of 5.57E11 $CO_{2\text{-equivalent}}$, acidification 3.014E9 kg of SO_2 $_{\text{equivalents}}$, and eutrophication 2.139E9 kg of phosphate $_{\text{equivalents}}$.

The given allocation of each of the facilities within the United States is provided by Figure 5.13 which illustrates the majority of the customers located at California and the majority of preprocessing facilities and biorefineries near Iowa State.

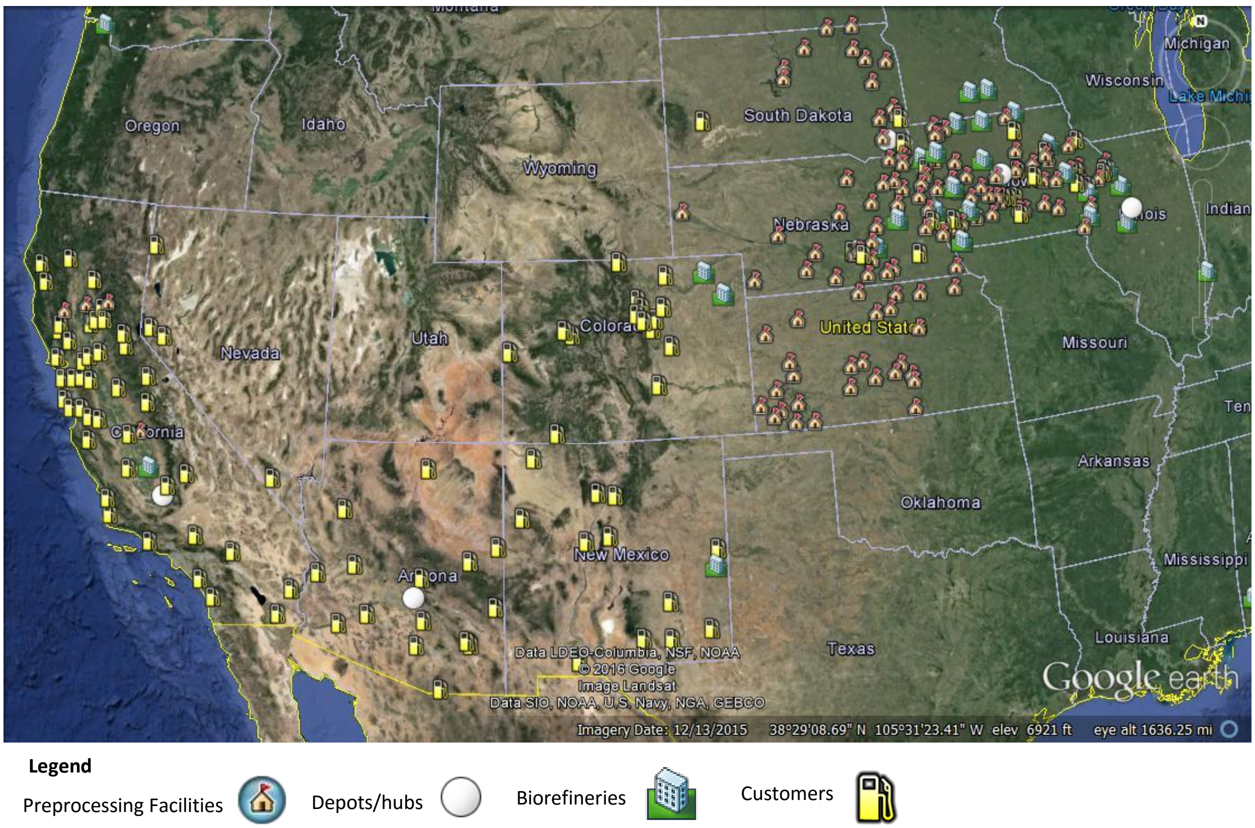

Figure 5.13: Solution 7 Facilities Location

Besides providing the planning configuration of the facilities allocation, the MOEA allocated the flow in the network, providing its optimal transportation design which is showed in Figure 5.14.

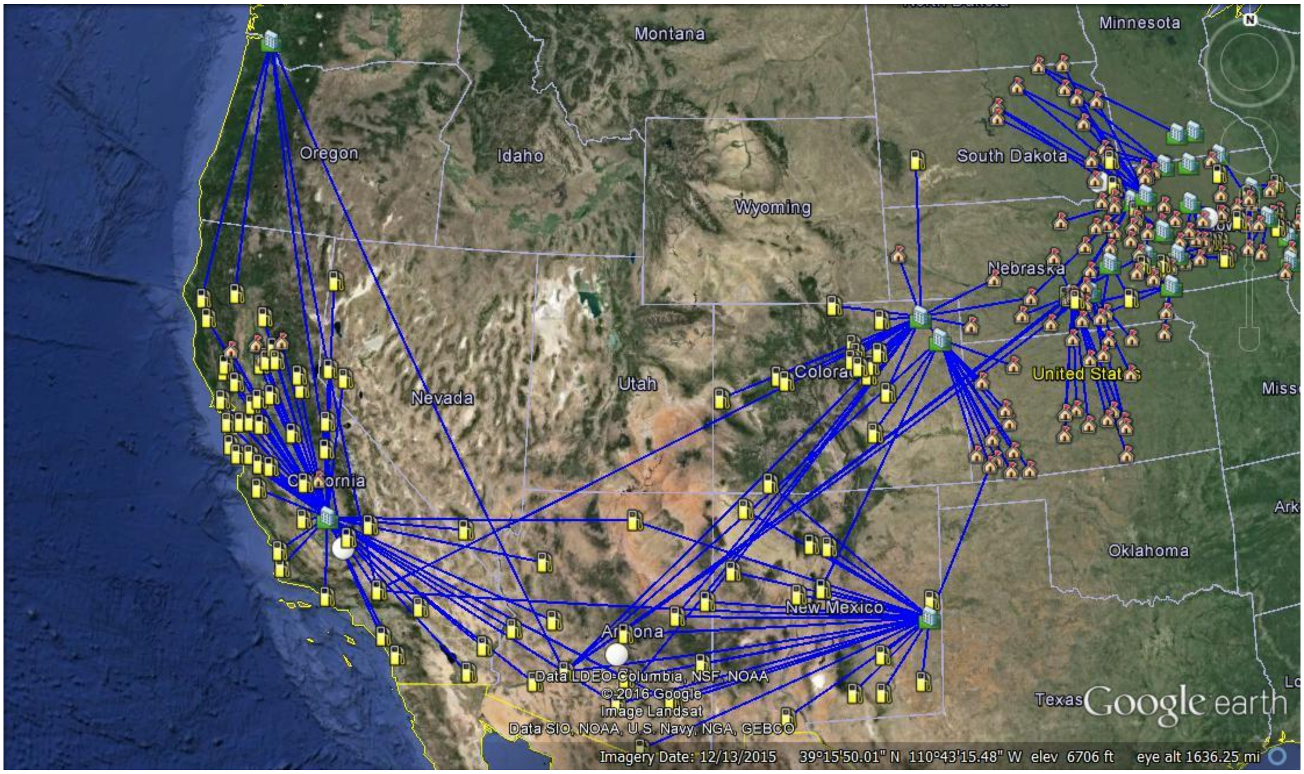

Figure 5.14: Solution 7 Routes for its Optimal Transportation Design

Chapter 6: Conclusions and Future Research

The environmental impacts of hub and spoke networks were analyzed with the purpose of developing sustainable transportation networks considering the studies of a centralized carrier collaboration and multi-hub location problem, in addition to a hub and spoke logistics design for a biomass to biofuel logistics system. Evolutionary algorithms were proposed as optimization tools to improve the corresponding objectives of each scenario, considering single and multiple objective evolutionary algorithms.

To first demonstrate the suitability of evolutionary algorithms in hub and spoke networks, a centralized collaborative carrier multi-hub location problem (CCCMLP) is introduced that provides a planning framework to analyze the benefits of a centralized multiple carrier collaborative network for the creation of hybrid hub-and-spoke system. It addresses the operational issues related to transfer locations and shipment consolidation by incorporating the concept of hybrid consolidation hubs from existing locations without the need to construct or invest new consolidation facility infrastructure. This is done by leveraging the current service locations of existing LTL collaborative carriers, synergized by novel opportunities provided through advances in ICT and e-commerce. An uncapacitated P-hub median location mathematical programming formulation was presented for a rate setting behavioral strategy for the collaborative system. The corresponding formulation was solved using a genetic algorithm for the single objective cases were cost and GWP were aimed to be minimized separately, whereas a Multiple Objective Evolutionary Algorithm was implemented for the bi-objective case were both objectives were considered simultaneously. The study results indicated that larger expected profit margins from the collaborative carriers under a revenue generating behavior would increase the likelihood of carriers collaborating. In addition, as the network size increases the effect of hybrid hub locational costs

was less. A key inference of this study is that carrier collaboration in terms of a collaborative hybrid hub-and-spoke system can become a critical strategy for small- to medium-sized LTL carriers to remain competitive. That is, by decreasing their operational costs when shipping across a point-to-point network.

Moreover, this first scenario consisted of a 10 node hybrid hub and spoke transportation network, where the algorithm was proved to be efficient as an optimization tool in transportation networks. When comparing the case of exhaustively searching for the optimal solution, a total of 1024 combinations were analyzed in a time of 199.94 s. whereas, the MOEA encountered the optimal solution through the evolution of fittest parents in a time of 47.449.

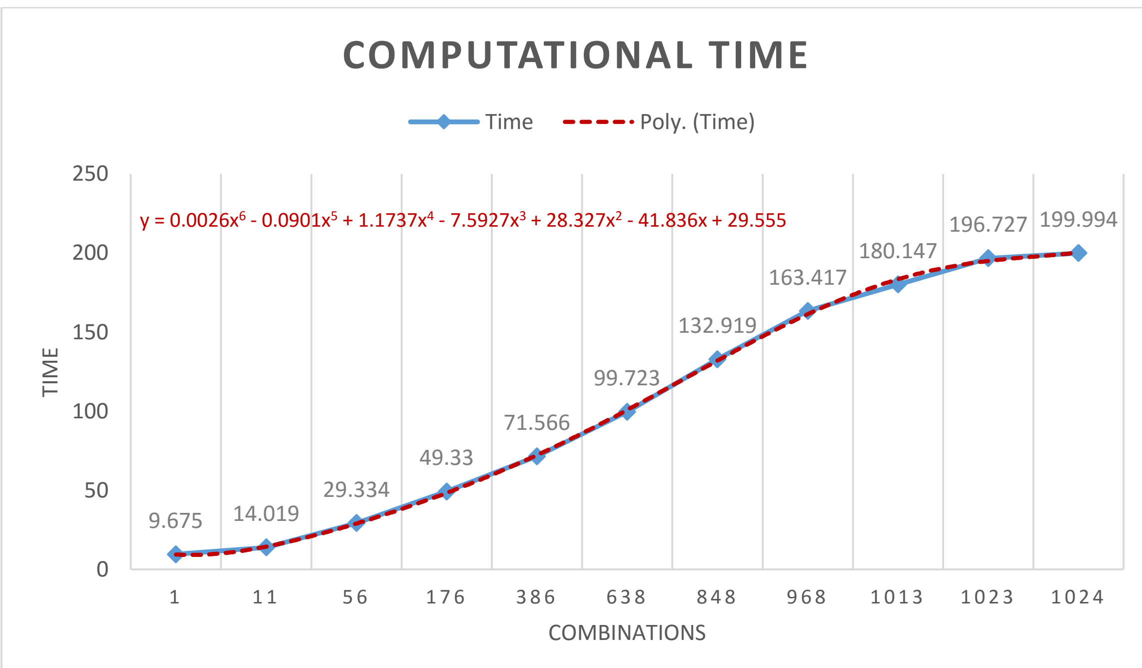

Figure 6.1: Exhaustive methodology computational time growth pattern

For the second scenario a network of 421 nodes was analyzed, considering only the binary configurations of the chromosome, a total of 5.415e126 minimum combinations are expected to be analyzed exhaustively, assuming the pattern for time will continue to be the same as shown in

Figure 6.1, a time of 2.774 e127s is expected to accomplish analyzing all possible combinations for this problem. Making exhaustive methodology impossible for the examination of this problem, and even more when the numeric portion of the chromosome is also considered. Therefore the second scenario can be considered an NP-hard problem. Proving the MOEA's accuracy and efficiency in the first scenario, demonstrates its suitability for the second scenario.

For the hub and spoke logistics biofuel design scenario, a multi-objective evolutionary algorithm was proposed as an approach to simultaneously address profit and environmental impacts. A mixed integer linear programming model was developed to provide the yearly operations configurations and transportation design of the biofuels supply chain. A single feedstock logistic system was considered efficiently allocating flow on network nodes to satisfy customer demands through a hybrid hub and spoke network. The additional objectives analyzed in this scenario are the minimization of the total global warming potential, minimization of acidification potential and the minimization of euthropication potential considering the best allocation of facilities and transportation route establishing depots as hubs that minimizes total greenhouse gas emissions measured in carbon dioxide equivalents (CO2equivalents), sulfur dioxide equivalents (SO2 equivalents) and phosphate equivalents (PO4 equivalents). The results obtained from the output of the MOEA consist of a Pareto-Optimal set conformed by diverse and robust solutions to serve as decision tool making for stakeholders according to their higher priorities.

Overall, the hub and spoke network can help in reducing the number of trips within a supply chain and hence minimizing the transportation cost and environmental impacts. Nonetheless, the proper arrangement of this networks can afford to improve even more their sustainability. In the meantime, the objective of this thesis is achieved by demonstrating the application of metaheuristic

approaches in transportation and supply chain problems and contributing to the transition of transportation sector into a more sustainable by implementing into the analysis three approaches stated by the U.S. Department of Transportation, which are, the use of low-carbon fuels, strategies to reduce the number of vehicle miles traveled, and optimizing the design of transportation networks to reduce trip frequencies. The different case studies in which the algorithms were employed reflected the performance of them by providing fast robust solutions and presenting improvements over their evolution.

The proposed thesis leaves an opportunity for future research that can potentially make improvements to the analysis. A possible expansion of the problem could entail the addition of a social objective function to develop a more comprehensive study of sustainability of hub and spoke networks. Social aspect can be measured by the number of local jobs resulting from the construction and operation of hub facilities within supply chains. Additionally, the biofuel study could be extended to more realistic case scenarios, by utilizing real data on the US railway network structure, as well as considering two or more carriers for the truck transportation incorporating a centralized carrier collaboration concept into the design. Moreover, other applications of hub and spoke networks can be studied regarding its environmental impacts, for instance in air and maritime industry. Furthermore, this problem could be solved making use of the variety of meta-heuristic algorithms, and an analysis comparing different algorithms could be implemented. Even though, Genetic Algorithms provided robust solutions, other methods could be employed, additionally Post-Pareto optimality methodologies could be implemented to reduce the size of the Pareto-optimal set to facilitate decision making.

References

[Abdinnour-Helm, Sue. (1998). A hybrid heuristic for the uncapacitated hub location problem. European Journal of Operational Research. Vol 106. SN 0377-2217]

[Abdinnour-Helm, Sue (2001),"Using simulated annealing to solve the p-Hub Median Problem", International Journal of Physical Distribution & Logistics Management, Vol. 31 Iss 3 pp. 203 – 220]

[American Association of State Highway and Transportation Officials (AASHTO). (2015). "2015 AASHTO Transportation Bottom Line Report". Executive Version.]

[Amit Kumar, Shahab Sokhansanj, Switchgrass (Panicum vigratum, L.) delivery to a biorefinery using integrated biomass supply analysis and logistics (IBSAL) model, Bioresource Technology, Volume 98, Issue 5, March 2007, pages 1033-1044, ISSN 0960-8524.]

[Arifa Sultana, Amit Kumar, Optimal configuration and combination of multiple lignocellulosic biomass feedstocks delivery to a biorefinery, Bioresource Technology, Volume 102, Issue 21, November 2011, Pages 9947-9956, ISSN 0960-8524]

[Aykin, T. (1995)3. Networking Policies for Hub and Spoke Systems with Applications to the Air Transportation System, Transportation Science, 29, 1995, pp. 201-221.]

[Aykin, Nuray A., Cunningham, James P. and Rotella, Joseph P. (1994) "Designing Operations Systems Interfaces That Are Easy to Use". AT&T Technical Journal. Vol 73. Blackwell Publishing Ltd. SN 1538-7305]

[B.D. Santer *et.al*. (1996). "A search for human influences on the thermal structure of the atmosphere," Nature vol 382, 4 July 1996, 39-46]

[B.D. Santer *et.al.* (2003). "Contributions of Anthropogenic and Natural Forcing to Recent Tropopause Height Changes," Science vol. 301 (25 July 2003), 479-483.]

[B. N. Gomes, A.X Martins, R.S. de Camargo, J.A. Ramirez. (2013). An efficient genetic algorithm for the design of Hub-and-Spoke Networks. IEEE Communications Letters. April 2013]

[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*, *21*(6), 640-650.]

[Caramia, M & Dell'Olmo, P. (2008). Multi-Objective Management in Freight Logistics Increasing Capacity, Service Level and Safety with Optimization Algorithms. XVI, 187 p.32]

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

[Charnes, A., Cooper, W. W., & Ferguson, R. O. (1955). Optimal estimation of executive compensation by linear programming. Management Science, 2, 138–151.]

[Chen, C. C. F., & Schonfeld, P. (2012). A hybrid heuristic technique for optimizing intermodal logistics timed transfer systems. Procedia-Social and Behavioral Sciences, 48, 2566-2576.]

[Chen, J.-F. (2007). A hybrid heuristic for the uncapacitated hub location problem. Omega The International Journal of Management Science, 35(2), 211-220]

[Cheng-Chang Ling and Sheu-Hua Chen (2008) An integral constrained generalized hub-and-spoke network design problem. Transportation Research Part E: Logistics and Transportation Review. Vol 44. 6. 986-1003]

[Cheng-Chang, Lin. (2010) The integrated secondary route network design model in the hierarchical hub-and-spoke network for dual express services. International Journal of Production Economics. Vol 123 (1)]

[Contreras, I., Cordeau, J., and Laporte, G. (2010). The Dynamic Uncapacitated Hub Location Problem. Transportation Science. Vol 45]

[Corne, David W., Knowles, Joshua D. & Oates, Martin J. (2000). The Pareto Envelope-Based Selection Algorithm for Multiobjective Optimization. Parallel Problem Solving from Nature PPSN VI: 6th International Conference Paris, France, September 18–20, 2000 Proceedings]

[Corne, David W., Jerram, Nick R., Knowles, Joshua D. & Oates, Martin J. (2011) PESA-II: Region-based Selection in Evolutionary Multiobjective Optimization. Proceedings of the Genetic and Evolutionary Computation Conference (GECCO). 2011. Pg 283-290. Morgan Kaufman Publishers]

[Cundiff, J., Dias, N. and Sherali, H. (1997) A linear programming approach for designing a herbaceous biomass delivery system. *Bioresource technology, 59*, 47–55.]

[Cunha, C.B., and M.R. Silva. A Genetic Algorithm for the Problem of Configuring a Hub-and-Spoke Network for a LTL trucking Company in Brazil. European Journal of Operational Research. 179(3), 2007, pp. 747-758.]

[D. Skorin-Kapov, J. Skorin-Kapov, M.E. O'Kelly, (1996) "Tight linear programming relaxations of uncapacitated p-hub median problems," European Journal of Operational Research, 94(3), 582-593.]

[DY Lin, KH Ng. (2012) The impact of collaborative backhaul routing on carbon reduction in the freight industry. Transportation Research Part D: Transport and Environment 17(8), 626-628]

[De Mello, P. F. B., & Frayret, J. M. (2014). Agent-Based Approach of Sustainability Assessment of Resource Sharing in Freight Transportation.]

[Ebery, J. Krishnamoorthy, M., Ernst, A., Boland, N. (2000) The capacitated multiple allocation hub location problem: Formulations and algorithms, European Journal of Operational Research, Volume 120, Issue 3, 1 February 2000, Pages 614-631, ISSN 0377-2217]

[Eckart Zitzler and Lothar Thiele (1998). An Evolutionary Algorithm for Multiobjective Optimization: The Strength Pareto Approach]

[Eckart Zitzler, Marco Laumanns & Lothar Thiele. (2001) SPEA2: Improving the Strenght Pareto Evolutionary Algorithm.]

[Eckart Zitzler, Lothar Thiele, Marco Laumanns, Carlos M. Fonseca and Viviane Grunert da Fonseca (2002) Performance Assessment of Multiobjective Optimizers: An Analysis and Review. IEEE Transactions on Evolutionary Computation, volume 7, pages. 117-132]

[Elhedhli, S., Wu, H., (2010) A lagrangean heuristic for Hub and Spoke System Design with Capacity Selection and Congestion, Informs Journal on Computing, 22(2), 282 - 296, 2010.]

[Environmental Protection Agency. (2014). “U.S. GHG Inventory (2014) Executive Summary.]

[Epplin, F., Clark, C., Roberts, R., and Hwang, S. (2007) Challenges to the Development of a Dedicated Energy Crop.*Am. J. Agr. Econ*. (2007) 89 (5): 1296-1302]

[Ernst, A.T., Krishnamoorthy, M. (1999). “Solution algorithms for the capacitated single allocation hub location problem”. Annals of Operations Research. Vol 86. SN 1572-9338.]

[European Commission. (2014). “Climate Action- Reducing emissions from transport”.]

[Fan Wang, Xiaofan Lai, Ning Shi, A multi-objective optimization for green supply chain network design, Decision Support Systems, Volume 51, Issue 2, May 2011, pages 262-269, ISSN 0167-9236]

[Fengli Zhang, Dana M. Johnson, Mark A. Johnson,(2012). Development of a simulation model of biomass supply chain for biofuel production, Renewable Energy, Vol. 44, August 2012, pages 380-391, ISSN 0960-1481]

[Fishburn, Peter C., (1974), Exceptional Paper--Lexicographic Orders, Utilities and Decision Rules: A Survey, *Management Science*, 20, issue 11, p. 1442-1471,]

[Gelareh, Shahin and Nickel, Stefan (2011) Hub location problems in transportation networks. Transportation Research Part E: Logistics and Transportation Review. Vol 47, Is 6, pp-1092-1111]

[H.Topcuoglu, F. Corut, M. Ermis, G. Yilmaz. (2005) “ Solving the uncapacitated hub location problema using genetic algorithms.Journal COmputers and Operations Research. Volume 32 Issue 4, April 2005, Pages 967-984. Elsevier Science Ltd. Oxford, UK, UK]

[Hernández, S., and S. Peeta.(2010) Less-Than-Truckload Static Single Carrier Collaboration Problem. CD-ROM. Transportation Research Board of the National Academies, Washington, D.C., 2010.]

[Hernandez, S., Unnikrishnan, A., & Awale, S. (2011a). A less-than-truckload centralized carrier collaboration multi-hub location problem: a hybrid hub-and-spoke network. Transport. Res. Rec. J. Transport. Res. Board, 2263, 20-28.]

[Hernandez, S., S. Peeta, and G. Kalafatas.(2011b) A Less-Than-Truckload Carrier Collaboration Planning Problem under Dynamic Capacities. Transportation Research Part E: Logistics and Transportation Review, 47(6), 2011, pp. 933-946.]

[Hernández, S., Unnikrishnan, A., & Awale, S. S. (2012). Centralized Carrier Collaboration Multihub Location Problem for Less-Than-Truckload Industry. Transportation Research Record: Journal of the Transportation Research Board, 2269(1), 20-28.]

[Horn, J.; Nafpliotis, N.; Goldberg, D.E.,(1994) "A niched Pareto genetic algorithm for multiobjective optimization," in *Evolutionary Computation, 1994. IEEE World Congress on Computational Intelligence., Proceedings of the First IEEE Conference on* , vol., no., pp.82-87 vol.1, 27-29 Jun 1994]

[Hyun Kim and M.E. O'Kelly. (2009). Reliable p-hub location problems in telecommunication networks. Geographical Analysis, 41(3), 283-306.]

[IPCC Fourth Assessment Report, Summary for Policymakers, p. 5]

[J.F. Campbell, G. Miranda, Jr., R.S. Camargo, R.S., M.E. O'Kelly, (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.]

[J. F. Campbell, A. T. Ernst and M. Krishnamoorthy. (2005). Hub Arc Location Problems: Part I- Introduction and Results. Management Science. 0025-1909.]

[Jaegler, A., & Burlat, P. (2012). Carbon friendly supply chains: a simulation study of different scenarios. *Production Planning & Control*, *23*(4), 269-278.]

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

[Jiuh-Biing Sheu, Yi-Hwa Chou, Chun-Chia Hu, An integrated logistics operational model for green-supply chain management, Transportation Research Part E: Logistics and Transportation Review, Volume 41, Issue 4, July 2005, pages 287-313, ISSN 1366-5545.]

[John Gunnar Carlsson, Fan Jia (2013) Euclidean Hub-and-Spoke Networks. Operations Research 61(6):1360-1382.]

[Judd, J., Sarin, S., Cundiff, J. S. and Grisso, R. D. (2010) An Optimal Storage and Transportation System for a Background / Motivation. In *2010 ASABE Annual International Meeting*; Vol. 0300, pp. 1–15.]

[Judd, J. D., Cundiff, J. S., Bish, D. R. and Koelling, C. P. (2012) Modeling and Analysis of a Feedstock Logistics Problem Modeling and Analysis of a Feedstock Logistics Problem, pp. 1–202.]

[Kalyanmoy Deb, Amrit Pratap, Sameer Arwal, and T. Meyariyan. (2002) A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II. IEEE Transactions on evolutionary computation, Vol. 6, No. 2, April 2002.]

[Kim, J.H., Soh, S. (2012) Designing hub-and-spoke school bus transportation network: a case study of wonkwang university. Promet-Traffic & Transportation 24(5),389-394]

[Klincewicz, J. G. (1992) Avoiding local optima in the p-hub location problem using tabu search and grasp , Annals of Operational Research, 40(1-4), 283 - 302, 1992.]

[Klincewicz J. G. (1996). A dual algorithm for the uncapacitated hub location problem. Location Science, 4(3);173-184.]

[Knowles, J.; Corne, D., (1999)"The Pareto archived evolution strategy: a new baseline algorithm for Pareto

multiobjective optimisation," in *Evolutionary Computation, 1999. CEC 99. Proceedings of the 1999 Congress on* , vol.1, no., pp.105 Vol. 1, 1999]

[Kumar, M., Hussian, M., Upreti, N., and Gupta, D. (2010) Genetic Algorithm Review and Application. *International Journal of Information Technology and Knowledge Management.* July-December 2010, Volume 2, No. 2, pp. 451-454]

[Li Sun, Lindu Zhao (2009) Research on optimization of hub-and-spoke logistics network design with impedance effect. POMS 20th Annual Conference.]

[M.E. O'Kelly (1987) A quadratic integer problem for the location of interacting hub facilities. European Journal of Operational Research, 32:393-404]

[M. E. , O'Kelly and Lao, Yong (1991) Mode Choice in a Hub-and-Spoke Network: A Zero-One Linear Programming Approach. Geographical Analysis. Vol 23. Blackwell Publishing Ltd. SN1538-4632]

[M.E. O'Kelly, (1992a) "Hub facility location with fixed costs," Papers in Regional Science, 71(3), 293-306.]

[M.E. O'Kelly, (1992b) "A clustering approach to the planar hub location problem," Annals of Operations Research, 40, 339-353.]

[M.E. O'Kelly. (2009) Rectilinear minimax hub location. Journal of Geographical Systems, 11(3), 227-241.]

[M.E. O'Kelly. (2012). Fuel Burn and Environmental Implications of Airline Hub Networks Transportation Research D, 17(7), 555-567.]

[M.E. O'Kelly, James F. Campbell, Ricardo S. de Camargo, Gilberto de Miranda Jr. (2014)Multiple Allocation Hub Location Model with Fixed Arc Costs. Geographical Analysis 47(1),73–96.]

[M.J. Kuby and R.G. Gray (1993) Hub network design problem with stopovers and feeders: Case of Federal Express. Transportation Research, 27:1-12]

[M. Mohammadi, R. Tavakkoli-Moghaddam, H. Rostami (2011) A multi-objective imperialist competitive algorithm for a capacitated hub covering location problem. International J Journal of Industrial Engineering Computations. ISSN 1923-2934. Vol. 2. Issue 3 pp.671-688]

[MacQueen, J. (1967) Some methods for classification and analysis of multivariate observations, Proceedings of the Fifth Berkeley Symposium On Mathematical Statistics and Probabilities, 1, 281-296.]

[Mapemba, L., Epplin, F., Taliaferro, F., and Huhnke, R.(2007) Biorefinery Feedstock Production on Conservation Reserve Program Land Lawrence D. Mapemba, Francis M. Epplin, Charles M. Taliaferro and Raymond L. Huhnke *Review of Agricultural Economics* Vol. 29, No. 2 pp. 227-246]

[Maurizio Bruglieri, Leo Liberti. (2008) Optimal running and planning of a biomass-based energy production process. *Energy Policy*. Volume 36, Issue 7, July 2008, Pages 2430-2438, ISSN 0301-4215]

[Meng, Qiang and Wang, Xinchang, (2011), Intermodal hub-and-spoke network design: Incorporating multiple stakeholders and multi-type containers, *Transportation Research Part B: Methodological*, 45, issue 4, p. 724-742,]

[Miranda Junior, Gilberto de, Camargo, Ricardo Saraiva de, Pinto, Luiz Ricardo, Conceição, Samuel Vieira, & Ferreira, Ricardo Poley Martins. (2011). Hub location under hub congestion and demand uncertainty: the Brazilian case study. Pesquisa Operacional, 31(2), 319-349.]

[Murata, T.; Ishibuchi, H., "MOGA: multi-objective genetic algorithms," in *Evolutionary Computation, 1995., IEEE International Conference on* , vol.1, no., pp.289-, Nov. 29 1995-Dec. 1 1995]

[Osman, Ibrahim H. & Laporte, Gilbert (1996). Metaheuristics: A bibliography. Annals of Operations research. Vol 63. Baltzer Science Publishers, Baarn/Kluwer Academic Publishers. SN0254-5330]

[Okan ¨Orsan ¨Ozener, Ozlem Ergun, Martin Savelsbergh (2007). Collaboration for Truckloads Carriers. Transportation Science]

[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, 143*(1), 86-94.]

[Pérez-Bernabeu, E., Juan, A. A., Faulin, J. and Barrios, B. B. (2015), Horizontal cooperation in road transportation: a case illustrating savings in distances and greenhouse gas emissions. International Transactions in Operational Research, 22: 585–606.]

[Pirkul H, Schilling DA (1998). "An efficient procedure for designing single allocation hub and spoke systems. Manage Sci 44(12):235-242]

[R.S. de Camargo, G. Miranda Jr., R.P.M. Ferreira, H.P. Luna. (2009) Multiple allocation hub-and-spoke network design under hub congestion. Journal Computers and Operations Research. Volume 36 3106Issue 12, December 2009, Pages 3097-3106]

[Rafay Ishfaq and Charles R. Sox (2012) Design of intermodal logistics networks with hub delays. European Journal of Operations Research. Vol.220 (3) pp. 629-641]

[Roni, Md Sadekuzzaman. (2013)Analyzing the impact of a hub and spoke supply chain design for long-haul, high-volume transportation of densified biomass. Pages144.]

[Roni, Md., Sandra D. Eksioglu, Kara G. Cafferty. (2014). *A Multi-Objective, Hub-and-Spoke Supply Chain Design Model For Densified Biomass.*Washington, D.C.: United States. Dept. of Energy.]

[Roni, Mohammad S., Eksioglu, Sandra D., Cafferty, Kara G., Jacobson, Jacob J. (2016). A multi-objective, hub-and-spoke model to design and manage biofuel supply chains. Annals of Operations Research.]

[Rosseuw, Peter J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics. Vol 20 pp. 53-65]

[Sadegheih, A., Drake, P. R., Li, D., & Sribenjachot, S. (2011). Global supply chain management under the carbon emission trading program using mixed integer programming and genetic algorithm. *International Journal of Engineering, Transactions B: Applications*, *24*(1), 37-53.]

[Sadegheih, A., Drake, P. R., Li, D., & Sribenjachot, S. (2011). Global supply chain management under the carbon emission trading program using mixed integer programming and genetic algorithm. *International Journal of Engineering, Transactions B: Applications*, *24*(1), 37-53.]

[Sara Thorsell, Francis M. Epplin, Raymond L. Huhnke, Charles M. Taliaferro. (2004) Economics of a coordinated biorefinery feedstock harvest system: lignocellulosic biomass harvest cost. *Biomass and Bioenergy*, Volume 27, Issue 4, October 2004, Pages 327-337, ISSN 0961-9534,]

[Shaofeng Liu, Dulekha Kasturiratne, Jonathan Moizer. (2012) A hub-and-spoke model for multi-dimensional integration of green marketing and sustainable supply chain management. Industrial Marketing Management 41 (2012) 581-588]

[Taboada, H. and Coit, D (2012). A New Multiple Objective Evolutionary Algorithm for Reliability Optimization of Series-Parallel Systems. *International Journal of Applied Evolutionary Computation*, 3(2).]

[Taha, Tarek T. and Taylor, G. Don and Taha, Hamdy A. (1996). A Simulation-Based Software System for Evaluating Hub-and-Spoke Transportation Networks. Simulation Practice and Theory. Vol.3, No. 6, pp 327-346]

[Tembo, G., Epplin, F. and Huhnke, R. (2003) Integrative investment appraisal of a lignocellulosic biomass-to-ethanol industry. *Journal of Agricultural and Resource Economics*, *28*, 611–633.]

[U.S. Department of Transportation,Center for Climate Change and Environmental Forecasting. (2007). "Transportation and Greenhouse Gas Emissions"]

[U.S. Energy Information Administration. (2015). “How much carbon dioxide is produced by burning gasoline and diesel fuel?” As of June 25, 2015; Preliminary estimates.]

[United Nations Department of Economic and Social Affairs/Population Division. (2015). World Population Prospects: The 2015 Revision, Key Findings and Advance Tables]

[Vidyarthi, N., Kumar, Rajesh, Jayaswal, S. (2013) Hub-and-spoke system design for freight transportation with priority consignment classes. Logistics and Transportation.]

[W. Alex Marvin, Lanny D. Schmidt, Saif Benjaafar, Douglas G. Tiffany, Prodromos Daoutidis. (2012). Economic Optimization of a Lignocellulosic Biomass-to-Ethanol Supply Chain. *Chemical Engineering Science*, Volume 67, Issue 1, 1 January 2012, Pages 68-79, ISSN 0009-2509,]

[World Health Organization (WHO). (2012). Our Planet, Our Health, Our Future. Human health and the Rio Conventions: biological diversity, climate change and desertification. Discussion Paper. 2012]

[Xiaozhou Xu.(2013) Collaboration Mechanism in the Horizontal Logistics Collaboration. Business administration. Ecole Nationale Supérieure des Mines de Paris, 2013. English.]

[Yancang, L., Lina Z., Shujing, Z. Review of Genetic Algorithm. *International Journal of Information Technology and Knowledge Management*. July-December 2010, Volume 2, No. 2, pp. 451-454]

[Yogendra N. Shastri, Zewei Miao, Luis F. Rodriguez, Tony E. Grift, Alan C. Hassen, and K.C. Ting. (2014) Determining optimal size reduction and densification for biomass feedstock using the BioFeed optimization model. Energy Biosciences Institute & Department of Agricultural and Biological Engineering]

[You, Fengqi and Wang, Belinda. (2011) Life Cycle Optimization of Biomass-to-Liquid Supply Chains with Distributed-Centralized Processing Networks. Industrial & Engineering Chemistry Research. Vol 50.]

[Yu, E. T., Wang, Z., English, B. C., Larson, J. A. 2014. “Optimization of Cost and Greenhouse Gas Emissions of a Dedicated Crop Supply System to Biorefineries in Tennessee”. Journal of Agricultural and Applied Economics, 46,3(August 2014):357–373.]

[Zhang J., Y. Wu, and P. Liu. (2007) Routing Problem for Hybrid Hub-and-Spoke Transportation Network: A Case Study of a LTL Carrier. Proceedings of the IEEE International Conference on Automation and Logistics, 2007.]

[Zhu, X. and Yao, Q. (2011) Logistics system design for biomass-to-bioenergy industry with multiple types of feedstocks. *Bioresource technology*, 102, 10936–45.]

[Zhu, X., Li, X., Yao, Q. and Chen, Y. (2011) Challenges and models in supporting logistics system design for dedicated-biomass-based bioenergy industry. *Bioresource technology*, *102*, 1344–51]

Appendix A: CCCMLP Data for Case Study 1

Table A.1: Binary Variables for Case study 1

Binary Variables	
Y_{ijklq}	Determines if a shipment is sent through a collaborative route
V_{ijq}	Determines if a shipment is sent through a non-collaborative route
X_k	Determines if the node at point k is to become a hub

Table A.2: Constants for case study CCCMLP

Constants	
W_q	Non-Collaborative Carrier Cost Matrix
C_q	Collaborative Carrier Cost Matrix
P_{kq}	Hub Establishment Cost Matrix
d_q	Demand Matrix

Table A.3: Assumed demand for Carrier 1

Demand Carrier 1										
Nodes	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	0	495	291	855	608	172	272	490	701	651
2	415	0	282	554	249	296	505	607	396	369
3	262	249	0	599	447	304	310	339	407	584
4	916	596	577	0	734	939	931	774	274	774
5	572	226	495	685	0	330	798	760	543	119
6	162	303	297	816	375	0	371	551	594	460
7	257	529	285	851	653	389	0	270	700	848
8	463	563	357	758	832	628	255	0	463	947
9	762	400	425	235	537	615	566	436	0	750
10	679	348	617	765	140	405	793	999	726	0

Table A.4: Assumed demand for Carrier 2

Demand Carrier 2										
Nodes	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	0	495	291	978	499	198	240	463	716	623
2	450	0	285	560	249	292	535	512	338	382
3	313	288	0	671	533	262	300	357	367	557
4	844	499	627	0	653	873	921	766	277	867
5	566	277	539	653	0	379	742	769	590	124
6	162	306	288	778	375	0	407	525	580	495
7	283	505	278	960	734	429	0	264	686	904
8	452	544	303	709	823	557	285	0	469	865
9	624	355	416	229	590	594	566	528	0	685
10	609	395	563	765	127	410	765	927	718	0

Table A.5: Hub establishing cost for carriers 1, 2.

Nodes	**Hub cost for carrier 1**	**Hub cost for carrier 2**
1	24685	24931.9
2	20644.9	21883.6
3	19934.5	16745
4	35574.2	29526.5
5	25805.3	22966.7
6	22662	20849.1
7	26899	27437
8	28702.9	24971.5
9	26668.6	23201.7
10	30581	34250.7

Table A.6: Non-collaborative shipping costs for Carrier 1

	Non-collaborative cost for Carrier 1									
Nodes	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	0	134	83	284	182	60	74	134	212	197
2	137	0	80	166	85	88	182	171	106	103
3	84	79	0	188	139	91	85	88	119	197
4	284	166	212	0	220	276	270	218	81	238
5	161	69	149	213	0	109	210	241	187	39

6	52	88	82	265	122	0	107	186	187	144
7	76	153	83	244	223	131	0	75	189	230
8	150	154	103	235	220	156	76	0	134	256
9	196	103	130	79	169	182	208	147	0	210
10	185	116	177	241	37	129	230	287	198	0

Table A.7: Non-collaborative shipping costs for Carrier 2

Non-collaborative cost for Carrier 2										
Nodes	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	0	134	92	300	150	58	76	157	190	202
2	126	0	78	151	81	93	181	182	101	104
3	78	90	0	193	131	85	90	107	119	167
4	253	153	216	0	210	273	279	193	81	263
5	172	79	134	227	0	113	230	217	193	35
6	56	107	89	256	109	0	126	190	176	126
7	71	159	86	267	218	123	0	90	187	254
8	157	159	89	223	217	177	86	0	139	269
9	206	119	118	74	179	172	195	134	0	213
10	174	112	169	230	34	140	274	253	196	0

Table A.8: Collaborative shipping costs for Carriers 1, 2.

Collaborative Cost										
Nodes	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	0	46	31	101	60	20	26	50	70	66
2	48	0	28	55	27	35	60	61	38	40
3	31	27	0	66	50	29	31	33	43	60
4	101	57	69	0	78	88	92	75	27	89
5	58	28	48	76	0	38	79	87	67	14
6	19	36	31	86	39	0	42	60	65	47
7	28	60	29	90	78	43	0	28	64	86
8	54	61	35	74	86	61	27	0	48	98
9	72	40	42	27	63	69	69	50	0	75
10	67	39	63	86	14	46	83	102	77	0

Appendix B: CCCMLP Data for Case Study 2

Table B.1: Constants for Case Study 2 Global Warming Potential

Constants	
W_q	Non-Collaborative Carrier GWP Matrix
C_q	Collaborative Carrier GWP Matrix

Table B.2 : GWP for Carrier 1

	GWP 100 kg-CO2 equivalent for Carrier 1									
\	1	2	3	4	5	6	7	8	9	10
1	0	0.043	0.026634	0.091134	0.058403	0.019254	0.023746	0.043	0.06803	0.063216
2	0.043963	0	0.025672	0.053269	0.027276	0.028239	0.058403	0.054873	0.034015	0.033052
3	0.026955	0.025351	0	0.060328	0.044604	0.029201	0.027276	0.028239	0.038187	0.063216
4	0.091134	0.053269	0.06803	0	0.070597	0.088567	0.086642	0.069955	0.025993	0.076373
5	0.051664	0.022142	0.047813	0.068351	0	0.034978	0.067388	0.077336	0.060007	0.012515
6	0.016687	0.028239	0.026313	0.085037	0.039149	0	0.034336	0.059687	0.060007	0.046209
7	0.024388	0.049097	0.026634	0.078299	0.07156	0.042037	0	0.024067	0.060649	0.073806
8	0.048134	0.049418	0.033052	0.07541	0.070597	0.05006	0.024388	0	0.043	0.082149
9	0.062896	0.033052	0.041716	0.025351	0.054231	0.058403	0.066746	0.047172	0	0.067388
10	0.059366	0.037224	0.056799	0.077336	0.011873	0.041396	0.073806	0.092097	0.063537	0

Table B.3: GWP for Carrier 2

GWP 100 kg-CO2 equivalent for Carrier 2										
\	1	2	3	4	5	6	7	8	9	10
1	0	0.041137	0.028243	0.092097	0.046049	0.017805	0.023331	0.048197	0.058328	0.062012
2	0.038681	0	0.023945	0.046356	0.024866	0.02855	0.055565	0.055872	0.031006	0.031927
3	0.023945	0.027629	0	0.059249	0.040216	0.026094	0.027629	0.032848	0.036532	0.051267
4	0.077668	0.046969	0.06631	0	0.064468	0.083808	0.08565	0.059249	0.024866	0.080738
5	0.052802	0.024252	0.041137	0.069687	0	0.03469	0.070608	0.066617	0.059249	0.010745
6	0.017191	0.032848	0.027322	0.078589	0.033462	0	0.038681	0.058328	0.05403	0.038681
7	0.021796	0.048811	0.026401	0.081966	0.066924	0.03776	0	0.027629	0.057407	0.077975
8	0.048197	0.048811	0.027322	0.068459	0.066617	0.054337	0.026401	0	0.042672	0.08258
9	0.06324	0.036532	0.036225	0.022717	0.054951	0.052802	0.059863	0.041137	0	0.065389

10	0.053416	0.034383	0.051881	0.070608	0.010438	0.042979	0.084115	0.077668	0.06017	0

Table B.4: Collaborative GWP

GWP 100 kg-CO2 equivalent for collaboration between Carriers 1 & 2										
\	1	2	3	4	5	6	7	8	9	10
1	0	0.020767	0.013995	0.045597	0.027087	0.009029	0.011738	0.022573	0.031602	0.029796
2	0.02167	0	0.012641	0.02483	0.012189	0.015801	0.027087	0.027539	0.017155	0.018058
3	0.013995	0.012189	0	0.029796	0.022573	0.013092	0.013995	0.014898	0.019413	0.027087
4	0.045597	0.025733	0.03115	0	0.035214	0.039728	0.041534	0.033859	0.012189	0.04018
5	0.026184	0.012641	0.02167	0.034311	0	0.017155	0.035665	0.039277	0.030248	0.00632
6	0.008578	0.016252	0.013995	0.038825	0.017607	0	0.018961	0.027087	0.029345	0.021218
7	0.012641	0.027087	0.013092	0.040631	0.035214	0.019413	0	0.012641	0.028893	0.038825
8	0.024379	0.027539	0.015801	0.033408	0.038825	0.027539	0.012189	0	0.02167	0.044243
9	0.032505	0.018058	0.018961	0.012189	0.028442	0.03115	0.03115	0.022573	0	0.033859
10	0.030248	0.017607	0.028442	0.038825	0.00632	0.020767	0.037471	0.046049	0.034762	0

Appedix C: Biomass-to-Biofuels: Data for Case Study 3

Variable Definition

Table C.1: Decision Variables

Decision Variables	
p	Preprocessing facilities (p =1:233)
b	Biorefinery (b =1:67)
d	Depots (d=1:8)
t	Transportation mode (t=1:truck, 2:train)
i	Origin of a shipment
j	Destination of a shipment
N	Total number of nodes in network
BFG	Gallons of biofuel
DTP	Dry tons of biomass processed
DTC	Consolidated dry tons

Table C.2: Binary Variables

Binary Variables	
Y_{ijklt}	1, if a shipment is sent from node i to node j via the hubs k and l by the transportation mode t. That is, if the shipment is sent through a multi-modal network. Otherwise, it will be equal to 0.
V_{ijt}	1, if a shipment is sent directly from node i to node j by transportation mode t, and 0 otherwise.
K_p	1, if the node of a preprocessing facility p is open, 0 otherwise.
Z_b	1, if the node of a biorefinery b is open, 0 otherwise.
X_k	1, if the node at point k will become a hub, and 0 otherwise.

Table C.3: Fixed Variables

Fixed Variables	
c	Customers (c=1:113)

ρ	Price gallon of biofuel
PC	Processing cost of DT of biomass
TC_t	Transportation cost per transportation mode t
OD	Cost of operating a depot d
OB	Cost of operating a biorefinery b
$BCAP_b$	Production capacity at biorefinery b
$SCAP_d$	Storage capacity of depot d
$PCAP_p$	Preprocessing Capacity at preprocessing facility p
d_c	Demand of customer c
ϑ_{kt}	Cost associated with additional tracks
$\emptyset_k$	Cost associated with turnout to allow rail cars to switch tracks
W_{ijt}	Cost of shipping directly

DATA

Table C.4: Fixed Costs, prices and fixed capacities used in case study 3

Fixed Costs and Prices	
ρ	\$1.8/gallon
PC	\$50/dry ton
$TC_{t=1}$	Truck: 0.4 per mile/dry ton
$TC_{t=2}$	Train: \$0.04 per mile/dry ton
ϑ_{kt}	\$717.80/yard
$\emptyset_k$	110,000.00
OD	\$60,000 per month
OB	\$10,000,000 per year
$SCAP_d$	200,000 dry tons/month

Table C.5: Preprocessing facilities and its capacity

ID	ORG_NAME	STATUS	COUNTY	ST	LAT	LON	$BCAP_b$
1	KAAPA Ethanol, LLC	Operational	Kearney	NE	41.424175	-97.2904	59
2	Advanced BioEnergy LLC/Heartland grain	Operational	Beadle	SD	41.424175	-97.2904	30
3	Glacial Lakes Energy, LLC	Operational	Codington	SD	40.815486	-98.60757	100
4	Glacial Lakes Energy, LLC	Operational	Edmunds	SD	40.974488	-101.3547	100
5	North Country Ethanol,LLC	Operational	Roberts	SD	41.520554	-96.07923	25
6	Nebraska Energy (Aventine)	Operational	Hamilton	NE	41.698984	-73.19228	50
7	Trenton Agri Products, LLC	Operational	Hitchccock	NE	39.203076	-93.47559	40
8	Midwest Renewable Energy, LLC	Operational	Lincoln	NE	44.109417	-93.67326	27
9	East Kansas Agri-Energy, LLC	Operational	Anderson	KS	39.007087	-87.62682	35
10	Western Plains Agri Energy, LLC	Operational	Logan	KS	40.531003	-90.0004	40
11	Mid America Agri Products/Wheatland	Operational	Perkins	NE	43.670726	-92.09721	40
12	Pinal Energy, LLC	Operating	Pinal	AZ	33.021418	-111.9925	50
13	Calgren Renewable Fuels, LLC (Pixley Ethanol)	Producing	Tulare	CA	36.002042	-119.3037	52.5
14	White Energy	Operational	Russell	KS	46.170377	-123.1524	53
15	Chief Ethanol Fuels, Inc.	Operational	Adams	NE	43.81409	-95.29521	62
16	Cargill, Inc.	operating	Wapello	IA	31.122697	-84.15445	35
17	Elkhorn Valley Ethanol, LLC (Louis Dreyfus)	Operational	Madison	NE	43.57525	-93.36842	40

18	Advanced BioEnergy LLC/Heartland grain	Operational	Brown	SD	40.589817	-98.33986	45
19	Dakota Ethanol, LLC	Operational	Lake	SD	40.86652	-98.03621	48
20	Plymouth Ethanol, LLC	operating	Plymouth	IA	42.733665	-96.25252	50
21	E Energy Adams, LLC	Operational	Gage	NE	46.28943	-96.06861	50
22	Siouxland Ethanol, LLC	Operational	Dakota	NE	39.738302	-94.84645	50
23	Redfield Energy LLC	Operational	Spink	SD	40.48849	-98.93216	50
24	Bonanza Bioenergy, LLC	Operational	Finney	KS	40.183432	-84.86248	55
25	POET Biorefinng	operating	Carroll	IA	42.507009	-93.27494	53.6
26	POET Biorefinng	Operational	Grant	SD	40.582508	-98.31735	75
27	Kansas Ethanol, LLC	Operational	Rice	KS	41.400898	-89.91134	55
28	Husker Ag, LLC	Operational	Pierce	NE	44.101479	-94.26894	60
29	POET Biorefinng	Operational	Davison	SD	42.060778	-97.38752	60
30	Big River Resources, LLC	operating	Des Moines	IA	40.125808	-102.6809	60
31	POET Biorefinng	operating	Webster	IA	41.526895	-94.37606	60
32	POET Biorefinng	operating	Hamilton	IA	42.74109	-92.61843	60
33	NuGen Energy LLC (bought Verasun)	Operational	Turner	SD	41.163336	-101.0859	110
34	ARKALON ETHANOL, LLC	Operational	Seward	KS	31.086315	-88.01764	110
35	POET Biorefining	operating	Adams	IA	40.968981	-94.79006	60
36	POET Biorefining	operating	Worth	IA	43.149673	-93.20014	45
37	POET Biorefinng	operating	Osceola	IA	42.640598	-92.02703	45
38	POET Biorefining	operating	Palo Alto	IA	45.809329	-119.7492	50

39	POET Biorefinng	Operational	Brown	SD	40.48018	-96.53933	50
40	POET Biorefinng	Operational	Turner	SD	40.782489	-99.7487	100
41	Cargill, Inc.		Washington	NE	44.543013	-94.33901	195
42	Sterling Ethanol, LLC	Operational	Logan	CO	40.63683	-103.1934	40
43	YUMA ETHANOL LLC	Operational	Yuma	CO	40.125808	-102.6809	40
44	Penford Products Co.	operating	Linn	IA	43.43312	-94.95419	45
45	Corn, LP	operating	Wright	IA	43.499654	-92.94909	50
46	Grain Processing Corp.	operating	Muscatine	IA	41.927371	-91.68551	50
47	Lincolnway Energy, LLC	operating	Story	IA	41.400167	-91.0635	50
48	Bridgeport Ethanol, LLC	Operational	Morrill	NE	41.709228	-98.02121	50
49	Amaizing Energy, LLC	operating	Crawford	IA	37.920669	-121.2936	55
50	Green Plains Renewable Energy, Inc.	operating	Page	IA	40.757178	-95.41655	55
51	Green Plains Renewable Energy, Inc.	operating	Dickinson	IA	43.43312	-94.95419	55
52	Little Sioux Corn Processors, LP	operating	Cherokee	IA	40.757178	-95.41655	92
53	ABSOLUTE ENERGY LLC	operating	Mitchell	IA	43.499654	-92.94909	100
54	Global Ethanol, LLC.	operating	Kossuth	IA	41.417751	-95.03887	100
55	Hawkeye Renewables, LLC	operating	Hardin	IA	40.84207	-91.17794	100
56	Homeland Energy Solutions	operating	Chickasaw	IA	43.40013	-94.07927	100
57	Advanced BioEnergy LLC	Operational	Fillmore	NE	41.005675	-97.53989	100

58	Golden Grain Energy, LLC	operating	Cerro Gordo	IA	41.985643	-95.40222	110
59	Valero Renewable Fuels	operating	Floyd	IA	43.102365	-92.74239	110
60	Valero Renewable Fuels	operating	Webster	IA	42.51949	-94.30489	110
61	Valero Renewable Fuels	operating	O'Brien	IA	42.8214	-95.76002	110
62	Valero Renewable Fuels	operating	Buena Vista	IA	42.733665	-96.25252	110
63	Hawkeye Renewables, LLC	operating	Buchanan	IA	41.816016	-90.21664	115
64	Hawkeye Renewables, LLC	operating	Guthrie	IA	41.154653	-92.63203	115
65	Hawkeye Renewables, LLC	operating	Butler	IA	42.73157	-93.91524	115
66	Platinum Ethanol LLC.	operating	Ida	IA	42.33738	-95.40058	
67	Abengoa Bioenergy Corp. (bought High Plains Corporation)		Roosevelt	NM	34.170381	-103.3698	

Table C.6: Biorefineries, and its capacity

ID	ORG_NAME	STATUS	COUNTY	ST	LAT	LON
1	KAAPA Ethanol, LLC	Operational	Kearney	NE	41.42418	-97.2904
2	Advanced BioEnergy LLC/Heartland grain	Operational	Beadle	SD	41.42418	-97.2904
3	Glacial Lakes Energy, LLC	Operational	Codington	SD	40.81549	-98.6076
4	Glacial Lakes Energy, LLC	Operational	Edmunds	SD	40.97449	-101.355
5	North Country Ethanol,LLC	Operational	Roberts	SD	41.52055	-96.0792
6	Nebraska Energy (Aventine)	Operational	Hamilton	NE	41.69898	-73.1923
7	Trenton Agri Products, LLC	Operational	Hitchccock	NE	39.20308	-93.4756

8	Midwest Renewable Energy, LLC	Operational	Lincoln	NE	44.10942	-93.6733
9	East Kansas Agri-Energy, LLC	Operational	Anderson	KS	39.00709	-87.6268
10	Western Plains Agri Energy, LLC	Operational	Logan	KS	40.531	-90.0004
11	Mid America Agri Products/Wheatland	Operational	Perkins	NE	43.67073	-92.0972
12	Pinal Energy, LLC	Operating	Pinal	AZ	33.02142	-111.993
13	Calgren Renewable Fuels, LLC (Pixley Ethanol)	Producing	Tulare	CA	36.00204	-119.304
14	White Energy	Operational	Russell	KS	46.17038	-123.152
15	Chief Ethanol Fuels, Inc.	Operational	Adams	NE	43.81409	-95.2952
16	Cargill, Inc.	operating	Wapello	IA	31.1227	-84.1544
17	Elkhorn Valley Ethanol, LLC (Louis Dreyfus)	Operational	Madison	NE	43.57525	-93.3684
18	Advanced BioEnergy LLC/Heartland grain	Operational	Brown	SD	40.58982	-98.3399
19	Dakota Ethanol, LLC	Operational	Lake	SD	40.86652	-98.0362
20	Plymouth Ethanol, LLC	operating	Plymouth	IA	42.73367	-96.2525
21	E Energy Adams, LLC	Operational	Gage	NE	46.28943	-96.0686
22	Siouxland Ethanol, LLC	Operational	Dakota	NE	39.7383	-94.8465
23	Redfield Energy LLC	Operational	Spink	SD	40.48849	-98.9322
24	Bonanza Bioenergy, LLC	Operational	Finney	KS	40.18343	-84.8625
25	POET Biorefinng	operating	Carroll	IA	42.50701	-93.2749
26	POET Biorefinng	Operational	Grant	SD	40.58251	-98.3174
27	Kansas Ethanol, LLC	Operational	Rice	KS	41.4009	-89.9113
28	Husker Ag, LLC	Operational	Pierce	NE	44.10148	-94.2689
29	POET Biorefinng	Operational	Davison	SD	42.06078	-97.3875
30	Big River Resources, LLC	operating	Des Moines	IA	40.12581	-102.681
31	POET Biorefinng	operating	Webster	IA	41.5269	-94.3761
32	POET Biorefinng	operating	Hamilton	IA	42.74109	-92.6184
33	NuGen Energy LLC (bought Verasun)	Operational	Turner	SD	41.16334	-101.086
34	ARKALON ETHANOL, LLC	Operational	Seward	KS	31.08632	-88.0176
35	POET Biorefining	operating	Adams	IA	40.96898	-94.7901
36	POET Biorefining	operating	Worth	IA	43.14967	-93.2001
37	POET Biorefinng	operating	Osceola	IA	42.6406	-92.027
38	POET Biorefining	operating	Palo Alto	IA	45.80933	-119.749
39	POET Biorefinng	Operational	Brown	SD	40.48018	-96.5393
40	POET Biorefinng	Operational	Turner	SD	40.78249	-99.7487

41	Cargill, Inc.		Washington	NE	44.54301	-94.339
42	Sterling Ethanol, LLC	Operational	Logan	CO	40.63683	-103.193
43	YUMA ETHANOL LLC	Operational	Yuma	CO	40.12581	-102.681
44	Penford Products Co.	operating	Linn	IA	43.43312	-94.9542
45	Corn, LP	operating	Wright	IA	43.49965	-92.9491
46	Grain Processing Corp.	operating	Muscatine	IA	41.92737	-91.6855
47	Lincolnway Energy, LLC	operating	Story	IA	41.40017	-91.0635
48	Bridgeport Ethanol, LLC	Operational	Morrill	NE	41.70923	-98.0212
49	Amaizing Energy, LLC	operating	Crawford	IA	37.92067	-121.294
50	Green Plains Renewable Energy, Inc.	operating	Page	IA	40.75718	-95.4165
51	Green Plains Renewable Energy, Inc.	operating	Dickinson	IA	43.43312	-94.9542
52	Little Sioux Corn Processors, LP	operating	Cherokee	IA	40.75718	-95.4165
53	ABSOLUTE ENERGY LLC	operating	Mitchell	IA	43.49965	-92.9491
54	Global Ethanol, LLC.	operating	Kossuth	IA	41.41775	-95.0389
55	Hawkeye Renewables, LLC	operating	Hardin	IA	40.84207	-91.1779
56	Homeland Energy Solutions	operating	Chickasaw	IA	43.40013	-94.0793
57	Advanced BioEnergy LLC	Operational	Fillmore	NE	41.00568	-97.5399
58	Golden Grain Energy, LLC	operating	Cerro Gordo	IA	41.98564	-95.4022
59	Valero Renewable Fuels	operating	Floyd	IA	43.10237	-92.7424
60	Valero Renewable Fuels	operating	Webster	IA	42.51949	-94.3049
61	Valero Renewable Fuels	operating	O'Brien	IA	42.8214	-95.76
62	Valero Renewable Fuels	operating	Buena Vista	IA	42.73367	-96.2525
63	Hawkeye Renewables, LLC	operating	Buchanan	IA	41.81602	-90.2166
64	Hawkeye Renewables, LLC	operating	Guthrie	IA	41.15465	-92.632
65	Hawkeye Renewables, LLC	operating	Butler	IA	42.73157	-93.9152
66	Platinum Ethanol LLC.	operating	Ida	IA	42.33738	-95.4006
67	Abengoa Bioenergy Corp. (bought High Plains Corporation)		Roosevelt	NM	34.17038	-103.37

Table C.7: Customers and its demand for case study 3

Customer No.	County	State	*FIPS code*	$demand_c$	Longitude(X)	Latitude(Y)
1	Apache	Arizona	04001	5941170	-109.236945	34.721205

2	Cochise	Arizona	04003	10911223	-110.0165	32.5
3	Coconino	Arizona	04005	11166671	-111.15	36.4394
4	Gila	Arizona	04007	4452430	-111.324348	33.995657
5	Graham	Arizona	04009	3091954	-110.115852	32.581441
6	Greenlee	Arizona	04011	700881.6	-109.3384	33.3656
7	La Paz	Arizona	04012	1702070	-114.2081	33.9864
8	Maricopa	Arizona	04013	3.17E+08	-112.82162	33.110062
9	Mohave	Arizona	04015	16629925	-113.480799	35.470604
10	Navajo	Arizona	04017	8926043	-110.0089	34.3966
11	Pima	Arizona	04019	81432770	-111.4753	32.4457
12	Pinal	Arizona	04021	31216104	-111.280402	33.024691
13	Santa Cruz	Arizona	04023	3939292	-110.840949	31.444972
14	Yavapai	Arizona	04025	17531011	-113.173182	34.22643
15	Yuma	Arizona	04027	16261499	-113.612635	32.85202
16	Alameda	California	06001	14095717	-122.090949	37.768455
17	Butte	California	06007	2053312	-121.409796	39.472159
18	Contra Costa	California	06013	9790799	-121.805304	37.820544
19	El Dorado	California	06017	1689857	-120.453986	38.67899
20	Fresno	California	06019	8684110	-120.058478	36.670532
21	Humboldt	California	06023	1256468	-123.771857	40.298296
22	Imperial	California	06025	1628911	-115.323371	32.950069
23	Kern	California	06029	7836475	-118.729133	35.61258
24	Kings	California	06031	1427817	-120.003547	35.886734
25	Lake	California	06033	603533.8	-122.772101	38.952914
26	Los Angeles	California	06037	91639367	-117.806281	34.553981
27	Madera	California	06039	1408059	-119.564093	37.455149
28	Marin	California	06041	2355793	-122.728156	38.18414
29	Mendocino	California	06045	819840.9	-123.519171	39.890686
30	Merced	California	06047	2387377	-121.080207	36.912479
31	Monterey	California	06053	3873825	-121.365851	36.401299
32	Napa	California	06055	1273838	-122.387579	38.610345
33	Nevada	California	06057	921787.8	-121.27796	39.213016
34	Orange	California	06059	28095209	-117.608527	33.557325
35	Placer	California	06061	3251998	-120.574835	38.952914
36	Riverside	California	06065	20436438	-114.949836	33.557325
37	Sacramento	California	06067	13241885	-121.299933	38.455652
38	San Benito	California	06069	515838.7	-121.409796	36.930045
39	San Bernardino	California	06071	18995097	-116.674689	34.282091
40	San Diego	California	06073	28889290	-117.213019	33.171949
41	San Francisco	California	06075	7515449	-122.464484	37.720674
42	San Joaquin	California	06077	6396123	-121.497687	37.785822

43	San Luis Obispo	California	06079	2516586	-120.596808	35.11757
44	San Mateo	California	06081	6705473	-122.310675	37.241173
45	Santa Barbara	California	06083	3956312	-120.486945	34.757318
46	Santa Clara	California	06085	16628487	121.695441	37.101104
47	Santa Cruz	California	06087	2448873	-122.068976	37.066046
48	Shasta	California	06089	1654064	-122.772101	40.511626
49	Solano	California	06095	3857838	-121.827276	38.18414
50	Sonoma	California	06097	4516148	-122.783087	38.816084
51	Stanislaus	California	06099	4801512	-120.552863	37.698945
52	Sutter	California	06101	884202.9	-121.7284	39.072424
53	Tehama	California	06103	592315.2	-121.816291	40.096897
54	Tulare	California	06107	4126961	-118.16883	35.944569
55	Tuolumne	California	06109	516734.7	-119.619025	38.050166
56	Ventura	California	06111	7684222	-119.190558	34.268474
57	Yolo	California	06113	1874571	-121.651496	38.322178
58	Yuba	California	06115	673439.7	-121.563605	39.178959
59	Adams	Colorado	08001	4495763	-104.479866	39.91507
60	Arapahoe	Colorado	08005	5823307	-104.721565	39.654255
61	Boulder	Colorado	08013	2998855	-105.204964	40.109502
62	Broomfield	Colorado	08014	568981	-105.0951	39.932821
63	Denver	Colorado	08031	6109941	-105.051155	39.637336
64	Douglas	Colorado	08035	2906192	-104.809456	39.450054
65	Eagle	Colorado	08037	531394.4	-107.094612	39.383054
66	El Paso	Colorado	08041	6334982	-104.304085	39.059629
67	Garfield	Colorado	08045	574071.2	-107.314338	39.484879
68	Jefferson	Colorado	08059	5441944	-105.248909	39.789458
69	La Plata	Colorado	08067	522608.5	-107.57801	37.228052
70	Larimer	Colorado	08069	3050399	-105.699348	40.952852
71	Mesa	Colorado	08077	1493721	-108.874397	39.025498
72	Pueblo	Colorado	08101	1619349	-104.710579	38.218674
73	Weld	Colorado	08123	2573898	-104.370003	40.611782
74	Black Hawk	Iowa	19013	1691148	-92.246589	42.488219
75	Cerro Gordo	Iowa	19033	569577.2	-93.136482	43.064793
76	Clinton	Iowa	19045	633629	-90.384407	41.851067
77	Dallas	Iowa	19049	853185.4	-93.927498	41.603037
78	Des Moines	Iowa	19057	520219.2	-93.548469	41.564003
79	Dubuque	Iowa	19061	1208186	-91.131447	42.631856
80	Johnson	Iowa	19103	1688465	-91.439094	41.637942
81	Linn	Iowa	19113	2724955	-91.373176	41.93285
82	Marshall	Iowa	19127	524386.2	-92.779426	41.965534
83	Muscatine	Iowa	19139	551438.9	-91.065559	41.522889
84	Polk	Iowa	19153	5555542	93.768196	41.572223

85	Pottawattamie	Iowa	19155	1201800	-95.635871	41.259143
86	Scott	Iowa	19163	2131499	-90.472297	41.76918
87	Story	Iowa	19169	1155151	-93.30677	42.161284
88	Warren	Iowa	19181	596333.2	-93.477058	41.16203
89	Woodbury	Iowa	19193	1318087	-96.229134	42.490245
90	Buffalo	Nebraska	31019	514953.6	-98.838774	40.820176
91	Douglas	Nebraska	31055	5776054	-96.27896	41.292384
92	Hall	Nebraska	31079	654632.8	-98.531157	40.720332
93	Lancaster	Nebraska	31109	3187960	-96.839262	40603658
94	Sarpy	Nebraska	31153	1774223	-95.960356	41.168447
95	Clark	Nevada	32003	18930673	-115.592925	36.047018
96	Lyon	Nevada	32019	504295.6	-119.240386	39.008648
97	Washoe	Nevada	32031	4088374	-119.85562	41.060846
98	Carson City	Nevada	32510	536253.1	-119.701811	39.174922
99	Bernalillo	New Mexico	35001	7239628	-106.223305	34.938356
100	Chaves	New Mexico	35005	717282.2	-104.683501	33.415539
101	Curry	New Mexico	35009	528589.3	-103.344887	34.541125
102	Doña Ana	New Mexico	35013	2286223	-107.144438	32.122274
103	Eddy	New Mexico	35015	588172.5	-104.683501	32.567794
104	Lea	New Mexico	35025	707251.5	-103.628813	32.752778
105	McKinley	New Mexico	35031	781170.6	-108.574379	35.369515
106	Otero	New Mexico	35035	697089.7	-105.430571	32.604821
107	Sandoval	New Mexico	35043	1437526	-106.50895	35.90524
108	San Juan	New Mexico	35045	1420950	-108.243071	36.688383
109	Santa Fe	New Mexico	35049	1575300	-106.045805	35.837995
110	Valencia	New Mexico	35061	836645.3	-106.860512	34.83021
111	Lincoln	South Dakota	46083	583626.2	-96.797035	43.115573
112	Minnehaha	South Dakota	46099	2206344	-96.75309	43.690271
113	Pennington	South Dakota	46103	1314266	-102.949379	44.007191

Vita

Ileana Delgado was born in Cd. Juarez, Chihuahua, Mexico. She received her B.S. in Industrial Engineering in May 2014, and after graduation she started a Master of Science in Manufacturing Engineering at The University of Texas at El Paso, USA. Since January 2013 to the date, she has been part of the research team at the Sustainability Engineering and Systems Optimization Lab (SESOL). During the summer of 2013, she was selected to participate in a funded international experience; the study abroad program in Peru titled "Global and Regional Sustainability Engineering." She has been an intern of EPA's Air Quality Program and Coca-Cola Refreshments. National winner of the USHCC and Liberty Power Bright Horizons Scholarship and recipient of a USDA's BGREEN grant. Certified by the Institute of Industrial Engineers on Green Belt Six Sigma. Her research interests include optimization, energy, sustainability, biofuels, supply chain, transportation, scheduling, and quality control. She had the opportunity to present her research on Multiple Objective Optimization of a Biomass to Bio-Refinery Logistics system at the 2015 Industrial and Systems Engineering Research Conference held in Nashville, Tennessee, USA, in addition to presenting Global Warming Potential and Cost Minimization for the Centralized Carrier Collaboration and Multi-Hub Location Problem at the 2015 International Conference on Operations Excellence and Service Engineering held in Orlando, Florida, USA which was recognized as the best track paper. Moreover, she gave a poster presentation on The Highly Complex Biofuel Supply Chain: An Optimization Algorithm to Solve a Biomass to Biorefinery Logistic System Considering Multiple Types of Feedstock at the Agri-Science Conference, held at Miami, Florida, USA. She is a member of IIE, Alpha Pi Mu (Honors Society for Industrial Engineers), and Alpha Chi (National Honors Society).

Contact Information: ileanadlgd7@gmail.com

This thesis/dissertation was typed by Ileana Delgado Camacho