Network Optimization for Middle Mile Delivery for an Oilfield Service Company

by

Brian Hinkamp
Bachelor of Science in Supply Chain and Operations Management, Miami University, 2015

and

Marwan Ismael
Bachelor of Science in Civil Engineering, University of Baghdad - College of Engineering, 2011

SUBMITTED TO THE PROGRAM IN SUPPLY CHAIN MANAGEMENT
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF APPLIED SCIENCE IN SUPPLY CHAIN MANAGEMENT
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2023

Signature of Author: ______________________________
Department of Supply Chain Management
May 12, 2023

Signature of Author: ______________________________
Department of Supply Chain Management
May 12, 2023

Certified by: ______________________________
Dr. Elenna R. Dugundji
Research Scientist, Supply Chain Management Program
Capstone Advisor

Accepted by: ______________________________
Prof. Yossi Sheffi
Director, Center for Transportation and Logistics
Elisha Gray II Professor of Engineering Systems
Professor, Civil and Environmental Engineering

Network Optimization for Middle Mile Delivery for an Oilfield Service Company

by

Brian Hinkamp

and

Marwan Ismael

Submitted to the Program in Supply Chain Management
on May 12, 2023 in Partial Fulfillment of the
Requirements for the Degree of Master of Applied Science in Supply Chain Management

ABSTRACT

The global transportation sector holds the top position as the primary source of greenhouse gas emissions, with road transportation, especially heavy-duty vehicles, being the primary source. Greenhouse gases significantly impact global warming by trapping heat in the Earth's atmosphere, causing rising temperatures. This phenomenon, the greenhouse effect, results in various climate change repercussions. In response to growing climate change concerns, international, national, and industrial communities have taken action. Companies globally are pursuing initiatives to achieve net-zero emissions, aiming for a climate-neutral world by mid-century. One approach to combat climate change involves optimizing a company's supply chain design. An efficient supply chain network design can lower transportation costs, reduce carbon emissions, and improve a firm's overall performance. This study explores the potential impact of introducing a middle-mile fulfillment center on transportation costs and greenhouse gas emissions within an oil field service company's supply chain network. By evaluating two candidate locations proposed by the project sponsor, the research assesses variable transportation costs and total carbon emissions generated during goods transportation to meet customer demand. This paper introduces a mixed-integer linear programming formulation as a solution to the single-period, multi-echelon supply chain network design problem. The model aims to minimize total transportation costs and greenhouse gas emissions, primarily from mobile sources within the supply chain. The research team found that neither of the proposed locations provides benefits in terms of transportation costs or carbon emissions. This study highlights the importance of integrating environmental considerations into strategic supply chain network design decision-making.

Capstone Advisor: Dr. Elenna Dugundji
Title: Research Scientist

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Motivation of the Study

Innovations in horizontal drilling practices and groundbreaking hydraulic fracturing technologies, commonly referred to as the "U.S. Shale Revolution," facilitated a notable surge in oil and gas production in the mid-2000s in the United States (The Strauss Center, The University of Texas at Austin). Tight shale formations, previously considered impermeable, accounted for approximately 66% of total U.S. crude oil production in 2021 (U.S. Energy Information Administration [EIA], 2022). This surge in oil production over the past fifteen years has significantly decreased the United States' reliance on imported oil, serving as a crucial contributor to the country's economy. The United States' crude oil production doubled, from 5.8MM bbl./day in 2000 to 11.6MM bbl./day at the end of 2021 (U.S. Field Production of Crude Oil, 2022). Despite the upward trajectory of growth in oil production, navigating volatile market conditions, and intense competition, while managing a capital-intensive cost structure, has proven an arduous pursuit. Transportation and logistics managers in the oil and gas industry must grapple with the increasing difficulty of consistently reducing costs, managing supply chain disruptions, and keeping up with the growing need for sustainable logistics operations to reduce collective carbon emissions.

In today's customer-centric, digitized, and global economy, significant challenges emerge across global supply chains. Climate change has increasingly become a significant concern for the global, national, and industrial communities. Corporations worldwide are announcing their initiatives to help reach net zero emissions to achieve a climate-neutral world by mid-century. Within those organizations, a focus on the biggest contributors to greenhouse gas emissions has become a focal point for reduction. Transportation is recognized as a significant contributor to greenhouse gas emissions, both domestically in the United States and globally. In 2020 the transportation sector accounted for 28% of greenhouse gas emissions globally, the highest of any sector (US EPA, 2022).

Burning fossil fuels for transportation purposes produces greenhouse gases (GHG), e.g., carbon dioxide, nitrous oxide, and methane, which are significant contributors to the rise in global temperatures leading to global warming (Climate.gov, 2022). As a result, investors, stakeholders, and customers are increasingly looking for enterprises that pursue profit and have strategies and objectives that focus on the firm's longevity and contribution to the planet's sustainability.

One approach to combat climate change is to consider a company's supply chain network design (SCND) to establish a distribution network that can help reduce greenhouse gas emissions, reduce transportation costs, and meet customer demand on time. The process of supply chain network design involves identifying the most efficient structure of a company's supply chain to effectively meet its strategic objectives. SCND involves analyzing the various components of the supply chain, including suppliers, transportation modes, warehouses, and distribution channels, to determine the most efficient and effective way to move products from the point of supply to the point of consumption.

One goal of SCND is minimizing costs while maintaining or improving service levels. This goal is achieved by identifying the most cost-effective transportation routes, optimizing inventory levels, and identifying the optimal location for warehouses and distribution centers. The SCND process typically involves the use of sophisticated analytical tools and software to model different scenarios and evaluate the impact of numerous factors on the supply chain. These factors include transportation costs, inventory holding costs, lead times, and demand variability. Overall, supply chain network design plays a critical role in helping companies optimize their supply chain operations, reduce costs, and improve customer service. Oilfield service companies are no exception to the growing number of companies pursuing supply chain network design to plan strategically while identifying the most cost-effective way to deliver its products and services.

The chemicals and tools required for shale exploration, drilling, and extraction of crude oil are often an overlooked aspect in the value chain required to fuel many aspects of our daily lives. Oil field service companies support major oil and gas ventures, often in a just-in-time manner, while trying to maintain an elastic cost structure and reduce the overall impact on the environment.

Meeting customers demand for tools and chemicals in distant and remote locations often requires quick transportation with little to no lead time, resulting in suboptimal supply chain planning and poor transportation network design. This demand variability can be attributed to the uncertainties encountered during the oil well drilling process, which necessitates unique chemical formulations and specialized tools to safely drill the wellbore, remove cuttings, and ensure well integrity.

Providing high levels of service under the aforementioned circumstances can lead to excessive costs, which can make it difficult for companies to maintain low operating costs and still meet non-negotiable service levels. In the oil field services market, single-digit profit margins force companies to tighten their budgets and strategically assess their operations. However, by considering the supply chain network design, companies can reduce costs, minimize their environmental impact, and eventually emerge as profitable industry leaders.

In this study, we assess the impact of introducing two candidate locations, selected by the project sponsor to serve as middle-mile distribution center, on the total transportation cost and the resulting carbon emissions. This study optimizes the multi-criteria objective of minimizing variable transportation costs and reducing carbon emissions. One key benefit of multi-criteria optimization is that it allows companies and their logistics managers to explore different scenarios and evaluate their outcomes under different assumptions. By comparing the costs and benefits of different scenarios, supply chain executives can identify the most efficient and effective solutions that achieve multiple objectives simultaneously. Multi-criteria optimization will help the project sponsor strike a balance between economic and environmental goals, ensuring that the transportation system remains sustainable and resilient in the long run. The result is a more holistic and sustainable approach to transportation planning that considers the economic and environmental impacts of different options.

1.2. Problem Statement and Research Question

Drastic drops in oil prices, widespread supply chain disruptions brought on by the COVID-19 pandemic, and imbalanced supply and demand in the oil market led to the recent oil and gas industry crash in 2020 (Investopedia, 2022). Oil field service providers were impacted by a negative ripple effect; the decline in oil prices resulted in reduced capital spending budgets, consolidation of facilities, and closures of distribution centers. The recent reconsolidation efforts have dramatically reshaped the sponsor company's entire supply chain network and subsequent distribution footprint in the United States. For the business unit (BU) in scope for this project, the existing supply chain network consists of approximately 209 supplier locations and twenty-four regional warehouses. In the last fiscal year alone, the company served hundreds of different customer locations, with over 18,000 shipments executed

The BU wants to reevaluate its supply chain network design in the wake of the most recent footprint consolidation. One option is considering a middle-mile distribution network. In the middle mile distribution network, products move from a suppliers' locations to fulfillment centers or regional distribution centers. Due to higher flexibility inherent to its design, the middle mile network is known to provide cost-reduction opportunities, which last mile delivery cannot offer. The project hypothesis is that establishing strategically located distribution centers (candidate locations provided by the BU) would generate cost efficiencies through a reduction in overall middle-mile traveling distance which can lead to lower transportation costs and lower carbon emissions. This research established a baseline, built an optimization model, and compared the proposed network's costs and emissions to the existing network's costs and emissions. Our research identified the optimal transportation network that will minimize the sponsoring company's transportation cost and decrease carbon emissions due to over-the-road freight transportation while maintaining existing customer service levels.

The "middle mile," which refers to the first tier of transportation of goods from suppliers to the regional distribution centers (RDCs), is a critical aspect of any company's distribution network. To optimize this middle mile network, our study employed mixed integer linear programming (MILP) models to assess the financial and environmental impact of introducing two candidate locations as middle mile distribution centers for a specific Business Unit within the sponsoring company. We ran multiple iterations of scenarios using objective functions to create hypothetical network structures and analyze cost and network movements. Each scenario involved tradeoffs between different variables and their corresponding sensitivities to the model, allowing for a comprehensive evaluation of potential outcomes.

The sponsoring company runs a many-to-many distribution network (see Figure 1). The distribution network consists of the company's chemicals suppliers, who ship their products to the company's regional warehouses. The regional warehouses store the goods until those products are either blended at the regional warehouse or shipped to the customer location. To determine the most efficient network configuration, the study employs an optimization approach and analyzes the outputs of three different scenarios. The first scenario represents the current state of the network, as shown in Figure 1. The second and third scenarios introduce the addition of a middle mile distribution center at two different candidate locations, as depicted in Figure 2.

Figure 1

Middle-Mile Distribution Network for the Sponsor Company

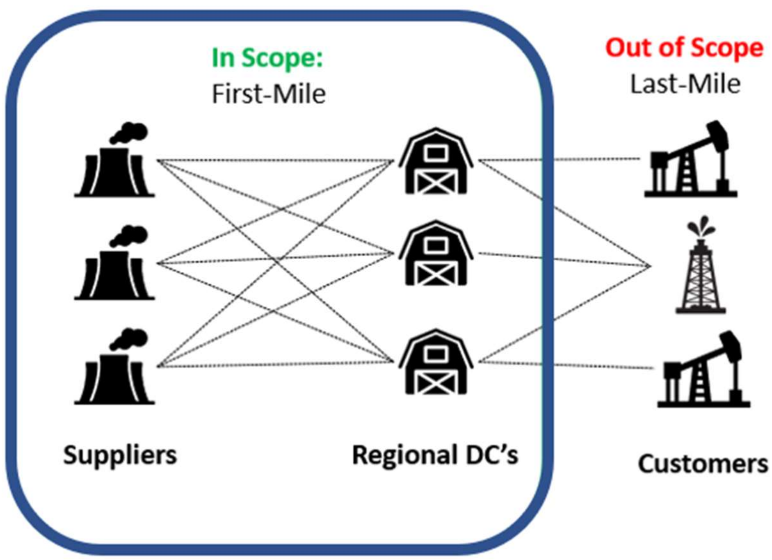
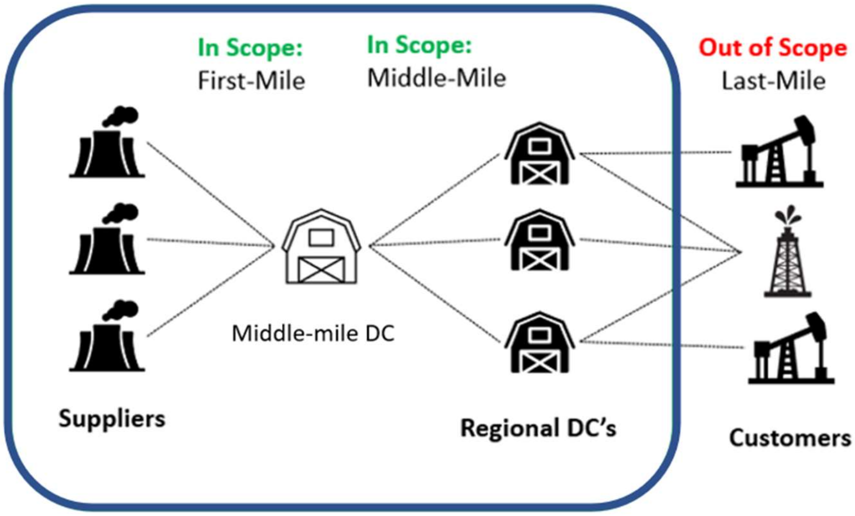


Figure 2

Potential Middle-Mile Distribution Network for the Sponsor Company



The network optimization analysis presented in this study aimed to answer several important questions:

1. This study evaluates the economic and environmental impact of introducing a middle mile distribution center.
2. What would be the most efficient and cost-effective arrangement for the nodes within this supply chain network that would lead to carbon efficiency improvements?
3. What are the most efficient transport routes that could minimize the total distance traveled from suppliers to middle-mile distribution centers and regional warehouses, and from these selected sites to the customers, thereby meeting demand more effectively?
4. Which transportation methods are suitable for each route to reduce both expenses and emissions?
5. Which scenario is the most optimal?

1.3 Project Goals and Expected Outcomes

This project aimed to improve the efficiency of the inbound transportation network of the sponsor business unit by introducing a new middle-mile fulfillment center. The study evaluated two potential candidate locations proposed by the project sponsor by analyzing the variable transportation costs and the total carbon emissions produced during the transportation of goods to meet customer demand. The research deliberately excluded fixed costs related to facility operations and inventory, as well as capital expenditures related to workforce, plant setup, or decommissioning. Notably, the scope of the research included only CO_2 emissions from mobile sources and did not consider other sources such as facilities.

Although we did not consider inventory in our study, supply chain managers within the business unit expressed interest in exploring the feasibility of establishing a vendor-managed inventory agreement (VMI) for the BU. VMI is a supply chain management strategy where the supplier assumes the responsibility of monitoring and replenishing inventory levels for the customer. Under this approach, the supplier keeps the inventory in stock on their books until it is consumed by the customer. In contrast to the traditional model where customers place orders and manage their own inventory levels, the VMI approach puts the supplier in charge of these tasks. This ensures that the customer always has optimal inventory levels without the need for constant monitoring and adjustments. VMI, if done well, can lead to improved inventory management, increased efficiency which can lead to significant cost savings.

The primary objective of our study is to assess the potential impact of introducing a middle mile fulfillment center on transportation costs and greenhouse gas emissions. Our findings will be crucial in determining the feasibility of introducing a middle mile distribution center to the network as part of the VMI feasibility project. We aim to quantify any potential impact, enabling the project sponsor to evaluate the tradeoff between the cost savings of the VMI model and the additional expenses associated with the introduction of the middle mile fulfillment center, if any. The hypothetical middle mile distribution center will function as the inventory consolidation location for the VMI service provider's regional distribution centers.

To develop our methodology and achieve the project's objectives, we drew upon literature across various areas, including traditional network design, carbon emissions in network design, facility location problem, methodologies for measuring carbon emissions in logistics networks, and the business impact of carbon-efficient network design. In the upcoming chapter, we present an in-depth literature review, providing a comprehensive understanding of the current state-of-the-art in these areas.

2. STATE OF THE ART

Our study aimed to investigate whether the implementation of a middle-mile fulfillment center in the sponsor company's network could effectively reduce transportation costs and the associated carbon emissions. To achieve this goal, we developed a model that incorporates carbon emission reduction objectives into the traditional Supply Chain Network Design (SCND) problem.

The State-of-the-Art chapter is organized into three main sections. Firstly, we examined the topic of network design, including research related to traditional SCND models and the Facility Location Problem (FLP), followed by a brief overview of literature on carbon emissions in network design models. Secondly, we delved into the measurement of carbon emissions in supply chains, with a focus on the Global Logistics Emissions Council (GLEC) framework. Lastly, we discussed the potential business impact of carbon-efficient network design in the third section.

2.1 Network Design

2.1.1 Traditional Network Design Models

Traditionally, supply chain network design (SCND) refers to the process of designing an optimal supply chain network that can deliver goods and services from upstream suppliers to downstream customers in an effective and efficient manner (Alamsyah & Purevdorj, 2021; Rezaee et al., 2017; Watson et al., 2013). The design of a supply chain network is a critical factor that influences decision-making for supply chain planning across all levels - operational, tactical, and strategic. (Alamsyah & Purevdorj, 2021; Peng et al., 2016; Rezaee et al., 2017). SCND involves identifying the best combination of suppliers, manufacturing facilities, distribution centers, warehouses, transportation modes, and delivery routes to meet customer demands while meeting other objectives like minimizing logistical costs and maximizing profits and service levels (Rezaee et al., 2017; Watson et al., 2013). The design of a supply chain network is crucial in determining the overall performance of a company's supply chain. An effective and resilient supply chain network can result in reduced logistical costs and improved service levels, which in turn can lead to increased revenues and profits. However, achieving a positive financial outcome is not a simple feat, as it requires careful consideration of various factors such as market demand, competition, and operational costs. Thus, a well-designed supply chain network is essential in achieving the desired financial outcome.

SCND encompasses a broad spectrum of problems, from basic to complex, and has been at the center of researchers' attention for many years (Peng et al., 2016). One key aspect of SCND is the Facility Location Problem (FLP). FLP is a strategic process that entails identifying the most advantageous locations for various types of facilities, including but not limited to manufacturing plants, mixing centers, cross-docking facilities, distribution centers, and retail outlets. The objective of the FLP is to determine the optimal locations that would provide significant advantages in terms of cost savings, revenue growth, and customer satisfaction. The goal is to minimize logistical costs or maximize profits and service levels. This problem is commonly referred to as the location-allocation problem. There are several types of FLPs in supply chain network design (Melo et al., 2009), including single facility location problem (e.g., Moradi & Bidkhori, 2009), multiple facility location problem (e.g., Tamir, 2001), capacitated facility location problem (e.g., Wu et al., 2006), non-capacitated facility location problem (e.g., Chudak & Shmoys, 2003), fixed charge facility location problem (e.g., Nozick, 2001), stochastic facility location problem (e.g., Turkeš et al., 2021) and lastly, green facility location problem (e.g., Martínez & Fransoo, 2017). How to construct these networks varies by type of research and goal. Our research problem is a non-capacitated facility location problem that will focus on fixed and variable transportation costs and carbon emissions from variable sources.

Researchers have taken multiple approaches to solving FLPs in SCND. The first approach uses mathematical programming techniques using mixed-integer linear programming and non-linear programming. Drezner and Hamacher (2004) comprehensively introduced mathematical programming approaches for facility location problems. The second approach uses heuristics and metaheuristics methods. Heuristic methods involve the development of approximation algorithms that provide near-optimal solutions in a reasonable amount of time. In contrast, metaheuristic methods are optimization algorithms that use an iterative process to search for optimal solutions by exploring the solution space. Farahani and Hekmatfar (2009) provide an overview of heuristic and metaheuristic approaches to solving facility location problems. Other less popular approaches to solving FLPs use simulation-based methods and other hybrid methods. One unique feature of any supply chain network is the presence of different facility types within the network. These facilities can be categorized based on their role in that supply chain (e.g., suppliers, manufacturing plants, distribution centers, or retail outlets). These categories are called echelons (i.e., layers). Facility location problems in SCND can be classified based on the number of layers (echelons): single-echelon, two-echelon, and multi-echelon SCND problems (Melo et al., 2009). The choice of the method depends on the specific problem requirements, the size and complexity of the problem, and the computational resources available. In our study, we will utilize MILP models to solve the optimal location for a middle mile distribution center for a multi-echelon supply chain network.

Aside from the number of echelons in the network, facility location problems (FLPs) can be classified further based on several factors, including demand and cost uncertainty, number of products, planning periods, product flow, objective functions, and solution space. FLP models can be classified as deterministic or stochastic in terms of demand and cost uncertainty. The planning period is also a factor in classifying FLPs, where problems with a single planning period are often easier to solve than those with multiple periods. FLP models can also include reverse logistics or not, depending on the nature of the business. The objective of the FLP model can vary, such as minimizing costs or maximizing service levels. Finally, the solution space can be discrete or continuous, depending on the number and nature of potential facility locations. Understanding these factors is crucial in selecting the appropriate approach to solve the problem (Alamsyah & Purevdorj, 2021; Peng et al., 2016; Rezaee et al., 2017). In our research, we focus on a single planning period of six months, consider deterministic demand, solving 169 different SKUs in the network, and only consider forward flow (excluding reverse logistics flow). Our FLP model has a dual objective function with a continuous space for solutions.

A multi-echelon SCND problem involves optimizing the configuration of facilities at various levels of a supply chain network (echelons), aiming to minimize total cost, or maximizing profits while meeting customer demand at the desired service levels. Solving a multi-echelon SCND problem includes determining the location, size, and number of facilities to open, inventory policies, and transportation routes while considering uncertainties in demand, supply, and transportation costs (Melo et al., 2009).

Several studies have addressed multi-echelon supply chain network design (SCND) problems with different objectives and constraints. For instance, Golpîra et al. (2017) tackled a multi-objective, multi-echelon SCND problem that incorporated demand, environmental uncertainties, and downstream risk attitude. To manage this uncertainty, a robust counterpart of the problem was formulated, and the sum constraint method was applied to transform the multi-objective MILP into a single-objective one. In another study, Tsiakis et al. (2001) proposed a MILP model for designing multi-product, multi-echelon supply chain network, considering fixed manufacturing plants, customer zones, and unknown locations for warehouses and distribution centers. Their objective was to minimize the total cost of the network, including infrastructure and operating costs. Watson et al. (2013) introduced a model for a multi-echelon supply chain that aimed to identify optimal warehouse locations while factoring in fixed plant and customer locations to minimize overall logistical costs. A thorough review of the literature was instrumental in developing the methodology used to tackle our research problem.

2.1.2 Carbon Emissions in Supply Chain Network Design Models

2.1.2.1 Accounting for Carbon Emissions in SCND

Sustainability in supply chain management has three main pillars: economic, social, and environmental (Labuschagne et al., 2005). Traditionally, the economic pillar has been the primary focus in supply chain network design (SCND), with cost minimization or profit maximization as the primary objective. However, contemporary supply chains must be optimized beyond their economics (Peng et al., 2016). SCND has emerged as a crucial factor in lowering the overall carbon footprint of supply chains. Consequently, research integrating carbon emissions into the SCND problem has gained momentum recently (Alamsyah & Purevdorj, 2021; Peng et al., 2016).

The supply chain network's (SCND) design plays a critical role in determining the efficiency of a supply chain, serving as a significant driver in achieving optimal performance. Companies worldwide are under mounting pressures from consumers, shareholders, and investors, to build more sustainable supply chains (Ageron et al., 2012). Therefore, the SCND is a critical starting point when optimizing any supply chain (Varsei & Polyakovskiy, 2017). Transportation is considered a significant source of Greenhouse gas emissions for any business. SCND plays a vital role in mitigating the effect of transportation emissions on the environment because the network's design significantly impacts supply chain transportation performance from cost and emissions standpoints (Martínez & Fransoo, 2017).

When solving an SCND problem, researchers have employed three primary approaches to integrate carbon emissions into SCND modeling and decision-making (J. Wang et al., 2020). The first approach revolves around reducing the overall carbon emissions in the supply chain, making it the main priority. To achieve this, Yang et al. (2016), Bouzembrak et al. (2011), and F. Wang et al. (2011) developed multi-objective MILP models that factor in both cost and emissions. These models encompass the emissions generated by transportation and facilities in their respective supply chains during the production and transportation of goods (Alamsyah & Purevdorj, 2021; J. Wang et al., 2020).

The second approach treats carbon emissions in the supply chain as a carbon cost to be incorporated into the economic objectives. Carbon costs are a function of supply chain activities associated with manufacturing, storing, and transporting products (Li et al., 2017). Rezaee et al. (2017) and Jiang et al. (2019) integrated carbon costs into the primary objectives of their multi-echelon SCND problems (Alamsyah & Purevdorj, 2021; J. Wang et al., 2020).

The third approach is to treat the carbon emissions as a constraint in the SCND model; these constraints can be introduced as a carbon emissions cap, carbon emissions tax, or emission permission trade among different parties in the supply chain. Zhou and Wen (2020), and Benjaafar et al. (2013) are examples of research in this field that explain the influence of these constraints on the SCND model decisions, which can lead to lower carbon emissions for the entire supply chain. (Alamsyah & Purevdorj, 2021; J. Wang et al., 2020).

We utilize the second approach in our research. In this study, we develop a mixed-integer linear programming (MILP) model that incorporates carbon emissions in the supply chain as a carbon cost within the economic objectives. Our model emphasizes the sequential minimization of transportation and carbon costs as primary objectives. A comprehensive analysis of the existing literature has been instrumental in expanding our knowledge of the state of the art, subsequently informing, and refining our methodological approach to address the research problem and meet the project objective.

Research that incorporates carbon emissions into the SCND problem is widely available (Peng et al., 2016). The existing literature on the topic includes a wide range of operational decisions that have been considered in the total emissions and cost objectives (Alamsyah & Purevdorj, 2021). For example, Peng et al. (2016) proposed a mixed integer linear programming (MILP) formulation as a solution to model and solve a multiperiod one-stage supply chain network design (SCND) problem. Their model solved two main objectives: minimizing the total logistics cost (transportation and handling costs) and minimizing the carbon emissions generated while transporting and storing the goods. While F. Wang et al. (2011) focused on capturing the trade-offs between the total cost and the implied carbon emissions in their model. In their study, F. Wang et al. (2011) created a multi-objective MILP model that considered what they referred to as "environmental investment decisions in the supply network design phase." Their model considered transportation, facility setup, warehousing, and environmental protection investment costs. For their total emissions objective, they took into account the emissions generated by both the facilities and the total distance traveled between nodes in their network.

Their findings suggest that expanding the capacity of the supply chain network and boosting the supply to facilities can lead to a reduction in both carbon emissions and the overall cost of the entire network. Another relevant study that considered carbon emissions while solving a traditional SCND problem is Jiang et al. (2019). In their research, Jiang et al. (2019) developed a mixed integer linear programming (MILP) model to find the optimal partners, select the optimal technology, and choose the best mode of transport for their goods. Their model considered the expenses and emissions associated with operational activities such as manufacturing, procurement, product distribution, and reverse logistics. While there were no exact matches between their work and our network, a comprehensive analysis of existing literature proved invaluable in developing the methodology employed to address our specific research question.

2.1.2.2 The Green Facility Location Problem (GFLP)

In addition to the traditional SCND and the integration of carbon emissions within it, another important strand of research relevant to our project is the Green Facility Location problem (GFLP), an extension of the traditional facility location problem (FLP) that considers carbon emissions reduction as the model's primary objective. The green facility location problem is a critical area of research that aims to optimize the location of facilities in the SCND problem while minimizing environmental impact. The goal is to design a facility location that reduces adverse environmental effects associated with transportation, waste management, and energy consumption. Dozens of studies have focused on developing models and algorithms for the green facility location problem, considering a range of environmental and economic factors, such as transportation costs, emissions, and energy consumption (Martínez & Fransoo, 2017).

Martínez and Fransoo's (2017) research sought to address the green facility location problem (GFLP), a crucial issue in supply chain sustainability. They introduced a bi-objective formulation of the GFLP that considers the costs and emissions associated with facility location decisions. Their distinctive approach focused solely on emissions and costs from mobile sources, such as trucks, while excluding those from stationary sources, like distribution centers and warehouses. This method enabled them to concentrate on the trade-offs between distance and utilization when making location-allocation decisions, ultimately developing a model that optimizes facility locations while minimizing transportation emissions and costs.

Martínez and Fransoo (2017) applied their model to a case study involving a logistics service provider in the Netherlands, assessing their approach's efficacy. They discovered that optimizing facility locations could decrease transportation emissions and costs without compromising service quality. Their study offers valuable insights into incorporating environmental sustainability into facility location decisions. Their mathematical model can be utilized to evaluate the environmental impact of various location-allocation decisions within a supply chain context and encourage more sustainable supply chain practices (Martínez & Fransoo, 2017). Although our problem was not identical to theirs, Martínez and Fransoo's (2017) work significantly influenced our methodology and approach to addressing our research problem.

2.2 Measuring Carbon Emissions in Supply Chains

Sustainable Supply Chain Management (SSCM) is a term used to describe the different actions companies undertake to ensure their supply chain operations are environmentally friendly, socially responsible, and economically viable. SSCM aims to minimize the environmental impact of an entity throughout its value chain by reducing greenhouse gas emissions, using natural resources in a sustainable manner, and reducing waste. Aside from the environmental impact, SSCM also refers to the promotion of corporate social responsibility (CSR) throughout the value chain, such as respecting human rights, fostering fair labor practices, and maintaining a safe working environment for employees (Seuring et al., 2008).

In terms of performance, SSCM refers to the supply chain's ability to function efficiently and sustainably in alignment with a company's economic, environmental, and social objectives. Achieving sustainability in supply chain management entails integrating economic, environmental, and social considerations into supply chain network design to achieve a coordinated and efficient flow of materials, services, and capital while meeting the economic objectives of the business (Rajeev et al., 2017). By prioritizing supply chain suitability, companies can drive operational excellence and enhance their overall sustainability performance while building a reputation for ethical and socially responsible business practices.

According to the Carbon Disclosure Project (CDP), measuring emissions in a company's supply chain involves quantifying the greenhouse gas emissions (primarily carbon dioxide or CO_2) resulting from activities related to the production, transportation, and storage of goods and the delivery of services offered by the company. This process is critical for assessing the environmental impact of a company's operations and identifying opportunities for improvement. By accurately measuring the carbon emissions in a company's supply chain, businesses can develop effective strategies to reduce their carbon footprint, minimize their environmental impact, and enhance their overall sustainability performance.

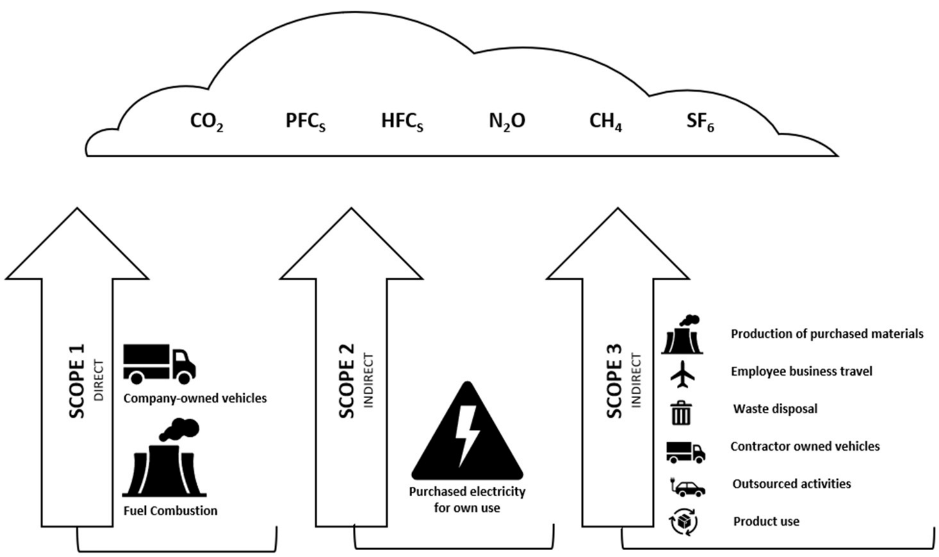
Greenhouse gases have varying environmental impacts, making it difficult to compare their overall effects. However, a unit of measurement known as CO_2e (or CO_2. equivalent) has been developed by researchers to compare the potential global warming impact of different greenhouse gases to the equivalent amount of CO_2. This unit of measurement allows for a standardized way of comparing and quantifying the environmental impact of different greenhouse gases, each of which has a unique potency and atmospheric lifetime. By converting all greenhouse gas emissions into CO_2e, it is possible to combine emissions from various sources into a single number, making it simpler to comprehend the overall environmental impact of a specific activity or industry (Alamsyah & Purevdorj, 2021; Boukherroub et al., 2017).

There are various methods for incorporating carbon emissions into business practices. The Greenhouse Gas (GHG) Protocol is particularly relevant in today's business environment. As an internationally recognized emissions accounting tool, the GHG Protocol establishes standards and guidelines for quantifying and managing greenhouse gas emissions. The Protocol was created through a collaboration between the World Resources Institute (WRI) - a global research organization focused on promoting sustainable development and addressing environmental challenges - and the World Business Council for Sustainable Development (WBCSD) - a CEO-driven global entity comprising over two hundred leading companies dedicated to fostering a sustainable future (WRI & WBCSD The Greenhouse Gas Protocol, 2011).

The assessment of a corporation's environmental impact requires consideration of three distinct scopes of emissions, as outlined by the GHG Protocol (WRI & WBCSD The Greenhouse Gas Protocol, 2011), as shown in Figure 3. The first scope, referred to as Scope 1, encompasses emissions that arise directly from the assets possessed or managed by the entity. These direct emissions result from the activities of the reporting entity, including the combustion of fossil fuels on-site or from vehicles owned by the company. On the other hand, Scope 2 emissions are indirect emissions generated during the process of producing the purchased energy (electricity, heat, or steam) consumed by the company. Lastly, Scope 3 emissions include all other indirect emissions in the company's value chain, including emissions from the production of purchased goods and services, transportation of goods, and the transportation of employees. By considering these three scopes of emissions, corporations can gain a comprehensive understanding of their environmental impact.

Figure 3

Overview of scopes and emissions across a value chain



Note. This figure provides an overview of scopes and emissions across a value chain. It was adapted from the GHG protocol.

Source: (WRI & WBCSD The Greenhouse Gas Protocol, 2011).

The contribution of each scope of emissions to a company's overall greenhouse gas (GHG) footprint can differ widely depending on numerous factors, such as the nature of the company's operations and the industry the company operates in. Direct emissions from sources controlled by the company (Scope 1) are the smallest portion of a company's total emissions, usually less than 20%. Indirect emissions from purchased energy (electricity, heat, or steam), known as (Scope 2) can account for a more sizable portion of a company's emissions, ranging from 20% to 70% or more, depending on the energy intensity of the company's operations and the carbon intensity of its purchased energy. The largest share of a company's emissions typically comes from all other indirect emissions in its value chain (Scope 3), ranging from 50% to 80% or more, depending on the company's industry and the design of its supply chain (The Carbon Trust, 2013, WRI & WBCSD The Greenhouse Gas Protocol, 2011). According to a survey conducted by CDP in 2019, the proportion of Scope 3 emissions originating from transportation can fluctuate based on the company and industry. Nonetheless, the survey determined that transportation was the most significant source of Scope 3 emissions for participating companies, accounting for 16.4% of the total emissions (CDP, 2019).

The carbon trust reported that Scope 3 emissions have the potential to make up as much as 90% of a company's carbon footprint and are often the most significant contributor to its overall carbon impact (Martínez & Fransoo, 2017). When addressing the challenge of finding the optimal location for distribution centers (i.e., solving the green facility location problem), Cholette and Venkat (2009) found that carbon emissions resulting from transportation activities can be up to ten times greater than those generated by stationary sources such as distribution centers and warehouses (Martínez & Fransoo, 2017). Therefore, our research will focus exclusively on Scope 3 emissions generated by transportation activities (mobile sources) and will not consider emissions originating from stationary sources such as supplier locations, distribution centers, and regional warehouses.

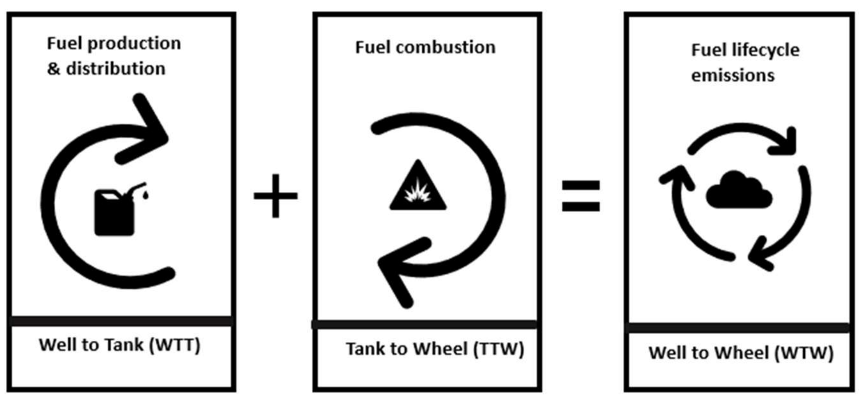
Incorporating transportation-specific carbon emissions into a supply chain network design (SCND) model is a critical step in enhancing the sustainability of logistics activities. Alamsyah and Purevdorj (2021) addressed this issue by employing the Global Logistics Emissions Council (GLEC) Framework. This framework comprises a set of guidelines and standards designed to assist companies in quantifying and reporting carbon emissions from logistics activities in a consistent manner. Developed by leading global organizations advocating sustainable logistics practices, including the World Business Council for Sustainable Development (WBCSD), the Smart Freight Centre, and the International Energy Agency (IEA), the GLEC Framework is widely adopted by logistics companies (carriers), shippers, and governmental agencies. This framework is a valuable tool for evaluating environmental performance and pinpointing areas for improvement. Like Alamsyah & Purevdorj (2021), we will be utilizing the GLEC framework in our transportation-specific carbon emissions accounting. The utilization of the GLEC Framework by Alamsyah and Purevdorj (2021) serves as a crucial source of inspiration in developing an approach to solving our research problem.

The GLEC Framework offers a comprehensive methodology for measuring and reporting carbon emissions, which earned it a widespread recognition and adoption by environmentally minded companies and organizations around the world (Smart Freight Centre, GLEC, 2021). Recognizing the significance of emissions associated with fuel production and distribution, the GLEC Framework incorporates these as the Well-to-Tank (WTT) emissions category in its analysis of a company's carbon footprint. For instance, when evaluating emissions associated with goods distribution, the GLEC Framework recommends the inclusion of both Tank-to-Wheel (TTW) emissions, stemming from fuel combustion within transport vehicle engines, and Well-to-Tank (WTT) emissions, originating from fuel production and distribution. Combining these emissions results in a CO2e factor known as Well-to-Wheel (WTW) (Greene & Lewis, 2019). Figure 4, adapted from the GLEC Framework, visually represents the fuel lifecycle for carbon accounting purposes.

While the GLEC Framework offers extensive guidance on logistics-related greenhouse gas emissions, encompassing transportation and warehousing, our project specifically concentrates on emissions generated exclusively by transportation activities, thereby excluding stationary sources such as fulfillment and regional distribution centers.

Figure 4

The Fuel Life Cycle for Carbon Accounting



Note*:* This figure illustrates the Fuel Life Cycle for Carbon Accounting. The figure was adapted from the GLEC Framework

Source: Global Logistics Emissions Council Framework for Logistics Emissions Accounting and Reporting Version 2.0 (p. 16) by S. Greene & A. Lewis, 2019, Smart Freight Centre.

The GLEC Framework is widely adopted by companies, industry associations, and governments worldwide due to its consistency, credibility, and relevance. Recognizing the challenging nature of the GHGs measuring and reporting task in the transportation sector, the GLEC Framework provides companies with an activity-based approach to calculate their transportation-specific (Scope 3) emissions, especially when primary data on fuel consumption per asset in the transportation fleet is unavailable (Greene & Lewis, 2019). This approach involves three main steps:

Step 1: Compute the total weighted distance (total tonne-kilometers (tkm)): Achieve this by recording the cargo weight and the distance traveled for all shipments carried out by the reporting company within the specified period.

Step 2: Determine the relevant fuel efficiency factor (CO_2e intensity factor) for each shipment: The CO2e for a shipment depends on the transportation mode, asset type, and shipment location, as different regions globally exhibit varying CO_2e factors.

Step 3: Transform the total weighted distance (total ton -kilometers (tkm)) into total GHG emissions: This can be accomplished by multiplying the total weighted distance (tkm) by the fuel efficiency or CO_2e intensity factors for each shipment. This calculation results in the total amount of GHG emissions (measured in kg CO_2 emissions) produced by transportation activities. When using CO_2e intensity factors, ensure the underlying activity data encompasses the entire Well-to-Wheel (WTW) cycle for precise emissions calculations. This step is essential for companies to evaluate their transportation-related emissions and pinpoint opportunities for emissions reduction.

In 2004, The U.S. Environmental Protection Agency (EPA) started a program called SmartWay. The SmartWay program is an initiative that aims to reduce greenhouse gas emissions and other air pollutants from the transportation sector. The program is a public-private partnership that includes the EPA, the freight industry service providers, and other shippers and stakeholders. SmartWay offers a range of tools and resources to help companies in the transportation sector improve their environmental performance, including a free, voluntary certification program that recognizes companies that meet certain emissions standards and performance criteria. Over the years, the EPA's SmartWay initiative has achieved success in reducing greenhouse gas emissions and other pollutants, making it a significant program in promoting sustainable practices within the transportation industry. The GLEC Framework utilizes the EPA's CO_2e estimations for transportation modes and asset types available in the US.

Through a collaborative effort with our project sponsor, we determined the most suitable CO_2e factors for every asset type used by the sponsor to be utilized in our network optimization model. This enabled us to ensure that our model accurately considers the environmental impact of the sponsor's network and effectively minimizes it.

2.3 Business Impact of Carbon Efficient Network Design

Over the past few years, there has been a significant increase in the prominence of sustainable supply chain management practices (Seuring et al., 2008). Market pressures are forcing companies to measure and report their carbon emissions and find ways to minimize their environmental and social impact. These pressures arise from varied factors, such as consumers' growing awareness of environmental and social issues, the increase in regulations, and the rising demand for environmentally and socially responsible products. Moreover, investors now use capital to pressure companies to adopt more eco-friendly strategies. Whether the decision-makers in companies believe in sustainable practices or not, these pressures have made sustainable supply chain practices a critical concern for companies across various industries globally (Hoffman & Woody, 2008; Vélazquez-Martínez et al., 2014). In 2022, over 20,000 companies disclosed their environmental data, including carbon emissions information, to the CDP (CDP, 2022). By doing so, these companies made a public pledge to decrease their carbon emissions, aligning with the goal of the Paris Agreement to limit the rise in global temperatures to 1.5 degrees Celsius above the pre-industrial era.

In their study, Vélazquez-Martínez et al. (2014) examined the roles of transportation costs and carbon emissions when solving the facility location problem for a manufacturing company. They developed a mixed-integer linear programming model that considers both factors in the objective function. Their research findings suggest that considering transportation costs and carbon emissions when deciding where to locate facilities can result in a more sustainable supply chain network design and greater operational efficiency. By optimizing the location of facilities based on transportation costs and carbon emissions, companies can reduce costs, minimize environmental impact, and enhance social and environmental performance (Vélazquez-Martínez et al., 2014). Overall, their research highlighted the importance of considering economic and environmental objectives in supply chain management decisions and provided a valuable framework for optimizing location decisions that balance these objectives. The sponsoring company has pledged to reduce scope one and two emissions by 50% and scope three emissions by 30% by 2030. In their "Road Map to Net Zero" announcement, the company committed to achieving net-zero greenhouse gas (GHG) emissions by 2050, aligning with the

Paris Agreement's 1.5°C target. Minimizing costs and greenhouse gas emissions remains a top priority in the company's optimization efforts.

This project seeks to assist the sponsoring company in identifying the ideal location for a new middle-mile distribution center. By leveraging the outcomes of our MILP program, the company can select a location that offers the most cost-effective and environmentally sustainable solution.

2.4 Summary

In recent years, extensive research has been conducted on Supply Chain Network Design (SCND), with a specific emphasis on addressing traditional SCND problems, carbon emissions accounting, facility location problem (FLP), and green facility location problem (GFLP). The literature review indicates a broad and diverse range of research that has positively contributed to advancing SCND practices. By thoroughly reviewing the literature, we have better understood the current State-of-the-Art in SCND and tailored our methodology to address our research problem and achieve our project's objective.

Our study adopts a multidisciplinary approach, aiming to minimize the total transportation cost and greenhouse gas (GHG) emissions. Specifically, we focus on determining the optimal location of the distribution center to be added to the middle mile of our sponsor company's supply chain. Our research will provide valuable insights for practitioners looking to reduce costs and environmental impacts while improving the efficiency of their supply chains. Our findings are expected to contribute to the body of knowledge on SCND, enabling stakeholders to make more informed decisions in the design and management of sustainable and cost-efficient supply chain networks.

3. DATA

Our project sponsor sought to determine whether introducing a middle-mile fulfillment center into their existing transportation network could lead to reduced transportation costs and decreased carbon emissions. To determine the optimal solution, we assessed various methodologies, each designed to produce the most favorable economic outcomes for this given problem. Our analysis included a comprehensive understanding of how to accurately account for transportation costing, measuring emission outputs from distinct asset types, as well as how to effectively mirror the sponsor company's supply chain for network evaluation. In this chapter, we provide an overview of the data collection process from the project sponsor and outline our general approach. In the subsequent chapters, we will delve into the specific methodology employed, discussing the network optimization strategy, business rules, and general assumptions made while building the optimization model.

Our study utilizes a single-period MILP model to analyze the network of a multi-product supply chain. By constructing a single-period, multi-echelon network model, we aimed to identify opportunities for reducing carbon emissions from transportation and overall costs, improving the sponsor company's business performance. This study optimizes the multi-criteria objective of minimizing variable transportation costs while reducing carbon emissions.

A few important considerations regarding our research:

1. Our scope of the SCND problem was limited to the middle-mile distribution.
2. The proposed research utilized a multi-objective approach to solve the SCND problem, which involves the parallel minimization of two objective functions, namely emissions and total cost.

3. Due to the limited availability of data and the scope of our research, our model concentrated solely on costs and emissions linked to transportation asset sources. Our analysis excluded other supply chain decisions, such as operational and energy expenses related to facilities.
4. In our research, we have selected the GLEC Framework as the preferred method for carbon emission accounting. This approach is not only the most comprehensive available, but it also aligns with the primary method utilized by the sponsor company.

3.1. Data Collection & Analysis

Our data collection process employed a structured approach to gather precise, relevant, and quantitative information about our project sponsor, ensuring a thorough understanding of the organization. The steps involved in this process included:

1. Identifying and defining the project goals and sponsor expectations for the study.
2. Establishing the scope of data collection by specifying the business units, timeframes, and products to be incorporated in the study, including any past or future analysis/implementation on the supply chain network.
3. Collecting historical data from the project sponsor's ERP, TMS, and ancillary systems.
4. Analyzing and interpreting the collected data sets.
    1. Processing, cleaning, and realigning unstructured data and outliers
    2. Created assumptions in data for missing, omitted, or unreachable data sets
5. Conducting interviews with business unit leaders to quantify any qualitative aspects not captured in the data.
6. Building a baseline information, goods, and monetary flow in a MILP optimization tool consisting of nodes/arcs and pre-existing business constraints to replicate historical data.

7. Receiving final agreement from the project sponsor that the data presented is an accurate representation of business rules, monetary flows, and product flows across all nodes and arcs.

Following this structured process ensured accuracy, relevance, and completeness in our study of the project sponsor. The following section will discuss our data processing at greater length.

3.1.1 Data Processing

To prepare the data for modeling, we took several steps to ensure the dataset's cleanliness, accuracy, and flexibility should new scenarios emerge, or should we face changes to the project's scope, or the business rules encompassed in the model. The process we followed can be described as follows:

1. Collect all relevant data and organize it into a structured format that the optimization model can easily understand.
2. Cleanse/evaluate - remove any data indicative of duplicates, inconsistencies, or outliers (infeasible shipment sizes, product classes that were out of scope, etc.)
3. Create assumptions to ensure the continuity of the model and fill in any missing data gaps. Missing values were imputed with historical averages, physical limitations, and input from business unit subject matter experts based on historical assumptions.
4. Feature selection - utilize Excel to standardize tables for the optimization software to digest. Standardizing the data helps reduce the problem's dimensionality and improve the performance of the optimization model
5. Validate the cleaned data before using it in the optimization model. Confirm historical spending, physical flows weight and quantity, and business rules that contributed to the

current state of the network (the baseline). This approach helps to ensure that the data is accurate and dependable.

The goal of cleaning data for a network optimization problem was to ensure that the data is accurate, consistent, and relevant to the problem at hand so that the optimization model can produce reliable and usable results. We extracted supplier shipments, inter-site shipments, product flows to customers, all logistics associated with shipments [inclusive of mode, asset type, asset class, and asset restrictions], and the corresponding supplier, distribution center/warehouse, and customer geographic locations. This data was used to develop a baseline model of the project sponsor's business. The model consists of aggregated physical, monetary, and information flows, representing averages and aggregated shipment data. The model did not simulate historical transactions but aggregates the total weight and cost moved across the network in the given time.

3.1.2. Model Inputs

The inputs selected for this network optimization model were designed to capture the intricate and multifaceted nature of the sponsoring company's supply chain. The model incorporated production policies, transportation assets, flow constraints, and internally defined emission calculations to reflect the sponsor's business rules accurately.

3.2. Products

The project sponsor's freight mix comprises two primary product categories: bulk and non-bulk products. Bulk products were goods typically transported and stored in copious quantities or volumes. These included powders like cement, barite, bentonite, and other industrial-grade liquid materials such as base oils, all used in various drilling, mining, construction, and manufacturing applications. The model did not aggregate products to a product family or parent product. Instead, the model considers each product to be distinct based on its own set of characteristics.

Due to their size and quantity, bulk products are historically often transported in specialized equipment designed for bulk transport, such as tankers, bulk carriers, and hopper cars. Bulk products are typically stored in large silos, whether mobile or stationery, designed explicitly for bulk storage. Non-bulk materials, known as general cargo, are goods typically transported and stored in smaller quantities or volumes and often require individual packaging or handling. This category includes palletized materials such as chemicals, parts, and other goods transported on pallets. Non-bulk materials are commonly transported in standard shipping equipment such as flatbed trucks and dry vans. Non-bulk materials can be quickly loaded and unloaded with conventional material-handling equipment such as forklifts and ramps. Unlike bulk products, non-bulk materials do not require specialized equipment for transport and storage, making them more accessible for smaller shipments and easier to manage during transport.

During our data collection process, we encountered a diverse range of packaging types for various products, including super sacks, bags, drums, totes, and an assortment of liquid tanks in varying sizes. Products were measured in pounds, kilograms, or tons. To streamline our model, we converted all products to a standardized weight in pounds. Consequently, in our product table, each item is assigned this standard weight, irrespective of its initial unit of measurement.

Our model focused solely on finished goods and does not account for product shelf life. All bulk products are assigned a standard weight in tons, with one ton equaling 2,000 pounds. Pounds are consistently used as the unit of measurement throughout the model. Additionally, we conducted conversions for bags, pallets, and other units to establish a uniform metric within the model.

3.3. Assets

The project sponsor's network comprises several transportation asset types that move goods across the network. Our model incorporated all these assets and only permitted products to be transported using the same historical assets they moved on. This approach ruled out any possibility of consolidating products across diverse types of equipment not historically associated with the products. Due to the specificity of each product's dimensions and weight, each asset's capacity was a weight function, with limited exceptions. In our historical data, there were instances where a product movement may be within an asset's stated weight and cubic metrics. Nevertheless, its' configuration was unique in that it was allocated an entire asset. Our model included these flows, costs, and goods movements. These shipments were constrained in both the baseline and subsequent optimization scenarios. These product/asset combinations did not allow for optimization on several vehicles, several products per vehicle, or reduction if the optimization algorithm attempts to divide the total weight across multiple assets. Although the asset and product combination are constrained and grouped in the model, that combination has an opportunity to optimize across different nodes. It was essential to include all historical shipments for the total cost to serve a purpose.

Our baseline and all subsequent scenarios presumed an infinite number of assets available in the network. Our project sponsor utilizes both open markets brokers and their rates, dedicated and a private fleet in some instances. Transportation assets in our model have a standard maximum weight based on DOT regulation. Any historical shipments with weights exceeding the standard maximum have been reduced to the maximum pounds per asset.

3.4. Transportation Policies – Costing, Carriers, and CO_2 consideration

3.4.1. Costing and Carriers

To accurately estimate transportation costs in the project sponsor's network, we implemented a standardized cost aggregation approach for each shipping lane (node-to-node connection) in our transportation cost model. Our model includes both fixed and variable costs. The fixed cost is calculated per shipment per asset type, while the variable cost is based on the distance traveled by each asset type. We have customized transportation policies to suit different products and assets. We performed a linear regression analysis on the historical data to estimate the fixed and variable costs. To enhance the performance of our linear regression model, we introduced a linehaul binary model. The binary model takes a value of 1 if the distance traveled is greater than 500 miles, and 0 if otherwise. The resulting linear regression equation is expressed as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$$

Here, Y is the dependent variable, representing the estimated shipping cost on a specific lane using a specific equipment type. β_0 is the intercept that represents the fixed shipping cost per shipment per equipment type. β_1 is the distance coefficients or slope, where X_1 is the distance traveled on that lane in miles. β_2 is the long-haul coefficient or slope, where X_2 is the binary value that takes 0 if the distance is less than 500 miles and 1 if otherwise. ε represents the error terms.

We consolidated the results of our various linear regression models into a table that contains the shipping cost components of each equipment type. This repository is then utilized as an input for our network optimization model.

Below are a few assumptions regarding how we constructed the model:

1. This model does not consider fixed transportation costs associated with packaging, handling, vehicle/route acquisition, brokerage fees, and insurance cost.
2. The historical data comprised of numerous carriers providing transportation services for the project sponsor across various lanes.
3. This model blended all carriers' rates into an average cost per mile per pound at the asset type level.
4. We did not utilize a split-by-ratio analysis, forcing a percentage of weight or shipments to be moved by differentiating carriers with historical rates. Instead, we blended all carrier rates across all product and asset-specific historical shipments. This averaging considers expedited/prioritized shipments, 'milk runs,' and everything in between (within an appropriate standard deviation of actuals).

3.4.2. Calculating Distance Matrix

To develop our network optimization model, we utilized the distance module of the Python Geopy Library to calculate the Euclidean distance between nodes in the supply chain network. By inputting the (latitudinal, longitudinal) coordinates of each node, we efficiently computed the distance in miles and kilometers to create a comprehensive distance matrix of all potential shipping lanes. We chose the Python Geopy Library for its widespread use and support, providing us with extensive documentation and community assistance throughout the development process.

To enhance the accuracy of our distance matrix, we trained a machine learning model using historical on-the-road distances to predict the estimated on-the-road distance based on the Euclidean distance for each specific lane. We utilized linear regression to predict the on-the-road distance for new lanes and introduced a long-haul binary variable that takes a value of 1 for Euclidean distances greater than 500 miles and 0 otherwise to improve the model's accuracy.

The linear regression equation that calculates on the road distance (OTR) based on Euclidean distance (ED) and the binary variable for long haul (BH) can be expressed as follows:

$$OTR = \beta_0 + \beta_1 ED + \beta_2 BH + \varepsilon$$

where:

β_0 is the intercept

β_1 is the coefficient of ED

ED is the Euclidean distance in miles

β_2 is the coefficient of BH

BH is a binary variable that takes a value of 1 for Euclidean distances greater than 500 miles and 0 otherwise

ε is the error term

3.4.3. CO_2 Consideration

Our optimization model employed the GLEC framework for carbon emissions accounting. Our model incorporates CO_2 emissions into transportation costs using a feature in the optimization software we are using called "greenhouse gas emissions modeling." This feature considered the carbon dioxide and other greenhouse gases generated by various transportation modes, including trucks, rail, ocean vessels, and air, integrating them into the overall transportation cost.

With a primary focus on carbon emissions from trucking operations, our model applied a carbon emission factor known as the CO_2 equivalent (CO_2e). Each asset type had a distinct CO_2e value, which aligns with the US EPA SmartWay program. In the model, the CO_2e is multiplied by the weighted distance of every shipment to estimate the carbon emissions associated with that shipment, expressed in kilograms of CO_2. The model aimed to minimize the transportation costs and the total GHG emissions in the network.

To determine the WTW (well-to-wheel) factors for CO_2 that we used use in our modeling, we consulted the GLEC (Global Logistics Emissions Council) framework, specifically Table 40, which contains emission intensity factors for North American roads. Each asset type utilized in our sponsor's network was assigned a corresponding SmartWay category that was deemed appropriate after discussions with the project sponsor, and we obtained the WTW value from the table for that category. This value, expressed in kg CO_2e/t-km, was be multiplied by the distance and weight of the shipment in question to determine the total amount of CO_2 emitted in kilograms because of the transportation of that shipment. Our model only considered emissions from mobile sources, such as trucks, and not from stationary sources, such as supplier manufacturing plants, supplier warehouses, RDCs, or fulfillment centers. In relation to CO_2 costing, our project sponsor has proposed utilizing a rate of $35 per ton of CO_2.

3.5 Demand

The demand inputs for our model are represented as an aggregated total demand for end customers. Demand is differentiated by individual physical products shipped to specific end customer sites within the network, and this model's demand covers each product type. Employing aggregated demand was an appropriate approach for tackling this network optimization problem, as it does concentrate on strategic network design rather than tactical operations.

3.6. Nodes

Our model does not cluster customers into zones as does Tsiakis et al. (2001). Instead, the customer's physical location is treated as a variable in the objective function for each scenario. Tsiakis et al. (2001) based customer zones on European sovereign borders, which limited transportation cost accuracy to statistical averaging of last-mile distribution. Our network comprises approximately one hundred end-customer locations, as depicted in (Figure 5). This model was focused on the transportation element CO2 reduction to capture the exact trucking costs and distance to the end customer. Our model sought to ensure more realistic transportation costs to end customers, and this also ensured higher accuracy regarding emissions from different modes and transportation asset types. Demand is assumed to be known, utilizing historical demand by well-site and overall longevity of oil extraction in each geographical region. Similarly, our suppliers and all company owned warehouses and regional distribution centers are represented in the model as distinct physical locations (Figure 6).

Figure 5

Distribution of Sponsor Company's U.S. Customer Base

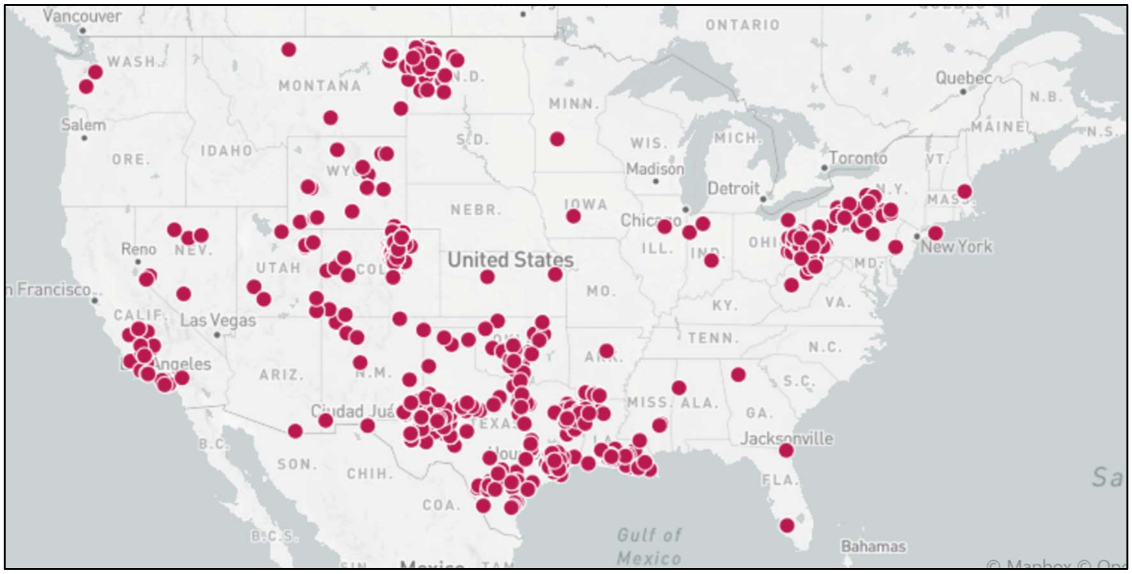


Figure 6

Sponsor Company's U.S. Network

Our model does include non-essential nodes and arcs, unrelated to the optimization of choosing a potential mid-mile distribution center. We included fixed constraints associated with port to supplier and port to regional distribution centers. Subsequently we included all outbound transportation of goods and costs associated with shipping products to the end customer, in anticipation of the model's future use by the project sponsor. These are fixed arcs in the model.

3.7. Network Flows

The "middle mile distribution network," an essential component of any company's distribution system, refers to the intermediate stage of a supply chain in which goods are transported from primary facilities, such as manufacturing plants or supplier warehouses, to regional warehouses (RDCs). Multiple iterations of scenarios with varying objective functions were executed to provide the business with different options. Each scenario in the Results section will reveal trade-offs between different model iterations about their corresponding variables.

3.8. Warehouse (Regional Distribution Center) – Costing and Policies

In our model, each regional distribution center is not subject to the following expenses:

1. Fixed operating costs required to achieve specific levels of throughput capacity
2. Costs associated with closing an existing site
3. Fixed startup costs for establishing a new facility
4. Handling costs for processing and managing goods
5. Production costs related to suppliers

3.9. Other Model Considerations

With any historical data set, there are incomplete or missing transactional records. In our exploratory data analysis, we discovered inconsistent volumes and, in many cases, non-existent shipments (that would be needed to complete a supplier-to-customer flow). We could not distinctly track a product from a supplier to the end customer through the project sponsor's network. Even on an aggregated level, specific node-to-node movements did not exist. This issue resulted in an inherent material imbalance in the network design model. Therefore, after discussions with the project sponsor, we decided to allow the historical movements of our existing products to dictate the good's movements, transportation costs, and corresponding financial transactions in the network.

Our model did not use a dummy location to fulfill linking constraints between sites or sites to customers. In some network optimization models, a 'dummy location' is used to fulfill inbound product to sites with no transactional volume. These sites have a fixed production quantity with no manufacturing cost and zero cost is applied to move these goods to the customer. This allows the model to have an instantaneous allocation of goods. This is also considered in some network optimization models as a superior approach instead of allowing nodes to carry or have initial inventory. Initial inventory assumes that the cost to transport goods to that node has already been incurred. Additionally, a modeling approach may be to include an inflationary transportation cost to the outbound lane to account for those missing inbound transportation cost. We included these physical movements that did not exist in the data and forced them to incur transportation cost because we knew that they existed historically and that our model optimizes transportation movements.

This model does not allow for inventory holding at the end of the time horizon. All products must be created, shipped, and fulfilled at the end customer in the single period.

Subsequently, sites and products passing through them do not incur a storage fee. Despite some of the advanced works in multi-period models with multi-objective functions, including Arntzen et al. (1995)'s MILP models, this research is confined to a single period model. Furthermore, the historical data reviewed does indicate any significant seasonality, lending itself to the decision to refrain from using a multi-period model. Similarly, all sites in the network have an unlimited capacity for throughput. There is no maximum constraint. This is an annualized model built for network design, and therefore capacity is not a critical element of the model. Additionally, no production or supply constraints exist in the baseline or any subsequent scenarios. Goods that are produced are forced to flow through the model in an optimal way to satisfy customer demand constraints.

4. METHODOLOGY

The methodologies section will cover the following topics:

1. **Modeling –** overview and scenario selection
2. **Scenarios** – baseline and the two mid-mile distribution centers
3. **Optimization** – calculations and objective functions
4. **Formulation** – parameters, decision variables, formulations

4.1 Modeling

The model consisted of a MILP that optimized demand scenarios and determined the optimal solution to including/excluding a new mid-mile objective function. It was constructed using a commercially available network optimization software package.

4.2 Scenario Selection

Scenarios are defined as unique collections of individual supply, demand, nodes, and arc representations of the project sponsor's data. Our project consists of several different scenarios to test against the project sponsor's baseline spend and flow of goods.

Table 1:

Model Scenarios

Scenario	Demand	Linking Constraints/Nuances
Chose mid-mile site (Houston)	2022 - 6 months of demand	Potentially include a binary fixed cost for managing/opening site
Chose mid-mile site (Oklahoma City)	2022 - 6 months of demand	Potentially include a binary fixed cost for managing/opening site

4.3 Optimization

After choosing relevant scenarios, MILP was used to prescribe an optimal trade-off between reducing the transportation cost and minimizing carbon emissions while meeting the demand from each scenario. Due to the inherently conflicting nature of reducing cost and reducing carbon emissions, a multi-criteria optimization engine was chosen to solve the problem. The following sections discuss defining the objective function, determining the decision variables, and formulating the logical constraints.

4.4 Optimization Calculation:

4.4.1 Multi-Criteria Objective Functions

Finding the best solution between two opposing objectives can be difficult. In many cases, optimizing one objective comes at the expense of another; therefore, there is no single "best" solution that can satisfy all objectives simultaneously. Instead, a set of solutions that represent the trade-offs between different objectives can be identified. Increasing transportation mileage typically leads to an increase in carbon emissions for those shipments. There are two distinct approaches to multi-criteria objective functions: sequential and parallel optimization. Sequential optimization lets you extend the optimization objectives beyond the traditional cost minimization and profit maximization. It allows the modeler to select from several additional objectives and constraints to use in solving your model. These objectives include all costs within the model such as transportation cost, inventory holding cost and fixed operating cost. Additionally, deviation of all hard constraints (conditions that must be satisfied exactly for a solution to be considered feasible) used in the model from their specified values can be accessed in sequential optimization.

These successive objectives and constraints are inclusive of all costs (transportation, manufacturing, inventory, and operating costs (fixed and variable). In addition, non-cost-based variables such as greenhouse gas emissions can be included in the optimization. Sequential optimization creates a hierarchy based on the users input and prioritizes each defined value based on its ranking. By extending the optimization objectives beyond the traditional minimize cost and maximize profit, we allowed the model to optimize on maximizing non cost specific objectives. As an example, our goal was to minimize the total transportation cost, then minimize total emissions output:

Table 2

Sequential Optimization

Objective	Priority	Weight
Total Transportation Cost	1	1
Total Emissions Output	2	1

In contrast, in parallel optimization, multiple optimization criteria are applied simultaneously to the problem, and their results are combined to obtain an improved solution. Additionally, instead of running sequential optimization, in which the optimization engine solves the first objective (minimize cost) and then subsequent objectives, we could assign weight to each objective. If we assigned an equal weighting of 0.5 to each objective it would treat them as equals.

Table 3

Parallel Optimization

Objective	Priority	Weight
Total Transportation Cost	1	0.5
Total Emissions Output	1	0.5

By default, weighted optimization artificially increases the value of a lower ranked objective. In the example below, if we assigned a weight of two to total emissions output cost, an emissions output cost of $100,000 in the model is treated as a cost of $200,000.

Table 4

Weighted Optimization

Objective	Priority	Weight
Total Transportation Cost	1	1
Total Emissions Output	1	2

Our model compared both types of solver approaches and we discuss them further in the succeeding sections.

4.4.2. Formulation:

To support the presentation of our mathematical model, we start by furnishing a verbal depiction of the said model in the following manner:

Objective Function: Minimize Total Cost

Total Cost = Transportation Cost + Carbon Emission Cost.

Subject to:

- Satisfying all customer demand
- Balancing the flow between different nodes in the supply chain network
- Nonnegativity and binary constraints
- Only one fulfillment center to be introduced at a time

Parameters:

O_i Capacity of supplier i ($i \in S$)

D_i: Demand of customer l ($l \in C$)

C_{ikn}: Cost to ship from supplier i to middle-mile distribution center candidate k using transportation mode n ($i \in S, k \in K, n \in N$)

C_{kjn}: Cost to ship from middle-mile distribution center candidate k to regional distribution center j using transportation mode n ($k \in K, j \in R, n \in N$)

C_{jln}: Cost to ship from regional distribution center j to customer l using transportation mode n ($j \in R, l \in C, n \in N$)

C_{iln}: Cost to ship from supplier i to regional distribution center j using transportation mode n ($i \in S, j \in R, n \in N$)

d_{ik}: Distance between supplier i and middle-mile distribution center candidate k ($i \in S, k \in K$)

d_{kj}: Distance between middle-mile distribution center candidate k and regional distribution center j ($k \in K, j \in R$)

d_{jl}: Distance between regional distribution center j and customer l ($j \in R, l \in C$)

d_{il}: Distance between supplier i and regional distribution center j (i ∈ S, j ∈ R)

e_n: CO_2e (Carbon Dioxide Equivalent) emissions per unit weight per unit distance for transportation mode n ($n \in N$)

E: Cost per kilogram of CO_2 emitted

U_n: Maximum shipment weight capacity for transportation mode n ($n \in N$)

M: A substantial number used in the middle-mile distribution center constraint

Decision variables:

X_{iknq}: Shipment quantity of product q from supplier i to middle-mile distribution center candidate k using transportation mode n.

Y_{kjnq}: Shipment quantity of product q from middle-mile distribution center candidate k to regional distribution center j using transportation mode n.

Z_{jlnq}: Shipment quantity of product q from regional distribution center j to customer l using transportation mode n.

V_{ilnq}: Shipment quantity of product q from supplier i to regional distribution center j using transportation mode n.

Binary variable: $W_k = 1$ if middle-mile distribution center candidate k is selected, 0 otherwise.

In mixed-integer linear programming (MILP) decision variables are the aspects that represent the quantities or values that we seek to optimize subject to the constraints of the problem. Decision variables can take on any numerical value within their respective domains, subject to certain conditions. Decision variables are the components that can change when solving the mixed-integer-linear-programming. These are the quantities that the program is trying to determine to solve the problem. Decision variables are the factors that influence cost in our model. They represent the quantities that we can control or adjust to achieve our objective.

The key feature of MILP is that some of the decision variables are restricted to taking on only integer values (e.g., 0, 1, 2, etc.), while others can take on any continuous value. This is in contrast to linear programming (LP), where all decision variables are continuous.

The term "mixed" in MILP refers to the presence of both integer and continuous decision variables in the same optimization problem. The integer variables can represent decisions that are either yes/no or choices between discrete options, while the continuous variables can represent quantities or amounts that can be any real number within a certain range. The objective is to find the combination of product quantities and choices that minimizes the total production cost while satisfying the resource constraints. This is a mixed-integer linear programming problem, where the decision variables include both continuous and integer variables.

Objective function:

Minimize Total_Cost = α * Transportation Cost + (1-α) * Carbon Emissions Cost

Transportation Cost = $\sum_{i=1}^{S} \sum_{k=1}^{K} \sum_{j=1}^{R} \sum_{l=1}^{C} \sum_{n=1}^{N} \sum_{q=1}^{Q} (C_{ikn} * d_{ik} * X_{iknq} * C_{kjn} * d_{kj} * Y_{kjnq} + C_{jln} * d_{jl} * Z_{jlnq} * C_{iln} * d_{il} * V_{ilnq})$

Carbon Cost = $(\sum_{i=1}^{S} \sum_{k=1}^{K} \sum_{j=1}^{R} \sum_{l=1}^{C} \sum_{n=1}^{N} (C_{ikn} * d_{ik} * X_{iknq} * C_{kjn} * d_{kj} * Y_{kjnq} + C_{jln} * d_{jl} * Z_{jlnq} * C_{iln} * d_{il} * V_{ilnq}))$* Carbon Cost per Kg

Constraints:
$\sum_{j=1}^{R} \sum_{n=1}^{N} \sum_{q=1}^{Q} Z_{jlnq} \geq D_i \ \forall l \in C$

Flow conservation constraint: $\sum_{n=1}^{N} \sum_{q=1}^{Q} (X_{iknq} - Y_{kjnq}) = 0, \forall \ i \in S, k \in K, j \in R$

Flow conservation constraint: $\sum_{n=1}^{N} \sum_{q=1}^{Q} (V_{ilnq} - Z_{jlnq}) = 0, \forall \ i \in S, j \in R, l \in C$

Binary constraint: $\sum_{k=1}^{K} W_k = 1$

Non-negativity constraint: $X_{iknq}, Y_{kjnq}, Z_{jlnq}, V_{ilnq} \geq 0, \forall \ i, k, j, l, n, q$

Shipment weight constraint: $\sum_{q=1}^{Q} X_{iknq} \leq U_n, \ \forall i, k, n$

Shipment weight constraint: $\sum_{q=1}^{Q} Y_{kjnq} \leq U_n, \forall \ k, j, n$

Shipment weight constraint: $\sum_{q=1}^{Q} Z_{jlnq} \leq U_n, \forall \ j, l, n$

Shipment weight constraint: $\sum_{q=1}^{Q} V_{ilnq} \leq U_n, \forall \ i, j, n$

Middle-mile distribution center constraint: $Y_{kjnq} \leq M * W_k, \forall \ k, j, n, q$ (Where M is a substantial number)

5. RESULTS

The key research question of this project was whether a mid-mile distribution center would reduce the total transportation cost and emissions discharged by manipulating different decision variables in a mixed-integer linear program network optimization model. Subsequently the goal was to determine whether opening one of the two candidate locations would be cost and emissions advantageous. The project sponsor also sought to assess the efficacy and impact of the network design changes considering vendor managed inventory.

In this chapter, we showcase the optimization results and articulate the limitations of the current model. Finally, we demonstrate the additional criteria that could be used to investigate further the cost, flows, and emissions of the network with the intended goal of making business sense of the network.

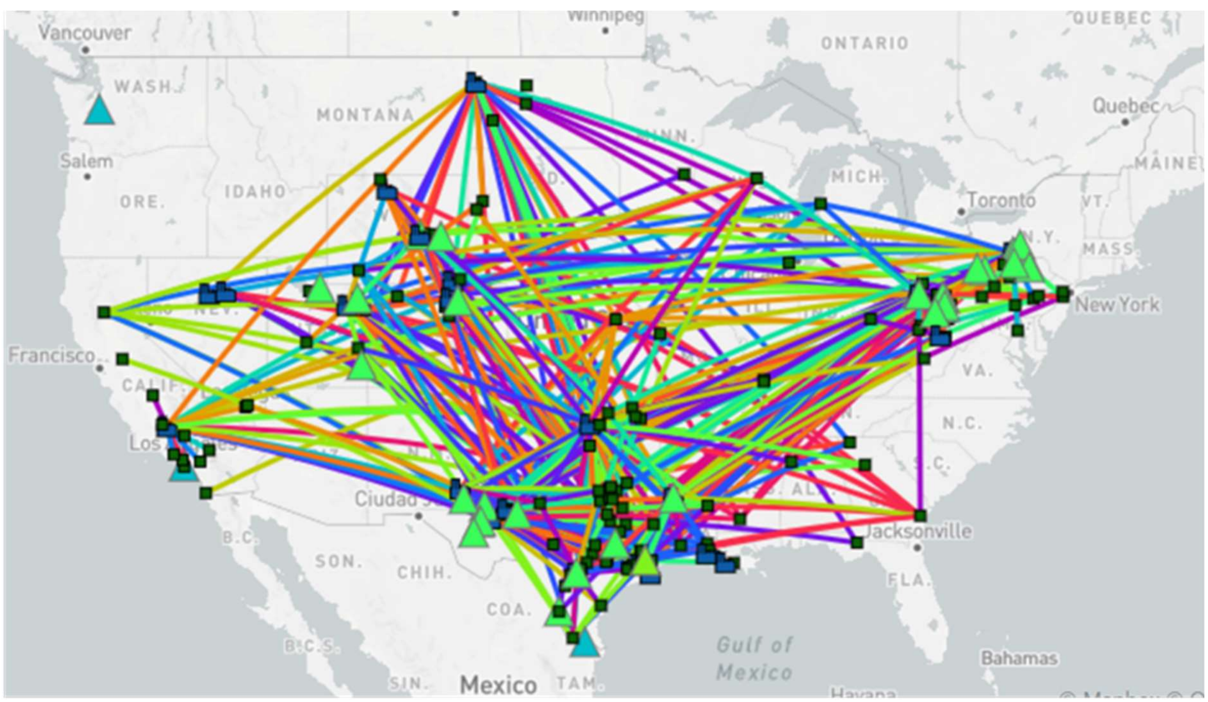
5.1 Baseline Model

The model consists of replicating the existing network and building a baseline model. In the baseline model, we inputted the location for the two candidate mid-mile distribution centers as well as the existing network (Figure 7). We established transportation lanes with subsequent fixed and variable costs for twenty-one modes/transportation asset types. These modes transport goods in historical shipment sizes. The baseline model forces historical flows of weight, distance, transportation mode, and product, irrespective of how non- or cost-efficient. The goal of the baseline model was to replicate the total spend, movement of goods, and business constraints that are historically accurate.

We exclusively used the shipment history obtained from the project sponsor to generate the inter-site movements of goods as well as the end customer demand. We cleaned the order data to include historical shipments with corresponding costs and emissions. We removed all outliers such as orders with order quantity of zero and shipments missing costs. With the values we used for transportation and operating costs, we generated a total five-month cost of $18.0M. The actual costs incurred by the project sponsor are $18.1M. The transportation cost of ~$18.0M included with the emissions cost of ~$.45M (at $35/ton/km) accounted for a total baseline cost of $18.1M.

Figure 7

Baseline network 2022 (colors represent different products)

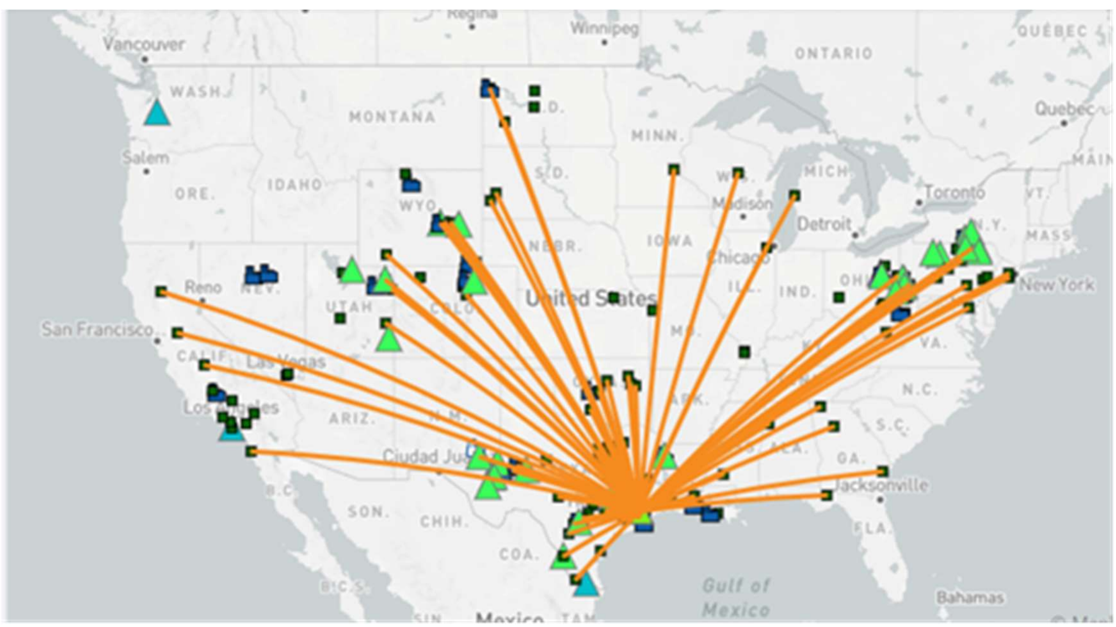


5.2 Scenario 1: Force all non-bulk product inter-site through Houston, Texas

In our first network design scenario we forced all non-bulk products from suppliers to a greenfield site in Houston, TX. The intent was to consider whether it was cost advantageous to have a distribution center consolidate all goods. These goods were then forced to leave the mid-mile distribution center in Houston and move to the next echelon in the network, our project sponsor's plants. From there, the plants were forced to ship the minimum quantities to other plants or to customers based on historical movements. Deliveries to customers are fixed from plants that historically performed those shipments. We did allow for the model to optimize based on production quantities (if suppliers that historically made/shipped products did not have a production capacity constraint). The model did not favor a single supplier for most of the production.

Figure 8

Houston, Texas – Scenario 1



This optimization scenario consolidated loads inbound/outbound of Houston. The model historically shipped goods in cost-inefficient loads across the network. With the optimization that forces goods in/out of Houston, the model has the opportunity to reduce costs by consolidating goods on specific modes to plants. This allowed the model to spread out the lowest cost across the greatest amount of weight. With the total mileage increasing due to being forced to ship goods to Houston, the total transportation cost moved across the network increased to $18.6M. In this scenario emissions were reduced by sending non-bulk products to Houston. The total emission cost incurred decreased to $40K from $45K (~11%). Despite the total cost increasing the model found opportunities to consolidate weight across fewer total shipments. Additionally, we ran the same scenario for forcing all non-bulk goods through a greenfield location in Oklahoma City, OK.

5.3 Scenario 2: Force all non-bulk product inter-site through Oklahoma City, Oklahoma

In this second scenario, we forced all non-bulk products from suppliers to a greenfield site in Oklahoma City, OK. Overall transportation costs and emissions increased from the baseline. The total transportation cost was ~$19.2M and the C02 emissions cost was ~$46k. Both the transportation and emissions costs were higher than the baseline or choosing Houston as a greenfield site. All other constraints are replications of the first scenario to Houston, TX.

Figure 9

Oklahoma City – Scenario 2

Table 5

Summary of scenario results

Scenarios	Demand	Total Transportation Cost	Total Emissions Cost*
Baseline	2022 - 6 months of demand	$17.9M	$0.45M
Chose mid-mile (Houston)	2022 - 6 months of demand	$18.6M	$0.40M
Chose mid-mile (Oklahoma City)	2022 - 6 months of demand	$19.2M	$0.46M

Note. **A standard cost of $35/ton was applied to each ton/km of C02 in our model*

As shown in Table 5, opening a mid-mile distribution center in either Houston or Oklahoma is not cost advantageous. Despite the model's outputs not being cost reducing, this project proved to be a useful building block for the organization to discuss the strategic importance of vendor managed inventory, as well as to begin measuring their emissions footprint. In the next section we discuss the results and their value to the project sponsor.

6. DISCUSSION

The results of the two scenarios and a corresponding sensitivity analysis made it evident that our initial hypothesis, which introducing a mid-mile distribution center in either Houston or Oklahoma City would be cost and/or emissions saving, was incorrect. Each of the scenarios resulted in an increase in total cost. This was primarily caused by the additional mileage that assets needed to incur to move goods to either of the mid-mile locations. While a large percentage of our project sponsor's inter-site movements exists in the central south, inefficiencies and corresponding additional costs emerge when producers and suppliers in a distant region have to incur travel to those central points.

Additionally, in our scenarios we forced historical plant to plant shipments to incur costs. To compare our historical baseline to each scenario, it was important that we considered the entire network, inclusive of the daily inefficiencies of operations. Allowing for a complete optimization to occur would provide results that were unrealistically optimal. Correspondingly, the model incurred those costs. Our primary focus was on the inbound from supplier lanes and their corresponding outbound lanes only. We encourage the project sponsor to consider running additional scenarios to review a fully optimized network design to comprehend the best-case scenario of include a mid-mile distribution center.

Although our hypothesis was incorrect, the outcomes of the modelling research have a multitude of advantages for our project sponsor. A replication of their business in a mathematical format that reflects their modes of transportation, nodes in the network, transportation costing and asset utilization is a foundational structure for dealing with the complexities of strategic design. The project sponsor can test ideas in a low cost, low risk environment. This can help the project sponsor avoid costly mistakes or unintended consequences that may arise from implementing these types of solutions in the real world. Moreover, this business unit now has the GLEC framework built around their business unit. Our hope is that the business continues to add additional business units into this model. Not only will the network design be more robust and have opportunities for consolidation and profit-sharing, but an opportunity to continue to discover the total emissions of the North American business.

We recommend that the project sponsor build on the multi-criteria objective function approach to network design. Multi-criteria optimization can also facilitate stakeholder engagement and participation in the decision-making process. By involving stakeholders in the selection of objectives and the evaluation of alternatives, our project sponsor can ensure that the optimization process reflects the values and priorities of the organization. This can help build trust and support for the transportation or overall supply chain planning process and improve the overall quality of decision making. Conclusively, multi-criteria optimization can help solve the tradeoff problem between transportation asset cost and carbon emissions by integrating multiple objectives, exploring different scenarios, and engaging stakeholders in the decision-making process.

Going forward we would like to provide recommendations for other researchers pursuing a network optimization project focused on reducing emissions and transportation costing. Firstly, consider advancing research in transportation costing by utilizing sophisticated machine learning techniques. Utilizing machine learning techniques to regress and prognosticate fixed and variable costs is a useful way to create costs for new transportation lanes that have not existed historically. Consider a regional, distinct asset type, multi-period approach to these types of modelling. We recommend random forest, support vector, or neural network regression.

Additionally, consider defining an industry standard cost of carbon. There is limited research in this space. Governing bodies across the world are still working on a standard definition of the true monetary cost of carbon. This differs by country, industry, and practices. Finally, consider carbon offset as a possibility instead of just carbon reduction. As more industries are coping with participating in carbon reduction, this is an introductory area to participate in.

6.1 Managerial Implications of Findings

The project sponsor's main goal was to be able to service their end customers while maintaining a cost-efficient network. Investing in a mid-mile distribution center did not prove to be a cost-efficient option in our modeling efforts. This model does not consider the cost implications of real estate, labor, inventory, and maintenance costs of opening and running a distribution center. The internal rate of return on an investment in a distribution center would need to fully consider all those costs to be comprehensive and eventually presented for a business case.

Despite this, by factoring in CO_2 emissions, we were able to balance cost and environmental impact for the project sponsor. Additionally, our transportation modeling can also help the project sponsor identify opportunities to reduce carbon emissions by optimizing their transportation routes and modes, reducing empty miles, and consolidating shipments. Once the project sponsor determines their true dollar cost of emitting carbon, they can utilize this model's structure and optimization to determine not only the baseline, but the different scenario's carbon costs. This does not contribute to the project sponsor's pledge of a 50% reduction in scope 1 and 2, and a 30% reduction in scope three emissions by 2030. Going forward we hope that the project sponsor will continue to refine emissions and all other elements in this network design model to develop and challenge the business-as-usual mentality.

With the evaluation of alternatives, our project sponsor and future teams can ensure that the optimization process reflects the values and priorities of the community. This can help build trust and support for the transportation or overall supply chain planning process and improve the overall quality of decision making. In summary, multi-criteria optimization can help solve the trade-off problem between transportation asset cost and carbon emissions by integrating multiple objectives, exploring different scenarios, and engaging stakeholders in the decision-making process.

One of the most advantageous aspects of our approach to modeling is the allowance of additional scenarios, developments, and evolvements of the project sponsor's supply chain network design. Our intent is that the sponsor will continue to utilize this model for testing additional elements of their cost structure and overall network. Ideally the project sponsor would maintain and update the historical flows, costs, and any other subsequent changes to the network to accurately test working hypotheses. This is a cost advantageous approach to strategic initiatives. Despite the efforts and results of network design, there are limitations of considering a mathematical model to be an altruistic faultless representation of the intricacies of a working business.

6.2 Limitations of the Current Model

In this optimization we utilized the five months of historical shipments that the project sponsor provided. All of this historical data is from May 2022 – October 2022. This model included incorrect 'eaches' in the shipment quantity field, forced the project team to develop a shipment weight based on total weight moved on a single shipment divided by the standard weight for each product. Additionally, the team had to develop standard weights for products. The project sponsor was not able provide an accurate or historical standard product weight, partially due to the differential chemical makeups of each product used at a distinct well-site. Initially the project team struggled with data integrity due to fusing from two ERP systems in the months following an ERP conversion. Subsequently there were outliers in the data set, as an example: shipment weights exceeding DOT vehicle regulation weights. We encourage the project sponsor to create standards regarding product weight, product naming, product formulations, and supplier to product relationships for business units interested in additional network design.

7. CONCLUSION

This project addressed the problem of whether a mid-mile distribution center would reduce emissions and transportation costs for an oil and gas services company. To address this problem, we researched industry best practices in network design and optimization, carbon emissions in supply chain networks, and ways to measure carbon emissions in transportation. From there, we performed data analysis and processing on our project sponsor's existing data. This included making sense of the business' rules, understanding and choosing a proper criterion for dealing with outliers, and making decisions on disparate or incomplete data. Then we had to develop transportation costing based on a regression analysis for each mode in our dataset for the model. Additionally, we calculated distances between all of our physical nodes in the network to create a more realistic depiction of over the road transportation distance. After developing a model with these inputs, we tested two scenarios after ensuring that the baseline matched the actual total weight and transportation moved. These scenarios led us to the conclusion that opening a mid-mile distribution center in Houston, TX or Oklahoma City, OK would cost more in terms of transportation costing and overall emissions cost than the current supply chain network design for the business unit.

At a strategic level our research helps the project sponsor determine that when considering transportation cost and emissions alone, opening a mid-mile distribution center in Houston or Oklahoma City is not advised. For further planning and execution, they now have a baseline mathematical model to test hypothetical scenarios on. These types of strategic scenarios can often be difficult to quantify and costing to implement in the real world. Our hope is that this working model will serve as a template for further extrapolation on the complexities of their business.

REFERENCES

Ageron, B., Gunasekaran, A., & Spalanzani, A. (2012). Sustainable supply management: An empirical study. *International Journal of Production Economics, 140(1), 168–182.* https://doi.org/10.1016/j.ijpe.2011.04.007

Alamsyah, A.-V., & Purevdorj, N. (2021). *Carbon Efficient Network Design: Evaluating The Trade-Offs Between Carbon Emissions, Transportation Cost and Delivery Time For a Middle-Mile Distribution Network*. [Capstone Research Project, Massachusetts Institute of Technology] https://dspace.mit.edu/handle/1721.1/130976

Benjaafar, S., Li, Y., & Daskin, M. (2013). Carbon Footprint and the Management of Supply Chains: *Insights From Simple Models. IEEE Transactions on Automation Science and Engineering, 10(1), 99–116*. https://doi.org/10.1109/TASE.2012.2203304

Boukherroub, T., Bouchery, Y., Corbett, C. J., Fransoo, J. C., & Tan, T. (2017). Carbon Footprinting in Supply Chains. In Y. Bouchery, C. J. Corbett, J. C. Fransoo, & T. Tan (Eds.), Sustainable Supply Chains: *A Research-Based Textbook on Operations and Strategy (pp. 43–64). Springer International Publishing*. https://doi.org/10.1007/978-3-319-29791-0_3

Bouzembrak, Y., Allaoui, H., Goncalves, G., & Bouchriha, H. (2011). A multi-objective green supply chain network design. 2011 4th International Conference on Logistics, 357–361. https://doi.org/10.1109/LOGISTIQUA.2011.5939315

CDP. (2019). CDP Global Supply Chain Report 2019. Retrieved from https://www.cdp.net/en/research/global-reports/global-supply-chain-report-2019

Cholette, S., & Venkat, K. (2009). The energy and carbon intensity of wine distribution: *A study of logistical options for delivering wine to consumers. Journal of Cleaner Production, 17(16), 1401–1413*. https://doi.org/10.1016/j.jclepro.2009.05.011

Chudak, F. A., & Shmoys, D. B. (2003). Improved Approximation Algorithms for the Uncapacitated Facility Location Problem. *SIAM Journal on Computing, 33(1), 1–25*. https://doi.org/10.1137/S0097539703405754

Climate Change: Atmospheric Carbon Dioxide | NOAA Climate.gov. (n.d.). Retrieved December 11, 2022, from http://www.climate.gov/news-features/understanding-climate/climate-change-atmospheric-carbon-dioxide

Drezner, Z., & Hamacher, H. W. (2004). Facility Location: Applications and Theory. Springer Science & Business Media.

Farahani, R. Z. (Ed.). (2009). Facility location: Concepts, models, algorithms and case studies. Physica-Verl.

Frequently Asked Questions (FAQs)—U.S. Energy Information Administration (EIA). (2022.). Retrieved December 11, 2022, from https://www.eia.gov/tools/faqs/faq.php

Golpîra, H., Zandieh, M., Najafi, E., & Sadi-Nezhad, S. (2017). A multi-objective multi-echelon green supply chain network design problem with risk-averse retailers in an uncertain environment. *Scientia Iranica, 24(1), 413–423*. https://doi.org/10.24200/sci.2017.4043

Greene, S., Lewis, A. (2019). Global Logistics Emissions Council Framework for Logistics Emissions Accounting and Reporting 2019. Amsterdam, Netherlands: Smart Freight Centre

Hoffman, A. J., & Woody, J. G. (2008). Climate Change: What's Your Business Strategy? *Harvard Business Press*. https://aiche.onlinelibrary.wiley.com/doi/10.1002/aic.10617

Jiang, Y., Zhao, Y., Dong, M., & Han, S. (2019). Sustainable Supply Chain Network Design with Carbon Footprint Consideration: A Case Study in China. *Mathematical Problems in Engineering, 2019, e3162471*. https://doi.org/10.1155/2019/3162471

Labuschagne, C., Brent, A. C., & van Erck, R. P. G. (2005). Assessing the sustainability performances of industries. Journal of Cleaner Production, 13(4), 373–385. https://doi.org/10.1016/j.jclepro.2003.10.007

Li, S., Li, X., Zhang, D., & Zhou, L. (2017). Joint Optimization of Distribution Network Design and Two-Echelon Inventory Control with Stochastic Demand and CO2 Emission Tax Charges. PLOS ONE, 12(1), e0168526. https://doi.org/10.1371/journal.pone.0168526

Martínez, J. C. V., & Fransoo, J. C. (2017). Green Facility Location. In Y. Bouchery, C. J. Corbett, J. C. Fransoo, & T. Tan (Eds.), Sustainable Supply Chains: A Research-Based Textbook on Operations and Strategy (pp. 219–234). Springer International Publishing. https://doi.org/10.1007/978-3-319-29791-0_9

Melo, M. T., Nickel, S., & Saldanha-da-Gama, F. (2009). Facility location and supply chain management – A review. European Journal of Operational Research, 196(2), 401–412. https://doi.org/10.1016/j.ejor.2008.05.007

Moradi, E., & Bidkhori, M. (2009). Single Facility Location Problem. In R. Zanjirani Farahani & M. Hekmatfar (Eds.), Facility Location: Concepts, Models, Algorithms and Case Studies (pp. 37–68). Physica-Verlag HD. https://doi.org/10.1007/978-3-7908-2151-2_3

Nearly 20,000 organizations disclose environmental data in record year as world prepares for mandatory disclosure—CDP. (n.d.). Retrieved March 3, 2023, from https://www.cdp.net/en/articles/media/nearly-20-000-organizations-disclose-environmental-data-in-record-year-as-world-prepares-for-mandatory-disclosure

Nozick, L. K. (2001). The fixed charge facility location problem with coverage restrictions. Transportation Research Part E: Logistics and Transportation Review, 37(4), 281–296. https://doi.org/10.1016/S1366-5545(00)00018-1

Peng, Y., Ablanedo-Rosas, J. H., & Fu, P. (2016). A Multiperiod Supply Chain Network Design Considering Carbon Emissions. Mathematical Problems in Engineering, 2016, e1581893. https://doi.org/10.1155/2016/1581893

Rajeev, A., Pati, R. K., Padhi, S. S., & Govindan, K. (2017). Evolution of sustainability in supply chain management: A literature review. Journal of Cleaner Production, 162, 299-314. https://doi.org/10.1016/j.jclepro.2017.05.026.

Rezaee, A., Dehghanian, F., Fahimnia, B., & Beamon, B. (2017). Green supply chain network design with stochastic demand and carbon price. Annals of Operations Research, 250(2), 463–485. https://doi.org/10.1007/s10479-015-1936-z

Seuring, S., Sarkis, J., Müller, M., & Rao, P. (2008). Sustainability and supply chain management – An introduction to the special issue. Journal of Cleaner Production, 16(15), 1545–1551. https://doi.org/10.1016/j.jclepro.2008.02.002

Smart Freight Centre, GLEC. (2021). Global Logistics Emissions Council Framework. Retrieved from https://www.smartfreightcentre.org/en/solutions/glec-framework

Tamir, A. (2001). The k-centrum multi-facility location problem. Discrete Applied Mathematics, 109(3), 293–307. https://doi.org/10.1016/S0166-218X(00)00253-5

The Carbon Trust. (2013). Scope 3 Indirect Carbon Emissions Reporting and Business Operations. Retrieved from https://www.carbontrust.com/resources/guides/carbon-footprinting-and-reporting/scope-3-indirect-carbon-emissions-reporting

The U.S. Shale Revolution. (n.d.). The Strauss Center. Retrieved December 11, 2022, from https://www.strausscenter.org/energy-and-security-project/the-u-s-shale-revolution

Tsiakis, P., Shah, N., & Pantelides, C. C. (2001). Design of Multi-echelon Supply Chain Networks under Demand Uncertainty. Industrial & Engineering Chemistry Research, 40(16), 3585–3604. https://doi.org/10.1021/ie0100030

Turkeš, R., Sörensen, K., & Cuervo, D. P. (2021). A matheuristic for the stochastic facility location problem. Journal of Heuristics, 27(4), 649–694. https://doi.org/10.1007/s10732-021-09468-y

U.S. Field Production of Crude Oil (Thousand Barrels per Day). (2022.). Retrieved December 11, 2022, from https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=MCRFPUS2&f=M

US EPA, O. (2016, March 3). SmartWay [Collections and Lists]. https://www.epa.gov/smartway

US EPA, O. (2022). Sources of Greenhouse Gas Emissions [Overviews and Factsheets]. https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions

Varsei, M., & Polyakovskiy, S. (2017). Sustainable supply chain network design: A case of the wine industry in Australia. Omega, 66, 236–247. https://doi.org/10.1016/j.omega.2015.11.009

Vélazquez-Martínez, J. C., Fransoo, J. C., Blanco, E. E., & Mora-Vargas, J. (2014). The impact of carbon emissions on facility location decisions. Transportation Research Part E: Logistics and Transportation Review, 66, 53-66.

Wang, F., Lai, X., & Shi, N. (2011). A multi-objective optimization for green supply chain network design. Decision Support Systems, 51(2), 262–269. https://doi.org/10.1016/j.dss.2010.11.020

Wang, J., Wan, Q., & Yu, M. (2020). Green supply chain network design considering chain-to-chain competition on price and carbon emission. Computers & Industrial Engineering, 145, 106503. https://doi.org/10.1016/j.cie.2020.106503

Watson, M., Lewis, S., Cacioppi, P., & Jayaraman, J. (2013). Supply Chain Network Design: Applying Optimization and Analytics to the Global Supply Chain. Pearson Education.

What Happened to Oil Prices in 2020. (2022). Investopedia. Retrieved December 11, 2022, from https://www.investopedia.com/articles/investing/100615/will-oil-prices-go-2017.asp

World Business Council for Sustainable Development & World Resources Institute. (2011). *Technical Guidance for Calculating Scope 3 Emissions*. Retrieved from https://ghgprotocol.org/sites/default/files/standards/Scope3_Calculation_Guidance_0.pdf.

Wu, L.-Y., Zhang, X.-S., & Zhang, J.-L. (2006). Capacitated facility location problem with general setup cost. *Computers & Operations Research, 33(5), 1226–1241.* https://doi.org/10.1016/j.cor.2004.09.012

Yang, J., Guo, J., & Ma, S. (2016). *Low-carbon city logistics distribution network design with resource deployment. Journal of Cleaner Production, 119, 223–228.* https://doi.org/10.1016/j.jclepro.2013.11.011

Zhou, P., & Wen, W. (2020). Carbon-constrained firm decisions: *From business strategies to operations modeling. European Journal of Operational Research, 281(1), 1–15.* https://doi.org/10.1016/j.ejor.2019.02.050