HEC MONTREAL

Facility Location and Sizing in Network Design: A Case Study with Bulk and Packaged Products

By

Julian Martinez

Global Supply Chain Management

A Thesis Submitted in Partial Fulfillment
of Requirements for the Degree of
Master of Science

August 2015

Résumé

L’objectif de cette thèse est d’identifier la configuration optimale d’une chaîne d’approvisionnements, afin de fournir un aperçu de l’emplacement et de la taille des entrepôts pour un ensemble de produits. Un cas réel dans l’industrie des lubrifiants est abordé ici, afin de représenter une chaîne d’approvisionnements dans des circonstances réalistes. Le présent cas prend en compte à la fois les produits en vrac et en lots, ce qui doit être expédié via différents modes de transport afin d’atteindre les clients. Divers scénarios jugés réalistes par l’entreprise sont évalués, avec pour objectif d’identifier la meilleure solution en termes monétaires. L’impact de la variation de certains coûts est analysée afin d’identifier la sensibilité des résultats face à ces variations de coûts, ce qui est difficile à prévoir dans l’horizon de dix ans présenté dans le cas étudié.

Afin de représenter la chaîne d’approvisionnements dans laquelle l’entreprise devra distribuer ses produits, un modèle de réseau logistique flexible fut adapté au cas étudié. Ce modèle optimise les coûts fixes et variables liés à l’entreposage et le flux des produits à travers toute la chaîne d’approvisionnements. Au regard du modèle dans son ensemble, plusieurs paramètres furent retenus tandis que d’autres furent rejetés, en gardant à l’esprit l’emplacement et la taille optimal en tant qu’objectif principal de cette étude.

Les résultats liés à l’implantation du modèle, à l’aide des données fournies par l’entreprise, mesurent le coût total de fonctionnement du réseau logistique avec les emplacements sélectionnés, tout en s’assurant que la demande demeure satisfaite. Ces emplacements sont présentés comme faisant partie intégrante de la solution, ainsi que le transfert de clients aux emplacements appropriés, l’affectation des produits aux emplacements correspondants, et le mode de transport choisi pour acheminer les produits entre les différents points. Enfin, le volume d’entreposage propre à chaque produit contribue à définir la taille des installations afin de faire fonctionner l’entreprise efficacement.

Abstract

The purpose of this thesis is to identify the optimal configuration of a supply chain in order to provide insights about the location and size of facilities for a set of products. A real case in the lubricants industry is addressed so as to represent a supply chain under realistic circumstances. The case considers both bulk and packaged products, which have to be shipped by different modes of transportation to reach the customers. Different scenarios deemed feasible by the company are evaluated with the aim of identifying the optimal solution in monetary terms. The impact of varying some costs is analyzed to identify the sensitivity of the results to cost variations, which are hard to predict in the ten-year horizon presented in the case.

To represent the supply chain that requires the company to distribute its products, a flexible logistics network design model was adapted to the case. The model optimizes the fixed and variable cost of storing and the flow of products through the entire supply chain. From the complete model some parameters were retained while others dismissed, having in mind the optimal location and size of facilities as the main objective of the study. Particularities of the industry and the case were embedded into the model to ensure a faithful representation and reliable results.

The output of implementing the model, with the data provided by the company, indicated the total cost of running the logistics network with the locations selected while making sure that demand is satisfied. It presents the locations that are part of the solution, the assignment of customers to locations, the assignment of products to locations, and the transportation mode to move products between the different nodes. Finally, the quantity to store for every product contributes to define the size of the facilities to efficiently operate the business.

Table of Contents

Abstract .... **i**

Table of Contents .... **ii**

List of Tables .... **v**

List of Figures .... **vii**

Acknowledgments .... **ix**

1. Introduction 1

1.1 Supply Chain General Overview 1

1.2 The Lubricants Case Description 5

1.2.1 Introduction 5

1.2.2 Objectives and Key Decisions 10

1.2.3 Type of Products and Modes of Transportation 13

1.2.4 Operational Costs and Constraints 15

1.3 Problem Statement 16

1.4 Thesis Outline 17

2. Literature Review 18

2.1 Supply Chain Network Design 18

2.1.1 Facility Location 19

2.1.2 Configuration of the Network 21

2.2 Research Background in Similar Cases 29

2.2.1 Proposals with Bulk Products 29

2.2.2 Proposals in the Lubricants and Oil Industry 31

3 General Optimization Model ... 33

3.1 Origin and Conception of the Model ... 33

3.2 Mathematical Model ... 35

3.2.1 Notation ... 36

3.2.2 Formulation ... 41

3.2.3 Extensions ... 43

4 Preparation of the Model to the Case ... 46

4.1 Selection of Decision Variables ... 46

4.2 Assignment of Parameters ... 48

4.2.1 Assignment of Parameters for Bulk Products ... 49

4.2.2 Assignment of Parameters for Packaged Products ... 55

4.3 Selection of Constraints and Extensions ... 59

4.4 Resulting Model ... 61

5 Data Set up in the Model ... 62

5.1 Model Implementation Overview ... 62

5.2 Matching Tool Development ... 64

5.3 Model Execution ... 71

5.4 Selected Scenarios to Evaluate ... 72

5.4.1 Bulk Products ... 73

5.4.2 Packaged Products ... 73

5.4.3 Data Sorting and Transformation ... 74

6 Results and Analysis 77

6.1 Description of the Output File 77

6.2 Results and Analysis for Bulk Products 78

6.2.1 Fixed Cost Sensitivity Analysis 84

6.2.2 Transportation Cost Sensitivity Analysis 89

6.2.3 Inventory Cost Sensitivity Analysis 91

6.3 Results and Analysis for Packaged Products 93

6.4 Combining Products Analyses 97

6.5 Comparison with the Existing Distribution Network 98

6.6 Limitations 100

7 Conclusions and Future Research 102

Bibliography 104

Appendix 107

List of Tables

Table 1: North America Lubricants Refinery Capacity 7

Table 2: Forecast of U.S. Lubricants Demand for 2018 9

Table 3: Types and Names of Products 14

Table 4: Summary of Notation for Sets 39

Table 5: Summary of Notation for Parameters 40

Table 6: Summary of Notation for Decision Variables 40

Table 7: Link of Decision Variables to the Case 47

Table 8: Extract of Detailed Demand for Bulk Products 50

Table 9: Extract of Fixed Cost for Bulk Products 51

Table 10: Transportation Cost from the Plant to Warehouses for Bulk Products .. 52

Table 11: Extract of Transportation Cost from Warehouses to Customers for Bulk Products 54

Table 12: Holding and Handling Cost per Warehouse for Bulk Products 55

Table 13: Extract of Detailed Demand for Packages Products.............................. 56

Table 14: Transportation Cost from the Plant to Warehouses for Packaged Products 57

Table 15: Extract of Transportation Cost from Warehouses to Customers for Packaged Products 58

Table 16: Holding and Handling Cost per Warehouse for Packaged Products 58

Table 17: Extract of Turnover Ratio per Product 60

Table 18: Extract of Flow Table Aggregated by Source 78

Table 19: Extract of Flow Table with Destination Detail for Bulk Products 80

Table 20: Extract of Flow Table with Product Detail for Bulk Products.................. 81

Table 21: Extract of Inventory Table Aggregated by Warehouse 82

Table 22: Total Cost Composition for Bulk Products 83

Table 23: Fixed Cost Values to Evaluate 85

Table 24: Impact of Fixed Cost Change on Products Assignment 86

Table 25: Total Cost Composition for Every Change in Fixed Cost....................... 88

Table 26: Impact of Transportation Cost Reduction on Flowing Quantities........... 90

Table 27: Inventory Cost in Ontario to Evaluate 91

Table 28: Extract of Flow Table for Packaged Products.. 94

Table 29: Extract of Inventory Table for Packaged Products................................. 95

Table 30: Total Cost Comparison for Bulk Products... 98

List of Figures

Figure 1: The Supply Chain Process 2

Figure 2: The Supply Chain Network Structure 3

Figure 3: Global Lubricants Demand by Region in 2013 6

Figure 4: U.S. Consumption of Lubricants in 2012 8

Figure 5: Plant and Potential Locations for New DCs 11

Figure 6: Network Structure to Optimize 12

Figure 7: Example of a Solution to a *p*-Median Problem with $p = 10$ 20

Figure 8: Solution Methods for Supply Chain Network Problems 25

Figure 9: Forecasting Demand for Bulk Products 49

Figure 10: Extract of C Language Programming of the Model 63

Figure 11: Display of Excel File and TXT Sheet 65

Figure 12: Display of Create File Button Functionality 66

Figure 13: Plants Structure Representation in the Spreadsheet 67

Figure 14: Display of Duplicate Data Button 68

Figure 15: Display of Warehouses, Customers, and Modes Tabs 69

Figure 16: Display of Complete File Button 70

Figure 17: Display of Export File Button 71

Figure 18: Scenarios to Evaluate for Packaged Products 74

Figure 19: Impact of Fixed Cost Change on Warehouses… 87

Figure 20: Impact of Inventory Cost Change on Products Concentration.............. 92

Figure 21: Total Cost of Scenarios Evaluated for Packaged Products 93

Figure 22: Adjustment of Results to Pack Sizes... 96

Acknowledgments

Much time, effort, and inspiration were devoted to come to the end of this voyage, but it is the people who backed me unconditionally who made it possible. I would like to express my sincere thankfulness and appreciation to the people who have supported and encouraged me during the pursuit of this thesis.

I would like to express the deepest gratitude to my advisors, Professors Jean-Francois Cordeau and Raf Jans for their support in guiding me through this process. I thank Professor Cordeau for sharing years of research and work with me; not only with his wise advice, but also with the tools that were a foundation in this study, and for entrusting the model that gave the first spark to undertake this journey. Thanks to him, I had the opportunity to immerse myself in the world of coding and modeling, which became one of my assets during the development of this thesis. Professor Jans supported me academically and emotionally since the starting point of this path. I appreciate his encouragement and patience during this challenge.

I would like to extend my thanks to the people in the Company that made this project possible, and contributed to it through the development and solution of this case. Their valuable time and accessibility helped my understanding of the case. I appreciate their enthusiasm and willingness for constructive and fruitful discussions, which led to the thorough analysis of results, and ensured the practical usefulness of this study.

1. Introduction

1.1. Supply Chain General Overview

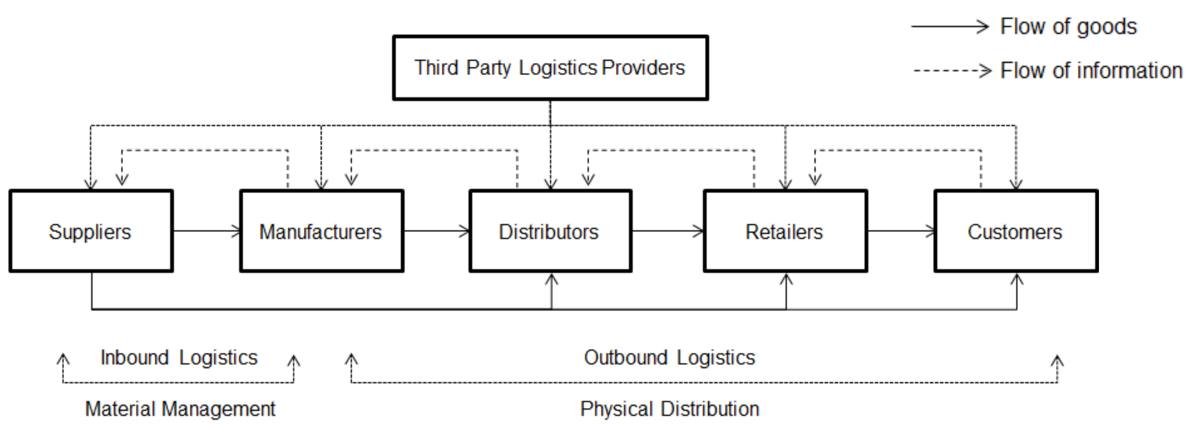
A supply chain refers to the entire network of entities taking part in the activities to make a service or good available to customers. These activities ensure the flow and transformation of goods from raw materials to finished products. It is becoming increasingly difficult to ignore the relevance of the linkage between the organizations in the network as a main source of benefit for the supply chain.

One can talk about supply chain management when there is a proactive relationship between buyer and supplier in all the tiers across the supply chain (Cox, 2004). Integration of independent entities leads to larger benefits to the whole network compared to individual performance optimization. The integration consists in the ability to share information efficiently, resulting in favorable aspects aligned with the main supply chain objective, which is to maximize the overall value generated. The benefits of integration have been largely discussed since the birth of the supply chain concept. The reduction of the so-called bullwhip effect, where variability in orders increases when we move upstream in the supply chain (Forrester, 1962), is one of the most widely known benefits. Other benefits are related to reduction in processes execution costs, lead-times and employee workloads (Groznik, 2006).

Representing a supply chain in order to capture the benefits mentioned above has become a central issue for logistics research in recent years. Among the steps required to draw the complete network, defining the entities and the connections to each other is the first and most relevant step. Typically, entities are in one of the following categories: suppliers, manufacturers, distributors, retailers or customers. The connection is made by modes of transportation depending on the flow of materials or goods. Furthermore, many third-party logistics providers may be involved in performing different activities throughout the supply chain. The result of such an analysis should be a graphic representation that allows the understanding of the network to continue with the analysis of the variables to optimize. The number of entities, tiers and links depend on the scope of the problem to solve. The dilemma between model complexity and reality should be addressed by the model builder in order to represent reality while keeping it simple enough to solve (Hokey & Gengui, 2002).

Independent of the type of products or nature of organizations interacting in the supply chain, two processes are always present to make it work: Inbound and outbound logistics. Inbound logistics has also been described as material management. It accounts for the activities concerning the acquisition of raw materials, such as storage, quality control, and warehousing of products (Johnson & Malucci, 1999). On the other side, outbound logistics has also been described as physical distribution, which encompasses the activities related to providing finished products to customers. Some of these activities are receipt ordering and processing, product handling, consolidation, and outbound transportation (Bowersox & Closs, 1996). A simple representation of the supply chain process is shown in Figure 1.

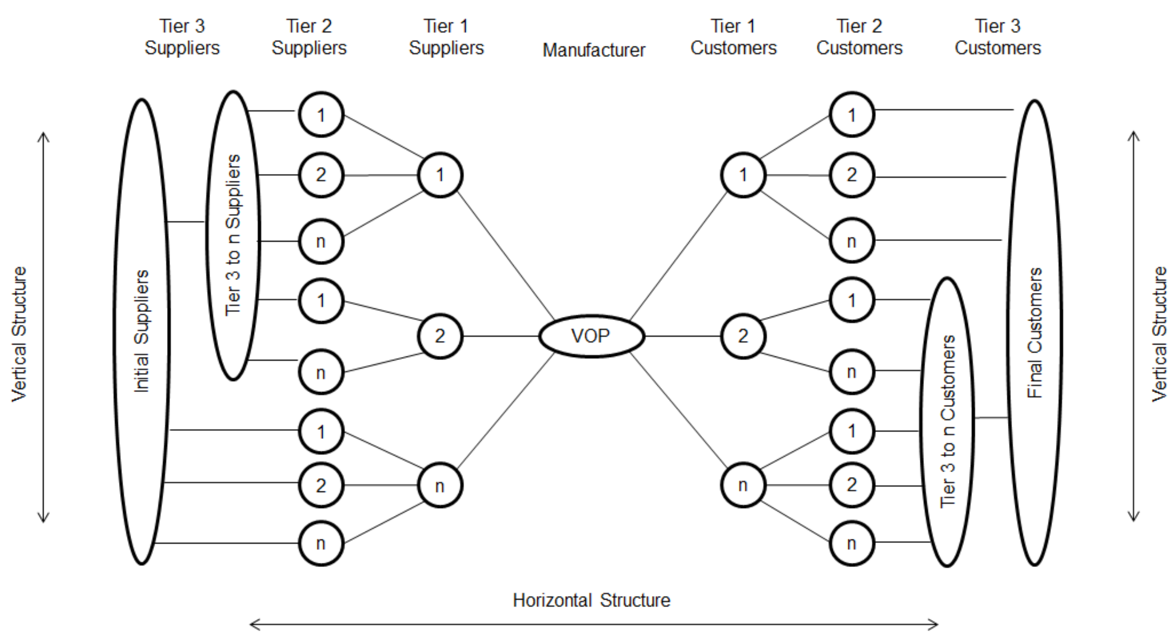
Figure 1: The Supply Chain Process



Source: (Hokey & Gengui, 2002)

The horizontal representation above may be complemented with a vertical structure. While the horizontal structure refers to the tiers forming the network, the vertical one responds to the number of entities at each tier (Lambert, Cooper, & Pagh, 1998). The possible connections between all the entities create a web that is sometimes difficult to represent graphically; nevertheless, it has to be modeled mathematically since that is the supply chain network itself. Such representation is known as supply chain modeling, which has been an object of research since the earliest efforts of Glover, Jones, Karney, Klingman, and Mote (1979) to build an integrated model. A two-structure representation of the supply chain is shown in Figure 2, where a manufacturer makes again the division between the tiers of raw material suppliers and the tiers of customers.

Figure 2: The Supply Chain Network Structure



Source: (Lambert et al., 1998)

The middle position for the manufacturer plays an important role in the figure above since it is the entity transforming the raw materials into finished goods. Although all the tiers add value to the product, after the manufacturer no further value is usually added. This middle point is also known as the value offering point

(VOP), where customers' allocated demand is aggregated, and quantities required of raw materials are known (Holmstrom, Hoover Jr., Eloranta, & Vasara, 1999). The VOP is also useful to define the scope of the problem to address. Some problems may consider the entire network from supplier to customers, whereas others can focus on only one segment. The first part of the supply chain, prior to the VOP, is denoted as the upstream segment. Similarly, the tier responsible for the consumption of the product is known as the downstream segment.

Modeling a supply chain network requires three key elements: decision variables, an objective function and constraints. Decision variables represent the main decisions, such as defining if and where facilities should be located, the number of facilities or equipment needed, the volume of a product to purchase, to produce or to store. The objective function defines the aim of the optimization problem expressed with the decision variables. Traditionally, objective functions attempt to minimize the costs or maximize the profit of the logistics network. Finally, the constraints represent the restrictions or limitations imposed over the range of decisions alternatives, and determine the feasibility of a solution (Hokey & Gengui, 2002).

Despite the increased interest in building optimization models to solve different aspects of the supply chain, the literature about the potential use of using existing models is limited. A variety of modeling tools are now available in the market. However, building a new model requires an expertise that is not common in enterprises that model their supply chains only once in a while. Companies may prefer to leverage on third parties rather than training their people in issues outside their core competencies. Re-using or adapting a supply chain model arises as an unexplored option that can yield benefits to companies. Modifying an existing model according to the requirements of the network has the potential to consume fewer resources than the traditional modeling fashion. Moreover, companies may focus their efforts in executing the model and analyzing results. Furthermore, companies can take advantage from the experience embedded in existing models (Bell, De Cesare, Lycett, Mustafee, & Taylor, 2007).

In this thesis, we used a flexible model that is capable of representing a wide variety of supply chains. The model was selected because it had the conditions required to solve the realistic supply chain problem in our case. As the human process of creating the model was already done, we focused on the preparation of data and analysis of results. For customizing the model easily accessible computer tools were used. The model was set to represent the supply chain of a lubricants company that requires identifying the optimal location for its distribution centers (DCs) within the United States. Moreover, the resulting configuration of the network indicates the assignment of customers to every DC for every product demanded throughout the country. The solution also provides the volume per product to flow between facilities and by which transportation mode. Finally, the volume per product to store in DCs is the input to calculate the size of facilities in a ten-year horizon. Therefore, defining locations, assignments of customers to DCs, assignments of products to DCs, determining the optimal flow through the network with the optimal transportation modes, and sizes of DCs are the main objectives of this thesis. Some particularities of the industry and the company make this case a unique attempt to provide solutions that facilitate decision-making for managers. A description of the case is presented in the next section.

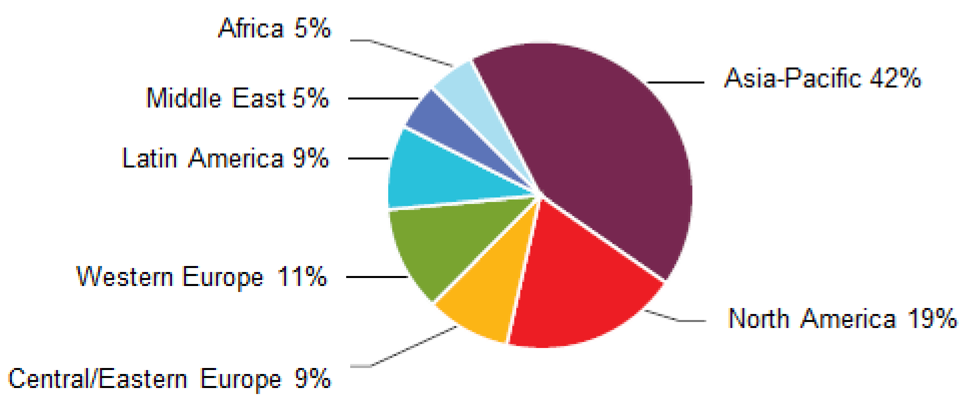
1.2. The Lubricants Case Description

1.2.1. Introduction

Many energy companies worldwide specialize in exploration and production of oil and gas among other sources of energy. Extraction of crude oil is one of the main activities of these companies with operations in the oil industry. After the refinery process, a set of valued products can be produced from crude oil such as gasoline, diesel, kerosene, and lubricants, among others. When a 42-gallon barrel of crude oil is processed, the expected average outputs are 19 gallons of gasoline, 12.1 gallons of diesel and 3.9 gallons of kerosene. These three products account for 83% of the initial barrel, while lubricants represent only 1% (LNG Publishing

Company, 2014). Companies in the lubricants industry have a significantly large portfolio of products to serve two target markets: commercial and industrial. The former demands car engine oils and transmission fluids, the latter demands industrial engine oils, processed oils, and general oils used in industries as diverse as chemicals, printing, mining, manufacturing, food, and pharmaceutical, among many others.

Figure 3: Global Lubricants Demand by Region in 2013



Source: Lubricants Industry Factbook (2014)

Figure 3 presents the percentage by region of the global demand of lubricants worldwide. The total demand for 2013 was around 35.3 million metric tons, where North America appears as the second largest market for this industry. Even if sold quantities in this region remained flat for the period between 2008 and 2013, the price for finished lubricants experienced a 30% increase in the same time period, confirming an upward trend that has been present since 2000. A market of this size represents a huge opportunity for lubricants manufacturers located in North America, where their geographical location gives them an advantage with respect to Persian Gulf countries or South Korea, the main external sources of lubricants for the region.

The case addressed in this thesis is based on one of those manufacturers, whose name is not mentioned in this thesis for confidentiality reasons. It will be simply called “the Company” in the coming sections. As a reference, it is worth mentioning some of the most considerable finished lubricants manufacturers in North America. Table 1 presents those companies sorted by their production capacity, and the number of refineries dedicated to lubricants they possess.

Table 1: North America Lubricants Refinery Capacity

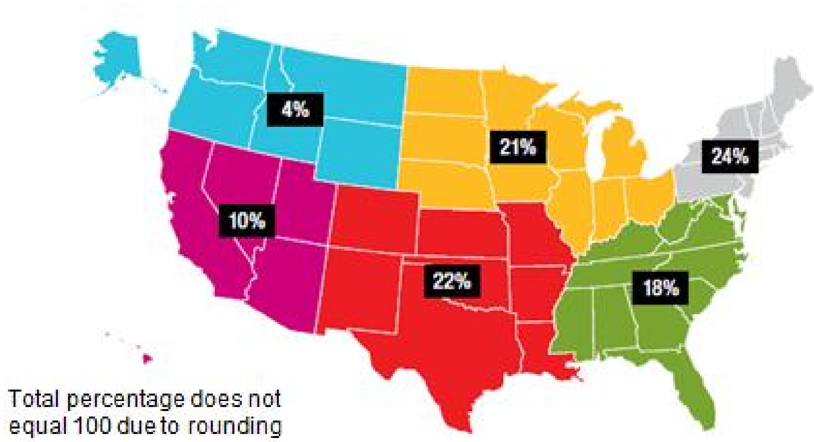
	Company	Locations	2014 Capacity (b/d)*
UNITED STATES	ExxonMobil	3	47 500
	Chevron	2	45 700
	Motiva	1	40 300
	Ergon	2	26 800
	Excel Paralubes	1	22 200
	Calumet	2	18 800
	PBF Energy	1	11 000
	HollyFrontier	1	9 500
	San Joaquin Refining	1	8 100
	Cross Oil	1	5 000
	LyndellBasell	1	4 600
	American Refining Group	1	2 400
	Valero	1	2 400
	Subtotal	**18**	**244 300**
CANADA	Petro-Canada	1	15 600
	Imperial Oil Ltd.	1	2 400
	Subtotal	**2**	**18 000**
	Total	**20**	**262 300**

* Barrels per Day

Source: Lubricants Industry Factbook (2014)

The supply chain network studied in this case is limited to the one serving the American market, namely the United States. After having studied the market in that country, the Company decided to undertake a project to efficiently fulfill the demand of its customers for the next 10 years. The objective is to optimize the logistics network to enable long term growth plans. The new network has to guarantee the minimum possible cost of operation, while ensuring full shipments to customers. The first phase of this project started in 2013 brainstorming ideas inside the Company. Concepts were transformed into data to be analyzed, and preliminary decisions about the scope of the project were made. Much data about the American market were analyzed in order to identify demand patterns and forecast it for the coming years.

Figure 4: U.S. Consumption of Lubricants in 2012



Source: U.S. Energy Information Administration (2014)

Figure 4 presents the distribution of the total demand in the United Stated, aggregated by region. North eastern states of the country represent the highest demand with 24%, where Pennsylvania is the leader of the region with a strong concentration in industrial lubricants. This state shows the third highest demand in the country. Afterwards, the two central regions account for 43% of the total, the north central states consumption being led by Ohio. In the south central states, Texas represents more than 40% of the demand of the eight states forming this region and it is the state with the highest demand in the country. With respect to the western regions, California accounts for more than 60% of the demand in all the states in the western area, and the second highest demand in the country.

Table 2: Forecast of U.S. Lubricant Demand for 2018

ITEM	2008	2013	F 2018	% Annual Growth 2008 - 2013	% Annual Growth 2013 - 2018
Engine Oils	1 150	1 125	1 150	-0.4	0.4
Process Oils	460	455	470	-0.2	0.7
Transmission Fluids	465	455	460	-0.4	0.2
Metalworking Fluids	148	141	145	-1.0	0.6
Other Lubricants	232	224	235	-0.7	1.0
Total Lubricants Demand	**2 455**	**2 400**	**2 460**	**-0.5**	**0.5**

Million Gallons

Source: Lubricants Industry Factbook (2014)

To finish with the general context of the industry addressed in phase 1 of the project, a forecast of the demand was analyzed. Table 2 presents real demand of lubricants on 2008 and 2013 in the United States, for products designed for both commercial and industrial purposes. The estimated demand for 2018 is also presented, and it is compared to the demand of 2013, which is the last year with real values in the table. From Table 2, we can see that the demand in the forecasted period is expected to experience an increase of less than 1% annually, but at least it shows a change in the trend seen from 2008 to 2013. This slight growth encourages the Company even more to improve its current distribution network, which is composed of eight DCs, since tough competition requires an efficient network to better meet customer expectations. Moreover, optimizing its supply chain will lead to cost reductions, which gives another advantage with respect to competitors, since demand can be boosted by the competitive price for its products. Finally, the environmental impact, a variable that is becoming increasingly important for customers, can be reduced by selecting the optimal mode of transportation.

1.2.2. Objectives and Key Decisions

The second phase of the project took place in early 2014, when the Company evaluated the relocation of its DCs that currently cover the American territory. During that phase, many possible locations for new DCs were assessed. Some attributes were taken into account, such as accessibility from the plant, location of customers, modes of transportation available and ease of building or renting suitable facilities for the industry. Quantitative aspects were also analyzed to finally define a narrow set of candidate locations that matched all the requirements of the Company. Identifying the optimal location of facilities among the candidates arises as the first main decision of this study. Regardless of their location, all the DCs have to be supplied from the unique lubricants plant located in Ontario. Unlike DCs, the plant cannot be relocated, thus is a fixed starting point in the supply chain. According to concepts presented in Section 1.1, in this case the plant represents the value offering point (VOP), and after this point, products do not undergo any transformation until reaching the final customer. As the case focuses on the distribution network of lubricants, we can say that we are coping with an outbound logistics problem. Relocating facilities entails a new assignment of customer to DCs. This means defining the DC responsible for delivering the products to each customer, once a shipping order is received. The answer to this question implies not only identifying the assignment of customers to DCs, but also the products that should be received, stored and dispatched from every DC. Furthermore, modes of transportation from plant to DCs and from DCs to customers must be selected among the possible options. These decisions related to the assignment of customers, assignment of products and selection of transportation modes set boundaries on the product flow in the network. Identifying the optimal flow of products through the supply chain becomes the second main decision of this study. Finally, the third decision of the case requires knowing the quantity of product flowing through the DCs every year in order to determine the volume needed in stock. This volume will allow the Company to perform calculations to set the size of the facilities to build or rent for the next ten years.

Figure 5: Plant and Potential Locations for New DCs

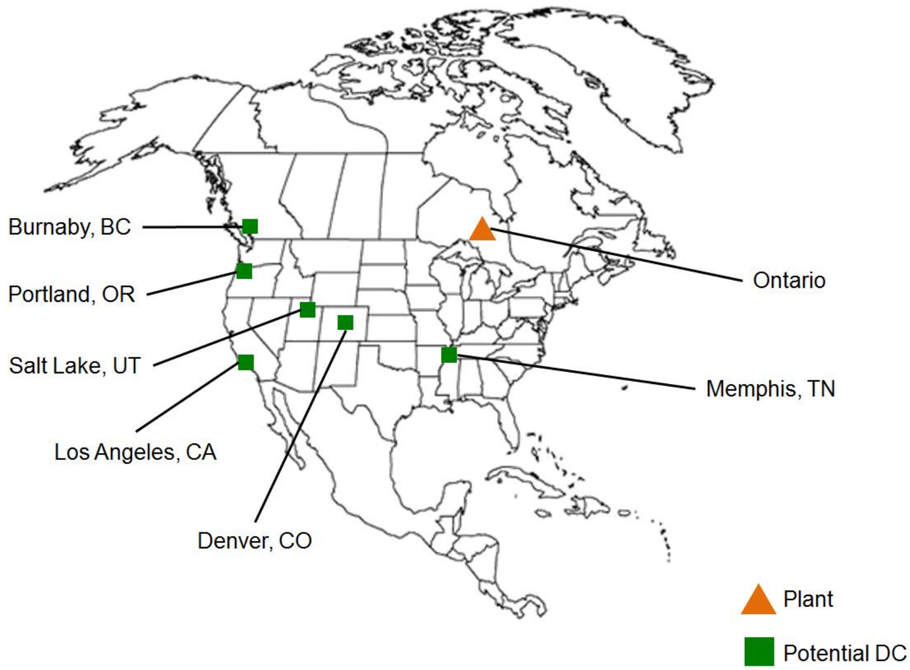


Figure 5 presents the location of the lubricants plant site in Ontario, the first point of the supply chain in this case. Additionally, the set of potential locations for future DCs is indicated. The geographical coordinates of these points as well as customer locations were given in order to have an accurate measure of distances within the supply chain. It is worth mentioning that geographical distances are just initial data that need to be turned into cost, because the optimal solution has to provide the configuration of the network that satisfies customer demand at the minimal cost.

Obtaining the configuration of the network involves determining the three main decisions stated before. These can be summarized as selecting the location of the DCs, identifying the optimal flow of products through the network, and defining the quantities stored in every DC. To the Company, having this level of detail is one of its objectives for the phase 3 of the project. From this, the Company will be able to design an implementation plan, which completes this phase. Phase 3 will be completed in 2016, to be continued with the implementation and monitoring, which is the last phase of the project, expected to end in 2017.

Figure 6: Network Structure to Optimize

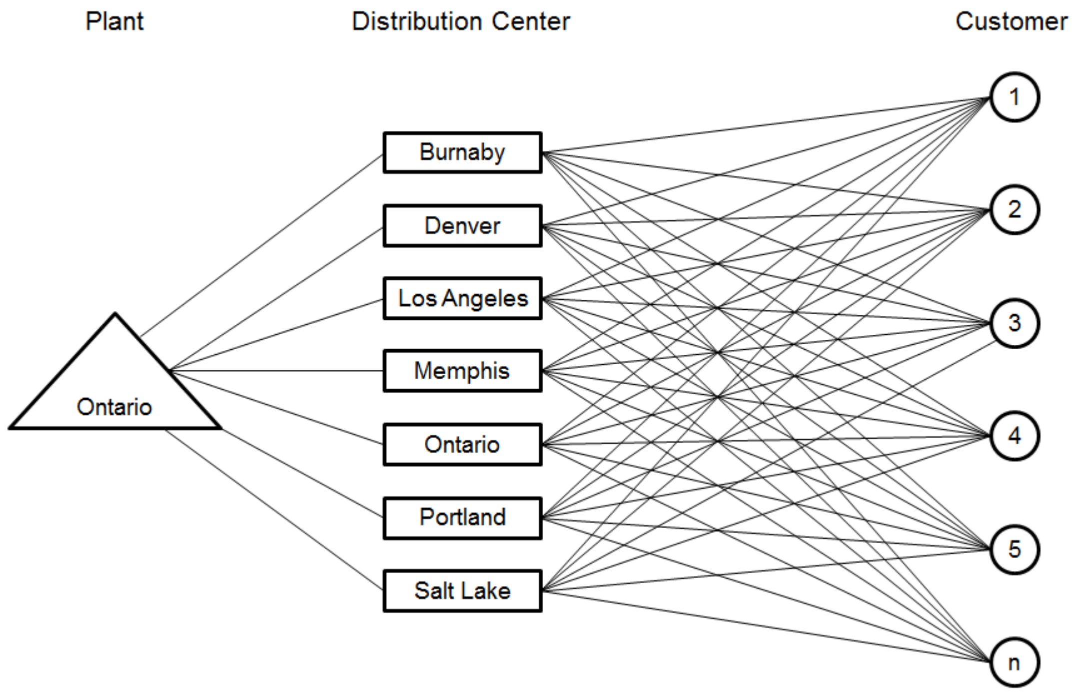


Figure 6 presents the network structure addressed in this case, with a graphical representation of the supply chain to optimize. Here, we can see all the possible paths a product can follow to reach the final customer. The solution of this problem involves identifying the least costly route to reach every customer while respecting the constraints. One way to do this is to evaluate all the possible routes in terms of costs for all the products demanded. In our case, we cope with scenarios holding more than 800 customers and more than 600 stock keeping units (SKUs). Every SKU has its own identification code and needs to be treated as an independent product. Moreover, when multiple modes of transportation exist between two points, each mode has to be evaluated to guarantee the optimality of the solution. The resulting number of combinations, in turn, increases with the number of years to consider in the problem. Before presenting the costs involved in this supply chain, it is important to go over the specifications of the products to distribute, and their main characteristics.

1.2.3. Type of Products and Modes of Transportation

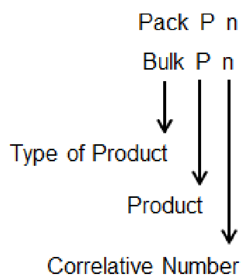
Two types of products have to be distributed by the resulting network: bulk products, and packaged products. Bulk refers to products manipulated in large volumes and without any sort of packaging process. These products are stored in independent tanks and transported by truck tanks or rail tank cars. Regarding the tanks, their size usually ranges between 9 000 USG[1] and 40 000 USG. Once a tank is dedicated to store one product, only that product can be put inside to avoid contamination. As for modes of transportation, there are two options to include in the supply chain depending on the origin and destination of the products: to ship products from the plant to any DC, the mode of transportation is by railway, with a maximum capacity of 23 483 USG per trip, and to deliver the products to a customer from any DC, the mode of transportation is by truck, with a full truck load capacity of 6 193 USG per trip.

Packaged products are available in a wide range of small packs ready for consumption, or in steel and plastic drums. They can be grouped in pallets for ease of transportation by forklift truck, and stored in common warehouses. Pallets in turn can be loaded into regular trucks and trains. With respect to the modes of transportation, there are three options to include in the supply chain: rail, truck, and a combination of both, known as intermodal. Similar to bulk products, packaged products travel by rail from the plant to any DC, and from any DC to the customer by truck. However, the intermodal option is available for the trip from the production plant to Los Angeles, and from the production plant to Denver. Regardless of the product or mode of transportation, packaged products have to be moved in pallets, which have a maximum capacity of 1 800 lbs. Although there are many package sizes and presentations, they can be consolidated to optimize the cubic volume of the pallet. Maximum product accumulation level and security specifications have to be respected in all the cases.

[1] U.S Gallon = 3.79 Liters

For the sake of simplicity, the names of the products are presented with a short structure in this thesis. It allows differentiating the type of product and the total number of products, which is 99 for bulk products, and 613 for packaged products. Table 3 presents the structure of the names used in the coming sections.

Table 3: Types and Names of Products



Bulk Products	Packaged Products
BulkP1	PackP1
BulkP2	PackP2
BulkP3	PackP3
BulkP4	PackP4
BulkP5	PackP5
…	…
…	…
BulkP99	PackP613

After understanding the concept of bulk and packaged products, we can see that the case might easily be extended to other industries distributing these two types of products. In general, bulk products are those that have completed all processing stages, but not including, final packaging. As they cannot be measured by eye, physical measurements such as volume and weight are needed to define quantities of the product. Storage is usually made in large quantities and special facilities. Transportation is ensured by large vehicles built to fit the specific requirements of the product. Moreover, bulk product carriers are limited to transport one bulk product during the haul to meet health and safety regulations issued by regulatory authorities. Examples of bulk products can be found in different industries. Bulk food products are those of the commodity grain market such as corn, soybeans, canola, and rice, among others. Bulk chemicals are usually liquids as ethylene, propylene, acetone and others used for the production in the chemical industry. In the construction industry, products such as sand, gravel, concrete, and coal are also examples of bulk products.

In packaged products, storage facilities and trucks can be filled with several different products that can be inventoried as individual units. Examples of these products are appliances or furniture. Other items are distributed and sold in the packaging unit in which they are handled. For example, paper is managed in reams and in boxes to make storing and handling more convenient. Overall, packaged products can be hand counted, which makes the main difference with respect to bulk products. Many different industries can distribute products in this category, which additionally allows mixing cargo to increase the efficiency of the truck.

By considering the broad concept of bulk and packaged product, the rest of this thesis could be seen as a general bulk and packaged products case, rather than a specific lubricants industry case.

1.2.4. Operational Costs and Constraints

As mentioned before, the optimal solution of this case has to provide the configuration of the network that satisfies customer demand at the minimal cost. Therefore, it is imperative to mention all the costs that are involved in the operation of the supply chain to be optimized in this case, and the constraints that the solution must satisfy. First, the cost structure for both bulk and packaged products is listed as follows:

- **Manufacturing cost:** The total cost of producing one USG of the product. It includes all the manufacturing costs of the plant in Ontario.
- **Transportation cost from the plant to DCs:** The total cost of preparing and moving one USG from the plant to the DC. It includes the freight rate of the carrier and the fuel surcharge (FSC), a charge that was designed to allow transportation companies to adjust to the uncertain prices of fuel.

- **Transportation cost from DCs to customers:** The total cost of preparing and moving one USG from the DC to the customer. The type of truck depends on the type of product. Some customers within a distance of 160 miles from the DC may not require delivery.
- **Handling cost:** The cost of receiving one USG of product in the DC and preparing a shipment for the customer. It includes movements of products and pick-up processes.
- **Holding cost:** The cost of storing and maintaining one USG of product in the DC for a period of time.
- **Tank preparation cost:** Initial fixed cost in period 1 for opening a tank for bulk products. There is no fixed cost for packaged products as per the information provided by the Company.

Regarding the constraints, the Company wants to make sure that customer demand is satisfied for the next ten years. Capacities of facilities and modes of transportation cannot be exceeded. Additionally, minimal quantities to store or transport are provided by the Company according to agreements with third-party logistics providers, infrastructure restrictions or regulatory compliance.

1.3. Problem Statement

In order to better state the aims of this thesis, the objectives are described as follows:

- Preparing and validating the data to represent the logistics network required by the Company to distribute lubricants throughout the United States.
- Adapting the existing optimization model to the specific case. This includes the following steps: presenting the values to assign to the needed parameters, accommodating the data into the optimization model, and selecting the

constraints of the model to ensure the accurate representation of the supply chain.

- Presenting and analyzing the results yielded by the model. This includes solutions for the following decisions: selecting the location of DCs, identifying the optimal assignment of customers to DCs, the assignment of both bulk and packaged products to DCs, determining the optimal transportation mode, and defining the quantity of product to store in every DC in order to estimate the size of the facilities.

1.4. Thesis Outline

After the introductory part of this thesis, the thesis is structured as follows: an overview of the most relevant literature related to the case is presented in Chapter 2, a description of the original mathematical model, of the variables, and of the possible extensions is given in Chapter 3. The preparation of the model to the case and cost structure is presented in Chapter 4. The accommodation and transformation of the data into the model is presented in Chapter 5. The results and the analysis of the numerical experiments are provided in Chapter 6, finally, the conclusion and future research directions are presented in Chapter 7.

2. Literature Review

This chapter presents a brief overview of the literature and previous research related to logistics network design problems. The objective is to make reference to the main contributions that allow a better understanding of the problem addressed in this thesis. This literature review provides a starting point for representing the logistics network of the lubricants company, and solving the problem previously described. Finally, the scarce literature about cases coping with bulk products, and attempts to solve problems in the lubricant industry or similar industries is discussed.

2.1. Supply Chain Network Design

In order to enhance customer satisfaction and control costs in the business, an efficient and effective supply chain network is needed. Integration of suppliers, manufacturers, distributors, and customers are essential for the efficient management of the supply chain. Planning activities have to ensure collaboration between all the entities in the supply chain, including customers, in order to meet their demand and increase the value of products and services offered. To achieve this main objective of the supply chain, a proper network design is required while costs and quantities in the entire network are optimized (Sha & Che, 2006).

The first efforts to optimize and design distribution systems attempted to identify the optimal warehouse locations by minimizing the transportation cost from the warehouse to the customer. Many contributions, extensions, and modifications have been proposed to this classical facility location problem.

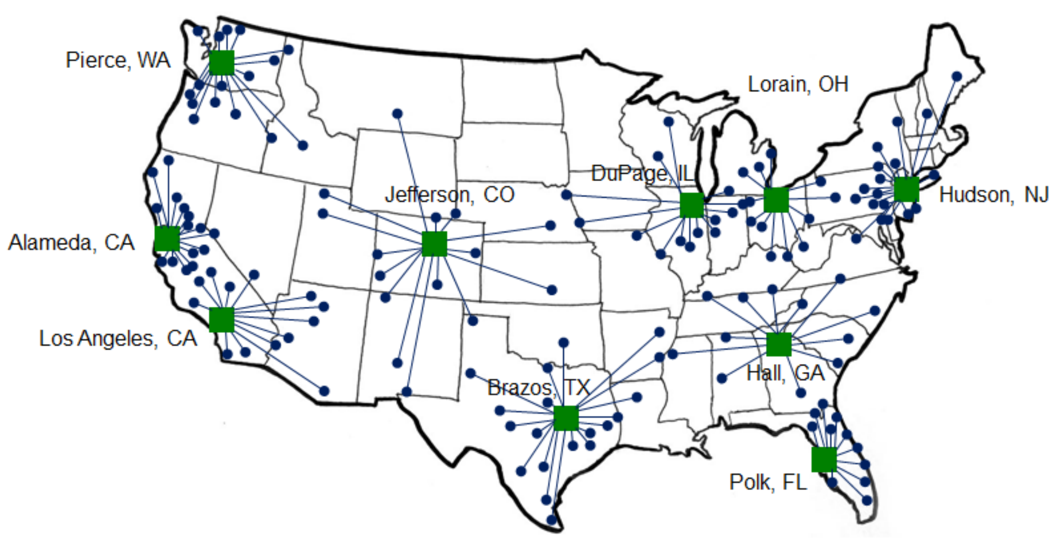
2.1.1. Facility Location

The first broad concept to address is related to facility location problems, which aim at finding the optimal location to carry out the purpose of the facility. Some applications are related to the location of fire stations, schools, waste disposal facilities or hospitals among a list of multiple real situations that require a solution (Daskin, 2008). In more typical situations, the objective is to serve a group of customers with certain demand. The basic problem consists in locating a single facility that minimizes the total travel distance to reach the customers in a continuous solution space (Weber, 1929). This is known as the Weber problem due to its author, who is considered as one of the pioneers of location modeling (Daskin, 2008). The problem becomes more complex when the number of facilities increases from one to several.

With many different types of problems and solution methods, a classification of possible situations arose as taxonomy of location models. Daskin (2008) presents a classification of models based on geographical assumptions about where the facilities can be located. The classification contains four types of models. The first one is the group of analytical models, which assume that customer demand is dispersed uniformly throughout an area and facilities can be placed anywhere inside the area. These models are the simplest but are not necessarily realistic. However, they provide the basis for the second category: continuous models. In this case, the models assume that the demand occurs at some specific points (nodes) in a region, thus lines (edges) connecting a single facility to customers can be drawn, and the solution is the location that minimizes the distance between the facility and the demand nodes. However, as analytical models, the facilities can be located anywhere in the plane. Even though this category is closer to real market situations, the solution is still unrealistic due to the lack of constraints on the locations. The third type consists of network models. They assume that both demand and facilities are located only on nodes or links of the network; thus the optimal solution is the location permitting the minimal distance between the facility and the customers. This represents an even more realistic situation. Edges are equivalent to the routes available to join a couple of customers between themselves or with the facility. Finally, the last category and most relevant for our case is the discrete models, where demand is present on the nodes and, unlike the third type of models, facilities can be selected only from a pre-defined set of potential locations (Daskin, 2008).

Among discrete models, we find the *p*-median model (Hakimi, 1964), where *p* facilities must be located to serve all the customers and minimize the total weighted distance between facilities and customers. The weight of a customer usually corresponds to its demand. Due to this weighted sum, the solution is likely to propose locations near the nodes with high demand. With respect to the constraints of the model, they guarantee that every customer is assigned to one potential facility, that the facility is open if customers are assigned to it, and that exactly *p* facilities are selected. Its solution, according to the objective function and constraints mentioned above, often looks like the one presented in Figure 7.

Figure 7: Example of a Solution to a *p*-Median Problem with *p* = 10



Source: (Daskin, 2008)

Although the *p*-median model provides a rational solution in several real contexts, it ignores that the fixed costs of opening and operating each location can be different. When including this aspect, a new problem arises: the problem is called the Uncapacitated Fixed Charge Location Problem (UFLP) (Balinsky, 1965). This problem minimizes the sum of the fixed cost and the variable transportation cost. This is done by simply adding a fixed cost for every facility that is open to the objective function. Moreover, a cost per mile can be added to the original formulation of the *p*-median problem which considers only distances. The constraints, also similar to those of the *p*-median model, ensure that every customer is assigned to one facility, and that the facility is open if customers are assigned to it. The number of facilities to open is usually not restricted in this case as the objective is to minimize the total cost.

2.1.2. Configuration of the Network

After the broad concept of facility location problems, the next step in supply chain design is the concept of network configuration. Apart from defining the optimal location of warehouses to serve customers, the logistics network considers a wider number of decisions. These decisions usually involve the number and capacity of plants and warehouses, the assignment of products to plants, the selection of suppliers and transportation modes, and the flow of products through the network. A first effort to design and optimize a distribution network was proposed by Geoffrion & Graves (1974). Their contribution aimed to represent a distribution system formed by 14 plants, 45 possible distribution centers (DCs), and 121 customer zones. Moreover, the linear programming model managed 17 commodities, which are the equivalent of finished products. The DCs to open are a decision variable to be selected by the model while the objective function minimizes the total cost. The optimal assignment of customers to DCs is also part of the solution, which determines the quantity to ship of every product from the plant to the DCs, and from the DCs to customers. The total cost results from the sum of fixed and variable costs. A fixed cost is incurred when a DC is opened. Variable costs are calculated according to the quantity of units produced at every plant, plus the cost of transporting the commodity to the DC and then to the customer. The constraints of this model ensure that customer demand is satisfied, that capacity of plants and DCs are respected, and that each customer is assigned to one DC (Geoffrion & Graves, 1974).

A couple of relevant features of the work described above are worth mentioning since they are still valid today. Firstly, the authors allow the case where commodities are shipped to customers directly from the plants rather than from DCs. This is done by adding a fictitious DC where the cost is adjusted so that only the cost of the real trip, i.e. from the plant to the customer, is accounted for. Therefore, the fixed cost and transportation cost from the plant to this DC is set to 0, and the transportation cost from this fictitious DC to the customer is now set to original transportation cost from the plant to the customer. Secondly, the model does not allow any customer to be served from more than one DC or one plant. This single sourcing constraint is imposed with binary variables that represent the assignment of customers to DCs.

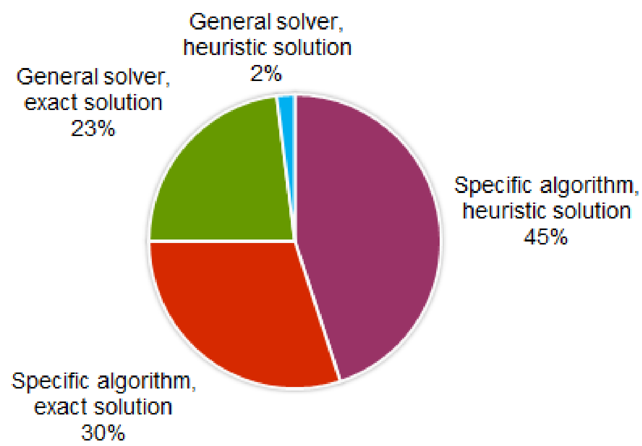
An extension of the distribution model presented before was proposed by Pirkul & Jayaraman (1996) where plants, and not only the DCs, were included as decision variables. Here, plants to be opened are selected among a set of candidates plants with different capacities, and each plant has a fixed cost for establishing and operating the plant. Therefore, the total cost to be minimized includes fixed cost for both plants and DCs, and all the variable costs to distribute products through the network. The authors claim that including plants in the problem benefits productivity in manufacturing activities since the results of the model can be used as an input for the manufacturing plan. This is possible because the solution of the problem includes the units to produce for every product at every plant. Moreover, this production plan ensures that total production of all the plants will satisfy customer’s demand, while previous models met this constraint only from the DCs (Pirkul & Jayaraman, 1996).

Going backward in the logistics network, the missing link to complete the chain is related to the selection of suppliers responsible for the procurement of raw materials to the plant. An integrated model that addresses procurement, manufacturing and distribution was presented by Cohen & Lee (1988). With regards to procurement and manufacturing, the two stages that make their contribution relevant, the authors consider the transformation of raw material into finished products. To make this transformation possible, a set of suppliers has to provide the items required for the bill of materials (BOM) defined for each product. The items must be in stock to pass through the production process, resulting in another stock of finished products. Multiple stocks at every stage lead to an inventory control system that impacts the entire network. The model considers also the cost of holding and handling inventory, as well as the shortage cost related to delays in the manufacturing stage (Cohen & Lee, 1988).

After the emergence of the key concept of integration, much literature has been written about supply chain network design. An effort to aggregate and analyze a huge portion of literature as a whole was done by Melo, Nickel & Saldanha-da-Gama (2009). The authors selected articles published between 1998 and 2008 to be analyzed. The selection criteria led to articles integrating multiple stages in the supply chain and excluding simple location facility articles. Moreover, only discrete models were taken into account, and finally a relevant research contribution in logistics was necessary to be in the scope. A resulting list of 98 articles from 19 journals was sorted and classified in order to obtain global conclusions. Of this list, 56 articles were published in 2004 or later and 44 articles were issued in the European Journal of Operational Research. A classification of optimization models showed that in the objective function 75% aim to minimize the costs involved in the model, 16% maximize profits, and 9% can be used for either minimizing or maximizing purposes (Melo et al., 2009). Under the profit maximization criterion, satisfying the demand might not be compulsory for the company. Companies could be willing to lose customers in order to increase profits, while satisfying the demand is a constraint in cost optimization models. Regarding the method used to solve the problems with a single objective, the literature was sorted in two main categories: specific algorithm and general solver. The former refers to algorithms tailored to solve problems with multiple objectives. The latter is based on mathematical optimization software that uses mixed integer programming models to solve a given formulation. Each category can solve a problem to optimality, which is classified as an exact solution, or in a heuristic way, which means that an optimal solution cannot be guaranteed.

According to the classification done by Melo et al. (2009), the category “specific algorithm, heuristic solution” is present in most of the literature in the study. This method is used to solve more complex problems with many variables in the objective function. Variables are related to the decisions to be made in the different layers of the logistics network. Although the heuristic method does not guarantee optimality, it provides a satisfying result for most common logistic network design problems. The category “specific algorithm, exact solution” is used in 30% of the literature. It includes solution techniques such as branch-and-bound and branch-and-cut, in which optimality can be reached. The third category is “general solver, exact solution”, where the software can find either, an optimal solution or a near optimal solution, for which a gap is defined in advance. In the “general solver, heuristic solution” the optimization software keeps running for the time specified. When the software stops the best solution found at that moment is taken as the result. The division of these categories among the literature analyzed by the authors is summarized in Figure 8.

Figure 8: Solution Methods for Supply Chain Network Problems



Source: (Melo et al., 2009)

The model we used in this thesis can be classified in the category general solver, exact solution. For our case, the optimality tolerance was set in 1%, which means that the mathematical solver is required to find a feasible solution within 1% of optimality. The model also presents an option to set the time limit to solve the problem. By setting this parameter, the model could be executed as a heuristic.

From the literature surveyed by Melo et al. (2009), the papers more relevant for our case are discussed next in further detail. Even if none of the articles deals with a lubricants industry case or similar, many of them present contributions in terms of solution approach and methodology.

A real-world case about the relocation of warehousing facilities is presented by Min & Melachrinoudis (1999). In the case, an American manufacturer of aluminum bars wants to relocate its current distribution facility from Boston to one of ten proposed sites. In the paper, the name of the manufacturer is disguised with the name “Alpha”, a request from the managers in order to avoid any impact on their employees’ morale. Alpha’s candidate locations are cities in the states of Maryland, Pennsylvania, and West Virginia, places that allow a good level of transit time to customers. Locations have also the land space needed and the suppliers to build the facility. Proximity to intermodal transportation, tax incentive packages, and access to skilled workers were attributes of the ten locations. The problem for Alpha is to select the location that maximizes long term profitability. This involves analyzing the trade-offs among all the variables, and evaluating the sensitivity of relocation decisions when priorities change. The authors used a decision-aid method known as analytic hierarchy process (AHP). This method required to identify 33 location factors that were aggregated in six categories: site characteristics, cost, traffic access, market opportunity, quality of living, and local incentives. Then, every category received a weight according to its relative importance in the relocation decision, and finally, facilities were ranked by each criterion. Calculations yielded a score per facility resulting in a general ranking of candidate locations. A sensitivity analysis was later performed by changing the weights of categories according to possible changes in priorities. Efficiency and simplicity of the proposal to support facility relocation decisions are the main contributions of this paper.

A mathematical model to optimize a multi-product multi-echelon distribution system was proposed by Lin et al. (2006). Here, the system is composed of plants, consolidation centers (CC), distribution centers (DC), and customers. As multiple products can be produced in independent plants dispersed in a region, the CC consolidates products that are then sent to DC. However, the inventory is held only at DCs. The decisions considered the number and location of CCs and DCs, the inventory level of products at the DCs, and the routing of shipments from plants to customers. The results include also a measure of service level as a percentage of customers covered within a certain time or distance. Due to this service measure, there is a trade-off with the number of DCs, since more facilities can provide a higher service level, which in exchange also results in a higher cost. The objective of the model is to minimize the sum of the transportation costs, the fixed costs for opening a CC or a DC, the inventory costs, and the penalty costs for customers out of the range of service. To define the inventory level, the authors set a constraint that ensures that for each product the DC holds enough inventory to avoid a stock out during the replenishment time. Stocking out is defined as the probability of having a higher demand than the amount in inventory during the needed lead time to replenish the DC. This integrated view of transportation cost, inventory cost and level of service, is the main contribution of the paper.

Another model that combines location decisions with inventory and flow variables was presented by Syam (2002). With a supply chain structure more similar to our case, the author copes with a case of three factories, three warehouses, and two destinations. The objective is to identify the optimal locations to be opened and the optimal paths to reach the final destination at the minimal distribution cost. The objective function includes the inventory ordering cost, holding cost, handling cost, transportation cost and manufacturing cost. There is also a fixed cost for opening a facility. The ordering cost is incurred every time a movement of products between facilities is done. A variable defined as “cycle time” represents the frequency of shipments per product between two facilities. Then, the holding cost is computed as a percentage of the value of the products in the warehouses. The percentage can be set differently for each warehouse. Finally, the handling cost results from a unit handling cost per product and per warehouse times the quantity of the product flowing through the warehouse. The focus on inventory costs by the author is the main relevant aspect for our case.

A production and distribution network with suppliers, plants, and warehouses is optimized with a combined model proposed by Jang et al. (2002). Decisions about the selection of suppliers, location of plants, and warehouses were addressed with three fragmented models. The first model optimizes the transportation cost for outbound logistics, which connects warehouses and customers. Its results are used in the second model, which focuses on inbound logistics. This model minimizes the cost of operating plants and transportation of raw material from suppliers to plants. Here, a bill of materials (BOM) is incorporated; therefore, a product can be produced by assembling multiple products in specific quantities.

Finally, a third model is composed of the two initial sub-problems, with more capacity constraints. This yields a better solution by removing the most expensive warehouse from the initial solution. The value of this paper is the use of fragmented models, which integrates production and distribution sub-problems that are difficult to solve simultaneously due to the number of integer and binary variables.

An aspect that was not considered in the four previous papers that have been discussed is the possibility of optimizing a supply chain with multiple periods. Melo et al. (2006) proposed a model to cope with the relocation of facilities over a long-term planning horizon with fluctuating demand. The model solves many decision variables simultaneously related to the location and capacities of warehouses. As the authors claim that network planning decisions should always consider several time periods, they focus their attention on decisions that can be influenced by future demand. The model offers the possibility of dynamic facilities than can be relocated from one period to another. It also considers capacity expansion or reduction for warehouses and the option to close or open a new location among a set of candidates. As the total cost to minimize is based on this flexible attribute, the model includes not only transportation and holding cost, but also fixed maintenance cost, relocation cost, installation cost, and shutdown cost. Installation costs, which are charged the period prior to the setup, account for property acquisitions, facility constructions, preparation of new infrastructure and traffic access. All these dynamic aspects appear when dealing with a multi-period problem, which are relevant for our case which considers a ten-year horizon.

Another aspect for which little literature is available is the use of multiple modes of transportation. Eskigun et al. (2005) proposed a model that considers the location of warehouses, lead times and transportation modes. The objective is to minimize the transportation cost, lead time related cost, and fixed cost of opening a facility. The outbound supply chain considers rail to send products from the plant to warehouses, and trucks to ship products from warehouses to customers. This transportation structure is similar to the one of our case. Additionally, in both cases, the products can be dispatched to customers from the plant by truck when profitable. The combination of the two modes of transportation, rail and truck, has a direct impact on the location of warehouses due to cost differences. As rail is usually cheaper, it is optimal to come as close to the customer as possible by this mean, and travel short distances by truck. However, for customers located between the plant and the warehouse, the solution might result in travelling first to the warehouse and then to the customer. This would be a longer distance but would also be a cheaper alternative. In the model, this trade-off between cost and lead time is optimized.

2.2. Research Background in Similar Cases

After having reviewed some of the abundant literature on network design, we now move on to discuss some of the scarce work done in supply chain optimization with bulk products. Literature pertaining to the lubricants industry and similar industries is then presented.

2.2.1. Proposals with Bulk Products

As mentioned in Section 1.2.3, our case might be extended to other industries with bulk and packaged products. Here, we make reference to literature with proposals to solve problems that deals with bulk products.

Huang et al. (2010) presented a proposal to minimize the cost of a future supply chain of bioethanol, an energy source obtained from the distillation of corn, rice, and wheat, among other biomasses. The case study considers eight candidate feedstock field locations in California. The aim was to identify the potential requirement of infrastructure to serve the state demand, which is aggregated by cities. Moreover, depending on the cost of operating the supply chain, the distributor could know if the gallon of bioethanol might be offered at a competitive cost. This aspect is relevant to define the feasibility of the project to reduce oil dependence in the long term. The resulting supply chain configuration provides location of feedstock, refineries, and storage centers to efficiently serve the cities. The case deals with two types of bulk products: first, bulk solids that have to go from the feedstock fields to the refineries, second, bulk liquids, which is the ethanol leaving from the plant until it reaches the final customers. The former are measured in tons and the latter in gallons. Every product requires a different truck, which entails different capacities, cost of transportation, and cost of loading and unloading. Besides all these variables that are included in the objective function, the long term is also considered in the model through the forecasted demand and future procurement of raw materials. When demand of ethanol exceeds the supply of the feedstock in one period, the ethanol has to be imported from other states. The multi-period attribute ensures that the cost is minimized for the entire planning horizon. With the resulting supply chain of this model, the authors showed that with the optimal configuration, the cost of producing and distributing one gallon of ethanol could be as low as $1.1 in the midterm. This cost makes the project feasible, a goal that according to the author is only possible with the optimization of the entire supply chain over the whole planning horizon.

Another case in literature is related to BP's strategy to introduce liquid hydrogen as a clean source of energy worldwide. Hugo et al. (2005) presented a model to evaluate the cost, operability, reliability, and environmental impacts of this alternative. The model provides the optimal solution to satisfy the forecasted demand with the minimal infrastructure investment level in the long term horizon. The model optimizes the trade-off between the investment and the impact on the environment. The distribution network is essential to achieve the goal of the project. Gaseous hydrogen can be distributed in pipelines, while liquefied hydrogen can be transported by ship, rail, or truck tanks. As it is for every bulk product, transportation modes and storage facilities require specific conditions. This means that investments in new infrastructure across the supply chain are the main criteria to consider the feasibility of a project of this magnitude. The case deals with hydrogen demand forecast from 2004 to 2038 for regions. Three stages are considered during this period of time: market introduction, wider commercialization, and maturity. The model selects the optimal configuration of the network for the long-term horizon, which includes the amount of liquid hydrogen to be shipped by truck, and the gaseous hydrogen moved by pipelines at every period.

2.2.2. Proposals in the Lubricants and Oil Industry

Regarding the lubricants industry itself, Fanchao et al. (2009) presented a model to optimize inventory lubricants costs while considering the level of service. The trade-off between costs and level of service is achieved by a what-if analysis. The model allows to represent a complete supply chain with suppliers, plants and warehouses, with all the costs related to hold inventory. Then, after setting up the variables for likely stochastic situations, the software runs multiple simulations for predefined scenarios, and yields the average resulting cost of each scenario. Finally, the model selects the network configuration with the minimal average cost while satisfying the desired level of service, which is measured as the fill rate of customers' orders. Although the model neither optimizes the entire supply chain nor supports location facility decisions, it provides a good approach to make decisions about inventory management such as inventory level and safety stock. The contribution to our case is the particular variables of the industry it takes into account. In order to simulate a real system, the model allows setting parameters as number of plants, cost factors, plant efficiency, assignment policies and scheduling policies. Moreover, since dealing with considerable-volume materials with slight differences among themselves, the customer orders can be defined by the types of products, the grades, the packaging, and the delivery windows. Products that customers receive out of the time window can have a penalty. Finally, as lubricants require expensive and sophisticated transportation equipment and storage facilities, the impact of technology can be included in the tool to be simulated.

Another approach that considers supply chain optimization in a similar industry is presented by Lasschuit & Thijssen (2004). The authors presented a tool based on a mixed-integer non-linear programming (MINLP) that supports decisions related to supply, manufacturing, and distribution in the oil industry. The novelty argued in the paper is that the model is designed for an industry where there are more products than raw materials, and prices are highly volatile. To manage this aspect a flexible price can be assigned to each product with an elasticity function. The objective of the model is to minimize the distribution cost by optimizing manufacturing volumes, rationalizing plants, warehouses and transportation modes. The costs included are transportation, handling, and holding costs among other variable costs. Fixed costs of processing lines, operating warehouses, and initial investments to start a facility are also taken into account. Although many useful aspects for our case are presented in the paper, the authors focus their attention on presenting the attributes and benefits of their tool to users, rather than explaining the mathematical model providing all those virtues. This fact can be attributed to the commercial purposes of the tool, rather than a research driven contribution.

3. General Optimization Model

This chapter describes the original model used to represent and optimize the logistics network of the Company. First, we comment the origin of the model. Then, the mathematical formulation is explained, and finally, a set of possible extensions to the original model are presented. This model is based on the work of Cordeau et al. (2006) and was later extended by Cordeau (2014).

3.1. Origin and Conception of the Model

Cordeau et al. (2006) introduced a flexible formulation for the logistics network design problem (LNDP). The model addresses a deterministic context and can solve single-country and single-period problems.

The flexible formulation presented in the paper allows to define the quantity, location, and capacity of suppliers, plants and warehouses. Also, multiple transportation modes with their own capacity can be used to link the points. Additionally, from the set of products flowing through the network, a range can be assigned to each facility and transportation mode. This functionality allows to represent that a specific raw material can be provided by only a set of suppliers, moved using only some transportation modes and stored at only some warehouses. The same applies to finished products that can be manufactured at only some plants or distributed to customers in a limited area. As the minimal size of a set can be one, a range of products could be forced to flow through a particular point or mode of transportation. All these variables make it possible to represent virtually any logistics network in order to optimize the total cost by using the model presented by the authors. Moreover, the model can support decisions related to the improvement of a logistics network. In their paper Cordeau et al. (2006) classify those decisions in three categories that can be summarized as follows:

Category	Decisions Involved
Strategic	Choices about the number, location and capacity of manufacturing plants and warehouses as well as the technology to use.
Tactical (Medium-Term)	Selection of suppliers, distribution channels and modes of transportation as well as assignment of products to plants and warehouses.
Operational (Short-Term)	Selection of raw materials, semi-finished and finished products flowing through the network.

Source: Summarized from Cordeau et al. (2006)

The LNDP provides a solution for all the categories while minimizing the sum of both fixed and variable costs needed to operate the logistics network. Satisfying demand is imposed as a constraint in the model to ensure useful results. The formulation considers the stages of procurement, production, warehousing and distribution simultaneously. Due to the integration of the solution, larger benefits are yielded for the entire network compared to the results when every stage is solved independently (Cordeau et al. 2006).

Cordeau (2014) extended the formulation of the LNDP to include more possible configurations, and more possible constraints. Firstly, among the new variables, the most relevant one refers to the amount of inventory to have in the warehouses at the end of each period. This variable quantity multiplied by the unit holding cost of the product is present in the objective function, leading to consider the inventory cost as a new tactical decision. Secondly, regarding the new possible configurations, the model includes the value of the gross revenue of selling the products to the customers. This allows the objective function to maximize profit instead of minimizing cost. Lastly, a limit on CO_2 emissions produced by the entire network can be imposed. This environmental constraint meets recent trends in logistics and anticipates coming governmental regulations, since maximum amounts of gas emissions can be adjusted for every future period in the model.

Among the improvements in the more recent model compared to the initial formulation of the LNDP, there are two aspects that are relevant to solve the case of the Company. First, the inventory available at the end of each period permits to create the link with the next period. This attribute makes the model a multi-period one instead of a single-period model. A key condition for our case is that we must consider a ten-year planning horizon. Second, bounds can be imposed on inventory levels for each individual product and also for a set of products that are stored together in the same warehouse. In our case, bulk products are stored individually in big tanks, while packaged products are stored on traditional pallets inside the same facility.

Due to the features mentioned above, the model of Cordeau (2014) is the one used in this thesis to represent the logistics network of the Company and to provide a solution to the points mentioned in Section 1.3. The detailed formulation is presented in the next section.

3.2. Mathematical Model

We now summarize the original model introduced by Cordeau et al. (2006) and extended by Cordeau (2014).

3.2.1. Notation

Different products flow through a supply chain. The raw materials going from suppliers to the plants are aggregated into a subset R. Two or more raw materials assembled in the plant may produce items that belong to subset A, which represents the assemblies. The combination of two or more assemblies, which are intermediate products in the manufacturing process, form another assembly or a final good. The subset of final goods is denoted as F. Additionally, combining assemblies with raw materials can also lead to final products. As final products are the items demanded by customers, they are the products going from the plant to the warehouses and then delivered to customers. The whole set of commodities within the supply chain is denoted as K, thus $K = R \cup A \cup F$.

Among all possible values for $k \in K$, some items $k \in R \cup A$ may be required to make an item $k \in A \cup F$. The subset of items required to make k is $B^k \subseteq R \cup A$. Similarly, for every raw material or assembly $l \in R \cup A$, there is a subset of products that require l *for* their production. This subset, denoted with $K^l \subseteq A \cup F$, allows to define the number of items requiring l in their bill-of-materials. In order to represent the units needed of $l \in B^k$ to produce one unit of $k \in K$, let b^{kl} be the quantity of units of l required to produce one unit of k. The option of having intermediate products in the bill-of-materials of other products makes it possible to recreate multi-level bill-of-materials, which are useful for complex manufacturing process.

Regarding the entities that participate in a supply chain, a notation is assigned to suppliers, plants, warehouses and customer. Let S be the set of suppliers that are in charge of providing raw materials. As only some suppliers can provide certain items, the subset $S^r \subseteq S$ represents the suppliers that can supply raw material $r \in R$ to the plants. The set of plants is denoted as P, where each plant $p \subseteq P$ can produce a set of items. The subset of plants that can produce item $k \in K$ *is* P^k. Alike, the set of warehouses is denoted as W and the subset of warehouses that can store and handle item $k \in K$ *is* W^k. Finally, let C be the set of customers demanding final goods as they are at the end of the supply chain. Every customer $c \in C$ may represent a point of demand with a specific location or a zone of demand if aggregation is considered adequate. Defining the correct level of the size of the zone depends on the optimization purpose, since aggregation might impact the accuracy of results.

In order to determine the feasible flow of commodities within the supply chain, points of origin and destination have to be defined. First, points of origin can be suppliers, plants and warehouses. If O is denoted as the set of points of origin, thus $O = S \cup P \cup W$. Second, the set of points of destination D, can be plants, warehouses and customers, thus $D = P \cup W \cup C$. Once the nodes of the network are defined, the next step requires indicating the subset of origins and destinations of every commodity $k \in K$. The subset $O^k \subseteq O$ and $D^k \subseteq D$ represent potential origins and destinations for item k, respectively. Additionally, the items $k \in K$ that can be received or dispatched from node $o \in O^k \cup D^k$, is represented as subset K_o. The adequate definition of these subsets ensures that items can flow through the nodes and reach their final destination.

As nodes in the network usually have a capacity, upper bounds need to be imposed. The capacity of one node is defined by two variables: the amount that can be received of each item, and the amount that can be dispatched. The output capacity of every node $o \in O$ is denoted as q_o, and u_o^k is the amount of units of capacity required by one unit of item $k \in K_o$ to be dispatched from node o. As the unit of measurement of the node can be different than that of the items, u_o^k may be used as a transformation factor to convert all measures into equivalent units. These variables allow assigning a maximum general capacity to the point of origin. However, a specific capacity for each item $k \in K$ is often required. This upper bound of the units of item k that can be dispatched by point of origin o is denoted as q_o^k. This value limits the output of the item for one node regardless of the destination. For situations where the amount of item k to be dispatched from origin o to destination $d \subseteq D$ must be limited, the value q_{od}^k can be used as a bound.

Once the points of origin and destination are defined as well as their capacity, the transportation modes linking the nodes of the network have to be included. Let M_{od} be the set of modes of transportation available between point o and d. To indicate if a mode of transportation $m \in M_{od}$ is selected among the options to link the two points, the binary variable Z_{od}^{m} takes the value of 1. Many modes m can be used to link two points. Although the selection depends on the cost, the load capacity available for every item might lead to use more than one mode. With respect to the cost, every mode of transportation may consider a fixed and a variable cost. Let c_{od}^{m} be the fixed cost of using mode m from origin o to destination d, while c_{od}^{km} represents the cost of transporting one unit of item k, between o and d. With respect to the capacity, the variable q_{od}^{m} represents the capacity of mode m from the origin to destination indicated, where u^{km} is the capacity required by one unit of item k in mode m. Additionally, $M_{od}^{k} \subseteq M_{od}$ represents the subset of transportation modes available between o and d for item k. Then, as the amount of products to move may change from one period of time to another, every period needs to be defined as $t \in T$. The number of units going from point o to d of item k by mode m at period t is defined as X_{od}^{kmt}. This value multiplied by the unit transportation cost c_{od}^{km} gives the total variable cost for every period. For example, the variable cost of transporting finished product f by mode m at period t from warehouse w to customer c is given by $c_{wc}^{fm}\ X_{wc}^{fmt}$.

When coping with node $w \in W$ it is important to consider that a warehouse is the only location that can hold stock from one period to the next. This entails a final inventory at the end of period t for which a holding cost has to be paid. To represent this, let I_{w}^{kt} be the amount of inventory of item k at the end of period t at location w. Similarly, let g_{w}^{k} denote the cost of holding one unit of item k for one period of time at location w. The inventory ensures the link between period $t\text{-}1$ and $t \in T$. The length of the period impacts the demand of the period. Let a_{c}^{ft} be the number of units of finished product f in period t demanded by customer c.

Finally, other binary variables take the value of 1 when the node is selected, which permits to include the fixed cost for opening a candidate location, which can be a plant or a warehouse. For node $o \in O$ the variable U_o is equal to 1 if the point of origin needs to be open; therefore, the fixed cost denoted as c_o has to be considered. Similarly, for supplier $s \in S$, the variable U_s take the value of 1 if supplier s is selected to be part of the network. Then, to identify the assignment of items to the nodes where they have to be produced or stored, the binary variable V_o^k equals 1 if item $k \in K$ is assigned to node o. The cost related to this assignment is denoted as c_o^k. A final binary variable Y_{od}^k takes the value of 1 if item $k \in K$ is dispatched from point of origin $o \in O^k$ to destination $d \in D^k$. The cost of activating that link between points o and d for the commodity k is denoted as c_{od}^k.

We conclude this section by providing a summary of the notation mentioned before. Table 4 contains the sets within the network, Table 5 the parameters required by the model, and Table 6 the decision variables.

Table 4: Summary of Notation for Sets

A	Set of assemblies
B^k	Set of items needed to make item k
C	Set of customers
C^f	Set of customers that require product f
D	Set of destinations. $D = P \cup W \cup C$
D^k	Set of potential destinations for item k
F	Set of finished products
K	Set of all items. $K = R \cup A \cup F$
K^l	Set of items that require item l
K_o	Set of items that may originate or be destined to o
M_{od}	Set of transportation modes between o and d
M_{od}^k	Set of transportation modes between o and d for item k
O	Set of origins. $O = S \cup P \cup W$
O^k	Set of potential origins for item k
P	Set of potential plant locations

P^k	Set of plant locations where item k can be made or used
R	Set of raw materials
S	Set of potential suppliers
S^r	Set of potential suppliers providing raw material r
T	Set of time periods
W	Set of potential warehouse locations
W^k	Set of warehouse locations where item k can be stored

Table 5: Summary of Notation for Parameters

a_c^{ft}	Demand of customer c for product f in period t
b^{kl}	Amount of item k needed in one unit of item l
c_o	Fixed cost of selecting origin o
c_o^k	Fixed cost of assigning item k to origin o
c_{od}^k	Fixed cost of providing item k to destination d from origin o
c_{od}^m	Fixed cost of using transportation mode m between o and d
c_{od}^{km}	Unit cost for providing item k to d from o with mode m
g_w^k	Cost of holding one unit of item k in node w during one period of time
q_o	Capacity of origin o in equivalent units
q_{od}^m	Capacity of mode m between o and d in equivalent units
q_o^k	Upper limit on the amount of item k shipped from origin o
q_{od}^k	Upper limit on the amount of item k shipped from o to d
u_o^k	Amount of capacity required by one unit of item k at origin o
u^{km}	Amount of capacity required by one unit of item k using mode m

Table 6: Summary of Notation for Decision Variables

I_w^{kt}	Inventory of item k in location w at the end of period t
X_{od}^{kmt}	Amount of item k provided by o to d with mode m in period t
U_o	= 1 if the origin o is selected, 0 otherwise
V_o^k	= 1 if item k is assigned to origin o, 0 otherwise
Y_{od}^k	= 1 if origin o provides item k to destination d, 0 otherwise
Z_{od}^m	= 1 if mode m is selected between o and d, 0 otherwise

3.2.2. Formulation

The original model to optimize the total cost of the supply chain presents the following formulation:

Minimize

$$\sum_{o\in O}\left[c_o U_o + \sum_{d\in D}\sum_{m\in M_{od}} c_{od}^{m} Z_{od}^{m}\right] + \sum_{k\in K}\sum_{o\in O^k}\left[c_o^k V_o^k + \sum_{d\in D^k}\left[c_{od}^k Y_{od}^k + \sum_{m\in M_{od}^k}\sum_{t\in T} c_{od}^{km} X_{od}^{kmt}\right]\right] + \sum_{k\in K}\sum_{w\in W^k}\sum_{t\in T} g_w^k I_w^{kt} \qquad (1)$$

Subject to:

$$\sum_{o\in O}\sum_{m\in M_{op}^l} X_{op}^{lmt} - \sum_{k\in K^l}\sum_{d\in D^k}\sum_{m\in M_{pd}^k} b^{kl} X_{pd}^{kmt} = 0 \qquad l\in R\cup A;\ p\in P^l;\ t\in T \qquad (2)$$

$$\sum_{o\in O^k}\sum_{m\in M_{ow}^k} X_{ow}^{kmt} - \sum_{d\in D^k}\sum_{m\in M_{wd}^k} X_{wd}^{kmt} + I_w^{k,t-1} - I_w^{kt} = 0 \qquad k\in K;\ w\in W^k;\ t\in T \qquad (3)$$

$$\sum_{o\in O^k}\sum_{m\in M_{oc}^k} X_{oc}^{kmt} = a_c^{kt} \qquad k\in K;\ c\in C^k;\ t\in T \qquad (4)$$

$$\sum_{k\in K}\sum_{d\in D^k}\sum_{m\in M_{od}^k} u_o^k X_{od}^{kmt} - q_o U_o \le 0 \qquad o\in O;\ t\in T \qquad (5)$$

$$\sum_{d\in D^k}\sum_{m\in M_{od}^k} X_{od}^{kmt} - q_o^k V_o^k \le 0 \qquad k\in K;\ o\in O^k;\ t\in T \qquad (6)$$

$$\sum_{m\in M_{od}^k} X_{od}^{kmt} - q_{od}^k Y_{od}^k \le 0 \qquad k\in K;\ o\in O^k;\ d\in D^k;\ t\in T \qquad (7)$$

$$\sum_{k\in K} u^{km} X_{od}^{kmt} - q_{od}^m Z_{od}^m \le 0 \qquad o\in O;\ d\in D;\ m\in M_{od};\ t\in T \qquad (8)$$

The objective function (1) minimizes the sum of both fixed and variable costs resulting from operating the supply chain with a specific configuration. The first parenthesis contains the parameters c_o and c_{od}^{m} to calculate the fixed cost of selecting a node in the network, and choosing a mode of transportation. The fixed cost is considered when the binary variable takes value 1, otherwise the cost becomes 0. The next parameters in the function are c_o^k and c_{od}^k , both fixed cost associated to their respective binary variables V_o^k and Y_{od}^k. The fixed costs c_o and c_o^k usually represents land acquisitions, facility building, equipment investment and government fees. Besides, they are assumed to be paid only once in period 1. Then, parameter c_{od}^{km} is associated with a continuous variable to yield the cost of moving items between two nodes. This parameter represents the variable cost per unit, as the total value depends on the amount of items flowing in the network. The cost c_{od}^{km} is not limited to transportation cost. It may also include the cost of raw materials when dealing with suppliers, the production cost when the origin is the plant, and handling cost when the warehouse delivers to customers. Finally, the variable cost g_w^k is also associated to a continuous variable. The result indicates the cost of carrying inventory in the warehouse, such as capital cost, storage space cost, inventory service cost, and inventory risk cost.

Constraint (2) is the first of the three constraints that ensure the adequate flow of items through the network. This constraint guarantees that the amount of items received in one plant from all possible sources is equal to the amount required by the plant to meet its production. Thanks to this, the items flow from supplier to plant in required quantities and to the right destination.

Constraint (3) ensures that the amount of items received in the warehouse plus the initial amount in inventory at the beginning of the period is equal to the amount leaving the warehouse plus the final inventory at the end of the period. This constraint ensures consistency in the inventory of the warehouse and guarantees the flow of items through the warehouse. Moreover, the fact that the final inventory of one period is equal to the initial inventory of the next period enforces the link between the periods. This link allows considering as many periods as needed for the network.

Constraint (4) ensures that the total amount of items delivered to a customer is equal to the demanded amount. This demand constraint guarantees that the items reach the final point of the supply chain that is the customer. Besides, since the objective function minimizes cost, demand constraints avoid that the solution takes the value of zero by simply circulating no item within the network.

Constraint (5) ensures that the capacity needed by suppliers, plants and warehouses to manage the assigned quantities to the node is less than or equal to the global capacity available at the node.

Constraint (6) is connected to previous constraint as both ensure that the capacity of the node is enough to manage the quantity assigned. Complementing global capacity, this constraint refers to availability of specific capacity for each item.

Constraint (7) ensures that the amount of items moved between two nodes is less than or equal to the upper limit imposed on the flow from one node to the other. The constraint also guarantees that the point of origin has been selected to dispatch to the point of destination.

Finally, constraint (8) ensures that the amount of items moved between two nodes by a mode of transportation is less than or equal to its capacity.

3.2.3. Extensions

In addition to the original model, Cordeau (2014) proposed several extensions to cope with different situations. Here, we present the extensions that are relevant to our case.

Lower Bounds

Similar to capacity bounds imposed by constraints (5) to (8), the amount of units moving between two nodes can also has a minimum. Many situations may require a lower limit on the amount to dispatch. For example, a production batch usually requires a minimum size to cover fixed cost, or warehouses may demand a minimum quantity in the orders they receive from their customer to justify delivery service. These constraints have a similar structure as constraints (5) to (8), but in order to represent these lower bounds, the inequality sign has to be reversed. Moreover, the parameters denoted with q are now substituted with l, indicating the minimum amounts. This results in these new constraints:

$$\sum_{k\in K}\sum_{d\in D^k}\sum_{m\in M_{od}^k} u_o^k X_{od}^{kmt} - l_o U_o \geq 0 \qquad o\in O;\ t\in T \qquad (9)$$

$$\sum_{d\in D^k}\sum_{m\in M_{od}^k} X_{od}^{kmt} - l_o^k V_o^k \geq 0 \qquad k\in K;\ o\in O^k;\ t\in T \qquad (10)$$

$$\sum_{m\in M_{od}^k} X_{od}^{kmt} - l_{od}^k\, Y_{od}^k \geq 0 \qquad k\in K;\ o\in O^k;\ d\in D^k;\ t\in T \qquad (11)$$

$$\sum_{k\in K} u^{km} X_{od}^{kmt} - l_{od}^m\, Z_{od}^m \geq 0 \qquad o\in O;\ d\in D;\ m\in M_{od};\ t\in T \qquad (12)$$

Inventory Level Constraints

Constraints (5) and (6) impose a limit on the total amount of units that can flow through each node in a period of time. However, warehouses are nodes with capacity to hold inventory between two periods. That capacity can also be subject to an upper or lower bound. Similar to capacity constraints (5) and (6), the upper limit on inventory capacity at the end of the period t for the warehouse w is represented by $\hat{q}_w^t$ as a specific capacity for each item $\hat{q}_w^{kt}$. Constraints (13) and

(14) ensure these upper bounds respectively.

$$\sum_{k \in K} \hat{u}_w^k \, I_w^{kt} \leq \hat{q}_w^t U_w \,. \qquad w \in W;\, t \in T \qquad (13)$$

$$I_w^{kt} \leq \hat{q}_w^{kt} V_w^k. \qquad w \in W;\, k \in K;\, t \in T \qquad (14)$$

In order to impose lower bounds, we need to determine the minimum amount to have in inventory at the end of the period. This amount may represent a safety stock or a point of replenishment, and it is defined with a turnover ratio denoted as α_w^{kt}. This parameter indicates the times the inventory of item k is "turned" or sold in period t at warehouse w. Therefore, the quantity in stock is the inverse of α_w^{kt} times the total amount of product flowing through the warehouse.

$$I_w^{kt} \geq \sum_{d \in D^k} \sum_{m \in M_{wd}^k} \frac{1}{\alpha_w^{kt}} X_{wd}^{kmt}. \qquad w \in W;\, k \in K;\, t \in T \qquad (15)$$

For example, if the total amount to be dispatched from the warehouse is 1200 units and the turnover ratio is 12, thus the amount to have in inventory on average will be 100 units. In this example the 100 items in stock are sold every month; therefore, we can say that the average inventory was sold 12 times during the period.

Single-sourcing Constraints

Finally, to ensure that the total demand of each customer for one product is dispatched from a unique warehouse, the following constraint can be imposed.

$$\sum_{o \in O^k} Y_{od}^k \leq 1. \qquad k \in K;\, d \in D \qquad (16)$$

4. Preparation of the Model to the Case

This chapter explains how the model presented in the previous chapter is fitted to represent the lubricants company case. After having analyzed the model and its possible extensions, we defined the values to assign to the selected parameters, while other parameters are dismissed. After assigning all the values of the case to represent the supply chain to optimize, the resulting model is summarized.

4.1. Selection of Decision Variables

As mentioned in Section 1.1, decision variables represent the possible options for the problem. In other words, the value obtained for the decision variable is the solution to the case. To make sure the variables will provide the answer needed by the Company, we firstly need to review the questions of the case and identify which variables will give the answer. Decision variables of the model were summarized previously in Table 5.

Regarding the questions of the problem, we defined that the objective of the case was to identify the configuration of the network that yields the lowest operating cost. The configuration entails selecting the optimal location of DCs, identifying the optimal assignment of customers to DCs, and selecting the mode of transportation between the nodes. Besides, the solution must consider the products to be stored at and dispatched from every DC, as well as the quantity of product to hold at the end of each period in every DC. For practical purposes, the DC will be treated as a warehouse.

Table 7: Link of Decision Variables to the Case

I_{w}^{kt}	The solution given by this variable will represent the quantity to have in stock in every w at the end of period p for every product k. As the index k will represent 99 products for bulk products with a ten-year horizon, we can expect 990 variables per DC. With 5 DCs among candidate locations, there are 4 950 such variables in the model for the bulk products. Similarly, packaged products dealing with 613 product and 4 DCs should yield up to 24 520 variables as part of the result.
X_{od}^{kmt}	The solution given by these variables will represent the quantity of each product to flow between two specific nodes by a mode of transportation in every period. If the amount of product going from a DC to a customer is higher than 0, it also indicates that the customer has been assigned to the DC. Also the mode of transportation has been selected as the optimal one between the two nodes. Assignment of products and customers might change from one period to another depending on changes in the demand.

Table 7 explains the decision variables that contribute to the solution of the case once the model is solved. It contains the two continuous variables in the model. Continuous means that they can take any fractional value in a certain range. The lower and upper bound of the range are imposed by the constraints. Note that the possible number of values that I_{w}^{kt} and X_{od}^{kmt} can take is infinite anyway since they are continuous variables. Unlike continuous variables, binary variables U_o , V_o^k, Y_{od}^k, and Z_{od}^m can take only the value of 0 or 1; however, they are also part of the optimization process and the output contributes to the solution of the case. For example, if the variable U_o takes the value of 1 when origin o represents the warehouse in Denver, we know that this location was selected to be part of the resulting network. Moreover, binary variables are linked to the continuous variables through some of the constraints. For example, the continuous variable X_{od}^{kmt} can take a strictly positive value when o represents the warehouse in Denver, only if variable U_o is equal to 1, due to constraint (7). This constraint ensures that units of item k are not transported from o to d unless origin o is selected as part of the network. Both continuous and binary variables are needed to answer the questions of our case.

4.2. Assignment of Parameters

Although all the decision variables are relevant to the solution of the case, not all the parameters have to be used to represent the supply chain of the Company. This occurs because the flexibility of the model allows to represent many different networks. The model can solve problems considering more levels in the supply chain and including more decisions to make, such as the selection of suppliers or assignment of products to multiple plants. Due to these reasons, the second step performed was identifying the parameters requiring input of data. All parameters were analyzed before dismissing any. As the data available in companies can usually exceed the data needed to solve the problem, attention was paid to avoid putting data in parameters that do not contribute to the solution of the case. Similarly, if a required variable is missing information, results will be far from optimality. Finding the right parameters enforced logical thinking, which is one of the major advantages of using models to optimize a supply chain. To verify the functionality of the model with the selected parameter, we used a simple example of the supply chain. In simple networks, results may be anticipated and total cost calculations may be performed.

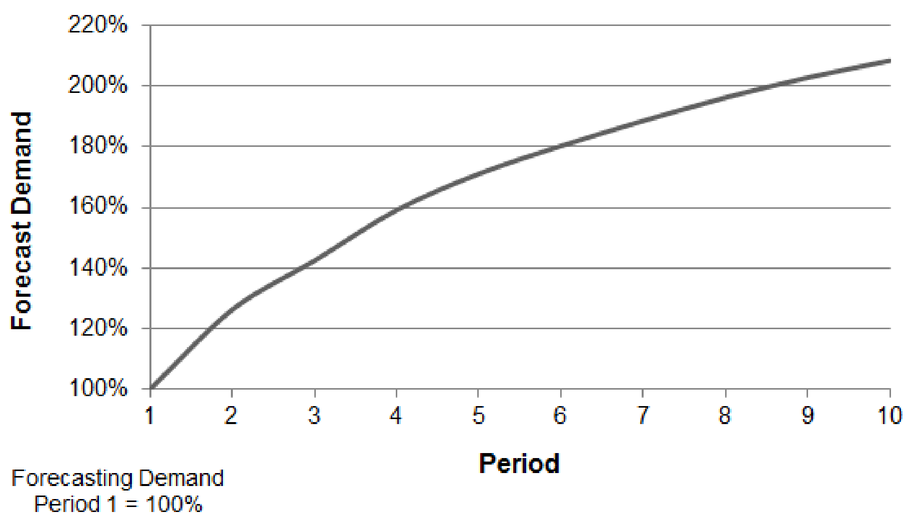
In the case of the Company, there are two types of products that must flow through the network: bulk and packaged products. As there is not any link or relation between them, we can indeed optimize the bulk and packaged goods separately. Even if bulk and packaged products can be in the same location, they cannot share facilities or modes of transportation; therefore, there are no costs subject to economies of scales, or decisions of one type affecting the other.

4.2.1. Assignment of Parameters for Bulk Products

In this section we select the parameters that are required to represent the supply chain of bulk products, and the data to be associated to each parameter. Parameters of the model were summarized previously in Table 4.

The first parameter to be used is a_c^{ft}, which represents the quantity demanded of each finished product at every period. This parameter is used in the demand constraint to ensure all customers can be served with the resulting network. The total forecast demand for the next ten years is presented in Figure 9:

Figure 9: Forecasting Demand for Bulk Products



Although total demand shows an expected increase in the demand, the detail per product and customer is required as an input to the parameter. Detailed data was provided by the Company so that the demand per customer and per product is available. As the solver can handle a large amount of input parameters and variables, we did not need to aggregate any data related to the demand. Thanks to this, we can expect higher accuracy in the results. For confidentiality reasons, just an extract of the data is presented in Table 8.

Table 8: Extract of Detailed Demand for Bulk Products

Customer	Product	Per1	Per2	Per3	Per4	Per5		Per10
Cus1	BulkP1	1 230	1 609	1 989	2 368	2 326	…	4 644
Cus2	BulkP2	216	228	-	252	264	…	324
Cus2	BulkP3	11 275	14 752	18 230	21 707	25 185	…	42 572
Cus2	BulkP6	820	1 073	1 326	1 579	1 832	…	3 096
Cus3	BulkP1	2 080	2 286	2 492	2 699	2 905	…	3 936
Cus3	BulkP3	9 635	10 590	11 546	12 501	13 456	…	18 233
Cus3	BulkP8	2 460	2 795	3 129	3 464	3 798	…	5 471
Cus3	BulkP15	21 840	19 413	16 987	14 560	12 133	…	-
Cus4	BulkP7	820	1 083	1 346	1 610	1 873	…	3 189
Cus4	BulkP8	384	507	631	754	877	…	1 493
Cus5	BulkP25	1 435	1 896	2 356	2 817	3 277	…	5 581
…	…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…	…
Cus168	BulkP99	4 160	5 495	6 831	8 166	9 501	…	16 178

USG

Reported values are not real but have been scaled to protect confidentiality.

For the total demand table for bulk products we cope with 99 products, which are demanded by 168 customers during the 10 periods. By a simple multiplication, we know that the table must contain 166 320 values. This is also the number of input values for the a_c^{ft} parameter. Even for customers not demanding a set of products, the value 0 must be indicated. In Table 8, there are examples of customers with increasing demand in some products, other customers demanding products only in some periods, and some customers whose demand for a product decreases with the time.

The next parameter we need to use is c_o^k, which is the fixed cost of assigning an item to an origin. In our case this fixed cost is related to the preparation of the tank. Although fixed cost usually refers to the construction cost of the facility, in this case the Company chooses to rent the facilities. One condition among candidate locations was to have a potential facility supplier with adequate infrastructure. In the case where the construction of a new tank is required, the supplier is responsible for the initial investment while preparation cost of the tank for the Company remains fixed. As the fixed costs may vary from one location to another, it may change the decision of the products to hold at every facility. The product assignment decision may also depend on the number of periods the product needs to be at the location. For example, it may be optimal economically speaking to have product BulkP1 in Denver from period 1 to period 3, but due to a reduction in the demand it is not from period 4 to period 10. As fixed costs would be sunk costs after period 1, the model takes this into account in order to evaluate the trade-off in the long-term. The total cost to optimize is not period per period, but the total cost for the ten- year horizon. Table 9 presents an extract of the fixed cost per product and per location. For comparison purposes, the basis cost equals 100 000.

Table 9: Extract of Fixed Cost for Bulk Products

Product	Ontario	Burnaby	Los Angeles	Denver	Memphis
BulkP1	100 000	104 500	122 000	108 500	116 300
BulkP2	100 000	104 500	122 000	108 500	116 300
BulkP3	100 000	104 500	122 000	108 500	116 300
BulkP4	100 000	104 500	122 000	108 500	116 300
BulkP5	100 000	104 500	122 000	108 500	116 300
…	…	…	…	…	…
…	…	…	…	…	…
BulkP99	100 000	104 500	122 000	108 500	116 300

US Dollars

Reported values are not real but have been scaled to protect confidentiality.

Then, the next parameter we need to incorporate is c_{od}^{km}, which is the unit cost of providing an item between two nodes by a mode of transportation. This is indeed the parameter for which we need to provide many input values, since it depends on four indices. Moreover, this parameter can represent the costs at the point of origin; for example, the production cost of the plant can be added to the trip from the plant to the warehouse. However, as it is a constant cost in our case and products are manufactured in only one plant, the manufacturing cost does not need to be considered. Therefore, it is not added to the parameter. Accordingly, parameter c_{od}^{km} represents only the transportation cost. First, the cost to move products from the plant to every warehouse is presented, followed by the cost to move products from warehouses to customers.

Table 10 contains the cost of transportation from the plant in Ontario to every warehouse by rail, which is the only mode of transportation to use in the case. This was another criterion that was taken into account when candidate locations were selected. For bulk products, the potential locations are Memphis, Los Angeles, Denver and Burnaby. Ontario is always considered as it is the main source of products. Although truck transportation is also an alternative to ship products from the plant, it should be used only for rush shipments arising during daily operation. As the supply chain configuration we pursue is designed for strategic and tactical purposes, short-term circumstances should not alter the result. Another aspect to consider is that cost structure has to be expressed in the same unit of measurement to be consistent. Until now, the demand has been expressed in USG, and the cost in US dollars, thus the transportation cost needs to be expressed in US dollars per USG.

Table 10: Transportation Cost from the Plant to Warehouses for Bulk Products

Origin	Destination	Freight per Mile	Rail Tank Miles	Cost per Trip	Cost per USG
Plant	Ontario	-	-	-	-
Plant	Burnaby	10.0880	2 699	27 227	1.1595
Plant	Los Angeles	9.8824	2 609	25 783	1.0980
Plant	Denver	13.4719	1 511	20 356	0.8668
Plant	Memphis	10.0880	985	9 937	0.4231

US Dollars

Reported values are not real but have been scaled to protect confidentiality.

As transportation rates are usually expressed in distance units, some calculations were needed to turn them into cost per USG. First, the cost per mile for a rail tank car was provided for the rail company according to the country and state of destination. Then, the total transportation cost with all the charges was obtained. Finally, the cost of every trip was divided by the capacity of the rail tank, which is 23 483 USG, in order to obtain the cost of moving one USG from the plant to a warehouse. As the cost remains the same for every product due to similar characteristics, the parameter may take only 5 different values, one for every warehouse. This is so because in c_{od}^{km} parameter k does not affect the cost, parameter m represents only rail, origin o is only one plant, and d is the one giving the value of the 5 warehouses. A last point to outline in Table 10 is that the cost from Plant to Ontario is 0. This is to represent that both the plant and warehouse are located in the same place; however, the plant needs to be represented as a warehouse to give it the attribute of holding inventory.

Similar calculations were performed to obtain the cost of transportation from warehouses to customers. In this case we had the geographical coordinates of the candidate warehouses and the 168 customers. By using a premium distance calculator software, we obtained the driving distance from every warehouse to every customer. With 5 points of origin to 168 points of destination, the resulting table contains 840 distances expressed in miles. The freight rate of the carrier depended on the state of origin and considers the go and return trip. Although some cases showed that the destination affected the rate, it happened in less than 5% of cases, and hence we considered that our approach was acceptable. Since we assume full truck loads, the total cost of every trip was divided by the capacity of the truck, which is 6 193 USG, in order to obtain the cost per USG. Table 11 presents an extract of the cost of transportation from warehouses to every customer by truck.

Table 11: Extract of Transportation Cost from Warehouses to Customers for Bulk Products

From / To	Ontario			Burnaby				Los Angeles	Denver	Memphis
	Distance	C. per Mile	C. per USG	Distance	C. per Mile	C. per USG		C. per USG	C. per USG	C. per USG
Cus1	572	13.2288	1.2218	2 143	16.1391	5.5847	…	4.7723	2.3115	3.1787
Cus2	1 062	13.2288	2.2685	1 022	16.1391	2.6634	…	3.1122	1.9323	2.6915
Cus3	1 876	13.2288	4.0073	1 224	16.1391	3.1898	…	0.8772	3.6760	2.2259
Cus4	2 057	13.2288	4.3939	139	16.1391	-	…	2.3267	3.1287	3.3143
Cus5	209	13.2288	0.4464	2 082	16.1391	5.4257	…	5.1093	5.4540	2.8880
Cus6	51	13.2288	-	2 118	16.1391	5.5196	…	5.4364	4.3904	3.8709
Cus7	520	13.2288	1.1108	1 919	16.1391	5.0010	…	4.3685	2.8908	2.3943
…	…	…	…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…	…	…	…
Cus168	2 132	13.2288	4.5541	1 035	16.1391	2.6972		-	1.8790	2.9012

US Dollars

Reported values are not real but have been scaled to protect confidentiality.

A final remark on Table 11 is the case of trips with zero cost. This occurs because most of the customers within a distance of 160 miles from the warehouse are willing to use their own fleet to pick up the product, leading to a cost of zero for the Company. This aspect may have a significant effect on the results, since the model will strive to assign products demanded by customers inside this ratio to the warehouses near the customers, motivated by this “zero cost”.

A last parameter to incorporate is g_w^k, which represents the cost of holding one unit of product k at location w for one period of time. However, as there is another cost related to handling activities, such as the manipulation of the products, the parameter $\hat{g}_w^k$ is used to represent the cost of handling one unit of product k at location w. These independent parameters are associated with every warehouse. Parameter g_w^k affects the final inventory at the end of each period, and $\hat{g}_w^k$ affects all the units flowing through the warehouse. The 5 different holding and handling costs are presented in Table 12.

.

Table 12: Holding and Handling Cost per Warehouse for Bulk Products

Warehouse	Holding Cost per USG	Handling Cost per USG
Ontario	0.4123	0.0820
Burnaby	1.0858	0.2062
Los Angeles	1.2676	0.1354
Denver	1.1273	0.1275
Memphis	1.2084	0.1134

US Dollars

Reported values are not real but have been scaled to protect confidentiality.

4.2.2. Assignment of Parameters for Packaged Products

Similar to bulk products, we had to select the required parameters to represent the supply chain of packaged products. As most of the costs refer to the same concepts, the parameters are briefly presented with the costs data to be associated to each parameter. The fixed cost of preparing the facilities we had in bulk products is not needed for packaged products. A more standard operation and facilities permit to avoid this cost to the Company. As for bulk products, the facilities will be rented from a third party having already the facilities. Although there is no direct interaction between the two types of products, some locations are able to store both bulk and packaged products. These locations are Ontario, Los Angeles, Denver and Memphis. Having the two types of product at the same location in any way contributes to simplify administration and managerial tasks.

To continue with the same order of Section 4.2.1, the expected demand to be associated to parameter a_c^{ft} is presented in Table 13.

Table 13: Extract of Detailed Demand for Packaged Products

Customer	Product	Per1	Per2	Per3	Per4	Per5		Per10
Cus10	PackP1	5 740	8 360	10 980	13 600	16 219	…	29 319
Cus10	PackP2	820	901	983	1 064	1 145	…	1 552
Cus23	PackP1	6 240	6 574	6 909	7 243	7 577	…	9 249
Cus23	PackP2	410	451	491	532	573	…	776
Cus23	PackP97	1 435	1 896	2 356	2 817	3 277	…	5 581
Cus52	PackP1	1 040	1 386	1 731	2 077	2 422	…	4 150
Cus52	PackP8	410	458	506	554	601	…	841
Cus103	PackP112	21 840	19 413	16 987	14 560	12 133	…	-
Cus103	PackP215	1 640	2 146	2 652	3 157	3 663	…	6 192
Cus221	PackP18	500	549	598	647	696	…	940
Cus315	PackP420	1 025	1 254	1 482	1 711	1 940	…	3 083
…	…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…	…
Cus894	PackP613	3 408	3 618	3 829	4 039	4 249	…	5 301

USG

Reported values are not real but have been scaled to protect confidentiality.

The cost of transportation from the plant to every warehouse is denoted by parameter c_{od}^{km}. Similarly to bulk products, the origin is always the production plant, while the destinations are the suitable warehouses for packaged products. The candidate locations for this type of products apart from Ontario are Los Angeles, Denver, Portland, and Salt Lake. In this case, the shipments that have Ontario as their destinations are always considered with zero cost. This occurs because there is no real distance between the plant and the warehouse in Ontario. Finally, to dispatch a product to Los Angeles and Denver, intermodal transportation is an available option. Thus the costs have to be taken into account. Table 14 presents the cost of transportation from the plant to every warehouse by rail or by intermodal means when the option exists.

Table 14: Transportation Cost from the Plant to Warehouses for Packaged Products

Origin	Destination	Freight per Mile	Rail Miles	Cost per Trip	Cost per USG
Plant	Ontario	-	-	-	-
Plant	Los Angeles	9.8824	2 699	26 673	1.1358
Plant	Denver	13.4719	2 609	35 148	1.4968
Plant	Memphis	10.2451	1 511	15 480	0.6592
Plant	Portland	16.2489	985	16 005	0.6816
Plant	Salt Lake	13.7852	1 511	20 829	0.8870

Origin	Destination	Freight per Mile	Intermodal Miles	Cost per Trip	Cost per USG
Plant	Los Angeles	11.7582	2 699	31 736	1.3514
Plant	Denver	26.1271	2 609	68 165	2.9028

US Dollars

Reported values are not real but have been scaled to protect confidentiality.

Parameter c_{od}^{km} also represents the cost of transportation from warehouses to every customer by truck. With 6 candidate warehouses and 894 customers for packaged products, the resulting table contains 5 364 distances expressed in miles, the cost rate per mile per USG, and the total transportation cost per USG. Table 15 shows the cost of moving one USG for a mile. In this case, transport suppliers provided the cost by USG since unlike bulk products, the truck might transport different products to maximize the truck load if regulation allows it. In order to maintain consistency in all the information, we obtained the cost per USG from all possible points of origin to all possible destinations. Table 15 presents an extract of the cost of transportation. Similarly to bulk products, customers within a distance of 160 miles from the warehouse are willing to use their own fleet to pick up the product, leading to zero cost for the Company.

Table 15: Extract of Transportation Cost from Warehouses to Customers for Packaged Products

From / To	Ontario Distance	Ontario C. per Mile per USG	Ontario Cost per USG	Los Angeles Distance	Los Angeles C. per Mile per USG	Los Angeles Cost per USG		Denver Cost per USG	Memphis Cost per USG	Portland Cost per USG	Salt Lake Cost per USG
Cus1	329	0.0048	1.5668	2 301	0.0053	12.1757	…	8.7045	2.2489	16.1655	8.2812
Cus2	673	0.0048	3.2051	1 491	0.0053	7.8896	…	3.9157	1.2435	10.6624	4.5771
Cus3	1 668	0.0048	7.9436	700	0.0053	3.7041	…	3.2543	2.8310	8.8368	3.1220
Cus4	433	0.0048	2.0621	1 732	0.0053	9.1649	…	5.2915	1.3758	12.0646	5.6090
Cus5	174	0.0048	0.8286	2 330	0.0053	12.3292	…	8.7574	2.6722	15.6628	8.2283
Cus6	1 332	0.0048	6.3435	829	0.0053	4.3867	…	-	2.5664	6.7731	1.6668
Cus7	1 928	0.0048	9.1818	236	0.0053	1.2488	…	3.4659	4.0745	5.1592	-
…	…		…	…		…	…	…	…	…	…
…	…		…	…		…	…	…	…	…	…
Cus894	1 069	0.0048	5.0910	1 227	0.0053	6.4927		3.2014	1.0848	10.5301	4.0745

US Dollars

Reported values are not real but have been scaled to protect confidentiality.

The values for the last parameters g_w^k, and $\hat{g}_w^k$, which represent the holding and handling cost respectively at every warehouse, are presented in Table 16.

Table 16: Holding and Handling Cost per Warehouse for Packaged Products

Warehouse	Holding Cost per USG	Handling Cost per USG
Ontario	0.2062	0.0825
Los Angeles	0.67757	0.2179
Denver	0.60259	0.1307
Memphis	0.64591	0.2179
Portland	0.55538	0.1478
Salt Lake	0.5554	0.1307

US Dollars

Reported values are not real but have been scaled to protect confidentiality.

4.3. Selection of Constraints and Extensions

After having assigned the costs of the supply chain to the parameters needed in the objective function and in the constraints of the original model, we selected the additional constraints needed for the case. The subset of constraints that can be removed while achieving a feasible solution was presented in Section 3.2.3. To ensure a faithful representation of the supply chain of the Company, we verified all the constraints before dismissing any of them. Selecting the right constraints ensures the solution will comply with the limits of the case.

Among the constraints of the original model, presented in Section 3.2.2, some of them are needed to ensure a feasible solution. Those constraints are: the inventory consistency constraint (3), and the demand constraint (4). It is worth mentioning that constraint (2) is a must for problems including the flow of items from suppliers to the plant to ensure the flow conservation. For our case, constraint (3) ensures also the flow of items from the plant to the customers. To properly model the supply chain of our case, we need to incorporate other four constraints.

The first set of constraints to include is the lower bound for the amount to transport between the plant and the warehouses for bulk products. As rail tank cars are used in this trip and as they are exclusive for one product, a minimum filled level of 10% is required for security reasons. This is 2 348 USG per product that must be assigned to parameter l_{od}^{m} in constraint (12). This minimum amount does not apply for the products dispatched from the plant to Burnaby, since shipments below the minimum level can be filled and then used to serve local market.

The second constraint to add is the inventory level constraint of the warehouse. The lower bound to impose to each item is presented in constraint (14). With this constraint we can impose the minimum desired capacity usage of the tank for bulk products, which is 9 000 USG. This value has to be assigned to parameter $\hat{q}_{w}^{kt}$ and has to be repeated for every item and period in the problem. Note that we need to reverse the inequality sign in the constraint to ensure we are imposing a lower bound instead of an upper one.

Although a minimum capacity usage is defined, we still need to set the level of inventory to have at the end of the period at every warehouse. We can do this by assigning the desired turnover ratio per product to parameter α_w^{kt} and enabling the constraint (15), related to inventory level. The expected turnover ratio has been provided by the Company according to its inventory policies. Table 17 presents an extract of the turnover ratio assigned to every product.

Table 17: Extract of Turnover Ratio per Product

Warehouse / Product	Ontario		Burnaby		Los Angeles		Denver	Memphis
	Inventory Days	Turnover Ratio	Inventory Days	Turnover Ratio	Inventory Days	Turnover Ratio	Turnover Ratio	Turnover Ratio
BulkP1	21	17.3810	30	12.1667	45	8.1111	24.3333	17.3810
BulkP2	21	17.3810	30	12.1667	40	9.1250	24.3333	17.3810
BulkP3	25	14.6000	35	10.4286	40	9.1250	30.4167	13.5185
BulkP4	25	14.6000	35	10.4286	42	8.6905	17.3810	24.3333
BulkP5	18	20.2778	25	14.6000	38	9.6053	20.2778	20.2778
…	…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…	…
BulkP99	21	17.3810	28	13.0357	30	12.1667	17.3810	15.2083

Then, the single sourcing constraint (16) is needed to ensure that customers will be served from only one warehouse for every product demanded. This is a requirement of the Company to provide a better level of service to customers, avoiding product consolidation and different time arrivals for the same item.

Finally, we need constraint (6) to correctly model the V_o^k binary variable, which in the objective function is associated with the fixed cost of preparing a tank, denoted by c_o^k.

4.4. Resulting Model

After having selected the decision variables, assigned values to the parameters to use, and selected the adequate set of constraints, the model is now ready to solve the problem. Even if the model we are going to run is the complete one presented in Section 3, we are using neither all the parameters nor all the constraints. Therefore, we can say that we are using a reduced version fitted to the case. Reduction occurs because when some parameters are dismissed, they need to take the value of zero or infinity in the original model. Indeed, one of those values needs to be assigned to every dismissed parameter to avoid any impact in the result. When parameters take the value of zero, some expressions disappear in the objective function. After assigning values to variables and parameters, we can summarize the simplification of the model as composed of the objective function (1), and the constraints (3), (4), (6), (12), (14), (15), and (16).

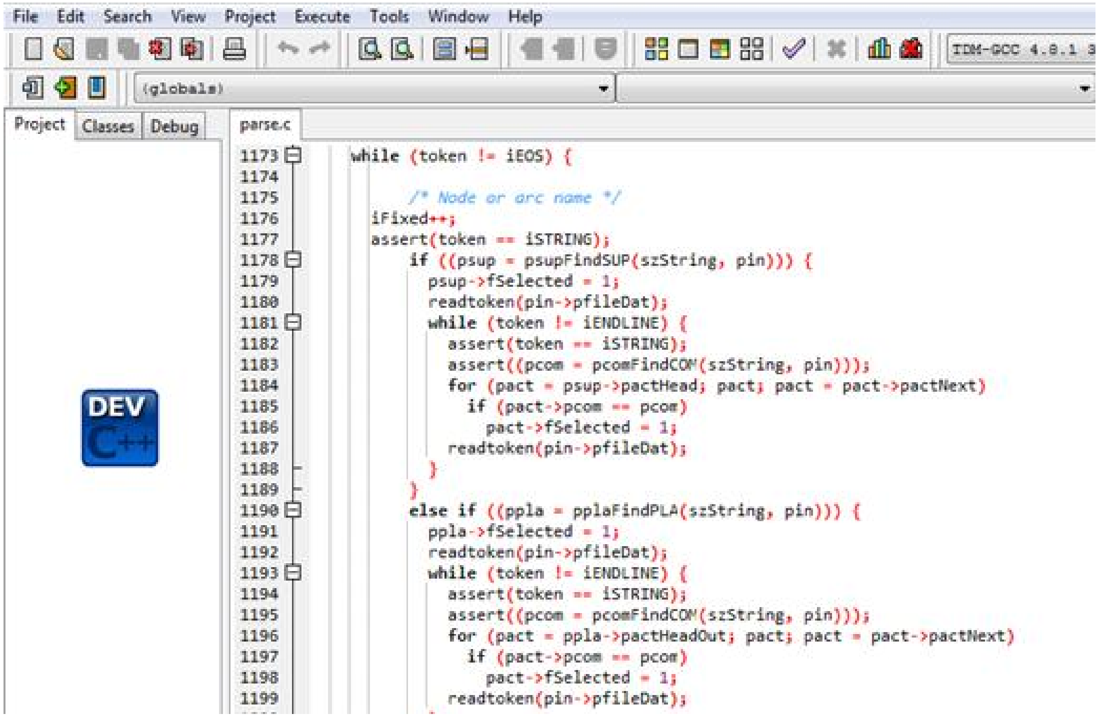
5. Data Set up in the Model

This chapter presents how the data assigned to every parameter in Chapter 4 is embedded into the model presented in Chapter 3. A user-friendly tool developed in standard software to accommodate data in the proper structure is presented. Once the data is sorted, the model is executed for the scenarios selected to evaluate. Finally, the transformation performed to the data before it could be used by the model is discussed.

5.1. Model Implementation Overview

When implementing an optimization model, a software is always needed to make possible the considerable amount of calculations in the problem. First, the mathematical model has to be written in the language required by the software. Then the data has to be put into the model to finally perform calculations and obtain a result. Cordeau (2014) implemented the model presented in Chapter 3 by using the C callable library of CPLEX. The C language is a general purpose programming language that allows building computer instructions through functions. CPLEX is a solver that optimizes integer programming and linear programing problems. Currently developed by IBM, CPLEX provides an interface with C to execute and debug optimization problems. In the model in reference, the C code was used to arrange the data of the problem, validate its structure, and make the link with the solver to upload the data to the problem.

Figure 10: Extract of C Language Programming of the Model



Source: Cordeau (2014)

Figure 10 presents an extract of the original C language code as a reference. More than 1 000 lines of code in this file are a good example of the work required to implement an optimization model. Some other files had to be created to upload the data into the model and make the link with CPLEX. As the modeling exercise has been done in advance since we are dealing with an existing model, no more attention will be paid on modeling matters. Because of this, during the execution of the model in this thesis, the developments in C and CPLEX were used as a black box. This means that we executed the model without the need to access the code to modify or adapt it to our case. All the assignment of values to parameters, and selection of constraints presented in Chapter 4 were performed only through the input data file to be read by the code. The input file is a text file with a specific order of characters that includes all the data of the problem to be solved. The format has also to respect a convention to make the data readable by CPLEX.

The input file is composed of eight sections: scenarios, periods, items, suppliers, plants, warehouses, customers and modes of transportation linking the nodes. Each section of the text file requires following the structure and format that is presented in the Appendix 1. The parameters have the same notation as the one used in the mathematical model, and summarized in Table 4 of Chapter 3.

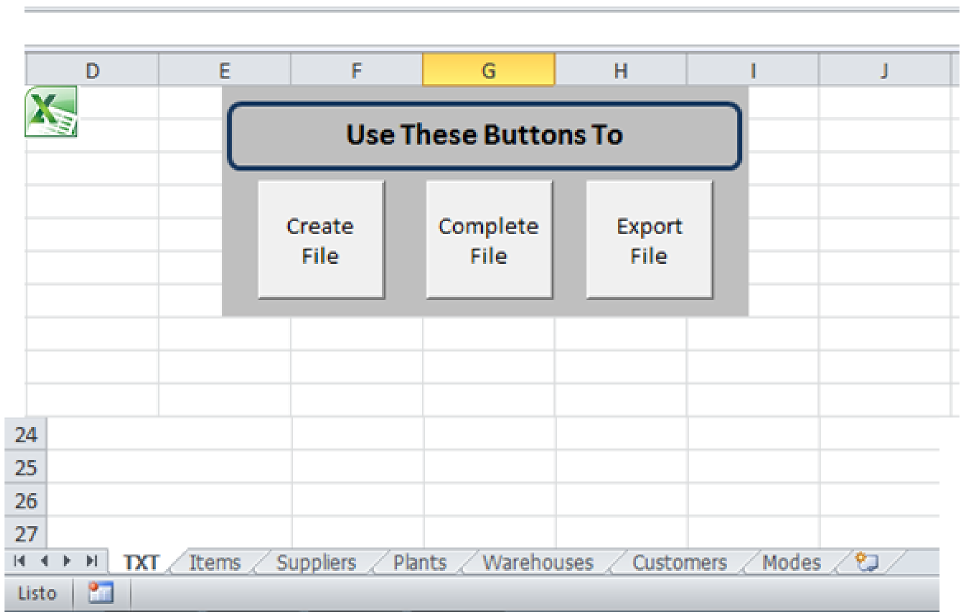
5.2. Matching Tool Development

Having analyzed the format required for the text file, which is presented in Appendix 1, we opted for building a spreadsheet to easily sort the data. A spreadsheet can be turned into a text file by simply changing the extension of the file to .txt. For practical purposes, we chose the standard version of Microsoft Excel 2010. Unlike working in C code environment or CPLEX, Excel is a software much more common to people, thus creating a tool like the one presented here is more practical.

The file follows almost the same structure as the eight sections demanded by the input file. One tab contains the structure of scenarios and periods. Then, there is an independent tab for each of the other sections, which are items, suppliers, plants, warehouses, customers and modes. The structure and functionality of each tab is presented here.

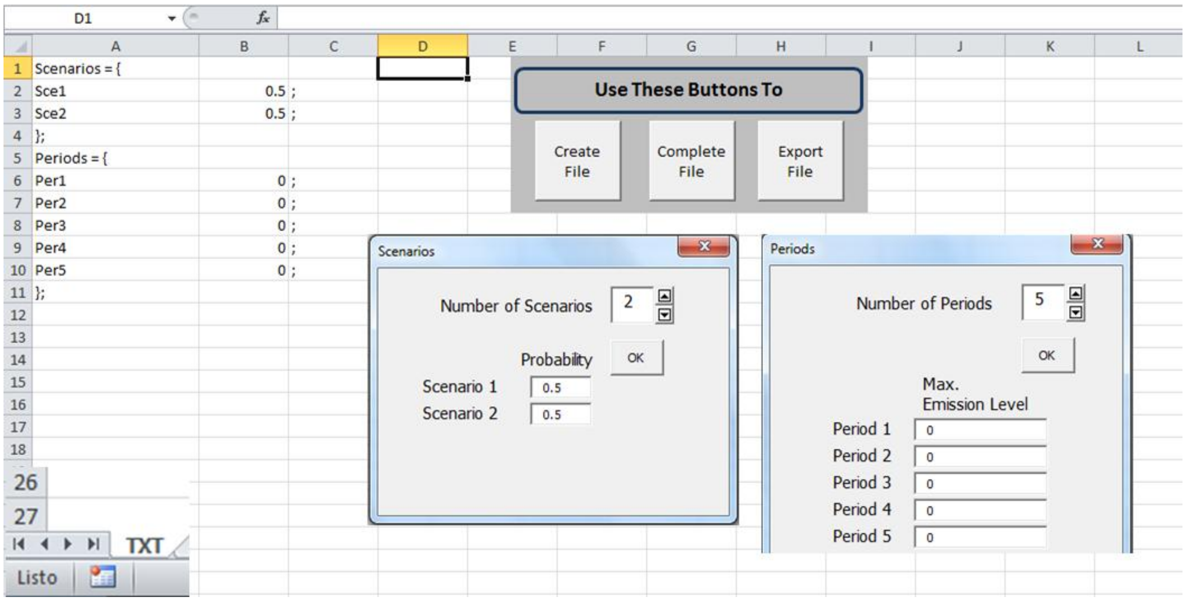
The first spreadsheet is called “TXT”, since this is intended to gather the data in the required format to then be converted into a txt file. In order to monitor the progress of the file, the commands of the file are also in this sheet. The commands summarize the three main steps required to build the txt file. The first step starts with the “Create File” button, the second step is performed with the “Complete File”, and finally “Export File” creates the input file with the required extension. Figure 11 presents a display of the file and the first sheet.

Figure 11: Display of Excel File and TXT Sheet



The "Create File" button calls on two new forms to select the number of scenarios and the number of periods in the problem. The sum of the probability of all the scenarios must be equal to 1, a simple validation that is done by the Excel file to avoid inconsistency in the data. The complete model including the possibility to have multiple scenarios is given in Cordeau (2014). In our further experiments, we will always consider only one scenario. Then, the form for the periods requires an emission level parameter to set the upper bound of emissions for each period. This parameter was not presented before as it belongs to a set of extensions that are not taken into consideration in our case. We entered zero to avoid any impact of this upper bound. Once the parameters have been entered, the structure required by the text file appears in the beginning of the sheet. Figure 12 shows the initial structure of the file when it is created and the instructions of the first button have been successfully executed. Although this structure could be written manually, here we built an Excel-Macro to do it faster and guarantee that the structure for the txt file is respected. The macro becomes useful when running many scenarios, so our attention may be focused on the consistency of the data rather than on format issues.

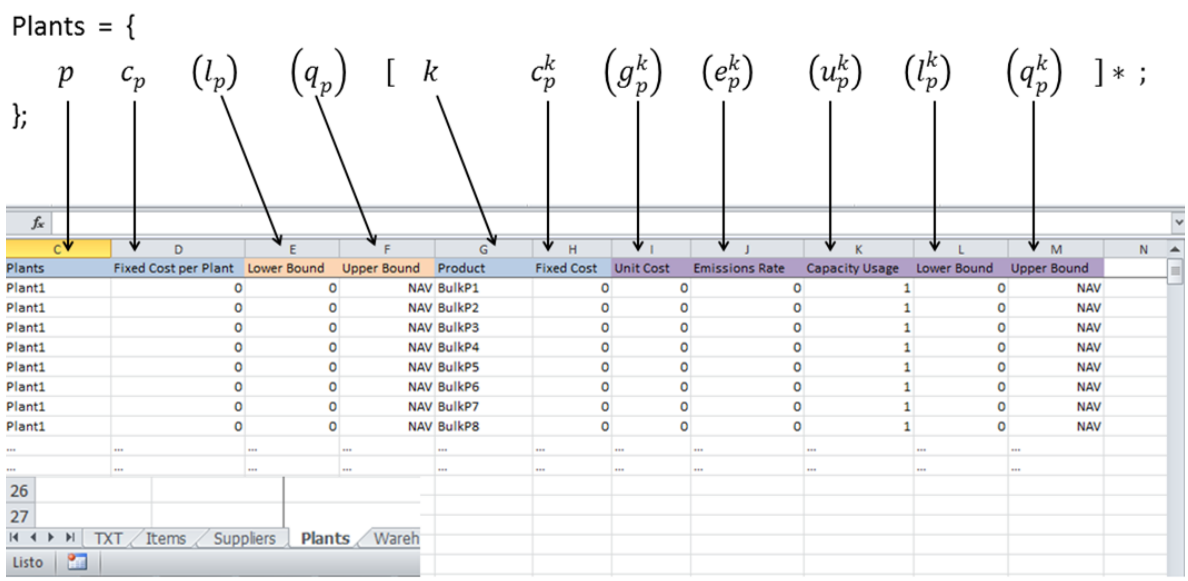
Figure 12: Display of Create File Button Functionality



The next command button, "Complete File", performs the instructions to accommodate the data of the missing sections in the right format. As the data is imported from the other six tabs, the content of those is first presented. As shown in Figure 11, the next tab from left to right is "Items". This tab contains only one column with a list of all the products to consider in the supply chain. No more columns are needed in this case, because billing of materials is not an issue in our case. The next tab is "Suppliers", whose structure is similar to the other four sheets, and it is available for future cases considering more tiers. In our case, we go directly to tab "Plants" to comment its structure.

The "Plants" sheet was built to serve two basic purposes: providing a familiar environment to easily accommodate the data of the case, and ensuring that data can be exported to the first tab of the file. To make sure that all the parameters related to plants are taken into account, the columns were placed in the same order as they are required by the text file. A short description of the parameter is in the title of the column. This structure is presented in Figure 13.

Figure 13: Plants Structure Representation in the Spreadsheet



Plants	Fixed Cost per Plant	Lower Bound	Upper Bound	Product	Fixed Cost	Unit Cost	Emissions Rate	Capacity Usage	Lower Bound	Upper Bound
Plant1	0	0	NAV	BulkP1	0	0	0	1	0	NAV
Plant1	0	0	NAV	BulkP2	0	0	0	1	0	NAV
Plant1	0	0	NAV	BulkP3	0	0	0	1	0	NAV
Plant1	0	0	NAV	BulkP4	0	0	0	1	0	NAV
Plant1	0	0	NAV	BulkP5	0	0	0	1	0	NAV
Plant1	0	0	NAV	BulkP6	0	0	0	1	0	NAV
Plant1	0	0	NAV	BulkP7	0	0	0	1	0	NAV
Plant1	0	0	NAV	BulkP8	0	0	0	1	0	NAV
...	...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...	...

The equivalent structures permit to easily put data into the cells, and let data with the order needed to be taken to the "TXT" sheet. Putting data into the cells is equivalent to assigning values to the parameters as presented in Section 4.2. In the example presented in Figure 13 we can see that all cells need a character. To avoid impact of dismissed parameters on results, either the value of zero or expression NAV must be entered into the cells of unused parameters. Then, as scenarios and periods have to be included in the text file, another Excel-Macro was created to consider these parameters. A button called "Duplicate Data", which is in the sheet "Plants", displays a form to select the number of scenarios and periods to consider. The initial data is assumed to be assigned to scenario one and period one. These values are over the titles of the parameters. Once the selection is entered into the form, the Macro duplicates the data into a new block, but the rows Scenario and Period are updated. For example, if the selection is 1 scenarios and 5 periods, 5 new blocks of data will be created with the same content. As the structure of the sheet is ready to enter new data, updating columns with new data for the new periods can easily be done.

Figure 14: Display of Duplicate Data Button

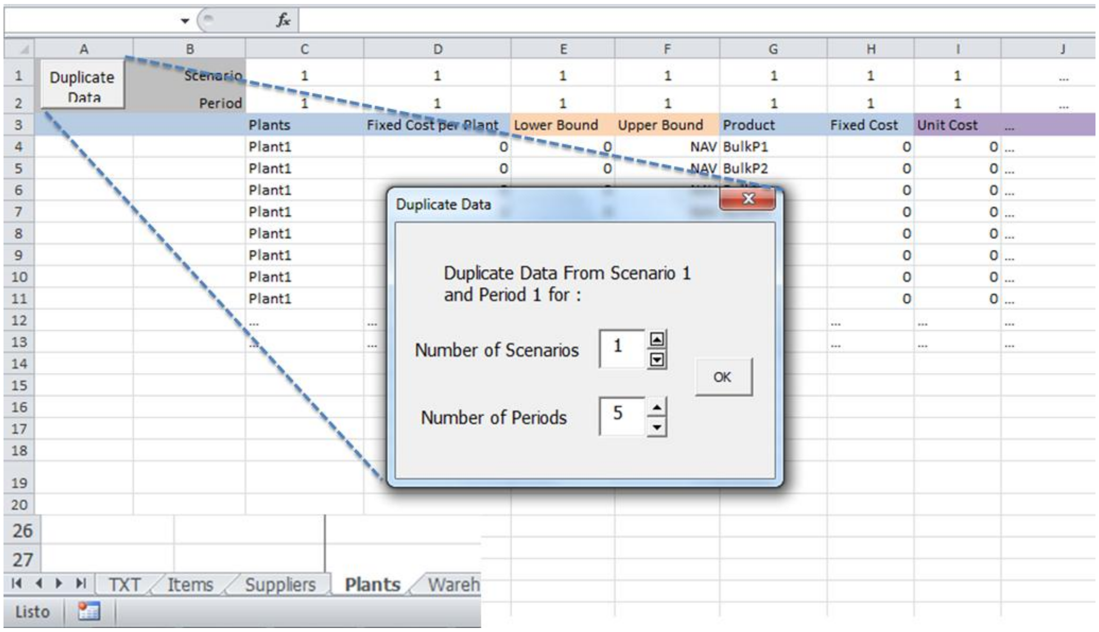


Figure 14 presents an example of the form created to execute the Macro described. Similarly to previous commands, the objective is to minimize the time of sorting the data manually, which is indeed the objective of a Macro. Moreover, duplicating the data for the next scenario or period offers an advantage to the user, since in many cases most of the data might not vary significantly from one year to the other. In the case of no variation in the data, the new blocks of data would be ready to be taken to the first tab. In our case, for example, data of the plant is used only to represent that products are manufactured and thus available to be shipped to warehouses; however, neither costs nor capacity bounds are added to the problem. In this case by simply executing the Macro, the data in sheet "Plants" is ready for the one scenario and ten periods we are considering in the case. The same structure and the same functionality are in tabs for warehouses, customers, and modes. In the "Customers" tab, for example, the values of the demand have to be updated, since the demand is different from period to period. However, updating the demand column is the only task that needs to be done manually. Figure 15 gives an example of the structure of the missing tabs.

Figure 15: Display of Warehouses, Customers, and Modes Tabs

	Warehouses	Fixed Cost per War.	Handling Lower	Handling Upper	Storage Lower	Storage Upper	Product	Fixed Cost per Product	Handling Unit Cost	Storage Unit Cost	Handling Emissions	Storage Emissions	Handling Capacity	Storage Capacity	Handling Lower
Scenario	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Period	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Ontario	100000	0	NAV	0	NAV	BulkP1	0	0.0820	0.4123	0	0	1	1	0
	Ontario	100000	0	NAV	0	NAV	BulkP2	0	0.0820	0.4123	0	0	1	1	0
	Ontario	100000	0	NAV	0	NAV	BulkP3	0	0.0820	0.4123	0	0	1	1	0
	Ontario	100000	0	NAV	0	NAV	BulkP4	0	0.0820	0.4123	0	0	1	1	0
	Ontario	100000	0	NAV	0	NAV	BulkP5	0	0.0820	0.4123	0	0	1	1	0
	Ontario	100000	0	NAV	0	NAV	BulkP6	0	0.0820	0.4123	0	0	1	1	0
	Ontario	100000	0	NAV	0	NAV	BulkP7	0	0.0820	0.4123	0	0	1	1	0
	Ontario	100000	0	NAV	0	NAV	BulkP8	0	0.0820	0.4123	0	0	1	1	0

Plants | Warehouses | Customers

	Customers	Product	Min. Demand	Max. Demand	Unit Revenue	Single Sourcing? No
Scenario	1	1	1	1	1	1
Period	1	1	1	1	1	1
	Cus1	BulkP1	1230	1230	0	1
	Cus2	BulkP2	216	216	0	1
	Cus2	BulkP3	11275	11275	0	1
	Cus2	BulkP6	820	820	0	1
	Cus3	BulkP1	2080	2080	0	1
	Cus3	BulkP3	9635	9635	0	1
	Cus3	BulkP8	2460	2460	0	1
	Cus3	BulkP15	21840	21840	0	1
	Cus4	BulkP7	820	820	0	1
	Cus4	BulkP8	384	384	0	1
			1435	1435	0	1

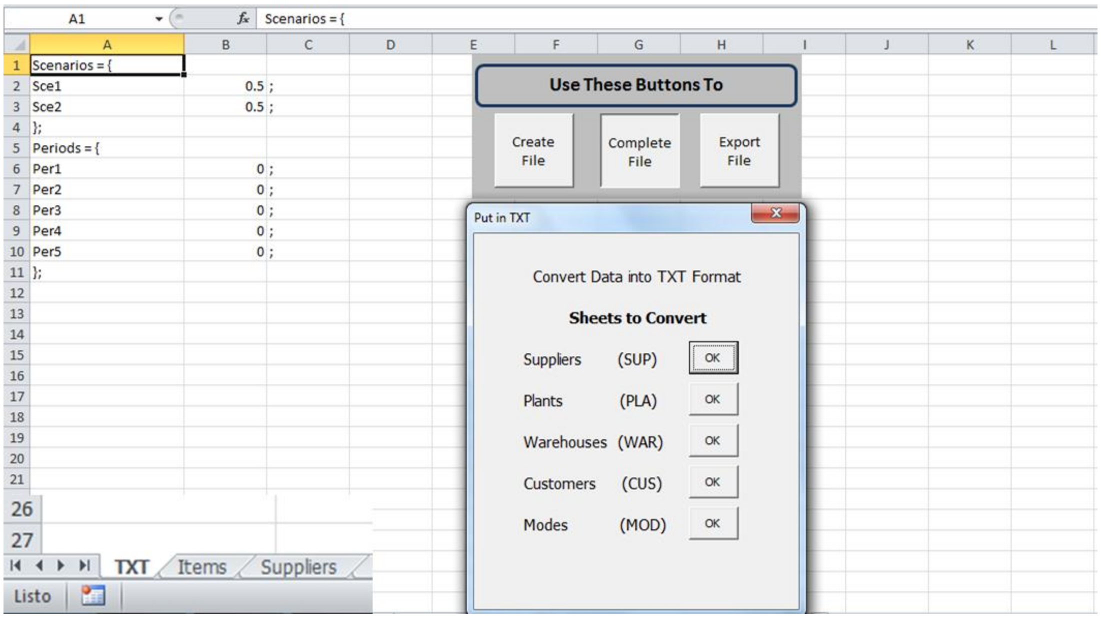
Warehouses | Customers | Modes

	Modes	Origin	Destination	Fix Cost per Mod.	Lower Bound	Upper Bound
Scenario	1	1	1	1	1	1
Period	1	1	1	1	1	1
	Plant_to_Ontario	Plant1	Ontario	0	0	NAV
	Plant_to_Ontario	Plant1	Ontario	0	0	NAV
	Plant_to_Ontario	Plant1	Ontario	0	0	NAV
	Plant_to_Ontario	Plant1	Ontario	0	0	NAV
	Plant_to_Ontario	Plant1	Ontario	0	0	NAV
	Plant_to_Ontario	Plant1	Ontario	0	0	NAV
	Plant_to_Ontario	Plant1	Ontario	0	0	NAV
	Plant_to_Ontario	Plant1	Ontario	0	0	NAV
	Plant_to_Ontario	Plant1	Ontario	0	0	NAV
			Ontario	0	0	NAV

Customers | Modes

Once the data of the problem is entered into the six tabs, the command button "Complete File" can be executed. This button displays a new form that presents the five sections that need to be imported to "TXT" sheet with its respective button to be executed. There is no button for the tab "Items", since instructions to import the list of items are in the "Suppliers" button. The buttons in the form have to be executed in the order they appear to ensure the text file will have the adequate structure to be read by CPLEX. Each button has an Excel-Macro that imports the data from the indicated tab to the "TXT" sheet. Accommodating the data in the right structure might be an endless task if performed manually. Note that the structure requires entering the values of one parameter per period and per scenario ($|T| \times |H|$). This implies that every line needs to import data from every period and scenario in the data sheet. This way to sort the data required developing the Macros, and overall developing this tool to adequately represent our case. Figure 16 presents the form created to execute the Macros described.

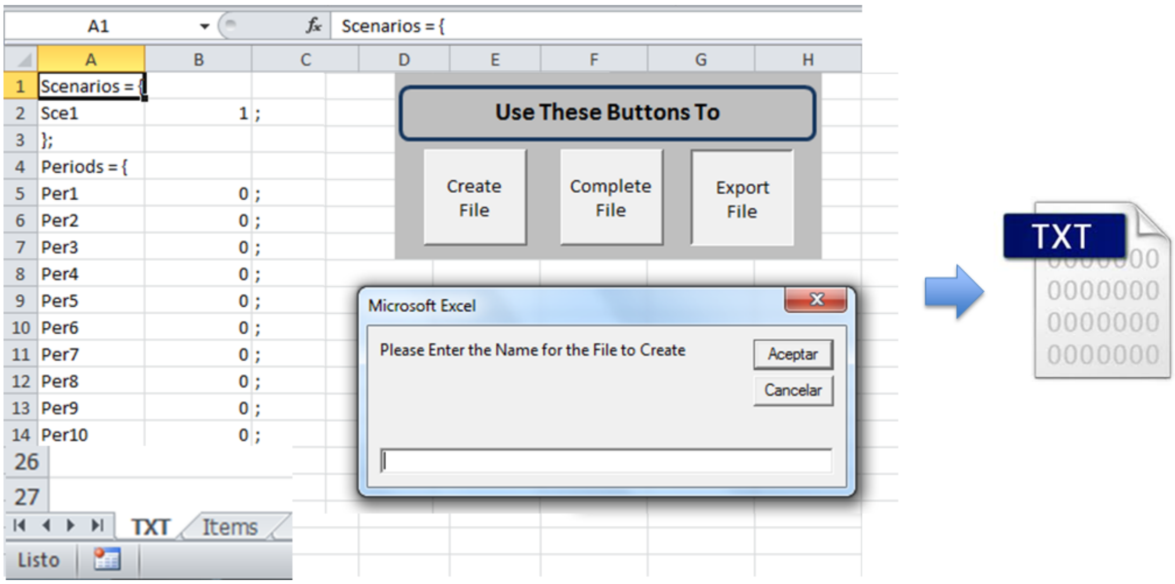
Figure 16: Display of Complete File Button



After executing the five buttons in the form, the TXT sheet will have all the data of the problem in the required structure. The buttons allow to import the data gradually from the other sheets, so that after executing every Macro, the data can be verified before executing the next section. Besides, if a change is needed in a specific sheet, the instruction for the sheet can be executed independently of the others. This provides flexibility when dealing with a significant amount of data.

Finally, the last command button can be executed. The Export File button activates an Excel-Macro to save the TXT sheet as a file with the extension .txt. A new form appears to enter the name desired for the resulting file, which will be then saved in the folder specified in the Macro. The name assigned to the file should be the name to identify the problem, since after executing the model in CPLEX, the output file with the solution will have the same name. Finally, the Macro also ensures that during the conversion, the data in different cells are separated by an empty space as required by the input file format. Figure 17 represents the functionality of the Macro described.

Figure 17: Display of Export File Button



5.3. Model Execution

Execution of the model is the process that uploads the data from the text file to CPLEX, runs the solver in order to optimize the problem, and provides a file with the solution. This process requires compiling the C code used for the implementation of the model. The code can be compiled in both Unix or Microsoft Windows environments, and for the execution, ILOG CPLEX 10.0 or newer has to be available to the computer. The standard instruction “make” runs the process taking the input data from the text file. After a few seconds some files are automatically created alerting about inconsistencies in the data, or exposing reasons that made the problem infeasible. If the problem is feasible, the CPLEX model is generated and solved, and a new output text file is created with the solution of the problem. The content of this solution file will be presented in Chapter 6.

As mentioned in Section 2.1.2, the model we used in this thesis can provide an exact solution, or can be executed as a heuristic by imposing a time limit to solve the problem. For all the problems presented in this thesis, the model was executed with no time limit. However, the tolerance to optimality was set to 1%, which we considered adequate for the purposes of this work. Regarding the computation times, it is worth mentioning a general range for both bulk and packaged problems. A standard bulk products problem contains around 325 000 variables, and the computation time ranged from 30 to 60 seconds. A standard packaged products problem contains around 2 200 000 variables, and computation time ranged from 120 to 180 seconds. The model was executed on a standard 4Mb-RAM laptop using Windows 7. No more detail about computation times will be provided, since these times are very low and do not have any impact on our objectives.

5.4. Selected Scenarios to Evaluate

With the model adapted to our case and the matching tool prepared, the final step before executing the model is defining the scenarios to evaluate. Each scenario represents a possible future event that we want to analyze. Scenarios are also the elements that determine the data to be entered into the model. Putting the right data ensures a right execution of the model and a reliable solution. In our case, the Company selected the scenarios to evaluate, which are considered as feasible from a qualitative standpoint. The scenarios to evaluate are described in the next two sections.

Note that we analyzed each scenario separately. This means that we prepared the input data and executed the model once for each scenario. The scenarios defined here are not related to the scenarios and probabilities that we defined before in the input file (see Figure 12). To each event to evaluate, we always assume only one scenario with a probability of 100%.

5.4.1. Bulk Products

For bulk products, one objective was to find the warehouses to open from the set of candidate locations. The selection of candidate locations satisfying requirements narrowed down the set to four warehouses apart from the plant. The locations are Burnaby, Denver, Los Angeles, and Memphis. Moreover, decisions related to the assignment of customers and products to warehouses, and quantities to flow and store in every warehouse needed to be solved, as mentioned in Section 1.2.2. From this, we can conclude that there was only one scenario to execute. Therefore, all the data collected needed to be put into the model to obtain the optimal solution.

Although there was only one scenario to execute with respect to locations, there was a parameter that required executing the model many times. The parameter carrying uncertainty was the fixed cost of opening a tank for each product, as it could change depending on the negotiations with facility suppliers for example. To cope with this uncertainty a range of possible values for the fixed cost was defined. To evaluate the impact of this variation on the results, the model needed to be executed once for each possible value that this fixed cost could take. By doing this, a sensitivity analysis was represented and more information to support decisions was obtained. The same had to be done for all the parameters for which a sensitivity analysis was performed.

5.4.2. Packaged Products

Unlike bulk products, there were different location scenarios to evaluate for packaged products. Even if there is a set of five candidate locations, the Company is looking to operate with three or maximum four facilities. Besides, some locations are mutually exclusive; this means that some locations cannot be opened simultaneously. To cope with this situation, we separated locations in conflict to avoid their representation together in one solution. Every scenario contains a possible combination of candidate locations, where the objective is to identify the most optimal in terms of cost. The selected scenarios to evaluate are presented in Figure 18.

Figure 18: Scenarios to Evaluate for Packaged Products

Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
Ontario	**Ontario**	**Ontario**	**Ontario**	**Ontario**
Memphis	**Memphis**	**Memphis**	**Memphis**	**Memphis**
Denver	**Portland**	**Denver**	**Salt Lake**	**Los Angeles**
Los Angeles	**Los Angeles**	**Portland**		

For each scenario, we accommodated the data and created the text file by using the matching tool as presented in Chapter 5. Every time we prepared a scenario, many data sorting tasks had to be performed. Section 5.4.3 is devoted to all the arrangements done to sort the data of this case. Once the data was ready to be used, and the txt files were generated with the Excel tool, the model had to be executed for every scenario in order to compare the resulting total cost of each scenario. The results provided by the model are presented in Chapter 6.

5.4.3. Data Sorting and Transformation

Evaluating different scenarios requires different input files to be read by the model. To create every txt file, data needs to be entered into all the spreadsheets of the Excel tool presented in Section 5.2. As data provided by companies usually comes from data warehouses or transactional systems, many changes have to be performed to obtain the right numbers. This situation, typical when coping with real data, was not an exception in this case.

To obtain most of the data presented in Section 4.2, we had to transform the data before it could be used. Among the main changes it is worth mentioning those related to products demand, fixed and variable costs, and turnover ratio.

Firstly, regarding demanded quantities, we needed to have the values for each of the 712 products (99 bulk and 613 packaged) per period. As the ten-year forecast did not require the same level of detail, the demand had been aggregated by family of products. Products are part of the same family when there is no significant variation in their use attributes. As a customer might replace one product by another in the same family according to its need, what is relevant for planning is the total demand of the family. In some cases the product is the same, and the only difference is the SKU; however, a different code makes it a different product for modeling purposes. To handle this issue we had to disaggregate the data per family to obtain values per product. Historical data from the Company's transactional system were taken into consideration. Also estimated growth factors per period and per family had to be applied to each product to ensure consistency in the data. Of course this required permanent validation to guaranty that total values were not affected. After splitting demand, for example, entire numbers became fractions, and a natural rounding process resulted in a total value that did not match the initial one. Even if it was a minor difference, we tried to avoid any inconsistency in the initial data.

Secondly, with respect to costs, all the values assigned to parameters are required to be expressed in the same unit of measurement, which is dollars per USG. However, transport suppliers rates are often given per mile. Moreover, depending on the state of origin, which is the location of the warehouse, the supplier may change from one state to the other. This entails to deal with different rate structures and formats that had to be unified. In some cases, where more than one supplier provided rates, they had to be aggregated according to the indications from the Company. Similarly, the holding and handling cost presented a different structure depending on the supplier. As we were dealing with new locations without previous rate experience, quotes were the main source of these costs. Some of these holding quotes were in pallets, squared feet, or pounds. All these had to be converted to USG to make them comparable among themselves. Apart from cost rates conversion, some formulas were needed to sort the data. For example, to identify the customers located within a 160 miles distance from the new warehouses, a validation was performed before assigning the zero cost of transportation to products shipped to the customer. As mentioned in Section 4.2.1, all the distances were calculated based on the geographical coordinates of candidate warehouses and the 1 062 customers (168 for bulk and 894 for packaged products). All these calculation were performed and validated in collaboration with the project team of the Company responsible for the project.

Finally, the turnover ratio for each product resulted from analyzing the historical data per product, the Company objectives and policies for the long term, and the level of service desired by customers. This shared analysis led to the turnover ratio that was used in the model. Apart from the data transformation examples we mentioned here, further calculations had to be performed to ensure integrality and consistency in the data entered into the model. When modeling a supply chain, quality data is a key aspect to ensure a reliable result.

6. Results and Analysis

This chapter presents the output yielded by the model for the scenarios presented in the prior chapter. Results are analyzed to identify the characteristics of each scenario and a sensitivity analysis is performed over some of the costs in the case. Finally, the results for packaged products scenarios are presented and compared among themselves.

6.1. Description of the Output File

Among the files obtained after executing the model, there is a text file with the same name as the input file, but identified with the extension .sol. This solution file can be opened as a text file and then converted into an Excel spreadsheet. The output file is divided in eight sections which are: suppliers, plants, warehouses, sources, modes, flows, inventories, and total cost. The first three sections list the nodes in the network that have been selected in the solution to perform an activity. In front of every node appears the product to supply, to manufacture, or to store, depending on the nature of the node. The section sources list all the nodes in the supply chain that have been selected to dispatch items. The destination and the products to be shipped from that node appear in front of every source. The section mode indicates the modes of transportation selected to operate within the supply chain. Section six, flows, presents the quantity of product flowing from every selected origin point to every active destination. It indicates the mode of transportation and quantity for every period in the problem. Section seven provides the quantity to have in inventory at the end of each period, and finally, the total cost of operating the entire supply chain for all the periods is presented in the last line.

6.2. Results and Analysis for Bulk Products

In order to summarize the results, the data from the sections “flow” and “inventory” were analyzed using pivot tables. These tables contain most of the relevant information we expected to obtain after executing the model. Here we mention again one of the objectives of this work, as it was stated in Section 1.3:

- Selecting the location of DCs, identifying the optimal assignment of customers to DCs, the assignment of both bulk and packaged products to DCs, determining the optimal transportation mode, and defining the quantity of product to store in every DC in order to estimate the size of the facilities.

First, the pivot table of flow data is analyzed to obtain aggregated information. Table 18 presents an extract of the resulting flow table aggregated by source.

Table 18: Extract of Flow Table Aggregated by Source

Sum of Quantity	Periods					
Source	Per1	Per2	Per3	Per4	Per5	...
⊞Burnaby	483 228	591 785	676 116	777 129	861 827	...
⊞Denver	740 327	930 688	1 078 568	1 255 698	1 404 223	...
⊞Los Angeles	857 059	1 147 271	1 372 722	1 642 763	1 869 197	...
⊞Memphis	1 926 262	2 577 170	3 082 823	3 688 486	4 196 347	...
⊞Ontario	14 827 746	18 540 978	21 425 594	24 880 718	27 777 907	...
Total	18 834 621	23 787 892	27 635 824	32 244 794	36 109 500	...

USG

Reported values are not real but have been scaled to protect confidentiality.

The flow table presents the total quantity of USG to dispatch from every warehouse at each period. Although the plant is also a source, it has been deselected in the pivot table. The reason is that, with only one plant, we know already that it has to be the only source for all the products, thus there is no a relevant decision related to the plant. The aggregated quantity in front of every warehouse considers the sum of all the products to be shipped from every node to all the customers. Even if the information is aggregated, the table allows drawing a first conclusion. We can see that the five candidate locations for the warehouses appear as a source for all the periods in the table. This, of course, implies that all the warehouses have to be opened as part of the solution, and this provides the answer to the first point. We observe also a concentration of 78.7% of products to be dispatched from Ontario. In Section 4.2.1, the transportation cost from the plant to the warehouses was presented in Table 10. In this table the cost from the plant to the warehouse in Ontario was zero, as they are located in the same place. This makes the warehouse more efficient, and it justifies in part the concentration of products in this facility. However, the concentration in Ontario depends also on the demand around that area, since in the end the total cost is taken into account and not only the cost from the plant to the DCs. In addition, Table 12 shows also that holding and handling costs are lower in Ontario than in any other location. This further supports the reasons that benefit Ontario over other facilities even if distance to customers is longer. However, the trade-off between inventory cost and transportation cost is sometimes in favor of opening another facility. This happens since it is sometimes better to store products at a facility to then cover a shorter path to reach the customers. For example, Burnaby was selected to dispatch 2.5% of the products to customers. This means that for 2.5% of the total volume it is cheaper to ship products to Burnaby by rail, and then to store and to handle them there before reaching the customers, than it is when dispatching directly from Ontario. Identify that small fraction that leads to a lower total cost is easily done by using CPLEX solver.

To continue with the analysis of the information yielded by the model, it is necessary to go deeper in the detail of the table. By using the functions of the pivot table we select to ungroup the data of Denver as an example. Table 19 presents an extract of the resulting flow table for this warehouse.

Table 19: Extract of Flow Table with Destination Detail for Bulk Products

Sum of Quantity	Periods					
Source	Per1	Per2	Per3	Per4	Per5	...
⊞Burnaby	483 228	591 785	676 116	777 129	861 827	...
⊟Denver	740 327	930 688	1 078 568	1 255 698	1 404 223	...
⊞Cus12	366 358	428 713	477 153	535 174	583 825	...
⊞Cus28	16 192	18 512	20 314	22 471	24 281	...
⊞Cus35	610	695	760	839	906	...
⊞Cus59	305	347	380	420	453	...
⊞Cus77	177 594	207 821	231 303	259 428	283 013	...
⊞Cus93	72 947	85 363	95 009	106 561	116 248	...
⊞Cus112	106 321	189 237	253 650	330 803	395 497	...
⊞Los Angeles	857 059	1 147 271	1 372 722	1 642 763	1 869 197	...
⊞Memphis	1 926 262	2 577 170	3 082 823	3 688 486	4 196 347	...
⊞Ontario	14 827 746	18 540 978	21 425 594	24 880 718	27 777 907	...
Total	18 834 621	23 787 892	27 635 824	32 244 794	36 109 500	...

USG

Reported values are not real but have been scaled to protect confidentiality.

When we include the destination of products in the pivot table, we can see the volume to be dispatched from the warehouses to every customer. In the example of Table 19, we see the volume to be dispatched from Denver to each customer assigned to this warehouse. This means that seven customers have been assigned to Denver; therefore, we have here the answer to the second point. The detail of every warehouse contains the set of customers to be served from every location. If we assume that there is a relationship between the total volume and the number of customers, we can suppose that most of the customers have been assigned to Ontario, which would be the warehouse responsible for dispatching the highest volume in the supply chain.

Another point we want to answer involves also selecting the mode of transportation that ensures the minimal cost. That information is present in both the modes section and flow section of the output file. However, for bulk products there is only one mode of transportation connecting two nodes. Thus, the output file just confirms that rail is the mode of transportation to use between the plant and the warehouses, and that the mode of transportation to use when shipping the products from the warehouses to the customers should be by truck.

Table 20: Extract of Flow Table with Product Detail for Bulk Products

Sum of Quantity	Periods					
Source	Per1	Per2	Per3	Per4	Per5	...
⊞Burnaby	**483 228**	**591 785**	**676 116**	**777 129**	**861 827**	...
⊟Denver	**740 327**	**930 688**	**1 078 568**	**1 255 698**	**1 404 223**	...
BulkP12	307 831	376 066	429 074	492 566	545 805	...
BulkP17	59 261	76 344	89 615	105 511	118 840	...
BulkP19	80 574	107 476	128 375	153 407	174 397	...
BulkP34	85 221	107 730	125 217	146 161	163 724	...
BulkP54	138 840	182 796	216 943	257 843	292 139	...
BulkP77	38 456	45 002	50 086	56 176	61 284	...
BulkP85	30 143	35 274	39 259	44 033	48 036	...
⊞Los Angeles	**857 059**	**1 147 271**	**1 372 722**	**1 642 763**	**1 869 197**	...
⊞Memphis	**1 926 262**	**2 577 170**	**3 082 823**	**3 688 486**	**4 196 347**	...
⊞Ontario	**14 827 746**	**18 540 978**	**21 425 594**	**24 880 718**	**27 777 907**	...
Total	**18 834 621**	**23 787 892**	**27 635 824**	**32 244 794**	**36 109 500**	...

USG

Reported values are not real but have been scaled to protect confidentiality.

Table 20 presents the flow table when including the products in the pivot table instead of destination. Here, we can see the products to be dispatched from Denver, to the customers assigned to this warehouse. This implies that those seven products have to be stored at that location, thus it would be necessary to have seven tanks in Denver. What we have here is an answer to another point of our problem, the assignment of products to the DCs. This is valuable information for our case, since we can now have a reference for the size of the resulting supply chain. We know the location and the number of tanks required at each location.

Moreover, we know the customers to serve from each facility, and the volume of products flowing through every DC. This allows the development of a short-term plan of distribution, and also helps in having an active negotiation with suppliers of transportation and facilities. Even if we know this configuration provides the minimal cost with current rates, the cost could be reduced more by negotiating, while taking support in this piece of information. The last part of the problem to figure out in the case is related to the amount of inventory required at every location. Inventory information is available in the section seven of the output report.

Table 21: Extract of Inventory Table Aggregated by Warehouse

Sum of Inventory	Periods					
Warehouse	Per1	Per2	Per3	Per4	Per5	...
⊞ **Burnaby**	**39 721**	**48 645**	**55 577**	**63 880**	**70 842**	...
⊟ **Denver**	**178 996**	**223 828**	**258 941**	**301 786**	**337 713**	...
BulkP12	66 492	81 230	92 680	106 394	117 894	...
BulkP17	21 968	28 301	33 220	39 113	44 054	...
BulkP19	19 451	25 945	30 990	37 033	42 099	...
BulkP34	20 155	25 478	29 614	34 567	38 721	...
BulkP54	32 836	43 231	51 307	60 980	69 091	...
BulkP77	9 095	10 643	11 845	13 286	14 494	...
BulkP85	9 000	9 000	9 285	10 414	11 360	...
⊞ **Los Angeles**	**125 796**	**160 134**	**187 739**	**221 223**	**250 155**	...
⊞ **Memphis**	**245 461**	**322 088**	**384 699**	**459 694**	**522 579**	...
⊞ **Ontario**	**1 218 841**	**1 524 068**	**1 761 184**	**2 045 195**	**2 283 344**	...
Total	**1 808 815**	**2 278 763**	**2 648 140**	**3 091 779**	**3 464 633**	...

USG

Reported values are not real but have been scaled to protect confidentiality.

Table 21 is a pivot table generated from the inventory data from the output report. Inventory represents the optimal volume to have in stock at the end of each period. To continue with the same example, the table presents Denver ungrouped to see the detail of the inventory level required per product at that facility. This information completes the objective we were looking for since the starting point of the case. Moreover, defining the size of each tank and being prepared for future upgrades becomes a simpler task with this forecast for the next ten years. Note that for

BulkP85 period 1 and period 2 the amount to have in inventory is the same. This is due to the lower bound imposed as a constraint for the tanks in Section 4.3. As the constraint is defining the quantity to have in stock for those two periods, it is reasonable to assume that the optimal quantity to store without this constraint is lower than 9 000 USG. In other words, if we removed the lower bound constraint, the result would indicate a lower volume of BulkP85 to have in stock in period 1 and period 2. This claim means that due to the constraint, Denver is carrying more inventory than required, which entails a higher holding and handling cost.

Finally, the last line of the output file presents the total cost of operating the entire supply chain. This value is a reference when comparing the results for multiple configurations of the network. With only one scenario performed for bulk products, we focused the analysis on the composition of the total cost.

Table 22: Total Cost Composition for Bulk Products

Type of Cost	Total Cost Fraction (%)
Transportation cost from DCs to customers	63.44
Handling cost	13.99
Holding cost	10.71
Transportation cost from the plant to DCs	8.73
Tank preparation cost	3.13
Total	100.00

Total cost = 100%

Table 22 presents the percentage that each type of cost represents of the total. The table was ranked in decreasing order. For practical purposes the total cost equals 100%.

Table 22 shows that more than 60% of the total cost of the supply chain corresponds to the freight rate of delivering the products to the customers. Then 24.7% of the cost is related to inventory activities, which take place in the DCs. Transportation cost to ship the products by rail from the plant to DCs accounts for

8.73%, and finally the fixed cost of opening a tank is the lowest in the ranking. Regarding the transportation cost, we see that the total cost for trips done by truck is more than 7 times the total cost for the trips done by rail. There are three main reasons that explain this difference. First, the cost of moving one USG by rail is significantly lower than moving it by truck. From Table 10, which presented the rail cost, we can calculate an average cost per USG of 0.88 dollars. This cost is higher than most of the cost per USG presented in Table 11. The difference can be attributed to the transportation mode rates offering by transportation suppliers, and to the economies of scale for moving larger volume of products. The second reason is that the transportation cost from the plant to the DC in Ontario is 0, since both the plant and the warehouse are located in the same place. This zero cost can also be seen in Table 10. If we observe the volume of product flowing through Ontario in Table 20, we can see that more than 78% of the total volume is dispatched from the warehouse in Ontario. This means that for 78% of the products, there is no transportation cost for the trip from the plant to the warehouse. The third reason that makes the transportation cost by truck higher than the transportation cost by rail is related to distances. The total distance to travel from the DCs to the customers is longer than the total distance to travel from the plant to the DCs in miles. A final remark of Table 22 is the low participation of the fixed cost. Although tank preparation cost is significant for the operation, this is a cost that is present only in period 1. For this reason, its participation seems minor in the ten-year horizon, when compared to the other types of cost.

6.2.1. Fixed Cost Sensitivity Analysis

Usually, the fixed costs for warehouses are related to facilities construction, which require huge investments of resources at the beginning of the project. Consequently, the fixed costs are a key factor when evaluating projects of this magnitude. In our case, due to the renting facility strategy chosen for this project, the fixed costs are related to tank preparation before holding a product. It is worth to mention that there is not fixed cost associated to the facility, but for every new tank required for the operation. Nevertheless, the impact of the variation of this fixed cost on the results has to be evaluated. Each value to evaluate entails running the model with a new value for parameter c_o^k. To prepare the text input file, only the section of warehouses needs to be executed in the Excel matching tool. Table 23 presents the values of fixed costs to be evaluated.

Table 23: Fixed Cost Values to Evaluate

	Variation	Burnaby	Los Angeles	Denver	Memphis
Possible Variation of Fixed Cost in the Case	30% Lower	73 150	85 400	75 950	81 410
	Current Fixed Cost	104 500	122 000	108 500	116 300
	30% Higher	135 850	158 600	141 050	151 190
Hipotetical Variation of Fixed Cost for Academic Purposes	50% Higher	156 750	183 000	162 750	174 450
	100% Higher	209 000	244 000	217 000	232 600
	200% Higher	313 500	366 000	325 500	348 900
	500% Higher	627 000	732 000	651 000	697 800
	1000% Higher	1 149 500	1 342 000	1 193 500	1 279 300

US Dollars

Reported values are not real but have been scaled to protect confidentiality.

The variations affect warehouses other than in Ontario, where the fixed cost remains the same since there is no uncertainty with respect to this location. Although possible variations in real circumstances are expected to range about 30% above and below the initial estimation, some additional values have been evaluated for academic purposes. First, the impact of reducing and increasing the fixed cost by 30% is presented. The effects of this change are evaluated in the inventory section of the output file. Table 24 presents how products assignment reacts when the fixed cost varies.

Table 24: Impact of Fixed Cost Change on Products Assignment

30% Lower

Sum of Inventory	Periods		
Warehouse	**Per1**	**Per2**	**...**
Burnaby	**39 721**	**48 645**	**...**
Denver	**191 127**	**237 281**	**...**
BulkP12	66 492	81 230	...
BulkP17	21 968	28 301	...
BulkP19	19 451	25 945	...
BulkP21	12 131	13 453	...
BulkP34	20 155	25 478	...
BulkP54	32 836	43 231	...
BulkP77	9 095	10 643	...
BulkP85	9 000	9 000	...
Los Angeles	**134 391**	**168 729**	**...**
Memphis	**254 056**	**330 683**	**...**
Ontario	**1 202 325**	**1 519 927**	**...**
Total	**1 821 620**	**2 305 265**	**...**

30% Higher

Sum of Inventory	Periods		
Warehouse	**Per1**	**Per2**	**...**
Burnaby	**39 721**	**48 645**	**...**
Denver	**169 996**	**214 828**	**...**
BulkP12	66 492	81 230	...
BulkP17	21 968	28 301	...
BulkP19	19 451	25 945	...
BulkP34	20 155	25 478	...
BulkP54	32 836	43 231	...
BulkP77	9 095	10 643	...
Los Angeles	**114 391**	**168 729**	**...**
Memphis	**234 056**	**330 683**	**...**
Ontario	**1 230 665**	**1 526 637**	**...**
Total	**1 788 829**	**2 289 522**	**...**

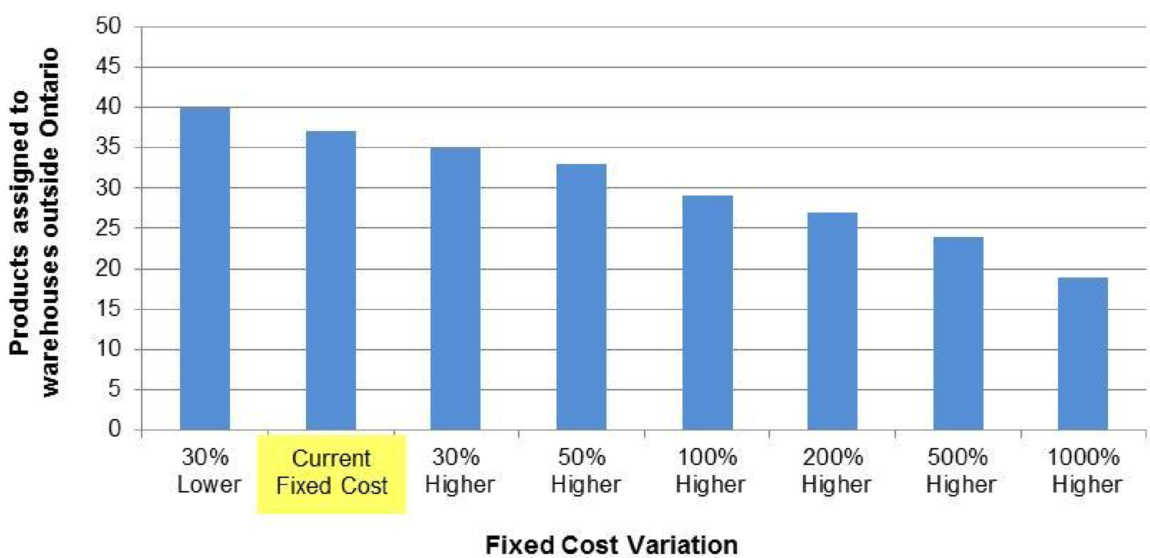
USG

Reported values are not real but have been scaled to protect confidentiality.

On the left side of Table 24 we see the effect of reducing the fixed cost for all the locations by 30%, while on the right side the effect of increasing the cost. The changes have to be compared with the information presented in Table 21, which presents the assignment of products to every warehouse with the initial fixed cost. By analyzing the changes in the products assigned to Denver, we can understand what occurs in the entire supply chain. When reducing the fixed cost, all the products initially at this location remain, but a new product is assigned to Denver. Product BulkP21 is now to be stored at this location, thus requiring a new tank for the product. An extra tank entails a higher fixed preparation cost. However, reducing the fixed cost makes opening a tank a better option than dispatching the product from another location in terms of cost. Assigning product BulkP21 to Denver implies that the quantity assigned, or perhaps the whole product, has been removed from another facility. As similar situations could happen to other products, it is difficult to track where the product was assigned before with the data presented in the tables. However, we know that Burnaby does not present any change in quantities, and that the other warehouses increased also the quantity to store. As the only warehouse reducing the total quantity in stock is Ontario, we can conclude that product BulkP21, was initially assigned to Ontario. Therefore, we can conclude that reducing the fixed cost encourages dispersing the product far from Ontario.

On the right side of Table 24, we can see that the assignment of products to Denver changes when the fixed cost increases. Fewer products are now assigned to this location reducing the total quantity stocked in Denver. Quantities in other locations also diminish while Ontario is the only one increasing, a fact that supports our conclusion. We can also say that increasing the fixed cost encourages concentration of products in Ontario.

Figure 19: Impact of Fixed Cost Change on Warehouses



To test our preliminary conclusion, the model was executed with all the values of fixed costs presented in Table 23. Figure 19 shows how the number of products assigned to other warehouses than Ontario reduces, when the fixed cost increases. This means fewer tanks to be opened and lower fixed cost; therefore, the initial conclusion regarding the variation in the fixed cost is true. The higher the fixed cost is, the fewer tanks to open in warehouses outside Ontario. At some point, some warehouses are not even part of the solution, and the locations in the solution go from five to four, and so on. We can predict that if the fixed cost were even higher, at some point all the products would be stored and dispatched from Ontario. It is worth mentioning that this is supported by the fact that the fixed cost in Ontario remains equal, due to the special condition of this location. Ontario is the only facility that is currently operating apart from being the location of the plant, thus the uncertainty level is lower. Continuing with the analysis of the result, with the magnitude in the variation of the fixed cost we simulated, it is easier to comprehend that any reassignment of product between warehouses changes the entire configuration of the network. Since products have to be relocated, the assignment of customers to warehouses changes as well as the quantities moved from the plant to the warehouses, and consequently the transportation cost. Finally, a point that arises from this analysis is that total quantities of USG to have in inventory in the entire supply chain are never the same. This can be observed in the total amount in inventory presented in Table 21 and Table 24. One could think that if products are simply relocated from one warehouse to another, the quantity removed from the first is equal to the quantity assigned to the second. However, this does not happen because the turnover ratio of every product varies from one location to the other, leading to a different level of inventory.

Finally, we present the resulting total cost composition for each of the fixed cost, which were presented in Table 23. Only the values intended for academic purposes were performed for this analysis.

Table 25: Total Cost Composition for Every Change in Fixed Cost

Type of Cost	Current Fixed Cost	30% Higher	50% Higher	100% Higher	200% Higher	500% Higher	1000% Higher
Transportation cost from DCs to customers	63.44	65.12	66.78	68.29	71.10	75.82	80.72
Handling cost	13.99	12.87	12.07	11.73	10.43	8.16	7.15
Holding cost	10.71	10.12	9.63	9.11	8.52	7.11	5.81
Transportation cost from the plant to DCs	8.73	8.65	8.33	7.91	7.12	6.62	4.29
Tank preparation cost	3.13	3.24	3.19	2.96	2.83	2.29	2.03
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Total cost = 100%

Table 25 presents the composition of the total cost, for every result obtained after executing the model with a different fixed cost. In this table, we can see that when the fixed cost is increased, the participation of the fixed cost in the total cost decreases. One could think that a higher fixed cost would increase its participation in the total cost. However, a smaller number of products assigned outside Ontario, ended up by reducing slightly the fixed cost participation. Note that except for the transportation cost from DCs to customers, all the other costs experienced a reduction in their participation. This is reasonable because fewer tanks outside Ontario means lower volume to transport to other locations, and lower volume to have in inventory. Like this, many other simulations related to fixed cost could be done by using the Excel matching tool and the model. Having more information about possible situations contributes to the supply chain decision-making process. Now, we present the impact of the results when varying other parameters.

6.2.2. Transportation Cost Sensitivity Analysis

Transportation cost is usually in all the tiers of the supply chain. In our case there is transportation from the plant to the warehouses, and from the warehouses to the customers. These two costs were presented in Table 10 and Table 11, respectively. As the first trip departs from Ontario, where the Company has currently the warehouse to replenish warehouses in the United States, there is almost no uncertainty in these rates. In contrast, the trip rates from candidate locations to customers may vary. An option to reduce uncertainty could be negotiating a standard rate with a corporative supplier. This could bring savings to the Company, and the carrier would ensure volume while being an exclusive supplier for the Company. To identify the impact of this option, we executed the model with a uniform rate for truck transportation. The selected rate is the cost per mile when departing from Ontario, which is 13.22 dollars. This rate is the minimal unit of Table 11. The resulting flow of products after executing the model with the flat rate is presented in Table 26.

Table 26: Impact of Transportation Cost Reduction on Flowing Quantities

Initial Transportation Cost

Sum of Quantity	Periods		
Source	Per1	Per2	...
Burnaby	**483 228**	**591 785**	...
Denver	**740 327**	**930 688**	...
BulkP12	307 831	376 066	...
BulkP17	59 261	76 344	...
BulkP19	80 574	107 476	...
BulkP34	85 221	107 730	...
BulkP54	138 840	182 796	...
BulkP77	38 456	45 002	...
BulkP85	30 143	35 274	...
Los Angeles	**857 059**	**1 147 271**	...
Memphis	**1 926 262**	**2 577 170**	...
Ontario	**14 827 746**	**18 540 978**	...
Total	**18 834 621**	**23 787 892**	...

Unified Transportation Cost

Sum of Quantity	Periods		
Source	Per1	Per2	...
Burnaby	**483 228**	**591 785**	...
Denver	**809 661**	**1 013 448**	...
BulkP12	307 831	376 066	...
BulkP17	59 261	76 344	...
BulkP18	13 456	16 416	...
BulkP19	80 574	107 476	...
BulkP34	85 221	107 730	...
BulkP54	138 840	182 796	...
BulkP61	21 567	25 233	...
BulkP67	19 876	24 156	...
BulkP77	38 456	45 002	...
BulkP81	14 435	16 954	...
BulkP85	30 143	35 274	...
Los Angeles	**1 147 916**	**1 547 271**	...
Memphis	**2 785 343**	**3 707 070**	...
Ontario	**13 608 474**	**16 928 318**	...
Total	**18 834 621**	**23 787 892**	...

USG

Reported values are not real but have been scaled to protect confidentiality.

In Table 26 we can see the impact in Denver, after unifying the truck transportation cost. With four new products to store in that location, the total volume of products to flow experienced an increase of more than 9%. Similarly, quantities in Los Angeles and Memphis increased, while Ontario is the only location reducing its volume. From this, we can conclude that the assignment of products to warehouses is very sensitive to changes in transportation costs. Besides, we can draw the conclusion that the concentration of products in Ontario responds partially to the lower transportation cost of this location. When the transportation cost rate is the same in every location, the optimal solution entails opening more tanks. What is interesting here is that every different transportation cost leads to a new configuration. We performed other scenarios with a flat rate equal to the cost per mile offered in other states, and there is not a visible trend. Every different cost yields a new supply chain configuration. This can be explained by all the interactions among the costs that make impossible to predict the outcome.

6.2.3. Inventory Cost Sensitivity Analysis

In the presented results we have seen a clear concentration of products flowing through Ontario. Among the reasons we mentioned the zero transportation cost from the plant to the warehouse, the demand concentrated around Ontario, and its low inventory cost with respect to other warehouses. In table 12, we saw that the holding cost of Ontario is less than a half of the cost of the next cheapest location. Similarly, the handling cost is significantly lower in Ontario than in the other warehouses. Here, we want to evaluate the impact of losing this cost advantage in Ontario. In other words, we wanted to know what would happen if the inventory cost in Ontario gets closer to the cost of the other warehouses. To evaluate this, we executed the model with some holding and handling costs ranging from Ontario's initial costs (Ontario 0 in Table 27) to the average inventory costs of the other warehouses. On the left side of Table 27 the average inventory cost is calculated. On the right side, the 5 different cost scenarios executed in the model are presented.

Table 27: Inventory Cost in Ontario to Evaluate

Warehouse	Holding Cost per USG	Handling Cost per USG
Burnaby	1.0858	0.2062
Los Angeles	1.2676	0.1354
Denver	1.1273	0.1275
Memphis	1.2084	0.1134
Average	**1.1723**	**0.1456**

→

Ontario Warehouse	Holding Cost per USG	Handling Cost per USG
Ontario 0	**0.4123**	**0.0820**
Ontario 1	0.5360	0.0984
Ontario 2	0.6597	0.1148
Ontario 3	1.1256	0.1312
Ontario 4 (Other Warehouses Average)	**1.1723**	**0.1456**

US Dollars

Reported values are not real but have been scaled to protect confidentiality.

The concentration of products is considered as the percentage of the quantity flowing through Ontario with respect to the total. The higher this percentage is, the lower the dispersion of products will be in the other warehouses. Figure 20 presents the scenarios evaluated, as well as the results.

Figure 20: Impact of Inventory Cost Change on Products Concentration

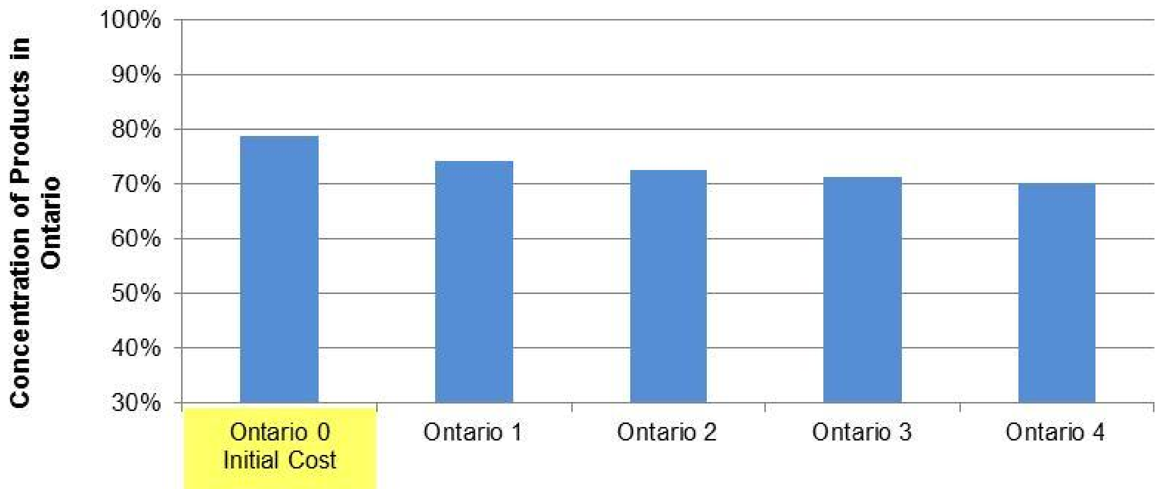


Figure 20 shows that in the initial situation, 79% of the products flow through the warehouse in Ontario, while the other 21% is distributed among the other warehouses. This is obtained from the resulting flow quantities of the solution file. The results with the initial inventory cost were presented in Table 18. By using the same data, the new quantities and percentages were obtained. New scenarios results show that when the inventory cost increases in Ontario, the concentration of products flowing through Ontario decreases. When inventory cost equals the average cost of the other warehouses (Ontario 4), the concentration decreases to 70%. This confirms that the concentration of products in Ontario is due, in part, to the low inventory cost; however, the slight variation demands more analysis. While the holding cost was increased 2.8 times, the concentration experiences a less than 10% decrease. This can be considered as almost not sensitive to the holding cost. Although inventory cost is a factor that makes Ontario be the optimal warehouse in most of the cases, it is not the main reason. Transportation cost produces a higher impact and becomes the main reason for product concentration in Ontario. In Table 26, by simply unifying the truck transportation cost, which represented a variation of 18%, the concentration of products in Ontario became 72%. Additionally, the zero cost from the plant gives another relevant advantage. Finally, the location of customers near to Ontario, explains why this warehouse is part of the optimal path.

6.3. Results and Analysis for Packaged Products

Results of packaged products contain the same structure as the one presented for bulk products. With more customers and more products, the time required to obtain the solution is slightly higher, but the solution file is much larger. Before going further in the data of the solution, we first need to identify the scenario yielding the minimum cost. The scenarios presented in Section 5.4.2 were executed with the model in order to obtain the total cost from the solution file. For analysis purposes, the total cost of the optimal scenario has been equaled to 100%. Figure 21 presents the results.

Figure 21: Total Cost of Scenarios Evaluated for Packaged Products

Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
Ontario	**Ontario**	**Ontario**	**Ontario**	**Ontario**
Memphis	**Memphis**	**Memphis**	**Memphis**	**Memphis**
Denver	**Portland**	**Denver**	**Salt Lake**	**Los Angeles**
Los Angeles	**Los Angeles**	**Portland**		
107.5 %	100.0 %	114.7 %	112.6 %	118.5 %

Scenario 2 presents the minimal cost, which has been adjusted to 100% to make the task of comparing the results easier. The next more economical option is scenario 1, which is 7.5% more costly than the optimal one. The other options are even higher. Scenario 5, for example, is 18.5% over optimality. This shows how options that could have been considered as interesting based on qualitative aspects, turn out to be options to be discarded. Even a few percentage points may be relevant when making long-term decisions about the supply chain to implement. However, it does not mean scenario 2 is the option to implement. When considering three warehouses the best option would be Scenario 4. Even if it is 12.6% higher than the optimal one, this scenario has to be considered by the decision makers. Dealing with three warehouses instead of four could produce operation benefits and hidden economies. Let's remember that the objective of a mathematical model is providing insights and not making decisions.

Regarding the solution details of scenario 2, we have the flow data and inventory data as main sources of information. From these, we can obtain most of the answers to the points stated in Section 1.3. First, we present an extract of flow pivot table aggregated by source.

Table 28: Extract of Flow Table for Packaged Products

Sum of Quantity	Periods					
Source	**Per1**	**Per2**	**Per3**	**Per4**	**Per5**	**...**
Los Angeles	**1 031 401**	**1 414 941**	**1 798 428**	**2 181 958**	**2 565 514**	**...**
Memphis	**6 549 152**	**7 408 830**	**8 268 497**	**9 128 255**	**9 987 948**	**...**
Ontario	**2 117 282**	**2 362 701**	**2 608 132**	**2 853 540**	**3 098 882**	**...**
Portland	**581 287**	**647 073**	**712 902**	**778 718**	**844 548**	**...**
Total	**10 279 123**	**11 833 545**	**13 387 959**	**14 942 471**	**16 496 892**	**...**

USG

Reported values are not real but have been scaled to protect confidentiality.

Count Destinatio	Periods					
Warehouse	**Per1**	**Per2**	**Per3**	**Per4**	**Per5**	**...**
Los Angeles	98	98	98	98	98	...
Memphis	525	525	525	525	525	...
Ontario	240	240	240	240	240	...
Portland	30	30	30	30	30	...

Quantity of Products

From the upper part of Table 28, we know the distribution of volume to flow through every warehouse, where the solution requires using the four locations. Due to the high number of customers demanding packaged products (894), the pivot table in the lower part was built to indicate the number of customers assigned to every warehouse. For example, the 1 031 401 USG to pass through Los Angeles in period 1, will be dispatched to 98 customers. We can also see that the number of customers assigned to every warehouse does not vary overtime. Additionally, from the solution file, we know that the transportation mode to replenish the warehouses is rail, while truck has to be used to take the products to the customers.

As another question is related to facility sizing, we need the inventory volume to have in stock at every facility. Let's remember that for packaged products, the facilities can be shared and the products can be stored together. An extract of the inventory pivot table aggregated by warehouse is presented in Table 29.

Table 29: Extract of Inventory Table for Packaged Products

Sum of Inventory Periods						
Warehouse	**Per1**	**Per2**	**Per3**	**Per4**	**Per5**	**...**
Los Angeles	**42 287**	**58 013**	**73 736**	**89 460**	**105 186**	**...**
Memphis	**268 515**	**303 762**	**339 008**	**374 258**	**409 506**	**...**
Ontario	**86 809**	**96 871**	**106 933**	**116 995**	**127 054**	**...**
Portland	**23 833**	**26 530**	**29 229**	**31 927**	**34 626**	**...**
Total	**421 444**	**485 175**	**548 906**	**612 641**	**676 373**	**...**

USG

Reported values are not real but have been scaled to protect confidentiality.

Count of Invento	**Periods**					
Warehouse	**Per1**	**Per2**	**Per3**	**Per4**	**Per5**	**...**
Los Angeles	286	286	286	286	286	...
Memphis	557	557	557	557	557	...
Ontario	410	410	410	410	410	...
Portland	225	225	225	225	225	...

Quantity of Products

From the upper part of Table 29, we know the distribution of volume to have in inventory at each of the four warehouses. In the lower part, the pivot table presents the number of products to store at each location. For example, the 42 287 USG to store in Los Angeles in period 1 consists of 286 products. As we know that the total number of packaged products is 613, it is easy to perform more calculations. For example, even if Memphis accounts for 63.7 % of the total volume, 91% of the products should be available at that location.

When analyzing the solution of the model per product, there is an issue to adjust related to quantities. As all the data of the model was entered in USG, the output is of course in the same unit of measurement. Unlike bulk products, packaged products have some defined pack sizes for every product that have to be respected when moving and storing the products.

Figure 22: Adjustment of Results to Pack Sizes

Warehouse	Per1	...		
Memphis	**268 515**	...	**Min SKU**	**Resulting Volume**
PackP2	513	...	50	=MROUND(Per1, Min SKU)
PackP3	11	...	20	20
PackP5	698	...	50	700
PackP6	77	...	40	80
PackP7	92	...	10	90
PackP8	130	...	20	140
PackP9	2 000	...	200	2 000
PackP11	370	...	80	400
PackP12	85	...	10	90
PackP14	42	...	20	40
PackP15	3	...	10	10
...	...	...	...	
...	...	...	...	

USG

Reported values are not real but have been scaled to protect confidentiality.

Figure 22 presents an extract of the products flowing through Memphis. On the left side, we can see the exact result yielded by the model for period 1; however, the quantities of the solution do not necessarily correspond to quantities that can be handled or transported. Unlike bulk products, packaged products have to be manipulated in the quantities defined by the pack size; therefore, an additional task needs to be performed in order to have a solution that can be implemented. It is important to mention, that this final task cannot be executed with the model. Although the model has lower and upper bound as possible constraints, it does not consider the size of the pack. This issue can be simply resolved by adjusting the result to the pack size that corresponds to every product. A new column Min SKU (minimal SKU) was added to each of the 894 products. This column represents the pack size of the product. Then, by using the Excel function MROUND (multiple round), the value of the solution is rounded up or down to the nearest multiple of the Min SKU. The resulting value is then presented in the right side of Figure 22. This method can be very useful when dealing with packaged products. The minimal SKU might be pallets, or any other measure defined as the minimal to manipulate the product.

6.4. Combining Products Analyses

In many sections we have mentioned that the interaction between all the costs in the problem is a key factor to reach optimality. In Section 4.2, we mentioned that the parameters were assigned independently to bulk and packaged products. Indeed, the problems themselves were executed separately. The reason argued was that there are no decisions of one type of product affecting the other. As it seems to be a slight contradiction, we opted for executing the model for both bulk and packaged products simultaneously. The result was a sum of the two independent solutions in terms of cost. Similarly, assignment of customers and products to warehouses were the same. Therefore, this confirmed that problems can be split when there is certainty of no interaction among the parameters.

However, it is worth mentioning, that for the Company, it is simpler to have fewer warehouses if possible. For that reason, some of the candidate locations were the same for the two types of products. Having the operations centralized in fewer facilities produces certain benefits that cannot be evaluated in mathematical models. Cultural differences in new locations, coordination cost, knowledge transfer cost, and decrease in short-term responsiveness are examples of hidden costs that appear when having more facilities.

6.5. Comparison with the Existing Distribution Network

Once the results have been presented and analyzed, a final question needs to be answered: What is the economic benefit of this study to the Company? In Section 1.2.1 we mentioned that optimizing the Company's current distribution network should lead to cost reductions. Now, we want to know the magnitude of the economies that might be achieved by implementing the resulting supply chain.

To evaluate the benefit, we compared the total cost of the supply chain resulting from executing the model, to the total cost of the current distribution network. We performed the calculations of the current situation with the forecasted demand, in order to compare the total cost in the ten-year horizon. We denoted the current situation as "Status quo".

Table 30: Total Cost Comparison for Bulk Products

	Total Cost Fraction (%)		
Type of Cost	**Optimal Solution**	**Status Quo**	**Difference**
Transportation cost from DCs to customers	63.44	58.56	-4.88
Handling cost	13.99	16.89	2.90
Holding Cost	10.71	19.02	8.31
Transportation cost from the plant to DCs	8.73	20.37	11.64
Tank preparation cost	3.13	0.00	-3.13
Total	100.00	114.84	14.84

Total cost of the resulting network = 100%

Table 30 presents the composition of the total cost of the optimal supply chain, as presented in Section 6.2. Again, the total cost of the optimal solution equals 100%. The next column, “Status quo” presents the composition of the total cost, if the distribution network would remain as it currently works. Finally, the difference column allows identifying the variation of each type of cost and the composition of the total cost difference, which is 14.84%. This percentage represents the difference between the total cost of the optimal solution and the total cost of the “status quo”. In other words, if the company decided to quit the project and keep the current distribution network, it would have to pay 14.84% more to distribute its products and serve its customers. This percentage can also be seen as the potential savings of implementing the optimal solution.

When we see the total cost composition of the “status quo”, and the differences with respect to the optimal solution, we can identify three types of cost that are higher (positive difference). First, the highest positive difference is in the transportation cost from the plant to DCs. With eight DCs to replenish, the products have to travel more miles before reaching their destination DCs. Besides, More DCs entails to desegregate the volume in smaller quantities, and this makes transportation rates more expensive. Additionally, some DCs are in locations where transportation modes have lower capacity. When rail capacity is limited, products have to be dispatched by truck, a mode much more expensive to the Company. The second relevant difference is in the holding cost. This is not only explained by the quantity of DCs, but also for the volume to have in stock. DCs far away from the plant require more stock to cover the replenishment lead time.

Regarding the type of cost with negative difference, we can see in Table 30 that there are two types of cost that are lower in the “status quo” than in the optimal solution. The transportation cost from DCs to customer is 4.88% lower due to the advantage of having eight DCs. With eight DC, the products are closer to most of the customers, and this leads to shorter distances and lower transportation cost. Finally, the next cost with negative difference is the tank preparation cost. This cost is zero because there is no need to prepare new tanks if the distribution network remains in the current state. Despite these two costs, which are 8.01% lower than in the optimal solution, the total cost of the “status quo” is 14.84% higher than the optimal supply chain yielded by the model.

Regarding the packaged products, potential savings of implementing the optimal supply chain are similar to those of bulk products. However, as we evaluated different scenarios, we opted for a different approach to estimate the savings. In Figure 21, we presented the resulting total cost of every scenario. As all scenarios were an option to the Company, we consider that the benefit of this study is to allow identifying the cost difference between each scenario. Therefore, the total cost difference can be seen as the potential savings. For example, the difference between scenario 2, which is the optimal one, and scenario 4, is 12.6%. As the Company could have opted for scenario 4, thus the benefit of implementing scenario 2 would be 12.6%. By following this approach, from Figure 21 we can conclude that the minimum potential saving is 7.5%, while the maximum is 18.5%.

The solution of this work was given to and validated by the Company, which used the information to provide insights for managerial decision.

6.6. Limitations

After having executed many scenarios in the model, and performed many analyses with its results, we now discuss some of the limitations of the model, and of our solution approach.

Firstly, regarding the model, we found a weakness to deal with the impact of stochastic demand on the inventory level. Even if our case considered a deterministic demand, a different problem could need some parameters to calculate a safety stock for example. Many demand scenarios can be entered into the model, as well as the turnover ratio needed in each scenario, but the stock-out probability cannot be associated to each turnover ratio. Overall, the turnover ratio is the only parameter defining the inventory level to have at the end of each period, and this could not be enough in some situations. Another aspect related to safety stock is the lead time, a parameter that is not available in the model. Perhaps, an upper bound to set a maximum time or distance traveled for every transportation mode in one trip, could contribute to guarantee a minimum service level in the solution. A final comment on the formulation that could be easily solved is about the single sourcing constraint. The model considers this constraint per product and per customer, but a customer cannot be limited to be served from only one warehouse. This is a common requirement for many cases, where the customer desires to be served by only one facility.

Secondly, regarding the solution approach, there is an issue to take into account about the long-term horizon. Although the model does not assume any length in the duration of one period, we considered every period as a year; however, this has some implications. When a year is the minimal unit of time, the fluctuation within the year cannot be analyzed. In other words, seasonality disappears, and the flow of products as well as the inventory level is assumed to be constant during the year, which is not necessarily true for all the products. This issue could be solved by using a month as the equivalent to one period; however, the forecasted demand should be expressed in the same unit of time for the ten-year horizon. Another limitation of our analysis is the assumption of full truck load transportation when moving the products. Although it is a reasonable approach, in daily operations it might not always be possible, and less-than-truck load (LTL) rates are usually higher. This could increase the cost of the operation, and perhaps the general network configuration results. Finally, ignoring the prices of final products can be considered as a weakness. A different configuration of the network could result when maximizing profits instead of minimizing cost. With this latter approach, we ignore if some products, customers, or locations are not profitable for the Company. With a profit maximizing approach, some customers could be left out of the supply chain, while increasing the benefits to the Company. This approach could be easily done in the short term, since the model considers this possibility.

7. Conclusions and Future Research

The purpose of the current work was to identify the optimal location and size of facilities for a real case in the lubricants industry. The case required to represent a supply chain to distribute bulk and packaged products throughout the United States. The optimization model should lead to the configuration of a supply chain providing the minimal cost for the Company over a ten-year horizon. The total cost of the resulting supply chain is approximately 15% lower than the current distribution network for bulk products. Benefits for packaged product are estimated to be at least 7.5%.

One of the main points that arose from the case with its solution is the relevance of obtaining an integrated solution. This means solving simultaneously a problem with different decision variables and many periods linked among themselves. The level of inventory required in every DC is closely related to the assignment of products to the DC, which in turn, is related to the assignment of customers to the DCs. Sometimes inventory issues are solved independently for a DC, but here we have shown how the inventory of one product in a single DC may affect the result for the other DCs. Besides, considering many periods at the same time allows obtaining favorable results for the entire planning horizon. Moreover, the impact of the fixed cost can be better appreciated in a long-term horizon. Initial investments in period one can affect the configuration of the network at period ten. Sensitivity analyses on fixed cost showed that the assignment of products to a DC changes with a slight change of the fixed costs. In general, the interaction among all the aspects of the supply chain creates a sensitive relationship between the various parameters that has an impact on the results.

Regarding the two types of products in the case, bulk and packaged products, we can conclude that problems can be solved independently only when there is no interaction between their components. Moreover, particularities of products must be considered not only when entering the data, but also when analyzing the results. Despite the optimality in the solution for packaged products for example, rounding was required to ensure consistency in the results. Besides, when thinking about the characteristics of bulk and packaged products, it is simple to understand that the case of lubricants can be extended to many other industries.

Another point to stand out is the significant influence of the collaboration required to represent adequately a supply chain with multiple tiers. Here, the level of detail in the data was possible due to the control of the Company over its supply chain. An incredible amount of data was needed to put together all the parameters, and consistency was a must to obtain a feasible solution for implementation purposes. Collecting the data from different sources, manipulating different criteria, or simply missing information may lead to a wrong analysis.

Future research pursuing new optimization models should consider this work as an alternative to efficiently solve supply chain problems. Using a flexible model allowed to represent the supply chain by simply varying the data in the input file. Moreover, for research leading to new proposals with regards to facility location and sizing in network design, this alternative should be considered as an example of integrality in the solution. The more tiers are present in the model, the better the opportunities to generate savings.

Finally, an interesting effort to go further with this work in the short term would be to consider the environmental parameters available as an extension of the model. Evaluating the emission level of the resulting supply chain, and optimizing it with limits on CO_2 emissions, would be a great step to continue this journey.

Bibliography

Balinsky, M. (1965). *Integer Programming: Methods, Uses, Computation.* Management Sciences, 12(3), 253-313.

Bell, D., De Cesare, S., Lycett, M., Mustafee, N., & Taylor, S. J. (2007). *Semantic Web Service Architecture for Simulation Model Reuse.* 11th IEEE Symposium on Distributed Simulation and Real-Time Applications, 129-136.

Bowersox, D., & Closs, D. (1996). *Logistical Management: The Integrated Supply Chain Process.* McGraw-Hill Companies.

Cohen, M., & Lee, H. (1988). *Strategic Analysis of Integrated Production-Distribution Systems: Models and Methods.* Operations Research, 36(2), 216-228.

Cordeau, J.-F. (2014). *A General Logistics Network Design Model.* Unpublished document.

Cordeau, J.-F., Pasin, F., & Solomon, M. M. (2006). *An Integrated Model for Logistics Network Design.* Annals of Operations Research, 144, 59–82.

Cox, A. (2004). *The Art of the Possible: Relationship Management in Power Regimes and Supply Chains.* Supply Chain Management: An International Journal, 9(5), 346-356.

Daskin, M. S. (2008). *What You Should Know About Location Modeling.* Naval Research Logistics, 55(4), 283-294.

Eskigun, E., Uzsoy, R., Preckel, P., Beaujon, G., Krishnan, S., & Tew, J. (2005). *Outbound Supply Chain Network Design with Mode Selection, Lead Times and Capacitated Vehicle Distribution Centers.* European Journal of Operational Research, 165(1), 182-206.

Fanchao, Z., Turner, S. J., & Aydt, H. (2009). *Symbiotic Simulation Control in Supply Chain of Lubricant Additive Industry.* Proceedings of the 2009 13th IEEE/ACM International Symposium on Distributed Simulation and Real Time Applications, 165-172.

Forrester, J. (1962). *Industrial Dynamics.* American Association for the Advancement of Science, 135, 426-427.

Geoffrion, A., & Graves, G. (1974). *Multicommodity Distribution System Design by Benders Decomposition.* Management Sciences, 20(5), 822-844.

Groznik, P. T. (2006). *Measurement of Supply Chain Integration Benefits.* Interdisciplinary Journal of Information, Knowledge, and Management, 1(38), 562.

Hakimi, S. (1964). *Optimun Locations of Switching Centers and the Absolute Centers and Medians of a Graph.* Operations Research, 12(3), 450-459.

Hokey, M., & Gengui, Z. (2002). *Supply Chain Modeling: Past, Present and Future.* Computers & Industrial Engineering,43(1-2), 231-249.

Holmstrom, J., Hoover Jr., W., Eloranta, E., & Vasara, A. (1999). *Using Value Engineering to Implement Breakthrough Solutions for Customers.* The Internation Journal of Logistics Management, 10(2), 1-12.

Huang, Y., Chen, C.-W., & Fan, Y. (2010). *Multistage Optimization of the Supply Chains of Biofuels.* Transportation Research, 46(6), 820-830.

Hugo, A., Rutter, P., Pistikopoulos, S., Amorelli, A., & Zoia, G. (2005). *Hydrogen Infrastructure Strategic Planning Using.* International Journal of Hydrogen Energy, 30(15), 1523–1534.

Jang, Y.-J., Jang, S.-Y., Chang, B.-M., & Park, J. (2002). *A Combined Model of Networking Design and Production/Distribution Planning for a Supply Network.* Computers & Industrial Engineering, 43(1-2), 263-281.

Johnson, G., & Malucci, L. (1999). *Shift to Supply Chain Reflects More Strategic Approach.* APICS-The performance advantage, 28-31.

Lambert, D., Cooper, M., & Pagh, J. (1998). *Supply Chain Management: Implementation Issues and Research Opportunities.* The international Journal of Logistics Management, 9(2), 1-19.

Lasschuit, W., & Thijssen, N. (2004). *Supporting Supply Chain Planning and Scheduling Decisions in the Oil and Chemical Industry.* Computers and Chemical Engineering, 28(6-7), 863-870.

Lin, J.-R., Nozick, L., & Turnquist, M. (2006). *Strategic Design of Distribution Systems with Economies of Scale in Transportation.* Annals of Operations Research, 144(1), 161-180.

LNG Publishing Company, I. (2014). *Lubricants Industry Factbook.* A Supplement to Lubes'n'Greases Magazine, 20(8), 7-26.

Melo, M., Nickel, S., & Saldanha-da-Gama, F. (2006). *Dynamic Multi-commodity Capacitated Facility Location: A Mathematical Modeling Framework for Strategic Supply Chain Planning.* Computers & Operations Research, 33(1), 181-208.

Melo, M., Nickel, S., & Saldanha-da-Gama, F. (2009). *Facility Location and Supply Chain Management - A Review.* European Journal of Operational Research, 196(2), 401-412.

Min, H., & Melachrinoudis, E. (1999). *The Relocation of a Hybrid Manufacturing/Distribution Facility from Supply Chain Perspectives: A Case Study.* Omega, 27(1), 75-85.

Pirkul, H., & Jayaraman, V. (1996). *Production, Transportation, and Distribution Planning in a Multi-commodity Tri-echelon System.* Transportation Science, 30(4), 291-302.

Sha, D., & Che, Z. (2006). *Supply Chain Network Design: partner selection and production/distribution planning using a systematic model.* Journal of the Operational Research Society (2006) 57, 52-62.

Syam, S. (2002). *A Model and Methodologies for the Location Problem with Logistical Components.* Computers & Operations Research, 29(9), 1173-1193.

U.S. Energy Information Administration. (2014). *Independent Statistics & Analysis*. Retrieved June 30, 2015, from www.eia.gov.

Weber, A. (1929). *Uber den standort der industrien, túbingen, English Translation, by C.J. Friedrich (1929), Theory of the Location of Industries.* University of ChicagoPress.

Appendix

Appendix 1: Structure Required by the Input Data File

As mentioned in Section 5.1 the input data file is composed of eight sections: scenarios, periods, items, suppliers, plants, warehouses, customers and modes of transportation linking the nodes. The beginning of each section is denoted with the opening brace " { ", while the end of the section is indicated with the closing brace and a semi-colon " }; ". To include any comment and avoid to be read as a character, the line needs to start with two diagonal lines " // ". Regarding the format of each section, the notation $[x]^*$ means that the expression inside the braces [ ] has to be repeated for every new parameter value that needs to be added to x. For example, to define the bill of material of commodity k, the expression is denoted as $k\ \ [l\ b^{kl}]^*$, meaning that the expression has to be repeated as many times as there are components l needed to produce one unit of item k. The parameters have the same notation as the one used in the mathematical model, and summarized in Table 4 of Chapter 3. Here, we present the general structure and format of each section of the input file

The format required for each section in the text file is as follows:

Scenarios

$\text{Scenarios} = \{$
$h\ p_h\ ;$
$\};$

The name of the scenario is represented by h and the probability assigned to each scenario is denoted by p_h. Each scenario is entered in a line, where the sum of the probabilities of the scenarios has to equal 1. In our case we coped with one single scenario that was entered as SCE1.

Periods

$$\text{Periods} = \{$$
$$t \;\; E_t \;;$$
$$\};$$

The name of the period is represented by t and each period has to be entered in a line. The parameter E_t represents the upper bound of emissions level for the period. This parameter was not presented in the thesis, as it belongs to a set of extensions that are not taken into consideration in our case. Although dismissed extensions do not require attention in this work, an expression has to be associated in the input file to be read by CPLEX. The expression NAV (Not A Value) is the one created to this purpose. In this case, parameter E_t received the input NAV.

Items

$$\text{Items} = \{$$
$$k \;\; [l \; b^{kl}]^* \;;$$
$$\};$$

The name of the item is represented by k and the expression in the braces is its bill of materials, where l denotes the component required, and b^{kl} the quantity of that component needed in item k. Raw materials need only parameter k.

Suppliers

$$\text{Suppliers} = \{$$
$$s \;\; c_s \;\; (l_s)_{\lceil T\rceil\times|H|} \;\; (q_s)_{\lceil T\rceil\times|H|} \; [\;\; r \;\; c_s^r \;(g_s^r)_{\lceil T\rceil\times|H|} \;\; (e_s^r)_{\lceil T\rceil\times|H|} \;\; (u_s^r)_{\lceil T\rceil\times|H|}$$
$$(l_s^r)_{\lceil T\rceil\times|H|} \;\; (q_s^r)_{\lceil T\rceil\times|H|} \;\;]* \; ;$$
$$\};$$

The name of the supplier is represented by s and every line represents a supplier. As we are not including suppliers in our problem we do not stop here to review the detail of this structure. We simply used a dummy supplier called Sup1 and assigned either a zero or NAV to parameters to avoid any impact on the result.

Plants

$$\text{Plants} = \{$$
$$p \;\; c_p \;\; (l_p)_{\lceil T\rceil\times|H|} \;\; (q_p)_{\lceil T\rceil\times|H|} \; [\;\; k \;\; c_p^k \;(g_p^k)_{\lceil T\rceil\times|H|} \;\; (e_p^k)_{\lceil T\rceil\times|H|} \;\; (u_p^k)_{\lceil T\rceil\times|H|}$$
$$(l_p^k)_{\lceil T\rceil\times|H|} \;\; (q_p^k)_{\lceil T\rceil\times|H|} \;\;]* \; ;$$
$$\};$$

The name of the plant is represented by p, the fixed cost of the plant is denoted by c_p, and lower and upper bound are indicated with l_p and q_p respectively. The notation $\lceil T\rceil \times |H|$ means that the parameter has to be entered for every period t and every scenario h. Then, the parameters in the braces [ ] will be repeated for every item k that is assigned to the plant. For each item a fixed cost c_p^k may be associated. Additionally, g_p^k denotes the production unit cost, e_p^k the unit emission rate, u_p^k the capacity usage needed from the plant, l_p^k the lower bound to be produced, and finally q_p^k the upper bound of item k to be produced at plant p.

Warehouses

$$\text{Warehouses} = \{$$

$$w\ c_w\ (l_w)_{\lceil T\rceil\times|H|}\ (q_w)_{\lceil T\rceil\times|H|}\ (\hat{I}_w)_{\lceil T\rceil\times|H|}\ (\hat{q}_w)_{\lceil T\rceil\times|H|}\ [\ k\ c_w^k\ (g_w^k)_{\lceil T\rceil\times|H|}\ (\hat{g}_w^k)_{\lceil T\rceil\times|H|} (e_w^k)_{\lceil T\rceil\times|H|}\ (\hat{e}_w^k)_{\lceil T\rceil\times|H|}\ (u_w^k)_{\lceil T\rceil\times|H|}\ (\hat{u}_w^k)_{\lceil T\rceil\times|H|}\ (l_w^k)_{\lceil T\rceil\times|H|}\ (q_w^k)_{\lceil T\rceil\times|H|}\ (\hat{I}_w^k)_{\lceil T\rceil\times|H|}\ (\hat{q}_w^k)_{\lceil T\rceil\times|H|}\ (\alpha_w^k)_{\lceil T\rceil\times|H|}\ (\beta_w^k)_{\lceil T\rceil\times|H|}\ ]*\ ;$$

$$\};$$

The name of the warehouse is represented by w and most of the other parameters have the same structure as the plant. However, there is a new notation defined with a "hat" (⌒) over the parameter. This symbol indicates that the parameter represents handling activities, while the other represent holding activities. Additionally, the parameter α_w^k represents the desired turnover ratio of the item, and β_w^k is used to convert units flowing through the warehouse into units of inventory.

Customers

$$\text{Customers} = \{$$
$$c\ \left[k\ (a_c^k)_{\lceil T\rceil\times|H|}\ (d_c^k)_{\lceil T\rceil\times|H|}\ (g_c^k)_{\lceil T\rceil\times|H|}\ y_c^k\right]*\ ;$$
$$\};$$

The name of the customer is represented by c and a line is required for every customer. The item demanded by the customer is denoted by k. Then, parameter a_c^k denotes the minimal quantity demanded, while d_c^k the maximal quantity. These parameters are also part of one of the extensions of the model. For our case, we entered the same value for both parameters. Similarly g_c^k is the revenue of selling one unit, and finally y_c^k has to be set to 1 to impose the single sourcing constraint or zero otherwise.

Modes of Transportation Linking the Nodes

$$
\begin{aligned}
&\text{Modes} = \{ \\
&m\ \ o\ \ d\ \ c_{od}^{m}\ (l_{od}^{m})\ \ _{\lceil T \rceil \times |H|}\ (q_{od}^{m})\ \ _{\lceil T \rceil \times |H|}\ \left[k\ \ (g^{km})\ \ _{\lceil T \rceil \times |H|}\ \left(e^{km}\right)\ \ _{\lceil T \rceil \times |H|} \right. \\
&\left. \left(u^{km}\right)\ \ _{\lceil T \rceil \times |H|}\right] *\ ; \\
&\};
\end{aligned}
$$

The name of the mode of transportation is represented by m and the two nodes linked by the mode of transportation are denoted by o and d. However, a line is required to represent every origin-destination link than can be done by the same mode of transportation. In other words, this section does not contain only the modes of transportation but also the edges between the nodes. Therefore, the same mode of transportation can be repeated as many times as paths are covered by the mode. The items transported by the mode are denoted by k, and the other parameters contain the same structure as the previous sections.