B2B Omnichannel Network Design and Inventory Positioning

by

Geoffrey Allen
Bachelor of Finance / Bachelor of Economics, James Madison University, 2004
Master of Business Administration, University of Virginia, 2012

and

Shoichi Ishida
Bachelor of Industrial and Management Systems Engineering, Waseda University, 2019
Postgraduate Certificate of Supply Chain Analytics, University of London, 2021

SUBMITTED TO THE PROGRAM IN SUPPLY CHAIN MANAGEMENT
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE IN SUPPLY CHAIN MANAGEMENT
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2023

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

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

Certified by: ______________________________
Dr. Eva Ponce
Executive Director, MITx MicroMasters Program in SCM
Capstone Advisor

Accepted by:

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

B2B Omnichannel Network Design and Inventory Positioning

by

Geoffrey Allen

and

Shoichi Ishida

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

ABSTRACT

The foodservice distribution industry in the United States is projected to experience a 40% growth from 2022 to 2025. As the industry expands, an increasing number of companies are adopting an omnichannel supply chain strategy. In this context, our sponsor company, a large U.S. foodservice distribution company, aims to optimize the end-to-end supply chain network, including omnichannel options, in order to minimize the overall supply chain cost. This objective is typically achieved through a technique known as supply chain network design, which has been extensively researched. However, while there is abundant literature discussing omnichannel supply chain strategies and network design, only 4% of the reviewed supply chain research specifically addresses B2B applications. Given the substantial differences between B2C and B2B businesses (e.g., lead-time requirements, network and cost structure, number of facilities and customers), a new customized model is required. Additionally, many previous studies have primarily focused on cost factors directly proportional to the volume of flows within the network, such as product cost, transportation cost, and warehouse handling cost. However, to achieve the most optimal solution, we must consider the impact of inventory positioning. Changes in inventory positioning can significantly influence the ability of the model to leverage inventory pooling, ultimately affecting safety stock costs. Since this pooling effect generally exhibits nonlinear behavior, many previous studies have either overlooked it or proposed separate models for network optimization and inventory positioning. To address these issues, we propose a mixed-integer non-linear programming (MINLP) model that simultaneously optimizes supply chain network flows and inventory positioning. We have solved this model by reformulating the square rooted term representing safety stock cost as a quadratic constraint, which can be solved by commercial solvers. Additionally, we have developed a tailored algorithm using outer approximation (OA) to expedite the solving process for our inherently complex model, known as NP-Hard. Depending on the specific products, our results demonstrate potential cost reductions of 3-9% in transportation, 2-8% in warehouse handling, and up to 50% in inventory costs. Furthermore, we have achieved more than nine times greater efficiency with our tailored algorithm compared to Gurobi's default solve. Lastly, our model exhibits the possibility of future expansion into a multi-item model, which would have a significantly greater impact on the company.

Capstone Advisor: Dr. Eva Ponce
Title: Executive Director, MITx MicroMasters Program in SCM

ACKNOWLEDGMENTS

Joint Acknowledgments:

We would like to thank our advisor, Dr. Eva Ponce, for continually inspiring us to dig deeper, to raise our standards to the highest possible level, to think cross-functionally, and to build upon the best foundations wherever we can. We greatly appreciate your thought partnership and mentorship over the past 6 months.

We would also like to thank our writing coach Toby Gooley for your dedication, your guidance, and countless thoughtful suggestions. This paper would have been an incoherent mess without you.

We thank Dr. Elenna Dugundji and Dr. Thomas Koch for encouraging us 9 months ago to build our own model for the learning opportunity, and then in supporting us with initial setup. You put us on a great path, and we have enjoyed the journey.

We also thank Dr. Sean Willems for his contagious passion for inventory management, for sharing his practical experience implementing new inventory strategies, and for his great advice to "simplify and accelerate."

Also, without the guidance and wisdom of Prof. Paul I. Barton at MIT Chemical Engineering and Dr. Byung-Gak Son at Bayes Business School, City, University of London, this paper would not have been possible. Thank you for your dedication to your students' success and willingness to take the time to discuss and debate ideas with us, even when it required going beyond the scope of the curriculum.

Lastly, we save a special thanks to our sponsor company who made all of this possible. For confidentiality reasons we will not name you, but our three executive sponsors were invaluable over the past 6 months by helping us remove barriers, challenging our thinking, and pushing us to be better. Thank you for the time and support.

Geoffrey Allen Acknowledgments:

I would like to thank my family for supporting my winding path through educational endeavors over the years and for always encouraging curiosity, practical or otherwise.

To my capstone partner, Shoichi Ishida, it truly has been a blessing and a pleasure working with you. I couldn't imagine going through this with anyone else.

Finally, I save a thanks for my early mentors, Chris MacCracken, John Blaney, Steve Fine, and Judah Rose who first taught me all I know about optimization and sparked my passion in this topic.

Shoichi Ishida Acknowledgements:

My foremost gratitude should go to my wife, Paulina Ishida, for her unwavering support, encouragement, and patience throughout highs and lows of this thesis and my entire life. Your love and understanding have been the anchor that kept me grounded and steady even through life's tumultuous waters. I also would like to express my heartfelt appreciation to my mother-in-law, Anna Antsuta, and my brother-in-law, Gorbunov Maksim, who have enabled my study journey at MIT and always been inspiration and wisdom. Their guidance and support have been invaluable.

To my capstone project partner, Geoffrey Allen, thank you for your dedication, hard work, and collaboration throughout this project. Your commitment to excellence has inspired me to strive for nothing less than the best.

Finally, I also wish to acknowledge the memory of my deceased professor at Waseda University, Prof. Takayuki Shiina, whose teachings on operations research continue to inspire and guide me. He will always be remembered for his contribution to my education and career.

TABLE OF CONTENTS

INTRODUCTION 7
1.1 Problem Statement and Research Questions 8
1.2 Scope: Project Goals and Expected Outcomes 8
2 STATE OF THE ART 10
2.1 Foodservice Supply chain and Omnichannel Strategy 11
2.2 Supply Chain Segmentation Techniques 13
2.3 Inventory Policy 15
2.4 Network Design with Integrated Inventory Policy 16
2.5 Omnichannel Distribution Models and Lateral Transshipment 18
2.6 Efficient Modelling 21
2.6.1 Computational Time for Supply Chain Problems 21
2.6.2 Efficient Methods for Large-scale NP-Hard Problems 22
2.7 State of the Art Summary and Conclusions 22
3 DATA AND METHODOLOGY 24
3.1 Scope and Process 24
3.2 Formulation Overview 25
3.3 Key Assumptions and Data Pre-processing 26
3.3.1 Approach to Data Collection 26
3.3.2 Item Selection 27
3.3.3 Product Cost, Distribution Site Location, and Warehouse Handling Cost 28
3.3.4 Product Demand 29
3.3.5 Inbound/Outbound Transportation Cost and Lane Availability 29
3.3.6 Inventory Cost Parameters 32
3.3.7 Direct Ship Cost 32
3.3.8 Cross-docking / Peer Replenishment Cost 33
3.3.9 Central Warehouse Cost 34
3.3.10 Summary of Cost Elements for Each Item 34
3.4 Formulation 36
3.4.1 Transportation Cost 37
3.4.2 Warehouse Handling / Inventory Cost 38
3.4.3 Product Cost 40
3.4.4 Formulation Overview 41

3.5 Tools Utilized 42
3.6 Omnichannel Fulfillment Options 42
3.6.1 Direct Ship 43
3.6.2 Cross-docking and Peer Replenishment 44
3.6.3 Central Warehouses 46
3.7 Algorithm 47
3.7.1 Gurobi Optimizer and Reformulation to Quadratic Programming 48
3.7.2 Outer Approximation for Inventory-integrated Network Optimization 49
4 RESULTS AND DISCUSSION 51
4.1 SKU1 Model Results 51
4.2 SKU2 Model Results 56
4.3 SKU3 Model Results 60
4.4 Omnichannel Fulfillment Options Insights 63
4.4.1 Direct Ship 63
4.4.2 Cross-docking and Peer Replenishment 64
4.4.3 Central Warehouses 65
4.5 Model Solve Time Insights 65
4.6 Expandability and Limitations 67
4.6.1 Expandability 68
4.6.2 Limitations 69
5 CONCLUSION 70

LIST OF FIGURES

Figure 1 General Supply Chain Network and Cost Factors for the Food Distribution Industry 12
Figure 2 Common Attributes for Segmentation 14
Figure 3 Research Process 25
Figure 4 ABC-XYZ Segmentation 27
Figure 5 Regression for Inbound Freight Rates 30
Figure 6 Approximation for Outbound Transportation Cost 31
Figure 7 Difference in Cost Factors between Normal Shipment and Direct Ship 43
Figure 8 Difference in Cost Factors between Normal Shipment and Cross-docking 46
Figure 9 Difference in Cost Factors between Normal Shipment and Central Warehouses 47
Figure 10 Network Flow Comparison for SKU1 53
Figure 11 Inventory Comparison for SKU1 54
Figure 12 Inventory Comparison by Site for SKU1 55
Figure 13 Network Flow Comparison for SKU2 57
Figure 14 Inventory Comparison for SKU2 58
Figure 15 Inventory Comparison by Site for SKU2 59
Figure 16 Network Flow Comparison for SKU3 61
Figure 17 Inventory Comparison for SKU3 62
Figure 18 Inventory Comparison by Site for SKU3 63

LIST OF TABLES

Table 1 SKU Selection 28
Table 2 Cost Elements for Each Item 35
Table 3 Cost Result Summary for SKU1 51
Table 4 Cost Result Summary for SKU2 56
Table 5 Cost Result Summary for SKU3 60
Table 6 Execution Time for Each Method 66
Table 7 Solution (Total Supply Chain Cost) Difference between Methods 66

INTRODUCTION

The trend towards omnichannel distribution, with multiple overlapping order methods and multiple fulfillment methods, have recently accelerated due to COVID-19. This trend has been present for B2C sales for over a decade, but recent research shows that B2B sales are now also permanently on a new trajectory towards omnichannel as well (McKinsey, 2021). B2B customers increasingly prefer flexibility in both how they buy and in the fulfillment options offered, inspired by what has been experienced in B2C. According to a 2021 report from Salesforce (2022), 72% of B2B buyers now expect "a similar experience on a B2B site as they get on a consumer website."

However, for companies used to a more traditional supply chain network design, implementing an omnichannel distribution network can be complex and potentially costly (Jocevksi, 2019). As Winkenbach et al. (2021) stated, "to address rising customer expectations with respect to delivery lead time, location, and flexibility… the strategic design of the underlying last-mile distribution network must align with the delivery service offerings and the associated omnichannel strategy." This may include opening new distribution methods and, to manage cost, creating a multi-tier network.

The sponsor for this project, a large B2B foodservice distribution company in the United States, is offering a comprehensive range of food-related products to diverse customers such as restaurants and schools through its extensive nationwide supply chain network. It is beginning to offer omnichannel fulfillment to customers via pilots started over the past year. While initial results have been promising, scaling these pilots across the United States will be challenging due to the current network design. The current fulfillment model ships from a vast network of local distribution sites to customers for next day delivery, with significant variation in product offerings, inventory policy, and fulfilment methods. However, customers do not actually expect all items to be delivered with the same lead time. This fragmentation, and the new opportunity of omnichannel fulfillment, has catalyzed our sponsor company to explore new network designs.

1.1 Problem Statement and Research Questions

Inventory positioning and fulfillment strategy at our sponsor company has been managed primarily in a decentralized approach, with the local distribution sites being the main determinant of which items to stock and which customers to serve from the different sites. However, this approach does not produce wholly optimal solutions. There is an opportunity to reduce total cost of fulfillment as well as reduce complexity of the network by considering a more holistic solution. In addition, new omnichannel fulfillment options are now demanded by customers, adding further complexity. Bell et al. (2014) stated that to win in an omnichannel world, companies must put the customer at the center of the omnichannel experience as well as foster cross-functional collaboration to succeed with omnichannel strategies. Taking this into account, this paper addresses the following questions:

1. What is the current state of the supply chain network in terms of suppliers, item flows, distribution centers, customer demand, and costs? What is the current network design strategy? What additional fulfillment options would be possible given existing facilities?
2. How do customer service and lead time requirements influence the network design?
3. Given the input above, what is the least-cost solution to satisfy customer demand? What are the most significant differences between the current state and the new solution?
4. What would changing to the new optimized design with new omnichannel fulfillment options be worth? Which design elements have the greatest influence on reducing cost?

1.2 Scope: Project Goals and Expected Outcomes

The overall goal of this project is to develop a proof-of-concept strategic model that solves for a national-level stocking strategy under different omni-channel fulfillment options. This resulted in a tool that could be used across functions at our sponsor company, including: Supply Chain, Merchandising, Finance, and Operations. We then demonstrate features of the tool with several example items. The model includes end-to-end supply chain costs such as product cost, inbound cost, warehouse handling cost, inventory cost, and outbound cost.

To address the questions raised by the project, this model considers different scenarios regarding customer fulfillment expectations (for example, different lead times) via different items. We reviewed existing research and literature to grasp the best way to consider omni-channel fulfillment and how to optimize service level and other trade-offs at the same time. We further discussed with our sponsor company the preferred approach to consider existing models, expected scenarios, and data usage (e.g., what inputs are needed, how to acquire them, and how to process them).

We hypothesized that to allow the model to be feasible at a national level, we would need to aggregate some details about the network. For example, we needed to cluster suppliers and/or customers into larger regions and make other approximations as appropriate. We have worked with the sponsor company to align on what level of aggregation strikes the right balance between detail and feasibility.

We further hypothesized that the new model will decrease the total network cost by optimizing the whole chain, which leads to improving P/L, and furthermore, may reveal opportunities to improve the balance sheet by identifying unnecessary assets or finding more efficient use of assets.

The deliverables of this capstone project include:

1. New strategic-level optimization model to explore end-to-end flows and inventory positioning strategy of the sponsor company supply chain network.
2. Compared to a historical baseline, recommendations of major changes to network design.
3. An estimate of financial implications of executing the optimal design.
4. Scenarios that illustrate different designs based on modifying key variables or constraints specifically for three items.

2 STATE OF THE ART

When designing supply chain distribution networks, there are several branches of theory that can be applied. Classic supply chain network design analysis optimizes flows between nodes in the network, typically to minimize transportation cost; inventory management analysis aims to minimize total inventory cost while meeting service level obligations; and newer omnichannel analytical techniques layer in multiple fulfillment modes as well as inventory pooling techniques. In reality, these areas of supply chain are highly interrelated and should be jointly considered while designing the overall supply chain strategy. A final complicating factor is the size of our sponsor company's distribution network: over 160,000 unique items distributed through a large national multi-echelon supply chain to over half a million customers. To help simplify this network we also examine common approaches using segmentation that will assist all areas of analysis mentioned above.

This chapter reviews the literature on these topics. First, we examine the context of the foodservice distribution industry and why supply chain strategy is important to the overall success of the companies involved (Section 2.1). Second, we look at common supply chain segmentation methods to help us simplify our sponsor company's data and focus on the most relevant aspects (Section 2.2). Third, we look at relevant inventory control policies (Section 2.3). Fourth, we review integrated inventory-transportation and inventory-location models that jointly optimize transportation and inventory costs in a way relevant to our sponsor company (Section 2.4). Fifth, we look at the latest research into omnichannel supply chains and how new models modify and/or build upon the integrated inventory-transportation models that existed before (Section 2.5). Sixth, we examine ways to reduce model complexity and decrease solve time (Section 2.6). Finally, we summarize the key lessons as they apply to this study (Section 2.7)

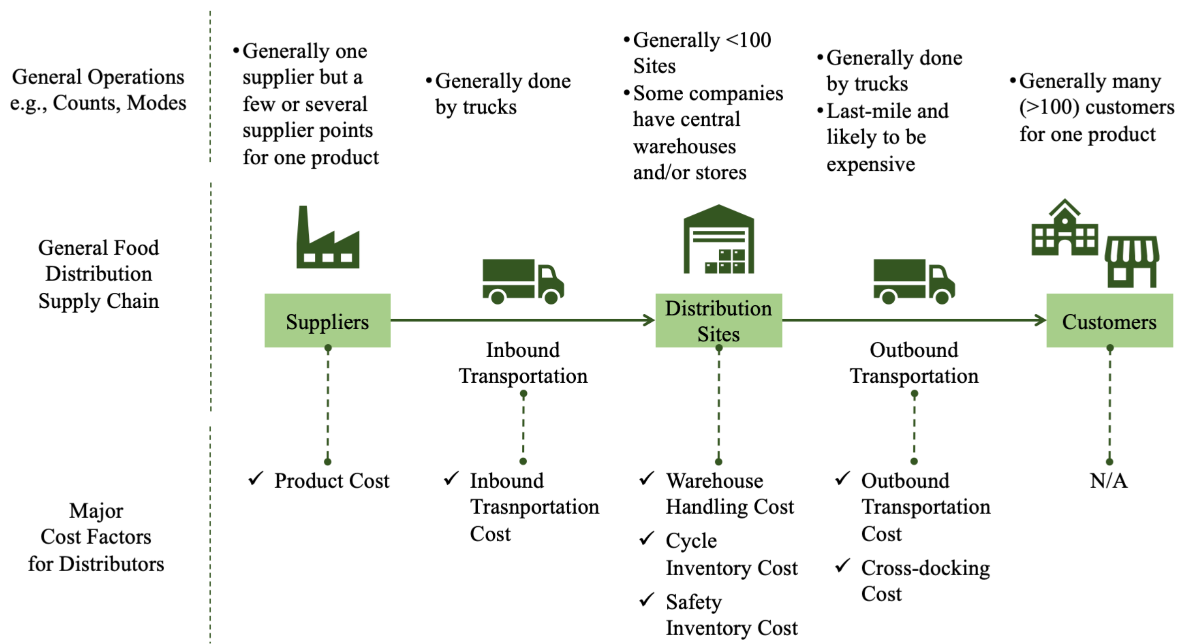
2.1 Foodservice Supply chain and Omnichannel Strategy

Foodservice distribution industry sales in the United States were approximately $300 billion in 2022 and are expected to grow to $422 billion by 2025 (International Foodservice Distributors Association, 2022). This industry is composed of large national public distributors, such as Sysco, US Foods, and Performance Food Group (PFG); one large national private distributor, Gordon Food Service (GFS); and a number of smaller regional distribution companies. The industry is not particularly concentrated, with Sysco the largest company at 17% market share (Sysco, 2022), US Foods at 11.4% (US Foods, 2022) and PFG at 8.9% (PFG, 2022). Though private, Gordon Food Service is estimated by Forbes (2022) to have revenue of approximately $19 billion, equivalent to 6.3% of market share.

Figure 1 illustrates the general supply chain for food distribution. In this process, each product is sourced from external suppliers, with the exception of a few unique cases such as own-brand products. Typically, a single supplier serves a specific product, although it is common for suppliers to have multiple supply points. The products then proceed to companies' distribution sites where they are stored and prepared for outbound delivery to customers. These warehouses are strategically located to achieve a balance between inventory pooling and lead-time requirements. While certain companies maintain physical stores where customers can collect their items, it is worth noting that this practice differs from B2C food distribution. In the B2B context, such store facilities are not as prevalent, as the majority of B2B distribution sites are typically situated in remote areas, away from urban centers. Ultimately, the products are delivered to a large number of customers, which are typically restaurants, schools, jails, military facilities, and hospitals. In the B2B food distribution industry, the prevailing norm is one-day delivery, a standard that is less common in B2C food distribution. Furthermore, B2B customers typically place larger volume orders compared to their B2C counterparts. It is worth noting that each of these activities incur distinct costs within the supply chain. Also, the product range offered by B2B food distribution companies is highly diverse. These products span from everyday food staples like vegetables, fruits, meat, and dairy products to various food-related supplies such as packaging materials, trays, gloves, and napkins.

Figure 1

General Supply Chain Network and Cost Factors for the Food Distribution Industry



At this point all four of the major companies have established digital ordering platforms as part of their omnichannel strategy. This paper does not focus on that aspect of omnichannel, but instead on their order fulfilment strategies and the proliferation of different modes and channels to serve demand. Each of these companies is pursuing a slightly different omnichannel fulfillment strategy. US Foods and GFS have both invested in foodservice focused grocery stores (specializing in items for restaurants and other food related business, for example) in addition to their distribution networks that deliver direct to customers. For GFS this has long been a core part of their strategy and they now have nearly 200 stores in 13 states (GFS, 2023). US Foods first opened a foodservice store more recently in 2012 and now has 88 locations across the US (US Foods, 2023). The supply chain strategy of these companies is similar to the omnichannel strategy emerging in the B2C space of large retail grocery companies. US Foods (2022) specifically mentions “leveraging omni-channel for additional growth opportunities” in their most recent Q4 FY2022 results, in addition to focusing on supply chain network optimization involving “further optimized inbound logistics” and “reduced routing mileage and increased cases per mile” (2022).

Sysco and PFG have both continued a distribution-only (no stores) strategy. Sysco focuses primarily on foodservice distribution, though also serves customers in Canada, the Caribbean, Central America, and Europe in addition to the core US distribution business. PFG focuses on the US, however, has diversified to distribution for convenience stores and movie theaters in addition to the core foodservice market. Both Sysco and PFG are pursuing omnichannel but primarily through new ordering channels via new digital platforms, fulfillment options, and subscription models. Neither has incorporated stores as a fulfillment method. Sysco recently mentioned developing “an omnichannel inventory system and distributed order management system to broaden access to what we carry” as part of their Recipe for Growth strategy (Sysco, 2022). Meanwhile, PFG has not emphasized omnichannel in its recent public presentations, instead focusing on “multiple delivery options” and “cross selling opportunities” that arise from its three core customer segments (PFG, 2023).

2.2 Supply Chain Segmentation Techniques

As previously mentioned, our sponsor company’s supply chain network is complex and not possible to accurately model with the high number of unique items, customers, and multiple echelons. We need to simplify and rationalize the most important aspects of the network while preserving relevant detail. Supply chain segmentation is a useful and widespread tool to assist with this.

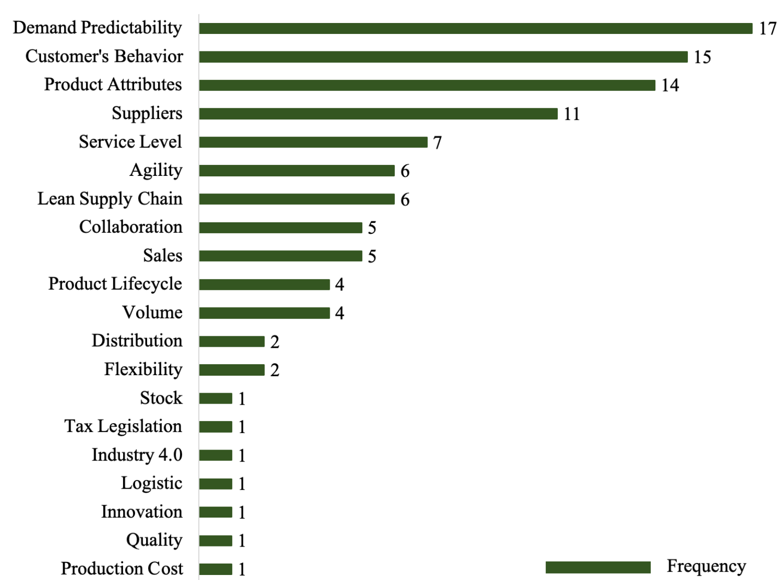
General Electric (GE) was one of the first companies to use the now common ABC segmentation approach (Stojanovic & Regodic, 2017). This approach uses one criterion, typically volume or revenue, to break items in a supply chain into three categories: A for high volume items, B for medium volume items, and C for low volume items. An alternative segmentation approach, XYZ analysis, also uses one criterion but focuses on variability. These two approaches can be combined into an integrated ABC-XYZ approach that categories items into a nine-box matrix of volume and variability. A typical weighting is below:

- Group A consists of the elements encompassing 0-70% of the obtained total value of the elements.
- Group B consists of the elements encompassing 70-90% of the obtained total value of the elements.
- Group C consists of the elements encompassing 90-100% of the obtained total value of the elements.
- Group X consists of the elements whose variation coefficient is less than 0.5.
- Group Y consists of the elements whose variation coefficient is between 0.5 and 1.
- Group Z consists of the elements whose variation coefficient is bigger than 1.

Other approaches also exist. Lee (2002) suggests using a 2x2 matrix of supply and demand variability. Christopher et al. (2006) suggest using demand predictability and replenishment lead time. McKinsey (Alicke & Forsting, 2017) uses a three-step approach that narrows down to 2-3 key performance drivers, analyzes the portfolio using the drivers, and then applies a tailored supply chain strategy to each segment. Given all the different techniques, Pereira et al. (2022) summarize a literature review with the most common attributes used for segmentation, the result of which is shown in Figure 2:

Figure 2

Common Attributes for Segmentation



Taking into account our understanding of the company's supply chain and most critical parameters, we aim to use a manageable but meaningful number of attributes to simplify the supply chain network. We then consistently use this segmentation throughout the network design and inventory policy analysis. Lastly, we use the segmentation framework to help select different items that demonstrate the robustness of our approach.

2.3 Inventory Policy

There are many different approaches to inventory policy. A common approach is to set inventory according to days of supply. Yet another looks at historical demand and sets inventory at the most conservative point where inventory is set at the average needed for order fulfillment plus a buffer of (maximum daily sales * maximum lead time) – (average daily sales * average lead time). Our sponsor company generally sets inventory levels using a periodic review method with a target fill rate. In this approach both a cycle stock level, related to normal order patterns, and a safety stock level, to cover uncertainty, for inventory can be estimated. A complicating factor is that for many products both demand from customers and lead time from suppliers are highly variable. Silver et al. (1998) noted the standard approach in this case with the following equation for safety stock:

$$z\sqrt{LT\sigma_D^2 + D^2\sigma_{LT}^2}$$

where z is the service level desired, LT is the combined average lead time and review period, σ_D is the standard deviation of demand, σ_{LT} is the standard deviation of lead time, and D is the average demand. This equation assumes that demand and lead time are normally distributed. Cycle stock in a periodic review system is then based on order frequency and can be calculated as:

Annual demand * Review period (weeks) / 52 / 2

Adding these two equations together then represents the average level of inventory expected over the period of time examined. While periodic review inventory models can result in higher levels of inventory compared to continuous review models, in which orders are triggered based on order points rather than time, the advantage of periodic review is due to a simpler execution. This is more important in large companies with significant SKU counts, like our sponsor company, as setting standard review cadences is more likely to prevent errors and can also help minimize freight costs if multiple SKUs are ordered at once.

Prak et al. (2017) note that the traditional approach, as discussed above, needs to be modified when looking to the future. Unless demand is relatively stable, most companies rely on forecasts rather than historical demand to plan for future demand. To adjust for forecast inclusion, historical average demand would be replaced by the forecast amount and standard deviation of demand would be replaced by forecast error, measured for example by the root mean square error (RMSE). However, Prak et al. (2017) also note that using common forecasting techniques such as simple moving average or exponential smoothing would result in forecast errors for different periods that are positively correlated. This would then likely result in setting safety stock targets that are too low.

This project relies only on historical demand to demonstrate proof of concept and so avoids this issue. However, any actual implementation should keep the previous warning in mind with adapting this methodology for forecast inclusion.

2.4 Network Design with Integrated Inventory Policy

While linear programming-based network optimization techniques, designed to optimize transportation flows, have existed since the 1940s, research on integrating inventory into these techniques is relatively new. The main stumbling block was that while many transportation-related costs scaled in a linear fashion, inventory costs such as safety stock cost scaled in a nonlinear way. Interest in this topic increased in the 1990s as companies increasingly globalized, extending supply chains and making the tradeoffs between transportation costs and inventory costs more important. At the same time, computing power was increasing, and researchers started applying different mathematical techniques to either approximate inventory costs into network optimization models or reformulate equations so they would be compatible with commercial solvers.

An early example of approximation of inventory costs was Nozick and Turnquist (1998), which integrated safety stock to network design by assuming a linear relationship between the number of DCs and the amount of safety stock. However, this assumption may not hold in cases where the relationship between variables is nonlinear, or when there is a pooling effect in the inventory across different products or locations, rendering the approach unrealistic for companies relying on their past records. Croxton and Zinn (2005) applied the "Square Root Law," which found natural behavior of safety stock nonlinearly increased with the number of DCs (Disney et al., 2006).

$$Safety\ Stock = SS_0\sqrt{\#\ DCs}$$

In this study, the initial value of safety stock SS_0 was determined using the standard safety stock equation, assuming a single distribution center (DC) and all demands were directed to this DC. While this method does incorporate the number of DCs, it does not capture the dynamic nature of safety stock, which is influenced by several other factors such as lead-time, lead-time variability, demand size, and demand variability as indicated by the standard equation. It is worth noting that even when two network designs have the same number of DCs, safety stock may vary considerably, especially when a certain DC has significantly higher or more variable demand compared to the other DCs.

Other early attempts addressed the issue by solving transportation models and inventory models sequentially and/or iteratively. This can be seen in Darmawan et al. (2021), which first solved the network design to fix the parameters needed to solve the safety stock problem and then solved safety stock optimization using these parameters. This can also be seen in some commercial software such as Coupa Supply Chain Design & Planning which offers both modules as separate entities. This methodology solves the larger location and transportation problem first then uses elements of that solution, such as allocation of demand and lead-time from suppliers, to feed a separate inventory module (Coupa, 2021).

However, this two-step approach does not address the interdependence between safety stock and network design, especially if holding costs are high. Changes in the safety stock amount for each DC, often incurred to take advantage of the pooling effect, lead to changes in optimal network design, and vice versa. Ignoring this impact may result in solutions that are far from optimal. Multi-echelon networks further complicate the issue.

Later research focused on transforming the nonlinear equations related to inventory cost to a form compatible with commercial solvers. For instance, a mixed-integer nonlinear programming (MINLP) model was developed by Ağralı et al. (2012) to incorporate safety stock considerations, which was solved using a customized algorithm based on generalized Benders decomposition and outer approximation. You and Grossmann (2008) used a different approach that transformed the objective function by moving the nonlinear parameters into constraints with quadratic equations. The transformed objective function was then linear. Puga et al. (2019) took a similar approach using transformation with conic quadratic equations, stating that "the main advantages of the CQMIP approach is that it is direct, efficient, and flexible, as it can be solved using standard optimization packages without the need for specialized algorithms."

With the release of version 9.0 in Nov 2019, the widely used commercial solver Gurobi gained the ability to solve optimization problems with linear objective functions and quadratic constraints. We take advantage of this new capability, combined with the more recent research approaches mentioned above, to create a model that can solve the integrated transportation-inventory problem in one step.

2.5 Omnichannel Distribution Models and Lateral Transshipment

Much of the literature of omnichannel distribution focuses on B2C and combining elements of store fulfillment vs online fulfillment. Taylor et al. (2019) extensively reviewed numerous past literatures pertaining to omnichannel fulfillment and arrived at the significant finding that "the integration of fulfillment channel inventory and resources is increasingly recognized as a crucial objective in fulfillment management." Furthermore, their research paper presented a summary of statistical data concerning published papers, highlighting the prevalence of analytical papers, which accounted for over 40% of all the papers in 2019. This observation suggests that both companies and the industry exhibit a greater emphasis on analytical research, prioritizing quantitative analyses over qualitative approaches.

The field of B2C omnichannel supply chain has witnessed a significant number of papers. Agatz et al. (2019) conducted research specifically on route optimization within the context of omnichannel strategies. Their study emphasized the effectiveness of considering shared capacity at facilities to accommodate multiple supply channels within the omnichannel network. Chopra (2018) examined how transportation costs, inventory costs, and other supply chain factors vary depending on the types of omnichannel strategies employed, such as showrooms, home delivery, and customer self-pickup. These papers provided valuable insights into B2C omnichannel supply chain dynamics. However, since our research focuses on the B2B domain, which involves different fulfillment methods that do not include stores, it is necessary to explore past literature that specifically addresses B2B supply chain scenarios, which are relatively less common.

Gerea et al. (2021) note in an omnichannel literature review, for example, that only 4% of papers reviewed address B2B specific applications. This section focuses on research specific and/or applicable to B2B and excluding stores. This focus includes opening multiple distribution fulfillment methods such as direct ship from supplier and drop ship. We also focus on review on lateral transshipments and the concept of virtual pooling of inventory.

Guerrero-Lorente et al. (2020) note that there is growing evidence that network designs with flexible fulfillment, defined as multiple possible fulfillment options by different echelons in the network, outperform more traditional network designs. This is due primarily to better facility utilization and the minimization of relevant fulfillment costs. They also note that there is the potential for better inventory management via the risk pooling effect across channels. However, the complexity of these new network designs can make finding optimal solutions more difficult.

Hubner et al. (2016) came to similar conclusions pointing out the inherent tradeoff between inventory pooling and delivery efficiency. They also note that those that adopt omnichannel fulfillment strategies can delay inventory allocation longer, taking advantage of more centralized inventory. They observe that a common strategy emerging is to stock fast moving items at regional DCs or stores close to customers with short lead times while stocking slow moving items at central DCs to take advantage of the pooling effect. They also note that in this strategy the regional DCs can also act as transshipment points (cross-dock) for the central DCs. Finally, they mention direct supplier shipments as a viable but limited option for fulfillment, which can have advantages in inventory and processing costs. However, there are potential downsides including longer lead times, data exchange issues, and fragmentation of orders.

While Winkenbach et al. (2021) mainly focus on omnichannel in a B2C context, there are several conclusions relevant to this project. First, that the network design and fulfillment options should be tailored to customer requirements, lead times, and demand characteristics such as density. This will result in more optimal solutions. Second, that solving the different supply chain design parameters in an integrated methodology, where possible, yields significantly better results than iterative and/or sequential solving methodologies.

As the number of nodes and possible fulfillment combinations increases, omnichannel supply chains begin to take on characteristics relevant from other branches of supply chain theory. There are two that we focus on in this paper: multi-echelon inventory management and lateral transshipments. Multi-echelon systems, as described by Graves and Williams (2000), can strategically locate inventory at different and potentially sequential points in the supply chain depending on net lead time and other network characteristics. The optimal solution in these networks may store inventory in multiple echelons in the supply chain. Lateral transshipment focuses specifically on nodes with duplicative inventory in the same echelon acting as potential sources of supply for each other. As discussed in Meissner and Senicheva (2018), this can be characterized as either reactive, due to stockouts, or proactive, which tend to focus on transshipments for inventory equalization. In general transshipment is worth the cost when the cost of stockout is higher than the cost of transshipment or when holding costs are high, and therefore tends to be reserved for either high margin or highly critical items such as spare parts. However, Paterson et al. (2011) note there is a growing body of research demonstrating the pooling effects from transshipments can lower overall system costs as well as increase service levels. Omnichannel is noted specifically as potentially a new area where lateral transshipment across channels could be used to improve performance.

2.6 Efficient Modelling

After discussing the essential components of omnichannel supply chain design models, another critical aspect to consider is model solve time due to the potential complexity of these models. In this chapter, we briefly explain the importance of model efficiency and discuss ways to improve it.

2.6.1 Computational Time for Supply Chain Problems

In supply chain management, the complexity of business rules and the numerous transactional datapoints necessitate the use of large models. This results in the need for high computational power and extended solving times. The slow-down in model solving can be attributed to various mathematical factors, such as binary variables, flow constraints, and non-zero constraints, according to Coupa (Strickler, 2021). Additionally, some classes of problems in supply chain management are known to be intrinsically difficult to solve. One good example is mixed integer linear programming (MILP), which is categorized as “NP-Hard”, “the complexity class of decision problems that are intrinsically harder than those that can be solved by a nondeterministic Turing machine in polynomial time (Atallah, 1999).” This observation naturally implies that comparatively intricate approaches, such as MINLP, are more challenging, or at least equally demanding. Due to the growing volatility in the supply chain industry, it has become increasingly common to make quick decisions by repeatedly running models with the large amount of data, which creates a conflict with the nature of this type of problem, highlighting the need for improved efficiency.

2.6.2 Efficient Methods for Large-scale NP-Hard Problems

The optimization of NP-Hard large-scale problems has been a focus of researchers who have dedicated considerable effort towards the development of faster algorithms to solve such "difficult" problems. One approach utilized involves breaking down the problem into smaller subproblems, based on binary variable values or other dominant factors of the problem, using a method such as Branch-and-Bound. The efforts to narrow down the scope of the problem lead to the development of more sophisticated methods such as Cutting Plane, Branch-and-Cut, and Column Generation. Dividing the problem based on the block problem structure has also been the focus of researchers who use various decomposition methods, including generalized Benders decomposition and Outer Approximation. The latter, Outer Approximation (OA), proposed by Duran and Grossmann (1986), is an intuitive approach that has shown promise in complex optimization problems such as MINLP. Compared to other decomposition algorithms, OA is relatively easy to implement, especially when the non-linear term is easily differentiable. This method decomposes a MINLP problem into an NLP problem and a MILP problem. The NLP problem solves the entire problem with fixed binary variables, while the MILP problem approximates the non-linear term as a linear function and obtains optimal values for the binary variables used in the NLP (Aloufi, 2021). OA then iteratively solves these problems to obtain the optimal solution or the solution whose upper bound and lower bound are close enough. The utilization of these sophisticated methods has significantly contributed to reducing the time taken to solve complicated large-scale optimization problems, including supply chain problems.

2.7 State of the Art Summary and Conclusions

Omnichannel supply chain strategies are increasingly important to the foodservice distribution industry. While some industry players are pursuing this through the integration of stores into existing distribution supply chains, others are experimenting with new fulfillment methods to meet customer demand while minimizing overall supply chain cost. New model techniques and commercial solver capabilities can explore how transportation and inventory costs change under varying scenarios in a single, integrated methodology. This capstone project demonstrates how these learnings apply to our sponsor company's distribution network by building an integrated transportation-inventory model with omnichannel inspired multiple fulfilment modes and lateral transshipments.

3 DATA AND METHODOLOGY

In this chapter, we outline the methodology and data utilized to develop the inventory-integrated omnichannel network design model for our sponsor, with a focus on minimizing the overall supply chain cost. We will focus on the scope of the project, the data processing required, model formulation, and techniques used to improve model solve time.

3.1 Scope and Process

To ensure the scope of our research was well-defined, we focused solely on the supply chain activities that the sponsor company can manage and control. These activities include procurement from suppliers, inbound transportation from suppliers to distribution sites, warehouse handling and holding, outbound transportation from distribution sites to customers, and other transportation activities, such as interfacility transportation. We excluded external activities that are not manageable by the sponsor, such as procurement activities from further tiers of suppliers (e.g., 2nd and 3rd tier suppliers), sales from customers to end-users, and indirect managerial activities like depreciation or taxation. As a result, our model incorporated product cost, inbound and outbound transportation cost, omnichannel cost (e.g., direct ship cost and cross-docking cost), warehouse handling cost, and inventory holding cost. We did not consider any other costs as we assumed they would not have a significant impact on the entire supply chain network of the company. We provide detailed definitions and formulations for each cost in Section 3.4.

To build the model, we followed the approach shown in Figure 3. Prior to model development, we processed the data to convert it into formats that fit our formulas (e.g., changing cost units from distance to cases) and established the current state of the network (baseline). Once we had a clear understanding of the network and our hypotheses were established, we developed the model using the formulas outlined in Section 3.4 and obtained the optimized network. We also tested multiple fulfillment options by adding new elements to the current network, such as new direct transportation lanes between suppliers and customers.

Figure 3

Research Process

Process	Explanation	Corresponding Chapter
Data Collection / Cleansing Data	• Collect the data necessary to build the model • Correct errors and calculate estimations for unrealized information	3 DATA AND METHODOLOGY
As-Is Analysis / Building Hypothesis	• Specify predicted improvements as the result of the network design • Visualize the As-Is supply chain network	1 INTRODUCTION / 4 RESULTS AND DISCUSSION
Modeling	• Define the objective functions, the constraints, the decision variables • Code the optimization model	3 DATA AND METHODOLOGY
Model Execution	• Run the model for each product with different scenarios	4 RESULTS AND DISCUSSION
Result Analysis	• Visualize the To-Be supply chain network for each scenario • Interpret the results and pull insights	4 RESULTS AND DISCUSSION

3.2 Formulation Overview

In order to find an optimal structure of supply chain network and inventory allocations with minimized cost, we formulated our problem as an optimization problem. We opted for optimization over simulation as we required "prescriptive" analysis that would provide us with ideas on how to restructure our network, rather than validating existing scenarios. To consider the non-linear safety stock cost, we developed a mixed-integer non-linear programming problem (MINLP) as our optimization model. The subsequent sections provide details on key assumptions, formulations for our model, and algorithms to solve it.

3.3 Key Assumptions and Data Pre-processing

Due to the difficulties of mathematical models accurately replicating reality, we made some assumptions in terms of timeframe and period, demand, regions, and items. Specifically, our model is a single-period model that optimizes total flows for a one-year period (52 weeks). Our research focused solely on the supply chain within the continental United States, excluding Alaska, Hawaii, and Puerto Rico, which were considered geographically remote from the mainland and expected to have no influence on the result. In addition, our model is a single-item model (only one item is considered at once for one-time run of the model), meaning only one SKU is considered at a time. However, it is expandable to a multi-item model by adding more decision variables and constraints.

In Sections 3.3.1 - Section 3.3.10 ahead we list the key assumptions included for our major model parameters, including: item selection, product cost, inbound cost, warehouse cost, inventory cost, outbound cost, direct ship cost, central warehouse cost, and cross docking cost. For all parameters mentioned we used the most recent data available from our Sponsor ranging from April 1, 2022 – March 31, 2023.

3.3.1 Approach to Data Collection

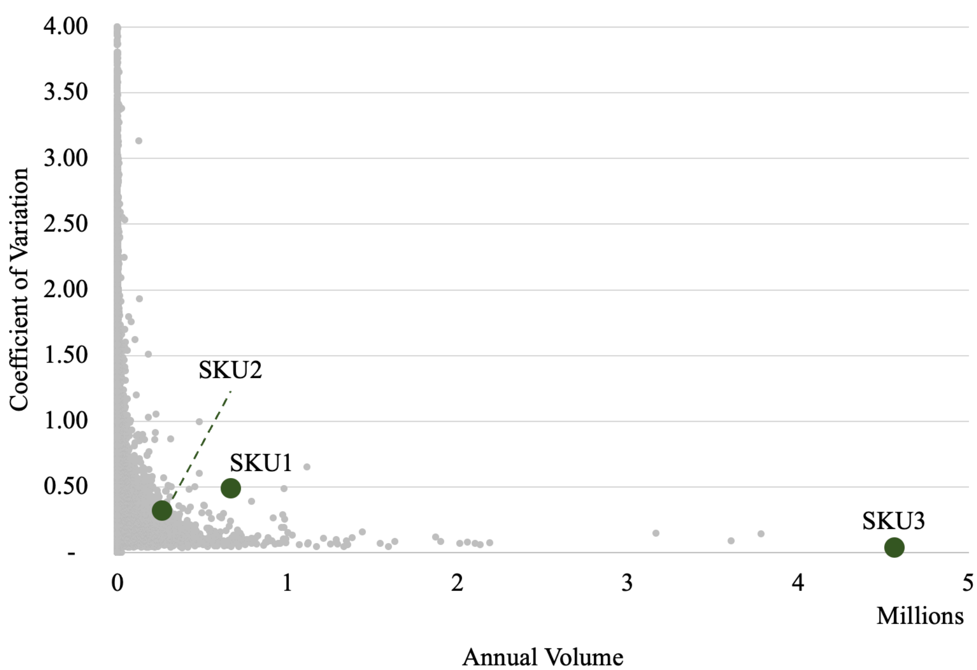
We relied on two primary sources for all data required: expert interviews and direct data collection from our sponsor company's databases. Over the past 6 months we interviewed subject matter experts (SMEs) at our sponsor company in the following functional areas: merchandising, inbound logistics, inventory management, network design, routing, supply chain finance, industrial engineering, and supply chain strategy. Data collection was handled via SQL queries that pulled annual data for relevant purchase orders, sales obligations, inventory levels, supplier characteristics, site characteristics, and customer locations. Additionally, we received financial summaries for all sites that had breakdowns by department and cost codes, for example: fuel, hourly wages, benefits, maintenance, and other relevant categories.

3.3.2 Item Selection

To determine which items to focus this project on, we analyzed and segmented our Sponsor company's entire SKU count for the annual period mentioned above using the ABC-XYZ approach. The product portfolio follows the common pattern of an "L" shape with a small number of products responsible for a significant portion of demand with low variability, while there is also a long tail of low volume products with high variability. We also added detail on item storage type, which could be either Dry, Cold (refrigerated), or Frozen. In addition, we added data on customer order lead time from historical sales information as well as inventory levels as of April 1, 2023. The combination of this data can be seen visually in Figure 4 below, with three items of interest highlighted.

Figure 4

ABC-XYZ Segmentation



These three items represent a cross-section of several of these important parameters that can help highlight how the model performs using different inputs. SKU1 is a dry item with moderate to high volume, relatively high variability of demand, a high percentage of customer orders with 2 days or more of lead time, and a moderate level of inventory. SKU2 is a cooler item with moderate volume, low variability, and also a high percentage of 2+ day customer lead time. This item has a higher overall inventory level. Finally, SKU3 is a frozen item with very high volume and very low variability, with close to 80% of orders from customers needing next day delivery. This item has a high inventory level. Table 1 below summarizes these factors.

Table 1

SKU Selection

Item Name	Storage Type	Annual Demand (Mil. Cases)	CV	2 + Days Order (%)	Model Lead Time	Inv Dollars (4/1/2023, $Mil)	Segment
SKU 1	Dry	0.7	0.51	75	2 day	1.5	AY
SKU 2	Cooler	0.4	0.21	61	2 day	2.5	AX
SKU 3	Frozen	4.5	0.07	22	1 day	2.2	AX

3.3.3 Product Cost, Distribution Site Location, and Warehouse Handling Cost

Product cost for all items was pulled from the Sponsor company's PO data during the time period of interest. An annual average $/case cost by supply point was used. Latitude and longitude of specific supply points were collected.

Warehouse handling cost was given by the Sponsor company's finance organization for each distribution site. This was also an annual $/case average for all cases flowing through the site. Latitude and longitude of specific distribution sites were collected.

3.3.4 Product Demand

Annual demand by item was aggregated into 3-digit zip code prefix demand nodes (approximately 900). We also calculated standard deviation of demand for each item by demand node. Lastly, we calculated latitude / longitude centroids for each of the nodes.

3.3.5 Inbound/Outbound Transportation Cost and Lane Availability

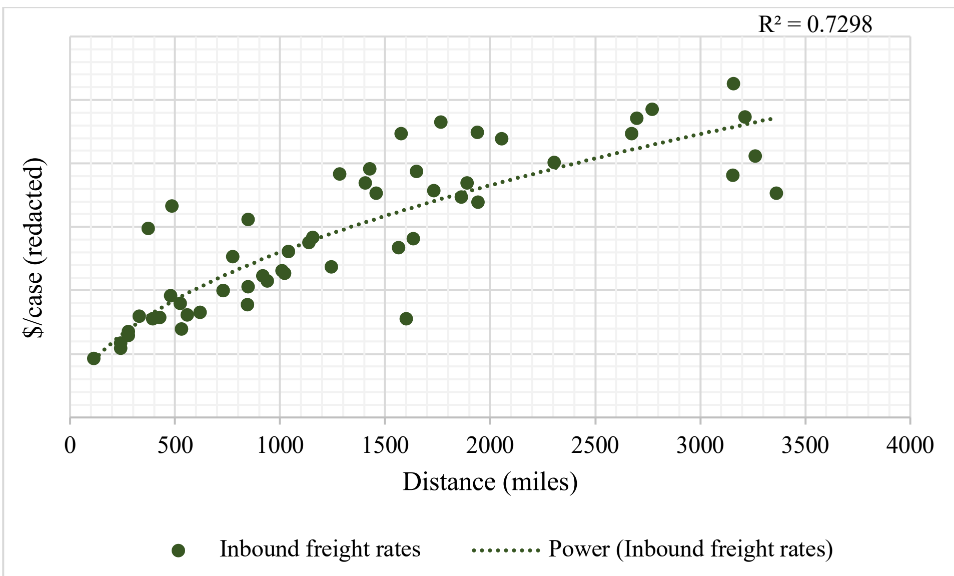
Given that the sponsor company operates a large distribution network with potentially millions of lanes and product combinations for both inbound and outbound, we used some simplifying assumptions to focus on the three items of interest. The most important of these is that as a general rule we limit the model to only existing lanes for both inbound and outbound across all products of the same storage type (including beyond our items of interest). For example, this means we would allow inbound lanes from the supply points of interest to distribution sites for all products the sponsor company currently receives from that supplier. On the outbound side, we allow only existing lanes of any product from distribution sites to demand nodes.

We made this simplifying assumption for three reasons in conjunction with feedback from the sponsor company. The first is that we want to avoid the issue of stranded items for either inbound or outbound. The sponsor company typically receives full truck load shipments from suppliers of multiple items to the distribution sites. Similarly, outbound flows are multi-item and multi-stop. Our cost model assumes a variable $/case cost for flows. Limiting the solution space to existing lanes would not require adding any fixed cost component of opening the new lane and would avoid, for example, the potential issue of the model suggesting the creation of a new outbound lane just to deliver one item, which is usually unrealistic. The second benefit of this approach is that it substantially constrains the possible solution space and allows for much faster solve times. The final reason is that limiting to existing lanes meant we could use existing cost data that would lead to more realistic conclusions without having to spend too much time estimating freight costs (not a core focus of this project).

Once we had the possible inbound lanes defined, we used annual historical cost information specific to that lane and item if it existed. If the lane existed but was for a different product, we used a simple regression model with distance as the independent variable to predict inbound rate data specific to the product of interest. Figure 5 below shows an example of this for SKU3. R-squared values for these regressions ranged from 0.55 to 0.73 depending on the item, which we deemed reasonable enough for the strategic level of this model. We used a similar regression approach to estimate lead times that was also based on distance.

Figure 5

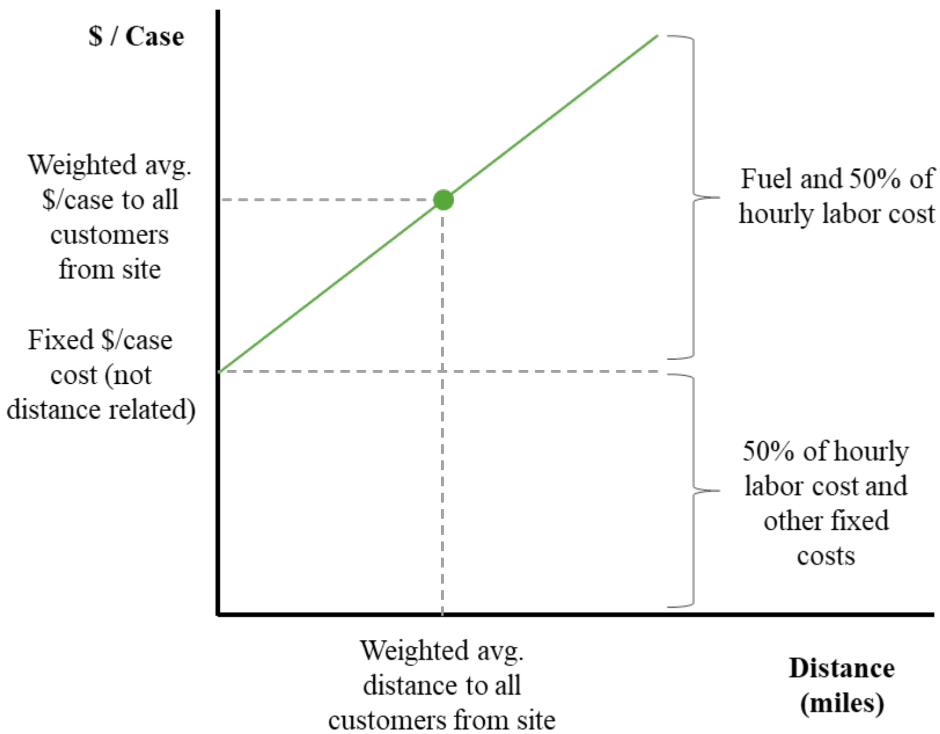
Regression for Inbound Freight Rates



Similarly for outbound lanes, once we had possible lanes defined, we used historical data to estimate $/case coefficients. In this case, we used a high-level approximation based on cost-to-serve data supplied by the sponsor company's finance team. This was broken down into fixed and variable components linked to each distribution site. The industrial engineering team of the sponsor company advised us that approximately 50% of a deliver partner's time is spent driving (to scale with distance), while the remainder is spent unloading and on other activities unrelated to distance. Linking that with the weighted average distance of volume for each site, we created a simple straight-line estimate of how outbound costs would scale depending on how far a demand node is located from a distribution site. See Figure 6 for an example of this approximation. Future improvements to this methodology could take more granular customer-level rather than site-level cost-to-serve data, which would then give more confidence in estimates further from the weighted average.

Figure 6

Approximation for Outbound Transportation Cost



Lastly, we used the well-known Haversine formula to estimate mile distance between two points using latitude and longitude and applied a universal 1.15 circuity factor to account for more realistic travel distances.

3.3.6 Inventory Cost Parameters

Our sponsor company uses a periodic review method for inventory replenishment. We collected historical review periods as weeks/order for each distribution site for each item. We also collected average lead time and standard deviation of lead time for each inbound lane using historical data. We calculated average annual inventory level for each item at each site using daily inventory numbers during the year of interest. Parameters for demand and demand variability would be calculated internally in the model based on flows chosen.

The sponsor company typically targets a fill rate of ~98% for items. To model target inventory levels, we created three views for comparison: actual average annual inventory, inventory based on actual historical flows using our inventory equation and an assumed fill rate of 98% (Type 2 service level), and a model estimated inventory level using a universal z-score of 1.282 (approx. 90% Type 1 service level). We adopted a universal z-score as it is significantly simpler to include in the model and roughly approximates the target fill rate given the sponsor company's typical order size and demand variability.

Lastly, we assume a 10% inventory holding cost percentage for all items.

3.3.7 Direct Ship Cost

To model a direct ship option from supplier to customer we made an exception to our assumption of no additional lanes. We created new lanes from these supply nodes to each demand node and estimated the approximate $/case cost of opening these new lanes using the cost profile of the nearest distribution center. This assumption was used to replicate local cost and logistical considerations. We then scaled these costs based on volume at the demand node and the assumption that customers would want at least once-weekly delivery of the item. This effectively made the cost "out of the money" for the vast majority of new direct ship lanes due to insufficient volume, however for a small group of high-volume demand nodes located close to supply nodes the costs were potentially in line with normal delivery methods.

3.3.8 Cross-docking / Peer Replenishment Cost

We modeled two variations of cross-site fulfillment: one, involving cross-docking between sites where the first site holds inventory and the second site only passes through items for delivery, and a second with peer-to-peer replenishment, in which the first site acts as a supplier to the second site and both sites hold inventory. Cross-docking was assumed to be universally lower in cost and thus preferred in every instance available. The model would be allowed to use each variant depending on the lead time requirement of the item and the distance between sites. Items with a 2-day lead time requirement would be given the option to cross dock for distances less than 2000 miles, whereas items with a 1-day requirement would only have time to cross-dock if the distance was less than 500 miles. Transfers between sites above the distance would still be allowed, however they would be required to use the more costly option of peer-to-peer replenishment.

In the first instance with cross-docking, we would need to model additional costs related to a site-to-site shuttle as well as a cross-docking handling cost at the second site. For each item, we opened up lanes from each site to every other site in the network set at 90% of the inbound cost estimated using our inbound regressions. This assumption is based on the premise that these shuttles would largely mirror similar inbound full truck load rates, that the shuttles would be part of a national cross-docking network that the sponsor company was considering establishing, and that a reasonable cost estimate for these flows would mirror the national inbound market rates with some discount based on not needing to support third-party profit margins. Looking at publicly available financial statements for large national freight companies, profit margins tended to be between 5-10%.

For the cross-docking warehouse handling cost, we used a simplifying assumption that this was equal to 50% of the historical warehouse handling costs already discussed in Section 3.3.2. The resulting $/case assumption was then verified with the sponsor company as reasonably similar to a cross docking pilot currently underway.

In the second instance of peer-to-peer replenishment, similar assumptions apply however the second site would be stocking inventory. As mentioned above, this would only be used if the customer lead time requirement was too short for the cross-site shuttle to arrive. In this case we applied the full warehouse handling cost of the second site, due to the need to do normal unload / storage / picking / loading, and we added a cycle stock adder to the flow. This adder was equivalent to:

Total cycle stock cost = Annual demand * Review period (weeks) / 52 / 2 * Holding cost

Simplified to cycle stock cost adder per case = Review Period / 52 / 2 * Holding cost

As a simplifying assumption, we model zero safety stock for this option. This is due to the short and predictable lead time for these shuttles (1-2 days) that effectively drive the safety stock requirement close to zero. In reality, implementation of this form of replenishment would likely have some nominal amount of safety stock, however, we did not find this simplifying assumption to substantially change model results.

3.3.9 Central Warehouse Cost

The sponsor company also has multiple central warehouses located across the US, which were typically used to drop ship items via third party carriers direct to customers. In some instances, the central warehouses also were used as inventory hubs for cross-docking into normal outbound lanes. We included central warehouses for both drop ship and as potential cross-docks sites in the model. Costs for drop ship were included from the sponsor company's finance organization. Cross-dock costs were modeled as discussed above in Section 3.3.8. We included new lanes to/from central warehouses if they did not exist using the estimation methods discussed previously.

3.3.10 Summary of Cost Elements for Each Item

As noted in section 3.3.2, the three items chosen have notably different demand profiles, customer lead times, and storage types. Additionally, they were chosen due to having different cost profiles as shown in Table 2 below. This figure summarizes all relevant cost elements for the annual period reviewed.

Table 2

Cost Elements for Each Item

Cost Summary (Thousand USD)	SKU 1		SKU 2		SKU 3	
	$K	%	$K	%	$K	%
Transportation Cost	2,832	7	3,593	6	15,467	20
Inbound Transportation Cost	1,410	4	2,761	5	5,871	8
Cross-dock Cost	-	0	-	0	1	0
Direct Ship Cost	-	0	-	0	-	0
Outbound Transportation Cost	1,422	4	832	1	9,595	13
Warehousing Cost	500	1	389	1	3,691	5
Inventory Cost	218	1	195	0	234	0
Cycle Stock Cost	71	0	70	0	61	0
Safety Stock Cost	147	0	125	0	173	0
Subtotal Logistics Cost	3,550	9	4,177	7	19,392	25
Product Cost	36,305	91	53,017	93	56,771	75
Total Supply Chain Cost	39,855	100	57,194	100	76,163	100

For SKU1, transportation cost is evenly split between inbound and outbound costs. Warehouse cost and inventory cost are minimal, and total logistics cost is approximately 9% of total supply chain cost. The evenly split transportation cost is due to having 5 supplier ship points across the US which limit total inbound miles. SKU2 is different in that inbound cost makes up 77% of transportation cost, due to having 2 supplier ship points both on the East Coast that need to deliver across the full network. Similar to SKU1, warehouse cost and inventory cost are minimal components of total supply chain cost. SKU3 is different once again in that outbound transportation cost makes up 62% of total transportation cost. Inventory cost again is minimal compared to other costs, however in this case warehouse cost is higher at 5% of total supply chain cost for the item. Product cost is also noticeably lower for this item compared to the others, at only 75% of total supply chain cost. These differences in cost profiles help demonstrate how the model optimizes differently as relative costs change.

3.4 Formulation

Based on the assumptions we made, we developed our MINLP model, with formulas that differ significantly by supply chain activity. In this section, we provide a summary of the formulas used for the model, with separate explanations provided for transportation, inventory, and procurement. One noteworthy element of our model is the introduction of decision variables x_{ikj}, which represents the fraction of demand at customer *j* fulfilled through distribution site *k* from supplier *i*. For example, if a customer 1 has a demand of 100 cases, and 60 cases are fulfilled through distribution site 1 from supplier 1, $x_{111} = 0.6 \left(= \frac{60}{100}\right)$. Thus, these are continuous decision variables ranging from 0 to 1, with the number of decision variables being equal to the number of possible transportation lanes among suppliers, distribution sites, and customers. The annotations, parameters, and decision variables used for our model are shown below.

Annotations:

$i \in S$	: set of suppliers
$k \in W$	: set of sites
$j \in C$	: set of customers

Parameters:

t_{ik}^{I}	: unit inbound transportation cost from supplier *i* to site *k*
t_{kj}^{O}	: unit outbound transportation cost from site *k* to customer *j*
t_{ij}^{D}	: unit direst ship cost from supplier *i* to customer *j*
w_k	: unit warehouse handling cost at site *k*
p_i	: unit product cost from supplier *i*
r_k	: weeks per order at site *k*
l_{ik}	: average lead time plus review period for lane *ik*
u_j	: standard deviation of demand at customer *j*
v_{ik}	: standard deviation of lead time for lane *ik*
D_j	: demand at customer *j*
h	: inventory holding cost percentage
P	: average product cost
Y	: weeks in a year (52 in our case)
z	: *z* score (in our case, 90% of cycle service level ($z \cong 1.282$))

Decision Variables:

x_{ikj} : fraction of demand at customer j fulfilled through distribution site k from supplier i
y_{ik} : whether supplier i serves distribution site k (binary)
x_{ij}^{D} : fraction of demand at customer j fulfilled directly from supplier i

3.4.1 Transportation Cost

In our optimization model, transportation cost was identified as the most significant component among all supply chain activities, as any changes in the supply chain structure directly impact it. For the sponsor company, transportation cost includes all the expenses associated with moving goods from one point to another, such as labor, fuel, and insurance. In our model, we considered three types of transportation costs: outbound transportation cost, inbound transportation cost, and other transportation costs (e.g., interfacility transportation cost and direct ship cost).

We defined inbound transportation cost as the product of the flow of cases from suppliers to distribution sites and the unit inbound transportation cost per case, which varies by lane (i.e., the combination of suppliers and distribution sites). This is expressed in the following formula:

$$\sum_{i \in S} \sum_{k \in W} \sum_{j \in C} t_{ik}^{I} D_j x_{ikj} \tag{1}$$

In the same way, we defined outbound transportation cost as the product of the flow of cases from distribution sites to customers and unit outbound transportation cost per case, which is:

$$\sum_{i \in S} \sum_{k \in W} \sum_{j \in C} t_{kj}^{O} D_j x_{ikj} \tag{2}$$

In our current model, we have considered only inbound and outbound transportation costs. However, in a later chapter, we have also included transportation costs for other lanes, such as direct shipping from suppliers to customers. To incorporate these additional costs, we have created new decision variables and constants based on the same principles as inbound and outbound transportation costs. We provide more details about these additional options in Section 3.6.

3.4.2 Warehouse Handling / Inventory Cost

When modifying the supply chain network, it is necessary to also adjust inventory positioning, as discussed in the state-of-the-art chapter. However, this can result in differences in the volume of items processed at each distribution site, leading to additional costs such as warehouse handling cost and inventory holding cost for the company. Warehouse handling cost refers to the cost of unloading, allocating, picking, packing, and loading items, which is primarily incurred by labor. To calculate this cost, we multiplied the number of cases processed by the unit warehouse handling cost, which varies by distribution site. The formula for warehouse handling cost is provided below:

$$\sum_{i \in S} \sum_{k \in W} \sum_{j \in C} w_k D_j x_{ikj} \tag{3}$$

Inventory holding cost is a significant factor to consider in strategic network design models, but it can be difficult to accurately account for all the associated cost factors. To address this, we introduced the concept of inventory holding cost percentage. This percentage represents the ratio of total yearly inventory holding cost to total yearly product cost, and for the sponsor's case, it was approximately 10%, as determined by recent years' data and input from the inventory manager. By using this simplified approach, we were able to incorporate inventory holding cost into our optimization model without the need for detailed cost breakdowns. With this concept, total inventory cost can be expressed as $Total\ Inventory\ Holding\ Cost = Inventory\ Holding\ Cost\ Percentage\ (\%) \times Average\ Product\ Cost\ \left(\frac{\$}{case}\right) \times Total\ Inventory\ Amount(case)$. In general, total inventory amount is summing of cycle stock, safety stock, and pipeline (WIP) stock.

Cycle stock is regarded as average amount of inventory to fulfill customers' demand, which is $CS = Total\ flow\ amount\ \times\ Weeks\ per\ order\ /\ 52\ /\ 2$. Total cycle inventory holding cost was therefore formulated as:

$$\frac{hP}{2Y} \sum_{i \in S} \sum_{k \in W} \sum_{j \in C} r_k D_j x_{ikj} \tag{4}$$

Safety stock is another type of inventory, and the standard safety stock equation with demand and supply variability considerations is:

$$\sqrt{LT\sigma_D^2 + D^2\sigma_{LT}^2} \tag{5}$$

Given that our decision variable x_{ikj} represents the fraction of demand for each customer, the combined variance of demand for distribution site *k* when using supplier *i* ($\sigma_{D_{ik}}$) can be written as below, assuming independent normal distributions for each customer demand:

$$\sigma_{D_{ik}}^2 = \sum_j u_j^2 x_{ikj}^2 \tag{6}$$

Note that a new demand for multiple combinations of *i* and *k* is a weighted sum of demands for all the served customers, and x_{ikj} should be squared to calculate the combined variance. Also, a square of weekly demand through supplier *i* and distribution site *k* (D_{ik}) is:

$$D_{ik}^2 = \sum_j \left(\frac{D_j x_{ikj}}{Y}\right)^2 = \sum_j \frac{D_j^2}{Y^2} x_{ikj}^2 \tag{7}$$

With these equations in mind, safety stock for every combination of supplier *i* and distribution site *k* is:

$$z\sqrt{l_{ik}\sum_j u_j^2 x_{ikj}^2 + v_{ik}^2 \sum_j \frac{D_j^2}{Y^2} x_{ikj}^2} \tag{8}$$

Assuming each distribution site must choose only one supplier, which is a very common situation for the sponsor company, the total amount of safety stock becomes the equation (9) as shown below. Note that without this assumption, safety stock at each distribution site can be duplicated for each supply lane.

$$z\sum_i \sum_k \sqrt{l_{ik}\sum_j u_j^2 x_{ikj}^2 + v_{ik}^2 \sum_j \frac{D_j^2}{Y^2} x_{ikj}^2} \tag{9}$$

Hence, the total safety stock cost is:

$$hPz\sum_i \sum_k \sqrt{l_{ik}\sum_j u_j^2 x_{ikj}^2 + v_{ik}^2 \sum_j \frac{D_j^2}{Y^2} x_{ikj}^2} = hPz\sum_i \sum_k \sqrt{\sum_j \left(l_{ik}u_j^2 + v_{ik}^2 \frac{D_j^2}{Y^2}\right) x_{ikj}^2} \tag{10}$$

The impact of pipeline (work-in-progress) stock is relatively low, and it is least likely to be affected by changes in the supply chain network for the sponsor due to differing payment terms. Therefore, after discussing with the sponsor company, we decided to ignore pipeline stock in our model.

3.4.3 Product Cost

Since the sponsor company procures items from external suppliers and sells them to customers, the cost of product procurement must be added for each item bought. Since product cost is differs by supplier, it can be formulated as below:

$$\sum_{i \in S} \sum_{k \in W} \sum_{j \in C} p_i D_j x_{ikj} \tag{11}$$

In the sponsor's case, all the related costs such as system transaction fee were included in product cost itself.

3.4.4 Formulation Overview

Next, with related constraints and a new binary decision variable y_{ik} meaning whether supplier *i* supplies to distribution site *k*, our entire inventory-integrated supply chain design problem (ISCDP) can be defined as follows:

(ISCDP)

$minimize$

$$\sum_{i\in S}\sum_{k\in W}\sum_{j\in C} t_{ik}^{I} D_j x_{ikj} + \sum_{i\in S}\sum_{k\in W}\sum_{j\in C} t_{kj}^{O} D_j x_{ikj} + \sum_{i\in S}\sum_{k\in W}\sum_{j\in C} w_k D_j x_{ikj} + \frac{hP}{2Y}\sum_{i\in S}\sum_{k\in W}\sum_{j\in C} r_k D_j x_{ikj}$$

$$+hPz\sum_{i\in S}\sum_{k\in W}\sqrt{\sum_{j\in C}\left(l_{ik}u_j^2 + v_{ik}^2\frac{D_j^2}{Y^2}\right)x_{ikj}^2} + \sum_{i\in S}\sum_{k\in W}\sum_{j\in C} p_i D_j x_{ikj} \tag{12}$$

$subject\ to$

$$\sum_{i\in S}\sum_{k\in W} x_{ikj} = 1, \qquad \forall j \in C \tag{13}$$

$$\sum_{i\in S} y_{ik} \le 1, \qquad \forall k \in W \tag{14}$$

$$x_{ikj} \le y_{ik}, \qquad \forall i,k,j \in S,W,C \tag{15}$$

$$x \ge 0, \ \ y = \{0,1\} \qquad (16),(17)$$

The first constraint (13) ensures demand for each customer is fulfilled. The second constraint (14) forces the model to choose at most one supplier for each distribution site. The third constraint (15) ensures that no flow occurs between supplier *i* and distribution site *k* unless supplier *i* supplies to distribution site *k*. The last constraints (16), (17) prevent negative flows and make *y* binary. As the objective function (12) shows, our problem is clearly a MINLP problem with the safety stock cost term and binary variables.

To simplify annotations for later use, we combined multiple terms with:

$$a_{ikj} = t_{ik}^{I} D_j + t_{kj}^{O} D_j + D_j w_k + \frac{hP}{2Y} r_k D_j + p_i D_j$$

$$b_{ikj} = l_{ik}u_j^2 + v_{ik}^2\frac{D_j^2}{Y^2}$$

$H = hPz$

Therefore, the entire problem can be written as follows:

(ISCDP')

$$minimize \sum_{i \in S} \sum_{k \in W} \sum_{j \in C} a_{ikj} x_{ikj} + H \sum_{i \in S} \sum_{k \in W} \sqrt{\sum_{j \in C} b_{ikj} x_{ikj}^2} \tag{18}$$

$subject\ to$

$$\sum_{i \in S} \sum_{k \in W} x_{ikj} = 1, \qquad \forall j \in C$$

$$\sum_{i \in S} y_{ik} \leq 1, \qquad \forall k \in W$$

$$x_{ikj} \leq y_{ik}, \qquad \forall i, k, j \in S, W, C$$

$$x \geq 0, \;\; y = \{0,1\}$$

3.5 Tools Utilized

To take advantage of modeling flexibility, our models were developed and implemented using Python 3 and Gurobi version 9.5.2, and executed on MIT's Supercloud, providing a secure and fast environment for running the code. Excel 16.72 and Tableau 2023.1 were used for further analysis and visualization of the results. Additionally, Microsoft Excel and Access were utilized for preliminary analyses, such as product segmentation.

3.6 Omnichannel Fulfillment Options

After developing the model outlined in Section 3.4, we created several options to help the sponsor company identify the best supply chain strategy. Specifically, we proposed four options to modify supply chain network structure for the targeted items:

a. Inventory-integrated model / The suggested formulation in Section 3.4.5.

b. Inventory-integrated model with direct ship

c. Inventory-integrated model with peer-to-peer cross-docking

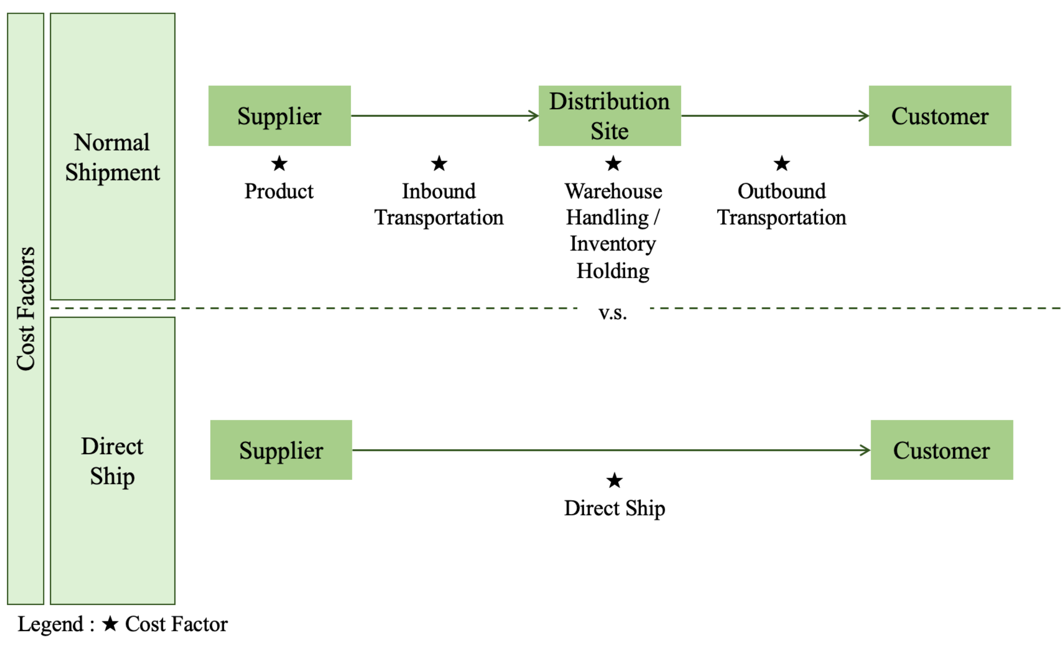
d. Use of Central Warehouse

3.6.1 Direct Ship

Direct ship literally means shipping products from suppliers directly to customers, for which we created new lanes (between supplier *i* and customer *j*). Introducing this option produced a new trade-off between multiple costs. As shown in Figure 7, choosing direct ship makes sense when direct shipping cost is lower than the sum of inbound and outbound transportation cost, warehouse handling cost, and inventory cost. Due to its higher costs, direct shipping is not typically preferred over standard transportation. Therefore, the potential benefits of using direct shipping lie in the savings generated from the elimination of distribution site-related expenses.

Figure 7

Difference in Cost Factors between Normal Shipment and Direct Ship



To deal with direct ship, we introduced new decision variables x_{ij}^D, the fraction of demand of customer *j* fulfilled with direct ship from supplier *i*.

$$\sum_{i\in S}\sum_{j\in C} t_{ij}^{D}\, D_j x_{ij}^{D} \tag{19}$$

Then the problem definition becomes:

$$minimize \sum_{i\in S}\sum_{k\in W}\sum_{j\in C} a_{ikj}x_{ikj} + H\sum_{i\in S}\sum_{k\in W}\sqrt{\sum_{j\in C} b_{ikj}x_{ikj}^{2}} + \sum_{i\in S}\sum_{j\in C} t_{ij}^{D}\, D_j x_{ij}^{D} \tag{20}$$

$subject\ to$

$$\sum_{i\in S}\sum_{k\in W} x_{ikj} + \sum_{i\in S} x_{ij}^{D} = 1, \qquad \forall j \in C \tag{21}$$

$$\sum_{i\in S} y_{ik} \leq 1, \qquad \forall k \in W$$

$$x_{ikj} \leq y_{ik}, \qquad \forall i,k,j \in S,W,C$$

$$x \geq 0,\ \ y = \{0,1\}$$

3.6.2 Cross-docking and Peer Replenishment

The next option we considered was cross-docking and peer replenishment, which involved creating new lanes for moving products from one site to another (lateral transshipment), something that the sponsor company is now experimenting with. While this would increase transportation costs and lead time and add a new cost element in the form of cross-docking costs, the company could benefit from lower warehouse handling or inventory holding costs at the first distribution site.

The cost structure associated with this fulfillment option is dependent on the distance between the first distribution site and the second distribution site. If the lead time is sufficiently short, the second distribution site can assume immediate shipment from it, thereby eliminating the need for inventory. In such cases, the cost of cross-docking is calculated as the sum of transportation costs from the first distribution site to the second distribution site, and from the second distribution site to the customer, warehouse handling and inventory cost at the first distribution site, and the cost of the cross-docking process at the second distribution site. However, if the lead time exceeds the threshold, the cost of handling and inventory at the second distribution site must also be considered. It is important to note that the variability of supply lead time for the second distribution site was disregarded for our case due to the stability of the lead time for internal shuttle shipment, which results in the assumption that the safety stock volume is negligible, with the fact that the coefficient of variation was also very small. To incorporate this option into our model, we developed the following algorithm:

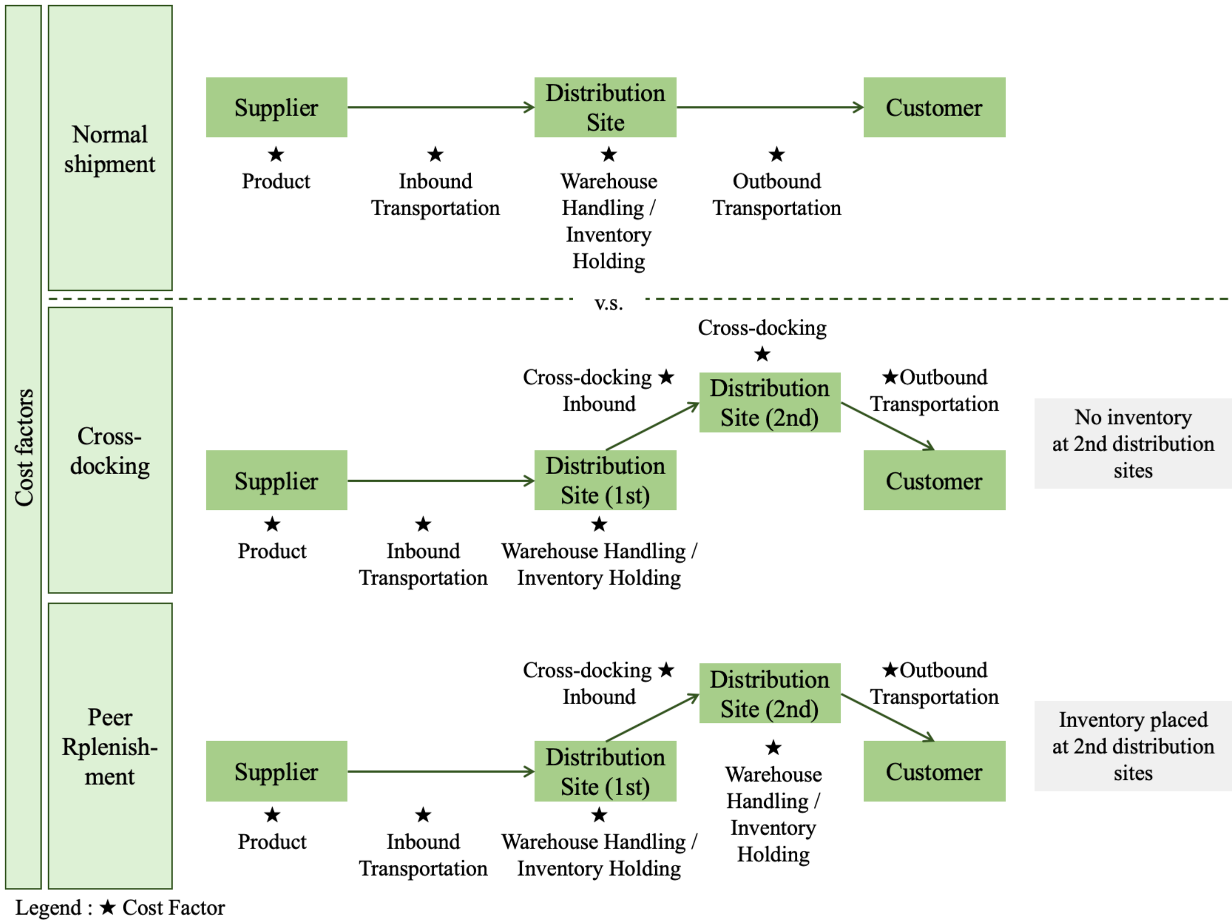
Algorithm 1 Cross-docking / Peer Replenishment Option Optimization

1: For each distribution site-customer combination, do:
  If the distance between the distribution site and the customer > the threshold,
    let $t'_{kj} =$ min {unit inbound transportation cost + unit outbound transportation cost + unit cross-docking cost for any cross-docking distribution sites possibly used for the specified distribution site-customer combinations},
  otherwise,
    let $t'_{kj} =$ min {unit inbound transportation cost + unit outbound transportation cost + unit warehouse handling cost and unit cycle stock cost for any cross-docking distribution sites possibly used for the specified distribution site-customer combinations}.
2:     Add s_{kj}, the intermediate distribution site with the minimum cost for cross-docking from distribution site k to customer j to the set W'.
3:   If $t^O_{kj} > t'_{kj}$,
    replace t^O_{kj} with t'_{kj}.
4: Execute optimization.
  If $x_{ikj} > 0 \ \forall s_{kj} \in W'$,
    keep s_{kj},
  otherwise,
    exclude s_{kj} from W'.
5: Return the optimization result and W'.

As shown in the algorithm and Figure 8, the cross-docking / peer replenishment model virtually reduces outbound transportation cost for specific lanes, which may incentivize the whole network to shift flows from other lanes to them to take advantage of cheaper transportation cost or cheaper warehouse handling / inventory cost at the first distribution site, especially with pooling effect of safety stock. From a practical business perspective, cross-docking would be advantageous in two situations: when large inbound rate differences exist or large cost structure differences between sites exist.

Figure 8

Difference in Cost Factors between Normal Shipment and Cross-docking

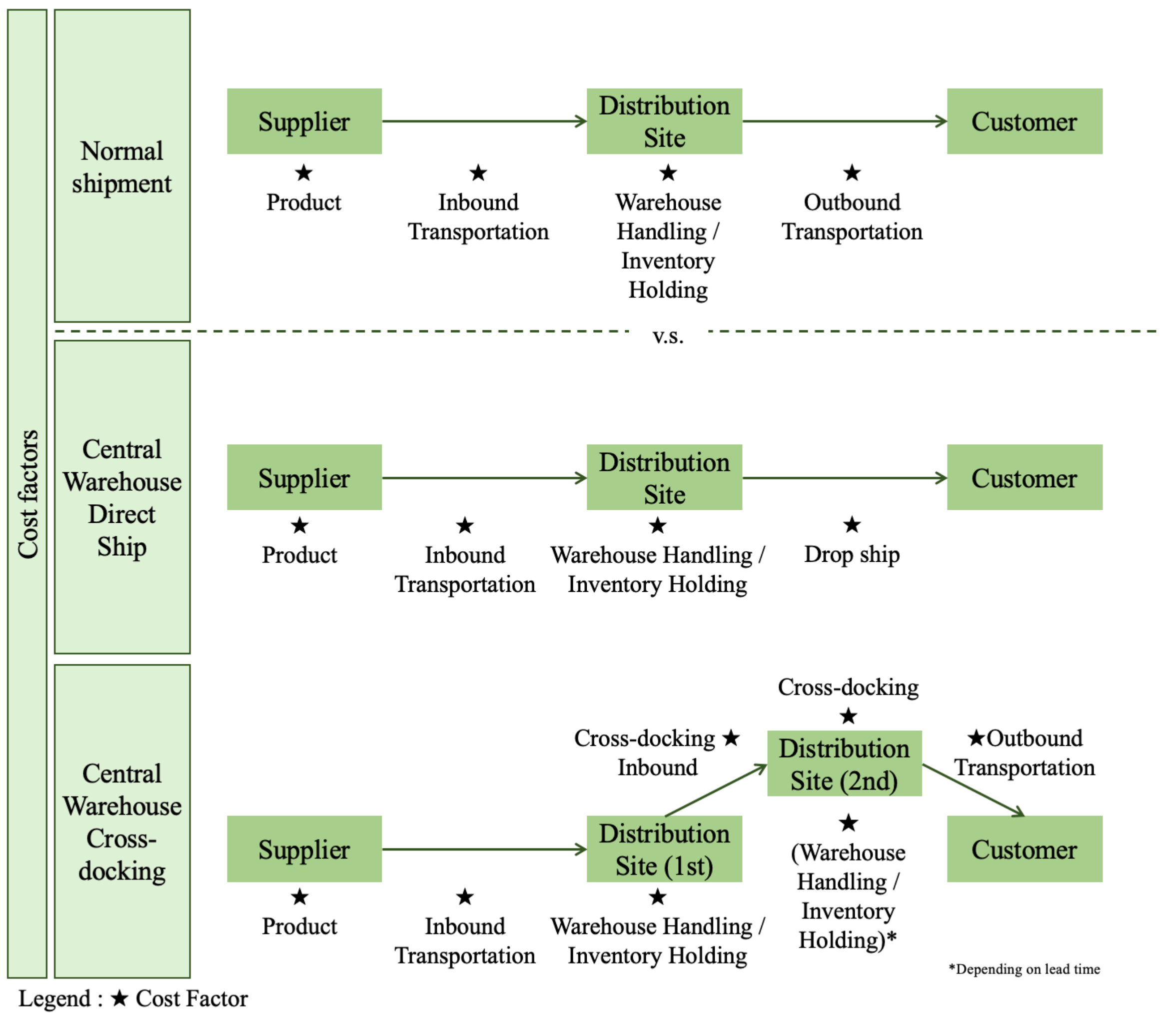


3.6.3 Central Warehouses

The third omnichannel option involves the use of central warehouses, which serve the same purpose as regular distribution sites but are typically larger and have lower warehouse handling and inventory costs. However, shipment methods from central warehouses are limited to two options: either through drop ship to customers using third-party logistics, which can be more expensive than standard outbound transportation, or by cross-docking at other distribution sites. Consequently, there is a trade-off between reduced warehousing or inventory costs and the potential for increased transportation expenses. Since central warehouses can be treated similarly to standard distribution sites, with the exception of their costing structure, no adjustments are needed to our model itself. We simply need to modify the parameters of our cross-docking model. The difference between normal shipment and shipment from central warehouses is summarized in Figure 9.

Figure 9

Difference in Cost Factors between Normal Shipment and Central Warehouses



3.7 Algorithm

Given that our problem is a MINLP, a computationally complex problem known to be NP-hard, it is crucial to discuss the various algorithms that can efficiently solve the problem with an acceptable level of optimality. In this chapter, we provide a comprehensive overview of the methods we employed to tackle the MINLP problem and achieve the desired outcomes.

3.7.1 Gurobi Optimizer and Reformulation to Quadratic Programming

Gurobi Optimizer is a mathematical optimization solver that solves LP, QP, and IP (Luce, 2022). To solve non-linear programming problems, non-linear terms must be quadratic ($ax^2 + bx + c$), which lead to reformulation of our objective function and constraints. Letting $Z_{ik}^2 \geq \sum_j b_{ikj} x_{ikj}^2 \geq 0$, our formulation became quadratic, and we call the problem "quadratic inventory-integrated supply chain design problem (QISCDP).

(QISCDP)

minimize

$$\sum_{i \in S} \sum_{k \in W} \sum_{j \in C} a_{ikj} x_{ikj} + H \sum_{i \in S} \sum_{k \in W} \sqrt{Z_{ik}^2} =$$

$$\sum_{i \in S} \sum_{k \in W} \sum_{j \in C} a_{ikj} x_{ikj} + H \sum_{i \in S} \sum_{k \in K} Z_{ik} \tag{22}$$

subject to

$$\sum_{j \in C} b_{ikj} x_{ikj}^2 \leq Z_{ik}^2, \qquad \forall i, k \in S, W \tag{23}$$

$$\sum_{i \in S} \sum_{k \in W} x_{ikj} = 1, \quad \forall j \in C$$

$$\sum_{i \in S} y_{ik} \leq 1, \quad \forall k \in W$$

$$x_{ikj} \leq y_{ik}, \quad \forall i, k, j \in S, W, C$$

$$x \geq 0, \;\; y = \{0,1\}$$

As the entire problem now comprises solely of quadratic and linear constraints, it can be efficiently solved using Gurobi optimization software. Furthermore, owing to the convexity of the objective function (proof of which can be found in the appendix, Proof 1), any local optimal solutions that are feasible are also globally optimal, and Gurobi can effectively identify and reach such solutions. By default, Gurobi implements branching algorithms such as branch-and-cut for integer decision variables, thereby providing a near-optimal solution for our problem (Gurobi Optimization, 2017). Our parameters for the mixed-integer programming thresholds (MIPGap, FeasibilityTol, and IntFeasTol) are all 1e-08. A similar reformulation can be found in You and Grossmann's research (2008).

3.7.2 Outer Approximation for Inventory-integrated Network Optimization

In our pursuit of a faster solution for the ISCDP problem, we devised a bespoke algorithm utilizing outer approximation, which effectively divides the problem into two distinct problems: an NLP master problem and a MILP subproblem. Since our problem can be viewed as a quadratic programming problem when the supplier binary decision variables are fixed, and the objective function is both convex and differentiable, except the point of $x_{ikj} = 0$, we recognized that leveraging outer approximation could significantly expedite the solving process.

First, the algorithm solved the NLP master problem, with initial fixed y_{ik} decision variables and obtained the optimal x_{ikj} and the upper bound. This, in turn, facilitated the reformulation of the objective function by introducing new decision variables ζ_{ik}^{OA} and linearized approximation constraints.

$$minimize\ \zeta_{ik}^{OA} \tag{23}$$

$subject\ to$

$$\zeta_{ik}^{OA} \geq f(x^*) + \nabla f(x^*)^T (x - x^*) \tag{24}$$

Except the point of non-differentiability $\sum_{j \in C} x_{ikj} = 0$, with x_{ikj}^* optimal solutions, the right-hand side of the constraint above can be revised as:

$$f(x^*) + \nabla f(x^*)^T (x - x^*) = \sum_{j \in C} a_{ikj} x_{ikj}^* + H \sqrt{\sum_{j \in C} b_{ikj} x_{ikj}^{*2}}$$

$$+ \begin{bmatrix} a_{ik1} + H \dfrac{b_{ik1} x_{ik1}^*}{\sqrt{\sum_{j \in C} b_{ikj} x_{ikj}^{*2}}} & a_{ik2} + H \dfrac{b_{ik2} x_{ik2}^*}{\sqrt{\sum_{j \in C} b_{ikj} x_{ikj}^{*2}}} \cdots & a_{ikN} + H \dfrac{b_{ikN} x_{ikN}^*}{\sqrt{\sum_{j \in C} b_{ikj} x_{ikj}^{*2}}} \end{bmatrix} \begin{bmatrix} x_{ik1} - x_{ik1}^* \\ x_{ik2} - x_{ik2}^* \\ \vdots \\ x_{ikN} - x_{ikN}^* \end{bmatrix} \tag{25}$$

Note that the square rooted term is a weighted l_2 norm. As for the non-differentiable point, we assumed sub-gradient of the second term of the objective function as zero (refer to Proof 2).

Solving MILP with the linearized term gives the optimal solutions for x_{ikj} and y_{ik} and the lower bound, which then were used as the termination criteria and the inputs for the next iteration. The entire algorithm is shown below:

Algorithm 2 Outer Approximation for ISCDP

1: Initialize $LB = -\infty, UB = \infty,$
$\epsilon = 0.001\ or\ 0.0001, k = 1, \quad y_{ik}^* \in Y\ (Y\ is\ an\ intial\ feasible\ set)$

2: While $(gap > \epsilon\ \ AND\ \ k \leq 100)$ do

3: Solve NLP master problem with y_{ik}^* and get x_{ikj}^*
Update UB

4: Add an outer approximation constraint to MILP subproblem:
If $\sum_{j \in C} x_{ikj} \neq 0,$
$\zeta_{ik}^{OA} \geq f(x^*) + \nabla f(x^*)^T (x - x^*),$
otherwise,
$\zeta_{ik}^{OA} \geq f(x^*) + \nabla g(x^*)^T (x - x^*)$
where $g(x) = a_{ikj} x_{ikj}$

5: Solve MILP subproblem and get $x_{ikj}^*\ and\ y_{ik}^*$
Update LB

6: Add an integer cut to prevent the same combinations of binary variables:
$IC^k = \left\{\sum_{b,b' \in B^k} y_{bb'}^* - \sum_{n,n' \in N^k} y_{nn'}^* \leq \left|B^k\right| - 1\right\}$
$where\ B^k\ is\ a\ basis\ set\ and\ N^k\ is\ a\ non\ basis\ set\ for\ y_{ik}^*$

7: Calculate $gap = (UB - LB)/LB$
Update $k = k + 1$

8: End while

9: Return UB and corresponding solutions $x_{ikj}^*\ and\ y_{ik}^*$

4 RESULTS AND DISCUSSION

This section walks through the model results for each item. To ensure an apples-to-apples comparison, we change only the flows between the "As-Is" baseline and the "Model" solution. The same cost logic is used to calculate total dollar reductions and the total amount of cases is the same. For the flows, we show the differences moving from the current results of the basic three-tier model to the new model with additional fulfillment options and cross-docking. For the inventory comparisons, we include three scenarios: the actual current inventory level, a target inventory level using the standard equation and "As-Is" flows, and the full model solution for inventory. Finally, we show for each item how the model suggests changes to inventory for each distribution site.

4.1 SKU1 Model Results

As a reminder, SKU1 is a dry item with moderate to high volume, relatively high variability of demand, a high percentage of customer orders with 2 days or more of lead time, and a moderate level of inventory. As can be seen in Table 3, inbound and outbound costs for SKU1 are relatively balanced in the historical data.

Table 3

Cost Result Summary for SKU1

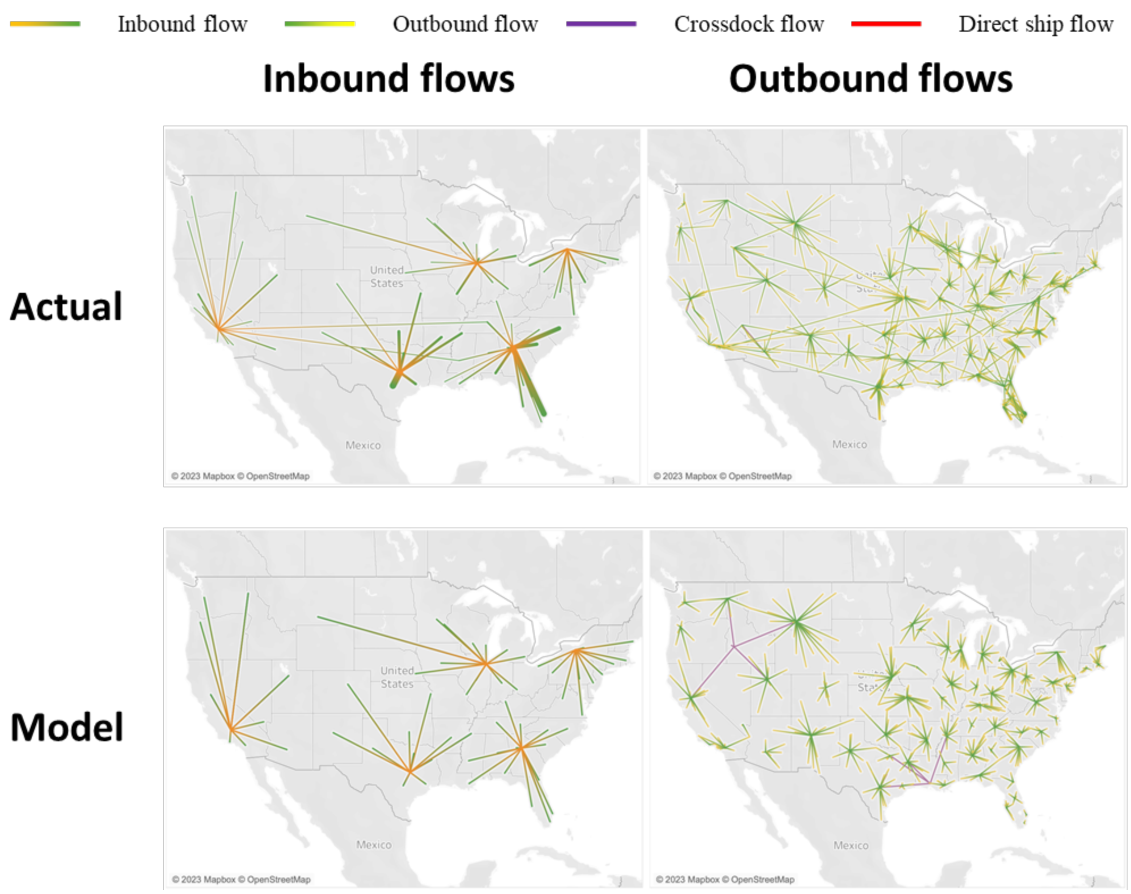
Cost Summary (Thousand USD)	As-Is ($)	Model ($)	Difference to As-Is ($)	Difference (%)
Transportation Cost	2,832	2,592	(240)	-8.5
Inbound Transportation Cost	1,410	1,356	(54)	-3.8
Cross-dock Cost	-	9	9	N/A
Direct Ship Cost	-	66	66	N/A
Outbound Transportation Cost	1,422	1,160	262	-18.4
Warehousing Cost	500	483	(17)	-3.5
Inventory Cost	218	209	(9)	-4.1
Cycle Stock Cost	71	66	(5)	-6.6
Safety Stock Cost	147	143	(4)	-2.9
Subtotal Logistics Cost	3,550	3,284	(267)	-7.5

Product Cost	36,305	36,273	(32)	-0.1
Total Supply Chain Cost	39,855	39,556	(299)	-0.7

The model results show an 8.5% reduction in transportation cost, which comes primarily through a large outbound transportation reduction of 18.4%. Major outbound realignments include a shift of volume towards the Central Florida site, which is centrally located in the state and minimizes outbound distance to customers, as well as a shift away from the high-cost Chicago site to the lower-cost Wisconsin site. Warehouse handling cost is also reduced due to this shift from high-cost to low-cost sites. Inventory cost is down 4.1%, primarily due to a new direct ship lane that opens up from the Texas supply point to a large demand node in Central Texas. This also has additional effects in lowering inbound, warehouse, and outbound costs in the Texas area as none of these cases need to flow through the sponsor company network. The model does open 2 cross-dock consolidation flows for this item, in the Louisiana site and the Idaho site. Inventory is no longer carried at these sites and instead is drawn from surrounding distribution sites. Figure 10 shows a national level visualization of the flows including historical actual flows and the new model flows.

Figure 10

Network Flow Comparison for SKU1



SKU1 has had some out-of-stock issues over the past year, and this is reflected in the inventory data shown in Figure 11. While the sponsor company usually targets ~98% fill rate to customers, SKU1 achieved ~88-92% depending on the distribution site. Using the inventory equation discussed in Section 3.4.2 with actual historical flows, target inventory levels are 76% higher than actual to achieve a 98% fill rate target. The integrated model solution, however, brings back inventory levels close to actual levels due to the flow realignments, cross-docking, and new direct ship lane.

Figure 11

Inventory Comparison for SKU1

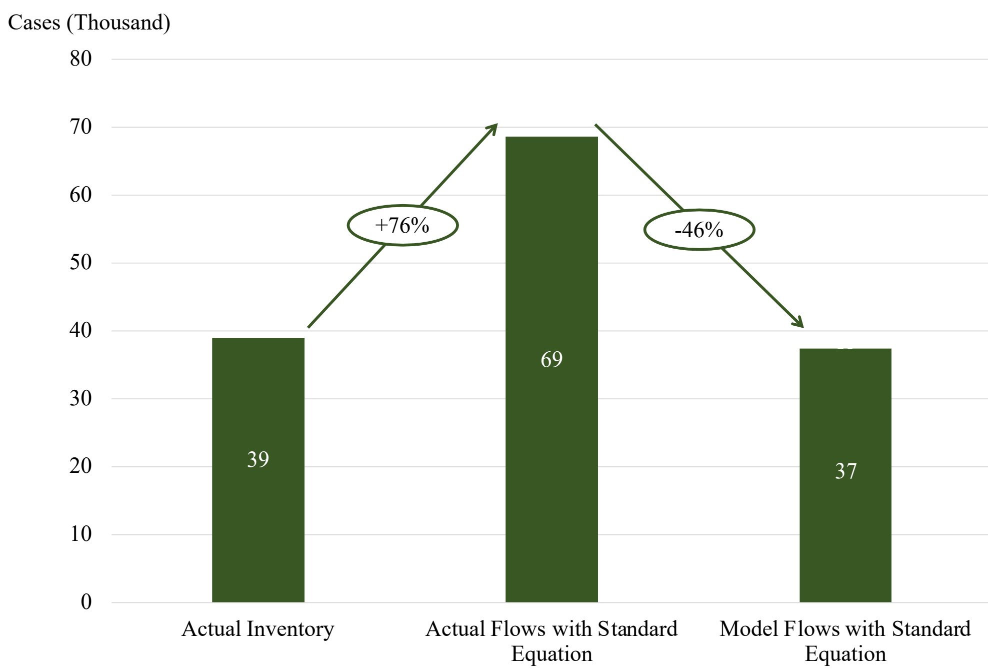
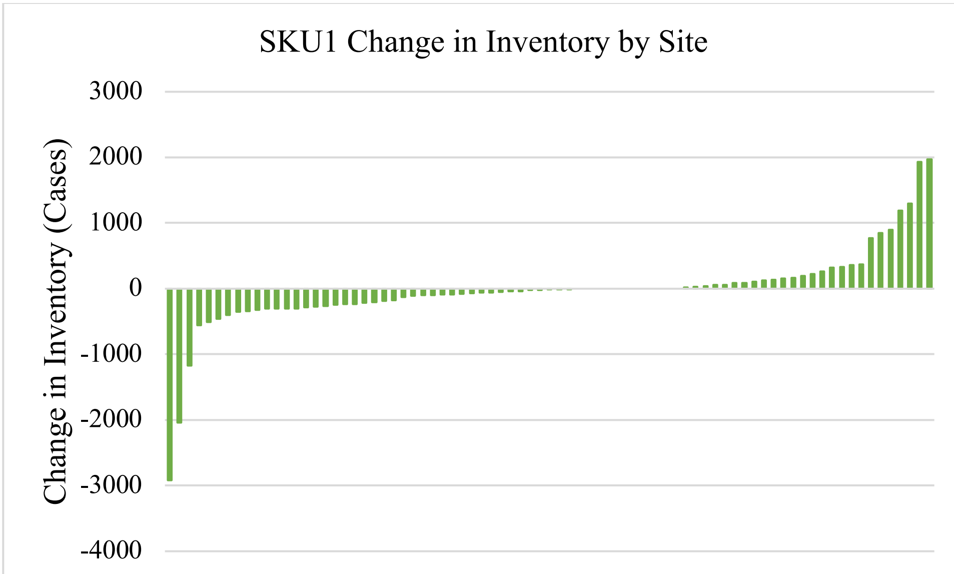


Figure 12 compares model inventory levels with historical actuals by site. Most sites have only minimal changes, however, the model chooses not to use central warehouses for this item. Currently, the sponsor company does have substantial inventory at central warehouses for infrequent shipments to distribution sites. This appears to be primarily backup storage and replenishment. The model adds over 1,000 cases of inventory to four distribution sites. These sites have historically had out of stock issues and have a mix of long supplier lead times combined with variable demand (coefficient of variation ranging from 0.8 to 2.0).

Figure 12

Inventory Comparison by Site for SKU1



4.2 SKU2 Model Results

As a reminder, SKU2 is a cooler item with moderate volume, low variability, and also a high percentage of 2+ day customer lead time. This item has a relatively high overall inventory level. As can be seen in Table 4, inbound transportation makes up the majority of total transportation cost.

Table 4

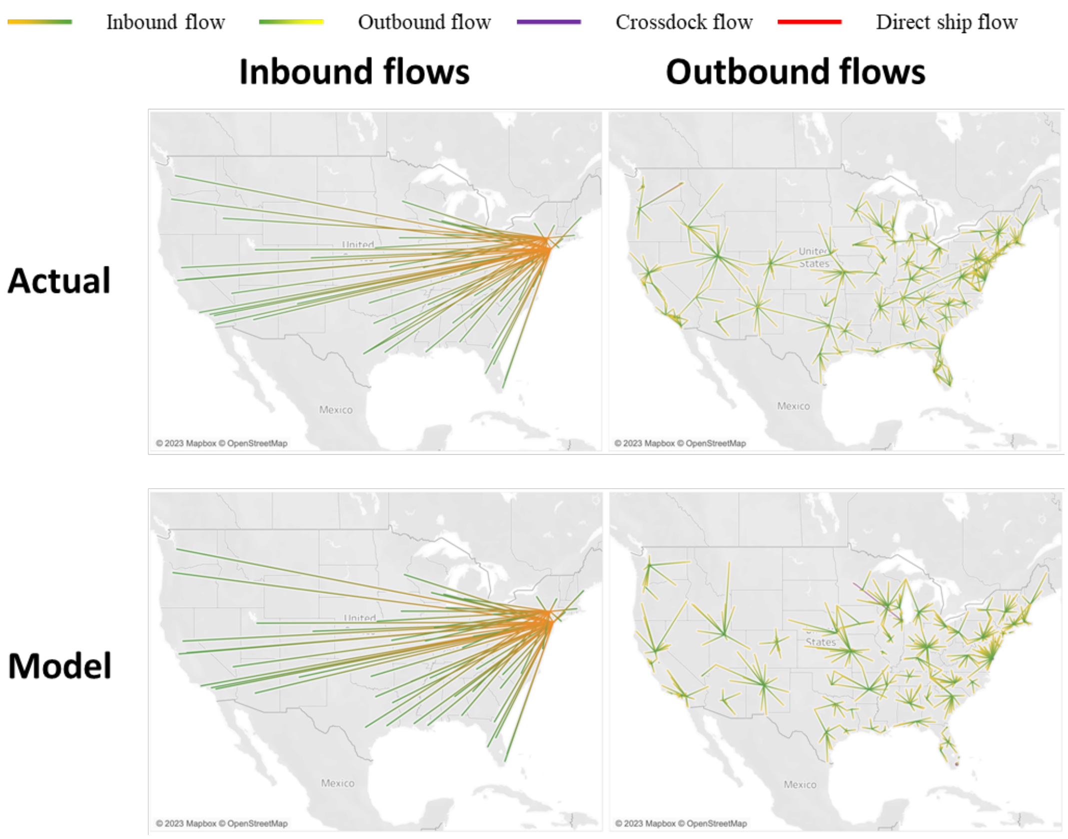
Cost Result Summary for SKU2

Cost Summary (Thousand USD)	As-Is ($)	Model ($)	Difference to As-Is ($)	Difference (%)
Transportation Cost	3,593	3,471	(122)	-3.4
Inbound Transportation Cost	2,761	2,568	(193)	-7.0
Cross-dock Cost	-	30	30	N/A
Direct Ship Cost	-	-	-	N/A
Outbound Transportation Cost	832	873	41	4.9
Warehousing Cost	389	379	(9)	-2.4
Inventory Cost	195	145	(50)	-25.7
Cycle Stock Cost	70	75	6	8.2
Safety Stock Cost	125	69	(56)	-44.6
Subtotal Logistics Cost	4,177	3,995	(182)	-4.3
Product Cost	53,017	53,011	(6)	0.0
Total Supply Chain Cost	57,194	57,006	(187)	-0.3

SKU2 has the smallest reduction in total logistics cost, with the majority of the savings coming from inbound transportation cost and safety stock reduction. As SKU2's product cost is significantly higher than the other SKUs, the reduction in safety stock had a greater impact. Although inbound transportation cost was reduced, outbound transportation cost increased as the model prioritized reducing the high inbound cost. Warehouse handling and product cost were minimally changed. The model did open two small cross dock lanes however these did not substantially reduce the overall supply chain cost. No direct ship lanes were opened, and central warehouse was not used. Figure 13 shows a national level visualization of the flows including historical actual flows and the new model flows.

Figure 13

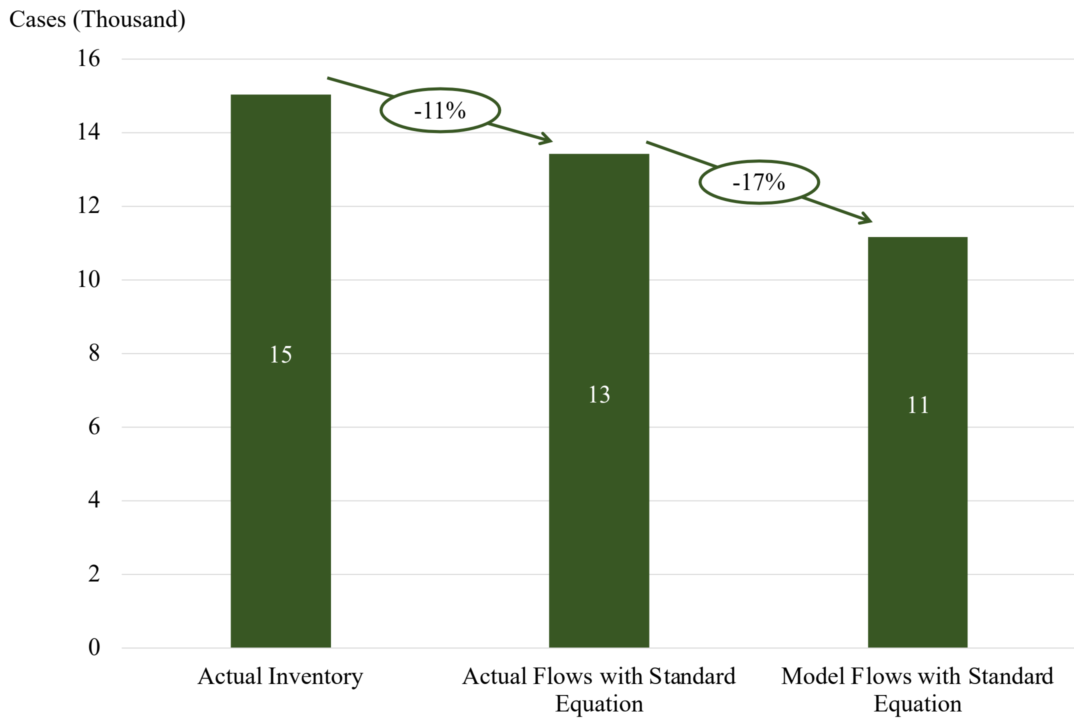
Network Flow Comparison for SKU2



SKU2 may have more inventory than needed to maintain the target service level, as indicated by Figure 14 below. Using the inventory equation in Section 3.4.2 with actual flows suggests a reduction of 11% is possible from actual inventory levels, and a further reduction of 17% is suggested by the integrated model with new flows.

Figure 14

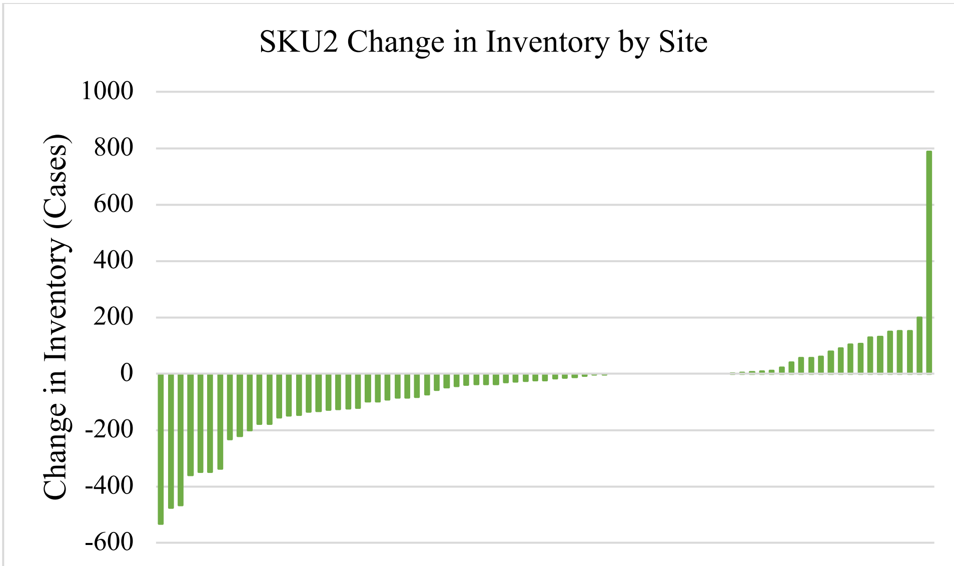
Inventory Comparison for SKU2



Looking at specific sites, the model suggests one site in Maryland to increase inventory by ~800 cases which is an outlier. This particular distribution site has a favorable inbound cost compared to other sites, and the model takes advantage of this by consolidating inbound flows at this location. A substantial portion of this inventory increase is due to cycle stock from the higher flows. Most sites, however, reduce inventory in the model. The most substantial decreases are in West Coast sites with high inbound cost from the East Coast suppliers for this item. Figure 15 illustrates these changes in inventory levels.

Figure 15

Inventory Comparison by Site for SKU2



4.3 SKU3 Model Results

Finally, SKU3 is a frozen item with very high volume and very low variability, with close to 80% of orders from customers needing next day delivery. This item has a high inventory level. As can be seen in Table 5, the majority of historical transportation cost for SKU3 comes from outbound transportation cost.

Table 5

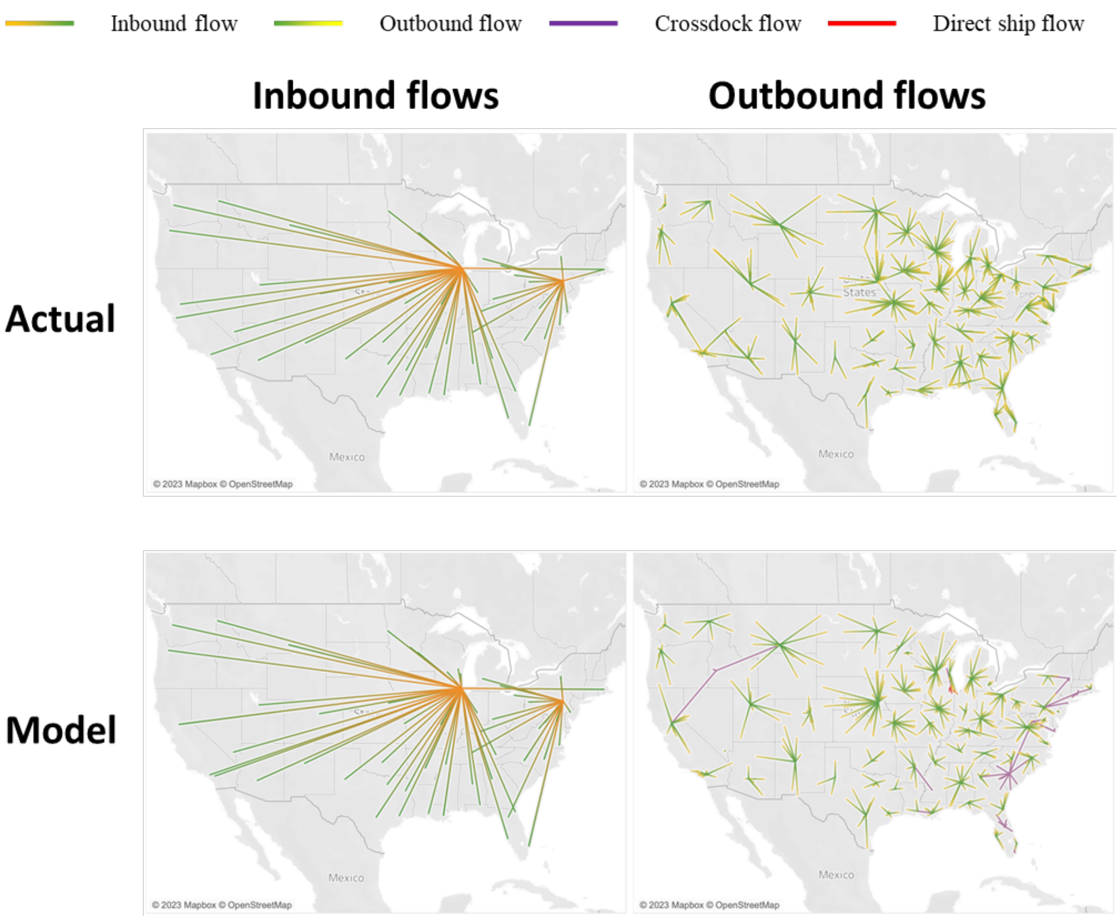
Cost Result Summary for SKU3

Cost Summary (Thousand USD)	AsIs ($)	Model ($)	Difference to AsIs ($)	Difference (%)
Transportation Cost	15,467	14,380	(1,088)	-7.0
Inbound Transportation Cost	5,871	5,635	(235)	-4.0
Cross-dock Cost	1	379	368	N/A
Direct Ship Cost	-	808	808	N/A
Outbound Transportation Cost	9,595	7,566	(2,029)	-21.1
Warehousing Cost	3,691	3,414	(276)	-7.5
Inventory Cost	234	118	(116)	-49.4
Cycle Stock Cost	61	66	5	8.6
Safety Stock Cost	173	52	(121)	-69.9
Subtotal Logistics Cost	19,392	17,912	(1,480)	-7.6
Product Cost	56,771	56,771	(0)	0.0
Total Supply Chain Cost	76,163	76,683	(1,480)	-1.9

This item had the largest reduction in both total supply chain cost and total logistics cost achieved among all three products, mainly due to the significant decrease in outbound transportation cost, safety stock cost, and the use of new omnichannel fulfillment options. Although there was a slight increase in cycle stock cost, all other types of costs were reduced, similar to SKU1. The success in reducing outbound transportation cost through the use of omnichannel fulfillment is notable, as the outbound cost was a significant burden for this particular item. This item has a large concentration of customer demand near the supply point in Chicago, which the model switches to direct ship to minimize cost. This item also uses both cross-docking lanes and peer replenishment in the model. Figure 16 shows a national level visualization of the flows including historical actual flows and the new model flows.

Figure 16

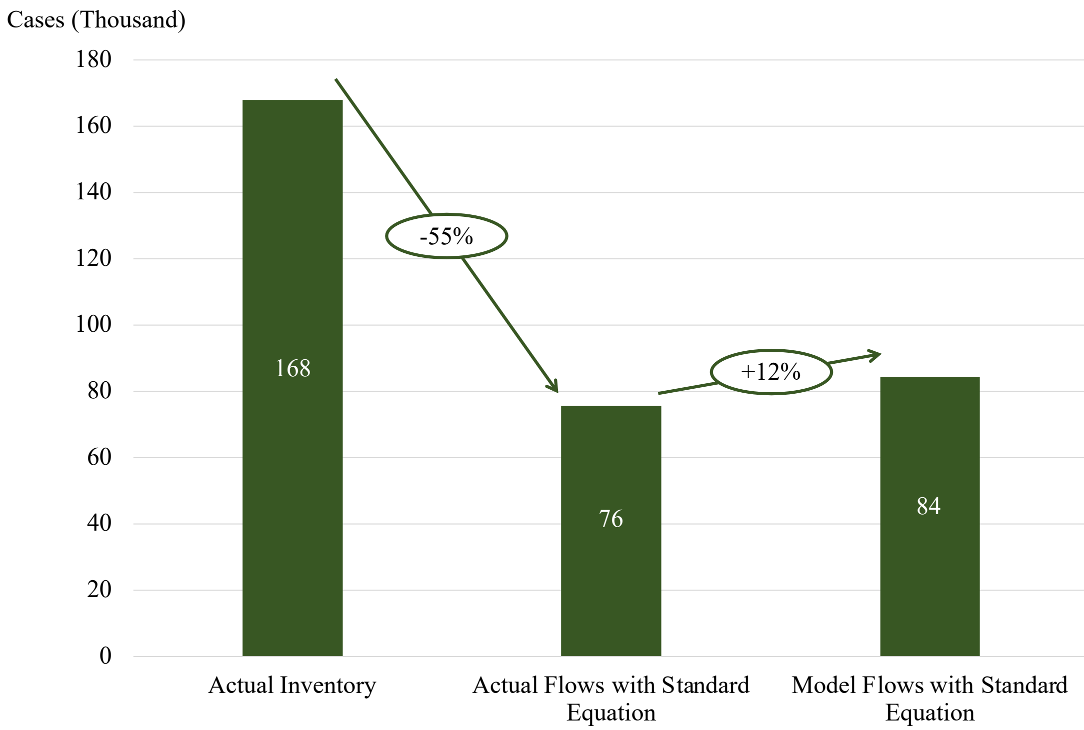
Network Flow Comparison for SKU3



This item has the largest reduction in inventory of the items examined. The sponsor company carries approximately 14 days of inventory for this item as measured across the entire network, however the coefficient of variation is 0.07, indicating very stable demand. Using the inventory equation in Section 3.4.2. and actual historical flows, inventory targets come in 55% lower than actual to maintain the target service level. The integrated model, with new flows, is 12% higher than this amount. Figure 17 compares these inventory levels.

Figure 17

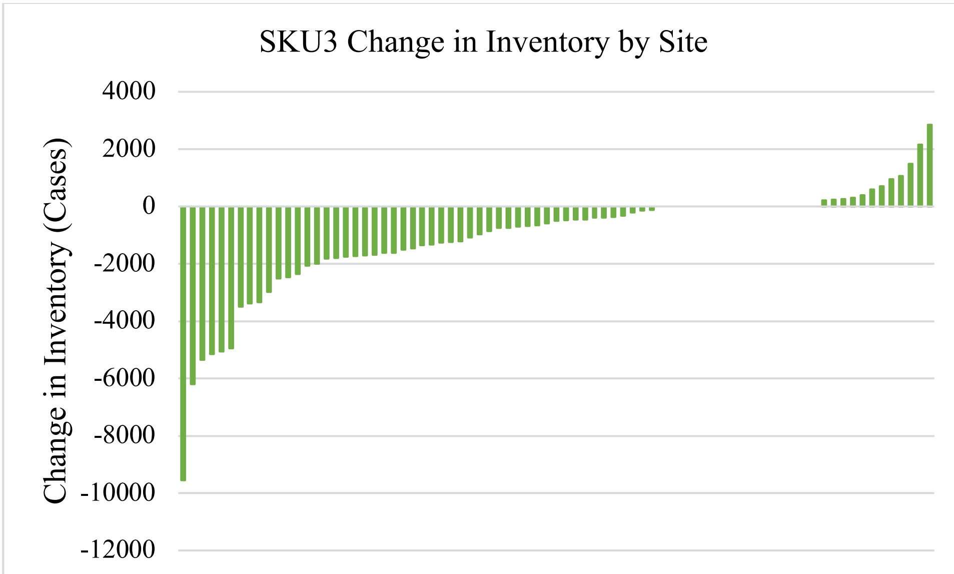
Inventory Comparison for SKU3



Though the model reduces inventory overall, four sites add over 1,000 cases of inventory due to changes in network flows. These distribution sites are all relatively low cost in the network. Fives sites reduce inventory levels by over 5,000 cases, due to the combination of new flows, direct ship taking inventory out of the network, and new cross-docking lanes. Most of these sites are near the Chicago area. Figure 18 shows a national level visualization of the flows including historical actual flows and the new model flows.

Figure 18

Inventory Comparison by Site for SKU3



4.4 Omnichannel Fulfillment Options Insights

This section provides a detailed analysis of the model's treatment of omnichannel options, including both numerical and business perspectives, following the more general overview of cost structure and inventory placements in the previous chapters (4.1 through 4.3).

4.4.1 Direct Ship

Regarding SKU1, the model utilized only one direct ship lane from a supplier to a nearby distribution site, as direct shipping is generally expensive for longer distances. However, the impact of this choice is not insignificant due to the large amount of flow, and it was a cost-effective decision for the model that attempted to shift flows whenever possible. Direct ship was not used for SKU2 due to lower overall demand and because demand is not located near supply points. For SKU3, more direct ship lanes were used compared to SKU1, possibly due to higher outbound costs or cheaper estimated direct ship costs. As well as SKU1, direct ship was used for customers closer to the suppliers, but for SKU3, another reason for the choice of direct shipping may have been the relatively high warehouse and transportation costs of the sites around the Chicago area.

In summary, the model's results suggest that direct ship is most advantageous for customers located close to suppliers, or in areas where transportation or warehouse handling costs are higher, such as in larger cities. However, the current method used for estimating direct ship costs may be too simplistic and may not accurately capture the actual future costs with involved in this new arrangement with suppliers and customers. Thus, further improvements to the cost estimation process are needed for more precise analyses.

4.4.2 Cross-docking and Peer Replenishment

Cross-docking / peer replenishment was implemented as another omnichannel option, with varying levels of usage across the different products. Surprisingly, despite the theoretical disadvantage of having a shorter lead-time threshold having to hold inventory than the other SKUs, SKU3 had the highest utilization of cross-docking lanes, while SKU2 had the lowest. One possible explanation is that outbound transportation cost, which is a major driver for cost savings with cross-docking, was relatively lower for SKU2 compared to the other products. Another possible explanation is that demand and supply variability for SKU2 might be more stable, resulting in less motivation to switch from existing outbound transportation to cross-docking, where safety stock can be consolidated at the first distribution sites. This interpretation is plausible given that SKU2 is a common item that is frequently used by customers.

In addition to these implications, upon analyzing the result maps, it becomes evident that the first leg of the cross-docking lanes (from the first distribution sites to the second distribution sites) is considerably longer than the second leg (from the second distribution sites to the customers). This observation suggests that outbound transportation cost is a significant challenge for the entire supply chain network, and it is better to minimize it by utilizing distribution sites that are close to customers and reducing the distance of outbound transportation. In this regard, cross-docking can play a crucial role in reducing the overall outbound cost by reallocating it to internal transportation cost, which is typically less expensive than outbound last-mile cost.

In summary, cross-docking is a beneficial solution when outbound transportation costs significantly impact the entire supply chain network or when pooling inventory has a higher impact due to factors such as volatile demand or variable supply lead times.

4.4.3 Central Warehouses

Drop shipping from the central warehouses was not utilized for any item, due to the higher cost. However, for SKU3, central warehouses were employed as origin points for cross-docking to other sites. The costs of outbound transportation, warehouse handling, and inventory for other distribution sites are crucial factors in determining whether central warehouses were more advantageous. Additionally, in practice, the use of central warehouses may be driven by capacity constraints, which we did not incorporate in this study. For a multi-item model, it is possible that distribution site capacity could be a limiting factor, and central warehouses could be leveraged to alleviate capacity constraints at other distribution sites especially for CZ segmentation items with low demand. These kinds of items would also benefit from a risk pooling effect for inventory.

4.5 Model Solve Time Insights

In this section, we present insights obtained from our model in terms of efficiency, in addition to the solutions themselves shown in the previous sections.

In addition to our focus on the projected cost savings, we aimed to create a model that could quickly and efficiently scale to real world problem sizes. This requires a fast solution time. As mentioned earlier, part of our effort to accomplish this focused on limiting the model variables to as small a scope as possible. In addition, we utilized a tailored outer approximation method to formulate and solve our MINLP problem and conducted a comparison with Gurobi's default solver. To obtain initial feasible solutions of binary decision variables for the outer approximation subproblem, we assumed that every distribution site is served from the first supplier (the smallest supplier number). Our computational experiments were performed on Intel Xeon Gold 6248 or Intel Xeon Platinum 8260 processors in the MIT supercloud environment, where the number of nodes and the CPU were automatically selected. Table 6 presents the execution times of each method for SKU1, based on 5 independent trials.

Table 6

Execution Time for Each Method

Method		Max (sec)	Min (sec)	Average (sec)
Gurobi Default		45.67761397	44.23880315	44.60165119
OA ($\varepsilon = 0.001$)	NLP Subproblem	2.677587986	2.509674072	2.579761171
	MILP Master	0.660363913	0.021645069	0.150467968
	Total	3.337951899	2.531970978	2.730229139
OA ($\varepsilon = 0.0001$)	NLP Subproblem	5.426824808	5.192915201	5.325380659
	MILP Master	1.026623964	0.1076231	0.293673038
	Total	6.348081112	5.305934191	5.619053698

To assess the quality of the solutions, we compared the outcomes generated by our outer approximation (OA) with those produced by the default algorithm of Gurobi. The difference between the solutions obtained by OA and Gurobi was calculated by taking one minus the absolute value of the solution from OA divided by the solution from the default algorithm. Table 7 presents the results of the comparison for SKU1. We are pleased to share we were able to significantly reduce the solve time without any meaningful loss of solution quality.

Table 7

Solution (Total Supply Chain Cost) Difference between Methods

	Solution Difference (%)
OA ($\varepsilon = 0.001$)	3.76e-1
OA ($\varepsilon = 0.0001$)	7.00e-9

Firstly, we were able to achieve a significant reduction in execution time, almost eight times faster than Gurobi's default solve time, by implementing our outer approximation algorithm that is effective for the block structure commonly seen in supply chain problems. It is worth noting that the MILP primary problem had a much lower execution time compared to the NLP subproblem, possibly due to the small feasible sets resulting from the limited number of suppliers and fewer branching times required for MILP. Additionally, the constraints that restrict the number of suppliers each distribution site can choose allowed Gurobi to efficiently solve the integer problem.

Secondly, we observed a significant impact of the termination criteria on run time, with a nearly doubled time when the criterion was tightened. Given the generally larger tolerance for optimality in supply chain network design problems, loosening the criteria can be a practical way to speed up the model, particularly when dealing with more extensive data or larger networks. However, considering the average cost reduction percentage for the entire supply chain, the solution difference when $\varepsilon = 0.001$ can be relatively impactful, this, a stricter criterion such as 0.0001 was preferable in our case, balancing optimality and speed effectively.

Overall, our tailored algorithm successfully reduced the execution time and proved to be highly effective for similar problems. Nevertheless, choosing the appropriate termination criteria, depending on problem characteristics, size, and computational power available, is crucial to achieving an optimal balance between speed and optimality. To further improve the overall performance, development of more sophisticated algorithm might be necessary.

4.6 Expandability and Limitations

This section outlines potential avenues for expanding our existing model into diverse applications and highlights major elements that could not be integrated due to constraints of time and workforce. By illustrating these factors, this paper aims to provide valuable insights for future researchers and projects undertaken by the sponsor company.

4.6.1 Expandability

Our model's expandability is a significant advantage for the sponsor company, as it can be easily extended to a multi-item model by introducing more granular decision variables (e.g., adding another dimension to x_{ikj} to consider item variation). This would enable optimization of the entire supply chain for all SKUs, albeit with some new constraints to be introduced. For instance, distribution site capacity constraints become a concern when changing flows for all items, as high-volume concentration at a single distribution site could potentially violate capacity. However, fortunately, it is possible to directly apply the model to other items even without changing its constraints or decision variables, which could still lead to new discoveries. Additionally, the model can accommodate other omnichannel options, including creating new larger distribution sites or utilizing shuttle yards (placing additional facilities between distribution sites and customer to deal with last-mile deliveries). These adjustments are generally easy, as the consideration of new distribution sites can be achieved through the introduction of new binary decision variables to control open / close conditions for them, as well as the use of cheaper inventory or warehouse handling costs to encourage inventory pooling.

Also, our model's incorporation of binary decision variables for each supplier-distribution site lane allows for the easy integration of fixed costs for selecting specific suppliers. This is particularly relevant when each supplier has a "minimum volume" requirement for the total flow from the supplier, ensuring that suppliers can maintain an economy of scale.

Lastly, our tailored outer approximation algorithm can be applied to omnichannel options without changing the original structure of the model. This approach not only accelerates the solving time for our omnichannel options but also proves to be highly effective when the number of decision variables or constraints increases.

In summary, these flexible features are made possible through the tailored development of the model using Python and Gurobi, which sets it apart from package software that typically requires a fixed structure for all model types.

4.6.2 Limitations

While our model is highly flexible, there are limitations that require improvement in the model structure or algorithm itself in some cases. One limitation arises when expanding the model to a multi-item model, where the number of decision variables and constraints increase significantly due to the structure of decision variables representing all possible lane combinations. This results in a higher computational power requirement or infeasibility, necessitating more efficient algorithms or changing formulations.

Multi-echelon inventory can also be problematic when considering more complicated omnichannel options or changing current assumptions. For instance, assuming safety stock is never placed at the second distribution site for peer-to-peer cross-docking might require reconsideration if demand variability is erratic or lead time for internal shipment is variable. As safety stock is placed for multiple tiers, new factors like each distribution site's coverage or demand propagation from downstream to upstream of the supply chain need to be considered, resulting in further complexity of the model's underlying equations.

Another limitation is the current model's assumption of limiting the number of suppliers each distribution site can accommodate to only one. While this is a realistic assumption for our sponsor's company, which rarely assigns two or more suppliers to a single distribution site for the same product, other industries or companies may have more flexibility. Changing this assumption will result in duplicating safety stock for different suppliers at the same distribution site, overestimating inventory amount or inventory cost. Moreover, as the model is specifically tailored to the sponsor company, other significant changes may be necessary when applying it to other companies' or organizations' cases.

Lastly, although our outer approximation successfully solves the problem faster than Gurobi's default solving method, we may need to seek even more efficient algorithms by combining other decomposition or branching methods or using other solvers for MINLP.

5 CONCLUSION

While omnichannel in B2B is not as widely studied as in B2C, it is an increasingly important opportunity for many companies. Our sponsor company, a large foodservice distribution company in the United States, is interested in expanding omnichannel fulfillment options for its customers. As we have shown in this paper, these options could be advantageous for both customers as well as the companies offering them when paired with a comprehensive review of the supply chain involved. However, this review will be most effective when it is end-to-end (from supplier to customer) and integrated (solved simultaneously across all flows and with inventory rather than sequentially).

In addition to the specific item level discussion contained in Chapter 4, we also offer the following generalized recommendations:

1) Consider applying the model to other items with a similar cost structure to SKU3. Our application of the model to three products demonstrated that network design can significantly reduce the total supply chain cost by optimizing facility use and product flows. This benefit is not limited to the products we tested, and the company can expand its use to more items. It is highly likely that products with a similar cost structure to SKU3 would be impactful, as it produced the most significant reduction in our case.
2) Integrate outbound and inbound management to optimize the entire supply chain network. Our model revealed that the current supply chain shows signs of local optimization for inbound and outbound logistics, which is not necessarily optimal for the entire network. By integrating these two functions and changing product flows, the company can minimize total transportation cost and improve the efficiency of the entire supply chain network based on the results produced in this project.
3) There may be an inventory reduction opportunity for items with low variability. Our model revealed that the current inventory levels may be excessive for these items. However, some customers may be concerned about level of service. Therefore, it is recommended to involve

customers in the implementation scheme to ensure a smooth transition. This collaborative effort can lead to a more optimal inventory level and improved supply chain efficiency.

4) Explore direct ship opportunities for supply points with high volume that are located near dense demand points. This fulfillment method avoids inbound, warehouse, and inventory costs. However, this may be partially offset by the cost of new dedicated routes.
5) Continue to pilot cross-docking and peer replenishment options. Our model revealed that for certain items these options can be competitive in terms of the total supply chain cost. However, this competitiveness is dependent on taking advantage of substantial differences in inbound costs between sites or other structural cost differences. Additionally, the fixed costs of setting up the initial cross-dock shuttle network lend itself to a “big bang” deployment to avoid large per unit costs related to shuttles.

B2B customers are now demanding more fulfillment options, similar to what they see in the B2C space. With continued research and experimentation in omnichannel fulfillment, our sponsor company can lead the industry with new ways to delight customers while minimizing cost. The model proposed can help our sponsor to evaluate trade-offs when considering these new options in a single, integrated methodology.

REFERENCES

Ağralı, S., Geunes, J., & Taşkın, Z. C. (2012). A facility location model with safety stock costs: analysis of the cost of single-sourcing requirements. *Journal of Global Optimization, 54*, 551–581. https://doi.org/10.1007/s10898-011-9777-z

Alicke, K., & Forsting, M. (2017). Supply chain segmentation scientific frameworks. Supply Chain Segmentation: Best-in-Class Cases, Practical Insights and Foundations, 5-13.

Aloufi, Y. (2021). *Outer-approximation (OA)*. Cornell University Computational Optimization Open Textbook. https://optimization.cbe.cornell.edu/index.php?title=Outer-approximation_(OA)#:~:text=The%20Outer%2DApproximation%20(OA),integer%20linear%20master%20problem%20successfully

Ansaria, S., Basdere, M., Lib, X., Ouyangc, Y., & Smilowitza, K. (2018). Advancements in Continuous Approximation Models for Logistics and Transportation Systems: 1996 – 2016. *Transportation Research Part B: Methodological, 107*, 229-252. https://doi.org/10.1016/j.trb.2017.09.019

Aouad, W. & Ganapathi, N. (2020). *An Omnichannel Distribution Model to Better Serve Online Customers*. [Master Thesis, Massachusetts Institute of Technology]. Massachusetts Institute of Technology. Department of Supply Chain Management. https://ctl.mit.edu/sites/ctl.mit.edu/files/theses/scm2020-aouad-an-omnichannel-distribution-model-to-better-serve-online-customers-capstone.pdf

Atallah, M. J. (Ed.) (1999). NP-hard. In *Algorithms and Theory of Computation Handbook*. CRC Press LLC.

Bell, D. R., Gallino, S., & Moreno, A. (2014). How to win in an omnichannel world. MIT Sloan Management Review.

Chandra, V., Gill, P., Kohli, S., Venkataraman, K., & Yoshimura, J. (2022, April 29). *The Next Horizon for Grocery e-commerce: Beyond the Pandemic Bump*. McKinsey & Company. https://www.mckinsey.com/industries/retail/our-insights/the-next-horizon-for-grocery-ecommerce-beyond-the-pandemic-bump

Chopra, S. (2018). The Evolution of Omni-Channel Retailing and its Impact on Supply Chains. *Transportation Research Procedia, 30*, 4-13. https://doi.org/10.1016/j.trpro.2018.09.002

Christopher, M., Peck, H., & Towill, D. (2006). A taxonomy for selecting global supply chain strategies. The International Journal of Logistics Management, 17(2), 277–287. https://doi.org/10.1108/09574090610689998

Coupa. (2021, December 13). *Inventory Optimization*. https://help.llama.ai/release/native/modeling/modeling-topics/Using_SafetyStockOptimization.htm#

Croxton, K. L., & Zinn, W. (2005). Inventory Considerations in Network Design. *Journal of Business Logistics, 26(1)*, 149-168. https://doi.org/10.1002/j.2158-1592.2005.tb00197.x

Darmawan, A., Wong, H., & Thorstenson, A. (2021). Supply chain network design with coordinated inventory control. *Transportation Research Part E: Logistics and Transportation Review, 145*, 102-168. https://doi.org/10.1016/j.tre.2020.102168

Deepak, A. & Naman, D. (2009). *Inventory positioning for a multi-echelon distribution network* [Master Thesis, Massachusetts Institute of Technology]. Massachusetts Institute of Technology. Engineering Systems Division. https://dspace.mit.edu/handle/1721.1/55338

Diabat, A., Jabbarzadeh, A., & Khosrojerdi A. (2019). A Perishable Product Supply Chain Network Design Problem with Reliability and Disruption Considerations. *International Journal of Production Economics, 212,* 125-138. https://doi.org/10.1016/j.ijpe.2018.09.018

Disney, S. M., McCullen, P., & Saw, R. J. (2006, July). *Dynamic Supply Chain Design: Square Root Law for Bullwhip.* 11th International Symposium on Logistics (11th ISL), Beijing, China. https://orca.cardiff.ac.uk/id/eprint/42586/1/42586%20Disney%202006%20Dynamic.pdf

Duran, M. A., & Grossmann, I. E. (1986). An outer approximation algorithm for a class of mixed-integer nonlinear programs. *Mathematical Programming, 36*, 307–339. https://doi.org/10.1007/BF02592064

Forbes. (2023). *Gordon Food Service*. https://www.forbes.com/companies/gordon-food-service/?sh=3b9855a66011

Gerea, C., Gonzalez-Lopez, F., & Herskovic, V. (2021). Omnichannel customer experience and management: An integrative review and research agenda. Sustainability, 13(5), 2824.

Govindarajan, A., Sinha, A., & Uichanco, J. (2017). *Inventory Optimization for Fulfillment Integration in Omnichannel Retailing* [Working Paper, University of Michigan]. University of Michigan. Ross School of Business. https://deepblue.lib.umich.edu/bitstream/handle/2027.42/136157/1341_Govindarajan.pdf?sequence=

Graves, S. C., & Willems, S. P. (2000). Optimizing strategic safety stock placement in supply chains. Manufacturing & Service Operations Management, 2(1), 68-83.

Guerrero-Lorente, J., Gabor, A. F., & Ponce-Cueto, E. (2020). Omnichannel logistics network design with integrated customer preference for deliveries and returns. Computers & Industrial Engineering, 144, 106433.

Harrison, L., Plotkin, C., Reis, S., & Stanley, J. (2021, December 15). *B2B sales: Omnichannel everywhere, every time*. McKinsey & Company. https://www.mckinsey.com/capabilities/growth-marketing-and-sales/our-insights/b2b-sales-omnichannel-everywhere-every-time

Hübner, A., Wollenburg, J., & Holzapfel, A. (2016). Retail logistics in the transition from multi-channel to omni-channel. International Journal of Physical Distribution & Logistics Management, 46(6/7), 562-583.

Ingalls, R. G. (2014). Introduction to Supply Chain Simulation. *Proceedings of the Winter Simulation Conference 2014*, 36-50. https://doi.org/10.1109/WSC.2014.7019876

International Foodservice Distributors Association. (2018). *We deliver: A Comprehensive Economic Impact Study of the U.S. Foodservice Distribution Industry*. https://www.ifdaonline.org/news-insights/research-insights/reports/foodservice-distribution-industry-economic-impact/ifda-economic-impact-study-final.aspx

International Foodservice Distributors Association. (2022). *The State of Foodservice Distribution* https://www.ifdaonline.org/news-insights/research-insights/reports/ifda-state-of-foodservice-distribution-2022

Janjevic, M., Merchán, D., & Winkenbach, M. (2021). Designing multi-tier, multi-service-level, and multi-modal last-mile distribution networks for omni-channel operations. *European Journal of Operational Research, 294 (3)*, 1059-1077. https://doi.org/10.1016/j.ejor.2020.08.043

Jocevski, M., Arvidsson, N., Miragliotta, G., Ghezzi, A., & Mangiaracina, R. (2019). Transitions towards omni-channel retailing strategies: a business model perspective. International Journal of Retail & Distribution Management.

Lee, T. (n.d.). *Optimization vs. Simulation*. Kisters North America. https://www.kisters.net/NA/fileadmin/KNA/Products/Optimization_vs_Simulation.pdf

Lee, H. L. (2002). Aligning supply chain strategies with product uncertainties. California management review, 44(3), 105-119.

Lim, S. F., Wang, L., & Srai, J. S. (2017, October). Wal-Mart's Omni-channel Synergy. *Supply Chain Management Review, September/October 2017*, 30-37. https://www.researchgate.net/profile/Stanley-Frederick-Wt-Lim-2/publication/326302770_Wal-Mart's_Omni-channel_Synergy/links/5c9a86c0a6fdccd4603cd08d/Wal-Marts-Omni-channel-Synergy.pdf

Lucas, P. (2019). *Sysco's Innovative E-Commerce Platform, Supplies on the Fly, Partners with Frieda's Specialty Produce*. Digital Commerce 360. https://www.digitalcommerce360.com/2019/08/23/us-foods-burnishes-its-online-strategy-with-us-foods-direct/amp/

Markowicz, F. (2017). *Optimizing Order-Routing Decisions: Leveraging Omni-Channel. Supply Chain Fulfillment*. [Master Thesis, Massachusetts Institute of Technology]. Massachusetts Institute of Technology. Department of Mechanical Engineering; Sloan School of Management. http://dspace.mit.edu/bitstream/handle/1721.1/111492/1003322576-MIT.pdf?sequence=1&isAllowed=y

Matheu, S. (2022, October 10). *US Foods Launches the Most Advanced E-commerce Application in the Food Service Distribution Industry*. Business Wire. https://www.businesswire.com/news/home/20221008005009/en/US-Foods-Launches-the-Most-Advanced-E-commerce-Application-in-the-Food-Service-Distribution-Industry

Meissner, J., & Senicheva, O. V. (2018). Approximate dynamic programming for lateral transshipment problems in multi-location inventory systems. European Journal of Operational Research, 265(1), 49-64.

Nozick, L. K., & Turnquist, M. A. (1998). Integrating Inventory Impacts into a Fixed-Charge Model for Locating Distribution Centers. *Transportation Research Part E: Logistics and Transportation Review, 34(3),* 173-186. https://doi.org/10.1016/S1366-5545(98)00010-6

Ozbun, T. (2022, February 22). *U.S. Food Retail Industry - Statistics & Facts.* Statista. https://www.statista.com/topics/1660/food-retail/#topicHeader__wrapper

Paterson, C., Kiesmüller, G., Teunter, R., & Glazebrook, K. (2011). Inventory models with lateral transshipments: A review. European Journal of Operational Research, 210(2), 125-136.

Paul, J., Agatz, N., & Savelsbergh, M. (2019). Optimizing omni-channel fulfillment with store transfers. *Transportation Research Part B: Methodological, 129*, 381-396. https://doi.org/10.1016/j.trb.2019.10.002

Performance Food Group (2022). *2022 Annual Report*. https://s22.q4cdn.com/140600076/files/doc_financials/annual/2022/Performance-Food-Group_AR_2022.pdf

Performance Foodservice. (2023). *Order Online*. https://www.performancefoodservice.com/Your-Operation/Manage-Your-Order/Order-Online

Pereira, L. X., Loures, E. D. F. R., & Santos, E. A. P. (2022). Assessment of supply chain segmentation from an interoperability perspective. International Journal of Logistics Research and Applications, 25(1), 77-100.

Ponce, E. (2017, June 8). *Omni-channel Distribution Strategies*. Massachusetts Institute of Technology. MIT Center for Transportation & Logistics. https://ctl.mit.edu/sites/ctl.mit.edu/files/attachments/tab%208%20Omni-channel.pdf

Prak, D., Teunter, R., & Syntetos, A. (2017). On the calculation of safety stocks when demand is forecasted. European Journal of Operational Research, 256(2), 454-461.

Puga, M. S., Minner, S., & Tancrez, J. S. (2019). Two-stage supply chain design with safety stock placement decisions. International Journal of Production Economics, 209, 183-193.

Stord Marketing. (2020, August 20). *Dark Warehouses: Are They in Your Supply Chain's Future?* https://www.stord.com/blog/dark-warehouses-are-they-in-your-supply-chains-future

Strickler, G. R. (2021, December 23). *Understanding Solve Time in MIP Supply Chain Models Using Community Data.* Coupa. https://www.coupa.com/blog/technology-innovation/understanding-solve-time-mip-supply-chain-models-using-community-data

Mangotra, D., Lu, J., & Tsao, Y. (2009). *A Continuous Approximation approach for the Integrated Facility Location-Inventory Allocation Problem* (ISyE Technical Report J-09-01). School of Industrial and Systems Engineering. https://smartech.gatech.edu/bitstream/handle/1853/31393/Mangotra-J-09-01.pdf?sequence=3&isAllowed=y

Matinrad, N., Roghanian, E., & Rasi, Z. (2013). Supply chain network optimization: A review of classification, models, solution techniques and future research. *Uncertain Supply Chain Management, 1*, 1-24. https://doi.org/10.5267/j.uscm.2013.05.003

Silver, E. A., Pyke, D. F., & Peterson, R. (1998). Inventory management and production planning and scheduling (Vol. 3, p. 30). New York: Wiley.

Stojanović, M., & Regodić, D. (2017). The significance of the integrated multicriteria ABC-XYZ method for the inventory management process. Acta Polytechnica Hungarica, 14(5), 29-48.

Sysco Corporation. (2022). *Annual Report 2022*. https://investors.sysco.com/~/media/Files/S/Sysco-IR/documents/annual-reports/Sysco_2022-Annual-Report_Web.pdf

Sysco Corporation. (2018, November 27). *Sysco's Innovative E-Commerce Platform, Supplies on the Fly, Partners with Frieda's Specialty Produce*. https://investors.sysco.com/annual-reports-and-sec-filings/news-releases/2018/11-27-2018-220018840

Sysco Corporation. (2023). *Supplies Your Way*. https://sysco.com/Products/Solutions-and-Services/Culinary-and-Product-Solutions/Supplies-on-the-Fly.html

Taylor, D., Brockhaus, S., Knemeyer, A. M., & Murphy, P. (2019). Omnichannel fulfillment strategies: defining the concept and building an agenda for future inquiry. *The International Journal of Logistics Management, 30(3),* 863-891. https://doi.org/10.1108/IJLM-09-2018-0223

Technomic. (2022, September). *U.S. Foodservice Industry*.

The Business Research Company. (2021, September). *Food and Beverages E-Commerce Global Market Report 2021: COVID-19 Growth and Change to 2030*. Report Linker. https://www.reportlinker.com/p06151609/Food-And-Beverages-E-Commerce-Global-Market-Report-COVID-19-Growth-And-Change-To.html?utm_source=GNW

US Foods Holding Corp. (2022). *Annual Report*. https://d18rn0p25nwr6d.cloudfront.net/CIK-0001665918/8e21dff5-394e-4919-9d8c-080a5a37ea48.pdf

You, F., & Grossmann, I. E. (2008). Mixed-Integer Nonlinear Programming Models and Algorithms for Large-Scale Supply Chain Design with Stochastic Inventory Management. *Industrial & Engineering Chemistry Research, 47(20)*, 7802-7817. https://doi.org/10.1021/ie800257x

Wallace, L. (2022) *How B2B Sales Teams Win with Digital Commerce.* https://www.salesforce.com/resources/articles/b2b-sales-wins-with-digital-commerce/

APPENDIX

Proof 1: Convexity of the Objective Function

Let $f(x) = \sqrt{\sum_{j\in C} b_{ikj}x_{ikj}^2} = \sqrt{\sum_{j\in C}(\sqrt{b_{ikj}}x_{ikj})^2}$, considering $b_{ikj} \geq 0$.

Also, let $g(x) = [\sqrt{b_{111}}x_{111}, \dots, \sqrt{b_{ikj}}x_{ikj}]$.

$f(x) = \|g(x)\|$ (the l_2 norm of $g(x)$)

Then,

$f(\lambda u + (1-\lambda)v) = \|g(\lambda u + (1-\lambda)v)\|$

$= \|\lambda g(u) + (1-\lambda)g(v)\| \leq \|\lambda g(u)\| + \|(1-\lambda)g(v)\| = \lambda f(u) + (1-\lambda)f(v)$

Therefore, $f(x)$ is convex. Considering $a_{ikj}x_{ikj}$ is a linear function and that summation of convex functions is convex, the entire function is convex. Note that *h*, *P*, and *z* are all non-negative constants, which keeps convexity of the function. Refer to Ağralı et al. (2012) for a similar proof.

Proof 2: Subgradient of the Function

For the point of $\sum_{j\in C} x_{ikj} = 0$, letting $c = 0$, $f(0) + c(x-0) = f(0) = f(x)$.

For the other points $\sum_{j\in C} x_{ikj} = \varepsilon > 0$, $f(0) \leq f(x)$, considering b_{ikj} and H are positive constants.

Therefore, $f(x) \geq f(a) + c(x-a)$ for $a = 0, c = 0$.