Perishable Inventory Routing Problem under Uncertainty

by

Ghazaleh Khalili

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Management Sciences

Waterloo, Ontario, Canada, 2023

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In an Inventory Routing Problem (IRP), a decision-maker decides the number of units delivered to each retailer and determines delivery routes, which becomes increasingly challenging as the network expands. Incorporating uncertainty and perishability into the IRP gives rise to a more complex problem known as the stochastic Perishable Inventory Routing Problem (PIRP). Traditional approaches, such as dynamic programming, often struggle to efficiently solve this problem. This is due to the curse of dimensionality, which grows exponentially with the number of retailers and the product's shelf life. In this work, we decompose the PIRP into a Perishable Inventory Problem (PIP) and a Vehicle Routing Problem (VRP) and address them sequentially in two distinct phases. By successfully determining the replenishment quantities first, we then solve the VRP using state-of-the-art algorithms. Consequently, our primary focus lies in identifying the optimal replenishment quantities for perishable products. To address the complexities of this problem, we propose a Direct Lookahead Approximation (DLA) policy designed for sequential decision-making problems under uncertainty. Specifically, we employ a two-stage approximation method that considers a limited number of sample paths while still achieving promising results. The problem is formulated as a mixed-integer programming (MIP) model with the objective of minimizing holding, shortage, wastage, and replenishment costs. In this context, a fixed cost is employed as an approximation for the routing costs of the second phase. To enhance the implementation of the DLA policy, we conduct a comprehensive analysis and recommend techniques such as incorporating linear cuts into the MIP model. To evaluate the effectiveness of the policy, we examine a blood supply chain focusing on perishable platelet units. Through extensive experiments, we demonstrate that the proposed policy can significantly outperform several known algorithms in the literature.

Acknowledgements

Pursuing higher education is undoubtedly challenging, but as an international graduate student, the difficulties multiply. It is not only academic pursuits but also navigating an independent life: finding housing, adapting to a new cultural environment that you've longed to call home, learning to cook while sometimes simply skipping meals, managing expenses, maintaining a social life, and, most importantly, being separated from family and loved ones—all these burdens are shouldered alone. On top of these, I had personal goals like staying healthy and disciplined. Balancing all these tasks alongside my academic work often felt overwhelming. However, the trust in myself and the support that my supervisors, Dr. Saeed Ghadimi and Dr. Hossein Abouee-Mehrizi, placed in me kept me going.

During moments of self-doubt, Saeed's appreciation of my hard work, generously dedicating his time to my numerous questions, and offering patient explanations provided much-needed reassurance. His feedback and support have been invaluable, not only in shaping this thesis but also in fostering my intellectual growth.

Besides, it was Hossein who consistently acknowledged my efforts and motivated me to continue striving for excellence. Knowing that my efforts were recognized by him brought me a renewed sense of motivation. His constant professional guidance has made this journey possible for me.

Therefore, I am truly grateful and wish to extend my deepest appreciation and heartfelt gratitude to them for their invaluable assistance and mentorship in every aspect. Additionally, I am thankful to Dr. Jim Bookbinder and Dr. Fatma Gzara for their generous time and patience in reviewing my thesis.

Dedication

I dedicate this thesis to my beloved family, whose unwavering faith in me has been a source of strength. Since I was a child, they taught me the value of ambition and instilled in me the belief that I would become a successful woman graduating from a prestigious university. My mother, whom I am incredibly proud of, has been my role model for commitment, while my father showed me the power of perseverance. They made sacrifices to ensure that my twin sister, Ghazal, and I could pursue our goals. I am immensely grateful to them, and although words cannot express my gratitude enough, I hope to bring them more happiness through further accomplishments.

This dedication is also extended to Ghazal and Mohammad Reza for their tireless support, both emotionally and academically, throughout this journey. Their presence and encouragement have made a profound difference.

Lastly, I dedicate this work to my dearest Mohi, as I cannot imagine how I could have navigated a single day of this remarkable journey without him by my side. His presence has been a constant source of comfort and solace to me.

This thesis is a testament to the love, support, and belief they have all shown me.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ADD Average Daily Demand

ADP Approximate Dynamic Programming

BSC Blood Supply Chain

DE Decomposition Algorithm

DI Decomposition-Integration

DLA Direct Lookahead Approximation

DP Dynamic Programming

EV Expected Value

FI Full Information

FIFO First-In, First-Out

IRP Inventory Routing Problem

LIFO Last-In, First-Out

MIP Mixed Integer Programming

PIP Perishable Inventory Problem

PIRP Perishable Inventory Routing Problem

SD Standard Deviation

SOS Special Ordered Sets

UL Up-to-Level

VMI Vendor-Managed Inventory

VRP Vehicle Routing Problem

List of Symbols

$[N]$ Set of natural numbers, where $[N] = \{1, 2, \cdots, N\}$ for any $N \in \mathbb{N}$

$[n, N]$ Closed interval of natural numbers, where $[n, N] = \{n, n+1, \cdots, N\}$ for any $n \in \mathbb{N}$

$\mathcal{N}$ Set of edges, where $\mathcal{E} = \{(i, j) : i, j \in \mathcal{N},\ i \neq j\}$ in a complete graph

$\mathcal{N}^0$ Set of retailers, where $\mathcal{N}^0 = \mathcal{N} \setminus \{0\}$

$\mathcal{E}$ Set of edges, where $\mathcal{N} = \{0, 1, \cdots, N\}$ in a complete graph

Π Set of policies

π Policy index

t Time index

l Remaining shelf-life index

i Customer index

T Last day of the planning horizon

L Shelf-life (lifespan) of the product

M Big number

Q vehicle capacity

$R^{max} = (R_i^{max})_{i \in \mathcal{N}^0}$ Total inventory capacity of retailer i

h Holding cost per unit of product

p Penalty cost per unit of product

e Wastage cost per unit of product

a Replenishment cost per unit of product

$\hat{g}$ Fixed cost

c_{ij} The travelling cost through the route $(i, j) \in \mathcal{E}, i \neq j$

$d_t = (d_{t,i})_{i \in \mathcal{N}^0}$ Vector of realized demand for retailer i at time $t \in [T]$

$D_t = (D_{t,i})_{i \in \mathcal{N}^0}$ Vector of random demand for retailer i at time $t \in [T]$

$\Omega_t = (\Omega_{t,i})_{i \in \mathcal{N}^0}$ Vector of sample space for retailer i at time $t \in [T]$

$\omega_t = (\omega_{t,i})_{i \in \mathcal{N}^0}$ Vector of sample path (scenario) for retailer i at time $t \in [T]$

$P(\omega_t)$ Probability of the sample path ω_t

W_t Random variable at time $t \in [T]$

S_t State variable at time $t \in [T]$

$R_t^l = (R_{t,i}^l)_{i \in \mathcal{N}^0}$ Vector of inventory levels at retailer i with a remaining shelf-life of $l \in [L]$ periods at time $t \in [T]$

$x_t = (x_{t,i})_{i \in \mathcal{N}^0}$ Vector of replenishment quantities for retailer i at time $t \in [T]$

$y_t = (y_{ij,t})_{(i,j) \in \mathcal{E}, i \neq j}$ Binary variable equals 1 if a vehicle takes the route (i, j) at time $t \in [T]$, and 0 otherwise

$q_{t,i}$ The remaining vehicle's capacity after visiting retailer $i \in \mathcal{N}^0$ at time $t \in [T]$

z_t Number of vehicles at time $t \in [T]$

u_t An auxiliary variable at time $t \in [T]$

v_t An auxiliary variable at time $t \in [T]$

o_t An auxiliary variable at time $t \in [T]$

b_t^1 A binary variable at time $t \in [T]$

$b_t^{2,l}$ A binary variable at time $t \in [T]$

$b_t^{3,l}$ A binary variable at time $t \in [T]$

x^+ Non-negative part, where $x^+ = \max\{x, 0\}$ for any value of the variable x

x^- Non-positive part, where $x^- = \min\{x, 0\}$ for any value of the variable x

Chapter 1

Introduction

A distribution network, comprising a supplier and several retailers, is referred to as a Retailer-Managed Inventory system when retailers are responsible for monitoring and replenishing their inventory. In this system, the routing decisions, which involve determining the delivery routes, are handled by the supplier (Archetti & Speranza, 2016). Conversely, if the supplier collaborates with the retailers, the system is referred to as Vendor-Managed Inventory (VMI), wherein retailers must promptly share relevant information with the centralized decision-maker to facilitate inventory decisions (Crama et al., 2018). VMI is often characterized as a mutually beneficial arrangement where both suppliers and retailers can achieve cost savings: Suppliers can reduce distribution and production expenses by aligning customer demands with shipments, while retailers can avoid allocating resources to control and manage their inventory (Coelho et al., 2012). The decision-maker also determines the most cost-effective routes for shipping the products to the retailers, further optimizing the system's overall efficiency.

Implementing VMI for a system with a set of retailers involves solving a challenging combinatorial problem known as the Inventory Routing Problem (IRP). This problem integrates *inventory* and *routing* decisions to minimize the total cost associated with a group of retailers (Bertazzi et al., 2015). In many real-world inventory systems, item spoilage and *perishability* are common phenomena. Ignoring these aspects can adversely impact the inventory decision-making process. Certain goods, such as dairy products, blood units, medicines, and chemical reagents with expiration dates, are vulnerable to perishability. This means that these goods need to be discarded once they exceed their limited shelf life. As such, it is vital to find an efficient inventory policy that can reduce costs for suppliers, retailers, and customers while ensuring the optimal usage of perishable products. This new factor changes the IRP to a more challenging problem called Perishable Inventory Routing Problem (PIRP).

The PIRP faces significant challenges when applied in the context of a Blood Supply Chain (BSC). Specifically, perishability is a critical factor that greatly impacts the design and management of the BSC. The short shelf life of blood products adds complexity to the supply chain processes. For instance, platelets, a crucial blood component, have a limited shelf life of up to 7 days (Service, 2023). Platelet transfusion is used to treat various medical illnesses, including cancer, organ transplantation, hematopoietic stem-cell transplantation, bone marrow failure, hepatitis, AIDS, cardiovascular surgery, and traumatology (Stroncek & Rebulla, 2007). In North America, hospitals receive the platelets they require from a centralized supplier. Examples of such suppliers include the Red Cross in the United States and the Canadian Blood Services in Canada (Abouee-Mehrizi et al., 2022). These suppliers are responsible for distributing platelet units to multiple hospitals within their coverage area. Consequently, they face the challenge of addressing perishability in managing their inventory and determining the optimal routes in their distribution networks.

To achieve efficiency, it is necessary to align the inventory policy with the demand across the supply chain. However, since demand is usually unknown in supply chains, decision-making occurs under *uncertainty*. This uncertainty is a new challenge to the problem and makes it to be a stochastic PIRP, which is the focus of our work.

To overcome these challenges, we employ a *two-phase* approach. The first phase focuses on solving the Perishable Inventory Problem (PIP), where the objective is to find optimal inventory decisions for perishable products. Subsequently, in the second phase, the Vehicle Routing Problem (VRP) needs to be addressed. The VRP is concerned with determining the most cost-effective routes for a fleet of vehicles that start and end their travels at the supplier's site and serve a set of retailers. Each retailer is visited only once, and the assignment of retailers to vehicles must adhere to the capacity constraints of the vehicles (Solomon, 1987). While this two-phase decision-making process may not lead to making optimal decisions, it provides a practical approach for effectively managing such a complex system (Crama et al., 2018).

In PIPs, there are four primary costs that significantly impact the decision-making process: *wastage*, *penalty* (shortage), *holding*, and *replenishment* (ordering) costs. Insufficient products in inventory can result in shortages and loss of demand. On the other hand, excessive products can result in waste. Besides, hidden costs may arise even when there is no over- or under-stocking. For example, failure to order products at the right time can alter their quality. Hence, determining the best inventory policy for perishable products is essential.

In addition to managing these costs in PIPs, it is essential to manage the wastage by implementing an appropriate issuing policy. In this work, we adopt the First-In, First-Out (FIFO) policy under which the oldest units in inventory are used or sold first, reducing the risk of spoilage. However, when freshness is a priority, the Last-In, First-Out (LIFO) policy may be preferred, as it first dispatches the most recently received products. It is worth noting that when the shelf life of products received by retailers is not fixed, the FIFO policy becomes the Oldest-Unit, First-Out policy. In this case, the oldest units in inventory are not necessarily the earliest received. Conversely, the Newest-Unit, First-Out policy prioritizes using the freshest products before the more aged ones.

Additionally, we have to handle the demand uncertainty in our PIP. In inventory management, unfulfilled demand can be handled in two ways: backlogged or lost. In a system with backlogs, unfulfilled demands are saved and fulfilled in the next periods, which can be thought of as a customer waiting for the item rather than seeking an alternative. On the other hand, lost sales systems lose the demand if they cannot be satisfied immediately. In this work, we focus on a lost-sale system.

Overall, the focus of this study is to design an inventory routing policy for a VMI system where a single perishable product should be delivered from a supplier to a set of retailers. The demand for the product is initially unknown and is realized over time. We develop a two-phase approach to solve the problem. The key contributions of this work can be summarized as follow:

- We propose a novel policy to solve the PIRP. This policy employs a two-stage approximation approach, allowing us to achieve cost-reduction solutions while utilizing a minimal number of scenarios. This represents a significant departure from existing literature, where a large number of scenarios is typically required. It is also worth mentioning that our policy involves solving a large-scale Mixed Integer Programming (MIP) model. To enhance the solution process, we provide several implementation techniques, including the addition of linear cuts, which significantly improve efficiency and effectiveness.

- We address the PIRP within a large-scale distribution network, which allows for practical consideration. Our research specifically aims to tackle the challenges of relatively larger distribution networks and address the limitations identified in previous studies. These studies have highlighted the need for improved policies that can effectively manage larger networks, such as those consisting of 20 retailers.

- We apply our policy to a case study within a BSC, where both replenishment and routing decisions must be made. To assess the performance of our proposed policy, we conduct extensive experiments and compare the results against well-established benchmark policies from the stochastic PIRP literature. This comparison provides valuable insights and highlights the strengths and advantages of our proposed policy.

The rest of the study is organized as follows. We review the related literature in Chapter 2, followed by a detailed analysis of a PIP and associated policies in Chapter 3. In Chapter 4, we extend the PIP to a PIRP by incorporating a second phase and discussing the new challenges this presents. We then introduce a novel policy to handle these challenges. In Chapter 5, we outline implementation matters related to our proposed policy and provide some techniques for improving its performance. To demonstrate the effectiveness of our policy, we conduct a BSC case study in Chapter 6 using platelet as a perishable product in a distribution network that includes a production site and multiple hospitals. Finally, we summarize our findings and suggest directions for future research in Chapter 7.

1.1 Notation

For any $N \in \mathbb{N}$, we use $[N]$ to denote the set $\{1, 2, \cdots, N\}$. With this notation, for any other $n \in \mathbb{N}$ the set $\{n, n+1, \cdots, N\}$ is shown by $[n, N]$. To simplify, we represent a vector $(x_1, x_2, \ldots, x_N)$ as $x_{[N]}$. Additionally, we define the function $x^+ = \max\{x, 0\}$, which returns the positive part of x, respectively.

Chapter 2

Literature Review

Inventory management, particularly for perishable products, and the complexities involved in shipping them within a distribution network, have drawn considerable attention in the literature. The high cost associated with perishable products and the presence of demand uncertainty have prompted numerous studies to optimize these systems. In this chapter, we review the related studies, exploring various approaches that have been proposed to improve the efficiency of such systems.

2.1 PIRP with Uncertainty (Stochastic PIRP)

Soysal et al. (2015) have developed a model for a PIRP with demand uncertainty in a VMI system. They have enforced a probability constraint to ensure a target service level, thereby transforming their formulation into a chance-constrained model. To handle the complexity of this constraint, they have created deterministic approximations that result in deterministic MIP models. These approximations have allowed them to obtain solutions using commercial solvers readily. The authors have implemented this multi-period model to make all decisions for the entire planning horizon at once. In a distribution network with less than 11 retailers, their models have successfully achieved significant savings in average total cost. However, when they extended the analysis to a larger distribution network with up to 20 retailers, they encountered higher complexity. In this case, optimal solutions have not been obtained, or the optimality of the solutions has not been proven within a five-hour time frame.

Crama et al. (2018) have suggested that incorporating uncertainty and combining inventory with routing decisions for perishable products can significantly increase net profit for retail chains in a VMI system. To address the challenge posed by the curse of dimensionality, they have approached the PIRP in two phases: the PIP and VRP. The Expected Value (EV) method serves as the main benchmark in their PIRP, where the daily expected demand is used to approximate the actual daily demand. In their network, they have proposed several policies to maximize profit. These policies include the up-to-level (UL) method, which considers a target service level, and the approximated dynamic programming (ADP) method, which independently determines the delivery quantity for each retailer. As their best policy, they have improved the ADP by applying a local search metaheuristic to consider real routing costs for only two periods ahead. However, it is important to note that their ADPs overlook three out of the four primary costs associated with inventory systems: holding, penalty, and wastage costs. For the second phase, they employ a classical VRP algorithm for each period.

The study conducted by Singla (2019) has focused on a PIRP that is built upon the framework originally proposed by Coelho and Laporte (2014), with the objective of maximizing profit. They have employed a two-stage stochastic programming approach, where the first stage involves determining routing decisions independent of demand realizations.

To handle uncertainty, they have utilized a finite set of demand scenarios, resulting in a scenario-based MIP model. To solve this model, they have utilized branch-and-cut and Bender's decomposition algorithms with the aid of a commercial solver. Their findings indicate that the branch-and-cut algorithm outperforms Bender's decomposition in terms of both computation time and solution quality. However, it should be noted that a significant number of demand scenarios are required to adequately capture the uncertainty, resulting in computational efforts. In order to address this limitation, they have incorporated robust optimization techniques. The robust model handles demand uncertainty by utilizing a smaller number of scenarios and aims to strike a balance between being conservative and considering worst-case scenarios. Comparison results have demonstrated that the robust model achieves average profit levels that lie between those obtained from the deterministic and stochastic models.

Onggo et al. (2019) have formulated the PIRP within a VMI system as a MIP model. In addition, they have proposed a simulation-based heuristic algorithm based on a local search metaheuristic to solve it. The algorithm utilizes Monte Carlo simulation not only to generate initial solutions but also to refine the solutions obtained through the local search process. By incorporating stochastic demand through Monte Carlo simulation, their approach distinguishes itself from previous similar studies.

In the study conducted by P. Liu et al. (2021), a mathematical model has been developed to optimize the PIRP with demand uncertainty in a BSC, utilizing the newest-unit, first-out issuing policy. To solve the robust optimization model with a scenario set, a three-phase heuristic algorithm has been proposed. In the first phase, routing decisions are made, while the remaining decisions are deferred to the subsequent phases. Through comparative analysis, it has been observed that the proposed algorithm, coupled with the inclusion of transshipment, results in reduced product shortages and lower inventory levels, ultimately leading to a decrease in overall costs.

Mousavi et al. (2022) have introduced a two-stage stochastic programming model for optimizing the total costs in a PIRP while incorporating production decisions to improve product freshness and reduce wastage. The model has been evaluated under both normal and pandemic conditions, showcasing its superior performance in dynamic and pandemic settings. To solve the model, the researchers have proposed a metaheuristic algorithm consisting of five phases, based on the decomposition technique proposed by Solyalı and Süral (2017) with certain modifications. In the computational experiments, the proposed method outperforms the CPLEX solver regarding solution quality while requiring less computational time. It is important to note that although the proposed model allows for shipping products of any age, the overall number of scenarios used in the algorithm is limited. This may not fully capture the true uncertainty present in real-world situations. For instance, in the computational experiments, a large network with up to 50 nodes is considered, but only 30 scenarios are utilized throughout the decision-making process. This limitation may not adequately capture the realistic range of uncertain situations for such a large-scale network.

2.2 PIRP without uncertainty

The majority of the literature on PIRPs primarily focuses on deterministic problems, assuming that all relevant factors, including demand, are known. This assumption allows researchers to explore different aspects of decision-making policies. However, it is important to recognize that real-world PIRPs involve complexities and uncertainties that are not captured in the deterministic setting. A notable study in this field is the work of Coelho and Laporte (2014), which has served as a foundational reference for numerous subsequent studies, for instance, the work of Singla (2019), as described earlier. In their study, Coelho and Laporte (2014) have proposed a model for joint replenishment and delivery of perishable products, with the aim of optimizing both inventory and routing decisions simultaneously. The model is formulated as a MIP and solved exactly using branch-and-cut techniques. The authors have investigated two variants of the PIRP, using oldest-unit, first-out and newest-unit, first-out selling priority policies. Additionally, they have introduced an optimized priority policy that determines which items to sell in each period, considering the trade-off between cost and revenue. The model incorporates several assumptions, such as having information on the age of products in both the supplier's and retailers' inventories, as well as their demand, making it a deterministic MIP model. Moreover, the revenue and holding costs vary with the product's age. The authors have incorporated linear cuts into their MIP and solved it using an exact classical branch-and-cut algorithm. However, the NP-hard nature of the problem has made it difficult to solve for realistic problem sizes. Overall, the paper presents an effective model for maximizing profits in medium-sized deterministic PIRPs. Furthermore, the authors have observed that the optimized priority policy outperforms both the oldest-unit, first-out and newest-unit, first-out policies.

Alvarez et al. (2020) have presented four mathematical formulations for the same deterministic PIRP as described in Coelho and Laporte (2014). Two of these formulations include a vehicle index, while the other two do not. To solve these formulations, the authors have employed branch-and-cut algorithms. Furthermore, the authors have proposed a hybrid heuristic approach that combines an iterated local search metaheuristic with two mathematical programming components. It manages routing decisions through the local search phase and the visit variables through a perturbation phase. The inventory variables are then optimized using a multi-commodity flow problem formulation for a given set of visit variables. Therefore, it performs like a two-phase approach where one decides about delivery quantities and the other works on routing decisions. Additionally, a MIP is used as a solution improvement step in the final phase of the method.

Alkaabneh et al. (2020) have addressed a PIRP within a VMI system, with a focus on estimating fuel costs and the supplier's ability to deliver products of any age, similar to the model proposed in Coelho and Laporte (2014). Their objective is to maximize the supplier's profit while minimizing costs associated with fuel consumption, inventory holding, and greenhouse gas emissions. To tackle this problem, the authors have proposed two different algorithms: Bender's decomposition and a two-stage metaheuristic, inspired by Archetti et al. (2017). In their approach using Bender's decomposition, they have incorporated various computational improvements, including the introduction of valid inequalities tailored to the routing variables. The computational results demonstrate that Bender's decomposition, when combined with several acceleration strategies, exhibits efficiency in handling small to medium-sized instances of the problem. On the other hand, their two-stage metaheuristic demonstrates the capability to handle larger instances of the problem.

In addition to the studies discussed above in relation to Coelho and Laporte (2014), there have been several studies that attempted to address PIRPs using various heuristics. The earliest one is the work by Le et al. (2013). They have developed a mathematical model for PIRP and used a column generation-based heuristic algorithm to optimize inventory and routing decisions in a deterministic PIRP. The formulation has been improved by incorporating linear cuts based on perishability and inventory bounds. The study demonstrates that integrating routing and inventory decisions can result in significant cost savings.

Mirzaei and Seifi (2015) have suggested a model that efficiently handles a PIRP by preventing overstocking and considering how the age of such inventory can negatively affect demand in a VMI. To solve the proposed MIP, the authors have employed a commercial solver for smaller cases and a metaheuristic approach for larger cases. The algorithm uses a predefined delivery quantity for each retailer and then a neighbourhood search to improve the delivery quantities. Afterward, a VRP sub-problem and a PIP are implemented consecutively within the metaheuristic framework to improve the solution in an iterative two-phase approach.

Rohmer et al. (2019) have introduced a PIRP and proposed a metaheuristic approach that combines an adaptive large neighbourhood search with a MIP formulation to solve it. They have tested three variants of the heuristic on different instances using a two-phase approach. In the first variant, the MIP is initially solved to obtain the solution for the PIP, followed by the adaptive large neighbourhood search to identify the optimal values for routing decisions. Conversely, in the second variant, the approach involves creating daily delivery routes first and then solving the PIP. The last variant is a hybrid of the first two and has demonstrated better performance on relatively larger networks. The results have shown that the cost structure plays a significant role in selecting the best variant.

Qiu et al. (2019) have addressed a PIRP integrated with a production planning problem. They have developed a MIP and have provided a branch-and-cut algorithm strengthened with logical, lot-sizing, and lifted MTZ-type valid cuts. The solution process follows a two-phase approach, with a sub-problem for production-distribution decisions and a VRP phase, for which they have employed a neighbourhood heuristic from the existing literature.

Liu et al. (2020) have addressed a PIRP and have adopted a two-phase approach similar to that of Crama et al. (2018). They have adopted this approach because they encountered computational challenges when attempting to solve instances with 20 hospitals in the network within a reasonable time frame. Their focus is on optimizing a BSC within a VMI system. In the PIP phase, they issue platelet units using a FIFO policy. They have made certain assumptions, such as assuming that expiration only occurs at the production center, as transportation to hospitals is based on realized demand. This results in a deterministic

MIP model. Like Crama et al. (2018), they have incorporated a path-related fixed cost to their MIP model to maintain the connection between PIP and VRP. After solving the multi-period MIP, they use the decided quantities in a multi-period VRP and apply an adaptive large neighbourhood search to solve this NP-hard problem. To reduce costs, they have designed a search procedure that integrates routing costs in the PIP by updating fixed cost parameters and repeating the two phases until a stopping criterion is reached. The authors argue against the exclusion of holding costs, as implemented by Crama et al. (2018), emphasizing that holding costs play an important role in decision-making for a BSC. They assert that disregarding holding costs is not a viable or practical approach, especially in the BSC of China. However, it should be noted that their VMI system does not account for stochastic demand, which limits its applicability to real-world situations.

2.3 IRP

Most of the earlier studies in the literature focused on inventory routing decisions for non-perishable or standard products. For instance, Hemmelmayr et al. (2009) have investigated the potential impact of introducing a VMI system in the BSC of the Austrian Red Cross. Their focus is on minimizing travel costs while disregarding inventory costs and not explicitly considering spoilage in inventory. To address spoilage, they adjust inventory capacities based on known demand and product shelf life. Hence, we categorize this study as an IRP. The authors have proposed three solution approaches to establish regularity in the delivery schedule: a basic heuristic, a MIP, and a variable neighbourhood. It is demonstrated that the latter two methods have significant potential for reducing routing costs. They have claimed that the proposed optimization techniques can handle non-stationary demand and multiple blood products. However, uncertain demand is a challenging setting that cannot be easily accommodated by their approaches, necessitating further investigation.

Coelho et al. (2012) have investigated an IRP that allows transshipment between retailers. They have developed an adaptive large neighbourhood search heuristic to determine delivery quantities, optimal routes, and the amount of transshipment utilizing a single vehicle. Their analysis incorporates uncapacitated and capacitated replenishment policies; however, the study does not consider stochastic demand.

Bertazzi et al. (2015) have focused on addressing the IRP with stochastic demand and transportation procurement in a VMI setting. They have found that finding an optimal solution for the stochastic problem is not straightforward. To overcome this challenge, they have proposed a heuristic approach that utilizes a rollout horizon algorithm based on the ADP method. They approximate different aspects of the value function by solving MIP models, enabling them to obtain near-optimal solutions. In order to handle the stochastic nature of the problem, they use scenarios generated by sampling from the demand distribution. In making routing decisions, they use a fixed-cost parameter as the transportation procurement cost in the MIPs. They have compared the effectiveness of their approach with the EV policy and emphasized the importance of incorporating the probability distribution of demand into the decision-making process. Although their approach shares similarities with our developed policy, their transportation fixed cost does not represent a realistic estimation of the routing costs and is not included in their metaheuristic.

Archetti et al. (2017) have presented a metaheuristic approach for solving the IRP with the objective of minimizing distribution costs. They have developed MIP models that ensure no stock-outs occur at retailer sites and vehicle capacity constraints are met. These MIP models incorporate a maximum-level replenishment policy, which enables the delivery of any quantity to retailers as long as the predetermined maximum level is not exceeded.

Sonntag et al. (2023) have presented an exact branch-price-and-cut method for addressing the stochastic IRP in an infinite horizon setting. Their method involves grouping retailers and determining fixed replenishment intervals and UL values for each group. This approach effectively balances transportation costs, emergency shipments, and holding costs. To ensure target service levels and respect vehicle capacity, the authors introduce two chance constraints. They convert the integer chance-constrained model into a MIP formulation and solve it using a column generation algorithm. The authors emphasize the importance of considering all cost components in the context of stochastic IRPs and compare the performance with an EV policy to illustrate the impact of uncertainty.

2.4 Inventory Problems in BSC

There are papers in the literature that employ multi-stage or two-stage programming techniques to address various inventory problems in the BSC. For instance, Dillon et al. (2017) have proposed a two-stage stochastic programming model for making inventory decisions in a hospital's BSC. The model is based on the multi-period multi-product lot-sizing problem and operates under an (R, S) replenishment policy. It considers demand uncertainty and the perishability of blood units. The authors have formulated a MIP model as the deterministic equivalent of the two-stage stochastic programming problem, and scenarios are generated using Monte Carlo sampling. They have used a commercial solver to find the optimal solution to the problem. Their experiments have shown that reducing the UL level, S, reduces the total cost without compromising service quality.

Hamdan and Diabat (2019) have introduced a two-stage stochastic programming approach for optimizing a red blood cell supply chain. The model incorporates four echelons: mobile blood facilities, local blood banks, regional blood banks, and hospitals. It considers various factors, such as uncertainty in both demand and supply, as well as the perishability of the product. In the first stage, the model determines the optimal number of mobile blood collection facilities to deploy, while the second stage focuses on inventory and production decisions. The authors have considered six scenarios, comprising various combinations of demand and supply scenarios. They have transformed the tri-objective problem into a single-objective MIP problem and solved it using a commercial solver.

Recently, Dillon et al. (2023) have proposed a two-stage stochastic programming model for a PIP in a platelet supply chain. In the first stage, the model determines a UL value, while the second stage focuses on allocating platelet units to hospitals without considering routing decisions. The model is formulated as a deterministic MIP using scenario sets derived from real-world data. The authors have employed the Progressive Hedging algorithm to solve the model, which facilitates a decomposed version of the model through the use of Lagrangian duality. One advantage of this approach is the ability to solve the model in parallel, ensuring a manageable computational workload. It should be noted, however, that this approach is most suitable when all decisions for a given planning horizon can be made at once. If frequent decision-making is required, such as on a daily basis, the computational steps involved may become too time-consuming.

Meneses et al. (2023) have addressed the challenges of defining optimal ordering policies for blood products in the presence of demand uncertainty. They have proposed a two-stage stochastic programming model to determine the optimal (R, S) policy parameters, considering factors such as perishability, ABO substitutions, and demand uncertainty. In the first stage of the model, the values for (R, S) are determined, and in the second stage, operational-level decisions, such as the quantity to be ordered each day, are made. The model is solved using a commercial solver. By replacing the unknown parameter with a set of discrete scenarios with known probabilities, different ordering policies are obtained.

The results show that the proposed model successfully reduces the UL level for various blood types across different policies while maintaining service levels, reducing total costs, and improving other relevant indicators.

Xu and Szmerekovsky (2023) have introduced a scenario-based multi-stage stochastic program to optimize the integrated platelet supply chain. The model focuses on decision-making regarding platelet collection, production, delivery, and transshipment among hospitals, considering different demand scenarios. To capture the uncertainty, a scenario tree is employed to represent all possible outcomes of discrete random events in the model. Specifically, two demand scenarios are generated for each day within the eight-day planning horizon. The model is solved using a commercial solver. The numerical results demonstrate the effectiveness of the proposed model in handling random daily demand, outperforming a two-stage stochastic model in a small network. However, it is important to acknowledge that their multi-stage programming approach faces challenges as the number of stages increases to 7 days. Additionally, the computational complexity increases exponentially as the number of scenarios grows with the number of time stages.

In the above-mentioned papers, the performance of the proposed approaches is compared to the deterministic formulations of their studied BSC. Moreover, routing decisions are not the primary focus of these studies.

Chapter 3

The Perishable Inventory Problem

While the primary focus of our work is on the stochastic PIRP, we believe it is essential to begin with a simpler version, the stochastic PIP, to provide a clear problem definition and develop a solid understanding of the applicable policy solutions. In particular, we discuss policies obtained by directly solving the related Dynamic Programming (DP) method, and the Direct Lookahead (DLA) models which offer practical solutions to the challenges encountered in DP. By establishing the foundation with the PIP, we can effectively address the integration of routing decisions into perishable inventory management. This will allow us to smoothly transition to the more complex PIRP and introduce our proposed policy in the next chapter.

3.1 Formulations

In inventory management systems, the decision at each period is how much to order to satisfy the demand while minimizing the total cost, including holding, penalty, wastage, and replenishment costs. The objective of such inventory problems can be to minimize the total cost or maximize the profit. The decision-making process typically occurs within a finite planning horizon, denoted by T. In this paper, we use the terms *time* and *period* interchangeably to denote the discrete period t, which is assumed to be one day.

When an inventory system deals with a product with limited shelf life, the problem is considered as a PIP. In this work, we consider a single perishable product with a shelf life of L periods, where $L > 0$. We assume that the demand in period t is a random variable denoted by D_t where the distribution of D_t is known, and the demand in periods $1, 2, \cdots, T$ are independent and identically distributed (i.i.d.). The realized demand in period t is represented by d_t, which corresponds to the actual value.

We adopt the common assumption from the literature that the finite planning horizon begins with zero inventory for all levels of $l \in [L]$. Additionally, we assume that there are no supply constraints. The sequence of events occurring in each period is as follows:

(i) The decision maker reviews the current state of the system, i.e. $S_t = (R_t^1, R_t^2, \cdots, R_t^{L-1})$ where R_t^l is the inventory level of products with a remaining shelf life of l periods at time t.

(ii) Based on the state of the system and the unlimited supply, an order with the unit price of a (if needed) is placed and fulfilled immediately. This implies that the lead time for receiving the product from the supplier is zero. As a result, the total inventory level is increased to $\sum_{l=1}^{L-1} R_t^l + x_t$, where x_t represents items received on day t, which comprise the freshest units in the inventory.

(iii) The demand is observed and satisfied using the available inventory. It is assumed that the issuing policy at the retailer is FIFO.

(iv) Any unmet demand is lost, resulting in a penalty cost of p per unit lost. Thus, our inventory system operates as a lost-sale system.

(v) Units expired in each period are discarded with a cost of e per unit expired.

(vi) All units available at the end of each period are carried to the next period, and the inventory levels are updated to R^l_{t+1} for all $l \in [L-1]$. The system incurs the holding cost of h per unit of inventory at the end of each period. We adopt a similar assumption to Coelho and Laporte (2014) that products reaching their maximum age ($l = 1$) at the end of a time period will not be stored in the regular inventory area. Rather, they will be segregated and stored separately to be discarded in the next period.

Let $C(\cdot)$ denote the cost function of the inventory system. Then, the cost incurred at time t can be written as,

$$C(S_t, x_t) = h\left(\sum_{l=1}^{L-1} R^l_t + x_t - D_t\right)^+ + p\left(D_t - \sum_{l=1}^{L-1} R^l_t - x_t\right)^+ + e(R^1_t - D_t)^+ + ax_t, \quad (3.1)$$

where the holding, penalty, wastage, and replenishment costs are captured by the first, second, third, and fourth terms, respectively. The dynamic of the system, which captures how the state of the system changes going from period t to period $t+1$, can be written as,

$$R^l_{t+1} = \left(R^{l+1}_t - \left(D_t - \sum_{j=1}^{l} R^j_t\right)^+\right)^+ \quad \forall l \in [L-2], \quad (3.2a)$$

$$R^{L-1}_{t+1} = \left(x_t - \left(D_t - \sum_{j=1}^{L-1} R^j_t\right)^+\right)^+. \quad (3.2b)$$

Equations (3.2) can be considered as the transition function $S_{t+1} = S^M(S_t, x_t, D_t)$ where $S^M(S_t, \cdot)$ is a function that takes the system state at time t, along with any other relevant information and returns the system state for the next period. We also assume that the inventory level in each period should remain below a fixed number of R^{max}. This means that we have a capacity constraint as

$$\sum_{l=1}^{L-1} R_t^l + x_t \leq R^{max}. \tag{3.3}$$

Clearly, this constraint indirectly puts an upper bound on the order quantity in each period. A lower bound may also be considered to achieve a predetermined *Target Service Level*, which reflects the desired level of service fulfillment and could be formulated deterministically as a fixed percentage or as a probabilistic metric in a stochastic context.

To determine the optimal replenishment decision, we need to repeatedly minimize the total cost of the system as new information is revealed at time $t \in [T]$. Since the demand in each period is unknown, the inventory R_t becomes a random variable, and so does the ordering quantity x_t. Due to this randomness, our multi-period problem turns into a sequential decision-making problem, in which the decision variable x_t is a function of S_t, the state of the system at time t. This relationship is commonly referred to as a policy, denoted by π, which holds the information that defines the decision-making function (Powell, 2022). Therefore, we aim to find an optimal policy that solves the following *base model*

$$\min_{\pi \in \Pi} \mathbb{E}\left[\sum_{t=1}^{T} C(S_t, X_t^\pi(S_t)) | S_0 \right], \tag{3.4}$$

where the initial state S_0 is assumed to be known, and the feasible region is expressed as

$$\mathcal{X}_t(S_t) = \{x_t | (3.2),\ (3.3),\ x_t \geq 0\}. \tag{3.5}$$

Note that by incorporating the randomness, we moved to a stochastic formulation of the problem by replacing the decision variable x_t with the policy function $X^\pi(S_t)$.

3.2 Policy Solutions

The policy function $X^{\pi}(S_t)$ can be any form of the two general classes categorized by Powell (2022). The first category is called *Policy Search*, where the parameters of a function or an optimization problem are tuned to make it efficient. In policy search, parametric models are commonly used to represent policies. Parametric models can be as simple as linear functions, which directly map a state to an action. In these analytical functions, parameters are coefficients of the independent variables. More complex structured models, such as (s, S) policies for inventory systems, also fall into this category. Finding the optimal values for parameters s and S is the main goal of the policy search here. Also, in the field of computer science, neural networks that require parameter tuning are considered within this category.

The second category is based on the *Lookahead* models, which capture the potential downstream effects of a decision made in the present on future outcomes. The basic statement of the Lookahead models would be

$$X_t^*(S_t) = \arg \min_{x_t \in \mathcal{X}_t(S_t)} \left(C(S_t, x_t) + \mathbb{E}\left[min_{\pi \in \Pi} \mathbb{E}\left[\sum_{t'=t+1}^{T} C(S_{t'}, X_{t'}^{\pi}(S_{t'})) | S_{t+1} \right] | S_t, x_t \right] \right), \tag{3.6}$$

where $X_t^{\pi}(S_t)$ is the Lookahead policy here (also referred to as *policy-within-the-policy*). Note that the expectations $\mathbb{E}[\cdots|S_t]$ and $\mathbb{E}[\cdots|S_{t+1}]$ in Equation (3.6) are the contracted forms of $\mathbb{E}_{S_t}\mathbb{E}_{W_{t+1}|S_t}$ and $\mathbb{E}_{S_{t+1}}\mathbb{E}_{W_{[t+2,T]}|S_{t+1}}$, respectively. These expectations are often written as $\mathbb{E}_W$, but we add $[\cdots|S]$ to emphasize the dependence on the initial state S.

Equation (3.6) is the well-known Bellman's equation which can be computed exactly using the DP method in certain cases. However, in many instances, the computational complexity necessitates approximating the value function in DP, which leads to the development of a group of methods known as ADPs. When ADPs cannot effectively approximate the Lookahead problem, DLA methods can be utilized to create a reasonable approximation of the future directly. We can further categorize the DLA methods into deterministic DLAs and stochastic DLAs (stochastic programming).

Since this chapter aims to develop a solid understanding of the solution approaches, we begin by examining DP for the PIP. Later in Chapter 4, we will extend DP and explain its complexities, which will lead us to the conclusion that DLAs are a more suitable approach.

3.2.1 Dynamic Programming

To develop a DP framework, the initial step involves identifying the action and state spaces. Based on the defined PIP, the state space can be represented by the set of vectors

$$\mathcal{S}_t = \{S_t | \sum_{l=1}^{L-1} R_t^l \leq R^{max}\}, \tag{3.7}$$

where $S_t = (R_t^1, R_t^2, \cdots, R_t^{L-1})$.

Since, in this chapter, we seek the optimal value for the replenishment decision, the action taken at time t would be to order x_t units of the product. In order to stay in the state space after each transition, the action space should be bounded from above by R^{max}. Hence, the feasible actions could be any element of the set

$$\mathcal{A}_t(S_t) = \{x_t | 0 \leq x_t \leq R^{max} - \sum_{l=1}^{L-1} R_t^l\}, \tag{3.8}$$

which clearly varies for each $S_t \in \mathcal{S}_t$. The value function can be then stated as

$$V_t(S_t) = \min_{x_t \in \mathcal{A}_t(S_t)} \left(C(S_t, x_t) + \mathbb{E}\Big[V_{t+1}(S_{t+1}) | S_t \Big] \right), \tag{3.9}$$

where $C(S_t, x_t)$ is defined in Equation (3.1). V_{t+1} represents the downstream impact of the action x_t taken in state S_t for all $t \in [T]$. Since time, states, and decisions are all discrete, the optimal solutions can be found using the *backward DP* technique. Interested readers can find more detailed information about the backward DP technique in Bertsekas (2005).

3.2.2 Direct Lookahead Approximations

As described earlier, when we cannot precisely solve the Lookahead model given in Equation (3.6), we can approximate the lookahead part (value function) by an ADP method. Nonetheless, accurately approximating the future cost poses several challenges. One such challenge arises from the dependencies and interactions among variables. For example, the replenishment cost may be influenced by the current inventory level. Furthermore, maintaining a precise estimate of the value function becomes challenging when the environment undergoes time-varying changes, such as fluctuating demand patterns. These changes require frequent updates to handle the non-stationary nature of the environment. Also, the *curse of dimensionality* presents another significant burden. As the problem's dimensionality increases, the size of the state and action spaces can grow exponentially, making it increasingly difficult to achieve an accurate approximation.

To address these challenges, we propose an algorithm based on DLA methods. As previously mentioned, DLAs directly approximate the impact of potential future outcomes on the current decisions. To gain a deeper understanding of this approach, we first explore the structure of sequential decision-making and Lookahead policies.

In a sequential decision-making problem, the decision-making process involves a sequence of states, decisions, and information that repeats until the end of the planning horizon. In a Lookahead model, we aim to make a decision at time t while considering its impact on future periods. This is akin to standing at the starting point of the sequence, represented as S_t. The information regarding the subsequent period, denoted as $W_{t,t+1}$, becomes available after the decision of the current day is made, i.e., x_t. When utilizing DLA methods, we approximate the entire sequence as

$$(S_t, x_t, \tilde{W}_{t,t+1}, \tilde{S}_{t,t+1}, \tilde{x}_{t,t+1}, \tilde{W}_{t,t+2}, \ldots, \tilde{S}_{tt'}, \tilde{x}_{tt'}, \tilde{W}_{tt'}, \ldots, \tilde{S}_{tT}, \tilde{x}_{tT}, \tilde{W}_{tT}), \tag{3.10}$$

where $\tilde{W}_{tt'}$, $\tilde{S}_{tt'}$ and $\tilde{x}_{tt'}$ are approximations of $W_{tt'}$, $S_{tt'}$ and $x_{tt'}$, respectively. The subscripts indicate that the decision is being made for the current time t while viewing its impact on future periods $t' \in [t+1, T]$.

We can simplify the Lookahead model using various approximation strategies, including *outcome aggregation*. According to this strategy, we use an aggregated set of scenarios instead of the entire set, i.e. $\tilde{\Omega} \subset \Omega$. By adopting this widely used strategy, we can express the DLA framework as

$$X_t^{DLA}(S_t) = \arg \min_{x_t \in \mathcal{X}_t(S_t)} \left(C(S_t, x_t) + \mathbb{E}\left[\min_{\tilde{\pi} \in \tilde{\Pi}} \tilde{\mathbb{E}}\left[\sum_{t'=t+1}^{T} C(\tilde{S}_{tt'}, \tilde{X}_{tt'}^{\tilde{\pi}}(\tilde{S}_{tt'})) | \tilde{S}_{t,t+1} \right] | S_t, x_t \right] \right), \tag{3.11}$$

where $\tilde{\Pi}$ is the modified set of Lookahead policies and $\tilde{\mathbb{E}}$ is the expectation over the aggregated set of random outcomes $\tilde{\Omega}$.

It is important to note that $X_t^{\pi}(S_t)$ is a policy-within-the-policy, which means any other type of policy can be employed within the DLA policy. As a case in point, Lookahead Cost Function Approximation is a hybrid policy specifically designed for sequential decision-making problems involving forecasts. As it falls beyond the scope of this study, we will not elaborate on it further. Interested readers can refer to Powell and Ghadimi (2022) for more detail.

In addition, there are other strategies to simplify the determination of the solution for X_t^{DLA}. One such strategy is known as *stage aggregation*, which involves combining stages in the decision-making process. Each stage consists of an information-disclosure process followed by decision-making. By employing this strategy, we can establish an alternative formulation for Equation (3.11), referred to as *two-stage approximation* of the sequential problem. In the following, we first introduce *multistage stochastic programming* and then apply the stage aggregation strategy.

Two-Stage Approximation

The PIP can also be formulated as a basic multistage stochastic program as follows:

$$\begin{aligned} \min_{x_t \in \mathcal{X}_t(S_t)} C(S_t, x_t) + \mathbb{E}\Bigg[\min_{x_{t+1} \in \mathcal{X}_{t+1}(S_{t+1})} C(S_{t+1}, x_{t+1}) + \mathbb{E}\Bigg[\cdots \\ + \mathbb{E}\Bigg[\min_{x_T \in \mathcal{X}_T(S_T)} C(S_T, x_T) | S_T) \Bigg] \cdots | S_{t+1} \Bigg] | S_t \Bigg]. \end{aligned} \tag{3.12}$$

In theory, Equation (3.12) is equivalent to the value function given in Equation (3.9) and should provide the optimal solution. However, practical challenges associated with DP, such as dependencies, non-stationary environments, and the curse of dimensionality, can still pose difficulties and restrict its efficiency. To overcome these challenges, we adopt the DLA methodology. Additionally, we employ different approximation strategies to further improve the efficiency of the algorithm. One such strategy is stage aggregation, where we approximate the problem in (3.12) as a two-stage problem by making the initial decision first. This strategy helps simplify the problem and mitigate the computational complexity associated with it. This is followed by observing all future events from an aggregated sample set $\tilde{\Omega}_t$. Afterwards, all remaining decisions $\tilde{x}_{tt'}$ for $t' \in [t+1, T]$ are made. This results in a change to the sequence in (3.10), such that

$$\Big(S_t, x_t, (\tilde{W}_{t,t+1}, \tilde{W}_{t,t+2}, \dots), (\tilde{x}_{t,t+1}, \tilde{x}_{t,t+2}, \dots) \Big). \tag{3.13}$$

The two-stage approximation of the sequential problem in Equation (3.12) can then be presented as

$$\min_{x_t \in \mathcal{X}_t(S_t)} C(S_t, x_t) + \mathbb{E}\left[\min_{\left(\tilde{x}_{tt'} \in \tilde{\mathcal{X}}(\tilde{S}_{tt'})\right)_{t'=(t,T]}} \sum_{t'=t+1}^{T} C\big(\tilde{S}_{tt'}, \tilde{x}_{tt'}\big) | S_t \right]. \tag{3.14}$$

The main idea behind this approach is to optimize an approximated problem across $|\tilde{\Omega}_t|$ number of scenarios and over the entire horizon, all at once, while only using the values of x_t at the end. This approach is distinct from the deterministic DLA only in terms of considering $|\tilde{\Omega}_t|$ samples instead of the actual value. Thus, we have

$$\Big(S_t, x_t, (\tilde{W}_{t,t+1}(\tilde{\omega}_t), \tilde{W}_{t,t+2}(\tilde{\omega}_t), \dots), (\tilde{x}_{t,t+1}(\tilde{\omega}_t), \tilde{x}_{t,t+2}(\tilde{\omega}_t), \dots)\Big), \tag{3.15}$$

where each $\tilde{\omega}_t \in \tilde{\Omega}_t$ represents a full sequence of the random variables $\tilde{W}_{[t+1,T]}$. Such a sequence is called a *scenario* or a *sample path*, and it occurs with a probability of $P(\tilde{\omega}_t)$.

The use of scenarios proves to be an effective method for capturing the inherent uncertainty of the problem. This involves assigning values to random variables, such as the amount of each period's demand in the PIP. By using scenarios, we can accurately account for the variability in the problem. This allows us to express the two-stage approximated problem in Equation (3.14) as follows:

$$\min \; C(S_t, x_t) + \sum_{\tilde{\omega}_t \in \tilde{\Omega}_t} P(\tilde{\omega}_t) \sum_{t'=t}^{T} C\big(\tilde{S}_{tt'}(\tilde{\omega}_t), \tilde{x}_{tt'}(\tilde{\omega}_t)\big) \tag{3.16a}$$

$$\text{s.t.} \qquad x_t \in \mathcal{X}(S_t), \tag{3.16b}$$

$$\tilde{x}_{tt'}(\tilde{\omega}_t) \in \tilde{\mathcal{X}}\big(\tilde{S}_{tt'}(\tilde{\omega}_t)\big) \qquad \forall t' \in [t+1, T]. \tag{3.16c}$$

Note that we use the tilde notation to distinguish the approximations. In fact, $\tilde{x}_{tt'}(\tilde{\omega}_t)$ does not represent the actual decision made for the future time t'. Instead, they are decisions associated with the random variables that appear in scenario $\tilde{\omega}_t$. Their sole purpose is to assist in making a more favorable decision for the current period by capturing the potential future outcomes. In a concise representation, first-stage decisions are denoted as x_t, while $\tilde{x}_{tt'}(\tilde{\omega}_t)$ represents the second-stage decisions.

In conclusion of this chapter, our exploration of the stochastic PIP has enabled us to define the problem and gain insights into diverse policy solutions. We have examined policies derived from the DP approach, as well as the application of DLA policies. With this groundwork established, we are well-prepared to explore the complexities of the PIRP and introduce our innovative policy in the following chapter.

Chapter 4

Perishable Inventory Routing Problem and the Proposed Policy

The goal of this chapter is twofold. First, we aim to expand the stochastic PIP introduced in Chapter 3 and formulate the stochastic PIRP. This will provide a comprehensive understanding of the problem at hand. Second, we explore the DP solution approach, building upon the foundations established in the previous chapter. We also examine ADP methods and their potential for addressing the stochastic PIRP. We then propose a novel algorithm based on the DLA methodology, which offers an effective and efficient solution to the problem.

4.1 Formulations

When routing decisions are involved, the PIP becomes more complex and is referred to as the PIRP. In our case, the PIRP is associated with a one-to-many distribution network, where a single supplier serves multiple retailers. We assume that the system operates as a VMI where the supplier should determine the optimal quantity of the perishable product to be delivered to each retailer while considering inventory constraints. Furthermore, the supplier needs to develop an efficient routing plan and assign retailers to different delivery routes, while considering the distribution constraints. To summarize, our PIRP involves the challenges of managing perishable inventory and the complexities of designing optimal routes for product delivery in a one-to-many distribution network that operates under the VMI system.

We define the problem using a complete graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ where the sets $\mathcal{N} = \{0, 1, \cdots, N\}$ and $\mathcal{E} = \{(i, j) : i, j \in \mathcal{N}, i \neq j\}$ represent the sets of nodes and edges, respectively. We let node 0 be the supplier, and the set $\mathcal{N}^0 = \mathcal{N} \setminus \{0\}$ stands for N retailers. The traveling cost through the edge (i, j) is also shown by c_{ij}. The deliveries are performed by a sufficiently large fleet of homogeneous vehicles with capacity Q where every travel begins and ends at the supplier, creating one route for the problem. We assume each retailer is assigned to only one route and served by a single vehicle in each period. This means that each vehicle can take one route at most, and there are no split deliveries.

We proceed under the assumption that the supplier does not hold any inventory and has the capability to meet all demand with fresh products in each period. This assumption aligns with the absence of supply constraint that we previously established in the PIP. This can be understood as the supplier having unlimited capacity, which enables replenishment quantities to be independent of one another. On the other hand, retailer i from $\mathcal{N}^0$ operates an inventory system with a certain capacity denoted as $R_i^{max} \leq Q$. This capacity assumption is consistent with perishable inventory systems, where fresh orders are typically not placed in large quantities.

In a PIRP, a similar sequence of events to that of the PIP occurs for each retailer, as described in Section 3.1. The only distinction is that a centralized decision maker, at each period t, determines the replenishment quantities, denoted by $x_{t,i}$, for every retailer $i \in \mathcal{N}^0$. Subsequently, the decision maker selects the delivery routes, represented by $y_t = (y_{ij,t})_{(i,j)\in\mathcal{E}}$. Additionally, the decision maker may establish additional criteria, such as defining a target service level for each retailer.

To provide a comprehensive understanding, we begin by separately defining the PIP and VRP formulations. Subsequently, we combine these formulations to form the PIRP as a unified problem. This approach enables us to set the foundation for the proposed policy solution, which will be discussed in the subsequent section.

4.1.1 PIP's Formulation:

We use similar formulations as in Section 3.1 with the only difference that the PIP's notations are appended by a subscript $i \in \mathcal{N}^0$ to differentiate between the multiple retailers in the network. Based on the stochastic nature of the demand, the problem can be expressed as follows:

$$\begin{aligned} C\left(S_t\right) = &\min\ C(S_t, x_t) \\ =&\min\ \mathbb{E}\Bigg[\sum_{i\in\mathcal{N}^0} h\left(\sum_{l=1}^{L-1} R^l_{t,i} + x_{t,i} - D_{t,i}\right)^+ + p\left(D_{t,i} - \sum_{l=1}^{L-1} R^l_{t,i} - x_{t,i}\right)^+ \\ &\qquad + e(R^1_{t,i} - D_{t,i})^+ + a x_{t,i}\Bigg] \end{aligned} \tag{4.1a}$$

s.t.

$$R_{t+1,i}^{l} = \left(R_{t,i}^{l+1} - \left(D_{t,i} - \sum_{j=1}^{l} R_{t,i}^{j} \right)^{+} \right)^{+} \quad \forall l \in [L-2],\ i \in \mathcal{N}^0, \quad (4.1b)$$

$$R_{t+1,i}^{L-1} = \left(x_{t,i} - \left(D_{t,i} - \sum_{j=1}^{L-1} R_{t,i}^{j} \right)^{+} \right)^{+} \quad \forall i \in \mathcal{N}^0, \quad (4.1c)$$

$$x_{t,i} \leq R_{i}^{max} - \sum_{l=1}^{L-1} R_{t,i}^{l} \quad \forall i \in \mathcal{N}^0, \quad (4.1d)$$

$$x_{t,i} \geq 0 \quad \forall i \in \mathcal{N}^0, \quad (4.1e)$$

where $x_t = (x_{t,i})_{i\in\mathcal{N}^0}$ and $S_t = (S_{t,i})_{i\in\mathcal{N}^0} = ((R_{t,i}^{l})_{l\in[L-1]})_{i\in\mathcal{N}^0}$. From this point, we define the feasible set of the PIP as,

$$\mathcal{X}_t\left(S_t\right) = \{x_t | (4.1b) - (4.1e)\}. \quad (4.2)$$

Equation (4.1a) aims to minimize the expected total cost of the PIP at time t. The constraints (4.1b)-(4.1c) capture the dynamics of the system, assuming that the retailers follow the FIFO issuing policy. Constraint (4.1d) ensures that orders are placed in a manner that does not exceed the inventory capacity of each retailer upon receiving the products. Lastly, constraint (4.1e) represents the non-negativity of the variables involved in the formulation.

4.1.2 VRP's Formulation:

With all assumptions described earlier in Section 4.1, the VRP seeks an optimal solution to the following formulation

$$G(x_t) = \min\ G(x_t, y_t) = \min \sum_{(i,j)\in\mathcal{E}, i\neq j} c_{ij} y_{ij,t} \tag{4.3a}$$

$$\text{s.t.} \quad \sum_{i\in\mathcal{N}^0} x_{t,i} \leq Q z_t, \tag{4.3b}$$

$$\sum_{i\in\mathcal{N}, i\neq j} y_{ij,t} = 1 \qquad \forall j \in \mathcal{N}^0, \tag{4.3c}$$

$$\sum_{j\in\mathcal{N}, j\neq i} y_{ji,t} = 1 \qquad \forall i \in \mathcal{N}^0, \tag{4.3d}$$

$$\sum_{i\in\mathcal{N}^0} y_{i0,t} = z_t, \tag{4.3e}$$

$$\sum_{j\in\mathcal{N}^0} y_{0i,t} = z_t, \tag{4.3f}$$

$$q_{t,j} - q_{t,i} \geq x_{t,j} - Q(1 - y_{ij,t}) \qquad \forall i, j \in \mathcal{N}^0, i \neq j, \tag{4.3g}$$

$$0 \leq q_{t,i} \leq Q - x_{t,i} \qquad \forall i \in \mathcal{N}^0, \tag{4.3h}$$

$$y_{ij,t} \in \{0, 1\} \qquad \forall (i,j) \in \mathcal{E}, i \neq j, \tag{4.3i}$$

$$z_t \in \mathbb{Z}^+, \tag{4.3j}$$

where $q_{t,i}$ is the remaining capacity after visiting node i. The set of feasible solutions for the VRP problem is

$$\mathcal{Y}_t(x_t) = \{y_t | (4.3b) - (4.3j)\}. \tag{4.4}$$

The objective function in (4.3a) minimizes the travel cost at time t. Constraint (4.3b) enforces the number of vehicles required for the delivery. Constraints (4.3c)-(4.3d) represent the flow conservation constraints governing the flow of the items between nodes. Constraints (4.3e)-(4.3f) guarantee that the vehicles return to the supplier after completing their travel. The constraints (4.3g)-(4.3h) are known as the MTZ constraints, which eliminate sub-tours while considering each vehicle's capacity. Lastly, constraints (4.3i)-(4.3j) indicate the type of the decision variables.

It can be seen that constraints (4.3b), (4.3g), and (4.3h) link the VRP to the PIP's decisions. Since demand is realized before this stage, the decisions derived from the PIP serve as inputs to the VRP. In other words, the routing decisions in the VRP are based on the earlier decisions made in the PIP.

4.1.3 PIRP's Formulation Over a Finite Horizon:

Ultimately, we integrate the PIP and VRP formulations in (4.1) and (4.3) over the entire horizon. The resultant is a multi-period stochastic PIRP as

$$\mathrm{PIRP}(S_0) = \min_{\pi \in \Pi} \mathbb{E}\left[\sum_{t=1}^{T} C(S_t, X_t^{\pi}(S_t)) + G(x_t, Y_t^{\pi}(x_t)) | S_0\right], \tag{4.5}$$

where X_t^{π} and Y_t^{π} are the policies that determine decision variables x_t and y_t based on the feasible sets $\mathcal{X}(S_t)$ and $\mathcal{Y}(x_t)$, respectively, for all $t \in [T]$.

4.2 Policy Solutions

In this section, we attempt to solve the stochastic problem given in Equation (4.5) by employing Lookahead policies where its basic statement for PIRP at time t would be

$$X_t^*(S_t), Y_t^*(x_t) = \arg \min_{x_t \in \mathcal{X}_t(S_t),\, y_t \in \mathcal{Y}_t(S_t)} \Bigg(C(S_t, x_t) + G(x_t, y_t) \\ + \mathbb{E}\Bigg[\min_{\pi \in \Pi} \mathbb{E}\Bigg[\sum_{t'=t+1}^{T} C(S_{t'}, X_{t'}^{\pi}(S_{t'})) + G(x_t, Y_t^{\pi}(x_t)) \\ |S_{t+1}, x_{t+1}) \Bigg] |S_t, x_t \Bigg] \Bigg). \tag{4.6}$$

As discussed in Chapter 3, DP, ADP, and DLAs are various forms of the Lookahead policy. While we can directly focus on the DLA approach, we choose first to establish the DP framework. This allows us to highlight the challenges associated with high dimensions and the limitations of DP in solving the stochastic PIRP. We then apply approximation techniques to show that the complexities persist even when using ADP methods and seeking optimal solutions through the backward DP technique. Consequently, a DLA methodology is more suitable for addressing the problem. Specifically, we adopt a two-stage approximation approach within the stochastic DLA framework as it offers superior capabilities in handling complexities more effectively than other methods.

4.2.1 Dynamic Programming

We start at T and proceed backward. The state space consists of $S_T = (S_{T,i})_{i \in \mathcal{N}^0}$ where $S_{T,i} = (R_{T,i}^l)_{l \in [L-1]}$. With the same notation, we show the replenishment decision by the vector $x_T = (x_{T,i})_{i \in \mathcal{N}^0}$ where $x_{T,i} \leq R_i^{max} - \sum_{l=1}^{L-1} R_{T,i}^l$. Moreover, one routing decision should be taken at this time, shown by y_T. The action space is then $\mathcal{A}_T(S_T) = \mathcal{X}_T(S_T) \cup \mathcal{Y}_T(x_T)$ and the value function can be written as,

$$\begin{aligned} V_T(S_t) &= \min_{x_T,\, y_T \in \mathcal{A}_T(S_t)} C(S_T, x_T) + G(x_T, y_T) \\ &= \min_{x_T \in \mathcal{X}_T(S_t)} \big\{ C(S_T, x_T) + \min_{y_T \in \mathcal{Y}_T(x_t)} G(x_T, y_T) \big\}, \end{aligned} \tag{4.7}$$

where $\min_{y_T \in \mathcal{Y}_T(x_t)} G(x_T, y_T)$ is a VRP that needs to be solved for every vector of the replenishment quantities, i.e., $\forall x_T \in \mathcal{X}_T(S_T)$. The same procedure is carried out for period $T-1$, with the difference that the value function now includes the expected cost of future periods, which is only period T here:

$$\begin{aligned} V_{T-1}(S_{T-1}) &= \min_{x_{T-1},\, y_{T-1} \in \mathcal{A}_{T-1}(S_{T-1})} C(S_{T-1}, x_{T-1}) + G(x_{T-1}, y_{T-1}) \\ &\qquad + \mathbb{E}_{D_{T-1}}\big[V_T(S_{T-1}, x_{T-1}, D_{T-1})\big] \\ &= \min_{x_{T-1} \in \mathcal{X}_{T-1}(S_{T-1})} \big\{ C(S_{T-1}, x_{T-1}) + \mathbb{E}_{D_{T-1}}\big[V_T(S_{T-1}, x_{T-1}, D_{T-1})\big] + g(x_{T-1}) \big\}, \end{aligned} \tag{4.8}$$

where

$$g(x_{T-1}) = \min_{y_{T-1} \in \mathcal{Y}_{T-1}(x_{T-1})} G(x_{T-1}, y_{T-1}). \tag{4.9}$$

In the same vein, the value function at time t becomes

$$V_t(S_t) = \min_{x_t \in \mathcal{X}_t(S_t)} \big\{ C(S_t, x_t) + \mathbb{E}_{D_t}\big[V_{t+1}(S_t, x_t, D_t)\big] + g(x_t) \big\}, \tag{4.10}$$

where the VRP is

$$g(x_t) = \min_{y_t \in \mathcal{Y}_t(x_t)} G(x_t, y_t). \tag{4.11}$$

It is straightforward to verify there are $\prod_{i \in \mathcal{N}^0} (R_i^{max} + 1)^{(L-1)}$ number of states that needs to be observed to find $N + 1$ optimal decisions at period t. Therefore, the curse of dimensionality makes the backward DP technique inefficient. Due to this complexity, approximations can be applied to different parts of V_t, namely the state space, action space, sample space, and lookahead part, resulting in an ADP policy. We apply all these approximations and use the backward DP technique to find the solutions. The backward DP technique iterates over the loops in Algorithm 1, where $S_t = (S_{t,i})_{i \in \mathcal{N}^0}$ and $x_t = (x_{t,i})_{i \in \mathcal{N}^0}$. To explain further, we start from the last period and calculate the value function for every possible state $S_T \in \tilde{\mathcal{S}}_T$. Then, we proceed backward and repeat the process for each prior period $(T \downarrow 1)$. This is why it is referred to as a backward technique. The crucial aspect is to consider all possible actions $x_t \in \tilde{\mathcal{X}}_t (S_t)$ and determine which one yields the best value for the value function $V_t (S_t)$ for all $S_t \in \tilde{\mathcal{S}}_t$. This evaluation takes into account various scenarios of the random demand $D_t \in \tilde{\Omega}_t$ with their corresponding probabilities $\mathbb{P} (D_t)$.

The ADP algorithm outlined in Algorithm 1 involves solving a VRP for each decision vector, which makes the problem to be still computationally intractable. To address this challenge, there are papers that attempt to approximate $g (x_t)$ as defined in Equation (4.11). One example of a policy that incorporates this approximation is the Decomposition algorithm (DE) proposed by Crama et al. (2018). In the DE, the VRP is eliminated, and instead, a fixed cost parameter is used for each retailer. This eliminates the need to solve the VRP repeatedly. Subsequently, the DP is decomposed into separate sub-problems, resulting in N individual DP instances, one for each retailer. The replenishment decisions obtained from this ADP are then utilized to solve the VRP only once. Thus, this approach addresses the problem in two phases, which ignores the true effect of routing costs on replenishment decisions in the first phase.

To improve the outcomes of the ADP methods, some papers apply (meta)heuristics. For instance, by implementing a neighbourhood search, Crama et al. (2018) improve their routing decisions made through the decomposition method, DE. As they search locally for

Algorithm 1 ADP with Backward Technique (Bertsekas, 2005)

Initialize: $K = 0,\ V_{T+1}(S_{T+1}) = 0\ \forall S_{T+1} \in \tilde{\mathcal{S}}_{T+1}$

```
for t = T ↓ 1 do
    for S_t ∈ S̃_t do
        V_t(S_t) ← M
        for x_t ∈ X̃_t(S_t) do
            K ← C(S_t, x_t)
            for D_t ∈ Ω̃_t do
                K ← K + ℙ(D_t) Ṽ_{t+1}(S_t, x_t, D_t)
            end for
            K ← K + g(x_t)
            if K < V_t(S_t) then
                V_t(S_t) ← K
            end if
        end for
    end for
end for
```

a better routing plan, they update their replenishment decision vector by solving an integer programming model. While this Decomposition-Integration algorithm (DI) appears to produce favourable results, it suffers from the same drawback as all other ADPs. To be more specific, an increase in the shelf life by one unit would result in having additional $R_i^{max}(R_i^{max}+1)^{L-1}$ number of states for retailer $i \in \mathcal{N}^0$. Thus, as the shelf life increases, the exponential growth in computational complexity becomes increasingly challenging to manage. In our preliminary numerical experiments, we observed that the computational complexity of the ADPs grows significantly and becomes impractical to solve the problem when the product's shelf life exceeds 5. In summary, for the majority of supply chains with perishable products, partially approximating the DP is inadequate for overcoming the curse of dimensionality. Therefore instead of ADPs, we propose employing DLA methods for stochastic PIRPs.

4.2.2 Direct Lookahead Approximation

To overcome the challenges associated with finding the optimal values for Equation (4.10), we explore adopting the two-stage approximation approach from the stochastic DLA category. Following the structure of the model in (3.16), we incorporate multiple scenarios in the stochastic PIRP to estimate the costs of future days. Hence, we aim to solve the following DLA model:

$$\begin{aligned}\text{DLA}\left(S_t\right) = \min\ & C(S_t, x_t) + G(x_t, y_t) + \sum_{\tilde{\omega}_t \in \tilde{\Omega}_t} P(\tilde{\omega}_t)\Big(\sum_{t'=t+1}^{T} C(\tilde{S}_{tt'}\left(\tilde{\omega}_t\right), \tilde{x}_{tt'}\left(\tilde{\omega}_t\right)) \\ & + G(\tilde{x}_{tt'}\left(\tilde{\omega}_t\right), \tilde{y}_{tt'}\left(\tilde{\omega}_t\right))\Big)\end{aligned} \tag{4.12a}$$

s.t.

$$x_t \in \mathcal{X}\left(S_t\right), \tag{4.12b}$$

$$\tilde{x}_{tt'}\left(\tilde{\omega}_t\right) \in \tilde{\mathcal{X}}\big(\tilde{S}_{tt'}\left(\tilde{\omega}_t\right)\big) \qquad \forall t' \in [t+1, T],\ \forall \tilde{\omega}_t \in \tilde{\Omega}_t, \tag{4.12c}$$

$$y_t \in \mathcal{Y}\left(x_t\right), \tag{4.12d}$$

$$\tilde{y}_{tt'}\left(\tilde{\omega}_t\right) \in \tilde{\mathcal{Y}}\big(\tilde{x}_{tt'}\left(\tilde{\omega}_t\right)\big) \qquad \forall t' \in [t+1, T],\ \forall \tilde{\omega}_t \in \tilde{\Omega}_t. \tag{4.12e}$$

Based on the problem structure, it can also be written as

$$\begin{aligned}
\mathrm{DLA}\left(S_t\right) = \min_{x_t, \tilde{x}_{tt'}(\tilde{\omega}_t),\ \forall t' \in [t+1,T],\ \forall \tilde{\omega}_t \in \tilde{\Omega}_t} \Bigg\{ & C(S_t, x_T) + \sum_{\tilde{\omega}_t \in \tilde{\Omega}_t} P(\tilde{\omega}_t) \sum_{t'=t+1}^{T} C(\tilde{S}_{tt'}\left(\tilde{\omega}_t\right), \tilde{x}_{tt'}\left(\tilde{\omega}_t\right)) \\
& + \min_{y_t, \tilde{y}_{tt'}(\tilde{\omega}_t)\ \forall t' \in [t+1,T],\ \forall \tilde{\omega}_t \in \tilde{\Omega}_t} \Big\{ G(x_T, y_T) \\
& + \sum_{\tilde{\omega}_t \in \tilde{\Omega}_t} P(\tilde{\omega}_t) \sum_{t'=t+1}^{T} G(\tilde{x}_{tt'}\left(\tilde{\omega}_t\right), \tilde{y}_{tt'}\left(\tilde{\omega}_t\right)) \Big\} \Bigg\}
\end{aligned} \tag{4.13}$$

$$\text{s.t.} \quad (4.12b) - (4.12e).$$

This decomposition is possible because the first two terms in Equation (4.13) do not include any routing decision variables y_t and $\tilde{y}_{tt'}$. Therefore, we can separate these terms and place the ones that involve the routing decisions into an inner minimization problem.

The last term in Equation (4.13) involves solving a VRP for every future day in each scenario. This imposes the same significant computational burden that was evident in Algorithm 1. At this stage, we adopt a two-phase approach that aligns with previous studies that have acknowledged the computational complexity of such problems (see e.g. Crama et al. (2018) and Liu et al. (2020)). To elaborate, we divide the problem into two subsequent phases. The first phase involves solving problem (4.1), addressing the PIP, while the second phase focuses on problem (4.3), targeting the VRP for each period.

Dividing the problem in (4.13) into two separate phases may overlook the effect of route plans on the replenishment decisions. Therefore, to bridge the gap between the optimal solutions of the two-phase approach and those of the PIRP, we use an idea from the existing literature (see e.g. Crama et al. (2018) and Liu et al. (2020)). Specifically, we introduce a fixed cost parameter as an estimation of the routing costs on each day of the planning horizon. This approximation allows the decisions made in the first phase to have insights into potential optimal solutions of the subsequent phase, thereby enabling more efficient decision-making.

With the implementation of the two-phase approach, the DLA policy focuses solely on the decision variables of the PIP. The routing decisions, on the other hand, are left to be determined by a VRP algorithm in the second phase. Consequently, there is no longer a need to make routing decisions for future periods, denoted as $\tilde{y}_{t[t+1,T]}\left(\tilde{\omega}_t\right)$. Therefore, the formulation used in the first phase with the presence of $\hat{g}_{t'}$ as the fixed cost, becomes

$$\begin{aligned} \widehat{\mathrm{DLA}}\left(S_t\right) = \min C(S_t, x_t) + \hat{g}_t b_t^1 + \sum_{\tilde{\omega}_t \in \tilde{\Omega}_t} P(\tilde{\omega}_t) \Big(\sum_{t'=t+1}^{T} & C(\tilde{S}_{tt'}\left(\tilde{\omega}_t\right), \tilde{x}_{tt'}\left(\tilde{\omega}_t\right)) \\ & + \hat{g}_{t'} \tilde{b}_{tt'}^1(\tilde{\omega}_t)) \Big) \end{aligned} \quad (4.14)$$

$$\begin{aligned} \text{s.t.} \quad & (4.12b) - (4.12c), \\ & x_t \le R^{max} b_t^1, \\ & \tilde{x}_{tt'}(\tilde{\omega}_t) \le R^{max} \tilde{b}_{tt'}^1(\tilde{\omega}_t) \quad \forall t' \in [t+1, T],\ \forall \tilde{\omega}_t \in \tilde{\Omega}_t, \\ & b_t^1,\ \tilde{b}_{tt'}^1(\tilde{\omega}_t) \in \{0,1\} \quad \forall t' \in [t+1, T],\ \forall \tilde{\omega}_t \in \tilde{\Omega}_t, \end{aligned}$$

where b_t^1 is a binary variable that equals one if an order is placed for the period t.

It is noteworthy that all variables pertaining to the future require the notation $v\tilde{a}r_{tt'}\left(\tilde{\omega}_t\right)$ to indicate their association with the sample path of $\tilde{\omega}_t$ in the approximated problem. Nonetheless, for the sake of simplicity, we will use $v\tilde{a}r_{t'}$ to represent the same concept.

Algorithm 2 DLA Policy

Step 1: Solve problem (4.14) and keep the optimal values of x_t

Step 2: Solve a VRP

Step 3: Observe the realized demand, d_t, and compute the actual cost via Equation (4.1a)

Step 4: Update the inventory of each retailer via Equations (4.1b)-(4.1c)

Algorithm 2 displays the steps for implementing the DLA policy at time t. The initial step involves determining the optimal solutions for the decision variables in problem (4.14). Afterwards, only the optimal values of x_t are considered, while the remaining values are disregarded. This is because the values of $\tilde{x}_{tt'}$ serve the purpose of capturing the potential impact of future scenarios. In other words, they are solely intended to evaluate the downstream effects of the current period's replenishment decisions on all future periods, as per the definition of the Lookahead policy. In the next step, a VRP needs to be solved, with the replenishment decisions of all retailers serving as input. Any well-developed VRP algorithm can be applied in this context. At this stage, the demand becomes known, and the actual cost of the network can be computed using Equation (4.1a). At the end of the procedure, all remaining inventory levels are updated based on Equations (4.1b) and (4.1c). Once these updates are completed, the entire process is repeated from the initial step on the next day. By iterating through this sequence until the final day, it is possible to determine the total cost incurred by the DLA policy throughout the planning horizon.

The problem (4.14) in **Step 1** of Algorithm 2 is nonlinear due to the inclusion of terms $(x)^+ = \max\{0, x\}$ in both the objective function and constraints. Consequently, in order to utilize a state-of-the-art solver (optimizer) such as Gurobi, it becomes necessary to linearize the problem. This linearization process results in the following expanded MIP model:

$$\text{MIP}\,(S_t) = \min \sum_{i\in\mathcal{N}^0} \Bigg(ax_{t,i} + \hat{g}_{t,i}b^1_{t,i} + \sum_{\tilde{\omega}_{t,i}\in\tilde{\Omega}_{t,i}} P(\tilde{\omega}_{t,i}) \Bigg(h\tilde{u_{t,i}} + p\tilde{v_{t,i}} + e\tilde{o_{t,i}} + \sum_{t'=t+1}^{T} \left(h\tilde{u}_{t',i} + p\tilde{v}_{t',i} + e\tilde{o}_{t',i} + a\tilde{x}_{t',i} + \hat{g}_{t',i}\tilde{b}^1_{t',i} \right) \Bigg) \Bigg), \tag{4.15a}$$

s.t. $\quad \forall i \in \mathcal{N}^0,\ \forall t' \in [t,T],\ \forall \tilde{\omega}_{t,i} \in \tilde{\Omega}_{t,i} :$

$$x_{t,i} = \tilde{x}_{t,i} \tag{4.15b}$$

$$b^1_{t,i} = \tilde{b}^1_{t,i} \tag{4.15c}$$

$$\tilde{u}_{t',i} \geq \sum_{l=1}^{L-1} \tilde{R}^l_{t',i} + \tilde{x}_{t',i} - \tilde{D}_{t',i} \tag{4.15d}$$

$$\tilde{v}_{t',i} \geq \tilde{D}_{t',i} - \sum_{l=1}^{L-1} \tilde{R}^l_{t',i} - \tilde{x}_{t',i} \tag{4.15e}$$

$$\tilde{o}_{t',i} \geq \tilde{R}^1_{t',i} - \tilde{D}_{t',i} \tag{4.15f}$$

$$\tilde{x}_{t',i} \leq R^{max}_i \tilde{b}^1_{t',i} \tag{4.15g}$$

$$\tilde{x}_{t',i} \leq R^{max}_i - \sum_{l=1}^{L-1} \tilde{R}^l_{t',i} \tag{4.15h}$$

$$\tilde{y}^l_{t',i} \geq \tilde{D}_{t',i} - \sum_{j=1}^{l} \tilde{R}^j_{t',i} \qquad \forall l \in [L-1] \tag{4.15i}$$

$$\tilde{y}^l_{t',i} \leq \tilde{D}_{t',i} - \sum_{j=1}^{l} \tilde{R}^j_{t',i} + M\tilde{b}^{2,l}_{t',i} \qquad \forall l \in [L-1] \tag{4.15j}$$

$$\tilde{y}^l_{t',i} \leq M\left(1 - \tilde{b}^{2,l}_{t',i}\right) \qquad \forall l \in [L-1] \tag{4.15k}$$

$$\tilde{R}^{l}_{t'+1,i} \geq \tilde{R}^{l+1}_{t',i} - \tilde{y}^{l}_{t',i} \qquad \forall l \in [L-2] \tag{4.15l}$$

$$\tilde{R}^{l}_{t'+1,i} \leq \tilde{R}^{l+1}_{t',i} - \tilde{y}^{l}_{t',i} + M\tilde{b}^{3,l}_{t',i} \qquad \forall l \in [L-2] \tag{4.15m}$$

$$\tilde{R}^{L-1}_{t'+1,i} \geq \tilde{x}_{t',i} - \tilde{y}^{L-1}_{t',i} \tag{4.15n}$$

$$\tilde{R}^{L-1}_{t'+1,i} \leq \tilde{x}_{t',i} - \tilde{y}^{L-1}_{t',i} + Mb^{3,L-1}_{t',i} \tag{4.15o}$$

$$\tilde{R}^{l}_{t'+1,i} \leq M\left(1 - \tilde{b}^{3,l}_{t',i}\right) \qquad \forall l \in [L-1] \tag{4.15p}$$

$$\tilde{x}_{t',i},\, \tilde{o}_{t',i},\, \tilde{u}_{t',i},\, \tilde{v}_{t',i} \geq 0 \tag{4.15q}$$

$$\tilde{R}^{l}_{t',i},\, \tilde{y}^{l}_{t',i} \geq 0 \qquad \forall l \in [L-1] \tag{4.15r}$$

$$\tilde{b}^{1}_{t',i} \in \{0,1\} \tag{4.15s}$$

$$\tilde{b}^{2,l}_{t',i},\, \tilde{b}^{3,l}_{t',i} \in \{0,1\} \qquad \forall l \in [L-1] \tag{4.15t}$$

Constraints (4.15b) and (4.15c) in the two-stage programming formulation are referred to as *nonanticipativity* constraints. These constraints address an important issue: the inclusion of $\tilde{x}_{t,i}$ for all possible future scenarios. However, this contradicts the definition as it would imply that the first-stage variables have access to future information and have different values for each scenario. Our objective is to find a single optimal value for the decision variable of the current period without knowledge of future demands. To ensure a unique optimal solution for $x_{t,i}$, we introduce nonanticipativity constraints. These constraints guarantee that all decision variables for the current period have the same value among all future scenarios.

Equation (4.15a) and Equations (4.15d)-(4.15t) are linearized forms of the objective function and constraints from the problem in (4.14). To linearize the term $y = \max\{0, x\}$, we introduce three additional constraints: $y \geq x$, $y \leq x + Mb$, and $y \leq M(1-b)$, where x is assumed to be non-negative, b is a binary variable, and M is a large constant. Regarding the nonlinear term inside the objective function, we define auxiliary variables, specifically $u_{t,i} = (\sum_{l=1}^{L-1} R_{t,i}^{l} + x_{t,i} - D_{t,i})^+$, $v_{t,i} = (D_{t,i} - \sum_{l=1}^{L-1} R_{t,i}^{l} - x_{t,i})^+$, and $o_{t,i} = (R_{t,i}^{1} - D_{t,i})^+$. However, since these terms are present in the objective function and are meant to be minimized, we simplify the linearization process by utilizing only the lower bound constraint ($y \geq x$). Therefore, including the constraints with big-M in the linearization of the objective function is unnecessary.

To sum up, in this chapter, we have formulated the stochastic PIRP based on the groundwork laid in Chapter 3. Additionally, we have explored policy solutions, including DP and ADP methods, leading to the proposal of our novel DLA policy. With this foundation, we will empirically evaluate our proposed policy's effectiveness in Chapter 6. Before that, we want to fix the implementation framework for it and offer some improvement techniques in the next chapter.

Chapter 5

Implementation Techniques

In this chapter, we analyze how the proposed algorithm can be efficiently implemented. Our goal is to shed light on the practical implementation of the algorithm and suggest viable ways to enhance its performance and runtime. We will illustrate the effectiveness of these approaches through numerical experiments in the next chapter.

5.1 Tuning the Number of Sample Paths

As expected, increasing the number of sample paths will yield improved results, but it comes at the cost of longer runtime. Inherently, there is a trade-off between the size of the approximated sample space and runtime. An important challenge is determining an appropriate number of sample paths that strikes a balance between these factors.

Scenario-based optimization approaches require a large number of sample paths to achieve high-quality solutions, which can be computationally intensive. Many studies in the literature have adopted greedy decision-making approaches that consider only selected sets of scenarios, which may overlook the true impacts of future events (see e.g. Bertazzi et al. (2015), Solyalı and Süral (2017), Hamdan and Diabat (2019)). However, recent papers have focused on developing techniques that can achieve the desired solution while using a smaller number of scenarios. Two popular techniques in this regard are Lagrangian relaxation and progressive hedging. For instance, Alvarez et al. (2021) have proposed a two-stage stochastic programming formulation using the progressive hedging algorithm to address an IRP with uncertain demand and supply. Similarly, as reviewed in Chapter 2, Dillon et al. (2023) have applied the progressive hedging algorithm to solve their two-stage stochastic programming model for a PIRP, where the first stage involves determining a UL value. Nguyen and Chen (2022) have accounted for stochastic supply and demand but in a PIP and developed an algorithm based on Lagrangian relaxation to obtain near-optimal UL levels. It is worth highlighting that the techniques mentioned above make decisions for the entire planning horizon in a single step.

In contrast, our proposed policy adopts a different approach by solving the model in each period. This eliminates the need for a large approximated sample space. Instead, we consider distinct scenarios each day with the true knowledge of the system state on that day. By making decisions on a daily basis, we ensure reliability and accuracy that may be compromised by making decisions for the entire horizon at once. Our findings suggest that a small number of scenarios is sufficient to achieve high-quality solutions. Furthermore, we have discovered that the size of the sample space can be determined by fine-tuning it in the simulation laboratory for each specific problem instance.

5.2 Decomposition and Parallel Processing

The model represented by Equation (4.15) can be classified as a large MIP due to the presence of a set of variables for each retailer. Besides, there are no constraints that directly link the decisions of one retailer to those of another retailer. This arises from assuming an unlimited supply discussed in Section 4.1, which enables the decision maker to place an order for a retailer without considering the needs of other retailers. Thus, we can decompose the MIP into N sub-problems and solve them independently.

This decomposition reduces the run time by enabling us to use parallel processing. That means we can determine the values of decision variables concurrently, resulting in a modification of Algorithm 2 to Algorithm 3. **Step 1** in Algorithm 2 is now expanded to two steps in Algorithm 3. In this modified algorithm, **Step 1** involves assigning each sub-problem to a separate core within a multi-core machine. State-of-the-art optimizers like Gurobi are then called to solve each sub-problem concurrently. In **Step 2**, Algorithm 3 waits for all sub-problems to be solved and combines their optimal decisions into a decision vector $x_t = (x_{t,i})_{i \in \mathcal{N}^0}$. The last three steps of the algorithm remain the same.

Algorithm 3 Parallel DLA Policy

Step 1: Send each sub-problem i to an idle core, solve it, and keep the optimal value of $x_{t,i}$

Step 2: Wait for all sub-problems to be solved in parallel, and assemble the vector $x_t = (x_{t,i})_{i \in \mathcal{N}^0}$

Step 3: Solve a VRP

Step 4: Observe the realized demand, $d_t = (d_{t,i})_{i \in \mathcal{N}^0}$, and compute the actual cost via Equation (4.1a)

Step 5: Update the inventory of each retailer via Equations (4.1b)-(4.1c)

To make the notation less unwieldy, we will exclude the index i when referring to the MIP model in each sub-problem going forward, even though it is still implicitly considered.

5.3 Linear Cuts

In the context of MIP, a practical approach for enhancing performance is to introduce valid linear cuts. These inequalities can help to tighten the MIP formulation in (4.15) by remaining a smaller part of the feasible set while preserving the optimal solution. First, we propose three linear cuts that are tailored to our specific problem structure.

Proposition 1. *The following linear cuts for all $t' \in [t, T]$ and $\tilde{\omega}_{t,i} \in \tilde{\Omega}_{t,i}$ are valid for the MIP formulation in* (4.15)*:*

$$\tilde{b}_{t'}^{2,l} \leq \tilde{b}_{t'}^{2,l+1} \qquad \forall l \in [L-1],\ \forall t' \in [t, T] \tag{5.1}$$

$$\tilde{b}_{t'}^{1} + \tilde{b}_{t'}^{3,L-1} \geq 1 \qquad \forall t' \in [t, T] \tag{5.2}$$

$$\tilde{b}_{t'}^{1} + \tilde{b}_{t'}^{2,L-1} + \tilde{b}_{t'}^{3,L-1} \leq 2 \qquad \forall t' \in [t, T] \tag{5.3}$$

The linear cut in Equation (5.1) ensures that after satisfying the demand at inventory level l, fresher products in higher inventory levels are not considered, and they will be carried over to the next period without resizing. This chain of linear cuts effectively accelerates the implementation of the FIFO policy.

The second cut in Equation (5.2) establishes a connection between $\tilde{b}_{t'}^{1}$ and $\tilde{b}_{t'}^{3,L-1}$ by analyzing the constraints in the MIP formulation (4.15). This linear cut ensures that if no order is placed, the inventory with a remaining lifespan of $L-1$ will be empty in the following period, leading to a more efficient representation of system dynamics.

By analyzing the constraints in the MIP formulation for $l = L-1$, we establish a relationship between certain binary variables, as shown in Equation (5.3). By incorporating additional constraints, we can eliminate certain permutations of these binary variables that are never valid. This linear cut significantly improves efficiency.

Moreover, by analyzing individual sample paths, we can consider the MIP formulation as a deterministic version of the PIRP, similar to the one in Le et al. (2013), where stock-outs are not allowed. Consequently, for any period of length L, at least one order should be placed to meet the demand during that range. This implies that if, for instance, an order is placed today and received immediately, it can cover today's demand and the demand for the subsequent $L-1$ days. However, it is important to note that this is not valid for the final $L-1$ days of the planning horizon. Therefore, the following inequality can be added to the set of linear cuts:

$$\sum_{\tau=t'}^{t'+L-1} \tilde{b}^1_\tau \geq 1 \quad \forall t' \in [t+1, T-L+1]. \tag{5.4}$$

It should be noted that we should exclude $\tilde{b}^1_t$ since it gets the same single value among all sample paths (see Constraint (4.15c)) and could potentially violate the linear cut of other sample paths. Therefore, we write the linear cut in Equation (5.4) starting from $t+1$.

In addition, we generate the following two sets of valid inequalities for the PIRP formulation, which draw inspiration from the lot sizing problem with fixed costs and inventory bounds investigated by Atamtürk and Küçükyavuz (2005).

Proposition 2. *The following linear cut for all $t' \in [t+1, T]$ and $\tilde{\omega}_t \in \tilde{\Omega}_t$ are valid for the MIP formulation in (4.15), where $\tau = \min\{t'+L-1, T\}$, $\tilde{ub}_{t'} = \min\{R^{max},\ \tilde{D}_{[t',\tau]}\}$, and $S \subseteq [t', \tau]$:*

$$\sum_{t'' \in S} \tilde{x}_{t''} \leq \sum_{t'' \in S} \tilde{D}_{[t'',\tau]} \tilde{b}^1_{t''} + \sum_{l=1}^{L-1} \tilde{R}^l_{\tau+1} \tag{5.5}$$

$$\sum_{l=1}^{L-1} \tilde{R}^l_{t'} + \sum_{t'' \in S} \tilde{x}_{t''} \leq \tilde{ub}_{t'} + \sum_{t'' \in S} \min\{\tilde{ub}_{t''+1} + \tilde{D}_{[t',t'']} - \tilde{ub}_{t'},\ \tilde{D}_{[t',\tau]} - \tilde{ub}_{t'},\ D_{[t'',\tau]}\} \tilde{b}^1_{t''} + \sum_{l=1}^{L-1} \tilde{R}^l_{\tau+1} \tag{5.6}$$

The linear cuts in Equations (5.5) and (5.6) are introduced by Atamtürk and Küçükyavuz (2005) as uncapacitated and capacitated inequalities for blocks of periods $S \subseteq [t', \tau]$. In our problem, τ must be exactly L days after the starting period t', if feasible, or the last day of the horizon, for all $t' \in [t+1, T]$. Upon further analysis, we discovered a means to streamline these linear cuts while preserving their effectiveness. This leads to the simplified version of these cuts presented in the following corollary.

Corollary 1. *The following linear cut for all $t' \in [t+1, T]$ and $\tilde{\omega}_t \in \tilde{\Omega}_t$ are valid for the MIP formulation in (4.15), where $\tau = \min\{t' + L - 1, T\}$ and $\tilde{ub}_{t'} = \min\{R^{max}, \tilde{D}_{[t',\tau]}\}$:*

$$\tilde{x}_{t'} \leq \tilde{D}_{[t',\tau]} \tilde{b}^1_{t'} + \sum_{l=1}^{L-1} \tilde{R}^l_{\tau+1} \tag{5.7}$$

$$\sum_{l=1}^{L-1} \tilde{R}^l_{t'} + \tilde{x}_{t'} \leq \tilde{ub}_{t'} + \min\{\tilde{ub}_{t'+1} + \tilde{D}_{t'} - \tilde{ub}_{t'},\ \tilde{D}_{[t',\tau]} - \tilde{ub}_{t'}\} \tilde{b}^1_{t'} + \sum_{l=1}^{L-1} \tilde{R}^l_{\tau+1} \tag{5.8}$$

Detailed proofs of the propositions and the corollary can be found in Appendix.

5.4 Specially Ordered Sets Constraints, and Optimizer's Parameters

Special Ordered Sets (SOS) constraints are a particular type of linear constraint that restrict the number of variables that can have non-zero values. These limitations can be used to simplify optimization problems, allowing the optimizer to focus on the most promising areas of the feasible region. By incorporating SOS constraints, optimization problems can be solved more efficiently, resulting in faster and more effective solutions. SOS1 and SOS2 are two different types of SOS constraints. At most, one variable out of a possible set of variables can have a value other than zero in SOS1. In SOS2, a set of variables is provided where no more than two adjacent variables may have non-zero values.

A linear constraint in the form of $y \leq M(1-b)$ can be considered a SOS1 constraint, where b is a binary variable. Through this constraint, there are three possible outcomes of $(y = 0, b = 0)$, $(y \neq 0, b = 0)$, and $(y = 0, b = 1)$. This suggests that y and b can form a set of variables in which, at most, one of them can have a non-zero value. This aligns with the definition of a SOS1 constraint. Accordingly, Constraints (4.15k) and (4.15p) can be considered as SOS1 constraints. Thus, we rewrite these constraints in a SOS1 form and use the Gurobi optimizer to handle it, as it is capable of solving a problem with SOS constraints. It is important to highlight that by taking this step, we can gain an extra benefit for our MIP, which is eliminating any negative impact that the value of the big-M may have on the problem through these two constraints.

We can further enhance the performance of the algorithm by modifying specific parameters of the optimizer. By default, the Gurobi MIP solver balances the search for new feasible solutions and proving that the current solution is optimal. Nonetheless, we have noticed that our model could easily discover high-quality solutions; therefore, we have instructed the solver to focus more on proving optimality by configuring $MIPFocus = 2$. It should be noted that the proper value for this parameter is highly dependent on the problem structure, including size, data set, and objective. Thus, modifying this parameter might reduce runtime in some circumstances.

In closing, this chapter has explored the practical implementation techniques for the proposed DLA policy. Our focus was to provide actionable ideas for an efficient implementation process, optimizing performance and runtime. These ideas will be put to the test in numerical experiments in the upcoming chapter, where we aim to validate their impact on the algorithm's effectiveness.

Chapter 6

Numerical Experiments

The goal of this chapter is to demonstrate the reliability and efficiency of the proposed policy based on a case study. We first briefly discuss the case study and the related data set. Afterward, the impact of the implementation techniques presented in Chapter 5 is examined. Then, we investigate the performance of the proposed policy compared to alternative approaches.

For computations, we engaged the Standard-F32s-v2 machine provided by Microsoft Azure Machine Learning Studio, which featured 32 cores, 64 GB RAM, and 256 GB disk. In the PIP phase, we used Gurobi 10.0 for optimization, and in the second phase, we utilized the VRP algorithm supported by Google OR-Tools, along with some modifications. We assume $Q = 120$, and the maximum travel distance is 230.

6.1 Case Study and Data Set

We evaluate the effectiveness of the proposed policy in a BSC tasked with satisfying the demand for platelet units from multiple hospitals. The shelf life of platelet units is seven days. The distribution network for our study centers around the Canadian Blood Services production site, located in Brampton, Ontario, Canada. This production site serves as the primary distribution center responsible for supplying hospitals located within the nearby cities. For our analysis, we designate this production site as the supplier within our distribution network. Furthermore, our study focuses on a chosen group of 20 hospitals that fall within the coverage area of the production site. These hospitals play the role of retailers in our distribution network. The traveling cost of $c_{i,j}$ for all $(i,j) \in \mathcal{E}$ is calculated based on the Euclidean distances.

We carefully selected hospitals for our study based on their likelihood of requiring platelet units. Factors such as specialty and the range of services provided were considered in the selection process. According to "NHLBI, Platelet Disorders, Thrombocytopenia" (n.d.), most platelet recipients are hospitals that primarily treat cancer (chemotherapy), bone marrow disorders, blood disorders (such as aplastic anemia and leukemia), liver disease, and platelet function disorders (such as idiopathic thrombocytopenic purpura). Moreover, massive blood loss due to injury or surgery often requires platelet transfusion.

To analyze the data effectively, we categorize the hospitals into three groups: A, B, and C. The categorization is based on the extent of the services they provide that require platelet units. Hospitals in category A have the highest demand, followed by those in categories B and C. Figure 6.1 illustrates the network of hospitals and their categorization. In this figure, the three categories, A, B, and C, are distinguished by different colors.

To estimate the average daily demand (ADD) of the hospitals, we rely on Mirjalili et

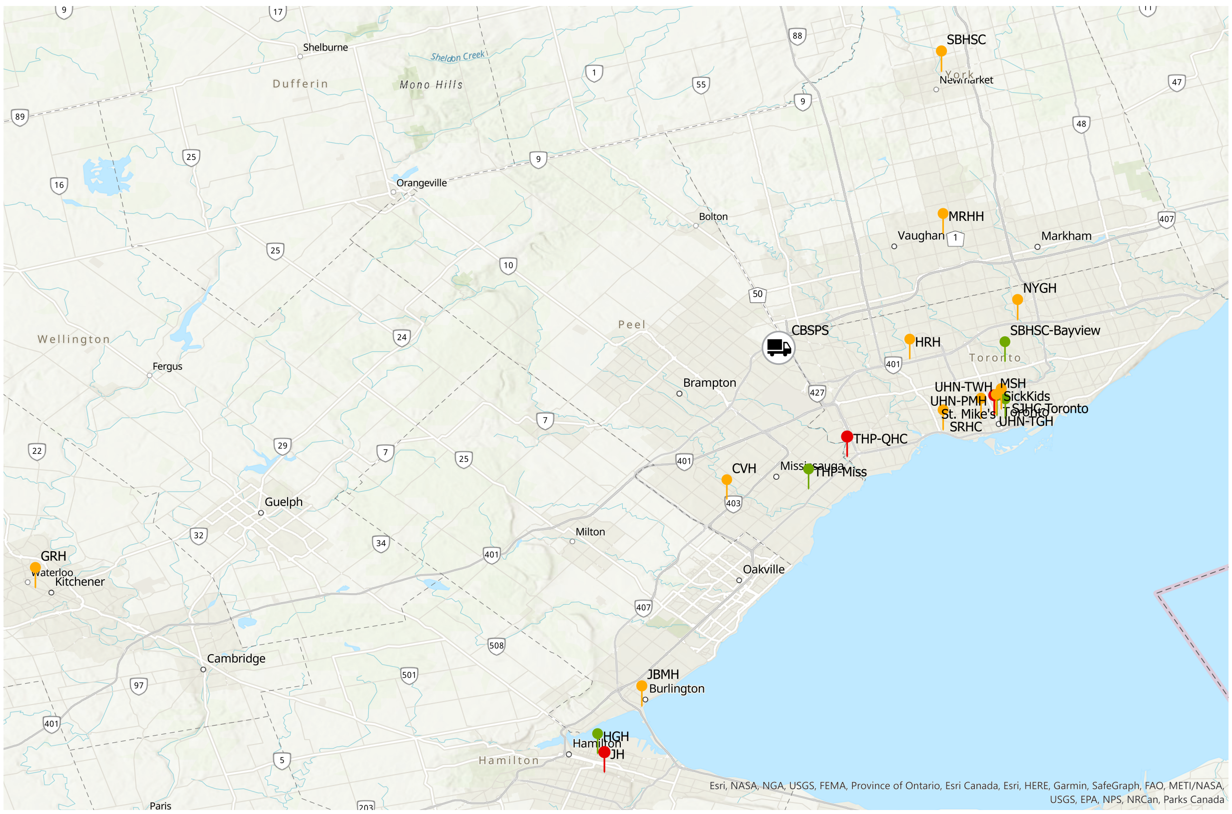


Figure 6.1: The studied network consisting of the Canadian Blood Services production site and 20 hospitals categorized as Ⓐ Ⓑ Ⓒ.

al. (2022) as our primary reference. This study analyzes the ADD and standard deviation (SD) of two hospitals, abbreviated as HGH and JH. According to our categorization scheme, HGH falls into Category A, while JH falls into Category C.

To obtain more precise estimations for the remaining hospitals, we implement additional clustering within each category. In this process, we utilize the number of beds in each hospital as a criterion for categorizing them into the following groups:

Group 1: $0 < \textit{number of beds} < 250$

Group 2: $250 \leq \textit{number of beds} < 500$

Group 3: $500 \leq \textit{number of beds} < 1000$

Group 4: $1000 \leq \textit{number of beds}$

Based on this clustering, both HGH and JH share the same group number, which is 2 in this case. Using the ADD and SD values from HGH and JH as reference points, we estimate the corresponding values for Group 2 within Category B. This estimation assumes that the ADD and SD values of Group 2 in Category B are equal to the average values of those of HGH and JH. Subsequently, we calculate the ADD and SD values for the other groups within each category relative to their group numbers. We present a heat map in Figure 6.2 to visualize the ADD values. The heat map utilizes proportional symbols to relatively represent the ADD values for all hospitals.

For the purpose of the numerical experiments, we require demand distributions from which we can extract samples to generate scenarios. In our case study, we assume that the demand distribution for each retailer follows a normal distribution with known mean and SD. Hence, we model the demand for retailer i as $D_i \sim \text{Norm}(ADD_i, SD_i)$. Regarding

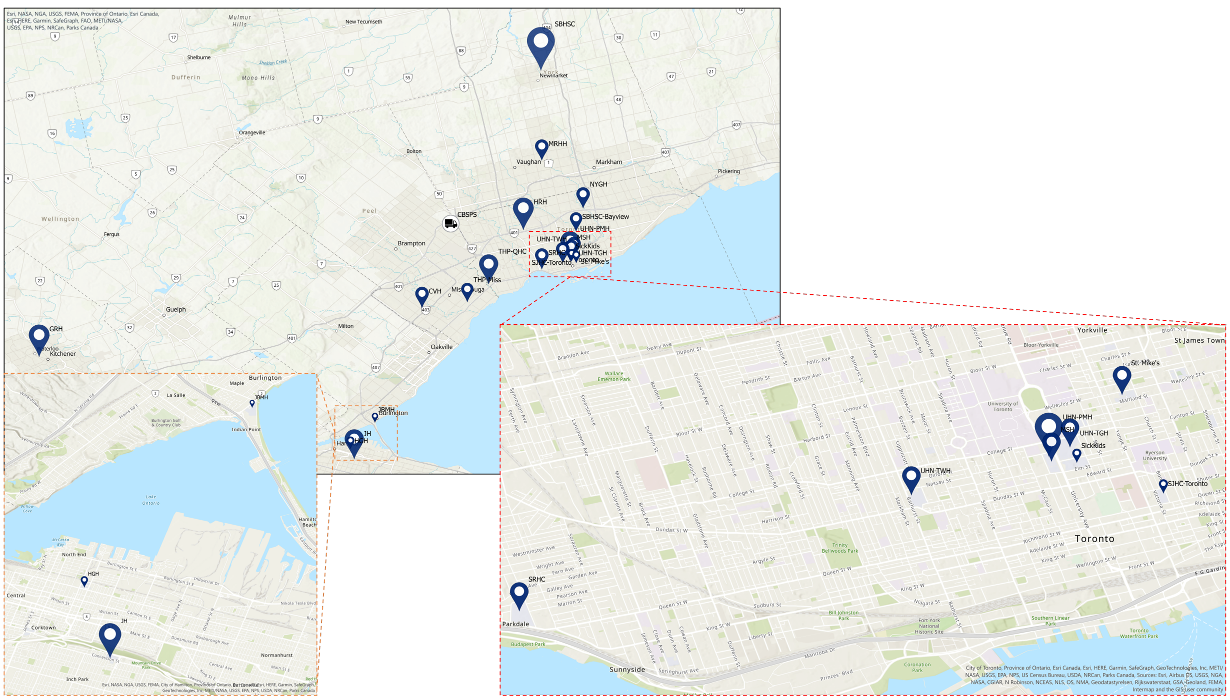


Figure 6.2: Proportional symbol map of ADD in the studied network.

inventory capacities, we set the maximum inventory level for each hospital as $R_i^{max} = L \times ADD_i$. This implies that each hospital is equipped to store units to serve the average daily demand for a period of L days. It is important to note that this value typically exceeds the three-sigma limit of the hospitals in this distribution network, which represents a conservative estimate of the maximum potential demand.

6.2 Benchmarks

We consider the policies proposed by Crama et al. (2018) as our benchmark, since this paper specifically addresses the stochastic PIRP, making it a relevant and comparable reference. Moreover, Crama et al. (2018) address the PIRP using a two-phase approach involving a PIP and a VRP. This is similar to our approach, where we also consider the PIRP in two phases. Hence, we can gain valuable insights and draw meaningful comparisons by examining the performance of the proposed DLA policy with the four policies suggested by Crama et al. (2018).

To solve the PIP phase, they first use a basic policy called EV, in which they ignore the random demand by using its expected value and replenish the inventory accordingly. The second policy is a simple heuristic of the classic up-to-level algorithm (UL) in which S from the (s, S) policy is established to meet a predefined target service level for the λ consecutive periods. Note that S may be referred to as the UL level, and λ is an adjustable parameter. To clarify further, when $\lambda = 1$, the order quantity, in addition to the current inventory, should be enough to meet the demand for the current day with a probability higher than the target service level. This ensures the inventory level stays above a specific threshold to maintain the desired service level for each period. The maximum value for the parameter λ is indeed $\lambda = L$, which corresponds to ordering once every L day while still meeting the target service level. These two instances are shown by UL_1 and UL_L, respectively.

The aforementioned DE policy in Subsection 4.2.1 is the third policy in Crama et al. (2018). This policy is an ADP with a fixed cost as an approximation to the routing cost, which is decomposed into N smaller ADPs, one for a retailer. A VRP algorithm should be subsequently performed in the second phase of these first three policies. Among the policies examined in their study, the fourth policy, denoted as DI, demonstrates the best performance. It is an enhanced version of the DE policy that incorporates a neighborhood search and refines the replenishment decisions if needed. Notably, Crama et al. (2018) claim that these ADPs could also be estimated by a (s, S) policy with insignificant alterations in their outcomes. In our subsequent discussions, when we mention ADPs in Crama et al. (2018) study, we specifically refer to the DE and DI policies. Furthermore, we refer to the policies proposed by this study as the *benchmark policies*.

Crama et al. (2018) demonstrate the effectiveness of their ADPs in numerical examples for shelf life values of $L \in \{2, 3, 4\}$. However, as discussed in Subsection 4.2.1, it is known that ADP methods face the curse of dimensionality when shelf life surpasses a certain threshold. In our experimental setups, we encountered computational difficulties when running the backward DP algorithm for $L = 5$ and beyond. Therefore, to compare the results of these policies and the DLA policy, we initially set the shelf life to $L = 3$. We evaluate and identify the best policies based on this initial setting. Subsequently, we compare the policies, except ADPs, in a dedicated section, maintaining the original setting with a shelf life of $L = 7$. By examining the performance of the policies under both settings, we can provide insights into their effectiveness and abilities to handle a longer shelf life.

To establish a lower bound for these policies, we adopt an offline policy that knows the actual demand. This policy supplies the product for $\lambda \leq L$ periods in the PIP phase, which results in zero wastage and shortage. Crama et al. (2018) call this policy the Full Information policy (FI). It is important to note that the optimal value of λ is problem-dependent for both UL and FI.

Regarding the fixed cost parameter, we utilize the distance-based cost used in Crama et al. (2018) in the DE, DI, and DLA policies. An alternative method called route-based cost is also introduced for determining the values of the fixed cost parameter. However, in their numerical experiments, the route-based cost showed improvement in only a few cases compared to the distance-based cost, despite being more complex. Therefore, for our analysis, we employ the distance-based cost. This choice is based on its overall effectiveness and simplicity compared to the route-based cost used in the Crama et al. (2018) study. The distance-based cost for the fixed cost considered by Crama et al. (2018) is

$$\hat{g}_{t,i} = \frac{\sum_{j \in J_i}}{c_{ij}} |J_i|, \tag{6.1}$$

where J_i refers to a group of retailers located near retailer i, and it is argued that J_i has a connection with the shelf life L. It is suggested that a suitable value for $|J_i|$ is $2L$, which was also employed in our experiments.

Furthermore, Crama et al. (2018) constrain the decision variables to a smaller feasible region by using a lower bound based on the target service level whenever an order needed to be placed in DE and DI. This value is denoted as $x^{(1)}_{t,i}$ and computed using UL_1. To provide a brief overview, the UL policy uses an adjustable parameter λ to determine the order quantity. This quantity is designed to fulfill the demands of λ consecutive days with a probability higher than the target service level. In this context, setting $\lambda = 1$ ensures that the lower bound, represented as $x^{(1)}_{t,i}$, is enough to meet the demand for the current period while achieving the desired service level. To ensure a proper comparison, we also incorporate this lower bound into our DLA policy, leading to adding the following constraint to the MIPs:

$$x_t \geq x_t^{(1)} b_t^1. \tag{6.2}$$

We will later explore how this lower bound affects the overall cost and investigate any potential trade-offs or benefits it may offer. Before doing experiments to compare the policies in the settings of $L = 3$ and $L = 7$, we do some experimental setups to determine the best number of simulations and fix the implementation framework for the DLA policy.

6.3 Experimental Setup

We perform these numerical examples specifically for platelet units that have a shelf life of $L = 7$ days and within a planning horizon of $T = 30$ days. The cost combination of $(h, p, e, a) = (2, 8, 8, 0)$ is chosen as the baseline for our experiments. Recall that (h, p, e, a) represents holding cost, penalty cost, wastage cost, and replenishment cost parameters, respectively. This choice is motivated by the consideration that a ratio of $\frac{p}{e} \geq 1$ is more applicable in the platelet supply chain, and shortages are typically considered to be more costly than wastage (Zhou et al., 2011). By selecting a ratio of 1, we ensure that penalty and wastage costs are treated equally. This approach prevents the model from being biased towards any particular goal and helps maintain a balanced perspective. Furthermore, in the literature of BSC, the holding cost is given the lowest priority if not ignored (see, e.g. Civelek et al. (2015)). Also, since the suppliers in BSC operate as non-profit organizations, no replenishment costs are involved ("About Us - Canadian Blood Services", n.d.).

To compare policy A and policy B, we use the *relative gap* between them, defined as

$$Gap(A, B) = \frac{P_A - P_B}{P_B} \times 100, \tag{6.3}$$

where P_A and P_B are performance measures obtained from policies A and B, respectively.

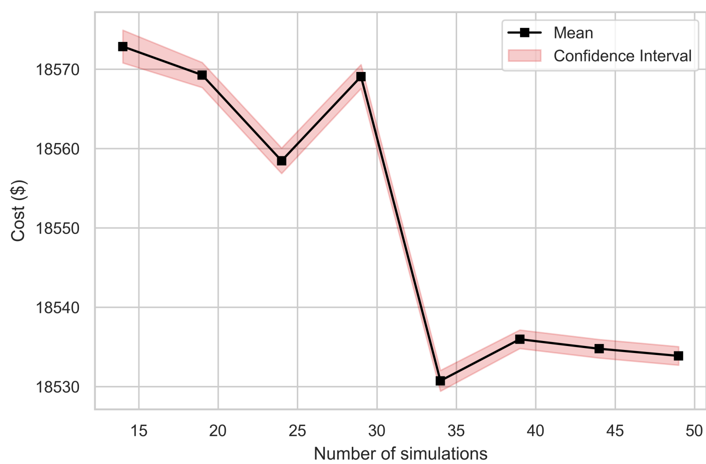
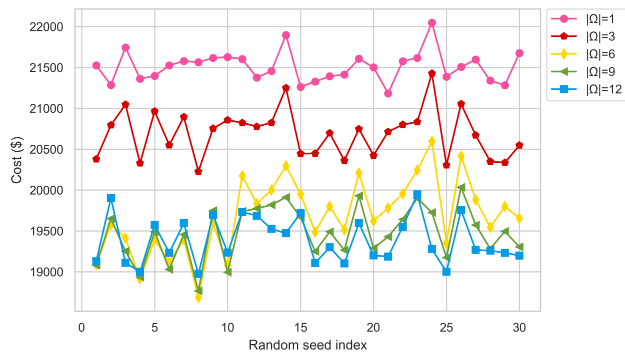


Figure 6.3: The mean and confidence interval of the DLA policy's total cost over different numbers of simulations

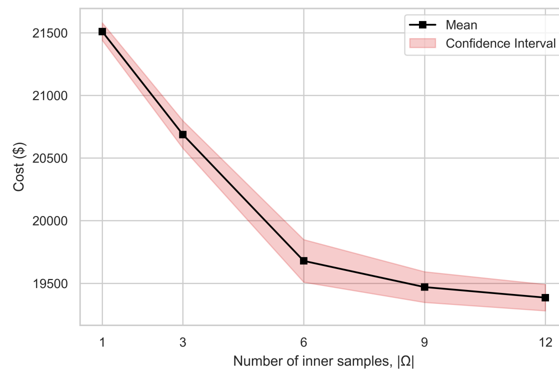
6.3.1 Determining the Number of Simulations

By running the policies multiple times and sampling the values of the random demand, we simulate the system's behaviour under various cases. Subsequently, we evaluate the performance of the policies by averaging the outcomes across the simulations, thereby obtaining a reliable estimate of their overall performance.

In this subsection, we aim to find the best number of simulations that are sufficient to evaluate the DLA policy. We ascertain the ideal number of simulations by gradually increasing the initial value from 15 to 50 simulations. Then we plot the mean and confidence interval of the total costs over each number of simulations in Figure 6.3. It can be observed that conducting 40 simulations is adequate to ensure sufficient reliability for all experiments. Further increasing the number of simulations beyond this point does not significantly change the mean and confidence interval of the total cost obtained from the DLA policy.



(a) The average total cost of the *DLA* policy over 40 simulations, through 30 different random seeds for generating the inner sample paths



(b) The mean and confidence interval of the average values in (a)

Figure 6.4: Tuning the Number of Sample Paths

6.3.2 Tuning the Number of Sample Paths (Scenarios)

The most important experimental setup is adjusting the number of sample paths or scenarios within the *DLA* policy. Specifically, we investigate how the average total cost varies for different values of $|\tilde{\Omega}_t|$ chosen from the set $\{1, 3, 6, 9, 12\}$. For each value, we conduct the experiment using 30 distinct sets of $|\tilde{\Omega}_t|$ number of sample paths. The goal is to determine whether obtaining the outcomes with a specific sample path is purely coincidental or if we can achieve similar outcomes with various sets of sample paths. Subsequently, we plot the mean and confidence intervals of these 30 distinct sets to visualize the degree of variation when employing different sets of random sample paths. Since these goals necessitate a large number of executions, we limit the implementation of the *DLA* policy to $L = 3$. However, we expect the same behaviour in the setting of $L = 7$.

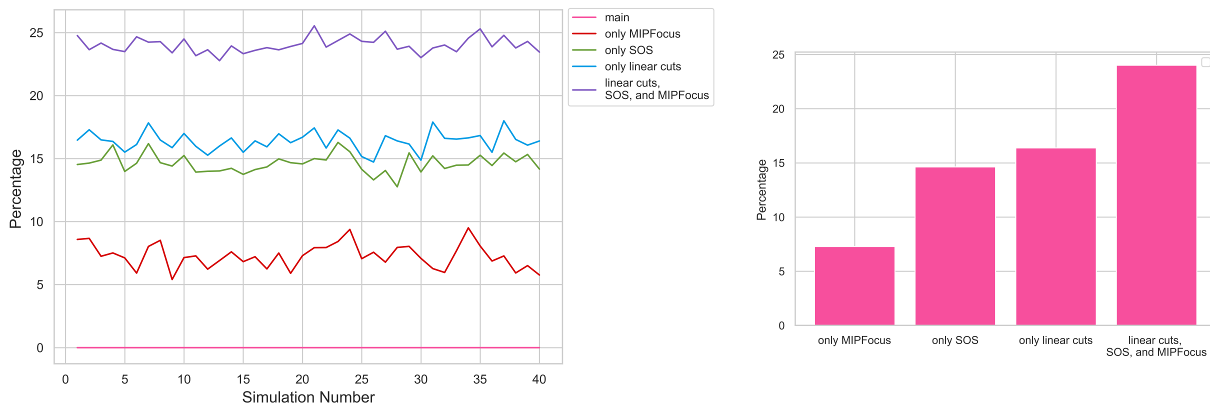
In Figure 6.4a, the x-axis represents different random seeds used for generating inner scenarios. Each point represents the average total cost of the *DLA* policy over 40 simulations. To generate the data points shown in Figure 6.4b, we calculated the mean and confidence interval of the average values obtained from Figure 6.4a. For example, the final data point in Figure 6.4b, corresponding to $|\tilde{\Omega}_t| = 12$, is the mean of 30 average values. Each average value is computed as the average of the total cost associated with the DLA policy within which a distinct set of 12 scenarios is used. In simpler terms, this point represents the average value for the line plot corresponding to $|\tilde{\Omega}_t| = 12$ in Figure 6.4a.

Based on Figure 6.4b, we can draw two conclusions. First, there is minimal variation among different sets of sample paths used by the DLA policy. In other words, any sets of sample paths yield close results as long as they contain a similar number of paths. This is evident from the small confidence intervals. Second, we observe that increasing the number of inner sample paths improves the performance of the policy, as it is approved by the decreasing trend. Based on this extensive experiment, using any 9 sample paths is sufficient for our specific case study. This is because, beyond that point, the mean and confidence interval show a minor decrease while the runtime increases exponentially.

It should be noted that the number of sample paths chosen for our experiments may be considered small. Nonetheless, the results presented in the subsequent sections demonstrate that the DLA policy can yield cost-reduction outcomes even with this limited number of sample paths. This suggests that the DLA policy effectively leverages the available information and makes effective decisions, even when working with a relatively small number of scenarios.

6.3.3 Implementation Techniques

In Chapter 5, we presented a set of implementation techniques aimed at enhancing the execution of the MIP formulation in the DLA policy. These techniques include decompo-



(a) The runtime improvement in each simulation

(b) The mean of runtime improvement over 40 simulations

Figure 6.5: The runtime improvement of the DLA policy by applying implementation techniques

sition and parallel processing, linear cuts, SOS constraints, and optimizer modification. To evaluate the impact of each of these techniques on the policy's runtime, we do experiments focusing on the baseline case where $(h, p, e, a) = (2, 8, 8, 0)$ for $L = 7$.

First, we conduct a runtime comparison between the original DLA policy (Algorithm 2) and the parallel DLA policy (Algorithm 3). The parallel DLA policy assigns one core of the machine to each hospital's MIP independently. Our findings reveal a remarkable reduction in the runtime when employing parallelism. Specifically, the parallel DLA policy demonstrates a runtime improvement of 92.03% compared to the original version. This translates to a speedup of approximately 7.97 times, indicating the effectiveness of Algorithm 3 in accelerating the computation process.

Figure 6.5 illustrates the runtime improvement achieved by implementing other techniques in our model, including the addition of linear cuts, the substitution of SOS constraints, the adjustment of the MIPFocus parameter, and a combination of all these techniques. On average, we observe a further runtime improvement of up to 25% by applying these implementation techniques. It is important to note that when analyzing the effect of adding linear cuts to the model, we disable the automatic cut generation feature of Gurobi by setting the cut parameter to zero, i.e., $Cuts = 0$.

6.4 The Setting of $L = 3$

Figure 6.6 displays the relative gap between the average total cost of each policy Π and FI over 40 simulations, for various combinations of cost parameters. The gap is computed based on $Gap(\Pi, FI)$ as defined in Equation (6.3). Recall that the UL_1 policy aims to place an order that makes the total inventory to be sufficient to meet the target service level for the current day. On the other hand, the UL_L policy aims to do this not only for today but also for the next $L - 1$ days.

It is observed that in this PIRP, the UL_1 policy performs better than the UL_L policy. This is due to the fact that the UL_1 focuses on replenishing more frequently with smaller quantities, which helps reduce holding, penalty, and wastage costs, although it increases routing costs. However, in the problem setting studied in Crama et al. (2018), the opposite was true since their objective function did not consider any of these costs, and routing cost played the main role. It is important to note that the difference between these two UL policies is the smallest when the holding cost is zero.

Figure 6.6 reveals an unexpected outcome, with DE and DI having the worst performance, even underperforming compared to EV. This contrasts with the findings reported in Crama et al. (2018). Conversely, the DLA policy exhibits the opposite behaviour, performing remarkably well in our PIRP. This suggests that (s, S) policies are not viable

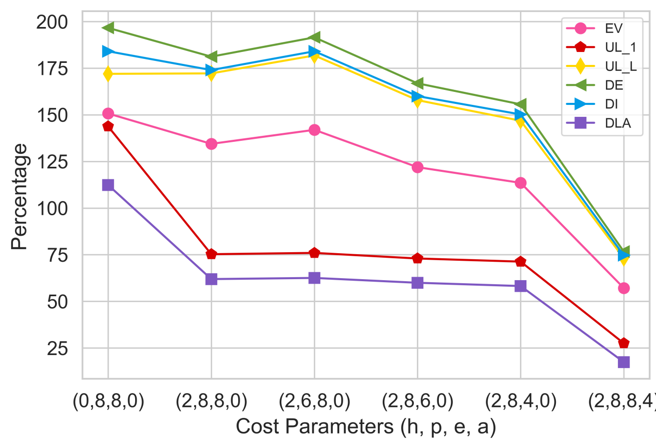


Figure 6.6: The relative gap between the average total cost of each policy Π and FI over 40 simulations in the setting of $L = 3$, i.e. $Gap(\Pi, FI)$, for various combinations of cost parameters

options when dealing with holding, penalty, and wastage costs along with the routing costs. Among these policies, only UL shows relative reliability due to its adjustable parameter λ, which ensures that the order quantity is sufficient to meet the demand for λ consecutive days with a higher probability than the target service level. Overall, we find that the algorithms studied in Crama et al. (2018) have a notable weakness in not considering standard inventory expenses. In contrast, the DLA policy consistently produces favourable outcomes in these scenarios.

Table 6.1 displays the average values of various performance indicators given in the Crama et al. (2018) study. The first indicator assesses the *Actual Service Level* and is defined by Equation (6.4a). In this equation, the variable ξ_L represents the average fill rate of demand across all hospitals over the planning horizon when the shelf-life is L. The second indicator pertains to the evaluation of the *Actual Freshness* and is expressed through Equation (6.4b) as ϕ_L, indicating the average actual freshness from the standpoint of patients. The term $(r^l)_{ti}$ denotes the number of units with a remaining shelf-life of l transferred to a patient in period t at hospital i. Lastly, Table 6.1 reveals the number of outdated products within the observations.

$$\xi_L = \frac{\sum_t \sum_i \min\{d_{ti}, \sum_{l=1}^{L-1} R_{ti}^l + x_{ti}\}}{\sum_t \sum_i d_{ti}} \tag{6.4a}$$

$$\phi_L = \frac{\sum_t \sum_i \{\sum_{l=1}^{L-1} l r_{ti}^l + L r_{ti}^L\}}{\sum_t \sum_i \sum_{l=1}^{L} r_{ti}^l} \tag{6.4b}$$

The results presented in Table 6.1a provide insights into the performance of different policies in terms of actual service levels. The DLA policy shows a relatively lower actual service level compared to the other policies; nonetheless, it consistently maintains an average above 90%. On the other hand, Table 6.1b demonstrates that the DLA policy excels in achieving the highest actual freshness level among all the other policies. Moreover, examining Table 6.1c, we observe that the DLA policy lead to the least amount of wastage. Conversely, the ADPs deliver higher service levels but result in a considerably larger number of outdated products. Overall, these findings highlight a trade-off between service level and product wastage, with the DLA policy achieving a better balance in minimizing wastage while maintaining a slightly lower actual service level.

6.5 The Setting of $L = 7$

We repeat the experiments with $L = 7$ to align with the shelf life of platelet units. Our results in the previous section revealed that the DLA policy outperformed all other policies. Nevertheless, we compare the DLA policy to benchmark policies again, except DE and DI, to better understand a longer shelf life's effects. This exclusion is due to the impracticality of solving a backward DP algorithm for a shelf life of this magnitude. The results obtained

(h,p,e,a)	$(0,8,8,0)$	$(2,8,8,0)$	$(2,6,8,0)$	$(2,8,6,0)$	$(2,8,4,0)$	$(2,8,8,4)$
EV	97	97	97	97	97	97
UL_1	95	95	95	95	95	95
UL_L	**99**	**99**	**99**	**99**	**99**	**99**
DE	**99**	**99**	**99**	**99**	**99**	**99**
DI	**99**	**99**	**99**	**99**	**99**	**99**
DLA	96	93	92	93	93	75
FI	100	100	100	100	100	100

(a) $\xi_3(\%)$

(h,p,e,a)	$(0,8,8,0)$	$(2,8,8,0)$	$(2,6,8,0)$	$(2,8,6,0)$	$(2,8,4,0)$	$(2,8,8,4)$
EV	2.1	2.1	2.1	2.1	2.1	2.1
UL_1	**2.4**	2.4	2.4	2.4	2.4	2.4
UL_L	1.9	1.9	1.9	1.9	1.9	1.9
DE	2.1	1.9	1.9	1.9	1.8	1.8
DI	2.1	1.8	1.8	1.8	1.8	1.8
DLA	2.3	**2.5**	**2.5**	**2.5**	**2.5**	**2.5**
FI	3.0	3.0	3.0	3.0	3.0	3.0

(b) ϕ_3

(h,p,e,a)	$(0,8,8,0)$	$(2,8,8,0)$	$(2,6,8,0)$	$(2,8,6,0)$	$(2,8,4,0)$	$(2,8,8,4)$
EV	77	77	77	77	77	77
UL_1	**10**	10	10	10	10	10
UL_L	89	89	89	89	89	89
DE	70	85	85	85	86	89
DI	70	89	89	88	86	89
DLA	21	**9**	**8**	**9**	**9**	**5**
FI	0	0	0	0	0	0

(c) Number of Outdated products

Table 6.1: Results of 40 simulations in the setting of $L = 3$ for various combinations of cost parameters

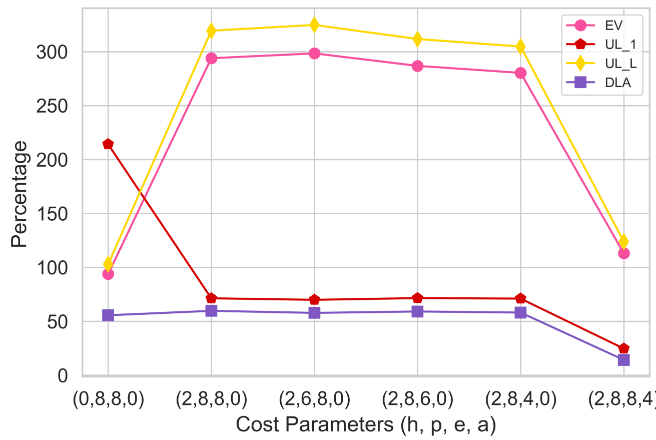


Figure 6.7: The relative gap between the average total cost of each policy Π and FI over 40 simulations in the setting of $L = 7$, i.e. $Gap(\Pi, FI)$, for various combinations of cost parameters

from this set of experiments demonstrate similar outcomes to those observed in the setting of $L = 3$. Specifically, the DLA policy shows superior performance in terms of the relative gap between average total costs, as illustrated in Figure 6.7.

Table 6.2a shows that the DLA policy maintains a lower actual service level than the alternative policies. However, it is important to note that this performance indicator has improved compared to the setting of $L = 3$. In other words, when the shelf life is extended from 3 to 7, the DLA policy performs better in terms of the actual service level, with an average exceeding 91%. Additionally, examining Table 6.2b, we find that the DLA policy consistently achieves the highest levels of freshness for the ordered products. This indicates that the products supplied through the DLA policy are generally fresher compared to those from other policies. Lastly, considering Table 6.2c, it is evident that the DLA policy incurs almost zero wastage, indicating its efficient utilization of perishable products.

(h,p,e,a)	$(0,8,8,0)$	$(2,8,8,0)$	$(2,6,8,0)$	$(2,8,6,0)$	$(2,8,4,0)$	$(2,8,8,4)$
EV	**99**	**99**	**99**	**99**	**99**	**99**
UL_1	93	93	93	93	93	93
UL_L	**99**	**99**	**99**	**99**	**99**	**99**
DLA	**99**	91	91	91	91	84
FI	100	100	100	100	100	100

(a) $\xi_7(\%)$

(h,p,e,a)	$(0,8,8,0)$	$(2,8,8,0)$	$(2,6,8,0)$	$(2,8,6,0)$	$(2,8,4,0)$	$(2,8,8,4)$
EV	4.3	4.3	4.3	4.3	4.3	4.3
UL_1	**6.4**	**6.4**	6.4	**6.4**	**6.4**	6.4
UL_L	4.1	4.1	4.1	4.1	4.1	4.1
DLA	4.6	**6.4**	**6.5**	**6.4**	**6.4**	**6.5**
FI	7.0	7.0	7.0	7.0	7.0	7.0

(b) ϕ_7

(h,p,e,a)	$(0,8,8,0)$	$(2,8,8,0)$	$(2,6,8,0)$	$(2,8,6,0)$	$(2,8,4,0)$	$(2,8,8,4)$
EV	44	44	44	44	44	44
UL_1	**0**	**0**	**0**	**0**	**0**	**0**
UL_L	48	48	48	48	48	48
DLA	9	**0**	**0**	**0**	**0**	**0**
FI	0	0	0	0	0	0

(c) Number of Outdated products

Table 6.2: Results of 40 simulations in the setting of $L = 7$ for various combinations of cost parameters

6.6 Comparison with the Best Benchmark Policy

We obtain the *relative cost improvement* of the DLA policy to the best benchmark policy, denoted as $-Gap(DLA, \Pi_{best})$ via Equation (6.3). In the initial setting of $L = 3$ in Section 6.4, the best benchmark policy was consistently found to be UL_1 across all combinations of cost parameters. On average, the DLA policy approach demonstrated a cost improvement of 8.5% compared to UL_1, with an SD of 0.89%, based on only 9 scenarios used within the DLA policy framework.

In the setting of $L = 7$ in Section 6.5, the DLA policy outperformed the best-performing policy, Π_{best}, with an average relative cost improvement of 10.0% and an average SD of 0.66%. When the holding cost is eliminated, the optimal policy UL_L achieves the highest relative cost improvement of 23.1%. However, in all other cases, UL_1 remains the best-performing benchmark policy. Overall, the best-performing policy consistently appears to be the UL policy. Therefore, we may refer to the best benchmark policy as the best UL policy.

Figure 6.8 illustrates the results. It is evident that the average of the relative cost improvements for $L = 7$ surpassed those for $L = 3$, accompanied by a lower SD. This observation suggests that DLA exhibits more cost-reduction outcomes as the shelf life of the product increases.

6.7 Sensitivity Analysis of the Cost Parameters

In this section, we perform a sensitivity analysis to examine the impact of small changes in each cost parameter on the total cost and the order quantity. Our analysis evaluates

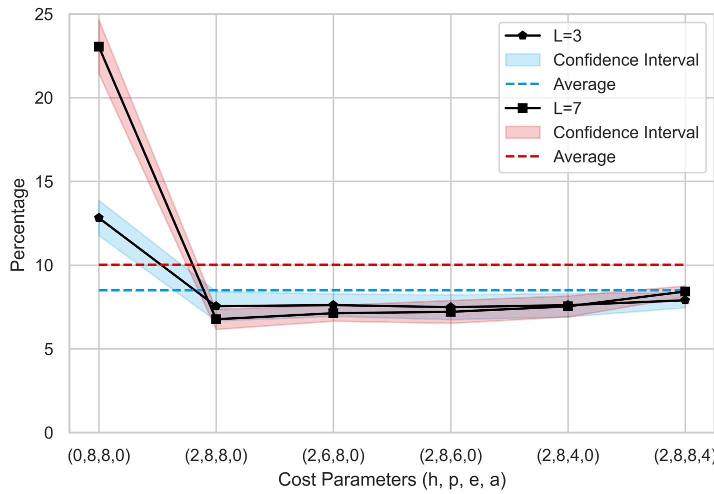
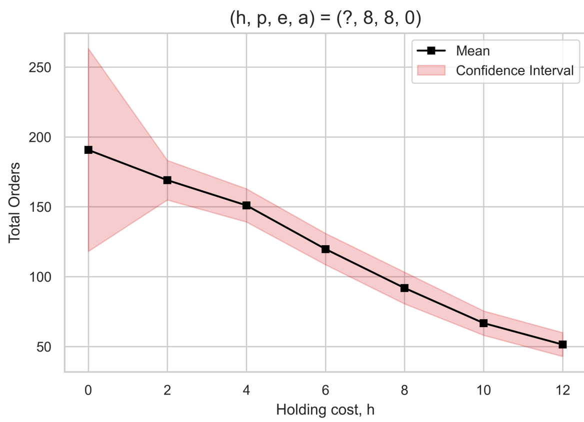


Figure 6.8: The mean of cost improvements of the DLA policy relative to Π_{best}, i.e. $-GAP(DLA, \Pi_{best})$, in the setting of $L = 3$ and $L = 7$

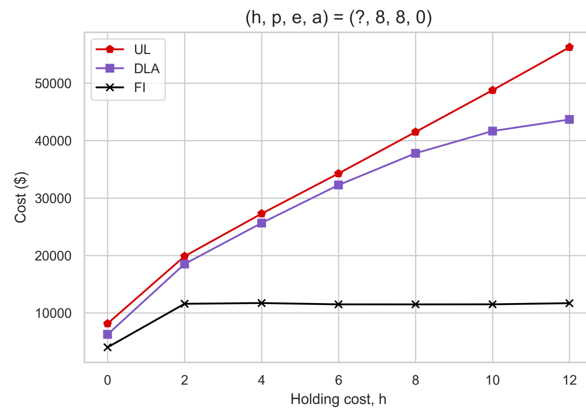
the sensitivity of the total cost and the ordering decisions to variations in the holding cost, penalty cost, wastage cost, and replenishment cost parameters.

6.7.1 Holding Cost

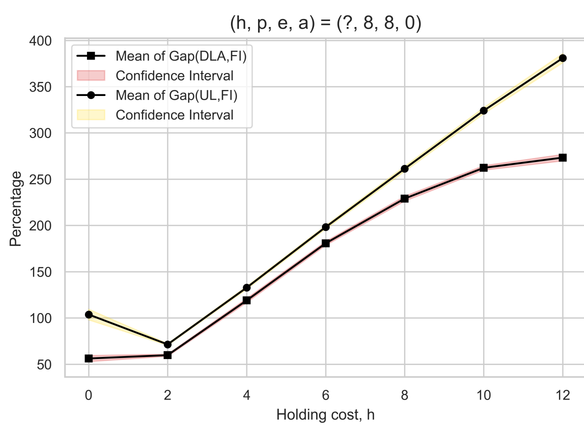
Figure 6.9a shows that when the holding cost, h, is set to zero, the DLA policy adopts a strategy of maximum ordering. The policy focuses on minimizing penalty and fixed routing costs by ordering as many products as possible. However, ordering a large number of products leads to high wastage. This is because a significant number of items are held in inventory, and over time, some may perish or expire. Nevertheless, it is important to note that the wastage is considerably lower than the unsatisfied demand. This aligns with the earlier understanding that shortages are generally considered more costly than wastage (Zhou et al., 2011). Therefore, while wastage cost is a concern, the impact of unsatisfied



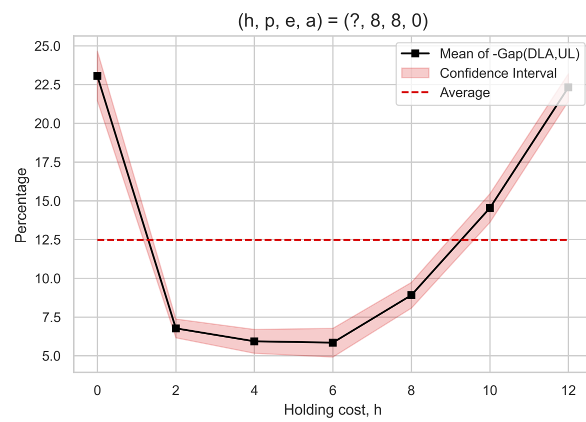
(a) The average total cost



(b) The average total orders by the DLA policy



(c) The gap with FI



(d) The cost improvement of the DLA policy relative to the best UL

Figure 6.9: Sensitivity Analysis of the Holding Cost

demand is typically more significant in the DLA policy's objective. This is why the highest orders are placed at this point compared to other cases of h.

It is important to note that the routing cost may be higher for larger order quantities. This occurs in specific cases. For example, when setting the holding cost to zero changes the decision from not placing an order to ordering a positive quantity for a hospital. Another case is when increasing the order quantities results in the need for more vehicles, leading to an increase in routing costs. The extent of this impact on routing costs depends on the sensitivity of routing costs to the vehicle capacity. Thus, the routing cost is the main portion of the total cost at $h = 0$ in Figure 6.9b.

As the holding cost increases, the total cost also increases due to the addition of a new cost term per product unit. This leads to a decrease in the order quantity, as shown in Figure 6.9a. Consequently, a larger proportion of demand will be lost after realization. This results in a shift in the total cost, where the penalty cost becomes more substantial than the holding cost, despite an expected increase in the holding cost due to the higher value of the parameter h.

Once the holding cost surpasses or equals the other cost parameters, which is $h \geq 8$ in our case, the DLA policy adjusts its ordering strategy accordingly. It tends to place more frequent orders, potentially on a daily basis, to meet the demand for each specific day while holding fewer units, as depicted in Figure 6.9a. The DLA policy's preference for more frequent orders significantly reduces the wastage cost. Because ordering fewer units at a time and receiving more daily fresh products minimizes the risk of excess inventory that may expire. Moreover, although the frequent ordering approach in the DLA policy may significantly increase the fixed routing cost, the actual rise in routing cost is usually milder. This is because the smaller number of ordered units requires fewer vehicles for transportation than in previous cases. This observation is evident in Figure 6.9b, where the proportional increase in the total cost becomes more gradual once the holding cost parameter reaches or exceeds a certain threshold, in this case, $h \geq 8$.

In Figure 6.9c, the behaviour of FI can be explained as follows. Based on the FI policy, the delivery quantity is precisely equal to the actual demand of all hospitals each day, resulting in no wastage or penalty costs. Therefore, when h is equal to zero, the only concern is the routing cost. In this case, ordering once every L day is preferred as it minimizes the routing cost. Ordering every L day represents the longest period in which the policy can fulfill the demand with a single batch of vehicles. Interestingly, even though larger order quantities might be expected to increase the number of required vehicles, it appears that the impact of less frequent transportation has a greater effect on reducing routing costs in this specific case study. This suggests that the vehicles' capacity may have minimal influence on the routing cost here. We will further analyze the effect of this parameter in the next section to gain a deeper understanding.

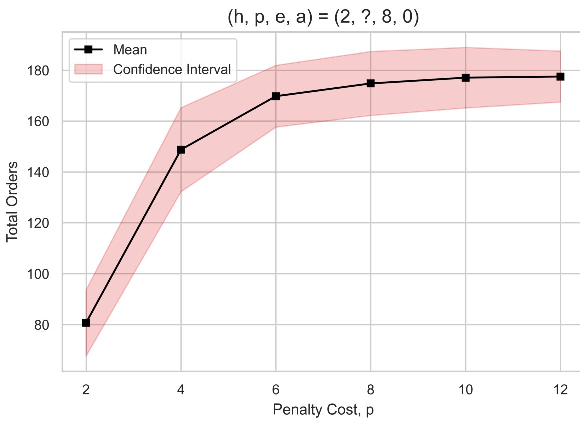
In contrast, when h is greater than zero, the policy shifts its approach and adopts a daily ordering strategy. The primary objective becomes minimizing the holding cost, even if it leads to higher routing costs. This shift occurs because, in our case study, the routing cost is relatively less than the other cost terms. In the daily ordering strategy, there is no product holding, and the vehicles follow the same routes to visit all hospitals and deliver their exact demand. Therefore, the routing cost remains constant and is unaffected by the variation in h. It is worth noting that due to the opposing strategies of the policies at $h = 0$ and $h > 0$, sudden increases are observed at $h = 2$ with the relative growth of $FI > DLA >> UL$. This characteristic is responsible for the atypical appearance of the plots in Figures 6.9c and 6.9d during the transition from $h = 0$ to $h = 2$.

Furthermore, Figure 6.9d shows a great improvement in the average cost of the DLA policy relative to the best UL policy at $h = 0$. This improvement can be attributed to the maximum ordering strategy employed by the DLA policy at $h = 0$, which enhances its robustness against lost demand. In other words, the policy addresses its weakness of having a lower service level, as observed in the previous section. It also displays minimal improvement when going from $h = 2$ to $h = 6$. The slight variations may be attributed to potential experimental errors or discrepancies arising from the VRP algorithm. However, it is noteworthy that beyond the point where $h = 8$, the DLA policy exhibits significantly superior performance compared to the best UL. This favourable performance allows the DLA policy to achieve a relative cost improvement of up to 23.1% compared to the best UL policy, as shown in Figure 6.9d. Overall, the relative cost improvement has an average of 12.5% with an SD of 1.8% in this sensitivity analysis.

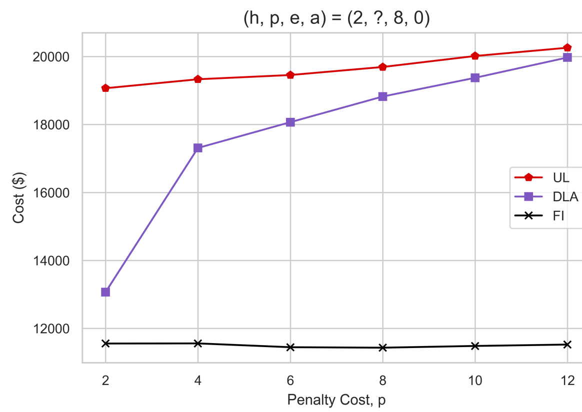
6.7.2 Penalty Cost

In our analysis, we excluded the case where $p < h$ from our study. This choice was made based on our observation in the case study, where we noticed that the model decided not to fulfill any demand in this particular case. This decision is because even ordering a small quantity of product units would result in higher holding and fixed routing costs, which would outweigh the minimal penalty costs associated with not fulfilling the demand.

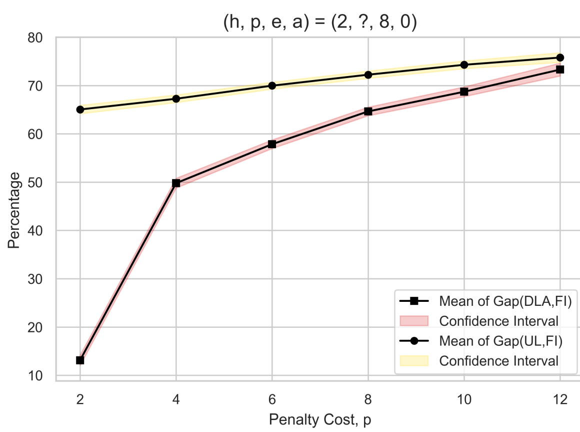
Moving forward, we observe that at $p = 2$ in Figure 6.10a, the DLA policy starts placing orders, but the replenishment quantities remain relatively low compared to the ones with higher penalty cost parameters. This is because incurring penalty costs is still considered more beneficial than incurring holding and fixed routing costs. The DLA policy also ensures no wastage occurs, as pushed by the relationship $p = h << e$. As a result, after fulfilling the demand, the total cost of the DLA policy at this point in Figure 6.10b mainly consists of minor penalty and routing costs. However, it is essential to note that



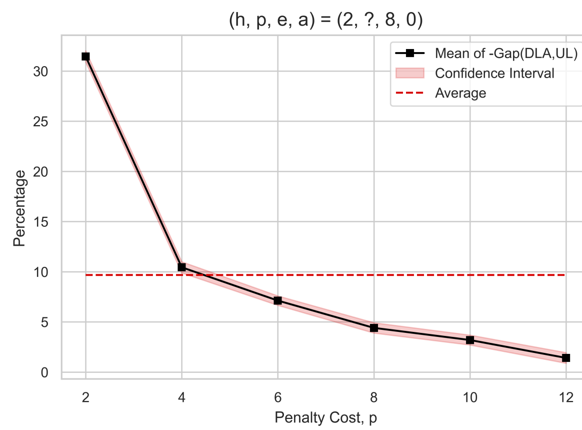
(a) The average total cost



(b) The average total orders by the DLA policy



(c) The gap with FI



(d) The cost improvement of the DLA policy relative to the best UL

Figure 6.10: Sensitivity Analysis of the Penalty Cost

the absence or minimal impact of holding and wastage costs, also contributes to the overall lowest total cost in this figure. The DLA policy's strategy of minimizing the order quantity while accepting some penalty and routing costs results in a smaller total cost compared to the UL policy, which may incur additional costs.

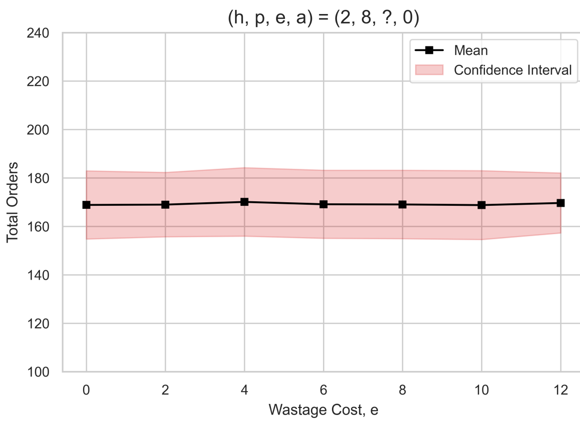
When the value of p exceeds h, the choice of not placing any or a few orders becomes less desirable in the DLA policy. The policy starts to maintain higher inventory levels by ordering more, as shown in Figure 6.10a. This results in higher holding costs, as more products are kept in inventory for longer duration. Furthermore, the higher inventory level also contributes to the wastage cost, as more products are at risk of spoilage. In addition to the rise in holding and wastage costs, there is a corresponding increase in routing costs, as more vehicles are required for product delivery. These factors collectively contribute to a significant overall increase in the total cost. Therefore, as the value of p surpasses h, the combination of a higher holding cost, an increased wastage cost, and augmented routing cost lead to a sharp rise in the total cost, as observed in Figure 6.10b.

When $p = e = 8$, the primary focus of the DLA policy shifts towards minimizing demand losses as much as possible. To achieve this, the policy places the maximum feasible order quantity. The approach remains consistent as we move beyond this point, but the increase in order quantities becomes less noticeable due to various limitations, such as capacity constraints. This behavior can be observed in Figure 6.10a. Consequently, the total cost experiences a minor increase for each unit increase in p compared to the cases when $p < 8$, as shown in Figures 6.10b and 6.10c. This outcome can be attributed to the fact that the penalty cost, which plays a significant role in the objective, is minimized at this stage. Recall that the FI policy benefits from having access to actual demand information, resulting in no demand loss. This characteristic ensures a stable total cost for the FI policy when p varies.

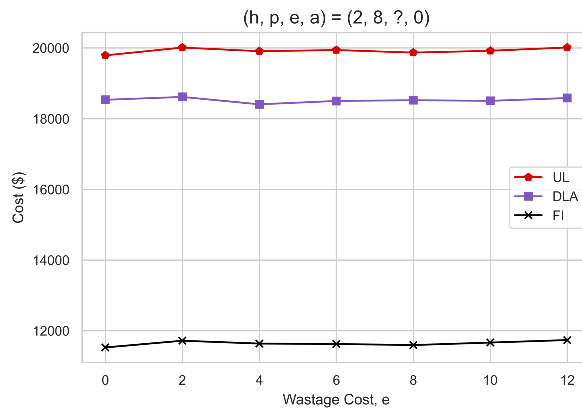
To compare the performance of the DLA policy with the best UL policy, we can refer to Figure 6.10d. Since the UL policy is functional and not directly influenced by the parameter p, it exhibits a consistently increasing trend. On the other hand, as the value of p increases, the total cost of the DLA policy approaches that of the best UL policy. This observation suggests that the DLA policy is more susceptible to changes in the penalty cost. In general, when $p < 8$, the DLA policy demonstrates a more significant improvement in cost compared to the best UL policy. However, as the value of p increases, the extent of this improvement diminishes. This observation emphasizes the importance of the penalty cost and underscores the sensitivity of the DLA policy to variations in this parameter. On average, the policy maintains a cost improvement of 9.7% with an SD of 1.1% relative to the best UL policy.

6.7.3 Wastage Cost

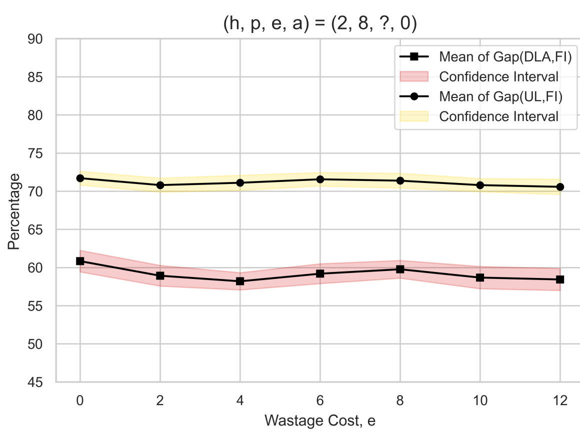
Examining Figure 6.11a, the cost associated with wastage is excluded at $e = 0$, and the DLA policy primarily focuses on minimizing penalty, holding, and fixed routing costs. Since $h < e << p$, the objective becomes minimizing lost demand by increasing replenishment quantities, despite increasing holding and routing costs. On the other hand, as the value of e increases, the policy adopts a more cautious approach by aiming to reduce the quantity of inventory held to lower the risk of wastage. However, the decision to reduce the order quantity introduces a higher penalty cost, which plays a significant role in the total cost. Upon examining the instances, no definite pattern emerges in the ordering quantities determined by the DLA policy due to this trade-off; see Figure 6.11a. The same observations hold when calculating the relative cost between the DLA policy and the best UL policy, as evident in Figure 6.11d. In this figure, the average and SD values are 7.0%



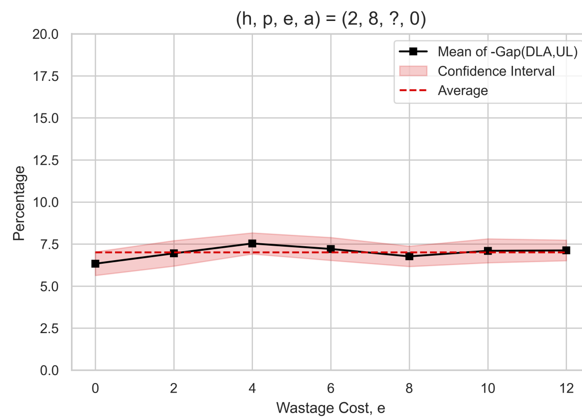
(a) The average total cost



(b) The average total orders by the DLA policy



(c) The gap with FI



(d) The cost improvement of the DLA policy relative to the best UL

Figure 6.11: Sensitivity Analysis of the Wastage Cost

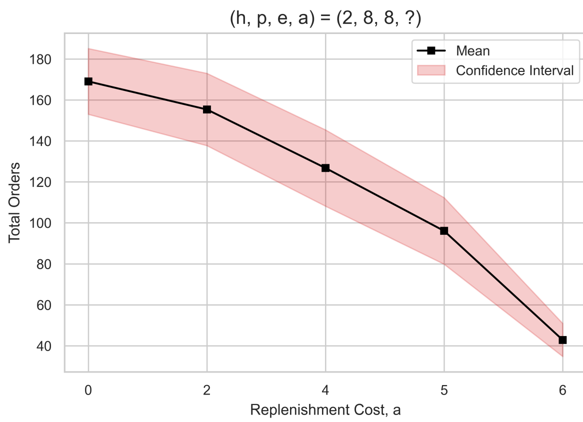
and 1.2%, respectively.

The FI policy is expected to maintain a constant total cost regardless of variations in e. As previously explained, the FI policy possesses knowledge of the actual demand, allowing it to avoid incurring wastage costs. Consequently, the wastage cost associated with the FI policy remains zero, resulting in almost the same cost as e varies. However, it is worth noting that Figure 6.11b displays slight fluctuations in the FI cost. These fluctuations should result from the specific implementation and execution of the VRP algorithm rather than reflecting any meaningful behavior of the FI policy to different values of e.

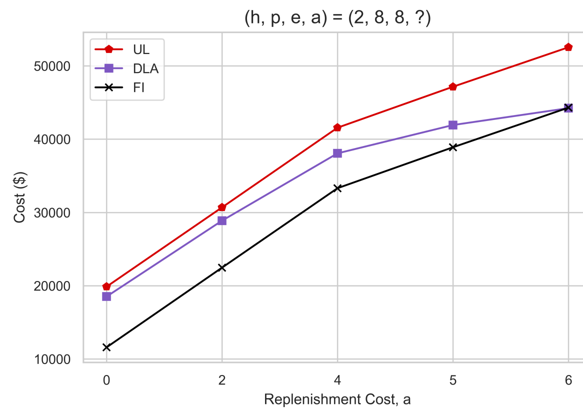
6.7.4 Replenishment Cost

As previously mentioned, due to the non-profit nature of BSC organizations, the replenishment cost is typically considered zero. However, the concept of this cost can be analogously represented by other hidden costs per product unit, such as advertising expenses, which could be incurred by the organization. If we incorporate this concept of replenishment cost into the PIP, a more complicated procedure emerges, wherein fewer orders might be placed. This can be observed in the decision-making pattern of the DLA policy in Figure 6.12a, which progressively decreases the number of orders placed as a increases from zero to 6. Ordering fewer units implies a higher likelihood of losing demand, leading to an increase in the penalty cost. This trend lasts until a point where no or minimal orders are placed, and the total cost solely consists of the penalty and possibly routing costs. In our experiments, this threshold is observed at $a = 6$ in Figure 6.12a.

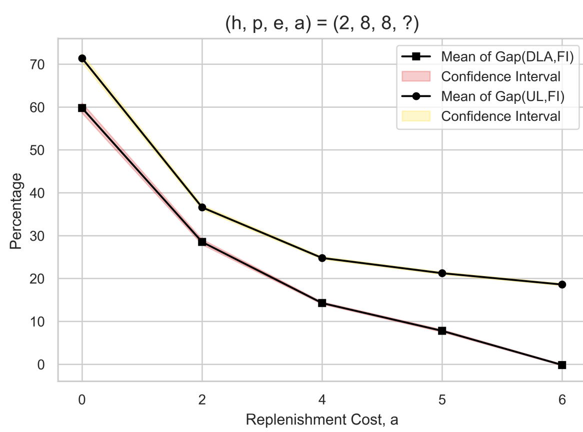
One notable observation in Figure 6.12b is that at $a = 6$, the total cost of the DLA policy aligns with that of FI. This similarity can also be observed in having a zero relative gap in Figure 6.12c. However, it is important to note that this convergence does not



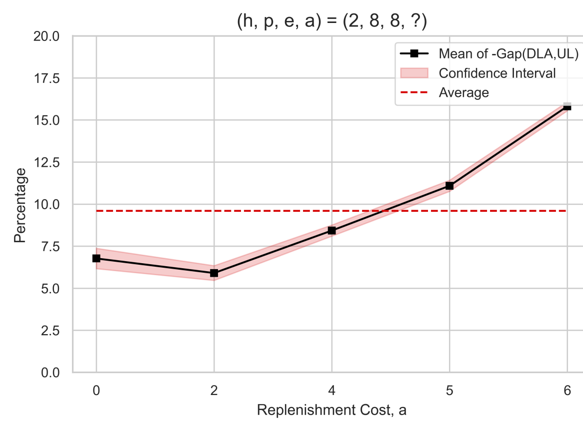
(a) The average total cost



(b) The average total orders by the DLA policy



(c) The gap with FI



(d) The cost improvement of the DLA policy relative to the best UL

Figure 6.12: Sensitivity Analysis of the Replenishment Cost

necessarily indicate that the DLA policy is making more optimal decisions. Instead, the FI policy aims to meet all demand regardless of the replenishment cost, resulting in the same routing cost and augmented replenishment costs. On the other hand, the DLA policy chooses not to fulfill the demand, resulting in penalty and routing costs. This coincidentally led to a total cost similar to that of the FI policy at this specific point.

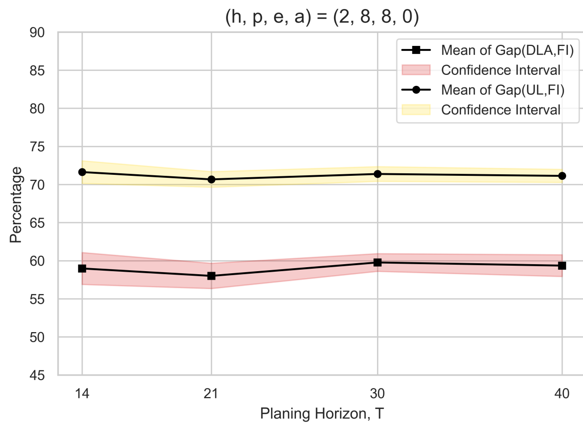
Furthermore, it is important to recognize that the cost improvement of the DLA policy relative to the best UL policy for $a > 4$ in Figure 6.12d should not be misinterpreted as well. While it may appear that the DLA policy exhibits significant improvement, this improvement is primarily a consequence of almost not satisfying any demand. This has been validated by looking into the ordering quantities in Figure 6.12a, which effectively eliminated all cost terms except for the penalty term. Therefore, it is more reasonable to consider the changes in relative cost improvement for $a \leq 4$, where we observe the average cost improvement of approximately 7.0% with an SD of 1.3%.

6.8 Sensitivity Analysis of Other Parameters

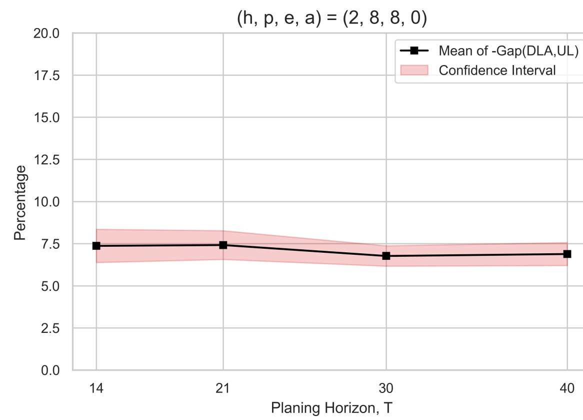
In this section, we analyze the sensitivity of the DLA policy and the best UL policy to some additional parameters. We examine how variations in the planning horizon, the capacity of vehicles, and the lower bound influence the performance of the policies.

6.8.1 Planing Horizon

Figure 6.13 depicts the variations in the total cost as T, the length of the planning horizon, changes using the baseline combination of parameters $(h, p, e, a) = (2, 8, 8, 0)$. Notably, the plot demonstrates high consistency in the cost changes across different planning horizons.



(a) The gap with FI



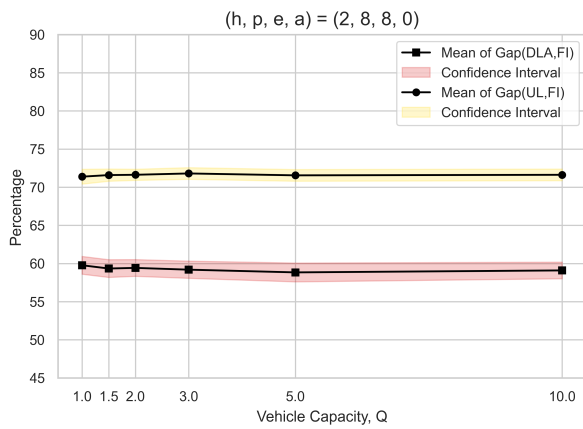
(b) The cost improvement of the DLA policy relative to the best UL

Figure 6.13: Sensitivity Analysis of the Planing Horizon

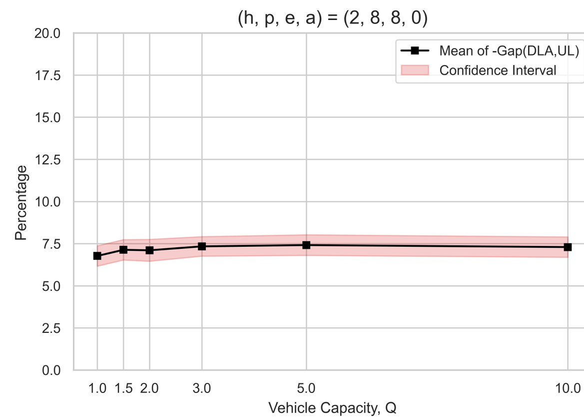
This consistency indicates the effectiveness and applicability of the proposed DLA policy across various planning horizons. It further underscores the robustness and reliability of the DLA policy in addressing PIRP across different time frames.

6.8.2 Capacity of Vehicles

Given that platelet units are small in size and capacity is not a limitation in reality, it makes sense to consider a large vehicle capacity in the VRP. However, practical considerations, such as limited daily working hours that necessitate maintaining a maximum travel distance limit, impact the feasibility of increasing capacity (Hemmelmayr et al., 2009). In our specific case study, we examined the effect of increasing vehicle capacity. Despite the potential benefits, such as reducing the number of vehicles required and lowering routing costs, the overall effect was insignificant. This can be attributed to the dispersed locations



(a) The gap with FI

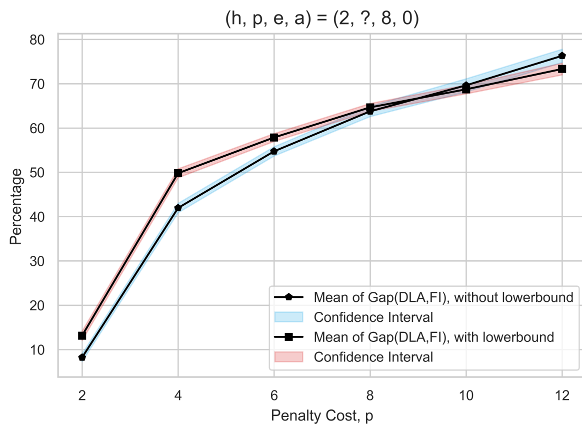


(b) The cost improvement of the DLA policy relative to the best UL

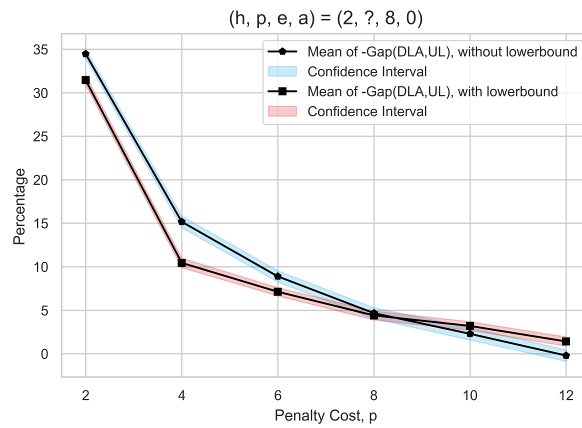
Figure 6.14: Sensitivity Analysis of the Vehicle Capacity

of hospitals within the network, which necessitate the use of multiple vehicles to ensure efficient delivery while adhering to the maximum travel distance limit. It is important to note that increasing vehicle capacity may yield more substantial advantages in denser network settings, where distance limitations have a less pronounced impact.

It is worth mentioning that the capacity of $10Q$ in Figure 6.14 can be interpreted as having no capacity constraint since it can satisfy the demand of almost all hospitals. Moreover, it should be recognized that the best UL policy does not show any improvement because it does not consider the routing cost from the second phase when determining the replenishment quantities.



(a) The gap with FI



(b) The cost improvement of the DLA policy relative to the best UL

Figure 6.15: Sensitivity Analysis of the Lower Bound

6.8.3 Lower Bound

In Section 6.2, we indicated that we utilize the lower bound from Crama et al. (2018) in the DLA policy, aiming to establish a fair comparison between the ADPs from Crama et al. (2018) and the DLA policy. In this subsection, our goal is to assess whether retaining this lower bound in the DLA policy can truly yield benefits in the setting of $L = 7$. To assess this, we conduct a penalty cost analysis without employing the lower bound and compare the results with the findings presented in Subsection 6.7.2. This choice was driven by our realization that the penalty cost term holds the highest impact among the other cost terms considered in the objective function.

Figure 6.15 demonstrates that the presence of a lower bound has a negligible impact when $p = 8 = e$. However, some improvements can be observed as p increases. In contrast, when $p < 8$, the absence of a lower bound can lead to an additional improvement of up to 3%. This reversal in behavior is due to the relevant importance of the penalty cost. When the penalty cost is of utmost significance, enforcing a lower bound and increasing the order quantity in the DLA policy can result in a lower total cost. This is because a higher order quantity enhances the robustness of the DLA policy against potential demand loss, mitigating the impact of penalties. However, when the penalty cost is relatively small compared to the wastage cost, and the wastage cost becomes the dominant factor, it becomes more reasonable to order fewer units. In such cases, the cost associated with the lost demand is lower, outweighing the potential increase in other cost terms such as wastage. Consequently, enforcing a lower bound and ordering more units in the DLA policy does not yield significant benefits.

Chapter 7

Conclusions

In this chapter, we begin by reviewing the PIRP and its key challenges. We then summarize the main findings and insights obtained from this study. Finally, we discuss potential directions for future research and development of the PIRP.

7.1 Summary of the Findings

In this study, we investigated a VMI system characterized by a centralized decision-maker who is responsible for making both inventory and routing decisions within a one-to-many distribution network. In each time period, the decision maker receives system states from retailers about the inventory level of a perishable product and then acts accordingly. We considered the presence of demand uncertainty to ensure more resilient decision-making in the supply chain. As a result, the problem under investigation can be classified as a stochastic PIRP.

To effectively make decisions for the planning horizon, we focused on policies grounded in Lookahead models. These models enable the assessment of the downstream impacts of a decision made in the current period on all future periods. In the VMI, this appeals to determining replenishment quantities and route plans while considering their potential effects on subsequent periods.

In Chapter 4, the basic statement of the Lookahead model for our PIRP resembles the widely recognized Bellman's equation. The Lookahead model can be solved optimally using the DP approach. However, in more complex instances, certain terms of the equation require approximations, leading to the application of ADP methods. We showed that when dealing with relatively larger distribution networks in this field, such as those involving 20 retailers, and when the product's shelf life exceeds a certain threshold, such as 5 days, the curse of dimensionality makes the problem inefficient to be solved using both DP and ADP methods. Consequently, we decided to explore DLA methods, which approximate the downstream impact on future decisions directly.

As a DLA policy, we developed a two-stage approximation-based model in which we consider multiple scenarios throughout the entire planning horizon in Chapter 4. Considering the structure of the problem, we employed a two-phase decomposition approach. Firstly, we incorporated the inventory decisions into the optimization problem; then, in the second phase, we address a single VRP.

In order to obtain optimal solutions in the DLA policy, we encountered the need for linearization. This resulted in a large MIP model that poses computational difficulties. In Chapter 5, we proposed implementation techniques to enhance the execution of the MIP, including leveraging parallelism and incorporating linear cuts.

To examine the performance of the proposed policy, we conducted a BSC case study based on a supplier located in Brampton, with platelets serving as a perishable product.

Before doing numerical experiments, it was necessary to determine the number of sample paths required for the DLA policy. To address this, we ran an extensive experiment, which revealed two key findings. First, there was minimal variation observed among different sets of random sample paths. Second, using only a few sample paths yielded desirable results. This can be attributed to the fact that the DLA policy was employed in each period, eliminating the need to consider a large approximated sample space.

Our numerical experiments in Chapter 6 indicated that the proposed policy consistently outperformed all other policies in terms of total cost and the freshness of the transfused platelet units, while achieving comparable or better results in terms of the number of expired units. It is noteworthy that the use of fresher blood products may contribute to enhanced treatment efficiency. Notably, the superiority of our proposed policy became more pronounced as the shelf life of the products increased. These findings highlighted the effectiveness of the DLA policy in managing inventory and routing decisions for perishable products under demand uncertainty. Furthermore, we conducted a series of sensitivity analyses, which demonstrated an average cost improvement of 9.1% with an SD of 1.3%. These results underscored the robustness and stability of our proposed policy across various settings.

7.2 Future Research

One limitation of the proposed policy was that it resulted in a slightly lower actual service level compared to other policies. As a future research direction, one can explore strategies to increase the service level, particularly by reducing the sensitivity of the policy to the penalty cost.

Furthermore, although our model focuses on perishable products with a fixed shelf life, it has the potential to be extended to include the products whose values gradually decrease over time. In such cases, the model can incorporate the cost of quality loss over time as an added component.

Additionally, investigating more implementation techniques to enhance the performance of the MIP model might be a promising area for further investigation.

Lastly, an interesting avenue for future research would be to study the performance and behavior of the proposed DLA policy in other supply chains, such as food supply chains, where the objective is to maximize profit. Investigating the applicability and effectiveness of the DLA policy in different supply chain contexts would contribute to a deeper understanding of its capabilities and potential for optimizing various industry-specific settings.

References

Abouee-Mehrizi, H., Mirjalili, M., & Sarhangian, V. (2022). Data-driven platelet inventory management under uncertainty in the remaining shelf life of units. *Production and Operations Management*, *31*(10), 3914–3932. https://doi.org/https://doi.org/10.1111/poms.13795

About us - canadian blood services [Accessed: June 30, 2023]. (n.d.). https://www.blood.ca/en/about-us

Alkaabneh, F., Diabat, A., & Gao, H. O. (2020). Benders decomposition for the inventory vehicle routing problem with perishable products and environmental costs. *Computers Operations Research*, *113*, 104751. https://doi.org/https://doi.org/10.1016/j.cor.2019.07.009

Alvarez, A., Cordeau, J.-F., Jans, R., Munari, P., & Morabito, R. (2020). Formulations, branch-and-cut and a hybrid heuristic algorithm for an inventory routing problem with perishable products. *European Journal of Operational Research*, *283*(2), 511–529. https://doi.org/https://doi.org/10.1016/j.ejor.2019.11.015

Alvarez, A., Cordeau, J.-F., Jans, R., Munari, P., & Morabito, R. (2021). Inventory routing under stochastic supply and demand. *Omega*, *102*, 102304. https://doi.org/https://doi.org/10.1016/j.omega.2020.102304

Archetti, C., Boland, N., & Grazia Speranza, M. (2017). A matheuristic for the multivehicle inventory routing problem. *INFORMS Journal on Computing*, *29*(3), 377–387. https://doi.org/10.1287/ijoc.2016.0737

Archetti, C., & Speranza, M. G. (2016). The inventory routing problem: The value of integration. *International Transactions in Operational Research*, *23*(3), 393–407. https://doi.org/https://doi.org/10.1111/itor.12226

Atamtürk, A., & Küçükyavuz, S. (2005). Lot sizing with inventory bounds and fixed costs: Polyhedral study and computation. *Operations Research*, *53*(4), 711–730. https://doi.org/10.1287/opre.1050.0223

Bertazzi, L., Bosco, A., & Laganà, D. (2015). Managing stochastic demand in an inventory routing problem with transportation procurement. *Omega*, *56*, 112–121. https://doi.org/https://doi.org/10.1016/j.omega.2014.09.010

Bertsekas, D. P. (2005). *Dynamic programming and optimal control* (3rd, Vol. 1). Athena Scientific.

Civelek, I., Karaesmen, I., & Scheller-Wolf, A. (2015). Blood platelet inventory management with protection levels. *European Journal of Operational Research*, *243*(3), 826–838. https://doi.org/https://doi.org/10.1016/j.ejor.2015.01.023

Coelho, L. C., Cordeau, J.-F., & Laporte, G. (2012). The inventory-routing problem with transshipment. *Computers Operations Research*, *39*(11), 2537–2548. https://doi.org/https://doi.org/10.1016/j.cor.2011.12.020

Coelho, L. C., & Laporte, G. (2014). Optimal joint replenishment, delivery and inventory management policies for perishable products. *Computers Operations Research*, *47*, 42–52. https://doi.org/https://doi.org/10.1016/j.cor.2014.01.013

Crama, Y., Rezaei, M., Savelsbergh, M., & Woensel, T. V. (2018). Stochastic inventory routing for perishable products. *Transportation Science*, *52*(3), 526–546. https://doi.org/10.1287/trsc.2017.0799

Diabat, A. H., Abdallah, T., & Le, T. H. (2016). A hybrid tabu search based heuristic for the periodic distribution inventory problem with perishable goods. *Annals of Operations Research*, *242*, 373–398.

Dillon, M., Oliveira, F., & Abbasi, B. (2017). A two-stage stochastic programming model for inventory management in the blood supply chain. *International Journal of Production Economics*, *187*, 27–41. https://doi.org/https://doi.org/10.1016/j.ijpe.2017.02.006

Dillon, M., Vauhkonen, I., Arvas, M., Ihalainen, J., Vilkkumaa, E., & Oliveira, F. (2023). Supporting platelet inventory management decisions: What is the effect of extending platelets' shelf life? *European Journal of Operational Research*, *310*(2), 640–654. https://doi.org/https://doi.org/10.1016/j.ejor.2023.03.007

Hamdan, B., & Diabat, A. (2019). A two-stage multi-echelon stochastic blood supply chain problem. *Computers Operations Research*, *101*, 130–143. https://doi.org/https://doi.org/10.1016/j.cor.2018.09.001

Hemmelmayr, V., Doerner, K. F., Hartl, R. F., & Savelsbergh, M. W. P. (2009). Delivery strategies for blood products supplies. *OR Spectrum*, *31*(4), 707–725. https://doi.org/10.1007/s00291-008-0134-7

Le, T., Diabat, A., Richard, J.-P., & Yih, Y. (2013). A column generation-based heuristic algorithm for an inventory routing problem with perishable goods. *Optimization Letters*, *7*(7), 1481–1502. https://doi.org/10.1007/s11590-012-0540-2

Liu, Ke, G. Y., Chen, J., & Zhang, L. (2020). Scheduling the distribution of blood products: A vendor-managed inventory routing approach. *Transportation Research Part E:*

Logistics and Transportation Review, *140*, 101964. https://doi.org/https://doi.org/10.1016/j.tre.2020.101964

Liu, P., Hendalianpour, A., Razmi, J., & Sangari, M. S. (2021). A solution algorithm for integrated production-inventory-routing of perishable goods with transshipment and uncertain demand. *Complex & Intelligent Systems*, *7*(3), 1349–1365. https://doi.org/10.1007/s40747-020-00264-y

Meneses, M., Marques, I., & Barbosa-Póvoa, A. (2023). Blood inventory management: Ordering policies for hospital blood banks under uncertainty. *International Transactions in Operational Research*, *30*(1), 273–301. https://doi.org/https://doi.org/10.1111/itor.12981

Mirjalili, M., Abouee-Mehrizi, H., Barty, R., Heddle, N. M., & Sarhangian, V. (2022). A data-driven approach to determine daily platelet order quantities at hospitals. *Transfusion*, *62*(10), 2048–2056. https://doi.org/https://doi.org/10.1111/trf.17080

Mirzaei, S., & Seifi, A. (2015). Considering lost sale in inventory routing problems for perishable goods. *Computers Industrial Engineering*, *87*, 213–227. https://doi.org/10.1016/j.cie.2015.05.010

Mousavi, R., Bashiri, M., & Nikzad, E. (2022). Stochastic production routing problem for perishable products: Modeling and a solution algorithm. *Computers Operations Research*, *142*, 105725. https://doi.org/https://doi.org/10.1016/j.cor.2022.105725

Nguyen, D. H., & Chen, H. (2022). An effective approach for optimization of a perishable inventory system with uncertainty in both demand and supply. *International Transactions in Operational Research*, *29*(4), 2682–2704. https://doi.org/https://doi.org/10.1111/itor.12846

NHLBI, platelet disorders, thrombocytopenia. (n.d.). https://www.nhlbi.nih.gov/health/thrombocytopenia#What-causes-thrombocytopenia?

Onggo, B. S., Panadero, J., Corlu, C. G., & Juan, A. A. (2019). Agri-food supply chains with stochastic demands: A multi-period inventory routing problem with perishable products. *Simulation Modelling Practice and Theory*, *97*, 101970. https://doi.org/https://doi.org/10.1016/j.simpat.2019.101970

Powell, W. B. (2022). *Reinforcement learning and stochastic optimization: A unified framework for sequential decisions*. John Wiley & Sons, Inc. https://doi.org/10.1002/9781119815068

Powell, W. B., & Ghadimi, S. (2022). The parametric cost function approximation: A new approach for multistage stochastic programming.

Qiu, Y., Qiao, J., & Pardalos, P. M. (2019). Optimal production, replenishment, delivery, routing and inventory management policies for products with perishable inventory. *Omega*, *82*, 193–204. https://doi.org/https://doi.org/10.1016/j.omega.2018.01.006

Rohmer, S., Claassen, G., & Laporte, G. (2019). A two-echelon inventory routing problem for perishable products. *Computers Operations Research*, *107*, 156–172. https://doi.org/https://doi.org/10.1016/j.cor.2019.03.015

Service, N. H. (2023). Blood products - transfusion guidelines. https://www.transfusionguidelines.org/transfusion-handbook/3-providing-safe-blood/3-3-blood-products

Singla, A. R. (2019). *Optimization models for the perishable inventory routing problem* (Master's thesis). University of Waterloo. Waterloo, Ontario, Canada. http://hdl.handle.net/10012/14987

Solomon, M. M. (1987). Algorithms for the vehicle routing and scheduling problems with time window constraints. *Operations Research*, *35*(2), 254–265. https://doi.org/10.1287/opre.35.2.254

Solyalı, O., & Süral, H. (2017). A multi-phase heuristic for the production routing problem. *Computers Operations Research*, *87*, 114–124. https://doi.org/https://doi.org/10.1016/j.cor.2017.06.007

Sonntag, D. R., Schrotenboer, A. H., & Kiesmüller, G. P. (2023). Stochastic inventory routing with time-based shipment consolidation. *European Journal of Operational Research*, *306*(3), 1186–1201. https://doi.org/https://doi.org/10.1016/j.ejor.2022.07.049

Soysal, M., Bloemhof-Ruwaard, J. M., Haijema, R., & van der Vorst, J. G. (2015). Modeling an inventory routing problem for perishable products with environmental considerations and demand uncertainty. *International Journal of Production Economics*, *164*, 118–133. https://doi.org/https://doi.org/10.1016/j.ijpe.2015.03.008

Stroncek, D. F., & Rebulla, P. (2007). Platelet transfusions. *Lancet (London, England)*, *370*(9585), 427–438. https://doi.org/10.1016/s0140-6736(07)61198-2

Xu, Y., & Szmerekovsky, J. (2023). The impact of transshipment on an integrated platelet supply chain: A multi-stage stochastic programming approach. *Computers Industrial Engineering*, *176*, 108991. https://doi.org/https://doi.org/10.1016/j.cie.2023.108991

Zhou, D., Leung, L. C., & Pierskalla, W. P. (2011). Inventory management of platelets in hospitals: Optimal inventory policy for perishable products with regular and optional expedited replenishments. *Manufacturing & Service Operations Management*, *13*(4), 420–438. https://doi.org/10.1287/msom.1110.0334

Appendix

Proof. Proof of Proposition 1. According to Constraints (4.15j) and (4.15k), the value of $\tilde{b}_{t'}^{2,l}$ is equal to 1 if the term $\tilde{D}_{t'} - \sum_{j=1}^{l} \tilde{R}_{t'}^{j}$ becomes non-positive, which results $y_{t'}^{l}$ to be zero. Moving on to the next level of the constraint and adding $\tilde{R}_{t'}^{l+1}$ leads to the same condition for $\tilde{b}_{t'}^{2,l+1}$ and $y_{t'}^{l+1}$; this is due to the fact that adding any additional non-negative values to the summation $\sum_{j=1}^{l} \tilde{R}_{t'}^{j}$ will keep the overall term non-positive. Hence, we can ensure that $\tilde{b}_{t'}^{2,j}$ is also equal to 1 for all $j \in [l+1, L-1]$. This can be achieved by applying the chain of linear cuts given in Equation (5.1).
Intuitively, to implement the FIFO policy, we begin by using the lowest inventory level and turn to higher levels only if there is any unsatisfied demand. Once all the demand has been met using inventory with a shelf life less than or equal to l, we no longer need to consider higher inventory levels. This means that any remaining inventory gets carried over to the next period without any change in size. The linear cuts in Equation (5.1) remove the need to check higher levels for fulfilling the demand once it gets satisfied at the current level.

We can discover the second cut in Equation (5.2) by examining Equation (4.15a) and Constraints (4.15n), (4.15o), and (4.15p). To elaborate, when $\tilde{x}_{t'}$ is set to zero, the minimization objective in Equation (4.15a) causes $\tilde{b}_{t'}^{1}$ to also be zero. Additionally, when $\tilde{x}_{t'}$ is zero, the expression $\tilde{x}_{t'} - \tilde{y}_{t'}^{L-1}$ in Constraint (4.15n) has a non-positive value. This makes

it necessary for $\tilde{b}_{t'}^{3,L-1}$ to equal 1, in order to ensure that $R_{t'+1}$ becomes zero by Constraints (4.15o) and (4.15p). The reverse is true when $\tilde{x}_{t'}$ is non-zero. Accordingly, if $\tilde{b}_{t'}^{1}$ is zero then $\tilde{b}_{t'}^{3,L-1}$ must equal 1, and vice versa.
The intuition behind this is that if no order is placed, no fresh product will be received, causing the inventory with a remaining lifespan of $L-1$ to be empty in the next period.

By investigating the Constraint (4.15g) and the Constraints (4.15j), (4.15k), (4.15o), and (4.15p) for $l = L-1$, we can establish a relation between the binary variables $\tilde{b}_{t'}^{1}$, $\tilde{b}_{t'}^{2,L-1}$, and $\tilde{b}_{t'}^{3,L-1}$. Specifically, out of the eight possible permutations of these three binary variables, $(1,1,1)$, $(0,1,0)$, and $(0,0,0)$ never occur. Therefore, we can eliminate these permutations by incorporating two constraints, $\tilde{b}_{t'}^{1} + \tilde{b}_{t'}^{2,L-1} + \tilde{b}_{t'}^{3,L-1} \geq 1$ and $\tilde{b}_{t'}^{1} + \tilde{b}_{t'}^{2,L-1} + \tilde{b}_{t'}^{3,L-1} \leq 2$. However, since the last two permutations are already excluded by the second linear cut in Equation (5.2), we only need to consider the second inequality as the third linear cut for our MIP. □

Proof. Proof of Proposition 2. Looking at each sample path separately, we can view the MIP formulation as a deterministic lot sizing problem with bounded inventory, which is considered in Atamtürk and Küçükyavuz (2005). In the MIP formulation in (4.15), the binary variable $\tilde{b}_{t'}^{1}$ acts as the fixed-charge variable for ordering in their problem. We assume that their other fixed-charge variable for inventory is set to zero. Additionally, the $\sum_{l=1}^{L-1} \tilde{R}_{t'}^{l}$ corresponds to the inventory at period t' in their formulation. Hence, with some modifications, we can use the inequalities introduced in Atamtürk and Küçükyavuz (2005). As the first modification, we exclude $\tilde{x}_t$ from consideration, given that it is shared among all sample paths, as indicated by Constraint (4.15b). Therefore, we initiate the linear cut for the period $t' \in [t+1, T]$. Furthermore, according to Atamtürk and Küçükyavuz (2005), this linear cut refers to the so-called (l, S) inequalities. We should consider the subset

$S \subseteq [t', t'+L-1]$ in our MIP model since the replenishment quantity at period t' can fulfill the demand for a maximum of $L-1$ future days due to the limited shelf life. It is important to note that after $t' = T - L + 1$, fewer than L periods remain on the planning horizon. Therefore, we adjust the block of periods to $[t', \tau]$, where $\tau = \min\{t' + L - 1,\, T\}$. These modifications lead to the creation of the linear cut in Equation (5.5), which is tailored to our MIP model.

The subsequent set of inequalities, introduced by Atamtürk and Küçükyavuz (2005), are capacitated inequalities for blocks of periods $S \subseteq [t', l]$, where l must be exactly L days after the starting period t' for all $t' \in [t+1, T]$. Employing the same modifications discussed earlier, we arrive at the linear cut in Equation (5.6). □

Proof. Proof of Corollary 1. We limit our linear cut in Equation (5.5) to the blocks of size 1, denoted by $|S| = 1$, for the MIP formulation in (4.15). Since we discovered in our preliminary numerical experiments that including larger subsets would not significantly improve the runtime and might even have adverse impacts. Therefore, adding the linear cut for $|S| = 1$ remains adequate to achieve optimal solutions for our MIP model while greatly reducing computational overhead. This alters Equation (5.5) to Equation (5.7).
It is worth mentioning that Atamtürk and Küçükyavuz (2005) have strengthened this inequality by replacing $\tilde{D}_{[t',l]}$ with $\min\{\tilde{D}_{[t',l]}, \tilde{D}_{t'} + \tilde{ub}_{t'}\}$, where $\tilde{ub}_{t'}$ denotes the upper-bound or capacity to their inventory. In our case, the upper-bound would be $\tilde{ub}_{t'} = \min\{R^{max},\, \tilde{D}_{[t',\tau]}\}$ at period t'. If the first term is true, adding it has no advantage since it is dominated by Constraint (4.15g). If the second term holds, then we have $\tilde{D}_{t'} + \tilde{ub}_{t'} = \tilde{D}_{t'} + \tilde{D}_{[t',\tau]}$, which is weaker than the original inequality in Equation (5.7). Therefore, this reinforcement is not beneficial to our MIP model.

With the same adjustment of $|S| = 1$, the Equation (5.5) is changed to

$$\sum_{l=1}^{L-1} \tilde{R}_{t'}^l + \tilde{x}_{t'} \leq \tilde{ub}_{t'} + \min\{\tilde{ub}_{t'+1} + \tilde{D}_{t'} - \tilde{ub}_{t'},\ \tilde{D}_{[t',\tau]} - \tilde{ub}_{t'},\ \tilde{D}_{[t',\tau]}\}\tilde{b}_{t'}^1 + \sum_{l=1}^{L-1} \tilde{R}_{\tau+1}^l.$$

We can further simplify this inequality by applying $\tilde{ub}_{t'} = \min\{R^{max}, \tilde{D}_{[t',\tau]}\}$. To explain more, four possible events could occur:

- if $\tilde{ub}_{t'+1} = R^{max}$ and $\tilde{ub}_{t'} = R^{max}$, then the minimum term equals $\min\{\tilde{D}_{t'},\ \tilde{D}_{[t',\tau]} - R^{max},\ \tilde{D}_{[t',\tau]}\} = \min\{\tilde{D}_{t'},\ \tilde{D}_{[t',\tau]} - R^{max}\}$
- if $\tilde{ub}_{t'+1} = \tilde{D}_{[t'+1,\tau+1]}$ and $\tilde{ub}_{t'} = R^{max}$, then the minimum term equals $\min\{\tilde{D}_{[t',\tau+1]} - R^{max},\ \tilde{D}_{[t',\tau]} - R^{max},\ \tilde{D}_{[t',\tau]}\} = \tilde{D}_{[t',\tau]} - R^{max}$
- if $\tilde{ub}_{t'+1} = \tilde{D}_{[t'+1,\tau+1]}$ and $\tilde{ub}_{t'} = \tilde{D}_{[t',\tau]}$, then the minimum term equals $\min\{\tilde{D}_{t'+L},\ 0,\ \tilde{D}_{[t',\tau]}\} = 0$
- if $\tilde{ub}_{t'+1} = R^{max}$ and $\tilde{ub}_{t'} = \tilde{D}_{[t',\tau]}$, then the minimum term equals $\min\{R^{max} - \tilde{D}_{[t'+1,\tau]},\ 0,\ \tilde{D}_{[t',\tau]}\} = 0$

Based on the above simplifications, we eventually have three different forms of the linear cut, in none of which the third term $\tilde{D}_{[t',\tau]}$ would appear as the final output. As a result, we can simplify the linear cut with $|S| = 1$ to the one shown in Equation (5.8). □

Glossary

Average Daily Demand Typical amount of products or services that customers require on a given day

Combinatorial Problem Arranging, selecting, or counting objects based on rules or constraints to find the optimal solution from a finite discrete set of feasible solutions

Curse of Dimensionality Strange phenomena and challenges that arise when analyzing data in high-dimensional spaces, such as data sparsity, computational complexity, and degraded algorithmic performance

Execution Time Total time spent by a system to complete a task, including both the direct task execution and any associated time for system service. Also known as run-time

Expectation of a Distribution A weighted average of all possible outcomes of a random variable, where the weights are determined by the probabilities associated with each outcome

Fine-tuning A process of making small and precise adjustments to a pre-trained model to maximize its effectiveness

Finite Horizon A limited timeframe within which plans, decisions, and actions are developed and evaluated, taking into account available resources and constraints

Heuristic Experimental and trial-and-error methods, utilizing experience, intuition, or domain-specific knowledge to guide the search for a solution

Linear Cuts Additional linear constraints incorporated into an optimization model to enhance efficiency by downsizing the feasible region and narrowing the search for optimal solutions

Lookahead The process of considering future outcomes and consequences before making a decision

Mixed Integer Programming Finding the optimal solution to a mathematical model with some or all variables restricted to be integers

Multi-period Decision-making Making decisions for multiple time periods, considering the inter-dependencies and trade-offs between decisions made in each period to optimize outcomes over the entire planning horizon

Local Search An iterative technique that starts with an initial solution and, in each iteration, moves to a neighbouring solution, aiming to find an improved solution within the local search space. Also known as Neighbourhood search

Platelets Small blood cells that help with blood clotting and wound healing

Sensitivity Analysis A method to evaluate how changes in parameters affect the outcomes of a model or system

Sequential Decision-making Making a series of interconnected decisions over a planning horizon, by considering uncertainty and the impact of each decision on future actions

Simulation Laboratory A controlled environment for conducting experiments or studies that replicate real-world scenarios

Standard Deviation A measure of how spread out the values are from the mean (average) of a data set

Stochastic Programming Modelling problems where some parameters or variables are subject to randomness or variability, and seeking optimal solutions that are resilient to uncertain future outcomes

Target Service Level A performance metric expressed as a percentage of a predefined goal, measuring the effectiveness of a system or service in meeting customer demands

Up-to-Level Placing an order to maintain a desired stock level