Health Care Management Science

Managing Perishable Blood Inventories: A Scenario-Based Rolling Horizon Approach for LTOWB under Zero-Inflated Poisson Demand

--Manuscript Draft--

Manuscript Number:	HCMS-D-25-00621
Full Title:	Managing Perishable Blood Inventories: A Scenario-Based Rolling Horizon Approach for LTOWB under Zero-Inflated Poisson Demand
Article Type:	Original Research
Keywords:	Low-Titer O Whole Blood (LTOWB); Scenario-Based Rolling Horizon; Zero-Inflated Poisson Demand; Perishable Inventory Management
Abstract:	Managing Low-Titer O Whole Blood (LTOWB) presents challenges due to the product's short shelf-life and irregular, zero-inflated demand. Classical inventory policies, such as the (s, S) rule, often prove inadequate under these conditions, whereas stochastic dynamic programming (SDP), though producing optimal results, is computationally impractical. This study introduces a scenario-based rolling horizon (Sc-RHA) framework calibrated with historical data from Vancouver General Hospital. Demand is modelled using a zero-inflated Poisson distribution. The method integrates scenario generation, clustering-based reduction, and mixed-integer linear optimization. Experimental benchmarking against SDP demonstrate Sc-RHA achieves near-optimal performance using only a small fraction of the initial scenario pool (≈15%) with a five-day planning horizon. Notably, Sc-RHA lowers the effective demand threshold for reliable performance from 1.0 unit/day (classical (s, S)) to 0.4 units/day, enabling efficient management of LTOWB programs in smaller hospitals.

Managing Perishable Blood Inventories: A Scenario-Based Rolling Horizon Approach for LTOWB under Zero-Inflated Poisson Demand

Abstract

Managing Low-Titer O Whole Blood (LTOWB) presents challenges due to the product's short shelf-life and irregular, zero-inflated demand. Classical inventory policies, such as the (s, S) rule, often prove inadequate under these conditions, whereas stochastic dynamic programming (SDP), though producing optimal results, is computationally impractical. This study introduces a scenario-based rolling horizon (Sc-RHA) framework calibrated with historical data from Vancouver General Hospital. Demand is modelled using a zero-inflated Poisson distribution. The method integrates scenario generation, clustering-based reduction, and mixed-integer linear optimization. Experimental benchmarking against SDP demonstrates Sc-RHA achieves near-optimal performance using only a small fraction of the initial scenario pool ($\approx 15\%$) with a five-day planning horizon. Notably, Sc-RHA lowers the effective demand threshold for reliable performance from 1.0 unit/day (classical (s, S)) to 0.4 units/day, enabling efficient management of LTOWB programs in smaller hospitals.

Keywords: Low-Titer O Whole Blood (LTOWB), Scenario-Based Rolling Horizon, Zero-Inflated Poisson Demand, Perishable Inventory Management

1 Introduction

Blood transfusion is an integral part of modern healthcare, essential for saving lives and improving patient outcomes across a wide range of medical scenarios. Globally, approximately 120 million units of blood are donated yearly, yet supply often falls short of demand, creating issues in emergency settings such as trauma and disaster response, where timely access to blood is critical [1]. These challenges have renewed interest in more efficient transfusion strategies, specifically the reintroduction of whole blood (WB) transfusion. This approach offers a more physiologically balanced alternative to the current component-based transfusion (CBT) method, in which donated blood is separated into red blood cells, plasma, and platelets and transfused separately. Emerging clinical evidence suggests that WB transfusion provides both logistical and therapeutic advantages, particularly in massive hemorrhage scenarios [2]. A significant advancement in this field is Low-Titer O Whole Blood (LTOWB), which has proven especially beneficial in resuscitating military patients under battlefield conditions. Due to its universal compatibility and demonstrated medical efficacy, LTOWB is attracting growing interest in civilian healthcare systems. However, managing LTOWB inventory presents substantial challenges. These include a limited shelf-life and uncertain, intermittent demand, herein referred to as 'spotty demand,' modelled as zero-inflated demand [3]. Conventional methods such as the (s, S) rule can fail under spotty demand. While exact methods for perishable inventory management exist (i.e., stochastic dynamic programming), they remain inaccessible in most instances due to complexity. This study introduces a tractable rolling horizon approach that can be implemented with modest data and computational requirements. Motivated by these challenges, the research addresses the following research questions:

- Can scenario-based rolling horizon inventory policies, explicitly designed for zero-inflated Poisson demand, manage LTOWB inventory and balance the conflicting objectives of minimizing wastage due to perishability and mitigating shortages during demand spikes?
- How do scenario configuration parameters (e.g., scenario sample size, planning horizon length) affect the performance of such policies?

This study provides both methodological advances and practical insights. Findings from the study indicate that (i) short look-ahead horizons (approximately five days) achieve near-optimal performance while maintaining manageable computational requirements; (ii) to stabilize outcomes, a modest scenario budget is sufficient; for the instances studied, about 15% of the total pool provided a reliable balance, with diminishing returns beyond that level; (iii) monitoring expected daily demand levels provides a decision threshold for the effective application of Sc-RH; and, (iv) tailoring inventory strategies to relative shortage and wastage penalties enables hospitals to align decisions with operational priorities.

2 Literature Review

2.1 Classical (s, S) and base-stock policies

Classical inventory control policies, such as the reorder point, order-up-to level (s, S) policy, have long been used to manage inventory under uncertainty. Under certain cost conditions (convex holding and shortage costs with fixed order costs), the (s, S) policy is provably optimal [4]. For perishable goods, however, ignoring product age makes such policies suboptimal. Thus, researchers have extended (s, S) policies to account for perishability. For example, Gürler and Özkaya studied a continuous review (s, S) system with random shelf lives and Poisson demand, showing performance losses can be "drastic" when shelf-life variability is high [5]. Other simple rules, such as base stock or (S-1,S) policies, remain attractive for their ease of use and are still applied in practice, especially in blood banks [6]. Yet these models depend heavily on demand assumptions and cannot handle intermittent usage well. Because perishable systems must track inventory by age, while (s, S) rules only consider totals, such simplifications yield suboptimal results [7]. Extensions, such as Barron and Baron and Kouki et al. introduced random shelf-life via Markov chains [8, 9]. These enhance realism, but remain limited when demand is erratic, motivating more advanced approaches such as stochastic dynamic programming.

2.2 Stochastic Dynamic Programming: Optimality and Computational Limits

Stochastic dynamic programming (SDP) models perishable inventory as a Markov decision process, solving Bellman equations over multi-period states. SDP provides an optimal result and can serve as a benchmark for other methods. For instance, Haijema et al. combined SDP with simulation for platelet inventory [10], Abdulwahab and Wahab extended it to multiple blood types [11], and Abouee-Mehrizi et al. addressed shelf-life uncertainty [12]. Despite these advances, SDP suffers from the curse of dimensionality. Since the state must capture the number of units at each remaining shelf-life, the size of the state space grows combinatorially with both the maximum shelf-life and storage capacity. For example, with a 14-day shelf-life and a capacity of six units per age category, the number of possible states exceeds 678 billion, making enumeration and exact policy computation infeasible. Even with GPU acceleration (e.g., Farrington et al. [13] and Ortega et al. [14]), scalability remains limited to reduced horizons or inventory capacities. As a result, SDP is primarily used as a benchmark in smaller instances, while heuristics, such as rolling horizon methods, provide more practical solutions for hospital scale problems [15–17].

2.3 Scenario-Based and Rolling Horizon Control Strategies for Blood Inventory

Rolling horizon (RH) and model predictive control (MPC) methods address complexity by repeatedly solving shorter horizon optimization problems. They are especially suited to perishable inventory because they allow frequent re-planning as new information arrives. Osorio et al. integrated simulation and optimization for donor scheduling [18]. Dalalah et al. applied a rolling horizon simulation optimization approach to platelets, reducing wastage and outdating [19]. Shih and Rajendran used a stochastic MILP [20], and Dural Selçuk et al. showed rolling horizon can deliver near optimal performance [21]. Scenario based approaches, such as Mohammadi et al., further highlight the value of embedding uncertainty into hospital blood bank planning [22]. More recently Hosseini, Motlagh et al. extended rolling horizon to collection planning with robust optimization [23]. In the broader control literature, Schildbach and Morari [24] introduced a scenario-based model predictive control framework that repeatedly solves rolling horizon optimization problems under sampled uncertainty scenarios. This approach demonstrates how scenario sampling can make MPC robust to demand and supply variability, closely aligning with the philosophy of the scenario based rolling horizon method proposed in this study. Other MPC applications in healthcare, from Braun et al. to Lejarza et al. [25–30], further demonstrate the adaptability of rolling horizon approaches. Yet most rely on Poisson or empirical demand assumptions, leaving intermittent, zero inflated demand unaddressed.

2.4 Inventory Models for Zero-Inflated and Intermittent Demand

Zero Inflated Poisson (ZIP) and Zero Inflated Negative Binomial (ZINB) models capture demand with long periods of inactivity punctuated by sudden spikes, a typical pattern for blood products in smaller settings . Dehghani et al. applied a ZINB distribution to hospital blood demand and incorporated it into a stochastic transshipment model, generating policies that were robust to periods of negligible demand such as weekends [31]. Similarly, Mohammadi et al. analyzed blood transfusion data using a bivariate zero-inflated Poisson model within a Bayesian framework, highlighting the flexibility of such approaches in capturing over-dispersion and excess zeros in transfusion datasets [32]. Earlier studies mainly emphasized demand estimation rather than control: Ding used Bayesian updating for ZIP parameter learning [33], Ünlü and Rossetti compared zero modified demand families [34], and Costantino et al. and Finco et al. applied ZIP frameworks to spare parts management [35, 36]. More recently, Safavi and Blake demonstrated that classical (s, S) policies perform poorly under ZIP demand for specialized blood products [3].

2.5 Evolution of the Literature and Comparative Synthesis

Table 1 gives examples of how the literature has moved from simple (s, S) rule to SDP for theoretical benchmarks, to rolling horizon and model predictive control methods for tractability, and finally to ZIP and ZINB models for improved demand forecasting. However, two gaps remain: (1) limited integration of zero inflated demand into operational inventory policies; and, (2) the absence of scenario-based rolling horizon models that explicitly accommodate intermittent demand. This study addresses both gaps by embedding a ZIP demand model directly into a scenario-based rolling horizon MILP for perishable inventory, incorporating explicit shelf-life tracking and FIFO usage. The method accommodates intermittent demand and delivers implementable replenishment thresholds.

3 Problem Description and Model Formulation

3.1 Problem Statement

Vancouver General Hospital (VGH), a major trauma centre in Canada, faces operational challenges managing an inventory of LTOWB . LTOWB, known for rapid and effective transfusion in life-threatening emergencies, is recognized for reducing transfusion complexity by delivering all necessary blood components to promote hemostasis in a single unit. Despite its clinical benefits, LTOWB inventory is inherently difficult to manage due to its short shelf-life and irregular demand. The distribution of LTOWB demand at VGH over a 723-day period is depicted in Figure 1. Analyzing historical data from VGH reveals significant demand irregularities: Out of 723 days, demand for LTOWB occurred on only 72 days, or approximately 10% of the time. Even on days with non-zero demand, the average request was low (mean of approximately 0.4 units per non-zero day), yet occasional demand surges reached six units per day. Traditional inventory models, based on normal or Poisson distributions, fail to accurately represent such demand patterns. While ZIP models capture these demand irregularities, they do not provide inventory policies. Without inventory policies tailored to zero-inflated scenarios, hospitals will experience costly wastage or critical stockouts during demand surges.

3.2 Model Formulation

3.2.1 Nomenclature

Indices:

t	Planning periods, $t \in \{1, \ldots, T\}$.
ℓ	Remaining shelf-life upon arrival, $\ell \in \{1, \ldots, S\}$, where $\ell = 1$ oldest and $\ell = S$ freshest.
m	FIFO threshold index, $m \in \{1, \ldots, S-1\}$.

Table 1: Comparative summary of the literature on perishable inventory models and the contribution of this study.

Category	Representative Studies	Demand Model	Perishability Handling	Method	Key Limitation
(s, S) policies	Gürler & Özkaya (2008); Barron & Baron (2019); Kouki et al. (2015)	Poisson or state-dependent Poisson	Random shelf-life, Markov age-tracking	Continuous review (s, S) extensions	Ignores age detail in decisions, poor under intermittent demand
Stochastic DP / ADP	Haijema et al. (2007); Abdulwahab & Wahab (2014); Abouee-Mehrizi et al. (2023); Farrington et al. (2025)	Poisson (with or without shelf-life uncertainty)	Explicit aging, multi-age categories	Exact DP, Approximate DP, GPU-accelerated SDP	Curse of dimensionality, infeasible for long horizons (e.g., 14-day LTOWB)
Rolling Horizon / MPC	Osorio et al. (2017); Dalalah et al. (2018); Shih & Rajendran (2020); Hosseini Motlagh et al. (2024); Schildbach & Morari (2016)	Empirical or Poisson	Age-differentiated demand, platelet aging	Rolling horizon simulation-optimization, MILP, robust RH, scenario-based MPC	Demand models too simplistic (no zero inflation), limited robustness
ZIP-based Forecasting / Policy Evaluation	Ding (2002); Ünlü & Rossetti (2011); Costantino et al. (2018); Finco et al. (2022); Dehghani et al. (2021); Safavi & Blake (2025)	Zero Inflated Poisson (ZIP), Zero Inflated Negative Binomial (ZINB)	Limited or empirical shelf-life tracking	Forecasting (Ding, Ünlü, Costantino, Finco); ZINB-based transshipment policies (Dehghani); ZIP (s, S) policy evaluation (Safavi & Blake)	Improved demand modeling, but no optimization-based replenishment policies
ZIP + Optimization	This study	Zero Inflated Poisson (ZIP)	Explicit shelf-life tracking, FIFO, wastage constraints	Scenario-based Rolling Horizon MILP	First to embed ZIP demand in optimization, delivers operational thresholds for practice

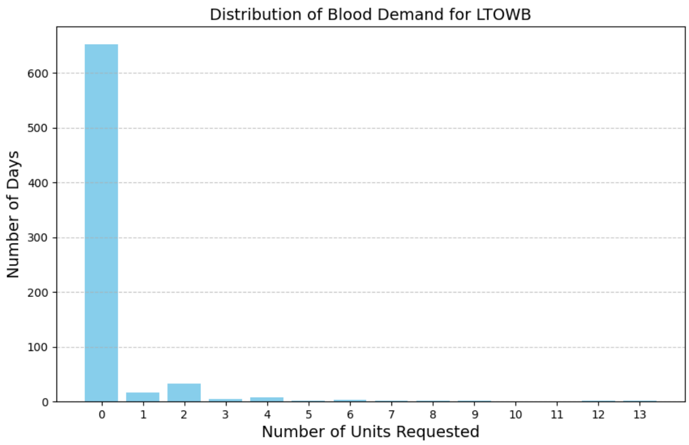


Fig. 1: Daily frequency distribution of LTOWB demand (in units) at VGH

Parameters:

S	Maximum shelf-life (periods).
C_o, C_s, C_w	Unit costs of ordering, shortage, and wastage, respectively.
d_t	Demand in period t.
$p_{t,\ell}$	Probability that an order placed in t arrives with ℓ periods of remaining shelf-life.
y_ℓ^{init}	Initial inventory available at the start of period 1 for classes $\ell \geq 2$.

Decision Variables:

x_t	Units ordered at the start of period t.
$i_{t,\ell}$	Units received in t with ℓ periods of remaining shelf-life.
$r_{t,\ell}$	Units used in t from class ℓ to satisfy demand.
$y_{t,\ell}$	End-of-period t inventory in class ℓ ($\ell \geq 2$).
w_t	Wastage at the end of period t.
s_t	Shortage at the end of period t.

3.2.2 Mathematical Model

Formulation (1)–(11) specifies the MILP model for the LTOWB inventory problem.

$$\min \ Z = \sum_{t=1}^{T} \big(C_o x_t + C_s s_t + C_w w_t\big) \tag{1}$$

$$i_{t,\ell} = \lceil p_{t,\ell}\, x_t \rceil \qquad \forall t,\ \forall \ell \tag{2}$$

$$\sum_{\ell=1}^{S} i_{t,\ell} = x_t \qquad \forall t \tag{3}$$

$$\sum_{\ell=1}^{S} r_{t,\ell} + s_t = d_t \qquad \forall t \tag{4}$$

$$0 \leq r_{t,\ell} \leq y_{t-1,\ell+1} + i_{t,\ell} \qquad \forall t,\ \forall \ell \tag{5}$$

$$\sum_{k=1}^{m} r_{t,k} \leq \sum_{k=1}^{m} \big(y_{t-1,k+1} + i_{t,k}\big) \qquad \forall t,\ \forall m = 1, \ldots, S-1 \tag{6}$$

$$\sum_{k=1}^{m} r_{t,k} \geq d_t - \sum_{k=m+1}^{S} \big(y_{t-1,k+1} + i_{t,k}\big) \forall t,\ \forall m = 1, \ldots, S-1 \tag{7}$$

$$y_{t,\ell} = y_{t-1,\ell+1} + i_{t,\ell} - r_{t,\ell} \qquad \forall t,\ \ell = 2, \ldots, S \tag{8}$$

$$w_t = y_{t-1,2} + i_{t,1} - r_{t,1} \qquad \forall t \tag{9}$$

$$y_{0,\ell} = y_\ell^{\text{init}}, \quad y_{0,1} = 0 \qquad \ell = 2, \dots, S \tag{10}$$

$$x_t,\ r_{t,\ell},\ y_{t,\ell},\ s_t,\ w_t,\ i_{t,\ell} \in \mathbb{Z}_{\geq 0} \qquad \forall t,\ \forall \ell \tag{11}$$

The objective function (1) minimizes the total cost of ordering, shortages, and wastage across the planning horizon. Constraints (2)–(3) define arrivals: each order quantity is distributed into shelf-life classes using the probability vector $p_{t,\ell}$. Since the ceiling operator in (2) is nonlinear, it is implemented through an SOS2 linearization in Appendix A [37, 38]. Constraint (3) ensures that arrivals sum exactly to the order quantity; although redundant, it is included for expositional completeness.

Constraint (4) enforces demand balance by requiring that total demand in each period is satisfied either through withdrawals from inventory or recorded as a shortage. Constraint (5) limits withdrawals from each age class by the available stock, including aged inventory and new arrivals.

Constraints (6)–(7) together enforce FIFO consumption: (6) ensures that cumulative withdrawals from the oldest m classes cannot exceed their cumulative availability, while (7) requires that these classes cover at least the portion of demand that newer classes cannot satisfy.

Constraint (8) maintains inventory balances for non-expiring classes ($\ell = 2, \dots, S$), reflecting ageing, arrivals, and withdrawals. Constraint (9) defines wastage as the unused portion of age-1 units after demand satisfaction. Constraint (10) specifies the initial inventory condition, setting available units at the start of period 1, with no stock in class $\ell = 1$. Finally, constraint (11) enforces non-negativity and integrality of all decision variables.

4 Solution Method and Results

A review of the literature shows that rolling horizon methods have proven effective for perishable products, yielding replenishment policies that balance demand satisfaction with ordering, shortage, and wastage costs. Prior studies have demonstrated their value in sequential decision-making under partial knowledge of future conditions [24–30]. This study adopts a Scenario-based Rolling Horizon Approach (Sc-RHA) embedded within a MILP framework. The approach models two key sources of uncertainty: demand fluctuations, captured by a ZIP distribution, and variation in the remaining shelf-life of incoming inventory units. Historical LTOWB demand data from VGH are used to calibrate demand trajectories through maximum-likelihood estimation of ZIP parameters. Since the VGH dataset does not include the shelf-life of arrivals, the empirical distribution for incoming-unit shelf-life is derived from shipments of platelet units in a different region.

The Sc-RHA procedure generates an initial scenario pool that captures plausible joint evolutions of demand and incoming-unit shelf-life over the planning horizon. A stratify-then-cluster reduction maintains tractability while preserving statistical representativeness. Shelf-life stratification precedes k-means clustering on ZIP demand features, and the nearest-to-centroid scenarios are retained to construct a representative reduced set (RRS). The full scenario pool remains available for evaluation, whereas the MILP only uses the RRS for decision-making. This design preserves coverage while ensuring computational feasibility. An adaptive clustering control adjusts the balance between scenario-budget size and performance, enabling exploration of how reduction granularity influences solution quality.

At each decision period, the model considers a look-ahead window of "L" periods and identifies the ordering actions that minimize expected total operational cost, including ordering, shortage, and wastage components. During each period, the MILP is solved once per scenario in the RRS, producing distinct ordering sequences. Each sequence is subsequently evaluated across the full scenario pool to measure expected performance under all possible realizations. The sequence with the lowest average cost is selected, but only its first ordering decision is implemented. The horizon then advances by one period, and the optimization–evaluation cycle repeats. This approach reflects standard practice in RH multistage stochastic optimization and underpins the average-cost selector used here. After implementation of the chosen order, realized demand and shelf-life outcomes are sampled from the empirical distributions. Inventory levels and age profiles update accordingly, and the process continues until the end of the planning horizon. Algorithm 1 summarizes the complete Sc-RHA procedure, including optimization, cross-scenario evaluation, and inventory updates based on simulated outcomes.

Algorithm 1 Scenario-Based Rolling Horizon Approach (Sc-RHA)

Require:

- Planning horizon T, look-ahead window length L
- Full scenario pool $\mathcal{S}_{\text{full}} = \{(d_t^m, p_{t,\ell}^m)\}_{m=1}^M$
- Adaptive clustering controls K, c (clusters, per-cluster coverage)
- Cost parameters C_o, C_s, C_w
- Initial inventory $i_{1,\ell}$ for $\ell = 1, \dots, S$

Ensure:

- Ordering policy $\{x_1, x_2, \dots, x_T\}$
- TotalCost over horizon T

$t \leftarrow 1$
Initialize inventory: $i_{t,\ell} \leftarrow i_{1,\ell}$ for all ℓ
TotalCost $\leftarrow 0$
$\mathcal{S}_{\text{RRS}} \leftarrow$ StratifyThenClusterSelect($\mathcal{S}_{\text{full}}, K, c$) ▷ Representative Reduced Set (RRS)
$N \leftarrow |\mathcal{S}_{\text{RRS}}|$
while $t \le T$ **do**
  $(K_t, c_t) \leftarrow$ UpdateReductionControls($t, i_{t,\ell}$) ▷ adaptive clustering control
  $\mathcal{S}_{\text{RRS}} \leftarrow$ StratifyThenClusterSelect($\mathcal{S}_{\text{full}}, K_t, c_t$)
  $N \leftarrow |\mathcal{S}_{\text{RRS}}|$
  for $s = 1$ to N **do**
    Use $(d_t^s, p_{t,\ell}^s) \in \mathcal{S}_{\text{RRS}}$ to solve MILP on $[t, t+L-1]$
    Let $x^s \leftarrow \{x_t^s, x_{t+1}^s, \dots, x_{t+L-1}^s\}$ ▷ ordering sequence from RRS
    for $j = 1$ to M **do**
      Evaluate Cost($x^s, d_t^j, p_{t,\ell}^j, i_{t,\ell}$) under $(d_t^j, p_{t,\ell}^j) \in \mathcal{S}_{\text{full}}$
    end for
    $\bar{C}^s \leftarrow \frac{1}{M} \sum_{j=1}^M \text{Cost}(x^s, d_t^j, p_{t,\ell}^j, i_{t,\ell})$ ▷ expected cost over full pool
  end for
  $s^* \leftarrow \arg\min_s \bar{C}^s$
  $x_t \leftarrow x_t^{s^*}$ ▷ implement first decision of best RRS policy
  Draw $d_t^{\text{real}} \sim \text{ZIP}(\lambda, \pi)$
  Draw $p_{t,\ell}^{\text{real}} \sim$ ShelfLifeDist
  $i_{t+1,\ell} \leftarrow$ InventoryUpdate($i_{t,\ell}, x_t, d_t^{\text{real}}, p_{t,\ell}^{\text{real}}$)
  $c_t \leftarrow$ Cost($x_t, d_t^{\text{real}}, i_{t,\ell}$)
  TotalCost $\leftarrow$ TotalCost $+ c_t$
  $t \leftarrow t + 1$
end while
return $\{x_1, \dots, x_T\}$, TotalCost

4.1 Scenario Generation and Reduction

Scenarios consist of a realization of demand for the blood product and a vector describing the remaining shelf-life of incoming units. Demand follows a zero-inflated Poisson (ZIP) distribution with two parameters: mean demand (λ) and zero-inflation probability (π). Historical data from Vancouver General Hospital (VGH) motivate these parameter values. Demand trajectories are generated by enumerating all (λ, π) combinations to ensure complete coverage of the parameter space.

The remaining shelf-life of incoming units follows an empirical distribution derived from historical records. For modelling purposes, this distribution is discretized into three operational categories: *high* (mostly fresh units), *medium* (middle-aged units), and *low* (near expiration). Combining the ZIP parameter grid with the three shelf-life profiles produces the initial scenario pool.

Based on the scenario-approach tail bound of Calafiore and Campi [39], with $\varepsilon = 0.05$ and $\theta = 0.05$, the theoretical minimum sample size is approximately 717 scenarios. To exceed this conservative requirement and guarantee broad coverage, the initial pool is set to $M = 1000$ scenarios.

Solving the MILP with all M scenarios is computationally prohibitive, so an adaptive reduction procedure is applied to preserve statistical representativeness while reducing computational burden. The pool is first partitioned into high, medium, and low strata according to empirical proportions. Within each stratum, ZIP-based demand features (mean and variance) are clustered using k-means, and from each cluster the nearest-to-centroid scenarios are retained. With K clusters and c scenarios retained per cluster, the reduced representative set (RRS) size is

$$N = 3 \times c \times K.$$

This structure allows targeting a desired fraction $f = N/M$ of the initial pool when studying budget–performance trade-offs. This methodology aligns with established scenario-based optimization practices, employing scenario generation and scenario reduction techniques common in the literature [39–42].

4.2 Benchmarking and Experimental Evaluation

To evaluate the performance of the Sc-RHA, its outcomes are benchmarked against a stochastic dynamic programming (SDP) model that serves as a lower-bound reference. The SDP model iterates over all future demand and shelf-life realizations, thereby enabling the computation of optimal ordering decisions that minimize the expected total operational cost. The model tracks inventory states across shelf-life categories, incorporates empirically derived shelf-life distributions, and utilizes a ZIP demand model calibrated to historical data. A detailed mathematical formulation is provided in Appendix A3 and [3].

The relative performance of Sc-RHA is quantified using an optimality gap, defined as:

$$\text{Gap}(\%) = \frac{C_{\text{Sc-RHA}} - C_{\text{SDP}}}{C_{\text{SDP}}} \times 100,$$

where $C_{\text{Sc-RHA}}$ and C_{SDP} denote the total operational costs of the Sc-RHA and the SDP benchmark, respectively. This gap quantifies the relative deviation of the Sc-RHA method from the optimal cost achieved by the SDP benchmark.

Following the benchmarking, an experimental analysis is conducted to investigate the sensitivity of Sc-RHA performance across a range of operating conditions. A full factorial experimental design is employed with the following parameters:

- Mean demand $\lambda \in \{0.5, 1, 1.5, 2, 3\}$,
- Probability of zero inflation $\pi \in \{0.2, 0.4, 0.6, 0.9\}$,
- Number of selected scenarios $N \in \{25, 50, 75, 100, 125, 150, 500, 750, 1000\}$,
- Look-ahead window length $L \in \{1, 3, 5, 7, 9, 11, 14\}$.

The parameter ranges were selected to balance empirical grounding with coverage of relevant operating regimes. The values of mean demand λ and zero-inflation probability π are calibrated from the VGH data, ensuring that experimental conditions reflect realistic demand magnitudes and spotty patterns observed in practice. The number of scenarios N is varied to examine the trade-off between computational effort and solution quality. The initial values increase in steps of 25 from 25 to 150, after which larger increments (500, 750, and 1000) are introduced to capture the performance of Sc-RHA across both practically tractable and exhaustive scenario sets. The look-ahead horizon length L begins with a purely myopic case ($L = 1$) and increases in two-day increments to capture progressive levels of foresight. The final setting ($L = 14$) matches the maximum shelf-life of the product, representing the longest feasible horizon under the system constraints.

For each parameter combination, the Sc-RHA method is executed over the full planning horizon, with performance evaluated based on total operational cost, including ordering, shortage, and wastage costs. The operational cost reflects established conventions in the blood inventory literature. A fixed ordering cost of 10 is assumed, while shortage penalties are set 100 times higher than wastage costs (1 unit) to represent the critical consequences of unmet demand. This 100:1 ratio aligns with modeling practices in the literature (Haijema et al., Stanger et al., Blake et al.). Holding costs are omitted due to their negligible impact [10, 43, 44]. Results for the experiments executed under a factorial design are compared via analysis of variance (ANOVA), Table 2.

Main Effects: Results indicate that both mean demand and probability of zero inflation strongly affect the optimality gap. Specifically, λ ($F = 30.4$, $p < 0.001$) and π ($F = 34.5$, $p < 0.001$) significantly influence the performance gap. Increasing λ reduces the gap, whereas increasing π increases it. The look-ahead window length L also shows a statistically significant, yet more moderate, effect

Table 2: Analysis of Variance results.

Parameter	DoF	SS	MS	F	p-value
λ	4	10251.9	2563.2	30.4	< 0.001
π	3	8719.7	2906.6	34.5	< 0.001
N	8	1542.5	192.8	2.3	0.020
L	6	4583.8	763.9	9.1	< 0.001
$\lambda \times \pi$	12	9118.2	759.8	9.0	< 0.001
$\lambda \times N$	32	3199.2	100.0	1.2	0.225
$\pi \times N$	24	3955.6	164.8	1.9	0.004
$\lambda \times L$	24	4633.1	193.0	2.3	0.005
$\pi \times L$	18	2928.4	162.7	1.9	0.011
$N \times L$	48	4388.5	91.4	1.1	0.330
$\lambda \times \pi \times N$	96	7784.0	81.1	1.0	0.585
$\lambda \times \pi \times L$	72	8504.7	118.1	1.4	0.021
$\lambda \times N \times L$	192	8905.2	46.4	0.5	0.999
$\pi \times N \times L$	144	8062.1	55.9	0.7	0.998

($F = 9.1$, $p < 0.001$), indicating that increasing L reduces the optimality gap. The number of scenarios N shows a statistically significant but relatively small effect ($F = 2.3$, $p = 0.020$), suggesting that increases in N beyond small values provide limited marginal improvement.

Interaction Effects: Significant two-way interactions are observed between λ and π, λ and L, π and L, and π and N, indicating that the effects of these parameters are strongly context-dependent (Figure 2). For λ and π, the gap decreases with higher mean demand, but this reduction is far stronger when π is high, reflecting the stabilizing effect of larger demand under intermittent conditions. For λ and L, longer look-ahead horizons reduce the gap primarily when λ is low, where small fluctuations have a relatively larger effect. At higher demand levels, extending the planning horizon has little added benefit. The π and L interaction is also significant: in higher π settings, demand variability is greater, thereby increasing the value of longer planning. These findings suggest that selecting an appropriate planning horizon length should consider both the mean demand level and its variability to improve inventory management effectiveness.

In cases where π is high, increasing the number of scenarios (N) can help reduce the gap by capturing more diverse demand patterns. Conversely, when π is low and the demand follows a steadier stream, adding more scenarios does not significantly improve performance, as fewer variations exist to consider. This indicates that generating additional scenarios is more beneficial when demand is uncertain or spotty.

The interactions highlight the importance of choosing model parameters that balance solution quality and computational effort. The analysis evaluates how the rolling horizon method responds to the scenario budget and the look-ahead window length L. For the scenario budget, results are reported by the fraction of the initial pool $f = N/M$ under a constant-coverage ($c = 10$ scenarios per cluster) adaptive-K design with $N = 3Kc$. For the horizon length, L varies while other settings remain fixed.

Figures. 3a and b display mean optimality gaps with 95% confidence intervals over 100 replications. Both panels show diminishing returns. For the scenario budget, Tukey's HSD finds no statistically significant improvement once the budget reaches about 15% of the initial pool ($\sim N = 120$ when $M = 1000$; Appendix A4). Beyond this threshold, enlarging the pool simply increases runtime; confidence intervals contract as expected, but the mean gap remains stable. For the look-ahead, pairwise contrasts reveal no mean gap improvement beyond a five-day window (Appendix A4), consistent with a diminishing decrease roughly proportional to $1/L$ [45].

Accordingly, the main experiments adopt a parsimonious configuration that uses approximately 15% of the initial scenario pool and $L = 5$, delivering good performance at lower computational cost.

The diminishing returns associated with longer planning horizons can be explained by the demand structure itself. When demand is frequent and stable (high λ, low π), the probability that at least one unit will be required within a few days increases rapidly, meaning that most relevant foresight is already captured with relatively short horizons. In contrast, under sparse and highly uncertain demand (low λ, high π), even a long horizon adds little predictive value. This saturation effect explains the shapes in performance metrics observed in Figure 3b and supports the finding that horizons beyond five days provide limited additional benefit while increasing computational cost. A formal

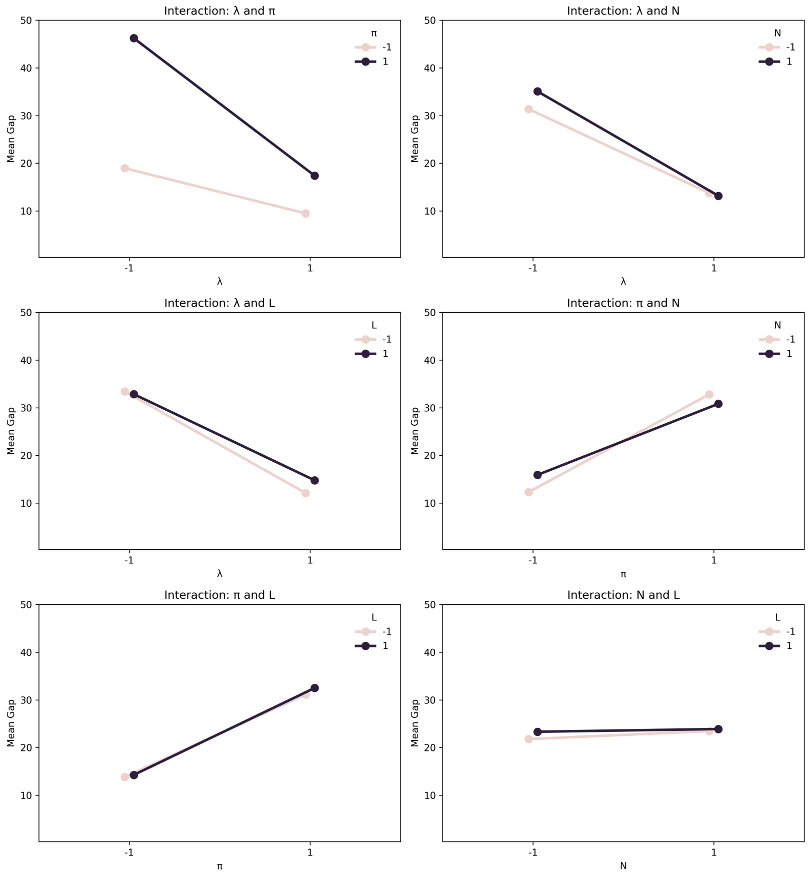


Fig. 2: Two-way interaction plots between λ, π, N, and L.

derivation and supporting probability estimates under varying demand conditions are provided in Appendix A5.

4.3 Expected Demand Threshold for Optimal RH Performance

Although tuning parameters such as N and L improves computational efficiency and contributes to performance consistency, these adjustments alone do not identify the conditions under which the Sc-RHA method reliably produces near-optimal solutions. To address this, we examine the role of expected demand as a key factor influencing solution quality and reliability.

ANOVA revealed that demand characteristics, particularly λ and π, influence the optimality gap. However, they do not directly inform when the Sc-RHA achieves near-optimal behavior. This section examines whether a threshold in expected demand,

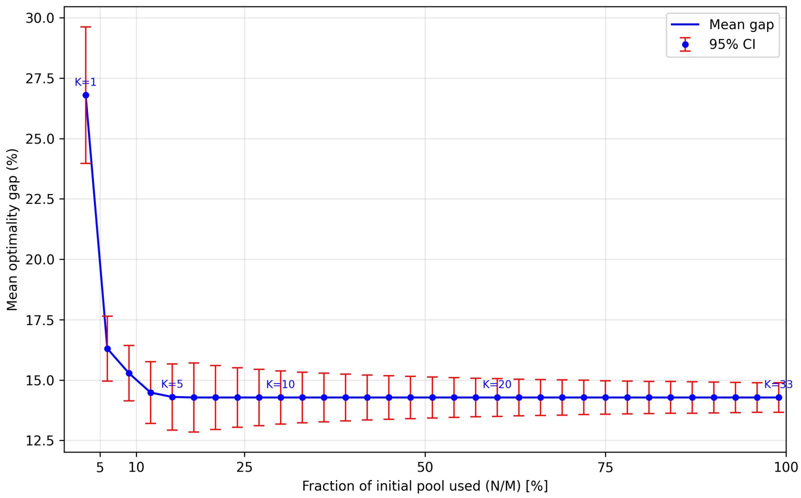
$$E[x] \;=\; \lambda\,(1-\pi),$$

can be identified beyond which the Sc-RHA delivers high-quality solutions with low performance gaps. Establishing such a threshold offers a data-driven criterion for practitioners to assess the suitability of rolling-horizon methods in real-world applications.

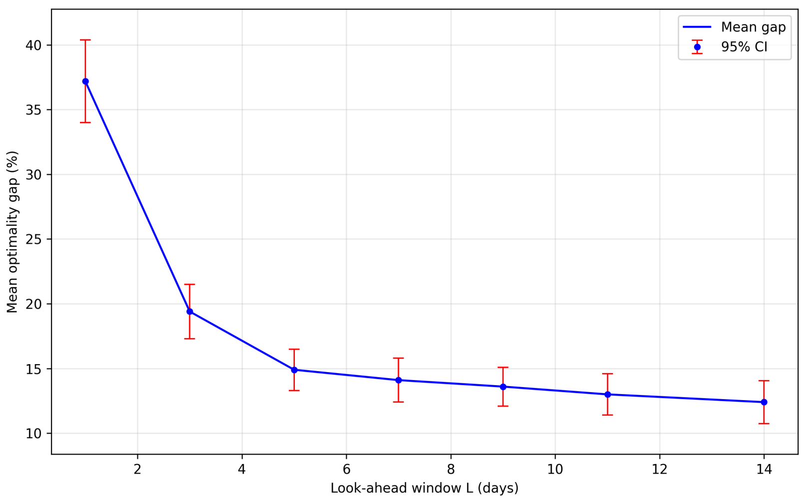
To empirically explore this possibility, the analysis considers how the optimality gap varies across a range of expected demand levels. The following boxplot (Figure 4) depicts the performance gap for different values of $E[x]$.

To facilitate analysis, the 95% confidence intervals (CIs) for the mean optimality gap were computed across four distinct demand groups, categorized based on their expected demand ranges: Very Low (0.05–0.20), Low (0.20–0.30), Medium (0.40–0.60), and High (0.80–1.60). These intervals, along with the respective demand ranges, are summarized in Table 3 and Figure 5.

The analysis of the optimality gap across varying demand groups reveals distinct patterns and statistically significant differences. For the "Very Low" demand group (expected demand between 0.05 and 0.20), the 95% confidence interval (CI) spans from 31.89% to 47.20%. The "Low" group



(a) Scenario-stability plot: Mean optimality gap (%) versus fraction of the initial scenario pool used (N/M).



(b) Performance sensitivity to horizon length: Mean optimality gap (%) versus look-ahead window

Fig. 3: Comparison of scenario-stability and horizon sensitivity results.

(0.20–0.30) exhibits a narrower CI of 36.93% to 46.22% (width = 9.29 points), yet this range falls entirely within the "Very Low" group's interval, indicating no evidence of a performance difference. Statistical tests support this observation: both the Mann–Whitney U test ($p = 0.39$) and Welch's t-test ($p = 0.52$) confirm the lack of significant difference between these two groups.

In contrast, a marked shift occurs at the "Medium" demand level (0.40–0.60), where the CI contracts to 10.73%–17.57%, with no overlap with lower-demand groups. This represents a statistically significant and practically meaningful improvement in Sc-RHA performance. The "High" group (0.80–1.60) continues this trend, with a CI of 7.65%–11.32% (width = 3.67 points). Although the improvement over Medium is smaller ($\approx$ 4.6 percentage points), it is statistically significant ($p < 0.01$), highlighting continued, but diminishing, returns.

The findings establish a performance threshold at $E[x] \geq 0.4$, beyond which the Sc-RHA method consistently produces statistically low and stable optimality gaps. This establishes a data-driven criterion for assessing when rolling-horizon optimization can serve as a feasible alternative to the computationally demanding SDP. The significance of this result lies in the 0.4 threshold, which indicates the method's capacity to manage the irregularities of zero-inflated demand while avoiding

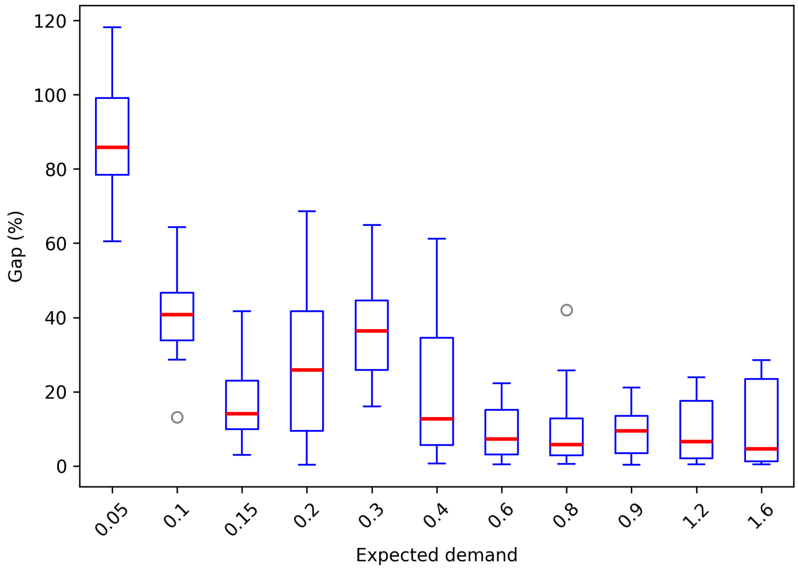


Fig. 4: Gap percentage for expected demands.

Table 3: Demand groups and 95% confidence intervals.

Demand Group	Min Expected Demand	Max Expected Demand	95% CI Lower	95% CI Upper
Very Low	0.05	0.20	31.89	47.20
Low	0.20	0.30	36.93	46.22
Medium	0.40	0.60	10.73	17.56
High	0.80	1.60	7.65	11.32

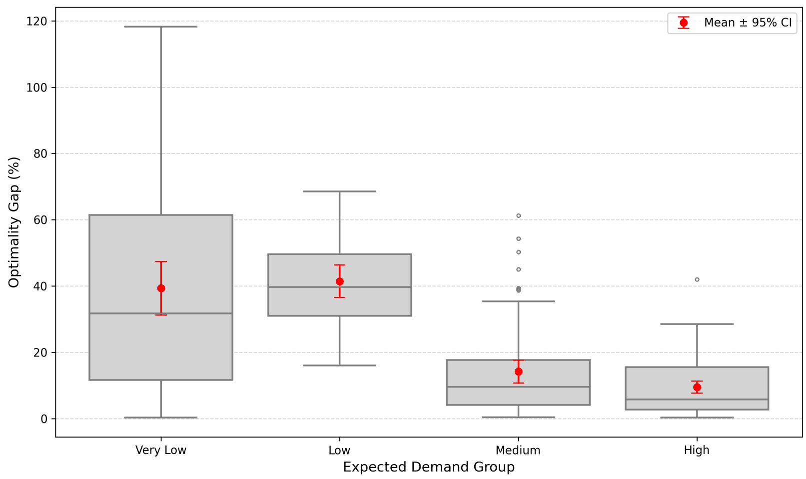


Fig. 5: Demand groups (based on expected demand) and Optimality Gaps(%)

excessive shortages or wastages. The model stabilizes inventory outcomes despite erratic and spotty demand patterns.

A direct comparison between the Sc-RHA and the classical (s, S) policy under identical zero-inflated demand conditions reveals a significant advantage of Sc-RHA: it reduces the effective demand threshold necessary for near-optimal performance. Table 4 illustrates that the (s, S) policy requires an expected daily demand of 1.0 unit/day for reliable results, whereas Sc-RHA reaches comparable performance at considerably lower thresholds. This difference is not merely numerical; it signifies a practical shift in the beneficiaries of a policy. Large regional centres, where demand consistently

Table 4: Comparison of effective demand thresholds for reliable performance under zero-inflated demand conditions using the classical (s, S) policy and the scenario-based rolling horizon (Sc-RHA) approach.

Method	Effective demand threshold $E[x]$	Suitable for
Classical (s, S) Policy	$\geq$ 1.0 units/day	Relatively large centres
Sc-RHA Policy	$\geq$ 0.4 units/day	Small to mid-size centres

surpasses one unit per day, can depend on the (s, S) inventory policy with minimal risk. Smaller hospitals, however, rarely face such demand. In these hospitals, Sc-RHA can be used to address the challenge, enabling safe and efficient LTOWB to function without large wastage costs.

These two methods approach risk management in different ways. The (s, S) rule is ineffective below the 1.0 threshold due to prolonged zero-demand periods that increase wastage, while sudden surges quickly deplete stock and create shortages. Sc-RHA mitigates this by continually adjusting order decisions within a rolling horizon, distributing the risk more evenly and reducing variability across replications. The result is not only a lower threshold for reliable performance but also greater consistency when demand is sparse and unpredictable.

4.4 Sensitivity Analysis of Threshold Demand under Cost Configurations

Of course, assumptions about operational costs influence the demand threshold at which the Sc-RHA achieves good performance. To investigate this relationship, a sensitivity analysis was conducted, and results were compared directly against the (s, S) policy. Table 5 reports the minimum expected demand required for both heuristics to yield solutions that are operationally comparable to the benchmark across six cost configurations that combine three shortage-to-wastage penalty ratios with two levels of fixed ordering costs.

Table 5: Minimum expected demand ($E[x]$) for near-optimal Sc-RHA and (s, S) performance under varying shortage-to-wastage ratios and fixed ordering costs.

Shortage : Wastage Ratio	Fixed Ordering Cost	Sc-RHA $E[x]$ (units/day)	(s, S) $E[x]$ (units/day)
50:1	10	0.1	0.6
50:1	20	0.2	0.8
100:1	10	0.4	1.0
100:1	20	0.8	1.6
200:1	10	0.9	1.4
200:1	20	1.2	1.6

Raising the shortage penalty from 50:1 to 100:1 produces a sharp increase in the Sc-RHA threshold demand, reflecting the system's stronger emphasis on avoiding stockouts and the consequent buildup of inventory buffers at moderate penalty levels. Increasing the penalty further from 100:1 to 200:1 continues to push the threshold upward, but at a slower rate. This tapering effect indicates that system constraints, such as finite storage and fixed order sizes, begin to cap further gains, limiting the marginal benefit of higher penalty ratios.

The influence of ordering cost is asymmetric. Doubling the ordering cost from 10 to 20 raises the Sc-RHA threshold substantially when penalties are moderate, sometimes doubling the effective demand requirement. However, under extreme shortage penalties, this impact weakens; unmet demand is penalized so heavily that stock availability dominates ordering efficiency, reducing the role of fixed cost considerations.

Compared with the classical (s, S) policy, Sc-RHA consistently requires lower demand thresholds across all tested shortage–wastage ratios and fixed ordering costs. The advantage is most pronounced at moderate penalty ratios (100:1), where Sc-RHA performs reliably at roughly half the demand level needed by (s, S) (0.4 vs. 1.0 and 0.8 vs. 1.6 units/day). Under very high penalty ratios (200:1), the gap narrows: Sc-RHA requires 0.9–1.2 units/day compared with 1.4–1.6 units/day for (s, S).

This convergence reflects the fact that when shortage penalties dominate, both policies adopt more aggressive ordering, though Sc-RHA still maintains a consistent advantage.

5 Managerial Insights

The results of this study provide insights for hospital blood banks and healthcare administrators managing perishable blood products under uncertain and zero-inflated demand:

- **Adopt Rolling-Horizon Optimization.**
  Static rules such as the classical (s, S) policy struggle with erratic demand patterns. By re-optimizing inventory decisions at regular intervals, the scenario-based rolling horizon approach (Sc-RHA) reduces both shortages and wastage, offering a more adaptive and resilient method for inventory control, with a moderate level of computational effort.
- **Use Sc-RHA in Irregular Demand Settings.**
  The (s, S) policy performs reliably only when daily demand exceeds one unit. In contrast, Sc-RHA remains effective at demand levels as low as 0.4 units per day, making it suitable for smaller hospitals where demand is sparse and unpredictable.
- **Plan with a One-Week Horizon.**
  A rolling horizon of about five days offers a sound balance between foresight and adaptability. Shorter horizons lead to reactive decisions and increase the likelihood of shortages, while longer horizons add computational burden without yielding notable performance improvements. Managers should therefore employ a rolling five-day forecast when planning LTOWB allocations, updating it regularly to reflect changes in demand and inventory.
- **Keep Scenario Sampling Practical.**
  Around 15% of all possible scenarios were found to be sufficient to capture the essential variability of zero-inflated demand while maintaining computational efficiency.
- **Stabilize Demand Where Possible.**
  Hospitals can strengthen the performance of Sc-RHA by coordinating blood use across departments or scheduling elective procedures to smooth consumption patterns. These practices elevate effective demand, reduce variability, and further improve inventory efficiency.

Taken together, the results suggest Sc-RHA is practical as an inventory management tool. Beyond large regional centres, where traditional (s, S) rules remain adequate, Sc-RHA enables smaller hospitals to manage low-volume, erratic blood demand safely and cost-effectively, without the computational burden of an SDP. This shift has important implications for expanding access to LTOWB programs and improving patient care in settings where demand is sparse but avoiding shortages is critical.

6 Conclusion

This study demonstrated the effectiveness of rolling-horizon and scenario-based methods (Sc-RHA) in blood inventory management under uncertain and intermittent demand conditions. Building on earlier research, which identified an optimal demand threshold of approximately 1.0 unit/day for traditional (s, S) policies, the current study demonstrates that Sc-RHA policies can lower this operational threshold to around 0.4 units/day. This shift expands the range of hospitals that can operate safe and efficient LTOWB programs. Experimental evidence further shows that mean demand and zero-inflation probability are the dominant drivers of performance, while planning horizon length and scenario count have secondary effects. Sensitivity analyses confirm that Sc-RHA is reliable across varied shelf-life distributions and adaptable to hospital-specific cost structures.

This work advances both theory and practice by demonstrating how rolling-horizon optimization can be embedded in blood inventory management. While grounded in the case of LTOWB, the approach generalizes to other perishable medical products such as platelets, vaccines, and temperature-sensitive pharmaceuticals. The integration of rolling-horizon optimization with zero-inflated demand models offers a practical, efficient, and computationally tractable alternative to exact stochastic dynamic programming, significantly broadening the applicability of heuristic inventory policies. Future research should extend this framework to multi-hospital coordination, where sharing inventories and information could further enhance efficiency and resilience across critical healthcare supply chains.

Acknowledgments

The authors gratefully acknowledge Vancouver General Hospital and Canadian Blood Services for providing access to empirical blood inventory data used in this study. Their support was essential for validating the models and analyses presented in the paper.

Statements and Declarations

The authors declare that they have no conflicts of interest.

Appendix A

A1 Linearized Formulation

To linearize the nonlinear term $\lceil x_t \, p_{t,\ell} \rceil$, we define a set of precomputed integers:

$$a_{t,j,\ell} := \lceil p_{t,\ell} \, j \rceil, \quad \forall t, \forall j, \forall \ell.$$

These coefficients represent the number of units expected to arrive with ℓ days of shelf-life when an order of size j is placed in period t. By computing $a_{t,j,\ell}$ in advance, we eliminate the ceiling operation and preserve linearity in the formulation.

To implement this linearization, we introduce auxiliary variables $\lambda_{t,j}$ which serve as convex combination weights over the discrete order size breakpoints $j \in J$. For each period t, the set $\{\lambda_{t,j}\}$ forms a Special Ordered Set of type 2 (SOS2), a well-established structure that restricts the solution to at most two adjacent nonzero weights. This allows both x_t and the corresponding arrivals $i_{t,\ell}$ to be reconstructed as linear combinations of known breakpoints and integer coefficients. The SOS2 formulation ensures integrality without introducing binary variables.

A1.1 Nomenclature

Indices:

t	Planning periods, $t \in \{1, \ldots, T\}$.
ℓ	Remaining shelf-life upon arrival, $\ell \in \{1, \ldots, S\}$.
j	Order size breakpoints, $j \in \{1, \ldots, X\}$.
m	FIFO threshold index, $m \in \{1, \ldots, S-1\}$.

Parameters:

S	Maximum shelf-life (periods).
X	Maximum admissible order quantity.
C_o, C_s, C_w	Unit costs of ordering, shortage, and wastage, respectively.
d_t	Demand in period t.
$p_{t,\ell}$	Probability that an order placed in t arrives with ℓ periods of remaining shelf-life.
y_ℓ^{init}	Initial inventory available at the start of period 1 for $\ell \geq 2$.
$a_{t,j,\ell}$	Precomputed integer arrivals: $a_{t,j,\ell} = \lceil p_{t,\ell} j \rceil$.

Decision Variables:

x_t	Units ordered at the start of period t.
$y_{t,\ell}$	End-of-period t inventory with ℓ periods of shelf-life.
$r_{t,\ell}$	Units used in t from class ℓ to satisfy demand.
$i_{t,\ell}$	Units received in t with ℓ periods of shelf-life.
w_t	Wastage at the end of period t.
s_t	Shortage at the end of period t.
$\lambda_{t,j}$	SOS2 convex weights for linearizing ceiling.

A1.2 Mathematical Model

Formulation (1)–(12) specifies the linearized MILP model for the LTOWB inventory problem.

$$\min \; Z = \sum_{t=1}^{T} \left(C_o x_t + C_s s_t + C_w w_t \right) \tag{1}$$

$$\sum_{j=1}^{X} \lambda_{t,j} = 1 \qquad \forall t \tag{2}$$

$$x_t = \sum_{j=1}^{X} j \, \lambda_{t,j} \qquad \forall t \tag{3}$$

$$i_{t,\ell} = \sum_{j=1}^{X} a_{t,j,\ell}\,\lambda_{t,j} \qquad \forall t,\ \forall \ell \tag{4}$$

$$\sum_{\ell=1}^{S} r_{t,\ell} + s_t = d_t \qquad \forall t \tag{5}$$

$$0 \leq r_{t,\ell} \leq y_{t-1,\ell+1} + i_{t,\ell} \qquad \forall t,\ \forall \ell \tag{6}$$

$$\sum_{k=1}^{m} r_{t,k} \leq \sum_{k=1}^{m} (y_{t-1,k+1} + i_{t,k}) \qquad \forall t,\ \forall m = 1, \ldots, S-1 \tag{7}$$

$$\sum_{k=1}^{m} r_{t,k} \geq d_t - \sum_{k=m+1}^{S} (y_{t-1,k+1} + i_{t,k}) \forall t,\ \forall m = 1, \ldots, S-1 \tag{8}$$

$$y_{t,\ell} = y_{t-1,\ell+1} + i_{t,\ell} - r_{t,\ell} \qquad \forall t,\ \ell = 2, \ldots, S \tag{9}$$

$$w_t = y_{t-1,2} + i_{t,1} - r_{t,1} \qquad \forall t \tag{10}$$

$$y_{0,\ell} = y_{\ell}^{\text{init}}, \quad y_{0,1} = 0 \qquad \ell = 2, \ldots, S \tag{11}$$

$$x_t,\ r_{t,\ell},\ y_{t,\ell},\ s_t,\ w_t \in \mathbb{Z}_{\geq 0}, \quad \lambda_{t,j} \in \mathbb{R}_{\geq 0} \qquad \forall t,\ \forall \ell,\ \forall j \tag{12}$$

A2 Scenario Pool Size Justification

A2.1 Notation

Symbols:

K	Number of sampled scenarios.
ε	Maximum allowable violation probability of chance constraints.
θ	Risk level for the probabilistic guarantee ($1-\theta$ is the confidence).
d	Support rank, i.e., effective number of independent constraints affected by uncertainty in the LP-relaxed model.
$V(\hat{x})$	Violation probability of the solution $\hat{x}$.
$B(K,\varepsilon,d)$	Binomial tail bound: $B(K,\varepsilon,d) = \sum_{i=0}^{d-1} \binom{K}{i} \varepsilon^i (1-\varepsilon)^{K-i}$.

A2.2 Formulation and Bounds

From Calafiore and Campi (2006) [39], the violation probability satisfies:

$$\Pr\{V(\hat{x}) > \varepsilon\} \leq B(K,\varepsilon,d). \tag{13}$$

Two sources of uncertainty exist in the model:

1. Demand balance: $\sum_{\ell=1}^{S} r_{t,\ell} + s_t = d_t$ (one independent uncertain direction per period).
2. Incoming shelf-life allocation: $\{p_{t,\ell}\}$ with $\sum_{\ell} p_{t,\ell} = 1$ (at most $S-1$ independent directions per period).

Over a rolling horizon of length L, the support rank is conservatively bounded by

$$d \leq L + (S-1). \tag{14}$$

FIFO and flow-balance equalities only redistribute inventory ages and therefore reduce effective dimensionality, so (14) is conservative. For $L = 14$ and $S = 14$, this gives

$$d \leq 27.$$

For $\varepsilon = 0.05$ and $\theta = 0.05$, the smallest K satisfying $B(K,\varepsilon,d) \leq \theta$ is

$$K_{\min}(d = 27) = 717.$$

Thus, with $d \leq 27$ under the implemented horizon and shelf-life structure, the theoretical minimum is $K_{\min} = 717$. Choosing $K = 1000$ strictly exceeds this requirement, preserves the probabilistic guarantee, and future-proofs the study against modest increases in support rank.

A3 Stochastic Dynamic Programming Formulation

A3.1 Nomenclature

Indices:

T	Stages (periods), $t \in \{1, \dots, T\}$.
I_k	Inventory level with k periods of remaining shelf life, indexed by i_k, for $k = 1, \dots, K$.
S	State space $s = [i_1, i_2, \dots, i_K]$ of all configurations across remaining shelf lives.
B	Batch ordering set, indexed by x.
$A_s \subseteq B$	Action/decision set for state s.

Parameters:

Cap	Capacity of storage.
D	Demand (modeled as ZIP distribution).
α	Fixed ordering cost.
β	Per-unit ordering cost.
θ	Per-unit shortage cost.
ϑ	Per-unit holding cost.
ω	Per-unit wastage cost.

Functions:

$O(\cdot)$	Ordering cost function.
$S(\cdot)$	Shortage cost function.
$H(\cdot)$	Holding cost function.
$W(\cdot)$	Wastage cost function.
$F_t(\cdot)$	Expected total cost function at stage t.

A3.2 Formulated Cost Function

For the final stage ($t = T$):

$$F_T(s,x) = O(x) + \mathbb{E}_D\big[S(s,x,D) + H(s,x,D) + W(s,x,D)\big], \tag{15}$$
$$F_T^*(s) = \min_{x \in A_s} F_T(s,x). \tag{16}$$

For $t = T-1, T-2, \dots, 1$:

$$F_t(s,x) = O(x) + \mathbb{E}_D\big[S(s,x,D) + H(s,x,D) + W(s,x,D) + F_{t+1}^*(s')\big], \tag{17}$$
$$F_t^*(s) = \min_{x \in A_s} F_t(s,x), \tag{18}$$

where s' is the state after action x and realized demand d.

The feasible action set:

$$A_s = \{\, x \in B \mid x + \textstyle\sum_{k=1}^{K} i_k - \min(D) \leq \text{Cap} \,\}. \tag{19}$$

A3.3 Cost Functions

$$O(x) = \begin{cases} 0, & x = 0, \\ \alpha + \beta x, & x > 0, \end{cases} \tag{20}$$

$$S(s,x,D) = \begin{cases} 0, & x + \sum_{k=1}^{K} i_k \geq D, \\ \theta\big(D - (x + \sum_{k=1}^{K} i_k)\big), & x + \sum_{k=1}^{K} i_k < D, \end{cases} \tag{21}$$

$$H(s,x,D)=\begin{cases}0, & x+\sum_{k=1}^{K} i_k \le D,\\ \vartheta\big(x+\sum_{k=1}^{K} i_k - D\big), & x+\sum_{k=1}^{K} i_k > D,\end{cases} \tag{22}$$

$$W(s,x,D)=\begin{cases}0, & i_1 \le D,\\ \omega(i_1 - D), & i_1 > D.\end{cases} \tag{23}$$

A3.4 Optimal Ordering Policy

$$x_t^*(s)=\arg\min_{x\in A_s} F_t(s,x), \qquad \pi^*=\big[x_1^*(\cdot),\, x_2^*(\cdot),\, \ldots,\, x_T^*(\cdot)\big]. \tag{24}$$

A3.5 Overview of the SDP Components

- **Stages and States:** The model operates over discrete time stages (e.g., days), denoted as $T=\{1,2,\ldots,T\}$. The state at each stage is represented by the vector $s=[i_1,i_2,\ldots,i_K]$, where i_k indicates the inventory level with k-periods of remaining shelf-life. The set S encompasses all possible configurations of inventory levels across products with different remaining shelf lives. Thus, each specific state $s\in S$ represents a distinct configuration of inventory distributed across items with varying shelf lives.
- **Action Set:** The action set A_s comprises possible batch order quantities x from a predefined set $B=\{0,1,2,\ldots\}$, constrained by storage capacity. Each action determines incoming inventory available to meet future demand.
- **Demand Model:** Demand is modeled exclusively using the Zero-Inflated Poisson (ZIP) distribution, which is particularly suitable for LTOWB. While the ZIP distribution is the focus here, this demand model can be modified or extended to follow a Poisson or other probabilistic distributions to accommodate different demand patterns.

A3.6 Cost Functions

- **Ordering Cost $O(x)$:** Incurred when placing an order, including fixed and variable costs based on order quantity.
- **Shortage Cost $S(s,x,D)$:** A penalty per unit of unmet demand when total available inventory is insufficient.
- **Holding Cost $H(s,x,D)$:** Represents storage costs when inventory exceeds demand; this cost is negligible for the current problem.
- **Wastage Cost $W(s,x,D)$:** Applied for units unused within their remaining shelf-life.

A3.7 Expected Total Cost Calculation

For each action x, the model calculates the expected total cost $F_T(s,x)$.

A3.8 Optimal Decision Policy

To minimize the expected total cost at each stage, the model selects an optimal action $x_t^*(s)$ for each state by minimizing $F_T(s,x)$ across all stages and states.

A4 Tukey Test tables

Tukey HSD pairwise comparisons for the scenario fractions (f) and look-ahead window (L). All p-values indicate no statistically significant differences between levels at $\alpha=0.05$.

Table 1: Tukey HSD p-values for Scenario Fractions (f)

Level 1	Level 2	p-value
25%	15%	0.94
50%	15%	0.97
75%	15%	0.96
100%	15%	0.95

Table 2: Tukey HSD p-values for Look-ahead Window (L)

Level 1	Level 2	p-value
7	5	0.50
9	5	0.49
11	5	0.29
14	5	0.13

A5 Supporting Probability Estimates under ZIP Demand

The probability that at least one unit is needed within an L-day window: The probability that at least one unit is needed within an L-day window:

$$P(\text{consumed} \leq L) = 1 - \left[\pi + (1-\pi)e^{-\lambda}\right]^L. \tag{25}$$

Table 3: Probability that a unit is consumed within L days under varying demand.

Exp. demand	λ	π	$P(\leq 5)$	$P(\leq 7)$	$P(\leq 14)$
0.05	0.5	0.9	18%	24.5%	43%
0.6	1.5	0.6	84.4%	92.6%	99.5%
2.7	3.0	0.1	99.9%	100%	100%

The key insight is that this probability increases faster when demand is more frequent and stable, i.e., when π is small and λ is large. While the formula itself does not directly use the product $\lambda(1-\pi)$, this value serves as a useful summary of expected daily demand and strongly correlates with the rate at which consumption likelihood increases over L days. For example, when expected daily demand exceeds 0.6, the probability of at least one demand occurrence within five days typically exceeds 80%. This indicates that most relevant consumption information is already captured in a short planning window. Extending L further yields minimal gain, while increasing model size and computational cost.

Conversely, under very sparse demand (e.g., $\lambda = 0.5$, $\pi = 0.9$), even a 14-day window yields only a 43% chance of demand, highlighting that longer foresight does not necessarily improve decision quality when uncertainty dominates. This dual behaviour (rapid saturation in high-demand settings and stagnation under low demand) explains the diminishing effect observed in performance metrics around $L = 5$ and aligns with the $\lambda \times L$ interaction observed in the ANOVA results.

References

[1] Blood transfusion safety. https://www.who.int/news-room/fact-sheets/detail/blood-safety-and-availability. Accessed October 2025

[2] Ngatuvai, M., Zagales, I., Sauder, M., Andrade, R., Santos, R.G., Bilski, T., Kornblith, L., Elkbuli, A.: Outcomes of transfusion with whole blood, component therapy, or both in adult civilian trauma patients: a systematic review and meta-analysis. Journal of Surgical Research **287**, 193–201 (2023)

[3] Safavi, S., Blake, J.T.: Inventory management of perishables under zero-inflated poisson demand. Operations Research, Data Analytics and Logistics **45**, 200482 (2025) https://doi.org/10.1016/j.ordal.2025.200482

[4] Clay, N.M., Abbasi, B., Eberhard, A., Hearne, J.: On the volatility of blood inventories. International Transactions in Operational Research **25**(1), 215–242 (2018)

[5] Gürler, Ü., Özkaya, B.Y.: Analysis of the (s, s) policy for perishables with a random shelf life. IIe Transactions **40**(8), 759–781 (2008)

[6] Prastacos, G.P.: Blood inventory management: an overview of theory and practice. Management science **30**(7), 777–800 (1984)

[7] Schmidt, C.P., Nahmias, S.: (s- 1, s) policies for perishable inventory. Management Science **31**(6), 719–728 (1985)

[8] Barron, Y.: A state-dependent perishability (s, s) inventory model with random batch demands. Annals of Operations Research **280**(1), 65–98 (2019)

[9] Kouki, C., Jemaï, Z., Minner, S.: A lost sales (r, q) inventory control model for perishables with fixed lifetime and lead time. International Journal of Production Economics **168**, 143–157 (2015)

[10] Haijema, R., Wal, J., Dijk, N.M.: Blood platelet production: Optimization by dynamic programming and simulation. Computers & Operations Research **34**(3), 760–779 (2007)

[11] Abdulwahab, U., Wahab, M.: Approximate dynamic programming modeling for a typical blood platelet bank. Computers & Industrial Engineering **78**, 259–270 (2014)

[12] Abouee-Mehrizi, H., Mirjalili, M., Sarhangian, V.: Platelet inventory management with approximate dynamic programming. INFORMS Journal on Computing (2025)

[13] Farrington, J., Li, K., Wong, W.K., Utley, M.: Going faster to see further: Gpu-accelerated value iteration and simulation for perishable inventory control using jax. arXiv preprint arXiv:2303.10672 (2023)

[14] Ortega, G., Hendrix, E.M., García, I.: A cuda approach to compute perishable inventory control policies using value iteration. The Journal of Supercomputing **75**(3), 1580–1593 (2019)

[15] Powell, W.B.: Perspectives of approximate dynamic programming. Annals of Operations Research **241**(1), 319–356 (2016)

[16] Campi, M.C., Garatti, S.: The exact feasibility of randomized solutions of uncertain convex programs. SIAM Journal on Optimization **19**(3), 1211–1230 (2008)

[17] Chaudhary, V., Kulshrestha, R., Routroy, S.: State-of-the-art literature review on inventory models for perishable products. Journal of Advances in Management Research **15**(3), 306–346 (2018)

[18] Osorio, A.F., Brailsford, S.C., Smith, H.K., Forero-Matiz, S.P., Camacho-Rodríguez, B.A.: Simulation-optimization model for production planning in the blood supply chain. Health care management science **20**(4), 548–564 (2017)

[19] Dalalah, D., Bataineh, O., Alkhaledi, K.A.: Platelets inventory management: A rolling horizon sim–opt approach for an age-differentiated demand. Journal of Simulation (2019)

[20] Shih, H., Rajendran, S.: Stochastic inventory model for minimizing blood shortage and outdating in a blood supply chain under supply and demand uncertainty. Journal of Healthcare Engineering **2020**(1), 8881751 (2020)

[21] Dural-Selcuk, G., Rossi, R., Kilic, O.A., Tarim, S.A.: The benefit of receding horizon control: Near-optimal policies for stochastic inventory control. Omega **97**, 102091 (2020)

[22] Mohammadi, N., Seyedi, S.H., Farhadi, P., Shahmohamadi, J., Ganjeh, Z.A., Salehi, Z.: Development of a scenario-based blood bank model to maximize reducing the blood wastage. Transfusion Clinique et Biologique **29**(1), 16–19 (2022)

[23] Hosseini-Motlagh, S.-M., Samani, M.R.G., Faraji, M.: Dynamic optimization of blood collection strategies from different potential donors using rolling horizon planning approach under uncertainty. Computers & Industrial Engineering **188**, 109908 (2024)

[24] Schildbach, G., Morari, M.: Scenario-based model predictive control for multi-echelon supply chain management. European Journal of Operational Research **252**(2), 540–549 (2016)

[25] Braun, M.W., Rivera, D.E., Flores, M., Carlyle, W.M., Kempf, K.G.: A model predictive control framework for robust management of multi-product, multi-echelon demand networks. Annual Reviews in Control **27**(2), 229–245 (2003)

[26] Conte, G., Pennesi, P.: Inventory control by model predictive control methods. IFAC Proceedings Volumes **38**(1), 85–90 (2005)

[27] Maestre, J., Fernández, M., Jurado, I.: An application of economic model predictive control to inventory management in hospitals. Control Engineering Practice **71**, 120–128 (2018)

[28] Fernandez, M.I., Chanfreut, P., Jurado, I., Maestre, J.M.: A data-based model predictive decision support system for inventory management in hospitals. IEEE journal of biomedical and health informatics **25**(6), 2227–2236 (2020)

[29] Hipólito, T., Nabais, J.L., Carmona-Benítez, R., Botto, M.A., Negenborn, R.R.: A centralised model predictive control framework for logistics management of coordinated supply chains of perishable goods. International Journal of Systems Science: Operations & Logistics **9**(1), 1–21 (2022)

[30] Lejarza, F., Venkatesan, S., Baldea, M.: Rolling horizon product quality estimation and online optimisation for supply chain management of perishable inventory. International Journal of Production Research **63**(10), 3709–3732 (2025)

[31] Dehghani, M., Abbasi, B., Oliveira, F.: Proactive transshipment in the blood supply chain: A stochastic programming approach. Omega **98**, 102112 (2021)

[32] Mohammadi, T., Kheiri, S., Sedehi, M.: Analysis of blood transfusion data using bivariate zero-inflated poisson model: A bayesian approach. Computational and mathematical methods in medicine **2016**(1), 7878325 (2016)

[33] Ding, X.: Demand estimation and optimal policies in lost sales inventory systems. PhD thesis, University of British Columbia (2002)

[34] Ünlü, Y., *et al.*: Zero-modified distributions for inventory control under intermittent demand. In: IISE Annual Conference. Proceedings, p. 1 (2011). Institute of Industrial and Systems Engineers (IISE)

[35] Costantino, F., Di Gravio, G., Patriarca, R., Petrella, L.: Spare parts management for irregular demand items. Omega **81**, 57–66 (2018)

[36] Finco, S., Battini, D., Converso, G., Murino, T.: Applying the zero-inflated poisson regression in the inventory management of irregular demand items. Journal of industrial and production engineering **39**(6), 458–478 (2022)

[37] Belotti, P., Kirches, C., Leyffer, S., Linderoth, J., Luedtke, J., Mahajan, A.: Mixed-integer nonlinear optimization. Acta Numerica **22**, 1–131 (2013)

[38] Williams, H.P.: Model Building in Mathematical Programming. John Wiley & Sons, ??? (2013)

[39] Calafiore, G.C., Campi, M.C.: The scenario approach to robust control design. IEEE Transactions on automatic control **51**(5), 742–753 (2006)

[40] Kaut, M., Wallace, S.W.: Evaluation of Scenario-generation Methods for Stochastic Programming. Humboldt-Universität zu Berlin Berlin, ??? (2003)

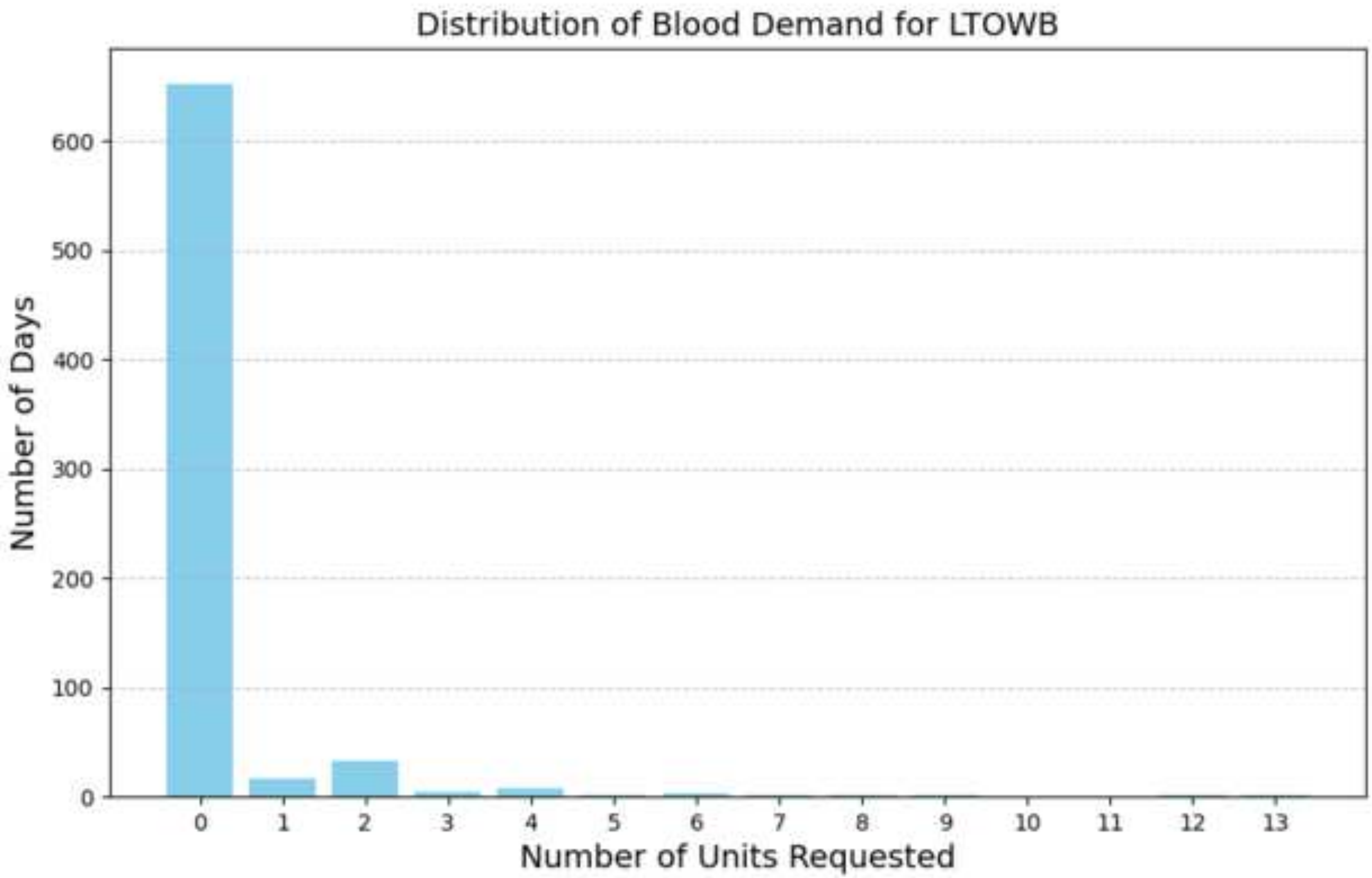
[41] Campi, M.C., Garatti, S.: A sampling-and-discarding approach to chance-constrained optimization: feasibility and optimality. Journal of optimization theory and applications **148**(2), 257–280 (2011)

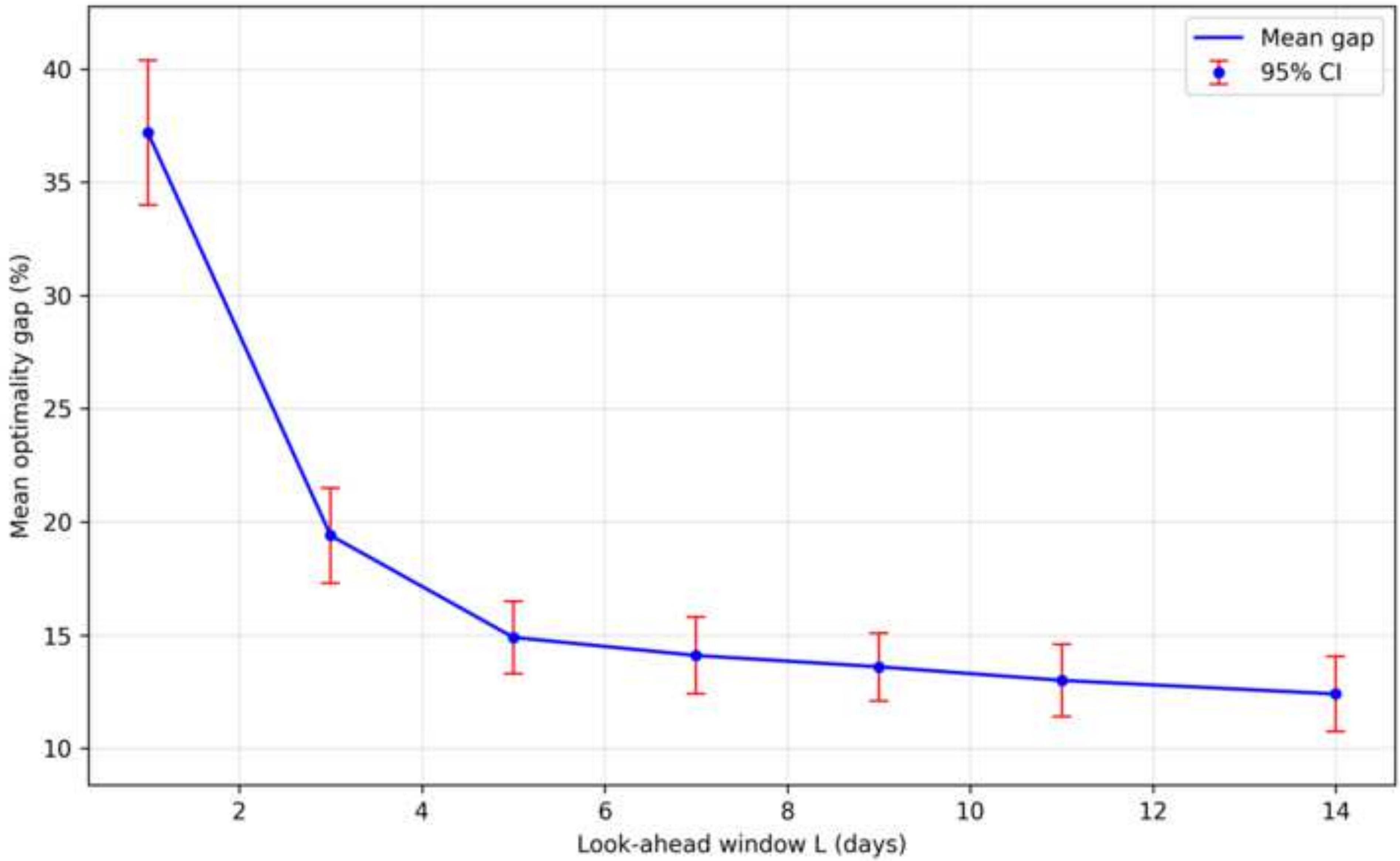
[42] Gaggero, M., Tonelli, F.: Optimal control of distribution chains for perishable goods. IFAC-PapersOnLine **48**(3), 1049–1054 (2015) https://doi.org/10.1016/j.ifacol.2015.06.222 . 15th IFAC Symposium onInformation Control Problems inManufacturing

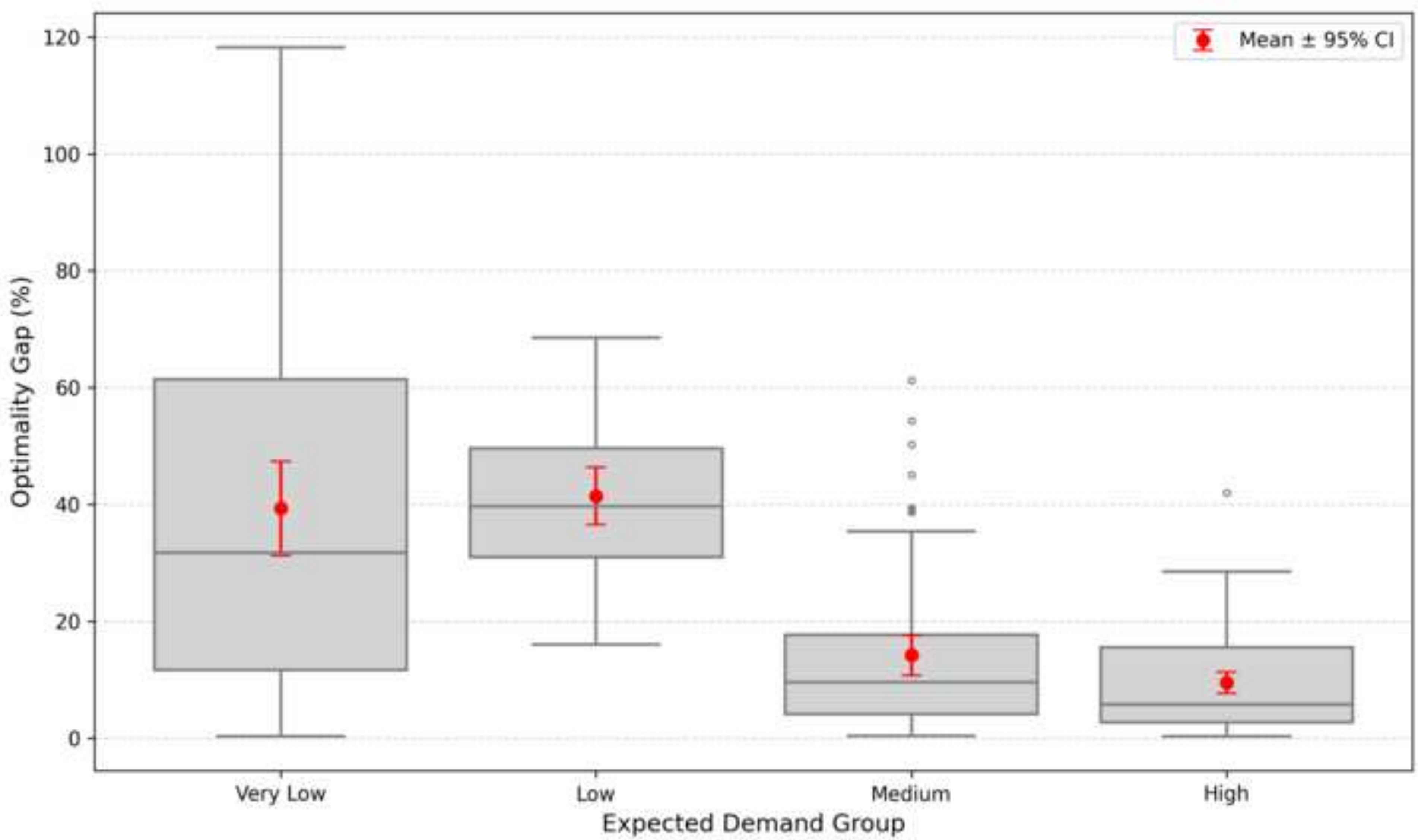
[43] Stanger, S.H.W., Yates, N., Wilding, R., Cotton, S.: Blood inventory management: Hospital best practice. Transfusion Medicine Reviews **26**(2), 153–163 (2012) https://doi.org/10.1016/j.tmrv.2011.09.001

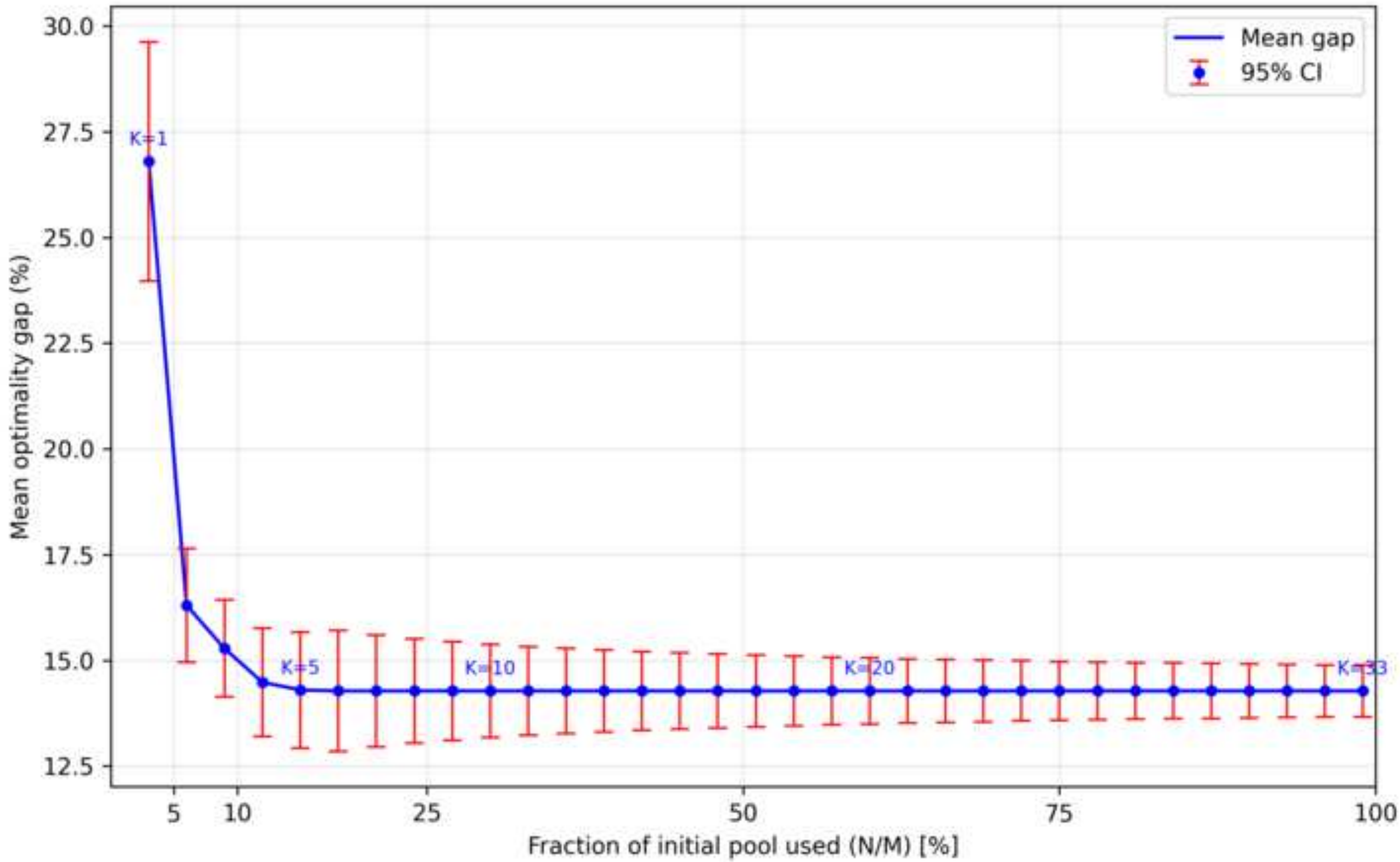
[44] Blake, J., McTaggart, K., Hardy, M.: Modelling a blood distribution network in the prairies with a generic simulation framework. INFOR: Information Systems and Operational Research **53**(4), 194–210 (2015) https://doi.org/10.3138/infor.53.4.194 https://doi.org/10.3138/infor.53.4.194

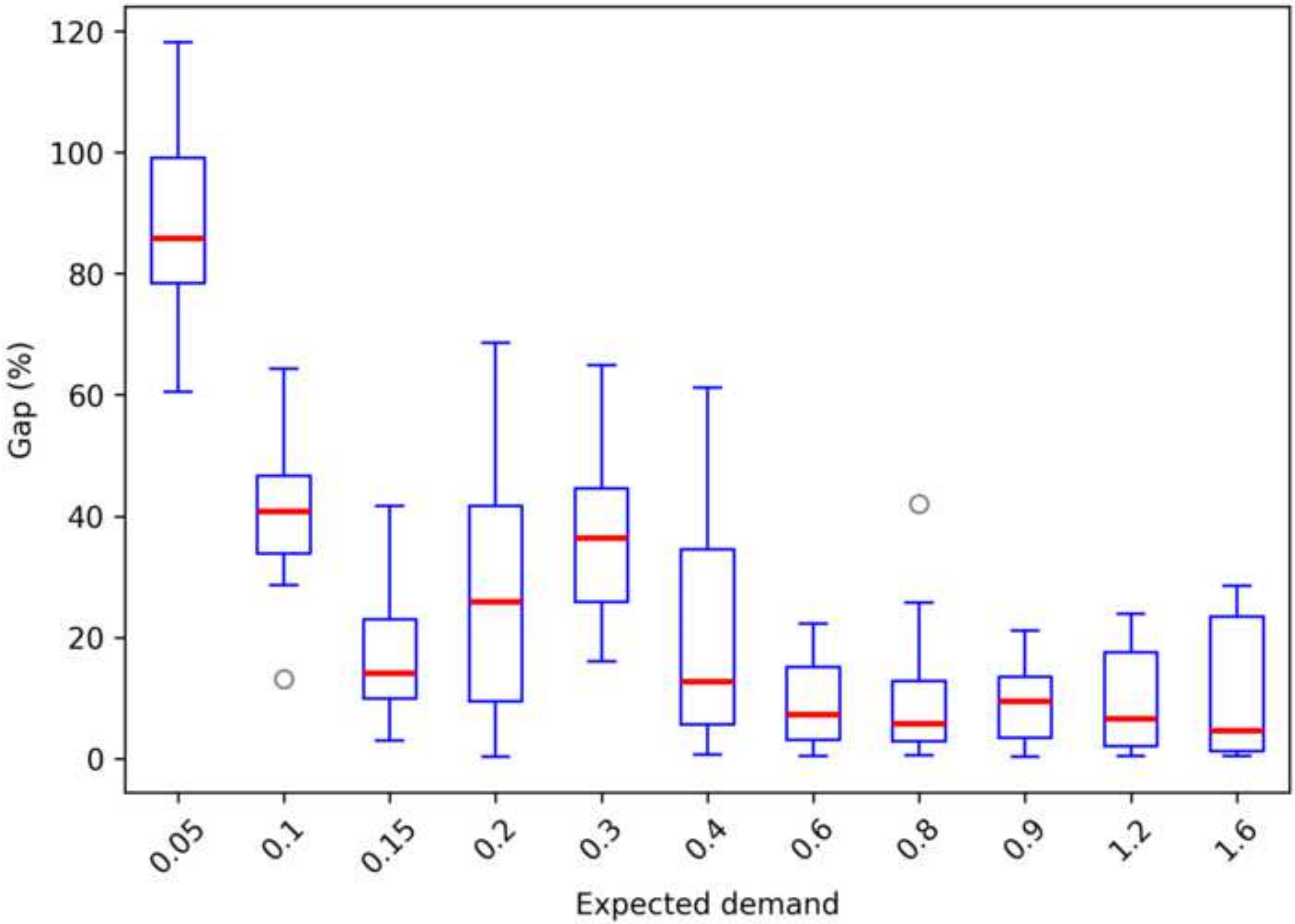
[45] Glomb, L., Liers, F., Rösel, F.: A rolling-horizon approach for multi-period optimization. European Journal of Operational Research **300**(1), 189–206 (2022) https://doi.org/10.1016/j.ejor.2021.07.043

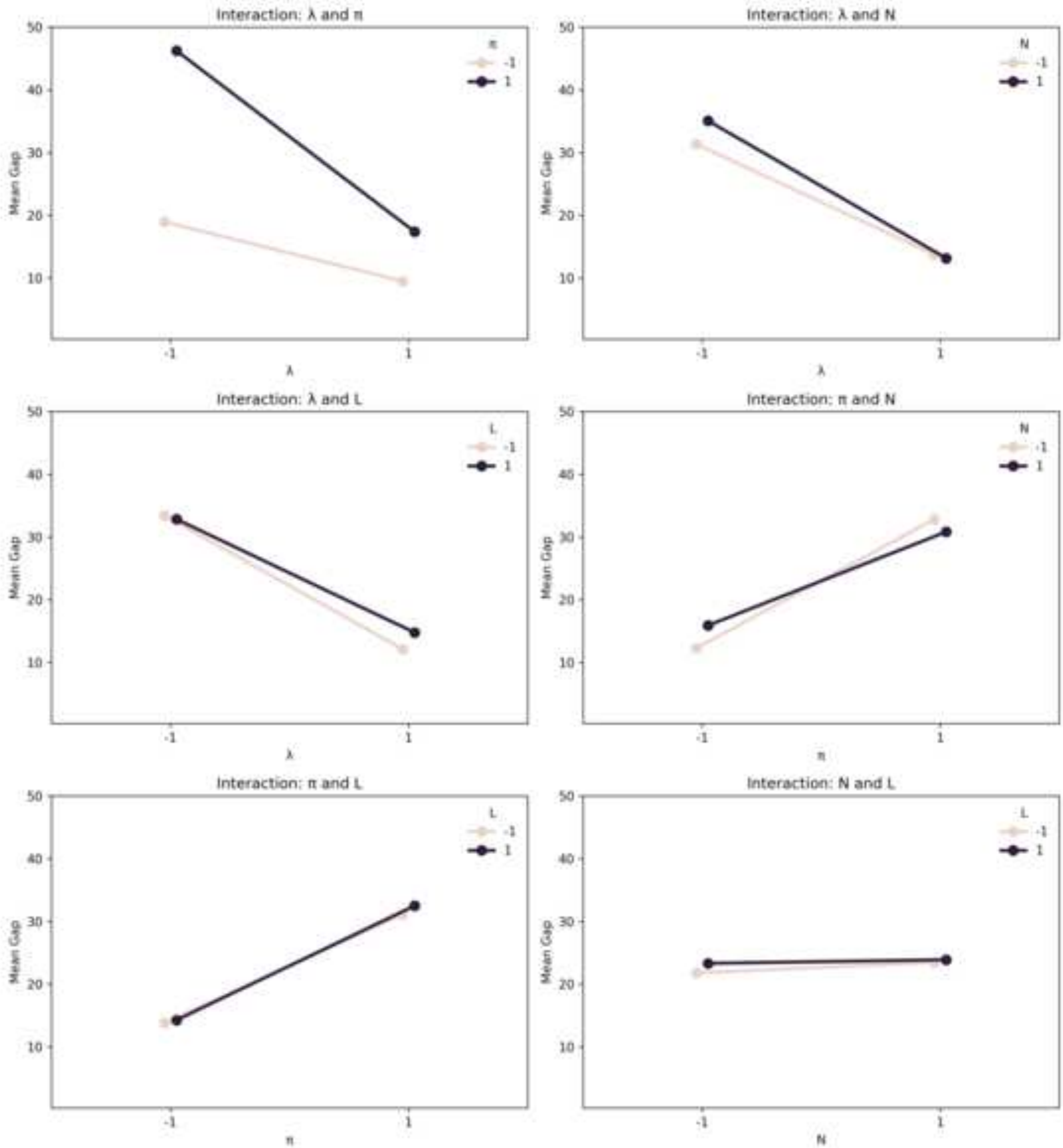

Distribution of Blood Demand for LTOWB
Number of Days
0
100
200
300
400
500
600
0
1
2
3
4
5
6
7
8
9
10
11
12
13
Number of Units Requested


Mean gap
95% CI
40
35
30
25
20
15
10
Mean optimality gap (%)
2
4
6
8
10
12
14
Look-ahead window L (days)


Mean ± 95% CI
Optimality Gap (%)
120
100
80
60
40
20
0
Very Low
Low
Medium
High
Expected Demand Group


Mean gap
95% CI
K=1
K=5
K=10
K=20
K=33
Mean optimality gap (%)
30.0
27.5
25.0
22.5
20.0
17.5
15.0
12.5
5
10
25
50
75
100
Fraction of initial pool used (N/M) [%]


Gap (%)
0
20
40
60
80
100
120
0.05
0.1
0.15
0.2
0.3
0.4
0.6
0.8
0.9
1.2
1.6
Expected demand


Interaction: λ and π
Interaction: λ and N
Interaction: λ and L
Interaction: π and N
Interaction: π and L
Interaction: N and L
Mean Gap
λ
π
N
L
-1
1
10
20
30
40
50

Click here to access/download
Supplementary Material
fig.eps