SCIENTIFIC OASIS

Spectrum of Operational Research

Journal homepage: www.sor-journal.org
ISSN: 3042-1470

Ethical AI for Crisis Response: A Dynamic Policy Orchestrator for Prescriptive Resilience in Pharmaceutical Supply Chains

Jamal Musbah[1,2]; Ibrahim Badi[1,*]

1 School of Engineering and Applied Sciences, Libyan Academy, Misrata, Libya
2 College of Technical Sciences, Bani Walid, Libya

ARTICLE INFO

Article history:
Received 7April 2026
Received in revised form 21 July 2026
Accepted 10 September 2026
Available online 13 September 2026

Keywords:
Adaptive Supply Chain; Machine Learning; Pharmaceutical Logistics; Risk Management; Humanitarian Operations; Ethical AI

ABSTRACT

Pharmaceutical supply chains in volatile regions must balance cost efficiency with continuity of life-saving supplies under changing demand, fuel, and security conditions. This study develops a hybrid AI-driven Policy Orchestrator that adapts target inventory days, rationing limits, and delivery radius to the observed operating context. A simulation-based framework generated 1,000 Latin Hypercube Sampling (LHS) contexts, each evaluated under three candidate regimes, yielding 3,000 context–policy records and 1,000 optimal labels. A multi-output Random Forest model learned the mapping from context to policy parameters, achieving R^2 values of 0.93, 0.82, and 0.77 for target inventory days, rationing, and delivery radius, respectively. The orchestrator was evaluated over a 90-day steady-state horizon and three separate 20-day shock episodes. In steady state, mean total cost increased by 12.3%, while routing distance and stockout penalty decreased descriptively by 6.9% and 15.0%, respectively; neither reduction was statistically significant. During the Security Crisis, mean modeled danger-node traversals decreased from 40.75 to zero per replication, accompanied by a mean increase of 13,618.5 USD in stockout penalties per replication. During the Fuel Price Surge, the delivery radius contracted from 20.0 to 14.2 km. The findings support context-aware, safety-constrained decision support while highlighting key trade-offs and the need for field validation.

1. Introduction

The distribution of essential pharmaceuticals in conflict-affected and highly volatile regions presents a fundamental dichotomy for supply chain managers: reconciling the exigencies of commercial cost-efficiency with the humanitarian imperative of reliable access [1]. In environments characterized by deep uncertainty—such as geopolitical instability, sudden epidemiological outbreaks, or severe macroeconomic shocks—supply chains operate under continuous duress [2]. Traditional supply chain optimization models, deeply rooted in the minimization of Total Landed Cost (TLC), heavily rely on static, deterministic, or mildly stochastic assumptions regarding demand patterns and transport accessibility [3]. However, when deployed in extreme crisis environments, these static policies can become materially brittle. A rigid adherence to cost-minimization during a sudden crisis can lead to excessive resource waste or severe pharmaceutical shortages [4].

[*] *Corresponding author.*
E-mail address: i.badi@lam.edu.ly

https://doi.org/10.31181/sor202786

© The Author(s) 2026 | Creative Commons Attribution 4.0 International License

A critical review of current operations management literature reveals a persistent bifurcation between commercial supply chain management (SCM)—which champions leanness, Just-In-Time (JIT) principles, and cost-efficiency [5]—and humanitarian logistics, which prioritizes speed, equity, and maximum coverage [6]. Consequently, a substantial theoretical and practical gap exists regarding "hybrid" environments that constantly oscillate between stability and crisis. Furthermore, while existing operations research (OR) models occasionally incorporate human safety (e.g., driver risk in conflict zones), they predominantly treat it as a rigid binary constraint rather than a dynamic variable within the objective function [7]. This neglect of the "ethical dimension" in quantitative modeling results in algorithmic decision-making that may inadvertently endanger personnel to meet predefined service targets.

To bridge this theoretical divide, this study introduces a novel Dynamic AI-Driven Policy Orchestrator. Rather than relying on a single static "all-weather" compromise policy [8], our framework employs prescriptive machine learning (ML) to dynamically govern the supply chain. By continuously retuning three critical operational levers—Target Inventory Days, Rationing Limits, and Delivery Radius—based on observed contextual signals, the orchestrator dynamically adjusts the network's policy behavior.

Specifically, this research aims to address the following critical research questions (RQs):

i. RQ1: How can prescriptive analytics and machine learning be utilized to dynamically transition a pharmaceutical supply chain between cost-focused, balanced, and resilient regimes under extreme uncertainty?
ii. RQ2: To what extent can a hybrid AI orchestrator operationalize safety-constrained logistics by prioritizing modeled human safety over commercial or service-level targets during acute security crises?

By addressing these questions, the contributions of this paper to the existing body of knowledge are fourfold:

Methodological: We propose a novel architecture that integrates Latin Hypercube Sampling (LHS) and simulation-based optimization with Random Forest regression, creating a context-aware policy-learning framework.

Theoretical: We introduce a context-dependent Regime-Switching framework that characterizes when the tested supply-chain policies shift from cost-oriented toward more resilience-oriented configurations under changing operating conditions.

Simulation-based: We provide an episodic simulation analysis comprising a 90-day steady-state horizon and three 20-day shock horizons, demonstrating how the orchestrator adapts across four operational scenarios.

Ethical: We operationalize an ethics-informed, safety-constrained decision layer by incorporating a security-dependent human-risk penalty and deterministic guardrails into the logistics objective, allowing modeled safety exposure to be reduced during severe security conditions.

The remainder of this paper is structured as follows. Section 2 reviews the relevant literature identifying the existing gaps. Section 3 details the mathematical formulation, simulation design, and AI training methodology. Section 4 presents the experimental results and performance analysis. Section 5 discusses the theoretical and managerial implications, and Section 6 concludes the study with avenues for future research.

2. Literature Review

The operational challenges inherent in pharmaceutical distribution within volatile regions intersect three distinct streams of literature: Supply Chain Resilience, Artificial Intelligence in

Operations Management, and Humanitarian Logistics. This section critically examines these domains to contextualize the proposed dynamic orchestration framework.

2.1. Supply Chain Resilience & Adaptability

The discourse on supply chain resilience has historically centered on designing networks capable of withstanding severe disruptions while maintaining core functions [9]. Traditional Operations Research (OR) heavily relies on stochastic programming and robust optimization to build this resilience, typically aiming to formulate a single, static policy that minimizes expected costs across a set of predefined scenarios [10].

However, recent literature argues that in environments characterized by "deep uncertainty"—such as geopolitical conflicts or sudden economic shocks—static robust policies become highly inefficient [11]. A policy designed to withstand extreme shocks is often too costly and bloated during stable periods, whereas a lean, cost-optimized policy becomes dangerously fragile during a crisis. Consequently, scholars have called for a paradigm shift from static robustness to "dynamic adaptability" (or agility), where supply chain parameters are continuously reconfigured in response to real-time environmental volatility [12]. Recent VMI research has also demonstrated that dynamic replenishment policies can outperform static and reactive alternatives in service-level performance under stochastic demand, while maintaining statistically comparable total costs [13]. Despite these advances, quantitative models capable of switching operational regimes in response to broader contextual signals remain scarce.

2.2. AI in Operations Management

To address the limitations of static mathematical programming, the integration of Artificial Intelligence (AI) and Machine Learning (ML) into operations management has accelerated. The literature documents a clear epistemological shift from predictive analytics (forecasting demand or disruptions) to prescriptive analytics, where algorithms recommend or autonomously execute complex operational decisions [14].

Techniques such as Reinforcement Learning (RL) and simulation-optimization have shown promise in dynamic routing and inventory control [15], [16]. In fact, recent studies in pharmaceutical supply chains have reported that rigorously optimized, parsimonious heuristics can absorb environmental shocks and reduce stockouts more effectively than highly complex, over-parameterized adaptive models [17]. Yet, a significant barrier to the adoption of AI in safety-critical logistics (like pharmaceuticals) is the "black-box" nature of many deep learning models [18]. Supply chain managers demand interpretability and operational guardrails. Furthermore, existing ML applications predominantly focus on single-echelon optimization (e.g., just routing, or just inventory) [19]. There is a notable dearth of research utilizing transparent AI (such as Random Forests coupled with expert heuristic rules) as a holistic "Policy Orchestrator" that simultaneously tunes inventory targets, rationing logic, and routing parameters.

2.3. Humanitarian Logistics & Ethics

The fundamental divergence between commercial SCM and humanitarian logistics lies in their objective functions. While commercial logistics optimizes for Total Landed Cost (TLC) and profit maximization, humanitarian operations prioritize speed, equity, and the minimization of "deprivation costs"—the human toll of unmet medical needs [20].

A critical, yet underexplored, area is the "hybrid" supply chain operating in developing or conflict-affected regions, which must routinely transition between commercial and humanitarian states [21]. More importantly, the ethical dimension of logistics—specifically the physical safety of drivers and personnel in hostile environments—is rarely modeled mathematically. Existing OR literature tends to treat security as a rigid binary constraint (e.g., rendering a route completely unavailable [7]). Very few studies have attempted to monetize or penalize human risk dynamically within the objective function, which supports explicit safety-prioritized trade-offs, such as accepting higher modeled stockout costs to avoid entering designated danger zones. Recent work has also explored LLM-assisted expert weight elicitation in pharmaceutical supply-chain decision models, suggesting a potential human-in-the-loop pathway for calibrating normative weights in future extensions [22].

2.4. Synthesis and Research Gap

A critical synthesis of the aforementioned literature reveals a tripartite gap. First, existing models lack the architectural fluidity to dynamically shift between cost-focused commercial logic and resilience-focused humanitarian logic. Second, there is a lack of interpretable, prescriptive AI frameworks capable of orchestrating multiple structural policies (inventory, rationing, routing) simultaneously in real-time. Third, the algorithmic governance of ethical trade-offs—specifically mathematically penalizing human danger to force "No-Go" safety protocols—remains unaddressed. This study directly addresses these gaps by proposing a dynamic, AI-driven policy orchestrator. By combining LHS-driven simulation optimization with Machine Learning, this paper introduces a quantifiable framework where ethics (human safety) and resilience (humanitarian coverage) are dynamically balanced against commercial efficiency.

3. Methodology

This section details the proposed hybrid framework, which integrates mathematical programming, discrete-event simulation, and machine learning. The methodology is executed in three sequential phases: (1) defining the mathematical formulation of the dynamic pharmaceutical environment, (2) generating high-fidelity operational scenarios using Latin Hypercube Sampling (LHS) for simulation-based optimization, and (3) training the AI-driven Policy Orchestrator.

3.1. Problem Description and Nomenclature

We consider a pharmaceutical distribution network modeled as a directed graph $G = (N, A)$, where N is the set of pharmacy nodes (including the central depot) and A is the set of transportation arcs, analyzed over a discrete planning horizon $t \in T$. To circumvent the methodological limitations of theoretical "toy networks" and retain geographic realism in the spatial routing dynamics, the simulation topology is grounded in a geographically realistic pharmacy network rather than a theoretical toy network. Specifically, the set N comprises a central distribution depot and 120 actual pharmacy locations in Misrata, Libya, represented by their corresponding latitude and longitude coordinates. These geographic inputs are real, whereas pharmacy-level demand, product profiles, storage capacities, and other operational attributes used in the simulation are synthetic because historical operational records were not available at the time of the study. Accordingly, the network provides geographic realism but does not constitute historical or live operational validation. Historical demand, product, and capacity records have been requested from a pharmaceutical supplier and, subject to data access, will support field-grounded historical validation in subsequent work. Furthermore, to capture product heterogeneity without overcomplicating the combinatorial action space, the aggregate stochastic patient demand $D_{i,t}$ is synthetically generated from three distinct priority tiers (high-velocity life-saving drugs, medium-velocity therapeutics, and low-velocity consumables), each parameterized with distinct Gamma distribution scales to reflect the tested operational variability. To ensure mathematical clarity, the indices, parameters, and decision variables used in the formulation are summarized in Table 1.

Table 1
Nomenclature

Symbol	Type	Description
Indices		
$i, j \in N$	Set	Pharmacy nodes and the central depot.
$t \in T$	Set	Discrete time periods (days) in the planning horizon.
Parameters		
$D_{i,t}$	Stochastic	Patient demand at node i at time t, modeled as $\Gamma(k, \theta)$.
δ_t	Dynamic	Time-varying demand stress factor shaping the Gamma distribution.
$C_{fuel,t}$	Dynamic	Fuel price per kilometer at time t.
$d_{i,j}$	Deterministic	Distance between node i and node j.
h	Deterministic	Unit holding cost per day.
p_t	Dynamic	Unit stockout (shortage/deprivation) penalty cost at time t.
σ_i	Dynamic	Security index at time t, $\sigma_t \in [0,1]$ (0 = hostile, 1 = safe).
C_r	Deterministic	Per-failed-delivery risk penalty; 100 USD in the base case.
θ_t	Binary	Humanitarian Emergency flag at time t (1 if emergency, 0 otherwise).
B_{perf}	Deterministic	Performance bonus rewarded for maximizing humanitarian coverage.
μ_i	Deterministic	Reference average daily demand at node i.
Q_{max}	Deterministic	Maximum vehicle payload capacity.
T_{max}	Deterministic	Maximum allowable shift duration for drivers; the effective limit is ηT_{max}, with η = 0.95.
Variables		
$q_{i,t}$	Continuous	Delivery quantity allocated to node i at time t.
$x_{i,j,t}$	Binary	Routing variable; 1 if arc $(i, j)$is traversed at time t, 0 otherwise.
$I_{i,t}$	Continuous	On-hand inventory level at node i at the end of period t.
π_t	Vector	AI-prescribed policy configuration (τ, ρ, R) at time t.

3.2. Mathematical Formulation

To capture the extreme volatility of crisis environments, patient demand $D_{i,t}$ is modeled as a stochastic process following a Gamma distribution, $D_{i,t} \sim \Gamma(k, \theta)$. The shape (k) and scale (θ) parameters are dynamic functions of the demand stress factor (δ_t), enabling the model to simulate sudden surges in medical needs without generating unrealistic negative demands.

Additional formulation notation: $\boldsymbol{\mathcal{R}}_t$ denotes the set of depot-originating and depot-returning routes constructed at time t; T_fixed and T_var are fixed and variable service-time components, v is the average vehicle speed, α is the target service-level probability, Ωrisk,i,j,t is the realized risk penalty on a served arc, and 0 denotes the depot node. The danger radius R_t is computed from the contemporaneous security context.

The central objective is to minimize the total logistics and risk cost (Z), defined as the sum of transport, holding, stockout, and security-risk terms, adjusted by the emergency performance incentive:

Minimize:

$$Z = \sum_{t \in T} \left[\sum_{i,j \in N} \left(C_{fuel,t} d_{ij} x_{ij,t} + \Omega_{risk,ij,t} \right) + \sum_{i \in N} \left(hI_{i,t} + p_t \max\{0, D_{i,t} - I_{i,t-1} - q_{i,t}\} \right) - \theta_t B_{perf} \right] \quad (1)$$

The security-risk term is realized in the execution model rather than as an independent deterministic arc cost. For a served arc ending at node j, $\Omega_{risk,i,j,t}$ represents the realized risk cost: it equals cr when the stochastic safety screen records a failed delivery at node j and zero otherwise, with cr = 100 USD per failed delivery in the base case. A danger-node failure event is generated only when

σ_t < 0.8, using the implemented probability $0.9(1 - d_{0,j}/R_t)(1 - \sigma_t)$, where $d_{0,j}$ is the node-to-depot distance and Rt is the computed danger radius. Thus, the risk term links the objective to modeled security exposure while remaining an operational penalty rather than a claim of autonomous moral reasoning. During a humanitarian emergency (θ_t = 1), the performance term provides an incentive for broader service, subject to the binding operational and safety constraints described below.

The optimization is subject to the following operational constraints:

$$I_{i,t} = I_{i,t-1} + q_{i,t} - \min\left(I_{i,t-1} + q_{i,t}, D_{i,t}\right) \forall i, t \tag{2}$$

$$\Sigma_{i\in r}\, q_{i,t} \;\le\; Q_{max}, \;\; \forall\, r \in \mathcal{R}_t, t \in T \tag{3}$$

$$\Sigma_{i\in r}\, (T\_fixed \;+\; T\_var \cdot q_{i,t}) \;+\; \Sigma_{(i,j)\in r}\, (d_{i,j}/v) \cdot x_{i,j,t} \;\le\; \eta T_{max}, \;\; \forall\, r \in \mathcal{R}_t, t \in T \tag{4}$$

$$q_{i,t} \le \rho_t \cdot \mu_i \forall i, t \tag{5}$$

$$\mathbb{P}\left(I_{i,t} \ge D_{i,t}\right) \ge \alpha \cdot (1 - \theta_t) \forall i, t \tag{6}$$

$$\Sigma_{j\in N}\, x_{ij,t} \;=\; \Sigma_{j\in N}\, x_{ji,t} \;\le\; 1, \;\; \forall\, i \in N \setminus \{0\}, t \in T \tag{7}$$

$$q_{i,t} \;\le\; Q_{max}\, \Sigma_{j\in N}\, x_{ij,t}, \;\; \forall\, i \in N \setminus \{0\}, t \in T \tag{8}$$

$$\Sigma_{j\in N\setminus\{0\}}\, x_{0j,t} \;=\; \Sigma_{i\in N\setminus\{0\}}\, x_{i0,t}, \;\; \forall\, t \in T \tag{9}$$

The above constraints jointly define the structural logic of the proposed optimization framework. Equation (2) preserves inventory balance, ensuring that stock levels evolve consistently with incoming replenishments and realized demand. Equation (3) limits the delivered quantity on each constructed route to the effective vehicle payload capacity. Equation (4) bounds the service and travel time of each constructed route. Equations (7)– (9) make route-flow continuity, delivery-to-route linkage, and depot return explicit: for each non-depot node, the number of selected inbound and outbound arcs is balanced and limited to one visit per period ($\Sigma_{j\in N}\, x_{ij,t} \;=\; \Sigma_{j\in N}\, x_{ji,t} \;\le\; 1$). Routes are constructed as depot-originating and depot-returning tours by the insertion heuristic; disconnected subtours are therefore excluded by construction rather than by a separate subtour-elimination constraint. Candidate node insertions are accepted only when routing, capacity, and route-duration conditions remain satisfied. Equation (5) operationalizes adaptive rationing, while Equation (6) defines the service-level condition and its relaxation during humanitarian emergencies. These layers separate policy adaptation from physical route feasibility, so that a policy recommendation is not treated as operationally executable unless its route remains feasible.

3.3. Scenario Generation and Simulation-Based Optimization

Unlike structural decisions (e.g., facility location), operational agility requires dynamic levers. We restrict the AI orchestrator's continuous action space to an operational configuration vector $\pi = (\tau, \rho, R)$ representing:

- **i.** Target Inventory Days (τ): Buffer against demand stress (Range: $[3.0 \;-\; 6.0]$ days).
- **ii.** Rationing Ratio (ρ): Crisis-response allocation limit per node (Range: $[0.6 \;-\; 1.0]$, where 1.0 implies no rationing).
- **iii.** Operational Radius (R): Spatial filter expanding or contracting operations based on safety (Range:$[15.0 \;-\; 30.0]$ km).

To train the AI on a broad representation of potential operating conditions, we used Latin Hypercube Sampling (LHS) to generate 1,000 distinct operational contexts. Each context was evaluated under the three candidate policy regimes, producing 3,000 context–policy records. As depicted in Figure 1, the contextual inputs were Demand Stress, Security Index, Fuel Price, and Emergency Flag.

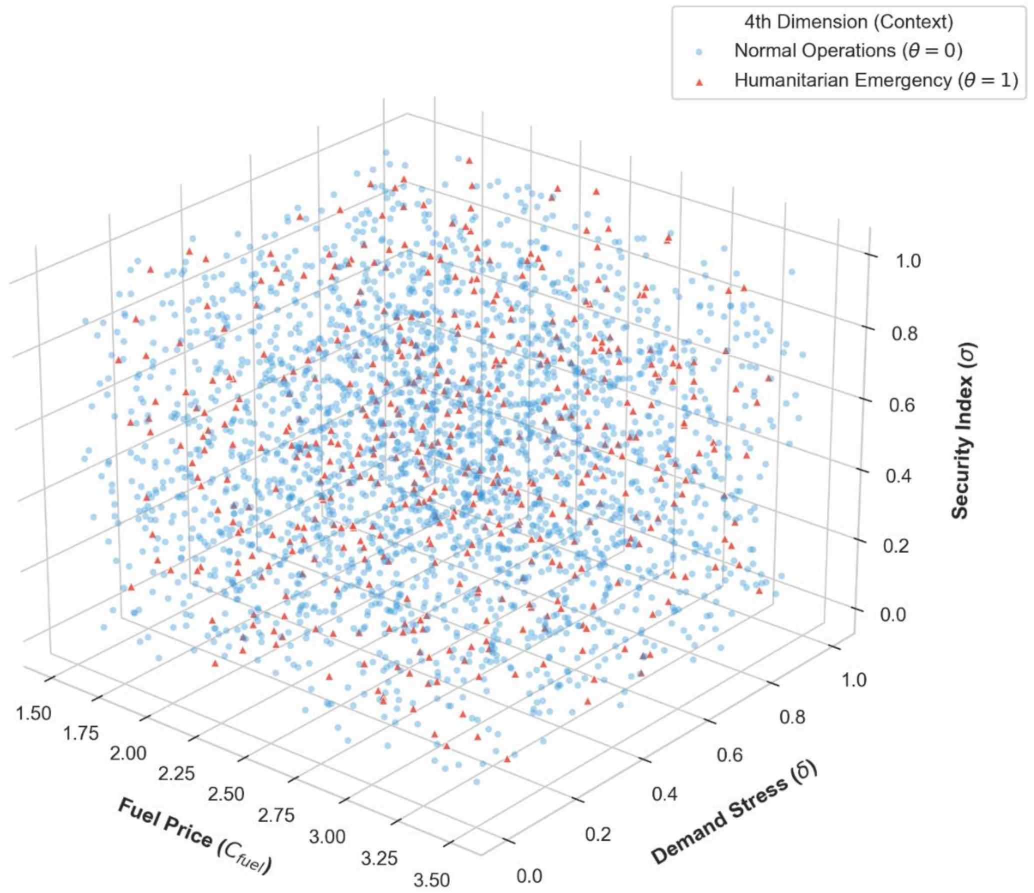

Fig. 1. Latin Hypercube Sampling Coverage of the Four-Dimensional Operational Context Space

The retained analysis file records the total cost obtained for each context–policy combination. For each of the 1,000 contexts, the policy with the lowest recorded Total_Cost was selected as the optimal label. This procedure produced 1,000 labeled observations for supervised learning.

$$\pi_k^* = \ argmin_{\pi \in \Pi} \, Total_{Cost} \, (S_k, \pi), \Pi = COST_{FOCUSED,BALANCED,RESILIENT} \tag{10}$$

This mapping transformed the stochastic outputs into a structured, labeled dataset:

$$\mathcal{D} = \{(S_k, \pi_k^*)\}_{k=1}^{1000} \tag{11}$$

3.4. AI Training and Hybrid Rule-Based Integration

A multi-output Random Forest Regressor was employed to map the four contextual inputs to three policy parameters: Target Inventory Days, Rationing Ratio, and Operational Radius. The retained training script uses 100 trees with $\mathrm{random}_{\mathrm{state}}$= 42 and a standard 80/20 train–test split with $\mathrm{random}_{\mathrm{state}}$ = 42. The 1,000 optimal labels therefore yield 800 training observations and 200 test observations. The optimized Random Forest hyperparameters are summarized in Table 2.

The raw model predictions are passed through a deterministic domain-logic guardrail layer before operational execution. The guardrail separates adaptive policy selection from non-negotiable safety restrictions. When a humanitarian emergency is detected ($\theta_t = 1$), the policy layer shifts toward service-oriented operation by permitting the removal of routine rationing restrictions and expansion of the operational search radius up to its predefined upper bound of 30 km. However, activation of the humanitarian protocol does not override an explicitly designated danger-zone restriction. When severe security conditions are detected, safety restrictions remain binding for affected nodes, and prohibited or infeasible service decisions are rejected at the operational execution layer. Thus, emergency adaptation increases service flexibility without authorizing the system to violate predefined safety constraints.

Table 2
Optimized Random Forest Hyperparameters

Hyperparameter	Value	Basis in retained script
Number of estimators	100	Matches the supplied RandomForestRegressor configuration.
Maximum depth	None (default)	No depth restriction was specified in the supplied script.
Minimum samples split	2 (default)	Default setting in the supplied RandomForestRegressor.
Minimum samples leaf	1 (default)	Default setting in the supplied RandomForestRegressor.
Max features	1.0 (default)	Uses all four contextual features at each split.
Bootstrap	True (default)	Bootstrap aggregation is enabled by default.

In this study, "Ethical AI" is used as an operational design concept rather than a claim of machine moral agency. It refers to three explicit mechanisms: (i) a modeled human-risk penalty embedded in the optimization objective, (ii) deterministic safety guardrails that constrain or reject unsafe operational decisions, and (iii) an auditable separation between adaptive policy selection and non-negotiable safety constraints. Under this definition, the ethical property being evaluated is the consistency of the prescribed decision behavior with the predefined safety and humanitarian priorities; the framework does not infer intentions, moral understanding, or autonomous ethical judgement from the machine-learning model.

3.5. Episodic Scenario Testing

To evaluate dynamic behavior without compounding transient shocks into a single continuous annual trajectory, we adopted an episodic discrete-event simulation design. The retained evaluation comprises a 90-day Steady-State Baseline followed by three separate 20-day shock horizons—Fuel Price Surge, Security Crisis, and Humanitarian Emergency—for a total of 150 simulated days. Each shock is interpreted as a separate stress-test episode rather than as a continuation of the preceding episode. The static comparator is a fixed-policy benchmark evaluated on the same network, demand realizations, cost structure, vehicle-capacity limits, and operating-time limits, but without context-dependent retuning of the policy levers. This static baseline is an internally generated comparator under the same simulation environment, not an external literature value.

The three Balanced, Resilient, and Cost-Focused regimes used during policy-label generation are internal candidate policies of the proposed framework, not independent implementations of external robust or stochastic optimization algorithms; consequently, they are not presented as evidence of superiority over the broader optimization literature.

Robustness was assessed at two complementary levels. First, context-space robustness was examined through LHS variation in demand stress, security index, fuel price, and emergency state, followed by evaluation of regime-selection frequencies and out-of-sample predictive performance. Second, a one-factor-at-a-time parameter sensitivity analysis was conducted over 20 independent replications per condition, varying the per-node risk penalty (50–200 USD), stockout penalty (2.5–10 USD/unit), and selective-AI safety threshold (0.60–0.90). The emergency performance bonus (0–900 USD) was assessed separately at the objective level because it is an additive term in the supplied implementation rather than a behavioral engine parameter. Accordingly, the sensitivity results are interpreted as local parameter sensitivity within the tested ranges; they are not presented as a global or factorial robustness analysis.

4. Results and Analysis

This section evaluates the performance of the proposed AI-driven Policy Orchestrator. We first assess the predictive fidelity of the machine learning model, followed by a spatial analysis of the regime-switching dynamics. Finally, we provide a comparative evaluation of the orchestrator against a static baseline policy across four episodic stress-tested operational scenarios.

4.1. Predictive Performance and Algorithm Fidelity

The retained Random Forest training results indicate a strong mapping from operating context to the policy parameters. Target Inventory Days achieved $R^2 = 0.93$ with MAE = 0.0286, Rationing achieved $R^2 = 0.82$ with MAE = 0.0090, and Operational Radius achieved $R^2 = 0.77$ with MAE = 0.386. The lower explanatory power for radius is consistent with the greater sensitivity of routing decisions to security and spatial conditions. These metrics should be interpreted as predictive performance on the retained test split, not as evidence of field-level validation. The resulting performance metrics are reported in Table 3.

Table 3
Model Predictive Performance Metrics

Parameter	MAE	R^2 Score	Interpretation
Target Days	0.0286	0.93	Strong predictive fit on the retained test split.
Rationing	0.0090	0.82	Good predictive fit on the retained test split.
Radius	0.386	0.77	Moderate predictive fit; routing is more context-sensitive.

4.2. Regime Dynamics and Strategic Topography

The generative capacity of the Latin Hypercube Sampling (LHS) engine produced a broad multi-dimensional operational space. The sampled parameter space is distributed across the tested context dimensions without obvious structural clustering, supporting the intended coverage of the simulated environments. The resulting three-dimensional strategy map of the retained optimal policy labels is shown in Figure 2, illustrating how the Balanced and Resilient regimes are distributed across the tested context space.

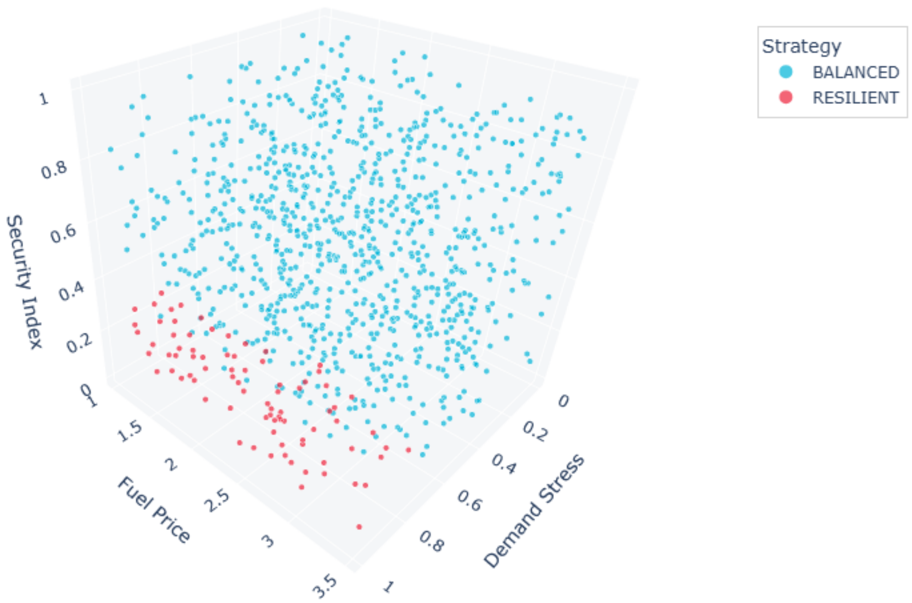

Fig. 2. Three-dimensional strategy map of the retained optimal policy labels across the tested context space

The retained optimization labels reveal two dominant regimes: "Balanced" operations account for 908 of the 1,000 optimal contexts (90.8%), while "Resilient" operations account for 92 contexts (9.2%). The Cost-Focused regime was not selected as the minimum-cost label in the retained dataset. This distribution indicates that, within the tested context space, the balanced configuration was generally preferred, with the resilient configuration emerging in a smaller subset of more demanding conditions.

The transition toward the Resilient regime is observed in contexts combining greater operating stress with less favorable security conditions. Rather than treating a single numerical cutoff as universal, the result is interpreted as an observed regime boundary generated by the tested context space and candidate policy set. The distribution of the resulting optimal strategic responses across Security Index (σ) deciles is shown in Figure 3.

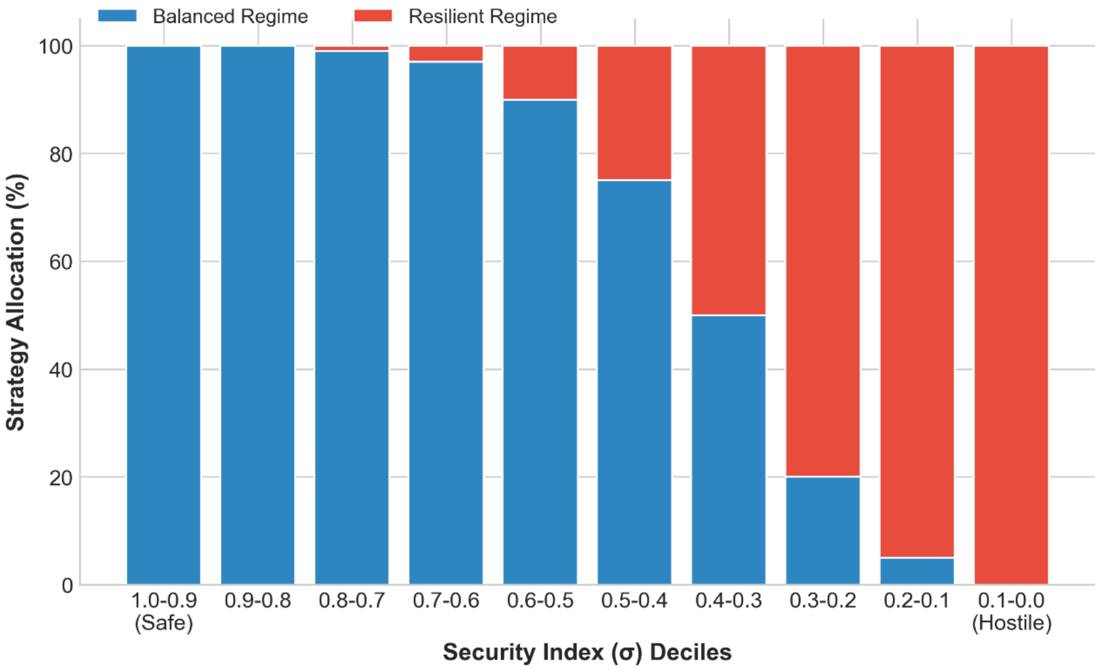

Fig. 3. Distribution of optimal strategic responses across Security Index (σ) deciles

To evaluate the orchestrator's dynamic adaptability without compounding different shocks into a single continuous trajectory, an episodic discrete-event comparison was used. The dynamic policy was compared with the fixed static benchmark over the 90-day steady-state horizon and three separate 20-day shock episodes. The benchmark uses the same network, demand realizations, cost structure, capacity limits, and operating-time constraints, while holding its policy configuration fixed rather than adapting it to the observed context. The three candidate regimes used in LHS policy labeling are internal alternatives within the proposed architecture and are not treated as independent external benchmark algorithms. The episodic comparison used 20 independent replications per policy. Scenario-specific Welch comparisons were used for the retained episodic performance outcomes. Results are reported as dynamic policy minus static baseline, together with Welch's t statistic, Welch–Satterthwaite degrees of freedom, two-sided p-values, 95% confidence intervals, and Cohen's d where estimable. The n = 10 replication count used during candidate-policy labeling is therefore distinguished from the n = 20 replications used for the final episodic policy comparison. Table 4 provides the scenario-specific comparison results.

Table 4
Scenario-specific Welch comparisons between the dynamic policy and fixed static benchmark

Scenario / Metric	Δ (Dynamic – Static)	Welch t	df	p (2-sided)	95% CI	Cohen's d
Steady-State — Total cost	+1000.96	+4.07	29.64	0.00031	[498.59, 1503.33]	+1.287
Steady-State — Transport	-10.90	-0.44	37.55	0.658	[-60.55, 38.73]	-0.140
Steady-State — Holding	+1014.12	+4.35	29.36	0.00014	[538.20, 1490.04]	+1.377
Steady-State — Stockout	-2.25	-0.177	34.11	0.860	[-28.06, 23.56]	-0.056
Steady-State — Distance (km)	-5.80	-0.569	37.99	0.572	[-26.43, 14.82]	-0.180
Fuel Price Surge — Total cost	+903.06	+3.74	31.63	0.0007	[412.07, 1394.05]	+1.185
Fuel Price Surge — Transport	-18.54	-0.473	37.92	0.638	[-97.80, 60.70]	-0.149
Fuel Price Surge — Holding	+908.86	+4.1749	30.96	0.00022	[464.85, 1352.87]	+1.320
Fuel Price Surge — Stockout	+12.75	+0.753	35.84	0.455	[-21.55, 47.05]	+0.238
Fuel Price Surge — Distance (km)	-5.81	-0.574	37.99	0.569	[-26.32, 14.69]	-0.181
Security Crisis — Total cost	+5994.54	+4.24	20.38	0.0003	[3053.47, 8935.62]	+1.342
Security Crisis — Transport	-2012.90	-11.79	19.03	3.40126e-10	[-2370.11, -1655.682]	-3.729
Security Crisis — Holding	-5611.05	-11.42	20.75	2.08006e-10	[-6633.42, -4588.67]	-3.611
Security Crisis — Stockout	+13618.50	+7.70	19.36	2.57195e-07	[9923.84, 17313.15]	+2.436
Security Crisis — Danger-node traversals	-40.75	-11.80	19.00	3.41365e-10	[-47.9748, -33.5252]	-3.733
Humanitarian Emergency — Total cost	-1559.52	-137.05	24.9850	1.82704e-37	[-1582.9620, -1536.0914]	-43.341
Humanitarian Emergency — Transport	+60.92	+12.23	27.5468	1.1869e-12	[50.7177, 71.1289]	+3.869
Humanitarian Emergency — Holding	-1620.45	-171.08	19.0239	8.67939e-32	[-1640.2732, -1600.6268]	-54.100
Humanitarian Emergency — Distance (km)	+27.41	+12.72	37.6999	3.25524e-15	[23.0517, 31.7794]	+4.022
Humanitarian Emergency — Trips	+0.90	+9.00	19.0000	2.79276e-08	[0.6907, 1.1093]	+2.8460

Note: Static-baseline values were generated within the same experimental environment and are reported for transparency and reproducibility; they were not imported from external studies.Δ denotes dynamic policy minus static baseline. Two-sided p-values are reported in the table. NE (not estimable) applies to fixed-value outcomes with zero within-group variance; such outcomes are not included in this inferential table. The candidate-policy labeling stage used n = 10 independent replications per candidate policy, whereas the final episodic comparison used n = 20 independent replications per policy.

During the 90-day Steady-State baseline, the retained comparison reports a static total logistics cost of 8,155 USD and holding cost of 7,948 USD. Using the corresponding mean differences in Table 4, the dynamic policy is 9,155.97 USD in total cost and 8,962.13 USD in holding cost, while routing distance decreases from 84.1 to 78.3 km. The implied changes are therefore +12.3% in total cost, +12.8% in holding cost, and −6.9% in routing distance; the stockout penalty decreases from 15.00 to 12.75 USD. This pattern is better characterized as a deliberate resilience–efficiency trade-off rather than as unconditional cost superiority. The resulting cost and routing trade-off during the steady-state baseline is visualized in Figure 4.

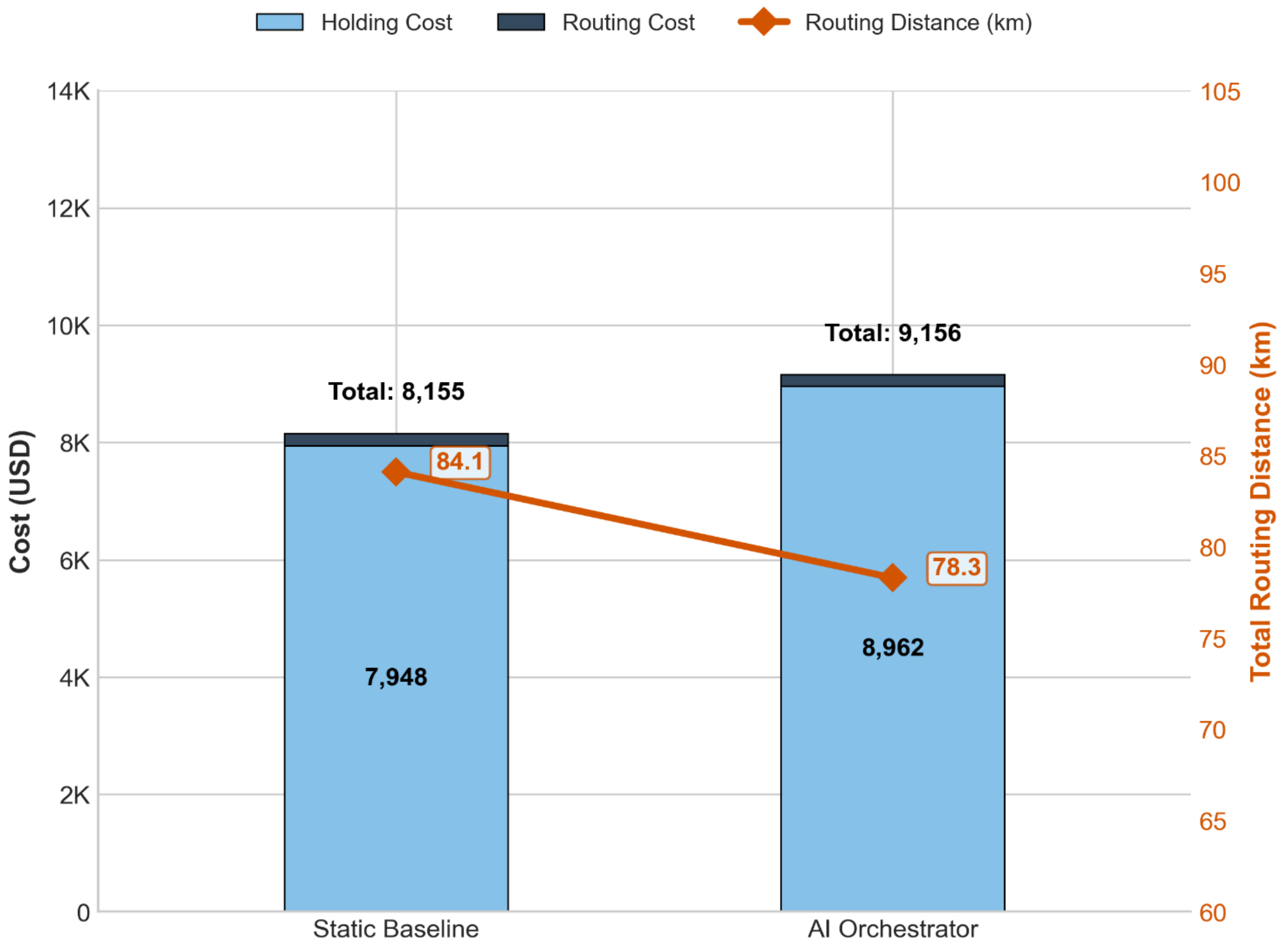

Fig. 4. Cost and routing breakdown comparison during the 90-day Steady-State baseline, showing the trade-off between higher inventory holding cost and a smaller routing footprint

The strategic divergence becomes clearer under the acute 20-day shock episodes. Under the retained 50% Fuel Price Surge test, the static policy maintained a 20.0 km delivery radius, whereas the adaptive policy contracted its effective radius to 14.2 km. As shown in Figure 5, this spatial contraction was accompanied by a lower cumulative transport-cost trajectory over the shock horizon.

The strongest modeled safety response occurred during the 20-day Security Crisis. The static baseline recorded a mean of 40.75 modeled danger-node traversals per replication and a mean modeled risk penalty of 1,610 USD per replication. The hybrid orchestrator instead activated the safety-oriented guardrail behavior, reducing modeled danger-node traversals to zero. This safety-oriented response was associated with a mean increase in stockout penalty of 13,618.5 USD per replication relative to the static baseline (Figure 6). These results illustrate the operational effect of the safety-prioritized objective and guardrail layer by reducing modeled danger-node exposure; they should not be interpreted as evidence of zero real-world harm or guaranteed personnel safety.

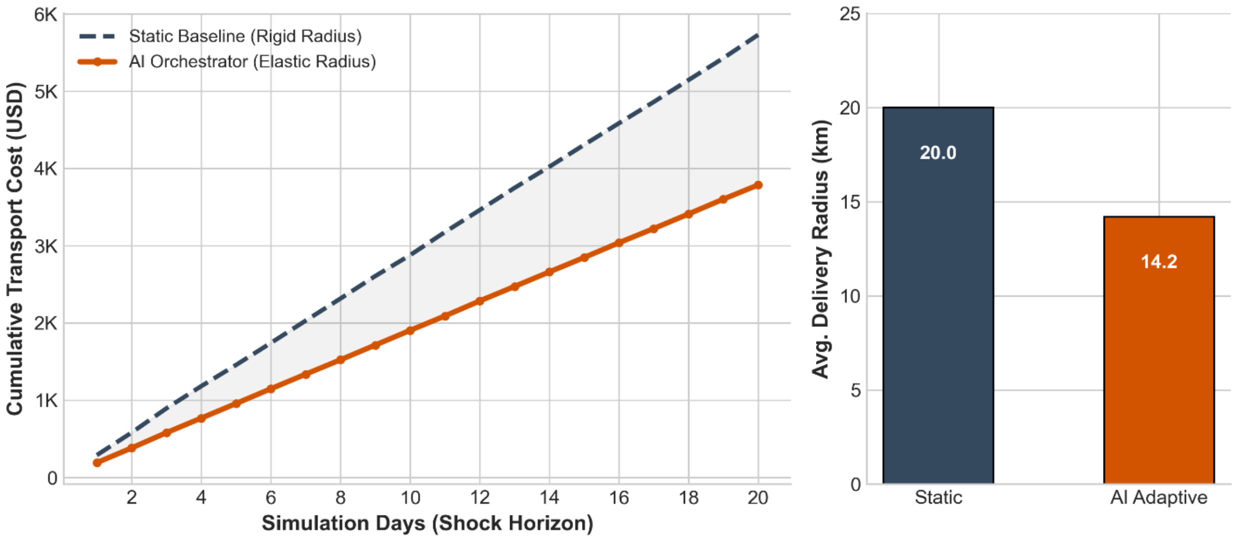

Fig. 5. Cumulative transport cost and delivery-radius comparison during the 20-day Fuel Price Surge

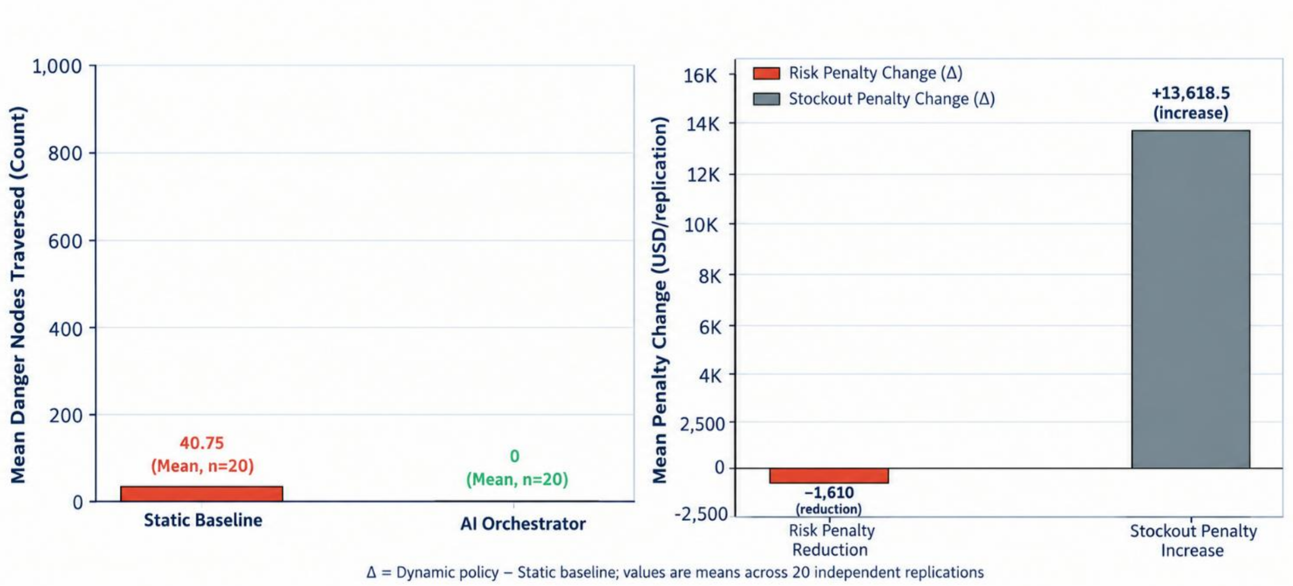

Fig. 6. Modeled safety-risk and stockout-penalty trade-off during the 20-day Security Crisis, showing the reduction of danger-node traversals under the safety-constrained policy

In the 20-day Humanitarian Emergency episode, the emergency flag changes the operating logic from routine commercial control toward service-oriented response. The episode spans 2,400 pharmacy-visits/service events (120 pharmacy locations × 20 days); this is an evaluation-opportunity count, not a count of unique pharmacies. The implemented policy space permits the removal of normal rationing restrictions and expansion of the delivery radius to its defined upper bound. Because the retained working evidence does not preserve a consistent pharmacy-level coverage denominator for this episode, no percentage coverage gain is reported here.

To address parameter-level robustness, a one-factor-at-a-time sensitivity analysis was conducted over the retained operating ranges. Each condition used 20 independent replications. The analysis varied the per-node risk penalty (50–200 USD), stockout penalty (2.5–10 USD/unit), and selective-AI safety threshold (0.60–0.90). The emergency performance bonus (0–900 USD) was examined separately at the objective level because it is additive in the supplied implementation and does not alter the simulated physical outcomes. The resulting local sensitivity patterns are summarized in Figure 7.

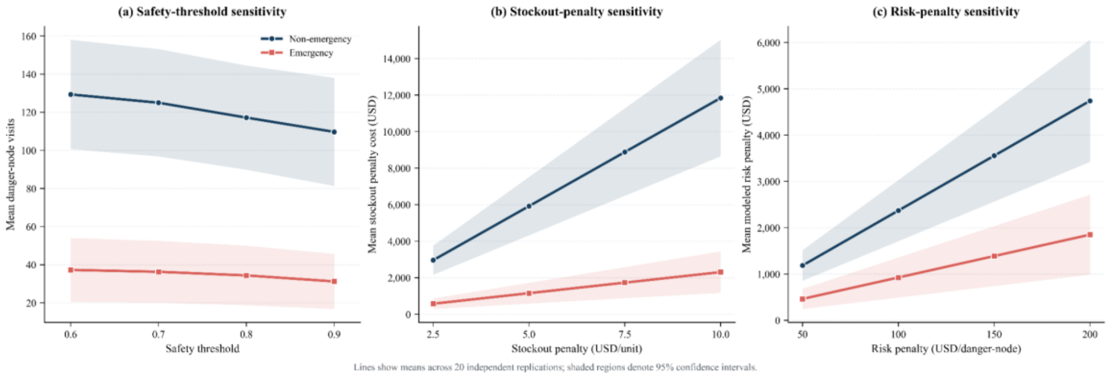

Fig. 7. One-factor-at-a-time sensitivity of modeled safety exposure, stockout penalty, and risk penalty across the tested parameter ranges. Lines show means across 20 independent replications; shaded regions denote 95% confidence intervals

The safety-threshold perturbation produces the clearest operational response because it changes the selective-AI admission rule for danger nodes. By contrast, the risk-penalty and stockout-penalty perturbations primarily change the monetary weight assigned to modeled outcomes, while the emergency-bonus perturbation changes only the objective value and therefore does not by itself alter routing, service, or exposure trajectories. These results provide local parameter sensitivity evidence within the tested ranges; they do not imply that the LHS-derived regime frequencies were re-estimated under every perturbation.

5. Discussion

The simulation findings of this study provide an operational perspective on the design and governance of high-stakes supply chains in volatile environments. By examining the interaction among commercial cost-efficiency, operational resilience, and modeled personnel safety, the study clarifies how a context-aware policy orchestrator can adapt competing operational priorities.

5.1. Theoretical Implications

From a theoretical standpoint, the findings support a shift from a single static policy toward context-dependent operational policies in environments characterized by deep uncertainty. The retained scenario labels show that the preferred policy is not constant across the tested context space: Balanced is selected in most contexts, while Resilient is selected in a smaller subset. This supports the use of regime-switching logic in which inventory, rationing, and routing parameters are functions of operating context rather than fixed constants.

The study also provides a quantitative illustration of ethics-informed logistics. By coupling a security-dependent risk mechanism with deterministic guardrails, the framework assigns a high modeled cost to unsafe delivery outcomes and can reject service decisions under designated danger conditions. In the retained Security Crisis episode, this mechanism reduced mean modeled danger-node traversals from 40.75 to zero per replication, while accepting a substantial stockout penalty relative to the static baseline. The finding demonstrates the operational effect of the modeled safety architecture; it does not imply that the algorithm possesses moral agency or establishes real-world personnel safety.

Finally, the steady-state comparison refines the interpretation of operational efficiency. Relative to the fixed benchmark, the dynamic policy increased mean total cost by 1,000.97 USD and mean holding cost by 1,014.13 USD, while reducing mean routing distance by 5.81 km. The corresponding changes are approximately +12.3%, +12.8%, and −6.9%, respectively. This pattern can be interpreted as a resilience-oriented trade-off in which additional inventory is exchanged for a smaller physical routing footprint.

5.2. Managerial and Policy Implications

For supply chain executives and humanitarian coordinators, the main managerial implication is not that AI should replace human decision-makers, but that a transparent policy-orchestration layer can provide context-dependent recommendations while retaining explicit safety guardrails. This is particularly relevant when demand, fuel availability, and security conditions can change faster than manually maintained static policies.

During acute shocks, the proposed architecture can reduce the need for ad hoc policy changes by linking observable contextual signals to predefined operational levers. The fuel-shock result illustrates spatial contraction as a response to transport-cost pressure, while the security-shock result illustrates the use of safety guardrails to restrict modeled exposure. These behaviors are decision-support mechanisms and should remain subject to operational oversight in real deployment.

The 90-day steady-state result also shows why resilience should not be evaluated only through immediate logistics cost. The dynamic policy incurred higher holding cost but reduced routing distance and reduced the recorded stockout penalty. Consequently, resilience investment should be evaluated against the full operational objective rather than against holding cost alone.

The hybrid architecture also addresses an important transparency requirement. Random Forest predictions are bounded by explicit domain rules in extreme contexts, making the final decision process more auditable than an unconstrained black-box policy. Nevertheless, the present study does not establish field-level safety assurance; the guardrails should therefore be viewed as operational safeguards requiring human oversight rather than as a substitute for real-world safety governance.

6. Conclusion and Future Research

This study examined the trade-off among commercial efficiency, operational resilience, and modeled personnel safety in a volatile pharmaceutical distribution setting. A hybrid Policy Orchestrator was developed by combining LHS-based scenario generation, simulation-based policy selection, and multi-output Random Forest regression. The retained dataset comprises 1,000 operational contexts evaluated under three candidate regimes, yielding 3,000 context–policy records and 1,000 optimal labels.

The retained episodic evaluation comprises a 90-day steady-state horizon and three independent 20-day shock horizons. The evidence supports three main conclusions: (1) adaptive policy selection can adjust routing behavior to changing operating conditions, while the retained steady-state results reveal a measurable cost–resilience trade-off; (2) the security-crisis episode demonstrates that explicit safety guardrails can reduce modeled danger-node traversals to zero within the simulated environment, while shifting cost toward unmet-demand penalties; and (3) the framework's principal value lies in transparent, safety-constrained decision support, with behavior bounded by predefined operational rules rather than autonomous moral agency.

Benchmark and robustness scope should also be stated explicitly. The retained study contains one fixed-policy external comparator rather than separate implementations of robust optimization, stochastic programming, or an independent simulation-optimization benchmark. The three candidate regimes used to generate supervised labels are internal policy alternatives and therefore do not constitute independent algorithmic baselines. Accordingly, the present results support comparison against a controlled static benchmark, not a claim of superiority over all alternative optimization paradigms. The robustness evidence combines LHS context-space variation with the one-factor-at-a-time parameter sensitivity reported in Section 3.5 and Figure 7; the latter provides local sensitivity evidence within the tested ranges rather than a full factorial or global robustness assessment.

Several limitations should be acknowledged. First, the final retained experimental evidence supports a 150-day episodic evaluation consisting of a 90-day steady-state horizon and three 20-day shock horizons. Second, the simulation-based policy-labeling stage used n = 10 independent replications for each candidate policy combination, with a documented 5% relative-error criterion for the 95% confidence-interval half-width of mean total cost, while the final episodic comparison used n = 20 independent replications per policy. Third, the supervised learning stage preserves 1,000 operational contexts, each evaluated under three candidate regimes, producing 1,000 optimal labels. Fourth, scenario-specific Welch statistics are reported for the retained comparative outcomes with Welch–Satterthwaite degrees of freedom, two-sided p-values, confidence intervals, and effect sizes. Fifth, although the network is geographically grounded in 120 actual pharmacy locations represented by real latitude/longitude coordinates, the operational attributes used in the simulation—including demand, product profiles, and capacities—are synthetic because historical operational records were unavailable at the time of the study. The present study therefore does not constitute historical or live operational validation; it provides simulation evidence on a geographically realistic network. Historical supplier records have been requested and, subject to data access, will support field-grounded validation in subsequent work. Sixth, the study uses one fixed-policy benchmark and does not implement independent robust, stochastic, or alternative adaptive optimization algorithms; conclusions are consequently limited to the controlled benchmark used here. Seventh, the robustness analysis combines LHS context variation with one-factor-at-a-time perturbations of risk, stockout, and safety-threshold parameters; the emergency bonus was evaluated at the objective level only. Full factorial or global sensitivity analysis remains future work. Finally, the humanitarian episode does not preserve a consistent pharmacy-level coverage denominator, so percentage coverage gains are not reported.

Overall, the results support the feasibility of a transparent, context-aware policy-orchestration approach for volatile pharmaceutical logistics. The contribution is best understood as a simulation-based demonstration of adaptive, safety-constrained decision support relative to the retained fixed-policy benchmark, rather than evidence that the orchestrator is universally superior, independently validated in the field, or capable of autonomous moral agency.

Acknowledgement
This research was not funded by any grant.

Conflicts of Interest
The authors declare no conflicts of interest.

References
[1] Kovács, G., & Falagara Sigala, I. (2021). Lessons learned from humanitarian logistics to manage supply chain disruptions. Journal of Supply Chain Management, 57(1), 41–49. https://doi.org/10.1111/jscm.12253
[2] Ivanov, D. (2020). Predicting the impacts of epidemic outbreaks on global supply chains: A simulation-based analysis on the coronavirus outbreak (COVID-19/SARS-CoV-2) case. Transportation Research Part E: Logistics and Transportation Review, 136, 101922. https://doi.org/10.1016/j.tre.2020.101922

[3] Chopra, S., & Meindl, P. (2019). Supply chain management: Strategy, planning & operation. In Das Summa Summarum des Management: Die 25 wichtigsten Werke für Strategie, Führung und Veränderung (pp. 265–275). Springer. https://doi.org/10.1007/978-3-8349-9320-5_22
[4] Zhu, G., Chou, M. C., & Tsai, C. W. (2020). Lessons learned from the COVID-19 pandemic exposing the shortcomings of current supply chain operations: A long-term prescriptive offering. Sustainability, 12(14), 5858. https://doi.org/10.3390/su12145858
[5] Milewski, D. (2022). Managerial and economical aspects of the just-in-time system "lean management in the time of pandemic". Sustainability, 14(3), 1204. https://doi.org/10.3390/su14031204
[6] Tomasini, R., & Van Wassenhove, L. (2009). Humanitarian logistics. Palgrave Macmillan. https://doi.org/10.1057/9780230233485
[7] Jabbarzadeh, A., Fahimnia, B., & Seuring, S. (2014). Dynamic supply chain network design for the supply of blood in disasters: A robust model with real world application. Transportation Research Part E: Logistics and Transportation Review, 70, 225–244. https://doi.org/10.1016/j.tre.2014.06.003
[8] Musbah, J., & Badi, I. (2026). Strategic selection of resilient pharmaceutical replenishment policies: A human-in-the-loop (HITL) Z-number SWOT-VIKOR framework. Science, Engineering and Technology, 6(S1), 38–55. https://doi.org/10.54327/set2026/v6.iS1.350
[9] Christopher, M., & Peck, H. (2004). Building the resilient supply chain. International Journal of Logistics Management, 15(2), 1–14. https://doi.org/10.1108/09574090410700275
[10] Snyder, L. V. (2006). Facility location under uncertainty: A review. IIE Transactions, 38(7), 547–564. https://doi.org/10.1080/07408170500216480
[11] Ivanov, D., & Dolgui, A. (2020). Viability of intertwined supply networks: Extending the supply chain resilience angles towards survivability. A position paper motivated by COVID-19 outbreak. International Journal of Production Research, 58(10), 2904–2915. https://doi.org/10.1080/00207543.2020.1750727
[12] Ivanov, D., & Dolgui, A. (2021). A digital supply chain twin for managing the disruption risks and resilience in the era of Industry 4.0. Production Planning & Control, 32(9), 775–788. https://doi.org/10.1080/09537287.2020.1768450
[13] Musbah, J., Badi, I., & Pamucar, D. (2026). Dynamic replenishment policies for vendor-managed inventory under stochastic demand: A simulation-based comparative study. Acta Technica Jaurinensis, 19(3), 195–205. https://doi.org/10.14513/actatechjaur.00960
[14] Lepenioti, K., Bousdekis, A., Apostolou, D., & Mentzas, G. (2020). Prescriptive analytics: Literature review and research challenges. International Journal of Information Management, 50, 57–70. https://doi.org/10.1016/j.ijinfomgt.2019.04.003
[15] Guan, Q., Cao, H., Jia, L., Yan, D., & Chen, B. (2025). Synergetic attention-driven transformer: A deep reinforcement learning approach for vehicle routing problems. Expert Systems with Applications, 274, 126961. https://doi.org/10.1016/j.eswa.2025.126961
[16] Oroojlooyjadid, A., Nazari, M., Snyder, L. V., & Takáč, M. (2022). A deep Q-network for the beer game: Deep reinforcement learning for inventory optimization. Manufacturing & Service Operations Management, 24(1), 285–304. https://doi.org/10.1287/msom.2020.0939
[17] Musbah, J., Badi, I., & Bouraima, M. B. (2025). Optimizing time-based heuristics for resilient VMI replenishment: A simulation-optimization approach. Journal of Industrial Intelligence, 3(2). https://doi.org/10.56578/jii030204
[18] Thejasree, P., Manikandan, N., Vimal, K. E. K., Sivakumar, K., & Krishnamachary, P. C. (2024). Applications of machine learning in supply chain management—A review. In K. E. K. Vimal, S. Rajak, V. Kumar, R. S. Mor, & A. Assayed (Eds.), Industry 4.0 Technologies: Sustainable Manufacturing Supply Chains: Volume 1—Theory, Challenges, and Opportunity (pp. 73–82). Springer. https://doi.org/10.1007/978-981-99-4819-2_6
[19] Govindan, K., Mina, H., & Alavi, B. (2020). A decision support system for demand management in healthcare supply chains considering the epidemic outbreaks: A case study of coronavirus disease 2019 (COVID-19). Transportation Research Part E: Logistics and Transportation Review, 138, 101967. https://doi.org/10.1016/j.tre.2020.101967
[20] Holguín-Veras, J., Pérez, N., Jaller, M., Van Wassenhove, L. N., & Aros-Vera, F. (2013). On the appropriate objective function for post-disaster humanitarian logistics models. Journal of Operations Management, 31(5), 262–280. https://doi.org/10.1016/j.jom.2013.06.002
[21] Friday, D., Savage, D. A., Melnyk, S. A., Harrison, N., Ryan, S., & Wechtler, H. (2021). A collaborative approach to maintaining optimal inventory and mitigating stockout risks during a pandemic: Capabilities for enabling health-care supply chain resilience. Journal of Humanitarian Logistics and Supply Chain Management, 11(2), 248–271. https://doi.org/10.1108/JHLSCM-07-2020-0061
[22] Musbah, J., & Badi, I. (2026). LLM-assisted virtual expert weight elicitation in pharmaceutical supply chains: A Z-number multi-agent framework. Intelligent Systems Research and Applications Journal, 2, 27–39. https://doi.org/10.59543/mmhqdg22