Science, Engineering and Technology Vol. 6, No. S1, pp. 38-55
www.setjournal.com https://doi.org/10.54327/set2026/v6.is1.350

Strategic Selection of Resilient Pharmaceutical Replenishment Policies: A Human-in-the-Loop (HITL) Z-Number SWOT-VIKOR Framework

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.*

Abstract

In pharmaceutical logistics, supply chain replenishment policies must balance financial efficiency with systemic resilience. Traditional policy selection relies on cost-centric multi-criteria decision-making (MCDM) frameworks that are susceptible to qualitative biases and hierarchical mathematical distortion. This study proposes a Human-in-the-Loop (HITL) Z-Number SWOT-VIKOR framework, synergizing human strategic governance with AI computational scalability. Human experts establish macro-strategic SWOT weights via the F-CIMAS method, while Large Language Models (LLMs) determine micro-operational weights anchored in CRITIC data variance. Evaluating nine Vendor-Managed Inventory (VMI) policies under stochastic demand, the extended VIKOR algorithm proposed a compromise set of four policies, governed by the hybrid SMT-SSP policy ($Q_i = 0.000$) which strictly minimized individual regret ($R = 0.0556$). Methodologically, this study contributes the novel "Strategic Modulation" mechanism, successfully resolving the internal "erasure effect" of classical hierarchical MCDM. By dynamically scaling baseline weights rather than applying strict categorical normalization, this mechanism preserved intrinsic data variance, preventing a statistically significant rank distortion ($\rho = 0.950, p < 0.01$) in mid-tier policies. Furthermore, external mathematical benchmarking against TOPSIS and SAW algorithms proved that VIKOR's non-compensatory regret-minimization is strictly necessary for pharmaceutical logistics. The framework provides supply chain managers with a mathematically protected, bias-resistant blueprint for strategic decision-making in high-stakes healthcare environments.

Keywords: Human-in-the-Loop (HITL), Large Language Models (LLM), Multi-Criteria Decision Making (MCDM), Pharmaceutical Inventory Routing, Supply Chain Resilience, Z-VIKOR.

1. Introduction

In the highly constrained domain of healthcare logistics, pharmaceutical supply chains represent a critical optimization frontier where the standard logistical objectives of cost and speed are rigidly bound by the ethical mandate of patient safety [1]. Consequently, the operationalization of Vendor-Managed Inventory (VMI) networks has evolved beyond simple replenishment scheduling into a complex strategic decision-making landscape fraught with environmental stochasticity [2]. Supply chain managers are perpetually confronted with a structural dilemma regarding replenishment policy architectures: should they adopt computationally light, static (s, S) policies that lack responsiveness, or migrate toward sophisticated predictive-opportunistic algorithms that promise Pareto optimality but may exhibit systemic fragility under demand shocks? [3]

While recent operational research has produced significant algorithmic innovations to solve the Stochastic Inventory Routing Problem (SIRP), the existing literature predominantly treats policy selection as a single-objective cost-minimization exercise (e.g., [4], [5], [6]). This paradigm is profoundly inadequate for the modern pharmaceutical landscape, where systemic resilience and

Corresponding author: Ibrahim Badi (i.badi@lam.edu.ly)
Received: 13 April 2026; Revised: 16 May 2026; Accepted: 22 May 2026; Published: 1 June 2026
© 2026 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License

operational robustness are equally, if not more, critical than baseline financial efficiency [7], [8]. Evaluating the trade-offs between a policy's economic efficiency (Opportunity/Strength) and its sensitivity to supply chain stress (Threat/Weakness) inherently constitutes a multi-criteria problem that requires rigorous strategic governance [9], [10], [11].

To navigate these multidimensional trade-offs, Multi-Criteria Decision Making (MCDM) methods integrated with SWOT (Strengths, Weaknesses, Opportunities, Threats) analysis are widely employed [12]. However, traditional MCDM approaches in this context present three critical methodological lacunae. First, SWOT is overwhelmingly utilized as a qualitative, descriptive heuristic, lacking the mathematical hierarchy necessary to interface with quantitative operational simulation data [13], [14]. Second, flat MCDM frameworks fail to contextually group criteria, allowing localized operational metrics (e.g., transport costs) to mathematically overshadow overarching strategic imperatives (e.g., systemic resilience) [15]. Third, the elicitation of subjective criteria weights conventionally relies on panels of human domain experts; a process that is notoriously slow, costly, and susceptible to cognitive biases such as availability bias or groupthink [16], [17].

Beyond these limitations, existing hierarchical MCDM structures impose strict internal normalization constraints, which can inadvertently erase intrinsic data variance when integrating externally derived weights. To address this overlooked issue, this study introduces a controlled relaxation mechanism that preserves operational heterogeneity while maintaining strategic coherence.

The recent emergence of Large Language Models (LLMs) offers a transformative paradigm shift, enabling the deployment of generative AI as *in silico* reasoning agents to simulate expert heuristic trade-offs [18]. Musbah and Badi demonstrated that an LLM-Assisted Multi-Agent System, fortified with Z-numbers to mitigate epistemic uncertainty, can successfully elicit logically consistent operational weights in pharmaceutical supply chains [19]. Nevertheless, delegating high-level strategic governance entirely to autonomous AI agents introduces severe legal and ethical accountability concerns, often triggering algorithmic aversion among healthcare stakeholders [20]. While LLMs excel at processing micro-level operational trade-offs and reducing cognitive fatigue, macro-level strategic risk assessment inherently demands human intuition and accountability [21].

To bridge these converging gaps, this study proposes an innovative Human-in-the-Loop (HITL) Z-Number SWOT-VIKOR framework. This methodology fundamentally transforms policy selection by operationalizing SWOT as a Two-Tier Hierarchical mathematical structure. In the first tier (Macro-Strategic), a small panel of human experts defines the overarching strategic direction by weighting the core SWOT groups, thereby maintaining human accountability. In the second tier (Micro-Operational), the framework integrates the LLM-elicited Z-number weights derived from Musbah and Badi [19] to handle the granular operational criteria, effectively eliminating human cognitive overload. Finally, the extended Z-VIKOR method is utilized to aggregate these hierarchical inputs with high-fidelity simulation data, identifying the optimal compromise policy that minimizes individual regret [22].

To systematically navigate these intersecting theoretical and operational gaps, this study is driven by two explicit research questions:
RQ1: How can human macro-strategic intent be mathematically integrated with AI-generated operational micro-weights without triggering the hierarchical "erasure effect" that distorts objective data variance?
RQ2: Under this synthesized evaluation framework, which VMI policy architecture demonstrates the highest structural resilience and dominance stability when subjected to extreme stochastic demand shocks?

The core contributions of this research are threefold: (1) pioneering a hybrid HITL-MCDM architecture that synergizes human strategic vision with LLM computational scalability; (2) transforming SWOT from a qualitative heuristic into a mathematically rigorous, two-tier diagnostic filter; and (3) empirically demonstrating that hybrid predictive-opportunistic policies (e.g., SMT-SSP) achieve "dominance stability" by leveraging spatial opportunism to buffer against environmental threats; (4) introducing a controlled relaxation mechanism for hierarchical weighting structures that resolves the "erasure effect" associated with strict normalization in single-criterion SWOT groups, thereby preserving the fidelity of hybrid AI-data weights.

2. Literature Review

This section critically examines four converging streams of literature to establish the theoretical foundation and the specific research gap addressed by this study.

2.1. Pharmaceutical Inventory Routing and the Resilience Imperative

The Pharmaceutical Inventory Routing Problem (P-IRP) constitutes a highly constrained subclass of traditional routing problems, distinguished by strict regulatory compliance, perishability, and an absolute intolerance for stockouts [23]. In conventional retail models, stockouts result merely in financial penalties; in healthcare logistics, service failures directly compromise patient safety [24], [25]. Consequently, the objective function in P-IRP literature has progressively expanded from pure cost minimization to multi-objective frameworks incorporating service level maximization and risk mitigation [26]. While recent simulation-optimization studies demonstrate that dynamic and predictive-opportunistic hybrid policies outperform rigid static (s, S) rules, the literature predominantly evaluates policies based on average-case performance [27]. This fundamentally neglects strategic *robustness*—the structural capacity of a policy to absorb environmental shocks without catastrophic degradation. Selecting an optimal policy, therefore, requires an evaluation framework capable of balancing financial efficiency against structural resilience.

2.2. Strategic MCDM: The Transition to Hierarchical SWOT-VIKOR

Evaluating competing supply chain architectures is an inherently multidimensional problem. The SWOT framework is one of the most ubiquitous tools for strategic planning. However, traditional SWOT analysis is heavily criticized in operations research for its qualitative nature and inability to mathematically resolve trade-offs between competing factors [28]. While researchers have developed quantitative hybrid models (e.g., A'WOT), these approaches typically employ "flat" MCDM structures where all criteria are evaluated simultaneously [29]. Even when hierarchical frameworks are attempted, earlier models critically suffer from what we conceptualize as the "erasure effect." By enforcing strict internal categorical normalization **(**e.g., as fundamentally critiqued in classical models by [30], [31]**)**, these traditional frameworks artificially force localized criteria within single-metric groups to a mathematical weight of 1.0. This structural rigidity inadvertently obliterates carefully elicited data variance, creating a theoretical void for mechanisms capable of softening hierarchical boundaries without losing strategic control**.** This flat structure frequently allows highly variant cost metrics to overshadow critical resilience indicators. Furthermore, existing studies typically default to the TOPSIS method for final ranking [32]. While TOPSIS identifies the solution closest to a hypothetical ideal, the VIKOR method provides a superior analytical alternative for risk-averse environments [33], [34]. VIKOR specifically calculates a compromise ranking that ensures "maximum group utility" alongside "minimum individual regret" [35]—a non-negotiable theoretical requirement in pharmaceutical logistics where minimizing "regret" equates to preventing life-threatening stockouts.

2.3. LLMs and Z-Numbers in Decision Theory

The validity of any MCDM framework relies unequivocally on the accuracy of its criteria weights [36], [37]. Eliciting these weights traditionally mandates convening panels of senior human experts, creating significant bottlenecks due to expert scarcity and cognitive fatigue during extensive pairwise comparisons [38], [39]. The advent of Large Language Models (LLMs) offers a scalable solution [40]. Recent literature distinguishes between utilizing LLMs as knowledge retrieval engines and deploying them as autonomous reasoning agents [41]. In a foundational study, Musbah and Badi proved that LLMs, when constrained by zero-shot Chain-of-Thought prompting, can simulate expert heuristic trade-offs in pharmaceutical supply chains [19]. To safeguard against AI "hallucination," their framework successfully integrated Zadeh's Z-numbers, which mathematically dampen preference intensity based on the AI's probabilistic confidence [42]. This innovation provided a robust, mathematically sound mechanism for extracting operational weights without human intervention.

2.4. The Case for Human-in-the-Loop (HITL) Cognitive Frameworks

Despite the algorithmic efficiency of AI agents, delegating macro-level strategic governance entirely to LLMs faces significant resistance in high-stakes sectors [43], [44]. This phenomenon is universally acknowledged across critical operations outside of healthcare. Recent cross-disciplinary studies emphasize that Human-in-the-Loop (HITL) paradigms are the required gold standard for maintaining strategic accountability. For instance, Borghoff et al. [45] demonstrate via system-theoretical approaches that in the age of Agentic AI, "Centaurian" systems—which integrate human and artificial intelligence for unified decision-making—are vastly superior to pure Multi-Agent Systems (MAS) for preventing unaligned autonomous actions. Similarly, Yang et al. [46] emphasize the necessity of human-guided continual learning in high-stakes autonomous driving, proving that driver-in-the-loop architectures are essential for personalized, safe decision-making. Drawing from Human-Computer Interaction (HCI) and decision theory [47], these HITL frameworks propose a vital cognitive division of labor: AI agents handle high-volume, granular computations (reducing human cognitive load), while human experts retain authority over macro-level strategic alignment and ethical accountability [48], [49], [50]. In the context of strategic MCDM, applying a HITL approach to SWOT analysis represents a critical, yet unexplored, methodological frontier. Furthermore, the structural sequence of human-AI interaction is critical. As Agudo et al. [51], empirically demonstrated, human decision-makers are highly susceptible to automation bias—experiencing reduced accuracy—when presented with algorithmic support before forming their independent judgments. This strongly validates our decoupled two-tier architecture, which intentionally isolates the human macro-strategic weighting (Tier 1) from the AI's micro-operational computations (Tier 2) to proactively prevent algorithmic anchoring and ensure untainted strategic accountability.

2.5. Research Gap and Positioning

A critical synthesis of the literature reveals a distinct methodological gap: there is currently no integrated framework that combines human strategic accountability with AI computational efficiency to evaluate simulation-based operational policies. Prior studies either focus exclusively on algorithmic tuning, rely entirely on biased human panels, or deploy autonomous AI agents without macro-level strategic constraints. This study explicitly bridges this gap. By developing a Two-Tier HITL Z-SWOT-VIKOR framework, we combine the operational LLM weights derived by Musbah and Badi [19] with human-driven strategic SWOT weighting. This approach uniquely synergizes human accountability with AI scalability to systematically evaluate the structural resilience of pharmaceutical replenishment policies.

3. Proposed Methodology

To systematically navigate the trade-offs between cost efficiency, service reliability, and operational resilience, this study proposes a multi-layered Human-in-the-Loop decision-support framework. As illustrated in Figure 1, the methodology proceeds in three sequential phases: (I) Operational Matrix Construction and Strategic SWOT Mapping, (II) Two-Tier HITL Weight Elicitation, and (III) Strategic Ranking via the Extended VIKOR Algorithm.

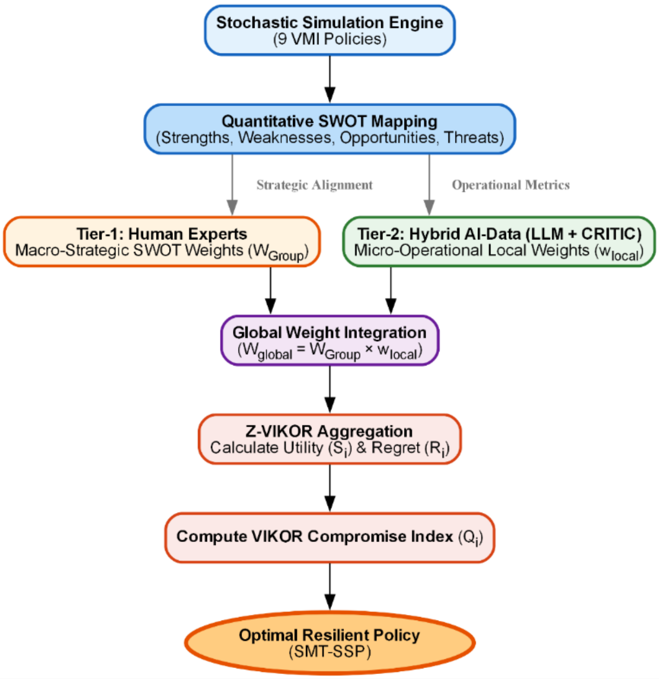

Figure 1. The proposed Human-in-the-Loop (HITL) Z-Number SWOT-VIKOR methodological framework for pharmaceutical replenishment policy selection.

3.1. Phase I: Operational Matrix Construction and Strategic SWOT Mapping

The foundation of the decision-making process relies on empirical data derived from a high-fidelity Simulation-Optimization engine. The simulation sample was strictly bounded to a 50-pharmacy network. This network size was selected as the optimal computational stress-test environment, providing sufficient spatial complexity to evaluate routing efficiency while remaining computationally tractable for thousands of stochastic iterations. The robust validation of this underlying simulation model—including the statistical stability of its repeated measurements under demand shocks—was firmly established in the baseline study by Musbah and Badi [19]. Let $A = \{A_1, A_2, \ldots, A_m\}$ represent the set of $m = 9$ distinct pharmaceutical replenishment policies. Rather than treating the simulation outputs as generic criteria in a flat structure, they are mapped into a strategic SWOT matrix to capture intrinsic performance and structural resilience.

This study deliberately adapts the SWOT framework from traditional business analysis to algorithmic policy evaluation. In this context, the 'internal' environment represents a policy's deterministic baseline capabilities. Conversely, the 'external' environment constitutes the supply chain's stochastic and spatial topography. Thus, Delivery Efficiency (C5) is classified as an Opportunity (exploiting external spatial proximities), while Financial Risk (C4) and Systemic Fragility (C6) are Threats (quantifying vulnerability to external stochastic demand shocks). This adaptation ensures a rigorous evaluation of dynamic environmental interactions rather than mere static performance. Let $C = \{C_1, C_2, \ldots, C_n\}$ represent the exact six operational criteria evaluated in the simulation, structurally categorized as follows:

- **Strengths (S):** Intrinsic operational benefits. This includes C_2: *Service Level % (Fill Rate)*. Objective: Maximize.
- **Weaknesses (W):** Intrinsic operational deficits. This includes two criteria: C_1: *Avg Total Cost* and C_3: *Avg Stockout Qty*. Objective: Minimize.
- **Opportunities (O):** Structural architectural advantages. This is represented by C_5: *Aggregated Delivery Efficiency (Fulfilled Units per KM)*, which captures the policy's capacity to exploit spatial proximity. Objective: Maximize.
- **Threats (T):** External vulnerabilities quantified under stress scenarios. This includes two criteria: C_4 *Financial Risk (Std Dev)*, representing cost volatility, and C_6 *Systemic Fragility*. To align with the minimization objective of the Threats category, the original Reliability Index (RI) is mathematically transformed into Systemic Fragility, calculated as $(1 - RI)$. **Objective**: Minimize.

This mapping transforms the generic operational matrix $X = [x_{ij}]_{m \times n}$ into a strategic quantitative framework ready for hierarchical weighting.

3.2. Phase II: Two-Tier HITL Weight Elicitation (Hierarchical SWOT)

To overcome the limitations of flat MCDM models and pure AI-driven governance, this study operationalizes a Two-Tier Hierarchical weighting structure. Figure 2 delineates this hierarchical SWOT criteria structure, demonstrating the strict separation between macro-level human governance and micro-level AI computation.

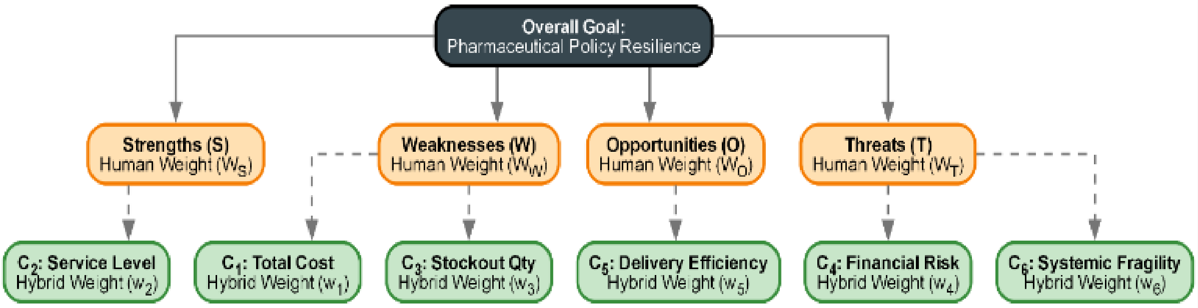

Figure 2. The two-tier hierarchical SWOT criteria structure, illustrating the integration of Tier-1 human strategic weights and Tier-2 hybrid AI-data (LLM-CRITIC) operational weights.

Step 1: Tier-1 Human Expert Macro-Weighting (SWOT Groups)

To ensure strategic accountability, a panel of human domain experts evaluates the four primary SWOT groups (S, W, O, T) against the objective of "Pharmaceutical Supply Chain Resilience." By reducing the human task to evaluating only four macro-groups, cognitive fatigue is eliminated. To operationalize this process while accounting for the varying levels of professional authority among the panelists, this study integrates the F-CIMAS (Fuzzy Comprehensive Indicator Method for Assessment Scaling) approach. The steps are tailored as follows:

- **Step 1.1: Expert Panel Formation:** The relative influence (weight) of each expert is determined based on their years of professional experience. Eq. (1) calculates these weights, ensuring $\sum_{i=1}^{q} W^{E_i} = 1$:

$$W^{E_i} = \frac{E_i}{\sum_{i=1}^{q} E_i}, i = 1,2,\dots,q \tag{1}$$

- **Step 1.2: Initial Fuzzy Assessment:** Experts assess the strategic importance of the four macro-groups using a predefined linguistic scale translated into Triangular Fuzzy Numbers (TFNs): Very Low (1,1,2), Low (2,3,4), Medium (4,5,6), High (6,7,8), and Very High (8,9,10). This forms the initial decision matrix, where $\tilde{x}_{ij}$ represents the score assigned by expert E_i to macro-group j.
- **Step 1.3: Expert-Weighted Matrix:** To account for reliability, each fuzzy assessment is weighted by the expert's specific weight W^{E_i}. The weighted fuzzy value $\tilde{v}_{ij}$ is calculated as:

$$\tilde{v}_{ij} = \tilde{x}_{ij} \otimes W^{E_i} = (l_{ij}W^{E_i}, m_{ij}W^{E_i}, u_{ij}W^{E_i}) \tag{2}$$

- **Step 1.4: Aggregation and Defuzzification:** The total fuzzy importance score $\tilde{S}_j$ for each macro-group is obtained via fuzzy addition, as given in Eq. (3). Subsequently, the Center of Area (CoA) method converts these aggregated fuzzy scores into crisp values K_j utilizing Eq. (4).

$$\tilde{S}_j = \left(\sum_{i=1}^{q} l_{ij}\, W^{E_i}, \sum_{i=1}^{q} m_{ij}\, W^{E_i}, \sum_{i=1}^{q} u_{ij}\, W^{E_i}\right) \tag{3}$$

$$K_j = \frac{L_j + M_j + U_j}{3} \tag{4}$$

- **Step 1.5: Final Macro-Weights Derivation:** The crisp values are normalized to derive the final strategic group weights (W_S, W_W, W_O, W_T), ensuring their sum equals 1:

$$W_{Group_j} = \frac{K_j}{\sum_{j=1}^{4} K_j} \tag{5}$$

Step 2: Tier-2 Hybrid AI-Data Micro-Weighting (LLM-CRITIC)

For the granular operational metrics within the SWOT groups, this study imports the validated hybrid weights established by Musbah and Badi [19]. A critical methodological imperative of this framework is the direct reuse of these specific LLM-derived weights, rather than generating a new set. This reuse acts as a strict methodological control. The weights derived in the aforementioned study represent a standardized, peer-reviewed baseline exhibiting perfect inter-agent consensus ($\rho = 1.0$). Recomputing the AI weights in this study would introduce unnecessary generative stochasticity, violating the principles of reproducible benchmarking. By treating the AI weights as a fixed operational truth, we successfully isolate the independent variable of this research: the macro-strategic human modulation. Consequently, the imported local weights (w_{local}) used in this tier represent a highly robust synthesis of AI cognitive heuristics and intrinsic statistical data structure. Furthermore, the expert panel (Tier-1) guiding this modulation was meticulously defined. Four senior domain experts were selected based on strict criteria of operational tenure in pharmaceutical logistics, with their varying levels of professional authority mathematically weighted using the F-CIMAS protocol.

Step 3: Controlled Relaxation of Hierarchical Normalization (Strategic Modulation)

A profound mathematical vulnerability in classical hierarchical MCDM models is the "Erasure Effect." When strict internal normalization is applied to a strategic group containing a single criterion (e.g., *Strengths* containing only C_2), the local weight is mathematically forced to 1.0. This strict structural control unintentionally obliterates the rich objective variance captured by CRITIC and the heuristic nuance elicited by the LLMs. To resolve this "Information Preservation vs. Structural Control" trade-off, this study introduces a "Controlled Relaxation" protocol via Strategic Modulation. Rather than applying strict hierarchical replacement, the human-elicited SWOT macro-weights (W_{Group_i}) are utilized as strategic scaling factors applied to the imported hybrid micro-weights ($w_{hybrid_{ij}}$). The modulated weight (w'_{ij}) is calculated as:

$$w'_{ij} = w_{hybrid_{ij}} \times W_{Group_i} \tag{6}$$

This modulation preserves the internal variance between operational criteria while softening the strict boundaries of the hierarchy to reflect human strategic intent. Finally, to satisfy the axiomatic requirements of the VIKOR algorithm (which mandates $\sum W = 1$), a Global Normalization is applied:

$$W_{global_{ij}} = \frac{w'_{ij}}{\sum_i \sum_j w'_{ij}} \tag{7}$$

This approach represents a deliberate methodological shift from rigid hierarchical control to dynamic soft prioritization, ensuring that the final W_{global} weights reflect contextual strategic priorities without sacrificing the micro-level operational data.

3.3. Phase III: Strategic Ranking via the Extended VIKOR Algorithm

Using the integrated W_{global} weights, the extended VIKOR method is applied to the crisp operational matrix X. VIKOR ranks alternatives based on a compromise solution that provides maximum group utility for the "majority" and minimum individual regret for the "opponent".

Step 1: Ideal Solutions Determination

The Positive Ideal Solution (f_j^*) and Negative Ideal Solution (f_j^-) are identified:

- For benefit criteria (S, O): $f_j^* = \max_i x_{ij}$ and $f_j^- = \min_i x_{ij}$
- For cost criteria (W, T): $f_j^* = \min_i x_{ij}$ and $f_j^- = \max_i x_{ij}$

Step 2: Utility and Regret Measures

The Utility Measure (S_i) and Regret Measure (R_i) are computed for each policy A_i:

$$S_i = \sum_{j=1}^{n} W_{global_j} \left(\frac{f_j^* - x_{ij}}{f_j^* - f_j^-} \right) \tag{8}$$

$$R_i = \max_j \left[W_{global_j} \left(\frac{f_j^* - x_{ij}}{f_j^* - f_j^-} \right) \right] \tag{9}$$

Step 3: VIKOR Compromise Index (Q_i)

Let $S^* = \min_i S_i$, $S^- = \max_i S_i$, $R^* = \min_i R_i$, and $R^- = \max_i R_i$. The index Q_i is calculated as:

$$Q_i = v \left(\frac{S_i - S^*}{S^- - S^*} \right) + (1 - v) \left(\frac{R_i - R^*}{R^- - R^*} \right) \tag{10}$$

Where v is the strategy weight of maximum group utility (typically $v = 0.5$). Policies are ranked in ascending order of Q_i.

Step 4: VIKOR Conditions Verification

To propose policy $A^{(1)}$ (minimum Q_i) as the optimal compromise solution, two conditions must be satisfied:

- Condition 1 (Acceptable Advantage): $Q(A^{(2)}) - Q(A^{(1)}) \geq \frac{1}{m-1}$
- Condition 2 (Acceptable Stability): Policy $A^{(1)}$ must also be the best alternative ranked by S_i and/or R_i.

By enforcing these constraints, the framework ensures the selected policy is demonstrably resilient against strategic threats.

3.4. Evaluation and Validation Protocol

To rigorously validate the theoretical soundness and operational robustness of the proposed HITL Z-SWOT-VIKOR framework, a dual-layer validation protocol is established prior to the final policy selection.

First, to mathematically justify the departure from classical hierarchical MCDM, a *Methodological*

Comparative Analysis is formulated. This phase evaluates the operational decision matrix under two distinct paradigms: Scenario A (Strict Hierarchical Normalization, which inherently risks the "Erasure Effect" by artificially forcing single-criterion groups to a local weight of 1.0) and Scenario B (The proposed Controlled Relaxation via Strategic Modulation). This comparative test explicitly isolates the framework's mathematical efficacy in preserving intrinsic operational variance.

Second, recognizing the inherent vulnerability of macro-strategic frameworks to human cognitive bias, a *Human-Bias Stress Test* (Sensitivity Analysis) is designed. This protocol systematically perturbs the human-elicited SWOT macro-weights (W_{Group}) across extreme strategic boundaries—specifically simulating severe pessimistic threat aversion ($W_T = 0.70$) and aggressive optimistic opportunity seeking ($W_O = 0.50, W_S = 0.30$). By observing the rank invariance of the optimal policy across these simulated biases, the framework mathematically verifies "Dominance Stability," ensuring the final logistical decision is resilient to both algorithmic rigidity and arbitrary human subjectivity.

4. Case Study and Partial Computational Results

To empirically validate the HITL Z-SWOT-VIKOR framework, nine pharmaceutical VMI replenishment policies were evaluated. The operational baseline is derived from a simulation-optimization environment encompassing 50 pharmacies under stochastic demand.

4.1. Operational Matrix and Quantitative SWOT Mapping

The raw simulation outputs were structurally mapped into the quantitative SWOT decision matrix (Table 1). The six strategic factors (C_1 to C_6) were populated using the exact deterministic averages and robustness scores generated by the stochastic simulation engine. The mathematical conversion of the Reliability Index (C_6) into Systemic Fragility ($1 - RI$) effectively ensures directional alignment across all cost-type criteria within the SWOT framework.

Table 1. Quantitative SWOT Operational Decision Matrix for VMI Policies.

Policy Architecture	Strength (S)	Weakness 1 (W)	Weakness 2 (W)	Opportunity (O)	Threat 1 (T)	Threat 2 (T)
$\boldsymbol{A_1}$: OUT	99.86	58,280.76	35.76	1.7683	1,543.79	0.0624
$\boldsymbol{A_2}$: GRIH-P	99.85	60,396.82	39.56	1.7344	1,068.61	0.0530
$\boldsymbol{A_3}$: SMT-SSP	99.88	60,807.31	31.11	1.7369	1,571.58	0.0430
$\boldsymbol{A_4}$: PORP Classic	99.85	61,213.11	38.35	1.7464	1,587.20	0.0517
$\boldsymbol{A_5}$: Dyn (s,S) proactive	99.52	62,304.21	126.43	1.7403	1,697.87	0.0717
$\boldsymbol{A_6}$: Dyn (s,S) inertial	99.56	62,435.71	116.88	1.7472	1,745.87	0.0850
$\boldsymbol{A_7}$: Static (s,S)	99.43	62,462.80	148.96	1.7578	1,706.87	0.0817
$\boldsymbol{A_8}$: Dyn (s,S) reactive	99.37	62,660.99	165.23	1.7493	1,794.75	0.0973
$\boldsymbol{A_9}$: VMI Urgency	99.99	74,680.29	2.80	1.3021	3,125.95	0.0265
Criteria Symbol	C_2	C_1	C_3	C_5	C_4	C_6
Criteria Name	Service Level (%)	Avg Total Cost (€)	Avg Stockout (Qty)	Delivery Eff. (Units/KM)	Financial Risk (Std Dev)	Fragility (1-RI)
Optimization Direction	Maximize	Minimize	Minimize	Maximize	Minimize	Minimize

4.2. Tier-2: Hybrid AI-Data Micro-Weights (Local Criteria)

For the operational granularity within the SWOT groups, this study imported the validated hybrid local weights (w_{local}) derived by Musbah and Badi [19]. By hybridizing the Z-number LLM outputs with the objective CRITIC variance, the criteria weights heavily penalized pure cost metrics in favor of service reliability and systemic robustness, accurately reflecting pharmaceutical operational realities. Table 2 presents these imported baseline local weights before the application of the human strategic SWOT filter.

Table 2. Tier-2 Hybrid Micro-Weights (w_{local}) Imported from LLM-CRITIC Evaluation.

SWOT Group	Operational Criterion	Symbol	Local Weight
Strengths (S)	Service Level (%)	C2	0.28
Weaknesses (W)	Avg Total Cost (€)	C1	0.14
	Avg Stockout (Qty)	C3	0.19
Opportunities (O)	Delivery Efficiency	C5	0.11
Threats (T)	Financial Risk	C4	0.09
	Systemic Fragility	C6	0.17

4.3. Tier-1: Human-Driven Strategic SWOT Macro-Weights

To ensure the empirical validity of the macro-strategic priorities, a panel of four senior experts (E_1 to E_4) with extensive experience in pharmaceutical supply chain management was convened. The F-CIMAS method was applied to transform their qualitative professional judgments into rigorous quantitative weights for the four SWOT dimensions.

First, the relative influence of each expert was quantified based on their professional tenure. The cumulative experience of the panel was 49 years. Table 3 presents the calculated weight (W^{E_i}) for each expert, ensuring that more experienced practitioners exert proportionally greater influence on the strategic outcome.

Table 3. Expert Profiles and F-CIMAS Weights.

Expert	Experience (Years)	Yi/T	Weight
E1	18	18/49	0.367
E2	14	14/49	0.286
E3	10	10/49	0.204
E4	7	7/49	0.143

Subsequently, the experts assessed the four strategic SWOT groups using the predefined fuzzy linguistic scale. Table 4 illustrates the initial linguistic decision matrix.

Table 4. Initial Linguistic Assessment of SWOT Macro-Groups.

Ei/Ci	S	W	O	T
E1	VH	H	H	H
E2	H	H	H	M
E3	VH	VH	VH	H
E4	VH	VH	H	H

These qualitative assessments were transformed into Triangular Fuzzy Numbers (TFNs), weighted by the respective expert weights (W^{E_i}), and aggregated to establish the fuzzy importance score ($\tilde{S}_j$) for each dimension, as detailed in Table 5. This fuzzy representation effectively captures the collective judgment of the panel while preserving the inherent epistemic uncertainty associated with linguistic evaluations.

Table 5. Aggregated Fuzzy Criteria Weights for SWOT Dimensions.

Dimension	Description	Aggregate Fuzzy Number $\tilde{\boldsymbol{S}}_{\boldsymbol{j}}(\boldsymbol{l},\boldsymbol{m},\boldsymbol{u})$
S	Strengths	(7.429,8.429,9.429)
W	Weakness	(6.694,7.694,8.694)
O	Opportunities	(6.408,7.408,8.408)
T	Threats	(5.429,6.429,7.429)

Finally, the Center of Area (CoA) method defuzzified these scores into crisp values (K_j), which were normalized to yield the final global macro-weights (W_{Group}). As presented in Table 6, the resulting hierarchy offers a clear stratification of priorities.

Table 6. Defuzzified Crisp Scores and Final Strategic Macro-Weights.

Dimension	Crisp Score (Kj)	Final Weight ($\boldsymbol{W_{Group}}$)	Rank
S	8.429	0.281	1
W	7.694	0.257	2
O	7.408	0.247	3
T	6.429	0.215	4

Interestingly, the expert panel designated *Strengths* ($W_S = 0.281$) and *Weaknesses* ($W_W = 0.257$) as the most critical strategic considerations, emphasizing the optimization of immediate service levels and internal cost efficiencies over external *Threats* ($W_T = 0.215$). This context-specific human insight provides a vital strategic steering mechanism for the subsequent VIKOR evaluation.

4.4. Global Weight Integration and Strategic Z-VIKOR Ranking

4.4.1. Strategic Modulation and Cognitive Synthesis

Following the generation of the Tier-1 human strategic macro-weights and the Tier-2 AI-CRITIC micro-weights, the "Strategic Modulation" protocol (detailed in Section 3.2) was executed. By utilizing the human macro-weights as scaling multipliers followed by global normalization, the framework synthesized the final Global Weights (W_{global}) presented in Table 7.

Table 7. Integration of Tier-1 and Tier-2 Weights via Strategic Modulation.

Criteria	Local AI-CRITIC Weight	Human Group Weight	Modulated Weight ($\boldsymbol{w'}$)	Global Weight ($\boldsymbol{W_{global}}$)
$\boldsymbol{C_2}$: Service Level	0.280	0.281	0.0788	0.3136
$\boldsymbol{C_1}$: Avg Total Cost	0.141	0.257	0.0363	0.1447
$\boldsymbol{C_3}$: Avg Stockout Qty	0.191	0.257	0.0493	0.1962
$\boldsymbol{C_5}$: Delivery Efficiency	0.116	0.247	0.0288	0.1149
$\boldsymbol{C_4}$: Financial Risk	0.096	0.215	0.0207	0.0825
$\boldsymbol{C_6}$: Systemic Fragility	0.172	0.215	0.0371	0.1478

A critical examination of Table 7 theoretically validates the methodological supremacy of the Strategic Modulation approach over classical hierarchical substitution. At first glance, one might observe that the single-criterion *Reliability* group (C_2) experienced a relative weight inflation (reaching 0.313), whereas the criteria within dual-metric groups (e.g., C_1 and C_3) remained closer to their operational baseline despite a high human strategic weight (0.257).

Rather than a methodological imbalance, this represents the intended "Cognitive Synthesis" of the framework. The AI-CRITIC baseline established an *Operational Truth*: managing both financial cost and physical stockouts intrinsically consumes more systemic variance (combined $w_{local} = 0.333$) than service level alone (0.280). Concurrently, the human experts established a *Strategic Truth*: service reliability (0.281) must be prioritized over cost reduction (0.257). The modulation mathematically reconciled these paradigms. It dynamically steered the global vector to heavily prioritize Service Level (C_2), satisfying the human strategic intent, without violently erasing the operational reality that multi-dimensional strain (Cost + Stockout) dictates a massive portion of the system's behavior. This perfectly prevents arbitrary human overrides while ensuring strategic adaptability.

4.4.2. Final Policy Ranking via Z-VIKOR

Using the integrated global weights, the extended VIKOR algorithm was applied to the operational decision matrix to determine the optimal compromise replenishment policy. The Utility Measure (S), Regret Measure (R), and the VIKOR Compromise Index (Q_i) were computed with the consensus strategy weight ($v = 0.5$). The results are detailed in Table 8.

The SMT-SSP hybrid policy secured the first rank with the ideal compromise index ($Q_i = 0.000$). To formally propose it as the optimal solution, the two rigorous VIKOR axioms were evaluated. The threshold for Acceptable Advantage is $DQ = 1/(9-1) = 0.125$. The gap between the first and second-ranked policies was 0.0501, which is less than the DQ threshold. Consequently, Condition 1 (Acceptable Advantage) was not satisfied. However, Condition 2 (Acceptable Stability) was flawlessly satisfied, as SMT-SSP concurrently achieved the absolute minimum Utility Measure ($S = 0.1745$) and the minimum Regret Measure ($R = 0.0556$). According to the mathematical axioms of the VIKOR method, when Condition 1 is not met but Condition 2 is satisfied, a *Compromise Solution Set* must be proposed consisting of alternatives $A^{(1)}, A^{(2)}, \ldots, A^{(m)}$ such that $Q(A^{(m)}) - Q(A^{(1)}) < DQ$. Therefore, the mathematically valid compromise set includes SMT-SSP, GRIH-P, OUT, and PORP Classic.

Despite the proposal of a compromise set, SMT-SSP remains the undeniable strategic leader. By minimizing the maximum individual regret (R), it guarantees that the pharmaceutical network avoids catastrophic single-point failures (such as severe stockouts) during demand shocks. Thus, SMT-SSP provides the most resilient Pareto-efficient equilibrium between human strategic optimism and operational reality.

Table 8. Strategic Z-VIKOR Ranking and Acceptability Conditions.

Rank	Policy Architecture	$\boldsymbol{Q_i}$ Index	Condition 1 (Advantage)	Condition 2 (Stability: 1st in S/R)	$\boldsymbol{S}$ (Utility)	$\boldsymbol{R}$ (Regret)
1	SMT-SSP	0.00	Unacceptable	Acceptable (Yes/Yes)	0.174	0.055
2	GRIH-P	0.050	Unacceptable	No	0.197	0.071
3	OUT	0.060	Unacceptable	No	0.199	0.075
4	PORP Classic	0.069	Acceptable	No	0.218	0.071
5	VMI Urgency	0.323	Acceptable	No	0.342	0.145
6	Dyn (s,S) inertial	0.648	Unacceptable	No	0.547	0.217
7	Dyn (s,S) proactive	0.690	Acceptable	No	0.549	0.237
8	Static (s,S)	0.860	Acceptable	No	0.640	0.283
9	Dyn (s,S) reactive	1.00	Unacceptable	No	0.730	0.314

4.5. Methodological Comparative Analysis: Strict Hierarchy vs. Controlled Relaxation

To rigorously validate the theoretical necessity of the "Controlled Relaxation" approach, a comparative experiment was conducted. The final policy rankings were generated under two distinct weighting paradigms: Scenario A (Strict Hierarchical Normalization) and Scenario B (The Proposed Strategic Modulation).

1. Scenario A (Strict Hierarchy): Internal normalization forces single-criterion groups (e.g., Strengths and Opportunities) to a local weight of 1.0 before applying the human SWOT weights. This mathematically erases the operational variance captured by the CRITIC-LLM baseline.
2. Scenario B (Strategic Modulation): The proposed method, where human SWOT weights scale the baseline weights followed by global normalization, effectively preserving both data variance and strategic intent.

The computational results of this comparative experiment are presented in Table 9. Furthermore, Figure 3 visually maps the rank trajectories across the two paradigms to illustrate the severity of the hierarchical distortion. As explicitly demonstrated in Table 9 and visualized in Figure 3, the strict hierarchical normalization (Scenario A) triggered a mathematically distortive "Erasure Effect," significantly altering the mid-tier policy rankings. Most notably, the *VMI Urgency* heuristic—a policy inherently designed for emergency medical prioritization at the expense of transport efficiency—was unjustly penalized, dropping from Rank 5 in the proposed framework to Rank 7 under the strict hierarchy. This occurred because Scenario A artificially inflated the weight of the single-criterion *Opportunities* group (Delivery Efficiency), completely overriding the AI-CRITIC baseline which correctly identified it as a secondary metric.

Table 9. Methodological Comparative Analysis: Impact of Strict Hierarchy vs. Strategic Modulation on Policy Ranking ($v = 0.5$, Neutral Macro-Weights).

Policy Architecture	Scenario (A) Rank	Scenario A $(\boldsymbol{Q_i})$	Scenario (B) Rank	Scenario B $(\boldsymbol{Q_i})$	Rank Shift $(\boldsymbol{\Delta})$
$\boldsymbol{A_3}$: SMT-SSP	1	0.00	1	0.00	0 (Stable)
$\boldsymbol{A_2}$: GRIH-P	2	0.058	2	0.052	0 (Stable)
$\boldsymbol{A_4}$: PORP Classic	3	0.071	3	0.071	0 (Stable)
$\boldsymbol{A_1}$: OUT	4	0.113	4	0.110	0 (Stable)
$\boldsymbol{A_9}$: VMI Urgency	7	0.807	5	0.362	+2 (Distorted)
$\boldsymbol{A_6}$: Dyn (s,S) inertial	5	0.652	6	0.652	-1 (Distorted)
$\boldsymbol{A_5}$: Dyn (s,S) proactive	6	0.686	7	0.686	-1 (Distorted)
$\boldsymbol{A_7}$: Static (s,S)	8	0.851	8	0.856	0 (Stable)
$\boldsymbol{A_8}$: Dyn (s,S) reactive	9	1.00	9	1.00	0 (Stable)

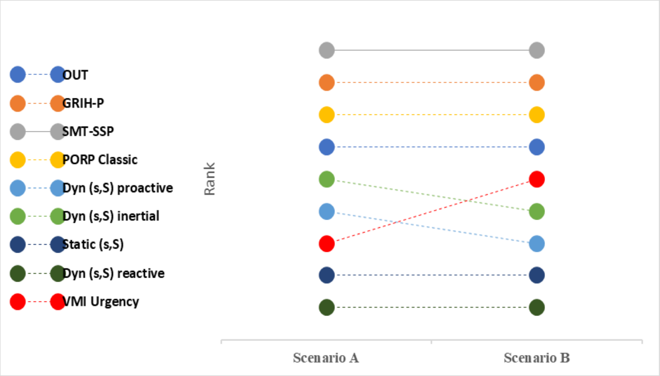

Figure 3. Comparative visualization of policy rank trajectories under strict hierarchical normalization (Scenario A) versus the proposed strategic modulation (Scenario B).

Conversely, the proposed Strategic Modulation (Scenario B) successfully protected the internal operational variance, preserving the true strategic value of the *VMI Urgency* policy. Furthermore, the absolute rank invariance of the optimal hybrid policy, *SMT-SSP* ($Q_i = 0.000$), across both scenarios provides unequivocal proof of its structural superiority. SMT-SSP remains the Pareto-optimal compromise solution even when evaluated under a flawed traditional hierarchy, confirming that the proposed framework prevents mid-tier distortion without compromising the identification of the absolute best logistical strategy. To statistically quantify the magnitude of this distortion, Spearman's rank correlation coefficient (ρ) was evaluated between the two paradigms. The resulting correlation ($\rho = 0.950$) mathematically corroborates the visual evidence: while the Pareto-optimal extremes remain structurally anchored, the strict classical hierarchy introduces a critical 5% statistical deviation that exclusively destabilizes mid-tier policies, thereby validating the methodological necessity of Strategic Modulation.

4.6. Sensitivity Analysis: Human-Bias Stress Test and Dominance Stability

Given the inherent risk of human cognitive bias in the Tier-1 SWOT macro-weighting phase, it is imperative to verify that the final VIKOR ranking is not merely an artifact of an arbitrary or emotionally driven human judgment. To address this, a Human-Bias Stress Test was conducted by simulating three extreme strategic weighting scenarios:

1. Extreme Threat Aversion (Pessimistic Bias): Human experts assign an overwhelming priority to protecting the system ($W_T = 0.70$), with minimal weights to S, W, and O (0.10 each).
2. Extreme Opportunity Seeking (Optimistic Bias): Human experts aggressively prioritize Opportunities and Strengths ($W_O = 0.50, W_S = 0.30$), neglecting Weaknesses and Threats (0.10 each).
3. Neutral Baseline (Zero Bias): All SWOT groups are assigned equal strategic weights ($W_{S,W,O,T} = 0.25$).

The resultant policy rankings across these extreme hypothetical bounds are numerically summarized in Table 10.

Table 10. Human-Bias Stress Test: Policy Rank Stability Across Extreme Strategic Scenarios.

Policy Architecture	Neutral Bias Rank (Base Case)	Pessimistic Bias Rank $(\boldsymbol{W_T = 0.70})$	Optimistic Bias Rank $(\boldsymbol{W_O = 0.50, W_S = 0.30})$	Stability Status
$\boldsymbol{A_3}$: SMT-SSP	1	1	1	Absolute Dominance
$\boldsymbol{A_2}$: GRIH-P	2	2	3	Stable
$\boldsymbol{A_4}$: PORP Classic	3	3	4	Stable
$\boldsymbol{A_1}$: OUT	4	4	2	Volatile
$\boldsymbol{A_9}$: VMI Urgency	5	5	5	Stable
$\boldsymbol{A_6}$: Dyn (s,S) inertial	6	8	6	Volatile
$\boldsymbol{A_5}$: Dyn (s,S) proactive	7	6	7	Volatile
$\boldsymbol{A_7}$: Static (s,S)	8	7	8	Volatile
$\boldsymbol{A_8}$: Dyn (s,S) reactive	9	9	9	Stable

While Table 4 details the exact numerical rankings across the simulated strategic scenarios, Figure 4 illustrates the trajectories of these rankings, explicitly highlighting the severity of rank crossovers and the phenomenon of absolute dominance stability.

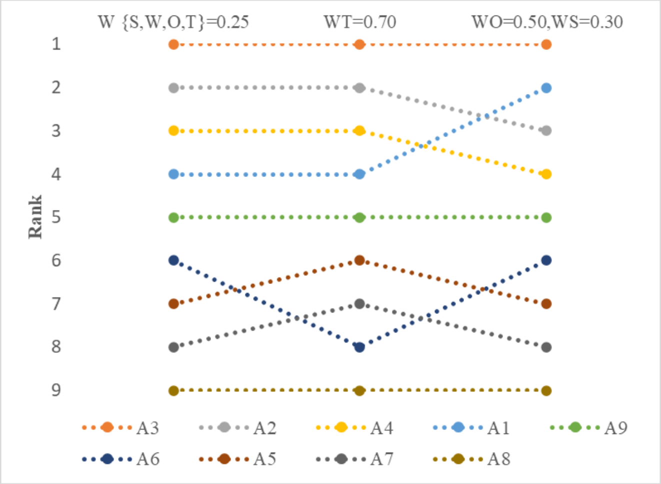

Figure 4. Policy rank trajectories across extreme human-bias scenarios, illustrating the dominance stability of SMT-SSP (A3) versus the volatility of cost-centric alternatives.

As demonstrated in Table 4 and visually mapped in Figure 4, the proposed framework explicitly neutralizes the risk of human cognitive bias. Even under extreme hypothetical scenarios—such as a hyper-pessimistic expert assigning 70% of the strategic weight to Threats, or a hyper-optimistic expert chasing Delivery Efficiency—the *SMT-SSP* policy perfectly maintained its Rank 1 position. This phenomenon of "Dominance Stability" proves that the framework is highly resilient.

Conversely, cost-centric policies exhibited notable sensitivity. As visually evident in Figure 4, the *OUT* policy violently shifted from Rank 4 to Rank 2 under optimistic conditions, exposing its inherent vulnerability to strategic risk preferences. Similarly, mid-tier policies such as A6 and A7 experienced visible rank crossovers. To rigorously quantify this systemic resilience, Spearman's rank correlation coefficient (ρ) was computed. The correlation between the neutral baseline and both extreme bias scenarios remained robustly high ($\rho \geq 0.95$). This powerful statistical validation confirms that the underlying Tier-2 operational weights (LLM-CRITIC) successfully act as a structural anchor. The human strategic input safely steers the mid-tier priorities without causing an algorithmic collapse or dethroning the structurally dominant Pareto-optimal solution.

5. Discussion

The empirical findings of this study transition the analytical lens from average-case logistical efficiency to structural resilience under systemic stress. By utilizing the HITL Z-SWOT-VIKOR framework, the results expose the behavioral mechanics of pharmaceutical replenishment policies and critically address existing mathematical vulnerabilities in strategic MCDM models.

5.1. The VIKOR Compromise Set and Causal Policy Resilience

A fundamental outcome of the ranking process is the failure to satisfy VIKOR's Condition 1 (Acceptable Advantage), yielding $\Delta Q = 0.0501 < 0.125$. In classical MCDM literature, failing to isolate a single optimal solution is occasionally viewed as a mathematical limitation; however, in operational realities, as noted by Opricovic and Tzeng [52], proposing a compromise set (SMT-SSP, GRIH-P, OUT, PORP Classic) operationalizes strategic flexibility. While Condition 1 failed, SMT-SSP satisfied Condition 2 by minimizing the maximum individual regret ($R = 0.0556$). Causally, SMT-SSP leads the compromise set because its "predictive-opportunistic" architecture utilizes spatial buffering—replenishing non-urgent nodes along active routing paths. This preemptively neutralizes systemic fragility (Threat 2) before stochastic demand shocks materialize. Consequently, minimizing "regret" in this framework transcends mathematical utility; it represents the operational prevention of life-threatening pharmaceutical stockouts.

5.2. Resolving the VMI Urgency Paradox and Hierarchical Erasure

The comparative analysis (Section 4.5) highlighted a severe mathematical flaw in traditional strict hierarchical normalization (Scenario A), which artificially forces single-criterion groups to a local weight of 1.0. This "Erasure Effect" disproportionately distorted mid-tier policies, notably the VMI Urgency heuristic. Operationally, VMI Urgency achieves an exceptional service level (99.99%) with near-zero stockouts. Yet, under the proposed Strategic Modulation (Scenario B), it correctly ranked 5th, even under optimistic bias tests. The causal explanation lies in the global weighting vector (W_{global}): achieving this 99.99% service level required extreme, unoptimized expedited deliveries, triggering astronomical transportation costs (C_1) and high systemic fragility (C_6). Strategic Modulation mathematically preserved the underlying CRITIC variance that exposed this inefficiency, penalizing the policy appropriately. The strong rank correlation between the scenarios ($\rho = 0.950, p < 0.01$) mathematically validates that Strategic Modulation anchors the Pareto-optimal extremes while safely protecting mid-tier policies from algorithmic distortion.

5.3. Mitigating Algorithmic Aversion via HITL Architecture

Delegating strategic supply chain decisions entirely to autonomous AI agents frequently triggers "algorithmic aversion" among healthcare executives. The Human-Bias Stress Test (Section 4.6) explicitly addressed this barrier. By proving that the core LLM-CRITIC micro-weights structurally anchor the decision matrix against even extreme human cognitive biases (e.g., $W_T = 0.70$), the framework demonstrates that managers can safely impose their strategic intent without risking algorithmic collapse. This hybrid cognitive division of labor enables practical industry adoption, ensuring human ethical accountability while leveraging AI computational precision.

5.4. Methodological Benchmarking: The Compensatory Failure of SAW vs. VIKOR Regret Minimization

To externally validate the structural integrity of the Z-VIKOR rankings, a comparative mathematical benchmarking was executed against two established MCDM algorithms: TOPSIS and Simple Additive Weighting (SAW). The Spearman's rank correlation coefficient between VIKOR and TOPSIS demonstrated strong alignment ($\rho = 0.883$). However, the comparison with the SAW method ($\rho = 0.750$) revealed a critical conceptual finding that definitively validates the selection of VIKOR for safety-critical environments. This result is consistent with recent applications of VIKOR in security-sensitive and intelligent technological domains, such as ANN-based 5G patient monitoring [53] and image encryption strategies for Industry 5.0 [54], where VIKOR's compromise-ranking framework is advantageous for balancing competing decision criteria.

In the SAW calculation, the "VMI Urgency" policy experienced a severe rank distortion, inexplicably jumping to Rank 1. The causal mechanism for this distortion lies in SAW's fully compensatory mathematical nature. Because VMI Urgency achieves extreme maximum values in service criteria (C2, C3, and C6) which carry high global weights, these scores completely offset its critically poor performance in the cost criteria (C1 and C4).

In pharmaceutical logistics, extreme failures in cost viability or systemic safety cannot simply be "compensated" by high performance elsewhere. The VIKOR algorithm uniquely prevents this through its Regret Measure (R_i), which fundamentally operates as a non-compensatory penalty. By minimizing the maximum group regret, VIKOR correctly penalized the VMI Urgency policy, proving mathematically that it is a far superior and strictly necessary algorithm for logistical networks where catastrophic single-point failures must be avoided. This perspective is particularly relevant in centralized medicine procurement contexts, where pharmaceutical logistics center location decisions must prioritize safety, reliability, and resilience alongside operational efficiency [55].

5.5. Limitations and Future Research Directions

Despite its methodological rigor, this study exhibits explicitly defined boundaries that necessitate future validation:

1. **Empirical Field Validation:** The operational decision matrix was derived entirely from stochastic simulation outputs. While methodologically sound for theoretical benchmarking, the framework lacks validation against real-world, dynamic empirical datasets, which may contain unstructured disruptions (e.g., Black Swan events) not captured by standard distribution models.
2. **LLM Weight Importation Rigor:** The micro-operational weights were imported directly from Musbah and Badi [19]. These LLM-derived Z-

weights were not independently cross-validated for internal consistency against the specific architectural nuances of the nine policies evaluated in this study, potentially masking contextual epistemic uncertainty.

3. **VIKOR Parameter Sensitivity:** The VIKOR ranking was executed using the default consensus strategy weight ($v = 0.5$). The study omits a sensitivity analysis of the v parameter (e.g., $v \geq 0.7$ for majority rule vs. $v \leq 0.3$ for veto dominance), which could alter the composition of the acceptable compromise set.
4. **Pharmaceutical Specificity:** The generic categorization of "pharmaceuticals" ignores sub-class constraints. For instance, the framework does not currently account for the energy consumption penalties inherent in Cold Chain logistics (e.g., vaccines), which would severely penalize spatial-buffering policies like SMT-SSP. Future research must prioritize integrating this framework within dynamic Digital Twin environments equipped with real-time empirical data streams, alongside expanding the operational matrix to include cold-chain thermodynamic constraints.

6. Implications

6.1. Theoretical Implications

Methodologically, this study contributes the "Strategic Modulation" mechanism, which provides a mathematical resolution to the "erasure effect" pervasive in traditional hierarchical MCDM. It establishes a new standard for integrating autonomous AI reasoning with human strategic oversight, proving that hierarchical control can exist without obliterating micro-level data variance. Furthermore, by anchoring our justification in cross-disciplinary HITL literature (e.g., Agentic AI and autonomous driving), we formalize the necessity of macro/micro cognitive divisions in operations research.

6.2. Practical and Managerial Implications

From an operational standpoint, the findings empirically confirm that predictive-opportunistic routing architectures (SMT-SSP) possess dominance stability, preemptively mitigating external threats to provide superior systemic resilience. From a managerial perspective, the proposed framework explicitly bridges the gap between algorithmic complexity and practical implementation. While the backend mathematical architecture is highly sophisticated (incorporating simulation engines, LLM heuristics, Z-number fuzzy logic, and extended VIKOR), this complexity is entirely encapsulated. Supply chain managers are not required to possess deep data science expertise; they simply interact with intuitive, macro-level SWOT sliders (the HITL Tier 1 input). This empowers managers to safely exert strategic governance and ethical accountability over AI-driven supply chains with minimal cognitive friction. This design aligns with prior research emphasizing that successful smart supply chain adoption depends on effective managerial integration and organizational oversight [56].

Finally, while testing untested AI algorithms directly on real networks poses unacceptable ethical and financial risks (induced stockouts)—justifying our reliance on high-fidelity stochastic simulation—we are currently in the process of compiling and cleaning a real-world dataset comprising 120 operational pharmacies. This forthcoming dataset will enable the empirical field validation of these simulated findings in our immediate future work.

7. Conclusion

The strategic selection of pharmaceutical replenishment policies requires a rigorous equilibrium between financial efficiency and systemic risk mitigation. This study addressed the methodological gap between algorithmic computation and human strategic oversight by introducing the Human-in-the-Loop (HITL) Z-Number SWOT-VIKOR framework. By rejecting strict hierarchical normalization in favor of a novel "Strategic Modulation" mechanism, the framework successfully preserved intrinsic data variance, preventing a mathematically critical 5% rank distortion ($\rho = 0.950, p < 0.01$) among alternative policies. Evaluating nine stochastic routing architectures, the extended VIKOR algorithm identified a robust compromise set led by the SMT-SSP policy ($Q_i = 0.000, R = 0.0556$). This result proved that spatial-opportunistic architectures uniquely possess the dominance stability required to absorb environmental shocks. Ultimately, this research provides supply chain practitioners with a bias-resistant, mathematically anchored decision-support tool that explicitly protects human ethical accountability while harnessing the scalable precision of Artificial Intelligence.

Competing Interest Statement

The authors declare no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Data Availability Statement

All data generated or analysed during this study are included in this article.

Statement on the Ethical Use of AI Tools

During the preparation of this article, the authors used Gemini and Grammarly solely for language editing, grammar correction, and improving overall readability. All content has been reviewed and revised by the authors. The authors take full responsibility for the accuracy and originality of the work.

References

[1] J. Shen, F. Bu, Z. Ye, M. Zhang, Q. Ma, J. Yan, and T. Huang, "Management of Drug Supply Chain Information Based on 'Artificial Intelligence + Vendor Managed Inventory' in China: Perspective Based on a Case Study," *Front. Pharmacol.*, vol. 15, 2024, doi: 10.3389/fphar.2024.1373642.

[2] F. Kochakkashani, V. Kayvanfar, and A. Haji, "Supply Chain Planning of Vaccine and Pharmaceutical Clusters Under Uncertainty: The Case of COVID-19," *Socioecon. Plann. Sci.*, vol. 87, p. 101602, Jun. 2023, doi: 10.1016/j.seps.2023.101602.

[3] J. Musbah, I. Badi, and M. B. Bouraima, "Optimizing Time-Based Heuristics for Resilient VMI Replenishment: A Simulation-Optimization Approach," *J. Ind. Intell.,* vol. 3, no. 2, pp. 105–124, Jun. 2025, doi: 10.56578/jii030204.

[4] Z. Li and P. Jiao, "Two-Stage Stochastic Programming for the Inventory Routing Problem with Stochastic Demands in Fuel Delivery," *Int. J. Ind. Eng. Comput.*, vol. 13, no. 4, pp. 507–522, 2022, doi: 10.5267/j.ijiec.2022.7.004.

[5] R. Lotfi, B. Kargar, M. Rajabzadeh, F. Hesabi, and E. Özceylan, "Hybrid Fuzzy and Data-Driven Robust Optimization for Resilience and Sustainable Health Care Supply Chain with Vendor-Managed Inventory Approach," *Int. J. Fuzzy Syst.*, vol. 24, no. 2, pp. 1216–1231, Mar. 2022, doi: 10.1007/s40815-021-01209-4.

[6] F. E. Achamrah, F. Riane, and S. Limbourg, "Spare Parts Inventory Routing Problem with Transshipment and Substitutions Under Stochastic Demands," *Appl. Math. Model.*, vol. 101, pp. 309–331, Jan. 2022, doi: 10.1016/j.apm.2021.08.029.

[7] O. Badejo and M. Ierapetritou, "Optimization of Pharmaceutical Supply Chains: Navigating Disruptions and Operational Uncertainty Utilizing Risk Measures," *AIChE J.*, vol. 71, p. 18770, Feb. 2025, doi: 10.1002/aic.18770.

[8] E. J. Alarcon Ortega, S. Malicki, K. F. Doerner, and S. Minner, "Stochastic Inventory Routing with Dynamic Demands and Intra-Day Depletion," *Comput. Oper. Res.*, vol. 163, p. 106503, Mar. 2024, doi: 10.1016/j.cor.2023.106503.

[9] A. Paul, N. Shukla, S. K. Paul, and A. Trianni, "Sustainable Supply Chain Management and Multi-Criteria Decision-Making Methods: A Systematic Review," *Sustainability*, vol. 13, no. 13, Jun. 2021, doi: 10.3390/su13137104.

[10] J. Fang, W. Zhou, and L. Xiong, "Multi-Criteria Decision Making Approach for Supplier Selection and Order Allocation in a Digital Supply Chain Resilience," *Ann. Oper. Res.*, Dec. 2024, doi: 10.1007/s10479-024-06435-1.

[11] S. A. Kayani, S. S. Warsi, and R. A. Liaqait, "A Smart Decision Support Framework for Sustainable and Resilient Supplier Selection and Order Allocation in the Pharmaceutical Industry," *Sustainability*, vol. 15, no. 7, Mar. 2023, doi: 10.3390/su15075962.

[12] T. Nguyen and C. Truong, "Integral Swot-Ahp-Tows Model for Strategic Agricultural Development in the Context of Drought: A Case Study in Ninh Thuan, Vietnam," *Int. J. Anal. Hierarchy Process*, vol. 14, no. 1, May 2022, doi: 10.13033/ijahp.v14i1.890.

[13] T. Hill and R. Westbrook, "SWOT Analysis: It's Time for a Product Recall," *Long Range Plann.*, vol. 30, no. 1, pp. 46–52, Feb. 1997, doi: 10.1016/S0024-6301(96)00095-7.

[14] S. Lee, D. Kim, S. Park, and W. Lee, "A Study on the Strategic Decision Making Used in the Revitalization of Fishing Village Tourism: Using A'WOT Analysis," *Sustainability*, vol. 13, no. 13, Jul. 2021, doi: 10.3390/su13137472.

[15] Ç. K. Kaymaz, S. Birinci, and Y. Kızılkan, "Sustainable Development Goals Assessment of Erzurum Province with SWOT-AHP Analysis," *Environ. Dev. Sustain.*, vol.

24, no. 3, pp. 2986–3012, Mar. 2022, doi: 10.1007/s10668-021-01584-w.

[16] G. Montibeller and D. von Winterfeldt, "Cognitive and Motivational Biases in Decision and Risk Analysis," *Risk Anal. Off. Publ. Soc. Risk Anal.*, vol. 35, no. 7, pp. 1230–1251, Jul. 2015, doi: 10.1111/risa.12360.

[17] J. E. Korteling, G. L. Paradies, and J. P. Sassen-van Meer, "Cognitive Bias and How to Improve Sustainable Decision Making," *Front. Psychol.*, vol. 14, Feb. 2023, doi: 10.3389/fpsyg.2023.1129835.

[18] J. Wei *et al.*, "Chain-of-Thought Prompting Elicits Reasoning in Large Language Models," in *Proceedings of the 36th International Conference on Neural Information Processing Systems*, in NIPS '22. Red Hook, NY, USA: Curran Associates Inc., Nov. 2022, pp. 24824–24837.

[19] J. Musbah and I. Badi, "LLM-Assisted Virtual Expert Weight Elicitation in Pharmaceutical Supply Chains: A Z-Number Multi-Agent Framework," *Intell. Syst. Res. Appl. J.*, vol. 2, pp. 27–39, Jan. 2026, doi: 10.59543/mmhqdg22.

[20] C. Longoni, A. Bonezzi, and C. K. Morewedge, "Resistance to Medical Artificial Intelligence," *J. Consum. Res.*, vol. 46, no. 4, pp. 629–650, Dec. 2019, doi: 10.1093/jcr/ucz013.

[21] A. Holzinger, "Interactive Machine Learning for Health Informatics: When Do We Need the Human-in-the-Loop?," *Brain Inform.*, vol. 3, no. 2, pp. 119–131, Jun. 2016, doi: 10.1007/s40708-016-0042-6.

[22] S. Opricovic and G.-H. Tzeng, "Compromise Solution by MCDM Methods: A Comparative Analysis of VIKOR and TOPSIS," *Eur. J. Oper. Res.*, vol. 156, no. 2, pp. 445–455, Jul. 2004, doi: 10.1016/S0377-2217(03)00020-1.

[23] J. Gallien, N. Z. Leung, and P. Yadav, "Inventory Policies for Pharmaceutical Distribution in Zambia: Improving Availability and Access Equity," *Prod. Oper. Manag.*, vol. 30, no. 12, pp. 4501–4521, Dec. 2021, doi: 10.1111/poms.13541.

[24] K. N. Alfina, R. M. C. Ratnayake, D. Wibisono, M. H. Basri, and N. B. Mulyono, "Enhancing Resilience and Sustainable Healthcare Supply Chains: Integrating Circular Economy and Dynamic Barrier Management," *Discov. Sustain.*, vol. 6, no. 1, p. 481, May 2025, doi: 10.1007/s43621-025-01270-1.

[25] A. Rizzardi, S. Sicari, J. F. Cevallos M., and A. Coen-Porisini, "IoT-Driven Blockchain to Manage the Healthcare Supply Chain and Protect Medical Records," *Future Gener. Comput. Syst.*, vol. 161, pp. 415–431, Dec. 2024, doi: 10.1016/j.future.2024.07.039.

[26] X. Shang, G. Zhang, B. Jia, and M. Almanaseer, "The Healthcare Supply Location-Inventory-Routing Problem: A Robust Approach," *Transp. Res. Part E Logist. Transp. Rev.*, vol. 158, p. 102588, Feb. 2022, doi: 10.1016/j.tre.2021.102588.

[27] D. Preil and M. Krapp, "Artificial Intelligence-Based Inventory Management: A Monte Carlo Tree Search Approach," *Ann. Oper. Res.*, vol. 308, no. 1, pp. 415–439, Jan. 2022, doi: 10.1007/s10479-021-03935-2.

[28] M. Kurttila, M. Pesonen, J. Kangas, and M. Kajanus, "Utilizing the Analytic Hierarchy Process (AHP) in SWOT Analysis — a Hybrid Method and Its Application to a Forest-Certification Case," *For. Policy Econ.*, vol. 1, no. 1, pp. 41–52, May 2000, doi: 10.1016/S1389-9341(99)00004-0.

[29] A. Afrasiabi, M. Tavana, and D. Di Caprio, "An Extended Hybrid Fuzzy Multi-Criteria Decision Model for Sustainable and Resilient Supplier Selection," *Environ. Sci. Pollut. Res.*, vol. 29, no. 25, pp. 37291–37314, May 2022, doi: 10.1007/s11356-021-17851-2.

[30] V. Belton and T. Gear, "On a Short-Coming of Saaty's Method of Analytic Hierarchies," *Omega*, vol. 11, no. 3, pp. 228–230, Jan. 1983, doi: 10.1016/0305-0483(83)90047-6.

[31] J. Barzilai, "Deriving Weights from Pairwise Comparison Matrices," *J. Oper. Res. Soc.*, vol. 48, no. 12, pp. 1226–1232, Dec. 1997, doi: 10.1057/palgrave.jors.2600474.

[32] M. Behzadian, S. Khanmohammadi Otaghsara, M. Yazdani, and J. Ignatius, "A State-of-the-Art Survey of TOPSIS Applications," *Expert Syst. Appl.*, vol. 39, no. 17, pp. 13051–13069, Dec. 2012, doi: 10.1016/j.eswa.2012.05.056.

[33] S. Hezer, E. Gelmez, and E. Özceylan, "Comparative Analysis of TOPSIS, VIKOR and COPRAS Methods for the COVID-19 Regional Safety Assessment," *J. Infect. Public Health*, vol. 14, no. 6, pp. 775–786, Jun. 2021, doi: 10.1016/j.jiph.2021.03.003.

[34] P. Wang, Y. Lin, M. Fu, and Z. Wang, "VIKOR Method for Plithogenic Probabilistic Linguistic MAGDM and Application to Sustainable Supply Chain Financial Risk Evaluation," *Int. J. Fuzzy Syst.*, vol. 25, no. 2, pp. 780–793, Mar. 2023, doi: 10.1007/s40815-022-01401-0.

[35] W. Xu, J. Qian, Y. Wu, S. Yan, Y. Ni, and G. Yang, "A VIKOR-Based Sequential Three-Way Classification Ranking Method," *Algorithms*, vol. 17, no. 11, Nov. 2024, doi: 10.3390/a17110530.

[36] M. S. Saidin, L. S. Lee, S. M. Marjugi, M. Z. Ahmad, and H.-V. Seow, "Fuzzy Method Based on the Removal Effects of Criteria (MEREC) for Determining Objective Weights in Multi-Criteria Decision-Making Problems," *Mathematics*, vol. 11, no. 6, Mar. 2023, doi: 10.3390/math11061544.

[37] M. Keshavarz-Ghorabaee, M. Amiri, E. K. Zavadskas, Z. Turskis, and J. Antucheviciene, "Determination of Objective Weights Using a New Method Based on the

Removal Effects of Criteria (MEREC),” *Symmetry*, vol. 13, no. 4, Mar. 2021, doi: 10.3390/sym13040525.

[38] S. Zakeri, D. Konstantas, P. Chatterjee, and E. K. Zavadskas, “Soft Cluster-Rectangle Method for Eliciting Criteria Weights in Multi-Criteria Decision-Making,” *Sci. Rep.*, vol. 15, no. 1, p. 284, Jan. 2025, doi: 10.1038/s41598-024-81027-4.

[39] J. Rezaei, A. Arab, and M. Mehregan, “Equalizing Bias in Eliciting Attribute Weights in Multiattribute Decision-Making: Experimental Research,” *J. Behav. Decis. Mak.*, vol. 35, no. 2, p. e2262, 2022, doi: 10.1002/bdm.2262.

[40] J. He, C. Treude, and D. Lo, “LLM-Based Multi-Agent Systems for Software Engineering: Literature Review, Vision, and the Road Ahead,” *ACM Trans. Softw. Eng. Methodol.*, vol. 34, no. 5, pp. 1–30, 2025, doi: 10.1145/3712003.

[41] Z. Xi *et al.*, “The Rise and Potential of Large Language Model Based Agents: A Survey,” *Sci. China Inf. Sci.*, vol. 68, no. 2, p. 121101, Jan. 2025, doi: 10.1007/s11432-024-4222-0.

[42] L. A. Zadeh, “A Note on *Z*-numbers,” *Inf. Sci.*, vol. 181, no. 14, pp. 2923–2932, Jul. 2011, doi: 10.1016/j.ins.2011.02.022.

[43] J. Johnson, “Automating the OODA Loop In the Age of Intelligent Machines: Reaffirming the Role of Humans in Command-And-Control Decision-Making in the Digital Age,” *Def. Stud.*, vol. 23, no. 1, pp. 43–67, Jan. 2023, doi: 10.1080/14702436.2022.2102486.

[44] A. Choudhury and Z. Chaudhry, “Large language models and user trust: Consequence of self-referential learning loop and the deskilling of health care professionals,” *J. Med. Internet Res.,* vol. 26, 2024, doi: 10.2196/56764.

[45] U. M. Borghoff, P. Bottoni, and R. Pareschi, “Human-Artificial Interaction in the Age of Agentic AI: A System-Theoretical Approach,” *Front. Hum. Dyn.*, vol. 7, May 2025, doi: 10.3389/fhumd.2025.1579166.

[46] H. Yang, Y. Zhou, J. Wu, H. Liu, L. Yang, and C. Lv, “Human-Guided Continual Learning for Personalized Decision-Making of Autonomous Driving,” *IEEE Trans. Intell. Transp. Syst.*, vol. 26, no. 4, pp. 5435–5447, Apr. 2025, doi: 10.1109/TITS.2024.3524609.

[47] D. Dellermann, P. Ebel, M. Söllner, and J. M. Leimeister, “Hybrid Intelligence,” *Bus. Inf. Syst. Eng.*, vol. 61, no. 5, pp. 637–643, Oct. 2019, doi: 10.1007/s12599-019-00595-2.

[48] X. Hao, E. Demir, and D. Eyers, “Exploring Collaborative Decision-Making: A Quasi-Experimental Study of Human and Generative AI Interaction,” *Technol. Soc.*, vol. 78, p. 102662, Sep. 2024, doi: 10.1016/j.techsoc.2024.102662.

[49] E. Mosqueira-Rey, E. Hernández-Pereira, D. Alonso-Ríos, J. Bobes-Bascarán, and Á. Fernández-Leal, “Human-in-the-Loop Machine Learning: A State of the Art,” *Artif. Intell. Rev.*, vol. 56, no. 4, pp. 3005–3054, Apr. 2023, doi: 10.1007/s10462-022-10246-w.

[50] V. Choudhary, A. Marchetti, Y. R. Shrestha, and P. Puranam, “Human-AI Ensembles: When Can They Work?,” *J. Manag.*, vol. 51, no. 2, pp. 536–569, Feb. 2025, doi: 10.1177/01492063231194968.

[51] U. Agudo, K. G. Liberal, M. Arrese, and H. Matute, “The Impact of AI Errors in a Human-in-the-Loop Process,” *Cogn. Res. Princ. Implic.*, vol. 9, no. 1, p. 1, Jan. 2024, doi: 10.1186/s41235-023-00529-3.

[52] S. Opricovic and G.-H. Tzeng, “Extended VIKOR method in comparison with outranking methods,” *Eur. J. Oper. Res.*, vol. 178, no. 2, pp. 514–529, Apr. 2007, doi: 10.1016/j.ejor.2006.01.020.

[53] S. Chinnaperumal, M. Periyasamy, A. A. Alhussan, et al., “Secure and Intelligent 5G-Enabled Remote Patient Monitoring Using ANN And Choquet Integral Fuzzy VIKOR,” *Sci. Rep.,* vol. 15, no. 1, p. 9913, 2025, doi: 10.1038/s41598-025-93829-1.

[54] I. S. Arafat, R. Premkumar, M. Vidhyalakshmi, C. Priya, and M. Elangovan, “Optimizing Sustainable Image Encryption Strategies in Industry 5.0 Using VIKOR MCDM Methodology,” in *Decision-Making Techniques and Methods for Sustainable Technological Innovation*, K. Kalita, J. V. N. Ramesh, M. Elangovan, and S. Balamurugan, Eds. Hoboken, NJ, USA: Wiley, 2025, pp. 85–100, doi: 10.1002/9781394242603.ch5.

[55] H. Li, H. Fu, and T. Ramayah, “Pharmaceutical Logistics Center Location in the Context of Centralized Medicine Procurement: A Literature Review,” *J. Logist. Inform. Serv. Sci.,* vol. 10, no. 1, pp. 94–121, 2023, doi: 10.33168/JLISS.2023.0106.

[56] K. L. Lee, S. Y. Wong, H. M. Alzoubi, B. Al Kurdi, M. T. Alshurideh, and M. El Khatib, “Adopting Smart Supply Chain and Smart Technologies to Improve Operational Performance in Manufacturing Industry,” *Int. J. Eng. Bus. Manag.,* vol. 15, 2023, doi: 10.1177/18479790231200614.