OPERATIONS AND SUPPLY CHAIN MANAGEMENT
Vol. 19, No. 4, 2026, pp. 825 – 841
ISSN 1979-3561 | EISSN 2579-9363

Prioritizing Risk Mitigation Strategies in Hazardous-Materials Logistics: An Integrated BWM-PROMETHEE II Approach with Evidence from Ammunition Transportation Operations

Nunik Endah Sulistiyawati
Interdisciplinary School of Management and Technology,
Sepuluh Nopember Institute of Technology (ITS), Surabaya, Indonesia
Email: 7032231002@student.its.ac.id

Bambang Syairudin
Department of Industrial and Systems Engineering,
Sepuluh Nopember Institute of Technology (ITS), Surabaya, Indonesia
Email: bambangsyairudin@gmail.com

Buana Maruf
Research Center for Hydrodynamics Technology,
National Research and Innovation Agency (BRIN), Surabaya, Indonesia
Email: buana.maruf@brin.go.id

ABSTRACT

Hazardous-materials logistics is highly vulnerable to disruptions at storage-transport interfaces, where handling failures can generate cascading safety, security, environmental, and operational consequences. This study applies an integrated Best-Worst Method (BWM) and PROMETHEE II procedure to prioritize five risk-mitigation alternatives for ammunition loading and unloading operations at a maritime depot operated by the Indonesian Navy. Nine purposively selected domain experts, organized into three functional deliberation groups, provided criterion-comparison and alternative-performance judgments across Safety, Security, Regulatory Compliance, Operational Efficiency, Environmental Impact, and Technical Readiness. Group-level BWM weights were aggregated using the arithmetic mean, while geometric-mean and median aggregation were examined as alternative specifications. Safety, Security, and Regulatory Compliance collectively account for 77.5% of the baseline decision weight. Under the corrected baseline PROMETHEE II specification, Integrated Security System Enhancement ranks first (net flow +0.210), followed by Automated/High-Performance Equipment (+0.073), Integrated Risk Management (+0.033), Intensive Training and Certification (-0.120), and Specialized Cranes and Dock Infrastructure (-0.196). The ordinal ranking remains unchanged across the tested aggregation, weighting, criterion-structure, preference-function, threshold, and one-at-a-time perturbation scenarios, although net-flow magnitudes and rank margins vary. The study therefore demonstrates scenario stability rather than conclusive robustness. Its contribution is a context-specific, resilience-oriented decision procedure. For managerial use, the risk-priority ranking must be followed by a separate cost and feasibility screening before implementation or procurement decisions are made.

Keywords*: hazardous-materials logistics, risk-mitigation prioritization, Best-Worst Method, PROMETHEE II, supply chain resilience, scenario stability*

1. INTRODUCTION

Hazardous-materials logistics constitutes one of the most consequential domains within modern supply chain management. Failures in handling, storage, or transport can rapidly escalate into catastrophic outcomes affecting personnel, surrounding communities, infrastructure, and the environment (Gholamizadeh *et al.,* 2025; Hsu *et al.,* 2024). These supply chains differ from conventional logistics systems because they operate under dense regulatory regimes, severe consequence asymmetry, and strong interdependence among warehousing, handling, transport, and emergency-response activities (Fan and Stevenson, 2018). Localized failures can therefore propagate across connected operations and generate ripple effects that extend beyond the initiating event (Dolgui and Ivanov, 2021; Ivanov and Dolgui, 2020; Morteza Beigi *et al.*, 2024).

The storage-transport interface is a particularly critical but insufficiently examined decision point. At this interface, dock-side handling, access control, equipment readiness, regulatory verification, and vessel operations converge within a spatially constrained zone. Empirical studies of hazardous cargo handling repeatedly identify procedural

violations, human-factor failures, coordination breakdowns, fire, explosion, and inadequate handling as major causes of high-consequence incidents during loading and unloading (Khan *et al.,* 2022, 2024; Peron *et al.,* 2023). Nevertheless, hazardous-materials research has concentrated predominantly on route optimization, network design, and transit-phase risk assessment (Guo and Luo, 2022; Mohri *et al.,* 2022; Tasnim *et al.,* 2023). Decision support for prioritizing competing mitigation actions at the interface level remains comparatively underdeveloped.

Prioritization in this setting is inherently multi-criteria. Decision makers must compare mitigation alternatives across Safety, Security, Regulatory Compliance, Operational Efficiency, Environmental Impact, and Technical Readiness. These criteria do not have equal consequence, and their relative importance cannot be inferred solely from financial or technical metrics. Multi-criteria decision-making methods provide a structured means of eliciting expert preferences and translating them into an auditable ranking (Fagundes *et al.*, 2020; Moktadir *et al.*, 2025). Previous research has also applied MCDM to prioritize risk-management strategies within supply chain operations (Enyinda, 2018). However, the value of an MCDM result depends not only on the selected method but also on aggregation rules, criterion overlap, preference functions, thresholds, and the composition of the expert panel.

This study uses supply chain resilience as the conceptual basis of the decision architecture. Resilience at the storage-transport interface concerns the capacity to absorb or contain a disruption, adapt operations under degraded conditions, and restore performance while incorporating lessons into subsequent procedures (Aldrighetti *et al.,* 2023; Gaudenzi *et al.,* 2023; Ivanov, 2024). The six evaluation criteria operationalize these capacities. Safety and Security primarily support absorptive capacity by preventing or containing accidental and intentional disruptions. Regulatory Compliance supports absorptive and restorative capacity through controlled response, auditability, corrective action, and institutional learning. Operational Efficiency supports adaptive capacity by maintaining critical throughput, while Environmental Impact supports consequence limitation and recovery. Technical Readiness supports both adaptation and restoration by enabling continuity, repair, and system recovery. BWM and PROMETHEE II are therefore analytical instruments within a resilience-oriented decision architecture rather than the theoretical contribution of the study.

Existing studies leave three unresolved issues, developed in Section 2.7: interface-level mitigation prioritization remains limited relative to route-level risk assessment; resilience capacities are rarely operationalized through tactical decision-support procedures; and many MCDM studies report a final ranking without examining whether the result is driven by aggregation rules, concentrated weights, criterion overlap, preference functions, or thresholds. The present study addresses these issues in a single-site ammunition-handling case while limiting the transferability of the numerical results.

Accordingly, the study pursues four objectives: (1) to derive the relative importance of six resilience-related risk-mitigation criteria at the storage-transport interface; (2) to rank five mitigation alternatives using an integrated BWM-PROMETHEE II procedure; (3) to determine whether the ranking is sensitive to alternative aggregation rules, weight structures, criterion structures, preference functions, thresholds, and expert-group composition; and (4) to translate the risk-priority order into an implementation pathway that explicitly considers cost, feasibility, organizational readiness, and resource constraints.

The study makes three bounded contributions. First, it offers a contextual methodological application of BWM and PROMETHEE II in a safety-critical maritime ammunition-handling environment. Second, it provides case-specific empirical insight into how experts balance safety, security, compliance, operational, environmental, and technical considerations. Third, it proposes a transparent two-stage decision process that separates risk-reduction priority from implementation feasibility. The study does not introduce a new MCDM algorithm, and its numerical weights and rankings are not claimed to be universally transferable.

2. LITERATURE REVIEW

This section critically synthesizes five related research streams: hazardous-materials risk management, ammunition and explosive logistics, storage-transport interface risk, MCDM in supply chain risk, and supply chain resilience. The review emphasizes unresolved decision problems and modeling limitations rather than presenting a study-by-study catalogue.

2.1 Risk Management in Hazardous-Materials Logistics

Risk management in hazardous-materials logistics has evolved from route-level exposure minimization toward multi-dimensional governance that incorporates consequence assessment, regulatory control, infrastructure condition, and coordination across logistics networks (Guo and Luo, 2022; Ke, 2022; Mohri *et al.,* 2022; Hsu *et al.,* 2024). More broadly, supply chain risk management has developed into an established research area, with resilience emerging as an important consideration and several research gaps remaining across the literature (Sahab and Oulfarsi, 2024). The conventional link-based risk expression is shown in Equation 1.

$$R_l = p_l \times c_l \times e_l \quad (1)$$

Here, p_l is the probability of an adverse event on link l, c_l is the magnitude of the consequence, and e_l is the quantity or level of exposure. This formulation is appropriate for transit and network-design problems, but it provides limited guidance when organizations must choose among mitigation alternatives that affect several non-equivalent operational dimensions.

2.2 Safety Risks in Ammunition Transportation and Storage Operations

Ammunition logistics is an extreme subset of hazardous-materials management because accidental initiation, deliberate interference, regulatory violation, and infrastructure degradation may interact within the same operational system (Al-Hajj *et al.,* 2021; Sari and Öztürk, 2025; Valsamos *et al.,* 2021). Compared with most civilian dangerous goods, ammunition handling requires stronger physical-security controls and more restrictive authorization

Table 1 Context-Specific Comparison of Alternative MCDM Approaches

Method	Relevant Strength	Limitation in the Focal Decision Context
AHP	Familiar hierarchical weighting and extensive use in practice	Requires more pairwise comparisons and can increase consistency burden as the criterion set expands.
TOPSIS	Produces a simple complete ranking based on distance from ideal solutions	Sensitive to normalization and uses compensatory distance logic that can obscure directional dominance.
VIKOR	Identifies a compromise solution under conflicting criteria	Less transparent for tracing pairwise directional preference among mitigation alternatives.
ELECTRE	Provides strong outranking logic and limits unrestricted compensation	May yield incomparability or an incomplete ranking, which is less suitable for a single initial-priority decision.
BWM-PROMETHEE II	Combines parsimonious, consistency-controlled weighting with a complete preference-flow ranking	Requires explicit choices regarding aggregation, preference functions, thresholds, and criterion structure.

Table 2 Mapping of Evaluation Criteria to Supply Chain Resilience Capacities

Criterion	Primary Resilience Capacity	Operational Rationale
C1 Safety	Absorptive	Prevents and contains accidental harm, fire, explosion, equipment failure, and unintended exposure.
C2 Security	Absorptive	Prevents intentional disruption, unauthorized access, theft, diversion, and sabotage.
C3 Regulatory Compliance	Absorptive / Restorative	Supports controlled response, audit traceability, corrective action, and institutional learning.
C4 Operational Efficiency	Adaptive	Maintains critical throughput, coordination, and resource use under constrained conditions.
C5 Environmental Impact	Absorptive / Restorative	Limits environmental consequences and supports remediation and recovery.
C6 Technical Readiness	Adaptive / Restorative	Enables continuity, equipment recovery, maintenance, and restoration of operational capability.

procedures. Incident causation cannot therefore be represented adequately by a single technical variable. Human error, aging equipment, environmental conditions, access control, communication, and compliance must be assessed jointly, and escalation pathways in comparable high-energy facilities are known to propagate through coupled rather than isolated failures (Men *et al.,* 2022).

Table 1 compares the principal MCDM approaches considered for the focal decision problem and clarifies the methodological rationale for combining BWM with PROMETHEE II. The comparison emphasizes the trade-off between elicitation burden, transparency of directional preference, completeness of ranking, and sensitivity to model specification. Table 2 then links the six evaluation criteria to absorptive, adaptive, and restorative resilience capacities, showing how the operational criteria translate the resilience concept into measurable decision dimensions at the storage-transport interface.

2.3 Risk at the Storage-Transport Interface

Loading and unloading form the functional junction between storage and transport systems. Personnel, vessels, cranes, vehicles, surveillance systems, documentation processes, and emergency-response resources operate concurrently at this junction. Studies of hazardous cargo ports show that procedural deviation and coordination failure are especially consequential during transshipment (Khan *et al.,* 2022, 2024). Despite this risk concentration, interface-level mitigation choices are often treated as operational details outside the analytical boundary of route and network models. Evidence from manufacturing supply chains also indicates that extensive communication can reduce operational risks associated with human failures and inadequate or failed internal processes (Wieteska, 2020).

2.4 Limitations of Conventional Risk Assessment

Checklists, audits, and probabilistic risk models remain essential (Gholamizadeh *et al.,* 2025; Tasnim *et al.,* 2023), but they do not by themselves establish which mitigation alternative should be implemented first when several objectives conflict. A route-level objective may minimize expected risk while ignoring security, compliance, readiness, or organizational feasibility (Liu *et al.,* 2025). The relevant decision problem is therefore not a replacement for probabilistic assessment. It is a complementary prioritization problem requiring explicit criteria, transparent preference elicitation, and sensitivity to modeling assumptions.

2.5 MCDM in Supply Chain Risk Management

Hybrid MCDM architectures combine a criterion-weighting procedure with an alternative-ranking procedure (Altubaishe and Desai, 2023; Moktadir *et al.,* 2025). BWM reduces comparison burden by requiring best-to-others and others-to-worst judgments, while retaining an explicit consistency assessment (Ecer, 2024; Liang *et al.,* 2020). PROMETHEE II produces a complete outranking order through positive, negative, and net preference flows (Ziemba and Gago, 2022). This combination is relevant to the focal setting because the expert population is specialized and limited, the alternatives exhibit asymmetric performance profiles, and the decision must remain auditable. Nevertheless, PROMETHEE II remains dependent on preference functions, thresholds, criterion structure, and weights. These assumptions must be examined rather than treated as neutral technical settings.

Table 3 Critical Synthesis of Prior Research Streams and the Bounded Position of the Present Study

Research Stream	Representative Studies	Established Contribution	Unresolved Issue	Position of this Study
Hazardous-materials routing and network risk	Guo and Luo (2022); Mohri *et al.* (2022); Tasnim *et al.* (2023)	Strong exposure and route-risk modeling	Limited guidance for prioritizing interface-level operational mitigations	Formulates a multi-criteria interface decision problem
Port and handling risk	Khan *et al.* (2022, 2024); Peron *et al.* (2023)	Identifies human, procedural, and coordination hazards	Mitigation alternatives are not compared through an auditable priority model	Evaluates five intervention alternatives at the handling interface
Hybrid MCDM in supply chain risk	Fagundes *et al.* (2020); Moktadir *et al.* (2025)	Structures expert-based weighting and ranking	Structural assumptions and criterion overlap are often insufficiently tested	Adds aggregation, weight, criterion, function, and threshold specifications
BWM-PROMETHEE applications	Kheybari *et al.* (2021); Taghipour *et al.* (2023)	Efficient weighting combined with complete outranking	Limited application to hazardous-materials interfaces	Provides a bounded contextual application, not a new method
Supply chain resilience	Aldrighetti *et al.* (2023); Gaudenzi *et al.* (2023); Ivanov (2024)	Develops absorptive, adaptive, and restorative capacities	Limited tactical operationalization at safety-critical interfaces	Maps resilience requirements to criteria and implementation decisions

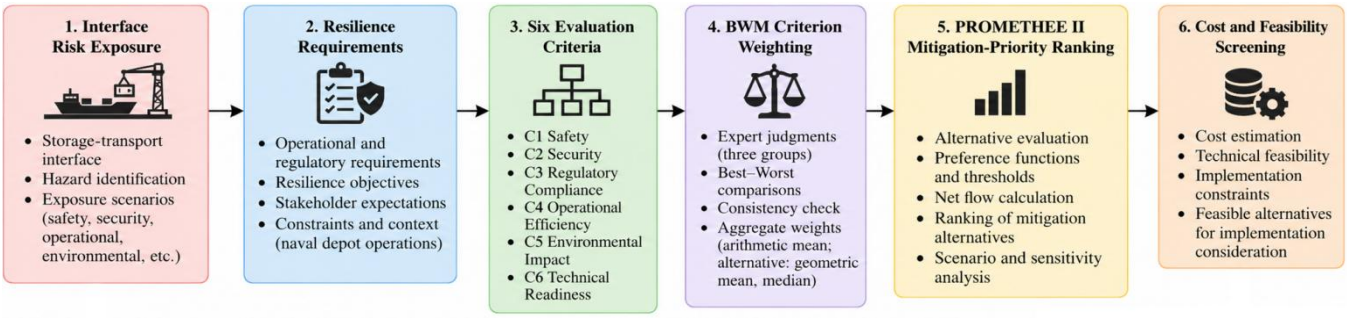


Figure 1 Resilience-Oriented Decision Architecture Linking Interface Risk Exposure, Resilience Requirements, Evaluation Criteria, BWM Criterion Weighting, PROMETHEE II Mitigation-Priority Ranking, and Cost-and-Feasibility Screening. Alternatives that pass the screening stage proceed to an implementation decision, and the model is periodically updated as operational evidence, risk conditions, costs, or technology readiness change.

The selected architecture is therefore not claimed to be universally superior. It is appropriate to the focal case because BWM limits elicitation burden for a small specialist panel, whereas PROMETHEE II provides a complete and auditable ranking while preserving directional preference information. The specification analysis is required precisely because these advantages do not remove dependence on modeling choices.

2.6 Supply Chain Resilience and Interface-Level Decisions

Supply chain resilience concerns the ability to absorb disruption, adapt while the disruption unfolds, and restore or redesign operations afterward (Hosseini *et al.*, 2020; Ivanov, 2023, 2024). Managing interconnected supply chain risks also requires governance mechanisms and coordination across organizational boundaries (Ahlqvist *et al.*, 2020). Strategic resilience research is extensive, but tactical operationalization remains more limited (Batista *et al.,* 2023). At a storage-transport interface, resilience is expressed through concrete capabilities: preventing initiation, containing impact, preserving critical throughput, maintaining authorized control, mobilizing emergency response, and updating procedures after an incident. In maritime port operations, human factors and organizational considerations influence the implementation of AI-driven digital collaboration platforms intended to strengthen port resilience (Lasse *et al.*, 2025). These capabilities provide the conceptual basis for the criteria used in this study.

2.7 Critical Synthesis and Research Positioning

Table 3 demonstrates that the research gap is not merely the absence of a particular hybrid method in a military context. For each of the five research streams it summarizes, the table identifies representative studies, the contribution they have established, the issue that remains unresolved, and the position adopted in response by the present study. The more substantive gaps concern interface-level prioritization, operationalization of resilience, criterion dependence, expert-judgment design, and robustness across modeling choices. Read across its final two columns, the table also shows that each position taken here answers a specific unresolved issue rather than claiming a new method.

3. METHODOLOGY

3.1 Research Design

The study adopts a structured quantitative single-case design. Figure 1 positions BWM and PROMETHEE II within a broader decision process composed of six sequential stages. The first stage defines interface risk exposure by delimiting the storage-transport interface, identifying hazards, and specifying safety, security, operational, and environmental exposure scenarios. The second stage

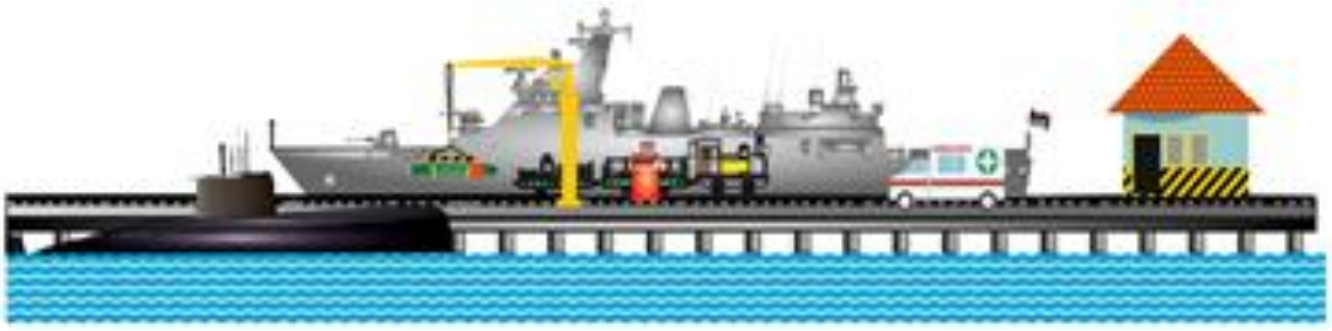


Figure 2 System Boundary and Principal Exposure Points at the Storage-Transport Interface of the Focal Depot

translates this exposure into resilience requirements that reflect operational and regulatory obligations, resilience objectives, stakeholder expectations, and the constraints of naval depot operations. These requirements are then operationalized in the third stage as the six evaluation criteria mapped in Table 2. In the fourth stage, BWM derives the criterion weights from the judgments of the three expert groups, including a consistency check and the aggregation of the group vectors. The fifth stage applies PROMETHEE II to evaluate the mitigation alternatives, specify preference functions and thresholds, calculate net flows, and subject the resulting order to scenario and sensitivity analysis. The sixth stage screens the ranked alternatives for cost, technical feasibility, and implementation constraints. This structure prevents the risk-priority result from being interpreted as an automatic procurement decision.

The analytical scope of the study ends at the sixth stage, which is why Figure 1 terminates at cost-and-feasibility screening. Implementation is nevertheless not the endpoint of the decision cycle. Alternatives that pass the screen proceed to an implementation decision, followed by implementation monitoring and performance review, after which the model should be periodically updated as operational evidence, risk conditions, costs, or technology readiness change, with new data and feedback returned to the definition of interface risk exposure. This cyclical extension lies outside the analytical procedure evaluated in the study, and its managerial form is elaborated in Section 5.5.

3.1.1 Case Site and Decision Boundary

The focal site is a maritime ammunition storage depot operated by the Indonesian Navy. It was selected as an extreme case because regulatory intensity, physical hazard concentration, operational accountability, and security requirements occur simultaneously. The analytical boundary covers vessel loading and unloading, dock-side cargo handling, storage access control, security monitoring, and emergency-response preparedness. The numerical findings therefore represent the institutional and operational conditions of this site.

Figure 2 provides a schematic side view of this system boundary and highlights the principal exposure points concentrated at the storage-transport interface. The illustration reads from the seaward side on the left to the landward side on the right. At the seaward end, a naval surface vessel is moored alongside the jetty while a submarine occupies the adjacent berth, indicating that ammunition may be transferred while other assets remain within the same exposure zone; a mooring failure, vessel movement, or fire on one platform can therefore escalate to neighboring platforms and to the surrounding water. Next to the vessel, a dockside jib crane hoists ammunition loads between the deck and the jetty. This lifting zone concentrates the most immediate accidental hazards, such as dropped or swinging loads, rigging or equipment failure, and operator error, and it is the zone most directly addressed by the Safety and Technical Readiness criteria and by the infrastructure and automation alternatives. Behind the crane, the jetty deck serves as a temporary staging and transport lane in which a tractor-drawn train of ammunition carts, together with auxiliary handling equipment, moves cargo between the vessel side and the depot. Exposure in this lane arises from the temporary accumulation of explosive quantities in a confined space, from vehicle movements close to personnel, and from fire or collision during transit. An ambulance stationed on the jetty represents the medical and emergency-response capacity that must remain available throughout the operation. At the landward end, a guard post with hazard-striped markings represents the access-control point that separates the jetty from the storage area. Personnel, vehicles, and cargo are authorized or refused entry at this point, which makes it the principal exposure point for the unauthorized access, theft, diversion, and sabotage covered by the Security criterion.

Taken together, these zones are shown as contiguous and interacting elements rather than independent locations. The figure is intended to clarify where personnel, cargo, equipment, and security-control activities converge during loading and unloading, and it shows that this convergence occurs within a single narrow strip of jetty. This spatial concentration explains why a localized handling failure, equipment malfunction, unauthorized access event, or procedural deviation can propagate across safety, security, compliance, environmental, and operational dimensions at the same time, and it provides the physical rationale for evaluating every alternative against all six criteria jointly. Training and governance measures are not tied to any single zone; they act across the entire interface by shaping how procedures are executed and how deviations are detected, reported, and corrected. Accordingly, Figure 2 establishes the physical scope within which the criteria and mitigation alternatives are evaluated. It is a conceptual boundary illustration rather than a scale-accurate engineering layout of the depot, and it does not reproduce site-specific dimensions, separation distances, or security-sensitive features.

Figure 3 summarizes the operational configuration. Personnel and equipment readiness, vessel arrival, mooring, area control, cargo transfer, load verification, surveillance, medical support, communication, and documentation form

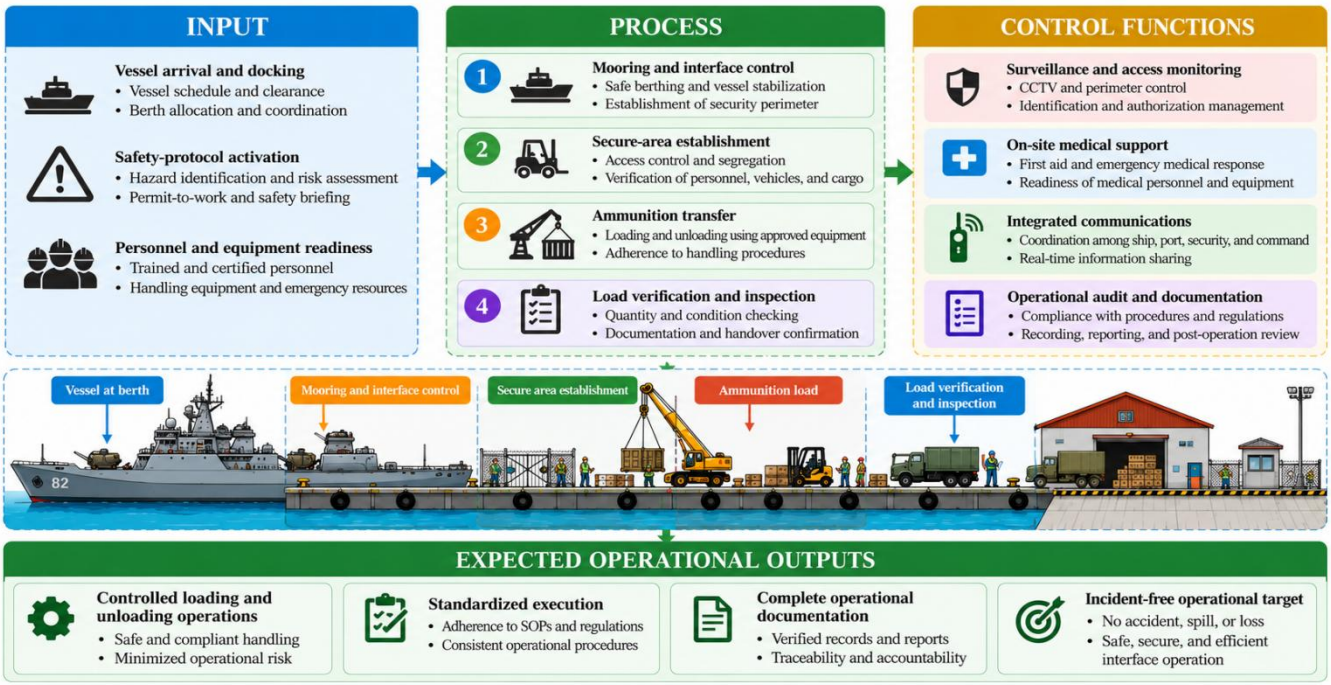


Figure 3 Operational Configuration of the Storage-Transport Interface and Associated Control Functions

Table 4 Expert-Panel Composition

Expert	Years	Role	Primary Expertise
E1	18	Senior Ammunition Depot Officer	Ammunition storage safety and SOP enforcement
E2	15	Logistics Operations Supervisor	Ammunition transportation and handling
E3	22	Military Safety and Compliance Officer	Regulatory compliance and risk management
E4	14	Technical Maintenance Manager	Handling equipment and infrastructure
E5	17	Training and Certification Coordinator	Personnel safety training
E6	12	Port Operations Officer	Maritime loading and unloading
E7	20	Security Systems Specialist	Physical security and monitoring systems
E8	16	Environmental and Hazard Control Officer	Environmental impact and emergency response
E9	13	Operational Planning Analyst	Risk assessment and operational planning

an interconnected control system. The figure arranges this system into four linked layers. The input layer specifies the conditions that must be in place before handling begins: vessel arrival and docking with schedule clearance and berth allocation, the activation of safety protocols through hazard identification, risk assessment, permit-to-work procedures, and safety briefings, and the readiness of certified personnel, handling equipment, and emergency resources. The process layer then sets out four sequential steps, namely mooring and interface control, secure-area establishment, ammunition transfer with approved equipment, and load verification and inspection concluded by a documented handover. Running in parallel with these steps, the control-function layer provides surveillance and access monitoring, on-site medical support, integrated communications among the ship, port, security, and command elements, and operational audit and documentation. The illustrated strip maps the process steps onto the physical sequence from the vessel at berth to the storage facility, while the output layer defines the intended results as controlled loading and unloading, standardized execution, complete operational documentation, and an incident-free operational target with no accident, spill, or loss. The mitigation alternatives described later in this section act on different parts of this configuration: integrated security enhancement reinforces surveillance and access monitoring, specialized cranes and automated equipment reinforce the transfer and verification steps, training and certification reinforce personnel readiness, and integrated risk management reinforces audit, documentation, and coordination. The figure is used to define the decision boundary rather than to imply that all controls operate at an identical maturity level.

3.2 Expert Panel, Elicitation, and Bias Control

Nine experts were selected purposively because of their direct operational, safety, regulatory, security, technical, and environmental responsibilities. Each expert had at least ten years of relevant experience. The panel is appropriate for an in-depth single-case decision analysis but is not statistically representative of military or civilian hazardous-materials facilities more broadly. Table 4 summarizes the panel composition, years of experience, organizational roles, and primary expertise, demonstrating the cross-functional coverage used for the elicitation process.

The experts were organized into three cross-functional groups. Each expert first completed provisional criterion comparisons and alternative-performance ratings independently after receiving standardized definitions, scoring anchors, and alternative descriptions. Differences

Table 5 Evaluation Criteria and Operational Definitions

Code	Criterion	Operational Definition	Type
C1	Safety	Capacity to prevent accidental physical harm, injury, fatality, fire, explosion, equipment failure, and unintended exposure during loading, unloading, and handling.	Benefit
C2	Security	Capacity to prevent intentional threats, unauthorized access, theft, diversion, sabotage, and deliberate interference.	Benefit
C3	Regulatory Compliance	Alignment with national military safety requirements, Indonesian Navy ordnance directives, applicable IMDG provisions, inspection requirements, and audit traceability.	Benefit
C4	Operational Efficiency	Capacity to preserve throughput, scheduling, coordination, and resource use without compromising control requirements.	Benefit
C5	Environmental Impact	Capacity to reduce contamination, fire-related environmental damage, and other environmental consequences of mishandling.	Benefit
C6	Technical Readiness	Availability, reliability, maintainability, and operational readiness of equipment, monitoring technology, and supporting systems.	Benefit

Table 6 Decision-Matrix Scoring Anchors

Score	Interpretation	General Anchor
1-3	Low	The alternative has minimal or indirect influence on the criterion and does not provide a system-level control.
4-6	Moderate	The alternative provides measurable but incomplete coverage of the criterion.
7-9	High	The alternative provides substantial or system-level coverage supported by operational controls.

exceeding two scale points were then discussed only to clarify meanings and evidence. Each group subsequently produced one post-deliberation consensus BWM comparison vector; consequently, the effective BWM decision units were the three groups rather than nine independent individual BWM models. The three resulting normalized group-level weight vectors were aggregated using the arithmetic mean for the baseline analysis. By contrast, the decision matrix represents the arithmetic mean of the nine final individual alternative-performance ratings. Group deliberation may reduce semantic ambiguity but can also introduce conformity or authority effects; it is therefore treated as a structured judgment process rather than a substitute for independent measurement.

Measurement validity was assessed at the level supported by the retained data. The rank ordering of the six criteria across the three group-level BWM profiles showed strong concordance (Kendall's W = 0.806; chi-square (5) = 12.10; p = 0.034, computed with mid-ranks because two groups produced tied weights). This statistic indicates agreement among the three group-level priority structures, not inter-rater reliability across nine experts. Because the retained research file contains consolidated group-level BWM outputs and the aggregated decision matrix rather than a complete anonymized individual-rating panel, a retrospective individual-level ICC or Kendall's W is not reported. This data-retention limitation is stated explicitly and prevents overinterpretation of the measurement evidence.

3.3 Evaluation Criteria and Measurement Anchors

Safety and Security are retained as separate criteria because they represent distinct threat mechanisms: Safety concerns accidental failure, whereas Security concerns intentional action. No inferential correlation between the two is reported, because five aggregated alternative means cannot support a credible estimate of criterion dependence. Potential overlap is instead examined through conceptual distinction together with the Safety-removed and Security-removed specifications described in Section 3.7. A mechanically combined criterion is not used as primary evidence, because additive aggregation preserves the baseline contribution algebraically and therefore cannot detect double counting. The operational definitions and benefit orientation of all six criteria are reported in Table 5, while Table 6 provides the 1-9 scoring anchors used to standardize interpretation of low, moderate, and high alternative performance across experts.

3.4 Mitigation Alternatives

The alternatives represent competing initial priorities rather than mutually exclusive long-term measures. A lower-ranked alternative may remain necessary as a complementary control, regulatory requirement, or rapidly deployable intervention. Cost and feasibility are not embedded in the baseline risk-effectiveness model; they are assessed in the second-stage implementation screen. Table 7 specifies the five mitigation alternatives, their operational scope, and their intervention domains, thereby defining the option set evaluated in the PROMETHEE II stage.

3.5 Best-Worst Method and Group Aggregation

For each effective expert-group decision unit, BWM identifies a best criterion C_B and a worst criterion C_W, constructs Best-to-Others and Others-to-Worst comparisons, and estimates a normalized weight vector. Each group-level vector was derived from one post-deliberation consensus comparison record. The linear formulation of the weight-derivation problem is shown in Equations 2 and 3.

$$\min \xi^L \tag{2}$$

$$\begin{aligned} &|w_B - a_{(Bj)} w_j| \le \xi^L, j \\ &= 1, \dots, n; |w_j - a_{(jW)} w_W| \le \xi^L, j \\ &= 1, \dots, n; \sum (j=1)^n w_j \\ &= 1; w_j \ge 0, j = 1, \dots, n \end{aligned} \tag{3}$$

Table 7 Mitigation Alternatives Evaluated in the Study

Code	Alternative	Operational Scope	Intervention Domain
P1	Integrated Security System Enhancement	Integrated access control, CCTV, biometric authentication, real-time monitoring, alerting, and command-and-control coordination.	Security architecture and system integration
P2	Provision of Specialized Cranes and Dock Infrastructure	Reinforced cranes, stabilized dock platforms, and mechanical handling aids intended to reduce manual exposure.	Handling infrastructure
P3	Intensive Training and Certification Programs	Competency development, certification enforcement, periodic drills, and human-factor awareness.	Personnel and procedures
P4	Use of Automated/High-Performance Equipment	Automated lifting aids, load sensors, equipment performance monitoring, and handling automation.	Automation and technical modernization
P5	Integrated Risk Management Implementation	Formal risk registers, monitoring dashboards, structured reporting, continuous improvement, and compliance tracking.	Governance and organizational control

Table 8 Group-Level BWM Nominations, Weights, Vector Sums, and Deviation Parameters

Group	Best Criterion	Worst Criterion	C1	C2	C3	C4	C5	C6	Vector Sum	ξ
Group 1	Safety (C1)	Technical Readiness (C6)	0.313	0.313	0.176	0.117	0.031	0.050	1.000	0.038
Group 2	Safety (C1)	Technical Readiness (C6)	0.363	0.204	0.204	0.035	0.136	0.058	1.000	0.083
Group 3	Regulatory Compliance (C3)	Technical Readiness (C6)	0.176	0.132	0.446	0.052	0.088	0.106	1.000	0.083

Table 9 Baseline BWM Weights Obtained as the Arithmetic Mean of the Three Group-Level Vectors

Code	Criterion	Arithmetic-Mean Weight	Rank
C1	Safety	0.284	1
C2	Security	0.216	3
C3	Regulatory Compliance	0.275	2
C4	Operational Efficiency	0.068	6
C5	Environmental Impact	0.085	4
C6	Technical Readiness	0.071	5

All three groups nominated Technical Readiness as the least important criterion. Groups 1 and 2 nominated Safety as the most important, whereas Group 3 nominated Regulatory Compliance, a difference consistent with the compliance and planning responsibilities concentrated in that group. Table 8 reports these nominations alongside the resulting weight vectors. Consistency is assessed through the deviation parameter returned by the optimization. Following Liang *et al.* (2020), the linear formulation is evaluated against input-based thresholds rather than through a ratio index derived from the nonlinear model, because the two formulations do not share a common scaling. The complete Best-to-Others and Others-to-Worst comparison vectors underlying each group solution are provided as supplementary material.

The baseline weight vector is the element-wise arithmetic mean of the three normalized group vectors. Because each group vector sums to unity, their mean does so as well, and no further normalization is applied at any point in the PROMETHEE II stage. Weights are reported to three decimals throughout, so the column total of 0.999 shown in Table 9 is a rounding artifact rather than a departure from normalization, and every downstream calculation retains full numerical precision. Aggregation, preference indices, and flows were generated with deterministic routines verified against a numerical tolerance of 10^{-9}.

Arithmetic aggregation is adopted as a transparent baseline rather than as a uniquely correct rule, and Section 4.1 reports geometric and median aggregation alongside it. To support auditability while respecting institutional restrictions, the de-identified group-level BWM records and the calculation workbook are available from the corresponding author on reasonable request, subject to institutional approval.

3.6 PROMETHEE II Specification

Let $f_{j(a)}$ denote the performance score of alternative a on criterion j. For each ordered pair of alternatives, the directional difference is calculated as shown in Equation 4.

$$d_j(a,b) = f_j(a) - f_j(b) \tag{4}$$

The baseline model uses the Type III V-shape preference function with $q = 0$. With preference threshold p_j, the function is shown in Equation 5.

$$P_j(a,b) = \begin{cases} 0, d_j(a,b) \le 0; \left(d_j(a,b)\right)/p_j, \\ 0 < d_j(a,b) \le p_j; 1, d_j(a,b) > p_j \end{cases} \tag{5}$$

The baseline uses $p = 8$ for all criteria as a transparent benchmark tied to the full 1-9 scoring range. It is not treated as an intrinsically optimal threshold. Additional scenarios use range-based criterion-specific sensitivity thresholds equal to the observed score ranges: 4.0, 4.1, 1.7, 1.5, 1.5, and 2.6 for C1-C6, respectively. These values are specification checks only and should not be interpreted as empirically validated operational indifference or preference thresholds. Type V linear and level functions use q and p values defined as proportions of the same observed ranges. The criterion-level intensities are then aggregated into the weighted

preference index defined in Equation 6, which expresses the overall degree to which one alternative is preferred to another and is asymmetric by construction.

$$\pi(a,b) = \sum (j=1)^n \, w_j P_j(a,b) \quad (6)$$

Illustrative audit calculation. For P1 relative to P2, the criterion score differences for C1-C6 are (4.0, 4.1, 1.7, -0.8, -0.3, and 2.6). Under the baseline Type III V-shape function with $p = 8$, the corresponding criterion-level preference intensities are (0.5000, 0.5125, 0.2125, 0, 0, and 0.3250). Using the arithmetic-mean weights reported in Table 9, the aggregate preference is $\pi\,(P1, P2)$ = (0.284000×0.5000) + (0.216333×0.5125) + (0.275333×0.2125) + (0.068000×0) + (0.085000×0) + (0.071333×0.3250) = 0.334563, reported as 0.335 in Table 12. In the reverse direction only C4 and C5 favor P2 over P1, giving $\pi(P2, P1)$ = (0.068000×0.1000) + (0.085000×0.0375) = 0.009988, reported as 0.010. Coefficients are shown to six decimals for legibility while the underlying computation retains full precision. The example illustrates both the criterion-level construction of the preference index and its directional asymmetry. Leaving and entering flows follow from Equations 7 and 8, and the net flow that orders the alternatives from Equation 9.

$$\varphi^+(a) = 1/(m-1) \sum (b \in A, b \neq a)\pi(a,b) \quad (7)$$

$$\varphi^-(a) = 1/(m-1) \sum (b \in A, b \neq a)\pi(b,a) \quad (8)$$

$$\varphi(a) = \varphi^+(a) - \varphi^-(a) \quad (9)$$

Internal computational validity requires the sum of positive flows to equal the sum of negative flows and the sum of net flows to equal zero, subject to rounding.

3.7 Robustness and Specification Analysis

The analysis distinguishes scenario stability from general robustness. Eight specifications are examined: arithmetic, geometric, and median aggregation; one-at-a-time weight perturbations of 10% and 20% in each direction; equal and capped weight structures; Safety-removed and Security-removed criterion sets; a mechanically merged Safety-Security diagnostic; four preference functions; uniform and range-based thresholds; and leave-one-group-out influence checks. Section 4.4 reports the parameters of each. The merged specification serves only as an algebraic diagnostic, since C1 and C2 order the alternatives identically and share a proportional threshold, so the composite reproduces the baseline contribution and cannot establish empirical independence. The removal scenarios therefore carry the internal diagnostic weight. None of these checks replaces external multi-site validation or a model that represents criterion interaction explicitly.

4. RESULTS

4.1 BWM Group Results and Aggregate Weights

Each group vector sums to unity, and the deviation parameters reported in Table 8 fall within the acceptable range for the linear formulation. These values indicate coherence of the comparisons within each group decision unit; they do not establish inter-rater reliability across nine experts. The group-level rank profiles nevertheless show substantial concordance (Kendall's W = 0.806; chi-square (5) = 12.10; p = 0.034, computed with mid-ranks because two groups produced tied weights), indicating broad agreement in the relative ordering of the six criteria despite differences in weight magnitude and in the criterion two of the three groups nominated as most important. With only three profiles, however, the statistic carries limited power and is reported as descriptive evidence.

Element-wise averaging of the three group vectors yields the baseline weights reported in Table 9. Because each group vector sums to unity, their mean does so as well; the displayed column total of 0.999 reflects three-decimal rounding only, and all PROMETHEE II calculations retain full precision.

The baseline gives Safety the highest weight (0.284), followed by Regulatory Compliance (0.275), Security (0.216), Environmental Impact (0.085), Technical Readiness (0.071), and Operational Efficiency (0.068). Safety, Security, and Regulatory Compliance collectively account for 77.5% of the total decision weight. This concentration represents the risk priorities of the panel, but it also creates a structural advantage for alternatives that perform strongly on those dimensions.

Figure 4 visualizes the baseline arithmetic-mean BWM weight distribution and makes the concentration of decision importance in Safety, Regulatory Compliance, and Security immediately visible. The bars fall into two clearly separated tiers: these three criteria each exceed 0.21, whereas Environmental Impact, Technical Readiness, and Operational Efficiency each remain below 0.09. Table 10 compares these baseline weights with geometric-mean and median aggregation. The geometric mean changes no weight by more than 0.01. The median, renormalized to unity, raises Safety to 0.341 and lowers Regulatory Compliance to 0.222 because it discounts the comparatively strong emphasis that Group 3 placed on Regulatory Compliance (Table 8). Both alternative rules nevertheless preserve the same three-criterion upper tier. Although the magnitudes shift across aggregation rules, the comparison is used as a specification check rather than as evidence that one aggregation rule is intrinsically superior.

4.2 Decision Matrix

Table 11 presents the aggregated expert performance scores used as the PROMETHEE II decision matrix. P1 records the highest scores on Safety, Security, and Regulatory Compliance. P4 performs strongly on Operational Efficiency and Technical Readiness, while P5 records the highest Environmental Impact score. P2 and P3 perform better than P1 on Operational Efficiency, and P5 performs better on Environmental Impact, which is why P1 records a non-zero entering flow rather than complete pairwise dominance.

Table 11 reports central-tendency inputs only. Cell-level standard deviations and interquartile ranges cannot be supplied retrospectively, because the anonymized nine-expert rating panel was consolidated before archiving and is no longer available in analyzable form. No dispersion statistics have been imputed. The aggregated means must therefore not be read as evidence of unanimity or low disagreement, and the limitation is carried into Section 5.6 and the data availability statement.

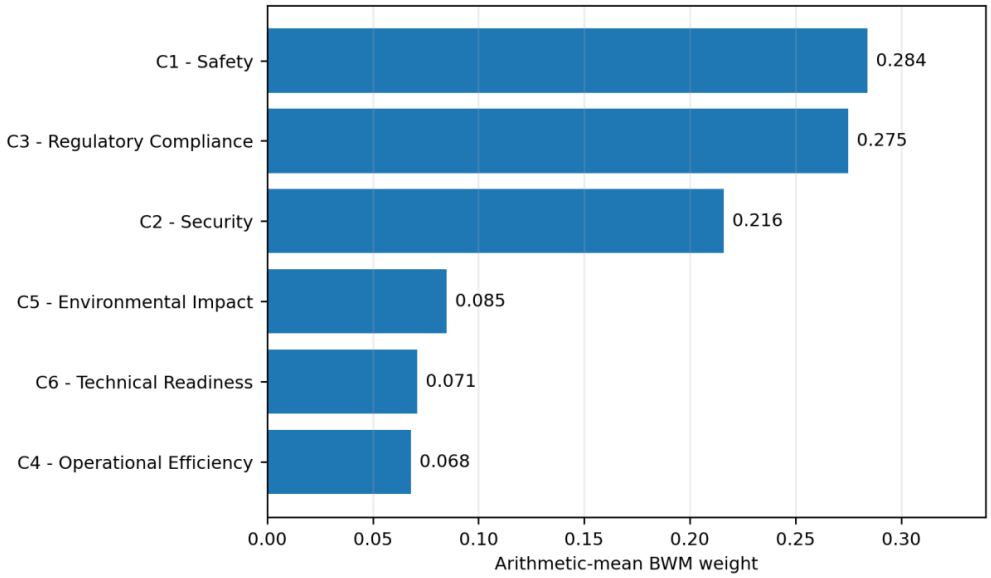


Figure 4 Arithmetic-Mean BWM Weight Distribution Across the Six Criteria

Table 10 Criterion Weights Under Alternative Aggregation Rules

Criterion	Arithmetic Mean	Geometric Mean	Median
C1 Safety	0.284	0.293	0.341
C2 Security	0.216	0.220	0.222
C3 Regulatory Compliance	0.275	0.272	0.222
C4 Operational Efficiency	0.068	0.064	0.057
C5 Environmental Impact	0.085	0.078	0.096
C6 Technical Readiness	0.071	0.073	0.063

Table 11 Aggregated Expert Performance Scores on the 1-9 Scale

Alternative	C1	C2	C3	C4	C5	C6
P1	8.8	8.6	8.2	7.0	6.5	7.8
P2	4.8	4.5	6.5	7.8	6.8	5.2
P3	5.2	4.8	7.3	8.2	7.0	5.8
P4	7.5	7.0	7.0	8.5	7.5	7.5
P5	7.0	6.5	7.2	7.5	8.0	7.0

Table 12 Pairwise Preference-Index Matrix Under the Baseline Type III V-Shape Specification (p = 8)

Alternative	P1	P2	P3	P4	P5
P1	-	0.335	0.279	0.133	0.162
P2	0.010	-	0.000	0.000	0.003
P3	0.016	0.061	-	0.010	0.009
P4	0.023	0.215	0.164	-	0.044
P5	0.020	0.185	0.131	0.012	-

4.3 PROMETHEE II Results

4.3.1 Pairwise Preference Matrix

Table 12 reports the baseline pairwise preference-index matrix under the Type III V-shape specification with p = 8. This matrix is the step at which PROMETHEE II consolidates the criterion-level comparisons into a single measure for every ordered pair of alternatives. Each off-diagonal element $\pi(a,b)$, obtained from Equation 6, expresses the weighted degree to which the row alternative a is preferred to the column alternative b. Because the weights sum to unity and each criterion-level intensity lies between 0 and 1, every element is bounded by 0, denoting no preference in that direction, and 1, denoting full preference on every criterion. The matrix is directional and therefore asymmetric: $\pi(a,b)$ and $\pi(b,a)$ are computed separately and are not complements. For example, P1 is preferred to P2 with an index of 0.335, whereas the reverse preference P2 over P1 is only 0.010, a residual that arises because P2 scores higher than P1 on Operational Efficiency and Environmental Impact. The elements also make each relation traceable. The zero entries for P2 relative to P3 and P4 show that P2 scores lower than both alternatives on all six criteria, whereas the largest entry, π (P1, P2) = 0.335, arises chiefly from margins of about four scale points on the heavily weighted Safety and Security criteria. These magnitudes are informative in relative rather than absolute terms. Under the uniform threshold, a criterion-level intensity reaches 1 only when one alternative scores 9 and the other 1, so the observed differences of at most 4.1 points generate partial intensities. The matrix finally supplies the inputs to the preference flows:

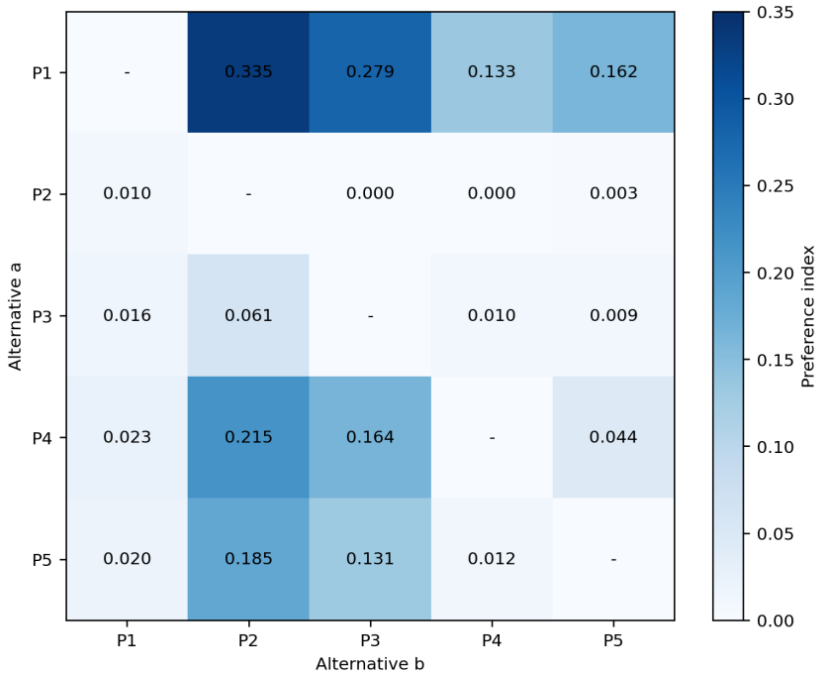


Figure 5 Pairwise Preference-Index Heatmap. The matrix shows directional preference rather than absolute effectiveness.

Table 13 PROMETHEE II Preference Flows and Final Ranking

Code	Alternative	Positive Flow	Negative Flow	Net Flow	Rank
P1	Integrated Security System Enhancement	0.227	0.017	+0.210	1
P4	Use of Automated/High-Performance Equipment	0.112	0.039	+0.073	2
P5	Integrated Risk Management Implementation	0.087	0.055	+0.033	3
P3	Intensive Training and Certification Programs	0.024	0.144	-0.120	4
P2	Provision of Specialized Cranes and Dock Infrastructure	0.003	0.199	-0.196	5
Total	-	0.453	0.453	0.000	-

the mean of the off-diagonal elements in each row gives the positive flow of that alternative, and the mean of each column gives its negative flow (Equations 7 and 8). The P1 row thus yields a positive flow of 0.227, and the P2 column yields a negative flow of 0.199. Comparing each pair of mirror elements further shows that the pairwise relations are transitive: in all ten unordered pairs, the larger index favors the alternative ranked higher in Table 13, and the weakest separation occurs between P4 and P5 (0.044 against 0.012).

Figure 5 reproduces the same matrix as a heatmap so that its directional structure can be read at a glance. Rows identify the preferring alternative a and columns the alternative b with which it is compared. Each cell is shaded on a sequential scale from white (0.00) to dark blue (0.35) and carries its numerical value, while the diagonal cells are marked with a dash because an alternative is not compared with itself. Reading along a row therefore shows how strongly an alternative outranks the others, and reading down a column shows how strongly it is outranked by them. Directionality appears as asymmetry about the diagonal, where a dark cell paired with a pale mirror cell indicates that preference runs predominantly in one direction. Three patterns stand out. The P1 row forms the darkest band, particularly against P2 (0.335) and P3 (0.279), while the P1 column remains uniformly pale, which shows that P1 outranks every alternative and is itself only marginally outranked. The P2 row is almost entirely white, whereas the P2 column is the darkest in the matrix, placing P2 at the opposite end of the preference structure. The P4 and P5 rows show moderate shading against P2 and P3 but only pale cells against each other, which signals the comparatively narrow separation between the alternatives ranked second and third. Because the heatmap encodes ordered pairwise relations rather than absolute performance, it should be interpreted together with the positive, negative, and net flows reported in the following subsection.

4.3.2 Positive, Negative, and Net Flows

Table 13 summarizes the positive, negative, and net PROMETHEE II flows that convert the pairwise preference relations into a complete ranking. P1 remains the highest-ranked alternative, with a positive flow of 0.227, a negative flow of 0.017, and a net flow of +0.210. It is followed by P4 (+0.073), P5 (+0.033), P3 (-0.120), and P2 (-0.196). The equality of total positive and negative flows and the approximately zero sum of net flows confirm internal computational consistency. P1 is therefore the strongest relative alternative under the baseline specification, but it is not perfectly dominant or universally superior. Figure 6 visualizes the resulting net flows around the zero-reference line, making the separation between alternatives with positive relative positions (P1, P4, and P5) and negative relative positions (P3 and P2) readily apparent. The chart therefore complements Table 13 by emphasizing relative separation within the evaluated set rather than absolute effectiveness.

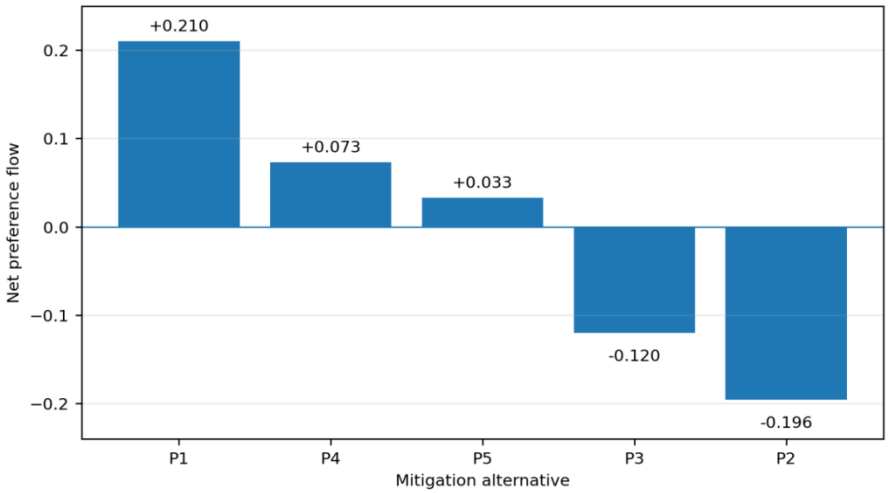


Figure 6 PROMETHEE II Net Flows. Positive and negative values indicate relative position within the focal alternative set, not absolute effectiveness.

Table 14 One-at-a-Time Weight-Perturbation Results for C1-C6

Criterion	Change	P1	P4	P5	P3	P2	Ranking
C1	-20%	+0.200	+0.068	+0.031	-0.111	-0.188	P1 > P4 > P5 > P3 > P2
C1	-10%	+0.205	+0.070	+0.032	-0.115	-0.192	P1 > P4 > P5 > P3 > P2
C1	+10%	+0.215	+0.075	+0.033	-0.124	-0.199	P1 > P4 > P5 > P3 > P2
C1	+20%	+0.220	+0.077	+0.034	-0.128	-0.203	P1 > P4 > P5 > P3 > P2
C2	-20%	+0.202	+0.070	+0.032	-0.114	-0.191	P1 > P4 > P5 > P3 > P2
C2	-10%	+0.206	+0.072	+0.033	-0.117	-0.193	P1 > P4 > P5 > P3 > P2
C2	+10%	+0.214	+0.074	+0.033	-0.123	-0.198	P1 > P4 > P5 > P3 > P2
C2	+20%	+0.219	+0.075	+0.033	-0.126	-0.200	P1 > P4 > P5 > P3 > P2
C3	-20%	+0.215	+0.081	+0.036	-0.130	-0.202	P1 > P4 > P5 > P3 > P2
C3	-10%	+0.212	+0.077	+0.034	-0.125	-0.199	P1 > P4 > P5 > P3 > P2
C3	+10%	+0.208	+0.068	+0.031	-0.115	-0.193	P1 > P4 > P5 > P3 > P2
C3	+20%	+0.206	+0.064	+0.030	-0.110	-0.190	P1 > P4 > P5 > P3 > P2
C4	-20%	+0.215	+0.072	+0.034	-0.122	-0.198	P1 > P4 > P5 > P3 > P2
C4	-10%	+0.213	+0.072	+0.033	-0.121	-0.197	P1 > P4 > P5 > P3 > P2
C4	+10%	+0.208	+0.073	+0.032	-0.118	-0.194	P1 > P4 > P5 > P3 > P2
C4	+20%	+0.205	+0.073	+0.031	-0.117	-0.193	P1 > P4 > P5 > P3 > P2
C5	-20%	+0.216	+0.073	+0.031	-0.121	-0.198	P1 > P4 > P5 > P3 > P2
C5	-10%	+0.213	+0.073	+0.032	-0.121	-0.197	P1 > P4 > P5 > P3 > P2
C5	+10%	+0.207	+0.072	+0.033	-0.119	-0.194	P1 > P4 > P5 > P3 > P2
C5	+20%	+0.204	+0.072	+0.034	-0.118	-0.193	P1 > P4 > P5 > P3 > P2
C6	-20%	+0.211	+0.072	+0.032	-0.119	-0.195	P1 > P4 > P5 > P3 > P2
C6	-10%	+0.210	+0.072	+0.032	-0.120	-0.195	P1 > P4 > P5 > P3 > P2
C6	+10%	+0.210	+0.073	+0.033	-0.120	-0.196	P1 > P4 > P5 > P3 > P2
C6	+20%	+0.210	+0.074	+0.033	-0.120	-0.196	P1 > P4 > P5 > P3 > P2

4.4 Robustness and Specification Results

The robustness assessment examines whether the ordinal order and the strength of preference change under alternative modeling choices. Across all reported specifications, the ranking remains P1 > P4 > P5 > P3 > P2. However, net-flow magnitudes and P1-P4 margins vary, demonstrating structural sensitivity in the strength of dominance. Table 14 reports the 24 one-at-a-time ±10% and ±20% weight-perturbation scenarios for C1-C6 and is used to assess local sensitivity of the ranking to individual criterion weights.

All 24 one-at-a-time scenarios retain the baseline order. Across these scenarios, the net flow of P1 varies only between +0.200 and +0.220, and its margin over P4 never falls below 0.13. Perturbations of the three dominant criteria produce the largest shifts, whereas perturbations of Technical Readiness leave all flows virtually unchanged. The narrowest adjacent separation arises between P4 and P5 when the weight of Regulatory Compliance is increased by 20% (+0.064 against +0.030), yet even this case preserves the baseline order. Taken together, these results demonstrate local ordinal stability under the tested weight changes. They do not establish general robustness because other modeling choices also influence the flow magnitudes.

P1 remains first under equal weights and the capped-weight specification, but its lead over P4 falls from 0.138 in the baseline to 0.049 under equal weights and 0.082 under capped weights. Removing Safety or Security also preserves the ranking while reducing P1's margin. The mechanically merged Safety-Security diagnostic exactly reproduces the baseline flows because C1 and C2 share the same directional ordering, the same Type III V-shape threshold, and an additive weighted structure. This algebraic equivalence is not evidence that the criteria are empirically independent.

Table 15 Weight-Structure and Criterion-Removal Specifications

Weight/Criterion Specification	Ranking	P1 Net Flow	P4 Net Flow	P1-P4 Margin
Baseline arithmetic aggregation	P1 > P4 > P5 > P3 > P2	+0.210	+0.073	0.138
Equal weights	P1 > P4 > P5 > P3 > P2	+0.133	+0.083	0.049
Top-three weights capped at 0.20	P1 > P4 > P5 > P3 > P2	+0.161	+0.079	0.082
Safety removed	P1 > P4 > P5 > P3 > P2	+0.161	+0.049	0.111
Security removed	P1 > P4 > P5 > P3 > P2	+0.168	+0.062	0.106
Mechanically merged Safety-Security (algebraic diagnostic)	P1 > P4 > P5 > P3 > P2	+0.210	+0.073	0.138

Table 16 Preference-Function and Threshold Sensitivity Specifications

Preference Specification	Ranking	P1	P4	P5
V-shape, uniform p = 8	P1 > P4 > P5 > P3 > P2	+0.210	+0.073	+0.033
Usual preference	P1 > P4 > P5 > P3 > P2	+0.694	+0.259	+0.051
V-shape, range-based criterion-specific p	P1 > P4 > P5 > P3 > P2	+0.484	+0.166	+0.091
Linear Type V, q = 0.25 observed range; p = observed range	P1 > P4 > P5 > P3 > P2	+0.416	+0.137	+0.075
Level, q = 0.25 observed range; p = 0.50 observed range	P1 > P4 > P5 > P3 > P2	+0.610	+0.263	+0.112

Table 17 Leave-One-Group-Out Influence Analysis

Influence Specification	Ranking	P1	P4	P5	P3	P2
Group 1 excluded	P1 > P4 > P5 > P3 > P2	+0.197	+0.064	+0.035	-0.108	-0.188
Group 2 excluded	P1 > P4 > P5 > P3 > P2	+0.206	+0.068	+0.027	-0.111	-0.190
Group 3 excluded	P1 > P4 > P5 > P3 > P2	+0.227	+0.086	+0.036	-0.140	-0.209

The Safety-removed and Security-removed scenarios are therefore the more informative internal diagnostics, while formal dependence modeling remains a future requirement. These structural checks are summarized in Table 15, which reports the ranking, P1 and P4 net flows, and the P1-P4 margin under alternative weight structures and criterion-removal specifications.

Table 16 compares the baseline with alternative preference-function and threshold specifications. The ordinal ranking remains unchanged across the Type III V-shape baseline, usual, Type V linear, level, and range-based criterion-specific threshold specifications. Nevertheless, net-flow magnitudes change materially. The range-based thresholds are sensitivity devices derived from the observed score dispersion; they are not validated operational thresholds. Accordingly, the result supports scenario stability of the order but not invariance of preference strength or empirical validation of a single threshold structure.

Table 17 reports the leave-one-group-out influence analysis. Excluding any one deliberation group does not change the ordinal order. The variation in net-flow magnitudes confirms that group composition influences the strength of preference. These checks are influence diagnostics and should not be interpreted as external validation

5. DISCUSSION

5.1 Interpretation of the Weight Structure

The arithmetic aggregation assigns the highest weight to Safety (0.284), followed closely by Regulatory Compliance (0.275) and Security (0.216). These three criteria collectively account for 77.5% of the baseline weight. The result is substantively consistent with a military ammunition-handling environment, where accidental initiation, unauthorized access, and non-compliance can produce irreversible consequences. Regulatory Compliance rather than Security occupies the second position, which reflects the weight this panel places on inspection readiness and audit traceability in a setting where ordnance handling is subject to layered regulatory oversight. Operational Efficiency receives the lowest baseline weight (0.068), followed by Technical Readiness (0.071). These values indicate relative prioritization within a safety-effectiveness decision, not organizational indifference to efficiency or equipment readiness.

The strong concentration in three criteria should be interpreted in two ways. It is an empirical expression of the panel's institutional priorities, and it is simultaneously a structural feature that can favor alternatives performing strongly on those dimensions. The equal-weight and capped-weight scenarios demonstrate that P1 remains first when the concentration is reduced, but the P1-P4 margin contracts materially. The rank order is therefore stable in the tested scenarios, while the apparent strength of P1's dominance is weight-dependent.

5.2 Interpretation of the Ranking and Structural Dominance

P1 ranks first because it performs strongly on Safety, Security, and Regulatory Compliance and maintains competitive Technical Readiness. Its net flow of +0.210 does not indicate perfect operational effectiveness. It represents relative dominance within the specified criteria, weights, alternative set, preference function, and thresholds. P1 has a non-zero negative flow because other alternatives outperform it on selected criteria. P4, for example, performs better on Operational Efficiency, and P5 performs better on Environmental Impact.

P4 ranks second because automation and high-performance equipment combine strong Safety, Security, Operational Efficiency, and Technical Readiness scores. P5 ranks third through balanced governance, compliance, environmental, and technical performance. P3 ranks fourth despite the importance of training because its performance on Safety and Security is lower than the system-level alternatives under the current scoring and weighting structure. P2 ranks fifth because its principal advantages are

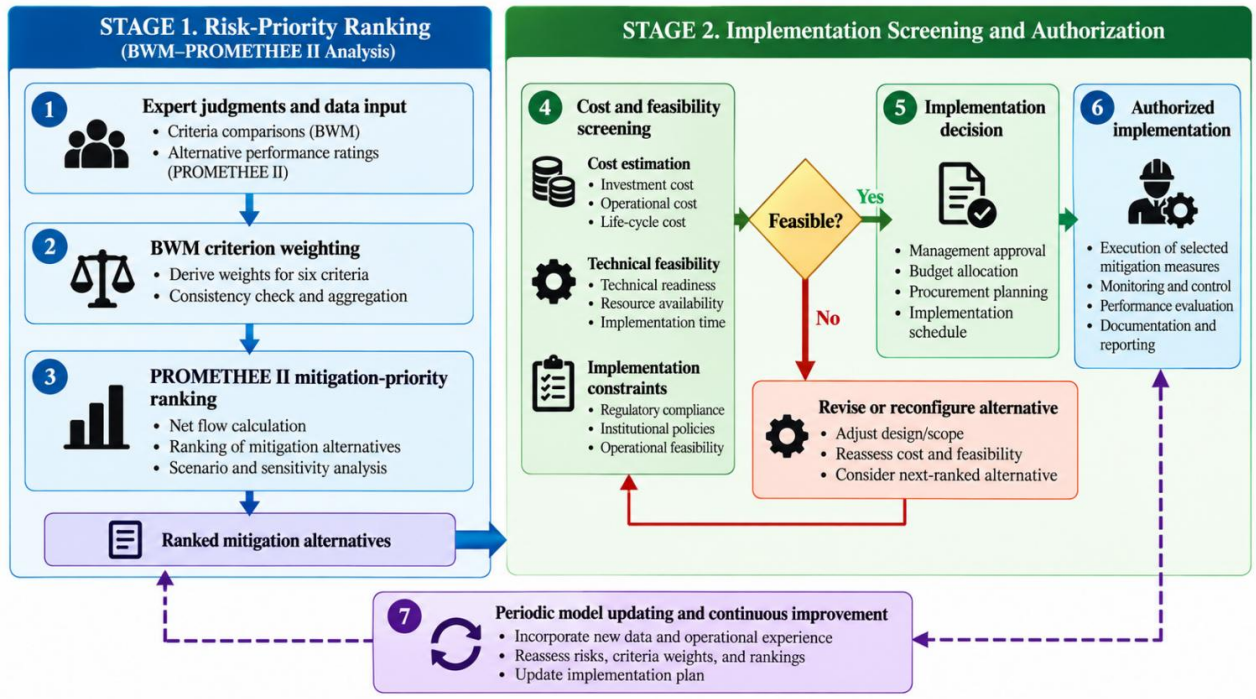


Figure 7 Two-Stage Decision Pathway Linking Mitigation-Priority Ranking, Cost-and-Feasibility Screening, and Authorized Implementation

concentrated in Operational Efficiency and infrastructure readiness, which receive comparatively low baseline weights. These positions identify relative risk-reduction priority; they do not establish that lower-ranked measures are unnecessary.

Safety and Security remain conceptually distinct because accidental and intentional threats require different controls. The Safety-removed and Security-removed specifications do not generate rank reversal, but both reduce P1's margin, confirming that the two criteria contribute materially to its dominance. A mechanically combined criterion is not interpreted as evidence against double counting because, under proportional preference and additive aggregation, such a construction can preserve the baseline contribution algebraically. Future research should undertake a dependence-specific elicitation of criterion interactions and/or apply an interaction-aware approach such as ANP, DEMATEL, a Choquet integral, or another non-additive model.

5.3 Methodological Interpretation: Scenario Stability, Not Conclusive Robustness

The unchanged ordinal order across aggregation, weight, criterion, preference-function, threshold, OAT, and leave-one-group-out specifications is meaningful, but it should not be described as conclusive robustness. The net-flow magnitudes and margins are not invariant. Under equal weights, the P1-P4 margin falls to 0.049, whereas alternative preference functions generate much larger flow magnitudes. The appropriate conclusion is that the ranking is stable within the examined scenario set while the strength of dominance is sensitive to model structure.

A stable order can arise because alternatives are well separated, because several assumptions reinforce the same alternative, or because the model is rigid. Reporting the ordinal order together with the changing margins prevents the sensitivity analysis from becoming a confirmation exercise, and robustness analysis is therefore treated here as specification diagnosis rather than as proof of model validity.

5.4 Contribution and Transferability

The contribution is deliberately bounded. Methodologically, the study demonstrates how established BWM and PROMETHEE II procedures can be applied transparently in a safety-critical interface decision. Empirically, it shows how the focal panel balances six criteria and how the ranking changes in magnitude under alternative specifications. Practically, it separates risk-priority assessment from implementation screening. The study does not claim a new MCDM algorithm or a general theory of ammunition logistics.

Transferability operates at three levels. The weights, expert scores, and ranking are specific to the focal depot. The criterion categories may suit comparable military or highly regulated maritime settings, but their definitions and weights require revalidation. The most transferable element is the analytical process itself, from decision-boundary definition through elicitation, weighting, outranking, and specification analysis to cost-feasibility screening.

5.5 Managerial Implementation Pathway

The PROMETHEE II result is a risk-effectiveness priority ranking, not a cost-effectiveness analysis or final investment decision. Decision makers should use it as the first stage of a two-stage process. Stage 1 identifies alternatives with the strongest expected contribution to risk reduction. Stage 2 evaluates shortlisted alternatives against implementation cost, technical feasibility, implementation duration, organizational readiness, resource availability, maintenance requirements, and regulatory authorization. Figure 7 depicts this two-stage pathway, separating the analytical risk-priority ranking from the managerial authorization process so that a high PROMETHEE II rank is not interpreted as an automatic procurement decision.

Table 18 Risk-Priority and Implementation-Feasibility Portfolio

Portfolio Position	Decision Treatment	Required Management Action
High priority / High feasibility	Proceed to pilot or controlled implementation	Confirm authorization, budget, responsibilities, and performance indicators
High priority / Low feasibility	Develop phased implementation	Complete engineering design, budget planning, and resource mobilization
Lower priority / High feasibility	Use as a complementary or mandatory control	Implement when inexpensive, rapidly deployable, or legally required
Lower priority / Low feasibility	Defer and monitor	Retain as contingency and reassess when conditions change

Read from left to right, Figure 7 shows that Stage 1 channels expert judgments and data inputs into BWM criterion weighting and the PROMETHEE II ranking, which together yield the ordered list of mitigation alternatives. In Stage 2, each shortlisted alternative is screened for cost, covering investment, operational, and life-cycle cost; for technical feasibility, covering technical readiness, resource availability, and implementation time; and for implementation constraints arising from regulatory compliance, institutional policies, and operational feasibility. Alternative judged feasible proceeds to an implementation decision involving management approval, budget allocation, procurement planning, and scheduling, and then to authorized implementation with monitoring, performance evaluation, and documentation. An alternative judged infeasible is revised or reconfigured by adjusting its design or scope, after which its cost and feasibility are reassessed or the next-ranked alternative is considered. The pathway closes with periodic model updating, in which new data and operational experience are used to reassess risks, criterion weights, rankings, and the implementation plan. These post-screening steps correspond to the implementation decision and periodic model updating that follow the analytical sequence summarized in Figure 1.

Under this pathway, P1 should not be purchased solely because it ranks first. It should proceed to technical, financial, organizational, and regulatory screening. P3 or P5 may still be implemented earlier when they are less costly, immediately feasible, or required as supporting controls. The ranking structures deliberation; it does not replace managerial accountability. Table 18 translates the combination of risk priority and implementation feasibility into four portfolio positions, linking each position to an appropriate decision treatment and management action. This portfolio is intentionally qualitative because cost, budget, engineering feasibility, and authorization data were not included in the baseline PROMETHEE II model.

5.6 Limitations and Future Research

The study has seven principal limitations. First, the panel consists of nine experts from one depot, so the numerical outputs reflect the institutional perspective of that site. Second, the effective BWM decision units are three post-deliberation groups rather than nine independent models, and deliberation may introduce conformity or authority effects alongside its clarifying function. Third, the deviation parameters and Kendall's W assess coherence and concordance among group profiles rather than individual-level inter-rater reliability. Fourth, the individual rating panel was not retained in analyzable form, so dispersion statistics and retrospective ICC estimates cannot be reconstructed. Fifth, Safety and Security may overlap conceptually, and removal or merging tests cannot fully model dependence. Sixth, the ranking remains conditional on aggregation, weights, preference functions, thresholds, and expert scores, and the range-based thresholds are sensitivity specifications rather than validated operational values. Seventh, cost and feasibility are excluded from the baseline model and addressed only through a second-stage screen.

Future research should use multi-site expert panels, retain anonymized individual pre- and post-deliberation judgments, report individual-level agreement and dispersion statistics, establish expert-calibrated operational thresholds, and undertake dependence-specific elicitation of criterion interactions. Interaction-aware approaches such as ANP, DEMATEL, and Choquet-integral models should be compared with the present additive specification. Future applications should also publish de-identified comparison vectors and executable calculation files where institutional security requirements permit.

6. CONCLUSION

This study applies an integrated BWM-PROMETHEE II procedure to prioritize five mitigation alternatives at a maritime ammunition storage-transport interface. The arithmetic aggregation assigns weights of 0.284 to Safety, 0.216 to Security, 0.275 to Regulatory Compliance, 0.068 to Operational Efficiency, 0.085 to Environmental Impact, and 0.071 to Technical Readiness. Safety, Security, and Regulatory Compliance collectively account for 77.5% of the baseline weight. Under the Type III V-shape PROMETHEE II baseline with $p = 8$, Integrated Security System Enhancement ranks first with a net flow of +0.210. It is followed by Automated/High-Performance Equipment (+0.073), Integrated Risk Management (+0.033), Intensive Training and Certification (-0.120), and Specialized Cranes and Dock Infrastructure (-0.196). P1 does not achieve perfect dominance. Its first-place position is conditional on the focal case, expert panel, criteria, weights, preference function, thresholds, and alternative set. The ranking remains unchanged across the tested aggregation rules, weight structures, criterion structures, preference functions, threshold specifications, one-at-a-time perturbations, and leave-one-group-out checks. However, net-flow magnitudes and rank margins vary. The evidence therefore supports scenario stability within the examined specifications rather than conclusive robustness. The main transferable contribution is the decision process, not the numerical ranking. The procedure links resilience requirements to criteria, expert weighting, outranking, specification analysis, and implementation screening. The resulting order should be treated as a risk-effectiveness priority ranking and must be followed by cost, feasibility, readiness, and authorization assessment before procurement or implementation decisions are made.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest with respect to the research, authorship, and publication of this article.

DATA AVAILABILITY STATEMENTS

The de-identified group-level BWM outputs, aggregated performance matrix, and calculation workbook are available from the corresponding author on reasonable request, subject to institutional and operational-security restrictions. The anonymized individual rating panel was not retained in analyzable form; cell-level standard deviations, interquartile ranges, and individual agreement statistics cannot be reconstructed, imputed, or inferred from the aggregated matrix.

REFERENCES

Ahlqvist, V., Norrman, A., & Jahre, M. (2020). Supply chain risk governance: Towards a conceptual multi-level framework. *Operations and Supply Chain Management: An International Journal, 13*(4), 382–395. https://doi.org/10.31387/oscm0430278

Aldrighetti, R., Battini, D., & Ivanov, D. (2023). Efficient resilience portfolio design in the supply chain with consideration of preparedness and recovery investments. *Omega, 117*, 102841. https://doi.org/10.1016/j.omega.2023.102841

Al-Hajj, S., Dhaini, H. R., Mondello, S., Kaafarani, H., Kobeissy, F., & DePalma, R. G. (2021). Beirut ammonium nitrate blast: analysis, review, and recommendations. *Frontiers in Public Health, 9*, 657996. https://doi.org/10.3389/fpubh.2021.657996

Altubaishe, B., & Desai, S. (2023). Multicriteria decision making in supply chain management using FMEA and hybrid AHP-PROMETHEE algorithms. *Sensors, 23*(8), 4041. https://doi.org/10.3390/s23084041

Batista, M., Ribeiro, J. P., & Barbosa-Póvoa, A. (2023). Supply chain resilience: Tactical-operational models, a literature review. In J. P. Almeida, F. P. E. Alvelos, J. O. Cerdeira, S. Moniz, & C. Requejo (Eds.), *Operational research—IO 2022—OR in turbulent times: Adaptation and resilience. XXII Congress of APDIO, 2022* (pp. 157–177). Springer. https://doi.org/10.1007/978-3-031-46439-3_12

Dolgui, A., & Ivanov, D. (2021). Ripple effect and supply chain disruption management: new trends and research directions. *International Journal of Production Research, 59*(1), 102–109. https://doi.org/10.1080/00207543.2021.1840148

Ecer, F. (2024). A state-of-the-art review of the BWM method and future research agenda. *Technological and Economic Development of Economy, 30*(4), 1165–1204. https://doi.org/10.3846/tede.2024.20761

Enyinda, C. I. (2018). Modeling enterprise risk management in operations and supply chain: A pharmaceutical firm context. *Operations and Supply Chain Management: An International Journal, 11*(1), 1–12. https://doi.org/10.31387/oscm0300195

Fagundes, M. V. C., Teles, E. O., de Melo, S. A. V., & Freires, F. G. M. (2020). Decision-making models and support systems for supply chain risk: literature mapping and future research agenda. *European Research on Management and Business Economics, 26*(2), 63–70. https://doi.org/10.1016/j.iedeen.2020.02.001

Fan, Y., & Stevenson, M. (2018). A review of supply chain risk management: definition, theory, and research agenda. International *Journal of Physical Distribution and Logistics Management, 48*(3), 205–230. https://doi.org/10.1108/IJPDLM-01-2017-0043

Gaudenzi, B., Pellegrino, R., & Confente, I. (2023). Achieving supply chain resilience in an era of disruptions: a configuration approach of capacities and strategies. *Supply Chain Management: An International Journal, 28*(7), 97–111. https://doi.org/10.1108/SCM-09-2022-0383

Gholamizadeh, K., Rostami, C., Zarei, E., Yazdi, M., & Moslem, S. (2025). Advancing hazardous materials transport safety: systematic insights on risks, challenges, and research gaps. *Journal of Safety Science and Resilience*, 100226. https://doi.org/10.1016/j.jnlssr.2025.100226

Guo, J., & Luo, C. (2022). Risk assessment of hazardous materials transportation: a review of research progress in the last thirty years. *Journal of Traffic and Transportation Engineering, 9*(4), 571–590. https://doi.org/10.1016/j.jtte.2022.01.004

Hosseini, S., Ivanov, D., & Blackhurst, J. (2020). Conceptualization and measurement of supply chain resilience in an open-system context. *IEEE Transactions on Engineering Management, 69*(6), 3111–3126. https://doi.org/10.1109/TEM.2020.3026465

Hsu, C., Yang, J., Chang, A., & Liu, G. (2024). A new hybrid MCDM approach for mitigating risks of hazardous material road transportation. *Mathematical Biosciences and Engineering, 21*(3), 4210–4240. https://doi.org/10.3934/mbe.2024186

Ivanov, D. (2023). Collaborative emergency adaptation for ripple effect mitigation in intertwined supply networks. *Annals of Operations Research*, 1–17. https://doi.org/10.1007/s10479-023-05408-0

Ivanov, D. (2024). Two views of supply chain resilience. *International Journal of Production Research, 62*(11), 4031–4045. https://doi.org/10.1080/00207543.2023.2253328

Ivanov, D., & Dolgui, A. (2020). Viability of intertwined supply networks: extending the supply chain resilience angles towards survivability. *International Journal of Production Research*, 58(10), 2904–2915. https://doi.org/10.1080/00207543.2020.1750727

Ke, G. Y. (2022). Managing rail-truck intermodal transportation for hazardous materials with random yard disruptions. *Annals of Operations Research, 309*(2), 457-483. https://doi.org/10.1007/s10479-020-03699-1

Khan, R. U., Yin, J., Mustafa, F. S., & Wang, S. (2022). Analyzing human factor involvement in sustainable hazardous cargo port operations. *Ocean Engineering*, 250, 111028. https://doi.org/10.1016/j.oceaneng.2022.111028

Khan, R. U., Yin, J., Mustafa, F. S., & Wang, S. (2024). Seaport hazardous cargo loading and unloading risk assessment using interval type-2 fuzzy sets and Bayesian networks. ASCE-ASME *Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering, 10*(1), 04023065. https://doi.org/10.1061/ajrua6.rueng-1120

Kheybari, S., Javdanmehr, M., Rezaie, F. M., & Rezaei, J. (2021). Corn cultivation location selection for bioethanol production: an application of BWM and extended PROMETHEE II. *Energy, 228,* 120593. https://doi.org/10.1016/j.energy.2021.120593

Lasse, R. T. N., Gurning, R. O. S., Baihaqi, I., & Wiradanti, B. (2025). Investigating digital collaboration: Human factors in artificial intelligence-driven collaboration platforms for resilient port. *Operations and Supply Chain Management: An International Journal, 18*(2), 185–200. https://doi.org/10.31387/oscm0610465

Liang, F., Brunelli, M., & Rezaei, J. (2020). Consistency issues in the best worst method: measurements and thresholds. *Omega*, 96, 102175. https://doi.org/10.1016/j.omega.2019.102175

Liu, L., Sun, S., & Li, S. (2025). Route optimization of hazardous material railway transportation based on conditional value-at-risk considering risk equity. *Mathematics, 13*(5), 803. https://doi.org/10.3390/math13050803

Men, J., Chen, G., Yang, Y., & Reniers, G. (2022). An event-driven probabilistic methodology for modeling the spatial-temporal evolution of natural hazard-induced domino chain in chemical industrial parks. *Reliability Engineering and System Safety, 226*, 108723. https://doi.org/10.1016/j.ress.2022.108723

Mohri, S. S., Mohammadi, M., Gendreau, M., Pirayesh, A., Ghasemaghaei, A., & Salehi, V. (2022). Hazardous material transportation problems: a comprehensive overview of models and solution approaches. *European Journal of Operational Research, 302*(1), 1–38. https://doi.org/10.1016/j.ejor.2021.11.045

Moktadir, M. A., Paul, S. K., Bai, C., & Santibanez Gonzalez, E. D. (2025). The current and future states of MCDM methods in sustainable supply chain risk assessment. *Environment, Development and Sustainability, 27*(3), 7435–7480. https://doi.org/10.1007/s10668-023-04200-1

Morteza Beigi, L., Hassini, E., & Soltani, N. (2024). A review of models for dependency of risks: Extension and applications to supply chains. *Operations and Supply Chain Management: An International Journal, 17*(2), 316–329. https://doi.org/10.31387/oscm0570430

Peron, M., Arena, S., Paltrinieri, N., Sgarbossa, F., & Boustras, G. (2023). Risk assessment for handling hazardous substances within the European industry: available methodologies and research streams. *Risk Analysis, 43(*7), 1434–1462. https://doi.org/10.1111/risa.14010

Sahab, S., & Oulfarsi, S. (2024). Trends and insights in supply chain risk management: A comprehensive review of literature. *Operations and Supply Chain Management: An International Journal, 17*(4), 402–416. https://doi.org/10.31387/oscm0590450

Sari, A., & Öztürk, S. (2025). Comprehensive risk assessment approach for storage tank facilities against hazards including fire, explosion, and earthquake. *Process Safety Progress, 44*(1), 70–85. https://doi.org/10.1002/prs.12666

Taghipour, A., Fooladvand, A., Khazaei, M., & Ramezani, M. (2023). Criteria clustering and supplier segmentation based on sustainable shared value using BWM and PROMETHEE. *Sustainability, 15(*11), 8670. https://doi.org/10.3390/su15118670

Tasnim, N., Tahmid, M., Jahan, N., & Syeda, S. R. (2023). Risk assessment framework for selecting the safer route for hazmat transportation based on accident database and vulnerability models. *ACS Chemical Health and Safety, 30*(5), 302–317. https://doi.org/10.1021/acs.chas.3c00044

Valsamos, G., Larcher, M., & Casadei, F. (2021). Beirut explosion 2020: a case study for a large-scale urban blast simulation. *Safety Science*, 137, 105190. https://doi.org/10.1016/j.ssci.2021.105190

Wieteska, G. (2020). The impact of supplier involvement in product development on supply chain risks and supply chain resilience. *Operations and Supply Chain Management: An International Journal, 13*(4), 359–374. https://doi.org/10.31387/oscm0430276

Ziemba, P., & Gago, I. (2022). Uncertainty of preferences in the assessment of supply chain management systems using the PROMETHEE method. *Symmetry, 14*(5), 1043. https://doi.org/10.3390/sym14051043

Nunik Endah Sulistiyawati is a postgraduate student at the Interdisciplinary School of Management and Technology, Sepuluh Nopember Institute of Technology (ITS), Surabaya, Indonesia. Her research focuses on risk management in ammunition transportation and storage operations, with particular attention to multicriteria decision making for prioritizing risk mitigation strategies in hazardous materials logistics.

Bambang Syairudin is a lecturer in the Department of Industrial and Systems Engineering, Sepuluh Nopember Institute of Technology (ITS), Surabaya, Indonesia, and also teaches in the Management of Technology postgraduate program at ITS. His research interests include industrial management, knowledge management, and industrial engineering. He holds the IPU (Insinyur Profesional Utama) and ASEAN Engineer professional titles.

Buana Ma'ruf is a researcher at the Research Center for Hydrodynamics Technology, National Research and Innovation Agency (BRIN), Surabaya, Indonesia. He previously served as a researcher at the Agency for the Assessment and Application of Technology (BPPT), which has since been integrated into BRIN. His research interests include ship and marine technology, ship production and management, and strategic management in the manufacturing industry. He is a Fellow of the Royal Institution of Naval Architects (FRINA), a Chartered Engineer (CEng), and holds the IPU professional engineer title.