Flow-Constrained Risk Management for OT Security: A Behaviourally Informed Framework for Critical Infrastructure Under Data Scarcity

A capacity-aligned, behaviourally informed risk-management model for improving OT security when data, budget, and evaluative capacity are limited.

Vsevolod Shabad

Department of Computer Science, University of Liverpool, Liverpool L69 3DR, United Kingdom

Corresponding author: Vsevolod Shabad, Department of Computer Science, University of Liverpool, Liverpool L69 3DR, United Kingdom.

Email: v.shabad@liverpool.ac.uk

ORCID: https://orcid.org/0009-0001-9332-6688

Date of this version: December 2025. This version supersedes the August 2025 working paper.

SSRN Working Paper — Not Peer Reviewed

This manuscript is an updated and extended working version submitted to the SSRN eLibrary. It is concurrently under submission to a peer-reviewed venue (IJCIP). Please cite the SSRN version unless a later version is posted.

Abstract

Critical-infrastructure operators face intensifying cyber threats while working under severe resource constraints, persistent shortages of reliable incident data, and governance processes that lag behind contemporary attack dynamics. Conventional security frameworks, including NIST SP 800-82, often assume parallel implementation of multiple controls and access to statistically meaningful incident datasets — assumptions that rarely hold in operational technology (OT) environments. Decision-making is further shaped by bounded rationality: incomplete information, limited evaluative capacity, and strong disincentives for experimentation.

This paper presents a flow-constrained risk-management (FC-RM) framework that aligns security improvement with actual organisational delivery capacity. The framework integrates (i) explicit work-in-progress limits for cybersecurity portfolio management; (ii) a progressive method for refining Annual Loss Exposure (ALE) under data scarcity; (iii) behavioural-bias

mitigation mechanisms embedded directly into prioritisation and decision processes; and (iv) empirical control-prevention measurement as an alternative to distribution-based risk inputs.

A composite scenario — triangulated from recurring patterns in peer-reviewed OT studies, practitioner literature, and industry/ regulator analyses — illustrates how FC-RM enables measurable, board-visible progress within constrained OT settings while avoiding the failure modes typical of traditional multi-year programmes.

The framework provides a practicable route for organisations lacking the data, budget, or evaluative capacity required by conventional quantitative risk methods. It also offers a basis for future empirical validation across OT sectors where regulatory expectations increasingly emphasise demonstrable improvement rather than policy conformance alone.

Version notes

- *v2 (December 2025):* Major revision. Expanded behavioural-bias integration; complete rewrite of composite scenario; formalised attack-attribution method; added mathematical validation and comparative framework analysis.
- *v1 (August 2025):* Initial working paper.

Keywords: operational technology security; critical infrastructure protection; risk modelling; constrained implementation; empirical data limitations; flow-constrained risk management

1 Introduction

Critical-infrastructure operators face an unprecedented escalation in cybersecurity threats while operating under severe resource constraints that traditional security frameworks struggle to address. State-sponsored Advanced Persistent Threat (APT) groups continue to intensify targeting of industrial organisations, with ransomware attacks rising by 87% in 2024 [1]. Recent joint guidance from the United States and allied cyber authorities highlights further escalation, warning that Chinese state-sponsored actors are actively compromising telecommunications, government, transportation and other critical infrastructure networks by exploiting edge-device vulnerabilities, modifying network configurations, and abusing built-in management functions to maintain long-term covert access [2]. This shift towards stealthy persistence across multiple sectors increases pressure on operators already constrained by limited security budgets and ageing architectures. Long-term underfunding further reduces defensive capacity, with HM Treasury acknowledging “years of chronic underinvestment” in UK infrastructure systems [3].

Existing operational-technology (OT) security frameworks — including ISA/IEC 62443, NIST SP 800-82r3 and NIST SP 800-39 — provide comprehensive technical guidance but do not fully address the implementation barriers that commonly arise in resource-constrained environments. Although they outline what “good” security looks like, they do not reflect the practical conditions many operators face, often expecting access to dedicated security

budgets, representative historical data for quantitative assessment, and sufficient capacity to apply multiple controls across large asset groups within the same implementation period. Recent empirical research with 72 OT practitioners shows that implementation gaps persist regardless of framework quality, with stakeholder priorities, operational values and workload pressures creating systematic obstacles to consistent security improvement [4].

This research addresses the operational and organisational constraints that combine to create recurring failure patterns in OT security implementation. Multiple factors — including workload pressures, entrenched operational values, legacy design assumptions and limited cybersecurity training — reinforce one another and limit the consistency of control deployment, irrespective of the formal framework adopted. Existing literature often discusses these challenges separately, treating them as independent problems to be addressed sequentially through “continuous monitoring”, “regular risk assessments”, or “incident response planning”, without recognising how their interaction constrains effective implementation [5].

1.1 Research Contributions

This work makes four contributions to operational technology security and risk-management practice by integrating established concepts into a unified framework tailored for resource-constrained critical-infrastructure environments.

1. **Integration of flow-constraint optimisation with cybersecurity risk management.** Prior work has applied lean manufacturing and Kanban principles to software development [6] and project management [7], whilst cybersecurity risk frameworks assume unconstrained deployment capacity [5, 8]. This research adapts work-in-progress (WIP) limitation principles to OT (operational technology) security portfolio management for resource-constrained critical infrastructure environments. Literature review indicates limited prior work linking Kanban-style WIP limits to OT security portfolio decision-making.
2. **Progressive Annual Loss Exposure methodology (ALE) addressing systematic data scarcity**. Whilst ALE calculations assume the availability of historical incident data [9], cybersecurity economics research documents systematic data scarcity in OT environments, where representative incident datasets don't exist [10, 11]. This research sets out a three-phase progressive refinement methodology that transitions from external validation (insurance estimates) through penetration testing to empirical control measurement, explicitly designed to function under data scarcity. The contribution lies in structuring these established estimation concepts into a progressive update mechanism aligned with OT operational realities.
3. **Explicit integration of behavioural bias mitigation into security framework design.** Behavioural economics documents cognitive biases that affect security investment [12-14], yet OT security frameworks assume rational stakeholder decision-making. This research embeds bias mitigation as a core design principle

through power-of-two voting, collaborative prioritisation, and progressive validation mechanisms. The framework architecture addresses the planning fallacy, optimism bias, not-invented-here syndrome, and anchoring bias. The contribution is the deliberate embedding of bias-mitigating design choices into a practical OT improvement workflow.

4. **Attack attribution methodology for empirical Annual Risk Occurrence (ARO) calculation in OT environments**. Security control effectiveness measurement typically uses proxy metrics (vulnerability counts, patching rates) or external threat intelligence (e.g., the CISA Known Exploited Vulnerabilities catalogue). This research sets out an attribution approach for estimating empirical ARO based on control-prevented attacks using distinct control activation records . The contribution is a practical attribution approach suited to environments where complete incident data is unavailable, but control-performance evidence exists.

These contributions draw on established concepts from lean operations, cybersecurity economics, and behavioural decision-making, but adapt and integrate them into a single OT-centric framework designed for constrained operational environments.

1.2 Article Structure

Section 2 reviews existing OT security frameworks, the evolution of the threat landscape, and the implementation barriers documented in the literature. Section 3 describes the framework design methodology combining cross-domain literature analysis, mathematical validation, and iterative refinement. Section 4 provides a detailed framework specification, including flow-constrained portfolio management, progressive ALE estimation, and stakeholder engagement processes. Section 5 presents mathematical validation and comparative framework analysis. Section 6 demonstrates the application's feasibility through a composite scenario constructed from recurring real-world patterns and triangulated with peer-reviewed empirical literature. Section 7 discusses practical implications, theoretical contributions, and limitations. Section 8 concludes with a research synthesis and future directions.

2 Background and Related Work

2.1 Existing OT Security Frameworks and Implementation Gaps

Current OT security frameworks provide comprehensive technical guidance but do not fully reflect the implementation constraints many operators face. **ISA/IEC 62443** establishes *security levels* (SL-1 through SL-4) and links their selection to risk assessments [15], yet it does not account for the limited time, staffing and system knowledge that often restrict the depth of such assessments in practice. **NIST SP 800-82r3** acknowledges critical OT constraints — including the need for careful planning of maintenance windows and the risk of operational disruption [5] — but offers limited guidance on how operators should prioritise or phase controls when only a subset can be applied during scheduled outages. **NIST SP 800-39** provides a structured approach for managing information security risk across

organisational tiers [8]. Still, because it focuses on governance and risk framing, it does not specify how strategic risk decisions should translate into phased control implementation under real-world OT operating conditions, such as limited maintenance windows or competing production priorities.

Table 1. Gap Analysis — Framework Assumptions vs. Organisational Realities

Assumed Capability	Actual Constraint (Critical Infrastructure)	Specific Gap Framework Addresses
Implement 90+ security controls across all OT assets even for LOW baseline [5, table 22]	Silva and Bobbert [16]: organisations struggle to deploy controls due to low prioritisation, unclear ownership, limited resources, and ICS-specific skill gaps; 60% cite lack of priority/stakeholder involvement	WIP limits match implementation pace to real organisational capacity and prevent overloaded transformation backlogs
Conduct comprehensive risk assessment, achieve target SL across all zones	Silva and Bobbert [16]: 50% rate current assessment/audit methods as ineffective; barriers include complex standards, insufficient ICS expertise, and difficulty performing full assessments	Progressive ALE enables more frequent, incremental assessment cycles, avoiding dependence on large, resource-intensive assessments
Access historical incident data for quantitative risk assessment	Cremer et al. [10]: limited, incomplete, and under-reported cyber-risk datasets; “lacuna in open datasets” limits quantitative modelling Silva & Bobbert [16]: lack of ICS incident data and structured knowledge	Progressive refinement: uses insurance estimates, penetration-test and harm modelling scenarios, and empirical control-efficiency measurement to approximate risk without historical OT dataset
Rational, unbiased stakeholder decision-making	Silva and Bobbert [16]: lack of priority, unclear ownership, poor ICS understanding, and insufficient stakeholder involvement inhibit	Bias-mitigation techniques such as collaborative co-authoring, structured voting, and external validation help reduce individual judgement

Assumed Capability	Actual Constraint (Critical Infrastructure)	Specific Gap Framework Addresses
	consistent, rational security decisions	errors and improve decision quality
Straightforward strategic approval for large-scale security programmes	Silva and Bobbert [16]: security initiatives require leadership buy-in, improved reporting, and clearer strategic alignment; lack of these factors reduces prioritisation of ICS security	Operational integration: breaks major transformations into quarterly increments, making improvements fundable, observable, and compatible with constrained approval cycles

Recent empirical research [4] based on interviews with 72 OT practitioners shows that cybersecurity work often needs to be scheduled around established operational priorities — particularly availability, production continuity and safety — because safety-critical systems require stable, predictable operation. Although cybersecurity ultimately reinforces long-term safety, these operational constraints often limit the time and flexibility available to implement security controls. The study indicates that the main implementation barriers stem from organisational and mindset factors rather than technical shortcomings. Workload pressures, legacy design assumptions, unclear ownership and limited cybersecurity training often undermine consistent control deployment, even when technical requirements are understood. As a result, organisational resistance and long-standing operational value systems continue to impede the adoption of recommended security practices, suggesting that technically comprehensive frameworks alone cannot overcome these constraints without complementary changes in organisational practice and prioritisation.

Forthcoming regulatory developments in the United Kingdom, reflected in the Cyber Security and Resilience Bill [17], indicate movement toward NIS2-aligned obligations and strengthened oversight for operators of essential services. These changes reinforce the need for practical methods that enable demonstrable security improvement under operational constraints.

2.2 Evolving Threat Landscape: Quantitative Evidence

Critical infrastructure operators face a rapid escalation in cyber threat volume and sophistication. Dragos reports ransomware attacks against industrial organisations rose 87% in 2024 [1]. The UK NCSC Annual Review 2025 highlights persistent blind spots in incident visibility, noting that voluntary breach reporting means national data "is not a true reflection" of cyber incidents [18]. Attackers can rapidly exploit discrete technical weaknesses, whereas defensive improvements require detailed system understanding, creating a cost–complexity imbalance favouring threat actors.

CISA's Known Exploited Vulnerabilities (KEV) catalogue shows the interval between vulnerability disclosure and active exploitation has compressed significantly, with ICS vulnerabilities appearing in threat intelligence days or weeks after disclosure [19]. This creates a temporal mismatch: exploitation cycles evolve quickly, whereas security improvements in critical-infrastructure organisations move through multi-stage approval processes spanning many months.

Independent analyses note long-term underinvestment in critical infrastructure [20], meaning even technically sound frameworks must operate within environments where resource scarcity, legacy constraints and deferred capital programmes limit defensive measures.

Operational-technology environments operate under conditions of bounded rationality [21, 22]. Decision-makers often face incomplete information, limited evaluative capacity and high penalties for operational failure, encouraging satisficing rather than comprehensive evaluation. Characteristic cognitive biases — optimism bias, anchoring, and normalcy bias — emerge as predictable responses to these structural constraints. Combined with broad control catalogues and detailed system knowledge requirements, these limits reduce realistic decision-making capacity. A sequenced, capacity-aligned improvement process therefore better reflects both the operational, behavioural, and informational constraints that shape real-world cyber-risk decisions.

2.3 Data Scarcity and Behavioural Constraints in OT Risk Assessment

The cyber-risk literature consistently highlights the absence of reliable data for quantitative OT risk assessment. Cook et al. characterise ICS cyber-attacks as high-impact, low-frequency events with volumes too low for statistically robust estimation [11]. Cremer et al. [10] identify persistent dataset scarcity, noting a 'lacuna in open databases' limiting quantitative assessment robustness. Morris [23] demonstrates that even purpose-built academic ICS datasets contain artefacts causing models to fail in real-world environments — exemplifying "garbage in, garbage out" risks, as described by Canbek et al. [24]. Anderson and Moore highlight the need for stronger empirical foundations, noting the lack of robust incident and loss data [25]. The empirical foundations for traditional risk quantification are weakest in the OT contexts where accurate modelling would provide the greatest value.

These data limitations amplify bounded-rationality conditions. Behavioural economics research demonstrates cognitive biases affecting investment decisions under uncertainty. Kahneman and Tversky's prospect theory shows decision-makers systematically deviate from rational choice when facing uncertain losses [12]. Korteling et al. [13] show executives systematically discount future risks, favour status quo options, and rely on external validation under uncertainty.

Critical infrastructure organisations exhibit distinctive patterns creating systematic resistance to security investment: **normalcy bias** (underestimating adverse event likelihood); **availability heuristic** (recalling operational disruptions but struggling with cyber probabilities lacking visible precedents); **anchoring bias** (initial estimates unduly influencing subsequent decisions); **overconfidence bias** (overestimating defensive capabilities relative to APTs); and **not-invented-here syndrome** (rejecting externally originated proposals as operational interference) [26]. These biases function as practical expressions of bounded rationality, where incomplete information, operational pressures and limited evaluative capacity shape stakeholder judgment.

Salzberger [14] identifies optimistic bias as a systematic factor in security investment, demonstrating that executives underestimate cyber risk likelihood whilst overestimating defensive capabilities. Even when external validation confirms threat reality, executives seek security resource reductions for "operational efficiency". This creates requirements for sustained demonstration of improvement progress through visible, measurable trends, maintaining stakeholder confidence and resisting budget erosion.

Quantitative cyber-risk models such as FAIR [27] presuppose statistically meaningful incident-frequency and loss-magnitude data. OT environments lack these datasets due to chronic under-reporting, low-frequency events and confidentiality restrictions [10, 11, 23]. As a result, FAIR cannot be applied directly in OT settings because its required inputs do not exist. This limitation motivates approaches that operate under data scarcity; the progressive ALE method developed in this work provides a structured refinement process. FC-RM extends this by substituting distribution-based inputs with progressive ALE refinement and empirical control-prevention evidence, enabling structured prioritisation even when historical incident data are unavailable.

2.4 Organisational Decision Constraints and Temporal Mismatch

Cybersecurity investment decisions are shaped by governance processes moving more slowly than contemporary threat activity. Multi-stage planning cycles — business-case development, stakeholder consultation, budget approval — span months, creating a *temporal mismatch* when exploitation unfolds over days or weeks.

Classical multi-criteria decision-making assumes decision-makers can provide complete evaluations across all alternatives [28]. In practice, complex choice sets overwhelm evaluators: *choice overload* reduces decision completion and increases avoidance [29]. As cybersecurity options proliferate, stakeholder evaluations become slower and less decisive.

Risk-based governance reinforces these dynamics. Under NIST SP 800-39, improvements must align with the organisation's *risk frame* [8]. leading to deferred mitigation when actions fall outside prioritised objectives. These factors produce a consistent lag between attacker adaptation and defensive change execution.

2.5 Research Gap: Need for Integrated Implementation-Focused Approach

Existing literature treats implementation challenges as separate problems: manufacturing research addresses resource constraints through lean principles [6, 30]; cybersecurity economics documents data limitations [10, 11]; behavioural economics reveals cognitive biases [12, 13]; operations research shows high WIP reduces responsiveness [31]. However, current approaches assume organisations can address gaps sequentially through "continuous monitoring" and "regular risk assessments" without acknowledging how challenges interact.

Freeman's stakeholder theory provides foundations for engaging groups with competing priorities [32], but does not model time and cognitive constraints. The Scaled Agile Framework introduces portfolio-flow practices, including WIP limits [7], yet these remain largely absent from cybersecurity governance, where extended planning cycles predominate.

Together, these challenges create the practical conditions of bounded rationality in OT environments, where resource, data, temporal and cognitive constraints shape risk decisions in interdependent rather than isolated ways.

The fundamental research gap

The existing literature lacks integrated frameworks that simultaneously address resource constraints, data limitations, behavioural biases, and agility requirements in operational technology security contexts. The need for integrated solutions acknowledging these gaps as interconnected constraints requiring simultaneous rather than sequential treatment represents the research gap this work addresses.

3 Methodology

3.1 Design Science Research Approach

This research employs design science methodology [33] to develop and validate an FC-RM framework that addresses implementation barriers to effective OT security enhancement. The approach emerged from practical necessity: the author, operating within critical infrastructure environments, faced urgent requirements to strengthen OT security against active nation-state threat actors whilst constrained by limited resources and funding allocations that made traditional strategic investment approaches infeasible.

Initial attempts to apply established security frameworks proved inadequate for addressing the simultaneous pressures of sophisticated attackers and severe resource limitations. Traditional approaches that implicitly assume comprehensive resourcing and extended implementation timelines were misaligned with operational realities in OT environments. This practical challenge necessitated developing an integrated approach to achieve systematic

security improvement within existing constraints, whilst providing demonstrable progress against documented threats.

Research Question

How can critical infrastructure operators achieve systematic OT security improvements within existing resource constraints, whilst providing measurable progress for executive oversight?

3.2 Cross-Domain Literature Analysis

A cross-domain literature analysis was conducted across five complementary bodies of work to identify implementation barriers and inform potential solution approaches:

1. **Operational technology security:** Technical foundations and recommended controls (ISA/IEC 62443, NIST SP 800-82r3) and documented implementation challenges in OT environments [4].
2. **Risk management:** Quantitative assessment approaches [34, 35] and persistent data-uncertainty challenges in OT cyber-risk estimation [10, 11].
3. **Behavioural economics:** Cognitive and decision-making biases affecting security investment decisions under uncertainty [12-14].
4. **Lean manufacturing and agile methodologies:** Resource-constrained flow principles [6, 7, 31] and queueing-theory insights showing that high WIP and utilisation increase delays and reduce responsiveness [31].
5. **Stakeholder management theory:** Collaborative decision-making frameworks for engaging diverse organisational groups (board, operations, safety, IT) [32, 36].

Analysis revealed that the existing literature typically treats implementation challenges as independent, isolated problems, whereas their interactions create reinforcing, systemic failure patterns. Four critical constraints requiring simultaneous treatment were identified:

1. Resource limitations prevent comprehensive control deployment.
2. Data scarcity that undermines quantitative risk assessment.
3. Psychological and behavioural constraints arising from bounded rationality that impede security-investment decisions.
4. Temporal mismatches between extended planning cycles and rapidly evolving threat environments.

These findings informed the design of a framework oriented toward integrated, incremental security improvement under real-world operational constraints.

3.3 Framework Component Development and Integration

The four critical constraints identified through cross-domain analysis were translated into specific design requirements using a structured mapping process:

Design Requirements:

- **Resource limitations:** Components must function within existing organisational capacity and constrained implementation bandwidth.
- **Data scarcity:** Approaches must operate effectively without reliance on comprehensive or high-quality historical OT incident datasets.
- **Psychological and behavioural constraints arising from bounded rationality**: Methods must support more precise evaluation, manageable decision load, and reduce resistance to security initiatives, addressing the reduction of evaluative capacity characteristic of bounded-rationality conditions.
- **Temporal mismatches:** Approaches must support incremental, responsive change while maintaining alignment with established strategic oversight processes.

Component Selection Criteria

1. Multi-constraint capability – Components must address multiple interacting barriers simultaneously, rather than treating them as independent problems.
2. Implementation feasibility – Each component must be achievable within typical resource, time and skills constraints found in OT environments.
3. Organisational acceptance – Components must be compatible with existing governance structures and decision-making processes to support adoption.
4. Measurable progress demonstration – Each component must provide visible, incremental evidence of security improvement to support executive oversight and counter bounded-rationality effects and associated psychological biases.
5. Theoretical rigour – Components must draw on established analytical, behavioural, or operational foundations to ensure conceptual soundness.

Selected approaches were adapted to account for differences between source domains and OT security contexts. This adaptation involved analysing the theoretical foundation, mapping each constraint to its operational implications, and modifying candidate techniques to address identified gaps. Preliminary validation was conducted through expert review to ensure conceptual relevance, implementation feasibility, and alignment with real-world OT security constraints.

3.4 Mathematical Foundation Selection and Validation

The mathematical approach selection evaluated frameworks for analytical capability, practical applicability, and integration potential. The research design combined:

- **Queuing theory principles**: Qualitative insights on WIP-lead time relationships informing governance constraints.
- **Multi-criteria decision analysis**: Power-of-two voting for improvements prioritisation.
- **Probability theory**: Progressive ALE refinement under data uncertainty.

The compatibility assessment evaluated the theoretical coherence among different mathematical approaches by formal analysis of their underlying assumptions, parameter requirements, and output interpretation methods. Mathematical validation included convergence analysis demonstrating theoretical validity under data-uncertainty conditions, sensitivity analysis examining robustness across parameter variations, and peer review by mathematical experts.

3.5 Validation Through Composite Scenario

Direct empirical validation through disclosed organisational case studies was not feasible due to confidentiality, contractual obligations, and national security restrictions that are common in OT environments. To address this, the framework was evaluated using a composite scenario constructed from well-documented operational patterns and triangulated with peer-reviewed empirical studies reporting similar organisational constraints across critical infrastructure sectors. This method demonstrates practical feasibility without revealing sensitive system information or operational data.

Practitioner Provenance Statement

The framework has been applied in the author's professional OT security practice (a large multinational mining, metallurgy, and energy conglomerate and critical infrastructure operator). Although specific quantitative results cannot be disclosed, practitioner observations consistently matched the framework's intended effects: stabilised improvement backlogs, interruption of recurring risk-decay cycles, strengthened stakeholder alignment, and sustained progress under tight resource and scheduling constraints. These observations are included solely for contextual orientation; empirical validation relies on the published research incorporated into the composite scenario.

To comply with contractual confidentiality obligations and national-security restrictions applicable to critical infrastructure operators, no client-specific operational data is disclosed in this manuscript. All quantitative parameters used in the composite scenario — including budget ranges, staffing levels, CVE counts, maintenance-window durations, ALE values, and throughput metrics — are presented as anonymised values bounded by the 25th–75th percentile ranges reported in the peer-reviewed OT/ICS literature. These values therefore represent realistic sectoral mid-ranges rather than invented figures, while ensuring that no sensitive or proprietary operational measurements are revealed.

Composite Scenario Construction

The composite scenario integrates four categories of independently sourced evidence:

1. **SCADA/OT system characteristics.** Derived from vendor documentation and established industry standards, defining realistic architectural boundaries, maintenance-window constraints, and operational dependencies.
2. **Threat and attack patterns.** Informed by MITRE ATT&CK for ICS [37], CISA advisories, Dragos OT threat intelligence, and recent empirical analyses [38], ensuring representation of contemporary adversary behaviours and exploitation timelines.
3. **Cost and impact estimation.** Informed by insurance-industry cyber-loss surveys and peer-reviewed impact analyses, providing plausible ranges for incident consequences and investment trade-offs.
4. **Organisational constraints.** Triangulated from Evripidou and Watson's qualitative study of OT practitioners [4] and wider empirical evidence documenting resource limitations, maintenance-scheduling restrictions, safety-security tensions, and implementation challenges [39-45].

Grounding scenario components in independently verified empirical sources ensures methodological transparency, while practitioner insight is used only for contextual refinement rather than as evidentiary support.

Limitations

The composite scenario demonstrates practical feasibility but cannot substitute for cross-organisational empirical studies using anonymised datasets. Future research should pursue structured pilots or comparative case studies — where confidentiality permits — to quantify the framework's performance and evaluate generalisability across sectors.

3.6 Research Limitations

This research acknowledges systematic limitations:

1. **Generalisation limitations**: The framework was contextualised using practitioner insight from a limited set of OT security environments. Although the core design is grounded in published empirical research, further validation is needed to assess its applicability across the full range of critical infrastructure domains.
2. **Lack of a controlled comparison**: No control group or parallel implementation of alternative approaches was available, preventing a direct comparative evaluation with traditional security-improvement methods.
3. **Qualitative validation evidence.** Validation relied primarily on qualitative indicators of progress (e.g., stabilised backlogs, reduced risk-decay patterns, stakeholder alignment) rather than standardised quantitative performance metrics.

4. **Single organisational provenance.** Practitioner observations originate from a single organisational environment and may not reflect the diversity of cultures, regulatory regimes, or resource constraints across critical infrastructure sectors.
5. **Researcher positionality.** The author's dual role as practitioner and researcher introduces potential bias. Reflexive documentation was used to mitigate this risk, but it cannot be entirely eliminated.

These limitations indicate that the framework should be regarded as a **promising, evidence-informed approach requiring broader empirical evaluation**, rather than a fully validated methodology for all critical-infrastructure contexts. Section 6.5 provides additional limitations specific to the composite scenario used for conceptual validation.

3.7 Evidenced Assumptions Underpinning the Composite Scenario

Each assumption used in the scenario is explicitly grounded in published findings or qualitative evidence consistently reported across multiple independent studies.

1. **Legacy OT/SCADA Components**: Studies across energy, manufacturing, and oil-and-gas consistently report substantial proportions of legacy equipment with outdated protocols and long-life-cycle hardware [46, 47]. Shewale [48] cites a 2021 survey indicating **61% of factories relied on outdated operating systems** (e.g., Windows XP/7) in OT environments.
2. **OT Cybersecurity Staffing**: Peer-reviewed literature highlights the scarcity of dedicated OT cybersecurity personnel and the difficulty of recruiting domain specialists [49, 50]. Professional guidance describes configurations where small teams support regional or multi-site operations [51].
3. **OT-Focused Budget Levels**: Empirical studies report OT-directed security budgets typically represent a minority fraction of cybersecurity expenditure, often **below 1% of operational expenditure** [49, 50].
4. **Maintenance Window Constraints**: Empirical studies in industrial reliability engineering describe strict maintenance-window regimes driven by safety, production, and process-continuity requirements [40, 41]. Upgrades are commonly tied to infrequent scheduled shutdowns, often months apart.
5. **Security Control Implementation Throughput**: Published research provides qualitative descriptions of slow control-deployment rates due to coordination costs, limited personnel capacity, and operational constraints [46, 47]. Quantitative throughput benchmarks are generally absent, but case studies consistently show incremental progress shaped by maintenance cycles and staffing limitations.
6. **Risk Quantification Constructs**: Studies on OT/ICS risk modelling increasingly use quantitative constructs such as loss expectancy and scenario-based impact assessments [41, 49, 52]. Combining insurance-based loss data with scenario modelling and external testing is a recognised approach.

These assumptions are evidence-constrained: they reflect convergent findings across independent sources while avoiding numerical assertions lacking empirical grounding.

Table 2. Empirical triangulation of scenario patterns

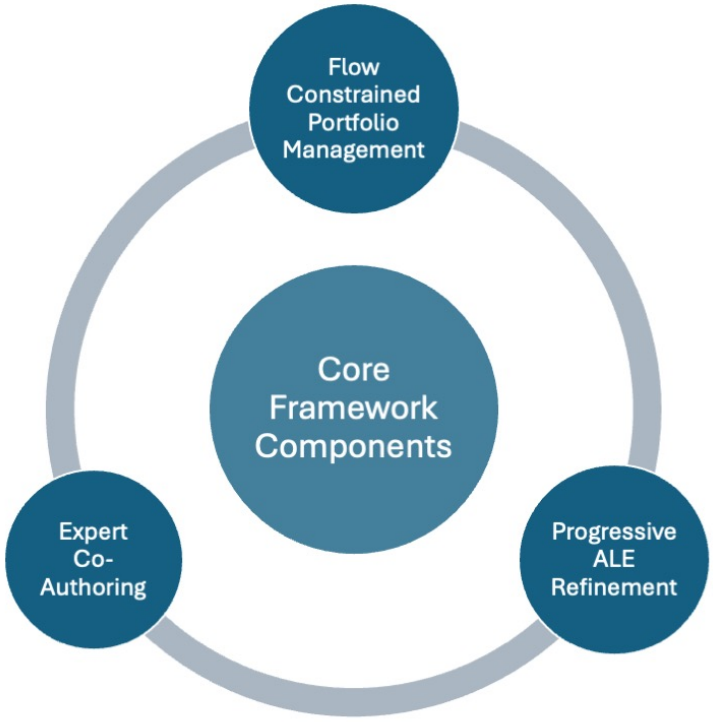
Scenario Pattern	Empirical Sources	Convergence assessment
Legacy OT/SCADA footprint	Anton et al. [38]; Badawy [46]; Chan & Zhou [47]; Shewale [48]	Independent studies consistently report high proportions of legacy systems and outdated OS deployments
Resource constraints (2–4 FTE; <1% OpEx)	Falco et al. [49]; Moon et al. [39]; Wai & Lee [50]; Ribeiro [51]	Multiple studies describe limited OT-security staffing and budgets, matching the scenario's mid-range values
Maintenance-window limitations	Cavalcante et al. [40]; Zhang & Yang [41]	Industrial reliability literature consistently documents tightly scheduled, infrequent maintenance periods
Safety–security tensions	Zhou et al. [42]; Agbo & Mehrpouyan [43]	Studies show operational and safety priorities frequently constrain cybersecurity change-windows
Framework–practice gaps	Evripidou & Watson [4]; Silva & Bobbert [16]; Sheikhi et al. [44]; Heinl et al. [45]	Empirical work consistently observes stakeholder, workload, and governance barriers to OT-security implementation

4 Framework Specification

4.1 Architectural Overview

The framework integrates three core components addressing fundamental implementation barriers.

Figure 1. Core Framework Components



1. Flow-Constrained Portfolio Management

Governs resource allocation and improvement selection, applying manufacturing optimisation principles [6] to prevent resource dilution whilst maintaining systematic progress. Constrains strategic security improvements to:

- **One major improvement per quarter per asset** (completing within quarter)
- **Maximum three minor improvements in parallel** (each completing within one month)

The '1 major + 3 minor' limits are calibrated governance heuristics derived from the empirical constraints documented in the OT literature (staffing 2–4 FTE, infrequent maintenance windows, and multi-stage approval cycles). They are not statistical optima, but

capacity-aligned boundaries designed to prevent overload under realistic implementation conditions.

2. Progressive ALE Refinement

Provides increasingly accurate risk prioritisation without requiring comprehensive historical datasets, transitioning through three phases:

- **Phase 1:** External validation using insurance survey worst-case scenarios.
- **Phase 2:** Penetration testing validation with realistic scenario modelling.
- **Phase 3**: Empirical measurement based on operational security control performance.

3. Multi-Expert Co-Authoring Process

Addresses systematic behavioural biases through structured collaborative decision-making adapted from Hohmann's Innovation Games methodology [36], ensuring operational personnel feel ownership of security decisions rather than compliance with headquarters mandates.

Although economic models for optimal security investment are well established, Gordon and Loeb [34] show that organisations struggle to implement them in practice because reliable estimates of breach probabilities and losses are unavailable. Rather than assuming organisations can overcome barriers through better economic analysis, the framework **operates within realistic constraints whilst building credibility through demonstrated results**.

Unlike conventional ALE approaches used in cyber-insurance modelling, NIST SP 800-30, or academic risk-estimation studies, this progressive ALE method is designed to operate under data scarcity and bounded-rationality constraints, converging towards accuracy through staged refinement rather than relying on complete historical datasets.

4.2 Flow-Constrained Portfolio Management

Building on Kanban WIP limitation principles [6, 30], the framework adapts manufacturing flow optimisation to OT security, addressing the *Planning Fallacy***,** in which organisations systematically underestimate task complexity [12]. The framework governs strategic security improvements (architecture implementations, control deployments, process redesigns), excluding routine operations (daily monitoring, standard patching, incident response).

Two-Layer Prioritisation: Enterprise-wide asset prioritisation selects **one asset per quarter** based on the highest ALE ranking. Within the selected asset, constraints apply: **one major improvement per quarter** (>1 month duration) plus a **maximum of three minor improvements in parallel** (each <1 month), enabling up to 10 discrete improvements per quarter whilst preventing resource overload.

Design Rationale: The WIP constraints are inspired by queuing theory principles [31]. Little's Law ($L = \lambda W$) demonstrates that in stable systems, lead time increases with WIP. Although cybersecurity environments lack steady-state conditions, the underlying insight remains valuable: **constraining concurrent initiatives reduces context-switching overhead and improves completion predictability**, even when strict mathematical assumptions do not hold.

WIP Enforcement: Monthly expert group reviews monitor major improvement advancement, assess completed minors, select up to 3 new initiatives (maintaining WIP limit), and adapt priorities based on operational learning and changing threats.

4.3 Progressive ALE Estimation Addressing Executive Credibility Challenges

The framework uses **Annual Loss Exposure (ALE)** as the primary metric for asset prioritisation and progress measurement:

$$ALE = ARO \times \text{SLE}$$

where:

- **ARO (Annualised Rate of Occurrence)** = Estimated attack frequency.
- **SLE (Single Loss Expectancy)** = Impact per incident.

Traditional ALE-style risk quantification is often impractical in OT environments because data scarcity and bounded-rationality constraints on executive decision-making create predictable limitations in how senior decision-makers interpret internally produced risk estimates. Research on cognitive biases [13] shows that decision-makers systematically discount future risks, underestimate low-frequency but high-impact events, and prefer maintaining existing operational patterns under uncertainty (normalcy bias, optimism bias, availability heuristic, status-quo bias). These biases reduce confidence in internally generated estimates and increase reliance on externally validated evidence, consistent with authority bias and social proof dynamics.

Rather than relying on static, one-off risk assessments, the framework employs a **progressive ALE methodology** where estimation accuracy improves through staged external validation and operational measurement.

Phase 1: External Validation Foundation

- **SLE Source**: Site-specific insurance risk-engineering reports for the relevant facilities (e.g., mining site, metallurgical plant, power generation unit). These reports incorporate local asset values, production dependencies, historical equipment failures, and worst-case operational disruption scenarios.

- **ARO Baseline**: Conservative initial values (e.g., 0.1–0.5 per year) applied uniformly across all enterprise assets, reflecting the high-impact, low-frequency nature of OT incidents and providing a comparable starting point before empirical refinement.
- **Purpose**: Establish a credible initial baseline grounded on *externally* validated, formally documented estimates that executives already trust as part of regular corporate risk-financing processes.
- **Limitation**: Insurance assessments often model broad operational disruptions (fires, natural disasters, equipment failures) rather than cyber-specific causal chains, leading to pessimistic SLE estimates for OT security-related events.

Phase 2: Penetration Testing Validation

- **Reality Testing**: External penetration testing provides concrete, independently verified evidence of exploitable vulnerabilities and attack paths.
- **Scenario Refinement**: Operational-impact modelling is updated based on the demonstrated technical attack vectors (e.g., ransomware deployment via SCADA misconfiguration → 7–10 days recovery). *These numerical examples are synthetic and intended to demonstrate modelling logic rather than report empirical incident values.*
- **Executive Confidence:** External validation from an independent security firm provides impartial, third-party evidence of credible attack pathways, increasing decision-maker confidence under uncertainty.
- **Safety Consideration:** Penetration testing should focus on the management and signalling planes in a non-invasive manner, avoiding any actions that could affect physical processes; plausible operational impacts are derived through controlled scenario modelling rather than live system manipulation.

Phase 3: Empirical Performance Measurement

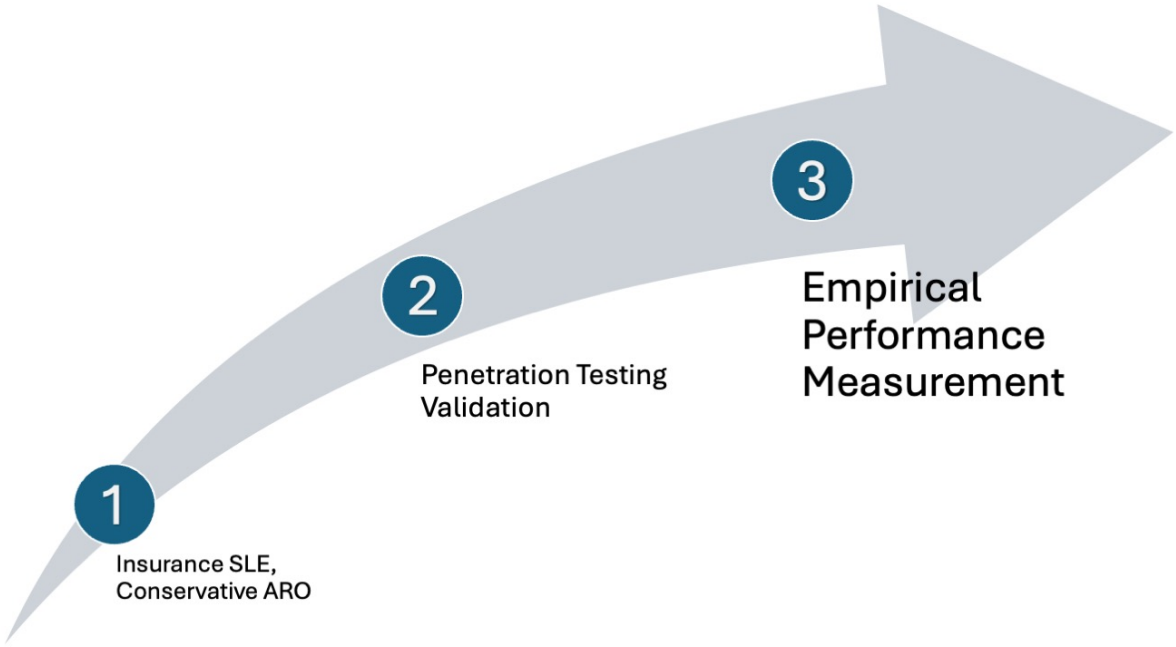
- **Control-Based ARO:** Estimate attempt frequency by tracking the number of verified attacks prevented by *non-redundant* security controls (e.g., intrusion detection blocks, firewall rule activations correlated with threat-intelligence indicators).
- **Operational Evidence:** Measurements reflect real, observed security-control activations that stopped attacks before causing harm.
- **Cultural Transformation:** The accumulation of verified, observable defensive successes provides internal credibility, supporting iterative investment and reducing scepticism driven by bounded-rationality constraints on executive decision-making.

Updated ARO Calculation (Phase 3):

$$\text{ARO} = \frac{\text{Real Incidents} \ + \ \text{Control} - \text{Prevented Attacks}}{\text{Measurement Period (years)}}$$

Control-Prevented Attacks represent verified attack attempts blocked by non-redundant security controls, meaning the organisation was effectively *one control failure away from a real operational-impact incident*.

Figure 2. Progressive ALE Refinement



Critical Limitation — ALE Reduction Not Guaranteed

Convergence is empirical rather than deterministic; estimates stabilise as more control-prevented events accumulate over a sufficiently long observation period. Whilst convergence toward measurement accuracy occurs, this does **not guarantee sustained ALE reduction**. Progressive refinement typically reveals:

1. Initial insurance-based SLE estimates are pessimistic (based on generic operational disruption scenarios).
2. Penetration testing provides more accurate, often lower, SLE (specific cyber-attack scenarios).
3. **Empirical measurement may reveal that actual attempt frequency is an order of magnitude higher than initial ARO assumption**s (e.g. tens of attempts per year rather than a fraction of one).

New vulnerabilities and threat actors may emerge faster than organisations can implement defensive improvements. These dynamics reflect bounded-rationality conditions: organisations can prioritise more accurately over time, but cannot expand implementation capacity or slow attacker evolution. The framework's value, therefore, lies in prioritising the highest-pressure risks at any given moment, rather than in promising an absolute ALE reduction across the entire asset base.

Board Reporting Strategy

Use trending rather than absolute ALE for executive communication:

- Report **control effectiveness percentages** (e.g., "controls implemented in Q2 blocked 37% of observed attack attempts").
- Provide **comparative analysis** (e.g. "Without Q1-Q2 improvements, site 'Alpha' would have experienced ~13 additional successful incidents annually").
- Emphasise **directional progress** (e.g. "ALE measurement accuracy increased through progressive validation; risk reduction evidenced through measured control performance").

Manual Adjustments for Non-Quantifiable Risks

Critical infrastructure sectors involve high-consequence risks in which safety, environmental harm, and societal trust cannot be reliably monetised [35].

Manual adjustment of SLE for such consequences is therefore necessary and must be transparently documented to avoid introducing subjective bias.

When resources are insufficient to mitigate all risks with significant non-quantifiable impacts simultaneously, the framework provides a **structured fallback** for empirical ALE-based prioritisation, ensuring systematic progress despite data limitations.

The progressive ALE refinement and the associated ARO attribution logic should be regarded as proposed measurement approaches that require multi-year observational validation to confirm their behaviour in real OT environments.

4.4 Empirical Security Control Performance Measurement

Unlike traditional risk-quantification approaches that require statistically meaningful loss-frequency data, the progressive ALE component provides a defensible update mechanism for data-scarce OT environments by leveraging empirical control-performance evidence and staged external validation.

Control-Based ARO Measurement Methodology

In OT environments where security controls often lack redundancy, each successful security control activation that prevents an attack represents empirical evidence of a genuine threat attempt. Methodology distinguishes between **security events** (often noise) and **attack attempts** (security-relevant) by counting control activations that prevented successful attacks rather than the volume of security events generated.

Attack Attribution Logic:

- **One attack campaign** = one ARO increment (regardless of events generated).
- **Multiple attack vectors** may represent separate increments **if different non-redundant controls stop each vector**.
- **Only attacks stopped by sole-capable controls** count toward the ARO measurement.

Example Application:

- Malware attempts lateral movement, that blocked simultaneously by endpoint detection AND network access controls = **redundant protection, no ARO increment.**
- Malware execution prevented by endpoint detection alone with no other capable controls = **non-redundant protection** (organisation "one control failure away" from incident) = **+1 ARO increment.**

ARO attribution in this framework relies primarily on observed control-performance evidence, but its accuracy increases when supported by external validation, such as penetration testing or red-team activity. Where such validation is unavailable, attribution remains approximate and is progressively refined over successive cycles.

Post-Implementation Validation

Following the security improvements, the methodology continues to measure control-prevented attacks to demonstrate actual risk reduction. Security improvements may reduce attack magnitude (better containment) or frequency (earlier detection):

- **Enhanced Control Redundancy**: Implementing outbound NACLs alongside endpoint detection creates defence-in-depth.
- **Attack Vector Mitigation**: Email sandboxing shifts detection from endpoint to gateway level, stopping attacks earlier in the kill chain.
- **Threat Sophistication Evolution**: Attackers resort to more sophisticated techniques, demonstrating that basic vectors have been eliminated.
- **Progress Measurement**:

$$ALE\ Reduction\ \% = \frac{Baseline\ ALE - Current\ ALE}{Baseline\ ALE} \times 100$$

SLE Refinement Through Operational Impact Modelling

Updated SLE calculations incorporate scenario analysis based on penetration testing results showing actual operational impacts specific to OT systems. When penetration testing successfully demonstrates control over critical operational systems, scenario modelling estimates:

- **Operational disruption duration** (time until human intervention or automatic safety systems activate).
- **Equipment damage costs**.
- **Production downtime costs**.
- **Potential safety-impact pathways** (e.g., loss of containment, hazardous energy release), **modelled qualitatively but not monetised**, in line with industry practice and regulatory expectations.

Asset Graduation Criteria

Graduation indicates that an asset has reached a point where its **remaining ALE** is no longer the highest in the portfolio, and where a curated external penetration test has confirmed that recently implemented controls prevent compromise of critical OT functions within the agreed scenario boundaries. Graduation **does not imply permanent completion**: if other assets improve more rapidly or if new vulnerabilities emerge, the graduated asset may again become the organisation's highest-risk OT system and will return to the front of the testing queue.

To minimise disruption to operational staff and remain within approved budget constraints, **external penetration testing is performed once per quarter**, targeting the **OT asset with the highest remaining ALE**. This provides an objective, repeatable prioritisation mechanism under resource limitations. Graduated assets remain under routine monitoring and vendor intelligence review, but they are not re-tested externally until their prioritisation score again exceeds that of other systems.

This approach ensures that scarce, high-quality external testing capacity is always applied to the **weakest OT asset at any given time**, while still allowing previously improved assets to re-enter the testing cycle as threat conditions evolve. It provides a pragmatic balance: demonstrable external validation of improvements, constrained testing frequency to protect safety and operational stability, and dynamic re-prioritisation that keeps attention aligned with real-world risk rather than static plans.

4.5 Expert Co-Authoring Process for OT Personnel Engagement

To surface operationally grounded improvements and reduce resistance from frontline OT personnel, the framework adapts Hohmann's *Buy the Feature* prioritisation exercise [36], with modifications derived from the author's experience applying similar techniques in enterprise and OT environments. The process ensures balanced participation by granting **equal voting weight** to all contributors, preventing hierarchical influence or dominance by more assertive experts.

Expert Group Composition

For each selected asset, the expert group consists of:

- 2 OT managers (SCADA systems, safety instrumentation).
- 1 industrial safety specialist.
- 1 plant/site manager (operational authority).
- 1 IT specialist (network architecture).
- 2 OT security specialists (headquarters and asset-level).

Equal voting weight ensures no participant — regardless of seniority, background, or confidence level — can outweigh the collective judgement of peers, creating a psychologically safe environment for cross-disciplinary prioritisation.

This co-authoring structure directly counteracts bounded-rationality pressures by reducing information asymmetry and distributing evaluative load across multiple experts.

Improvement Identification Process

1. **Brainstorming:** Inventory potential vulnerabilities and feasible security improvements (preventive, detective, reactive).
2. **Decomposition:** Break improvements into the smallest implementable components, classified as:
    a. *Minor* — feasible within one month.
    b. *Major* — requiring more than one month but deliverable within one quarter.
    c. *Excluded* — requiring more than one quarter (outside flow constraints).
3. **Feasibility Filtering:** Remove improvements that exceed resource constraints or conflict with operational requirements.

When a high-value improvement appears "too large" for a single quarter, the issue is almost always insufficient decomposition, not the actual size. The framework, therefore, requires that any such improvement be decomposed and design-resolved in the first quarter, producing the smallest viable, deliverable increments. If decomposition initially appears impossible, this is treated as an architectural or planning deficiency — and the act of resolving that deficiency becomes the valid improvement for the quarter.

Collaborative Prioritisation: Power-of-Two Voting

The prioritisation step uses **power-of-two voting cards** (1, 2, 4, 8, 16) rather than Fibonacci sequences, which are common in agile practice. This refinement emerged from repeated use in enterprise and OT contexts where clear differentiation between initiatives proved essential.

The power-of-two progression provides: **sharper distinction between priorities** (each step doubles, avoiding ambiguous mid-range scores); and **clear signalling of exceptional importance** (the highest card conveys strong preference without allowing single-expert dominance). Participants must "spend" limited cards on improvements they genuinely value, surfacing collective priorities resilient to individual bias.

Priority Score Calculation

Each expert allocates voting cards across shortlisted improvements, with values summed to produce a **Priority Score**:

$$PriorityScore = \sum Card\ Values\ Assigned\ by\ All\ Participants$$

This produces three advantages: **resistance to individual overweighting** (identical card sets prevent hierarchical dominance); **sharp separation between priorities** (doubling sequence makes low-intensity votes unlikely to outweigh high-intensity preferences); and **transparent, defensible ranking** (clear numerical justification avoiding false precision).

The final prioritised list is then used to construct a quarterly improvement plan, ensuring that resource allocation reflects both operational feasibility and multi-disciplinary expert judgement rather than top-down directives or subjective preferences.

Voting Process

Each expert receives **power-of-two voting cards** (1, 2, 4, 8, 16), each allocable to **no more than one improvement**. Card scarcity forces deliberate trade-offs mirroring real-world resource constraints. Votes are revealed simultaneously to avoid anchoring or hierarchical influence. The mechanism captures disagreement without reconciliation discussions: differing views appear in card placement distribution.

The exponential spacing of the card values prevents low-intensity preferences from outweighing a small number of strong preferences, ensuring that genuinely important improvements rise to the top based on aggregated expert judgment.

Experts are not given a formal scoring rubric; each participant draws on their own integrated understanding of security, operational, safety, and resource considerations. This approach deliberately avoids formal weighting to respect domain expertise and tacit professional judgement, allowing power-of-two voting to translate these qualitative assessments into a transparent prioritisation signal without forcing them into artificial numeric constructs.

Where strong disagreements persist after voting, the facilitator convenes a brief reconciliation discussion; safety-related matters follow the safety specialist's documented veto authority, and outstanding conflicts are escalated to plant management for final decision.

Psychological Design for Resistance Reduction

The co-authoring structure addresses **not-invented-here (NIH)** responses [26], where personnel reject solutions perceived as externally imposed. By involving OT, safety, and operational experts in prioritisation, every participant becomes an active co-author of the quarterly plan, substantially reducing NIH tendencies.

Psychological safeguards include: **equal voting weight** (no participant dominates regardless of seniority); **simultaneous vote reveal** (prevents anchoring and hierarchical signalling); **power-of-two scoring** (forces deliberate trade-offs, makes differing assessments explicit); and **quarterly re-voting cycles** (reduces defensiveness by allowing correction in subsequent iterations).

These design choices shift decision-making from positional authority toward collaborative analytical reasoning, increasing local ownership and enhancing OT personnel's willingness to support risk-reduction measures in operationally conservative cultures.

*The voting design does not entirely eliminate **intentional group behaviour** such as informal coalitions or coordinated strategic voting — risks inherent to any collaborative prioritisation where functional domains have different incentives. The framework constrains such effects through equal voting weights, limited card budgets, simultaneous revelation, and quarterly re-voting, thereby reducing strategic manipulation payoffs. Detecting subtle coordinated behaviour remains an open challenge. Additional safeguards — anonymous pre-voting, rotating facilitation, and justification requirements for high-value votes — represent promising avenues for future research.*

4.6 Implementation and Progress Management

Quarterly Asset Selection

The enterprise selects the asset with the highest *currently validated* ALE for focused improvement. Following successful risk reduction — demonstrated through penetration testing or empirical control-effectiveness measurement — the programme shifts to the next-highest ALE asset.

Monthly Expert Group Reviews

- Track quarterly initiative advancement, identifying and resolving blockers.
- Assess completed minor improvements and select up to three new initiatives (maintaining the WIP limit of three).
- Adjust priorities based on operational learning and the evolving threat landscape.

The WIP limit ensures throughput stability and prevents resource dilution across too many parallel initiatives. Where an unplanned incident response temporarily consumes all available capacity, planned improvements may be deferred and then reviewed and rebalanced at the subsequent monthly/quarterly review.

Enterprise Progress Tracking

Monitor **aggregate enterprise ALE reduction quarterly**, providing board-level visibility without requiring detailed technical discussion. When penetration testing fails to compromise

an asset after improvement work, ALE can be significantly reduced, demonstrating measurable progress. This methodology provides executives with concrete evidence of the effectiveness of security improvements, addressing the **sunk cost fallacy** [53] and strengthening continued investment.

Ongoing Threat Intelligence Integration

Systematic monitoring of CISA's Known Exploited Vulnerabilities (KEV) catalogue [19] may trigger re-evaluation of previously secured assets. When CISA publishes a KEV entry affecting products deployed within the organisation's OT estate, the organisation must verify that at least two independent controls continue to adequately mitigate the associated attack vector, ensuring defence-in-depth even if one control fails. Although CISA publishes a notable monthly volume of KEV entries, only a small subset typically applies to any given operator's deployed industrial products. This filtering keeps the workload manageable and allows thorough analysis without generating operational urgency.

5 Theoretical Analysis

5.1 Mathematical Validation

The framework integrates established mathematical principles from queuing theory, multi-criteria decision analysis, and probability theory to justify its structural choices.

Queuing Theory Inspiration

Little's Law states that in stable queuing systems, average lead time equals work-in-progress divided by throughput ($L = W/\lambda$). Cybersecurity environments violate steady-state assumptions — threat arrivals are bursty and non-stationary — so the framework does not claim rigorous application of queuing theory.

Instead, the framework draws on the **qualitative insight** that constraining concurrent work reduces lead time variability and improves completion predictability. The WIP limits (one major, three minor improvements) are **governance heuristics** informed by this principle, not mathematical optimisation results.

Multi-Criteria Decision Analysis Structure

Expert prioritisation employs power-of-two voting (1, 2, 4, 8, 16):

$$PriorityScore = \sum_{i=1}^{n} (expert_i\ vote\ \times\ 2^{preference\ level})$$

Exponential weighting ensures that high-priority consensus cannot be mathematically overridden by numerous low-priority votes, thereby preserving expert judgement and

producing defensible, auditable rankings. The absence of ambiguous mid-range values (common in Fibonacci sequences) supports more precise comparison of competing improvements.

Probability Theory Application

The progressive ALE method transitions through external estimates, validated operational scenarios, and empirical measurement phases:

$$ALE_{Phase1} = ARO_{conservative} \times SLE_{insurance}$$

$$ALE_{Phase2} = ARO_{conservative} \times SLE_{pentest-validated}$$

$$ALE_{Phase3} = ARO_{empirical} \times SLE_{pentest-validated}$$

Here, $\mathbf{ARO_{empirical}}$ represents a frequency-based approximation of threat pressure, derived from observed control activity and prevented attack attempts — appropriate for OT environments where conventional probability estimation for high-impact, low-frequency events is infeasible due to the lack of representative datasets.

The approach maintains probabilistic coherence while acknowledging these inherent data limitations.

5.2 Convergence Properties

Progressive ALE Refinement Convergence

The three-phase methodology converges toward *measurement accuracy* rather than *ALE reduction*. Phase 1 establishes conservative baselines using insurance-derived SLE (typically pessimistic). Phase 2 incorporates penetration-testing insights for realistic, often lower SLE estimates. Phase 3 introduces empirical measurement through observed security-control performance.

Convergence arises as parameter uncertainty is progressively replaced with measured values. However, empirical observation frequently reveals that actual threat pressure exceeds initial assumptions. Convergence toward accuracy does not imply sustained ALE reduction — new vulnerabilities may emerge faster than improvement capacity. The framework's value lies in ensuring continuous focus on the highest-priority risks**.**

Expert Consensus Convergence

Expert co-authoring typically shows convergence as priority differences become explicit through power-of-two scoring. Structured discussion clarifies assumptions, and subsequent

rounds often show reduced variance. The method achieves convergence through explicit prioritisation choices**,** bounded card budgets, and periodic quarterly re-evaluation.

WIP Optimisation Convergence

The flow-constrained design enables convergence toward optimal resource utilisation through learning effects. Teams observe actual completion times and refine planning accuracy over successive quarters. WIP constraints prevent overload that would obscure learning signals.

Mathematically, convergence rests on principles of bounded system behaviour: constraining concurrent work prevents uncontrolled complexity growth and supports more predictable throughput. The framework stabilises organisational improvement workflow, not the cyber-threat environment itself.

5.3 Mathematical Limitations and Assumptions

Cybersecurity environments violate equilibrium assumptions: threat arrival rates consistently exceed remediation capacity. The framework, therefore, does not claim mathematical optimality or equilibrium achievement.

The WIP limits represent empirically informed governance heuristics rather than theoretically derived optima. Their value lies in improving lead time predictability and preventing resource dilution under inherently non-stationary conditions.

Similarly, progressive ALE demonstrates improved measurement accuracy through operational experience, but rigorous convergence proofs with defined error bounds exceed the intended scope and may not be meaningful in adversarial environments.

5.4 Comparative Framework Analysis

OT security standards provide strong technical guidance but embed organisational assumptions that rarely hold in resource-constrained critical infrastructure settings.

Table 3 summarises these assumptions and shows how the proposed flow-constrained model complements, rather than replaces, them.

Table 3. Comparison of Framework Assumptions and Gaps Addressed

Framework	**Key Built-in Assumptions**	**What the Framework Measures**	**Gap Filled by Flow-Constrained Model**
NIST SP 800-82r3	Organisation can implement full OT control baselines;	Binary compliance; % of controls implemented	Aligns improvement pace to real capacity (WIP limits); quarterly visibility; bias-

	stable budget; specialist team available		resistant prioritisation; no need for multi-million strategic cases
ISA/IEC 62443	Organisation can conduct full zone–conduit analysis, achieve SL targets, and coordinate vendors	SL attainment; compliance %	Enables continuous improvement without long assessment pauses; reduces OT resistance via co-authoring and iterative ALE
NIST SP 800-39	Organisation can perform enterprise-wide risk analysis and define risk appetite at board level	Risk-register status; annual/bi-annual reviews	Eliminates annual-cycle delay; enables immediate quarterly improvements; provides trending evidence that avoids strategic re-approval cycles
Flow-Constrained Framework	Small security team; availability for monthly sessions; improvements funded from operational budget	ALE trend; control-effectiveness evidence; delivered improvements	Addresses all four systemic barriers: resource limits, data scarcity, behavioural bias, and strategic-approval bottlenecks

Table 4. Implementation Feasibility

Dimension	**Traditional (NIST/ISA/IEC)**	**Flow-Constrained**	**Advantage**
Start-up time	6–12 months (enterprise assessment + business case)	2-week asset ranking; immediate start	×10–25 faster initiation
Approval route	Requires multi-year CAPEX case; high rejection rate	Uses existing operational budget	Removes approval barrier
Stakeholder behaviour	Top-down mandate; resistance common	Co-authoring; explicit bias safeguards	Eliminates not-invented-here / anchoring
Delivery model	Broad, multi-year roadmap	One asset per quarter; strict WIP discipline	Ensures completion; avoids dilution
Visibility	Binary transformation status	Quarterly ALE trend; empirical control data	Sustains executive confidence
Adaptation speed	Changes require new approval (6–12 months)	Reprioritisation each quarter; monthly learning	Matches threat evolution

Integrative Discussion

Traditional frameworks assume:

- stable and substantial security budgets,
- abundant skilled personnel,
- comprehensive assessments before action,
- rational decision-making without behavioural constraints, and
- multi-year planning stability.

These conditions seldom exist in operational OT environments facing safety constraints, maintenance windows, and competing production pressures.

The flow-constrained model acts as an *implementation bridge*, enabling organisations to make measurable progress **without requiring the idealised organisational infrastructure** assumed by existing standards.

It therefore **complements established frameworks**, providing a pragmatic execution layer that enables progress where traditional approaches typically stall.

6 Composite Scenario Application

This section applies the proposed framework to a **composite scenario** synthesised from recurring operational patterns observed in the author's multi-country OT-security work across energy, mining, and metallurgical sectors. **Individual client data cannot be disclosed** due to contractual and national security restrictions. However, every structural pattern represented here — resource constraints, maintenance-window limitations, safety–security trade-offs, stakeholder resistance, and cognitive-bias effects — is consistently documented in the peer-reviewed literature [38-41, 44, 45], indicating that these issues are systemic rather than organisation-specific.

All numerical values in this scenario are anonymised examples consistent with the evidence-constrained assumptions documented in Section 3.7; they are not client measurements.

6.1 Composite Scenario Context

Global Energy & Resources Conglomerate (GERC) — Composite Case

Structure and Asset Base

– Multinational energy/mining/metallurgy conglomerate
– 12 major industrial sites across four jurisdictions
– Approx. 45 OT production systems (SCADA, PLC networks, safety instrumentation)

Baseline Security Posture (Evidence-Aligned)

Each quantitative element below is bounded by the evidence-constrained assumptions documented in Section 3.7; no organisation-specific measurements are used.

- **Legacy SCADA footprint:** approx. 60% (>10 years old), aligning with evidence that legacy ICS devices remain prevalent globally [38, 48].
- **Known severe vulnerabilities:** approx. 100–150 OT-relevant CVEs (CVSS ≥7.0) across deployed systems, consistent with mixed-vendor ICS vulnerability analyses [38].
- **Security spending:** aligned with empirical evidence of sub-1% OpEx allocations reported across constrained industrial operators [39].
- **OT-Security team:** approx. 2–4 FTE, consistent with documented staffing scarcity across resource-constrained operators [39, 51].
- **Attempted ISA/IEC 62443 SL-2 programme** stalled for 18 months due to scope, vendor dependencies, and resource limitations.

Operational Constraints

- Production cannot pause during the nine peak-demand months.
- Formal maintenance windows: one 2-week shutdown every six months per site, consistent with documented industrial maintenance scheduling constraints [40, 41].
- Capital expenditure exceeding £500K requires 8–12 months board approval.
- An operational budget allows one major improvement per quarter across the entire enterprise.

Threat Environment

- The sector is listed in the UK NCSC Annual Reviews as a **priority target for state-sponsored APT groups.**
- MITRE ATT&CK for ICS identifies multiple APTs actively targeting **energy, mining and extraction** operations.
- Peer organisations recently experienced ransomware incidents, resulting in production losses of more than £40 million.
- Multiple GERC OT assets assessed — via vendor advisories and threat intelligence services — as active targets of foreign state-backed cyber operations.

Interpretation: These conditions typify the structural gap between widely used OT-security frameworks and the operational realities of critical infrastructure.

6.2 Baseline State and Failure of Traditional Approach

GERC initially attempted a **comprehensive ISA/IEC 62443-aligned transformation**.

Phase 1 (Months 1–6): Enterprise Risk Assessment

- 342 gaps identified across all sites.
- Draft five-year roadmap.
- Estimated cost: £28M.

Phase 2 (Months 7–14): Business Case Development

- The ROI analysis failed due to insufficient quantitative risk data.
- The board judged the justification insufficient.
- No improvements were deployed after 14 months.

This is not a literal historical snapshot but a composite baseline reflecting empirically reported delay patterns in OT environments, where staffing limits, maintenance-window cycles and approval dependencies routinely extend implementation times.

Why the Traditional Approach Failed (Cognitive-Bias Mechanisms)

Empirical studies show that OT-security investment decisions frequently deviate from rational-choice assumptions [13]. GERC exhibited the following patterns:

– **Planning fallacy** — underestimation of delivery effort; implicit belief that full deployment was feasible despite capacity for <6 improvements/year.
– **Optimism bias** — discounting of threat likelihood despite clear APT interest in the sector.
– **Analysis paralysis** — inability to prioritise 342 gaps without quantitative data led to continued deferral.
– **Availability heuristic** — operational incidents overshadowed cyber-risk signals, lacking recent local precedent.
– **Anchoring bias** — the £28 million estimate anchored executives to an "excessive cost" perception, blocking exploration of incremental paths.
– **Temporal mismatch** — multi-year planning cycles could not align with adversary evolution; urgency was masked.

This pattern reflects a composite of recurring findings reported across multiple critical-infrastructure studies rather than a single organisation.

Conclusion: Under these conditions, technically sound frameworks become operationally unworkable.

Catalyst for Change

Month 15: Two SCADA-relevant vulnerabilities were added to the CISA KEV catalogue, prompting board demand for an immediate, evidence-driven action plan. The CISO proposed the flow-constrained framework.

6.3 Framework Implementation: Two-Quarter Illustration

Quarter 1: Insurance-Based ALE, Prioritisation, Delivery

Initial conservative ALE ranking identified **Site 'Alpha'** as the highest risk (£3.6M estimated ALE, derived from insurance-based SLE and conservative ARO assumptions).

A seven-member expert group (operations, safety, IT, OT security) generated 23 improvement options. Using **power-of-two voting**, the group selected:

- **1 major improvement**: SCADA–corporate segmentation via data diode.
- **3 minor improvements**: MFA for remote access; KEV-aligned patching; security logging enhancement.

Strict WIP constraints ensured delivery within 12 weeks without operational disruption.

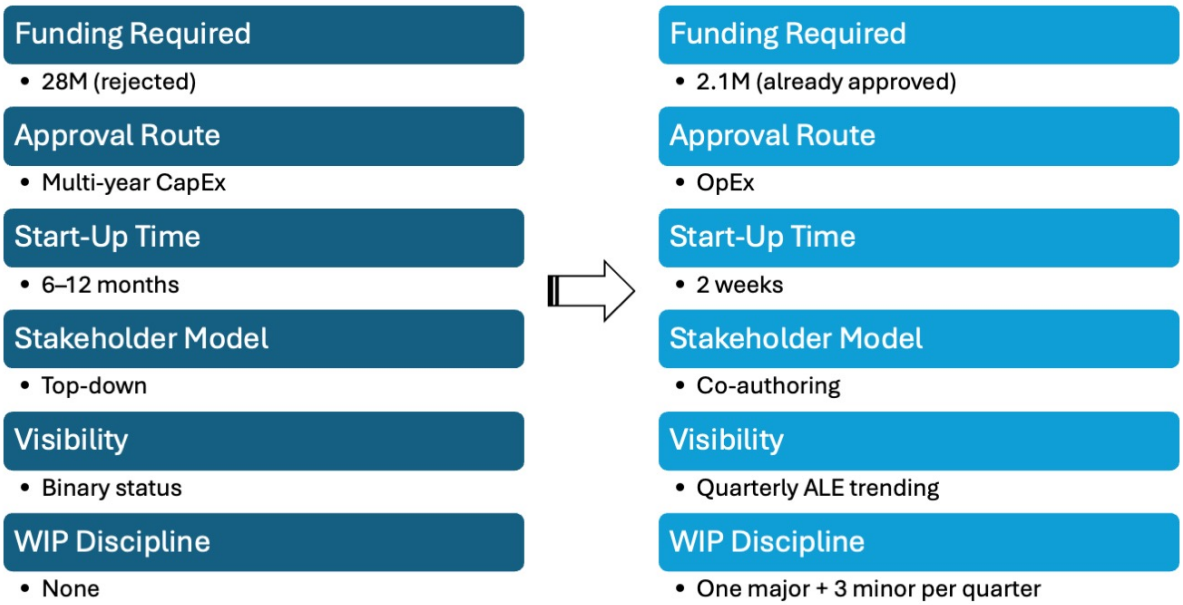
Quarter 2: Pentest Validation & Progressive ALE Refinement

External penetration testing confirmed:

- Lateral movement is blocked by the new segmentation.
- Credential-stuffing is blocked by MFA.
- An additional HMI vulnerability was identified.

Refined scenario modelling reduced ALE from £3.6M to £1.26M.

Figure 3. Traditional vs Flow-Constrained Approach



Quarter 2 delivered the required HMI patch plus three further minor improvements (anti-phishing training, engineering workstation hardening, backup-validation uplift).

6.4 Quantitative Outcomes

Table 5 reports activity and resource metrics used to quantify implementation velocity; security outcomes are assessed separately through progressive ALE refinement and empirical control-performance validation.

Table 5. Implementation Velocity (12-Month Equivalent)

Metric	Traditional (14 months)	Flow-Constrained
Improvements implemented	0	16 across 2 sites
Operational disruption	N/A	0 hours
Funding required	£28M (rejected)	£2.1M (approved quarterly)

Pentest results provided objective evidence of security improvement, shifting stakeholder sentiment: OT managers (+3.1/7), plant managers (+2.3/7), safety specialists (+2.7/7). Initial sceptics became advocates once improvements showed operational safety was not compromised.

The total reflects two quarterly cycles under fixed WIP limits and is the method's calculated output rather than an industry benchmark.

6.5 Lessons Learned and Limitations

Observed Benefits

- **WIP constraints** prevented overcommitment (3 minor vs 7 initially desired).
- **Progressive ALE** provided increasingly credible measurement through external and empirical validation.
- **Expert co-authoring** within the multi-stakeholder group reduced "not-invented-here" reactions and improved cross-disciplinary cooperation.
- **Quarterly cadence** maintained momentum and executive confidence.

Observed Challenges

- Penetration testing revealed higher-than-expected actual risk.
- Legacy equipment required vendor consultation.
- Attribution logic for prevented attacks required careful documentation, consistent with known attribution challenges in OT environments noted in prior studies.

6.6 Limitations of the Composite Scenario

Although grounded in converging practitioner observations and empirical literature, this scenario has inherent constraints:

1. **Controlled validation required** — generalisation requires cross-sector pilots with comparison groups.
2. **Threat-sophistication variance** — although the scenario accounts for targeting by state-backed cyber-warfare actors, modelling is limited to publicly documented TTPs; undisclosed capabilities or unobserved zero-days may exceed the representable threat profile.
3. **Regulatory variation** — nuclear/defence sectors may impose constraints requiring adaptation.
4. **Organisational culture** — highly dysfunctional environments may require additional change management support.
5. **Technical environment** — fully air-gapped systems require alternative forms of empirical validation.

Quantitative elements follow the evidence-constrained assumptions documented in Section 3.7; they illustrate representative mid-ranges rather than sectoral benchmarks.

Generalisation Guidance

Core principles — **WIP constraints, progressive ALE, and expert co-authoring** — appear sector-agnostic, given replication of the underlying constraints across energy, manufacturing, mining, and industrial automation contexts. Implementation details, however, must adapt to regulatory regimes, organisational maturity, threat environment, and local resource availability.

7 Discussion

7.1 Practical Implications

The framework enables security improvement within realistic organisational constraints through: (1) operational budget integration, eliminating 12-18 month strategic approval delays, (2) progressive credibility building from external validation to empirical measurement, (3) expert engagement protocols overcoming not-invented-here syndrome through power-of-two voting and collaborative prioritisation, (4) visible progress metrics maintaining executive confidence through quarterly ALE trending.

7.2 Theoretical Contributions

The research contributes a **novel integration of flow optimisation with cybersecurity risk management**. Four specific contributions: (1) first systematic application of WIP-limited

flow optimisation to cybersecurity portfolio management for resource-constrained critical infrastructure, (2) three-phase progressive ALE refinement methodology addressing systematic data scarcity, (3) behavioural-bias mitigation addressing the manifestations of bounded rationality in OT decision-making as a core framework architecture explicitly addressing planning fallacy, optimism bias, not-invented-here syndrome, and anchoring bias, (4) attack attribution methodology for empirical ARO calculation based on control-prevented attacks.

7.3 Limitations and Boundary Conditions

The framework assumes basic security-monitoring data to support empirical ARO estimation, regular cross-domain participation in monthly prioritisation sessions, and a multi-month observation window to establish an initial ARO baseline. It further relies on improvements that are decomposable into discrete initiatives with measurable effects and may require adaptation in sectors where regulatory structures, distributed operations, or vendor-controlled OT limit the applicability of the flow-constrained approach.

ALE and ARO estimates carry inherent uncertainty due to sparse monitoring data, parameter subjectivity, and variability in attacker behaviour. Confidence intervals were not calculated because of data sparsity; instead, uncertainty is acknowledged qualitatively. These epistemic limits affect the precision of early estimates but do not undermine the comparative value of progressive refinement across phases.

When vendor patch cycles or firmware updates are outside operator control, remediation timelines are dictated by supplier release schedules rather than organisational prioritisation. Operators remain accountable for the residual risk on legacy or vendor-controlled OT assets, even when direct software remediation is infeasible. In these situations, compensating controls — such as tightened network segmentation, read-only or interface-restricted HMIs, enhanced anomaly detection, and monitoring uplift — become the primary means of reducing exploitability. The framework's prioritisation logic still applies, directing limited resources toward the compensating controls that most effectively reduce attack paths or contain propagation risk. These adaptations enable FC-RM to operate effectively in environments dominated by legacy or vendor-controlled OT systems, where remediation options are inherently constrained.

The mechanisms generalise to other CI sectors (water, transport, telecom) where similar operational constraints and approval cycles apply.

7.4 Future Research Directions

Priority research areas: (1) controlled comparative studies evaluating framework against traditional approaches with comparison groups, (2) multi-case studies across telecommunications, water treatment, transportation, nuclear sectors identifying sector-specific adaptations, (3) longitudinal validation over 3-5 years assessing sustained

effectiveness, (4) integration with threat intelligence feeds enabling automated priority adjustment, (5) empirical validation during actual cyber incidents. Future research should employ multi-case designs with control groups, standardised quantitative metrics, and broader organisational sampling across sectors. Further work should also examine methodological refinements to empirical ARO estimation, including improved false-positive filtering, attribution-accuracy evaluation, and assessment of inter-operator consistency to strengthen analytical robustness.

A further research direction is the integration of bias-informed decision-making principles into future regulatory assurance models. Current assessment regimes, including the NCSC Cyber Assessment Framework [54], concentrate on evidencing the state of controls at defined points in time. They provide limited visibility of how organisations prioritise, sequence, and resource improvements under operational constraints. Embedding elements of prioritisation discipline, WIP-limited delivery, and structured multi-expert decision processes into future assurance criteria may offer a more realistic representation of how critical-infrastructure operators progress security enhancement over time. Exploring how such mechanisms could complement existing compliance frameworks represents a substantive avenue for future study.

8 Conclusion

This research presents an FC-RM framework addressing systematic implementation barriers in resource-constrained critical infrastructure environments through four integrated contributions: (1) WIP-limited flow optimisation for cybersecurity portfolio management, (2) progressive ALE methodology addressing data scarcity, (3) explicit behavioural bias mitigation, addressing bounded-rationality constraints, as core architecture, (4) empirical ARO calculation based on control-prevented attacks.

The mathematical foundation provides defensible justification for security resource allocation whilst acknowledging analytical limitations characterising operational technology environments, making systematic security improvement accessible to organisations lacking resources for sophisticated quantitative analysis. Collaborative stakeholder engagement addresses critical implementation challenges through involving operational personnel in security decision-making whilst providing structured analytical processes.

Application feasibility demonstrated through a composite scenario — constructed from recurring real-world patterns observed in the author's multi-country OT security experience and triangulated with peer-reviewed empirical literature — shows how the framework enables board-level visibility whilst avoiding contentious budget negotiations. Quantitative values reflect a plausible mid-range composite scenario based on recurring patterns observed in practice and convergent findings in the empirical literature. They are intended to illustrate how the framework operates under realistic constraints, rather than to provide predictive or statistically derived estimates.

For critical infrastructure operators, the framework offers an actionable methodology for measurable security enhancement within realistic constraints. For researchers, it demonstrates how operational research principles can enhance cybersecurity effectiveness. Future empirical validation across diverse sectors will enable refinement of sector-specific adaptation requirements.

9 Declaration of Competing Interest

The author declares no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

10 Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work, the author used ChatGPT (OpenAI) to improve readability and language quality as a non-native English speaker. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the publication's content.

11 References

[1] Dragos. 2025 OT/ICS cybersecurity report: 8th annual year in review. [Internet]. Hanover (MD): Dragos Inc.; 2025 [cited 04 Dec 2025]. Available from: https://hub.dragos.com/hubfs/312-Year-in-Review/2025/Dragos-2025-OT-Cybersecurity-Report-A-Year-in-Review.pdf.

[2] United States National Security Agency (NSA), United States Cybersecurity and Infrastructure Security Agency (CISA), United States Federal Bureau of Investigation (FBI), et al. Countering Chinese state-sponsored actors compromise of networks worldwide to feed global espionage system. Joint cybersecurity advisory. [Internet]. Washington (DC): US Government; 2025 [cited 04 Dec 2025]. Available from: https://media.defense.gov/2025/Aug/22/2003786665/-1/-1/0/CSA_COUNTERING_CHINA_STATE_ACTORS_COMPROMISE_OF_NETWORKS.PDF.

[3] HM Treasury. UK infrastructure: A 10-year strategy. [Internet]. London: HM Treasury; 2025 [cited 04 Dec 2025]. Available from: https://assets.publishing.service.gov.uk/media/6853c606df3015b374b73656/UK_Infrastructure_A_10_Year_Strategy_TEXT_PRINT.pdf.

[4] Evripidou S, Watson JDM. Understanding operational technology personnel's mindsets and their effect on cybersecurity perceptions: A qualitative study with operational technology cybersecurity practitioners. In: Proceedings of the 2024 European Symposium on Usable Security, Karlstad, Sweden; 2024. ACM, pp.137–154. doi:10.1145/3688459.3688472.

[5] National Institute of Standards and Technology. NIST SP 800-82r3: Guide to operational technology (OT) security. [Internet]. Gaithersburg, MD: NIST; 2023 [cited 04 Dec 2025]. Available from: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-82r3.pdf.

[6] Anderson DJ. Kanban: Successful evolutionary change for your technology business. Sequim, WA: Blue Hole Press; 2010.

[7] Knaster R, Leffingwell D. SAFe 5.0 distilled: Achieving business agility with the Scaled Agile framework. Boston, MA: Addison-Wesley Professional; 2020.

[8] National Institute of Standards and Technology. NIST SP 800-39: Managing information security risk, organization, mission, and information system view. [Internet]. Gaithersburg, MD; 2011 [cited 26 July 2025]. Available from: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-39.pdf.

[9] Chapple M, Stewart JM, Gibson D. ISC2 CISSP certified information systems security professional official study guide. Chichester: Sybex; 2024.

[10] Cremer F, Sheehan B, Fortmann M, Kia AN, Mullins M, Murphy F, Materne S. Cyber risk and cybersecurity: A systematic review of data availability. Geneva Pap Risk Insur Issues Pract. 2022;47:698–736. doi:10.1057/s41288-022-00266-6.

[11] Cook A, Smith R, Maglaras L, Janicke H. Measuring the risk of cyber attack in industrial control systems. In: 4th International Symposium for ICS & SCADA Cyber Security Research 2016, Swindon; 2016. Electronic Workshops in Computing. BCS Learning & Development, pp.103–113. doi:10.14236/ewic/ics2016.12.

[12] Kahneman D, Tversky A. Prospect theory: An analysis of decision under risk. Econometrica. 1979;47:263–291. doi:10.2307/1914185.

[13] Korteling JE, Paradies GL, Sassen-van Meer JP. Cognitive bias and how to improve sustainable decision making. Front Psychol. 2023;14:Article 1129835. doi:10.3389/fpsyg.2023.1129835.

[14] Salzberger A. The optimistic bias in cyber risk perception of German enterprises: Do organizational and personal moderators matter? Organizational Cybersecurity Journal: Practice, Process and People. 2025. doi:10.1108/ocj-02-2024-0003.

[15] International Society of Automation. ISA/IEC 62443 series of standards. [Internet]. Research Triangle Park, NC: International Society of Automation; 2025 [cited 05 Dec 2025]. Available from: https://www.isa.org/standards-and-publications/isa-standards/isa-iec-62443-series-of-standards.

[16] Silva ADPd, Bobbert Y. Cybersecurity readiness: An empirical study of effective cybersecurity practices for industrial control systems. Sci J Research & Rev. 2019. doi:10.33552/sjrr.2019.02.000536.

[17] UK Parliament. Cyber security and resilience (network and information systems) Bill. [Internet]. London: UK Parliament; 2025 [cited 05 Dec 2025]. Available from: https://bills.parliament.uk/bills/4035.

[18] National Cyber Security Centre. NCSC annual review 2025. [Internet]. London: National Cyber Security Centre; 2025 [cited 27 Nov 2025]. Available from: https://www.ncsc.gov.uk/files/ncsc-annual-review-2025.pdf.

[19] Cybersecurity and Infrastructure Security Agency. Known exploited vulnerabilities catalog. [Internet]. Washington, DC: Cybersecurity and Infrastructure Security Agency; n.d. [cited 05 Dec 2025]. Available from: https://www.cisa.gov/known-exploited-vulnerabilities-catalog.

[20] House of Commons Library. Infrastructure in the UK. [Internet]. London: House of Commons Library; 2025 [cited 05 Dec 2025]. Available from: https://researchbriefings.files.parliament.uk/documents/SN06594/SN06594.pdf.

[21] Simon HA. A behavioral model of rational choice. Q J Econ. 1955;69:99–118. doi:10.2307/1884852.

[22] Simon HA. Bounded rationality. In: Eatwell J, Milgate M, Newman P (Eds.). Utility and probability. London: Palgrave Macmillan; 1990. pp.15–18.

[23] Morris T. Industrial control system (ICS) cyber attack datasets. [Internet]. Mississippi State, MS: Mississippi State University; 2015 [cited 05 Dec 2025]. Available from: https://sites.google.com/a/uah.edu/tommy-morris-uah/ics-data-sets.

[24] Canbek G, Temizel TT, Sagiroglu S. Gaining insights in datasets in the shade of "garbage in, garbage out" rationale: Feature space distribution fitting. WIREs Data Min Knowl Discov. 2022;12:e1456. doi:10.1002/widm.1456.

[25] Anderson R, Moore T. The economics of information security. Science. 2006;314:610–613. doi:10.1126/science.1130992.

[26] Katz R, Allen TJ. Investigating the not invented here (NIH) syndrome: A look at the performance, tenure, and communication patterns of 50 R&D project groups. R D Manag. 1982;12:7–19. doi:10.1111/j.1467-9310.1982.tb00478.x.

[27] Freund J, Jones J. Measuring and managing information risk: A FAIR approach. Burlington, MA: Butterworth-Heinemann; 2015.

[28] Keeney RL, Raiffa H. Decisions with multiple objectives: Preferences and value tradeoffs. Cambridge: Cambridge University Press; 1993.

[29] Iyengar SS, Lepper M. When choice is demotivating: Can one desire too much of a good thing? J Pers Soc Psychol. 2000;79:995–1006. doi:10.1037/0022-3514.79.6.995.

[30] Stadnicka D, Bonci A, Lorenzoni E, Dec G, Pirani M. Symbiotic cyber-physical Kanban 4.0: An approach for SMEs. In: Proceedings of the 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Vienna, Austria; 2020. IEEE, pp.140–147. doi:10.1109/ETFA46521.2020.9212073.

[31] Little JDC. A proof for the queuing formula: L = λw. Oper Res. 1961;9:383–387. doi:10.1287/opre.9.3.383.

[32] Freeman RE. Strategic management: A stakeholder approach. Cambridge, MA: Cambridge University Press; 1984.

[33] Hevner AR, March ST, Park J, Ram S. Design science in information systems research. MIS Q. 2004;28:75–106. doi:10.2307/25148625.

[34] Gordon LA, Loeb MP. Budgeting process for information security expenditures. Commun ACM. 2006;49:121–125. doi:10.1145/1107458.1107465.

[35] Aven T. Risk assessment and risk management: Review of recent advances on their foundation. Eur J Oper Res. 2016;253:1–13. doi:10.1016/j.ejor.2015.12.023.

[36] Hohmann L. Innovation games: Creating breakthrough products through collaborative play. Boston: Addison-Wesley Professional; 2006.

[37] MITRE Corporation. MITRE ATT&CK® matrix for ICS. [Internet]. McLean, VA: MITRE Corporation; 2025 [cited 27 Nov 2025]. Available from: https://attack.mitre.org/matrices/ics/.

[38] Anton SDD, Fraunholz D, Krohmer D, Reti D, Schneider D, Schotten HD. The global state of security in industrial control systems: An empirical analysis of vulnerabilities around the world. IEEE Internet Things J. 2021;8:17525–17540. doi:10.1109/JIOT.2021.3081741.

[39] Moon S, Hou L, Han S. Empirical study of an artificial neural network for a manufacturing production operation. Oper Manag Res. 2023;16:311–323. doi:10.1007/s12063-022-00309-0.

[40] Cavalcante CAV, Scarf P, Melo YR, Rodrigues AJS, Alotaibi N. Planning maintenance when resources are limited: A study of periodic opportunistic replacement. IMA J Manag Math. 2024;35:573–593. doi:10.1093/imaman/dpae015.

[41] Zhang Z, Yang L. Postponed maintenance scheduling integrating state variation and environmental impact. Reliab Eng Syst Saf. 2020;202:107065. doi:10.1016/j.ress.2020.107065.

[42] Zhou C, Li X, Yang S, Tian YC. Risk-based scheduling of security tasks in industrial control systems with consideration of safety. IEEE Transactions on Industrial Informatics. 2020;16:3112–3123. doi:10.1109/TII.2019.2903224.

[43] Agbo C, Mehrpouyan H. Conflict analysis and resolution of safety and security boundary conditions for industrial control systems. In: 2022 6th International Conference on System Reliability and Safety (ICSRS); 2022. pp.145–156. doi:10.1109/ICSRS56243.2022.10067393.

[44] Sheikhi S, Eceiza M, Arellano C, López O, Kelnberger S, Lindner R, Partanen J, Lovén L. Bridging theory and practice: Addressing current cybersecurity gaps in industry 5.0. IEEE Access. 2025;13:92891–92905. doi:10.1109/ACCESS.2025.3569130.

[45] Heinl MP, Pursche M, Puch N, Peters SN, Giehl A. From standard to practice: Towards ISA/IEC 62443-conform public key infrastructures. In: Computer Safety, Reliability, and

Security: 42nd International Conference, SAFECOMP 2023, Toulouse, France; 2023. Lecture Notes in Computer Science, vol. 14281. Springer, pp.196–210. doi:10.1007/978-3-031-40923-3_15.

[46] Badawy M, Sherief NH, Abdel-Hamid AA. Legacy ICS cybersecurity assessment using hybrid threat modeling—an oil and gas sector case study. Applied Sciences. 2024;14. doi:10.3390/app14188398.

[47] Chan ACF, Zhou J. Non-intrusive protection for legacy SCADA systems. IEEE Communications Magazine. 2023;61:36–42. doi:10.1109/MCOM.003.2200564.

[48] Shewale V. Securing legacy SCADA systems: Practical strategies for the oil and gas industry. World Journal of Advanced Research and Reviews. 2025;26:341–346. doi:10.30574/wjarr.2025.26.2.1575.

[49] Falco G, Caldera C, Shrobe H. IIoT cybersecurity risk modeling for SCADA systems. IEEE Internet Things J. 2018;5:4486–4495. doi:10.1109/JIOT.2018.2822842.

[50] Wai E, Lee CKM. Seamless industry 4.0 integration: A multilayered cyber-security framework for resilient SCADA deployments in cpps. Applied Sciences. 2023;13. doi:10.3390/app132112008.

[51] Ribeiro A. OT security skills gap is a major challenge for industrial, manufacturing organizations. [Internet]. 2021 [cited 05 Dec 2025]. Available from: https://industrialcyber.co/news/ot-security-skills-gap-is-a-major-challenge-for-industrial-manufacturing-organizations/.

[52] Adekoya OA, Atlam HF, Lallie HS. Quantifying the multidimensional impact of cyber attacks in digital financial services: A systematic literature review. Sensors. 2025;25. doi:10.3390/s25144345.

[53] Arkes HR, Blumer C. The psychology of sunk cost. Organ Behav Hum Decis Process. 1985;35:124–140. doi:10.1016/0749-5978(85)90049-4.

[54] National Cyber Security Centre. Cyber assessment framework v4.0. [Internet]. 2025 [cited 06 Dec 2025]. Available from: https://www.ncsc.gov.uk/files/NCSC-Cyber-Assessment-Framework-4.0.pdf.

Revision History:

December 2025: Major update to the August 2025 working paper, incorporating empirical control-prevention measurement, revised flow-constraint integration, expanded behavioural-bias mitigation mechanisms, and additional methodological detail to support peer-review submission.