Quantifying Operational Cost Risk in Industrial Plants Using the Industrial Actuarial Risk Management Framework

Michael Heath Lewis
IRX, LLC
San Marcos, TX, USA
info@costriskreliability.com

***Abstract*—Industrial plants face a persistent gap between engineering judgment and financial accountability. Reliability engineers characterize failure modes and maintenance needs while financial officers manage capital allocation and budget risk, yet these disciplines rarely share a common analytical language. This paper introduces the Industrial Actuarial Risk Management (iARM) framework, which applies collective risk modeling, Monte Carlo simulation, and actuarial exposure metrics to industrial operations. By treating total annual plant cost as a compound stochastic process with both event frequency and severity modeled probabilistically, the framework produces loss distributions from which expected loss, percentile thresholds, and conditional tail loss metrics are derived. A severity decomposition model explicitly incorporates repair, labor, downtime, collateral, and regulatory cost components, connecting physical failure outcomes to their full financial consequences. The framework operates through four pillars: Quantification, Prioritization, Mitigation, and Communication. Together, these transform reliability investment from anecdotal justification into risk-adjusted capital management, enabling engineering and finance to govern operational uncertainty from a shared, quantified perspective.**

***Index Terms*—actuarial risk management, industrial reliability, Monte Carlo simulation, collective risk model, maintenance capital allocation, loss distribution, tail risk**

I. Introduction

At some point in the life of every industrial plant, a reliability engineer and a chief financial officer sit across a table and talk past one another. The engineer understands which assets are deteriorating and which interventions would reduce exposure. The financial officer understands what capital costs and what the budget can absorb. Both are correct about what they know, yet neither has the tools to fully translate the other's perspective into actionable decisions. As a result, multi-million-dollar capital choices are made on authority, habit, or the most recent failure event rather than on a shared picture of economic reality.

This gap exists because industrial operations have lacked a common analytical language capable of translating engineering uncertainty into financial exposure with the rigor both disciplines require. The application of actuarial techniques to operational risk has precedent in the broader risk management literature [5], and the iARM framework extends this tradition specifically to the industrial plant environment. Actuarial science, developed over centuries to characterize uncertainty as a distribution of outcomes [1], translates directly to the plant setting where failure frequency, severity dispersion, and cost volatility follow the same compound loss structure as insurance portfolios.

The author declares no funding for this work.

II. The Industrial Plant as a Microeconomy

An industrial plant functions as a contained economic system: it deploys capital, consumes labor and materials, and converts production output into financial performance under conditions of uncertainty. Like any economy, its health cannot be assessed by averages alone, because true resilience lies in the *distribution* of outcomes, not their mean.

Traditional operational metrics such as mean time between failures, availability, and maintenance spend describe performance averages, not financial exposure. Consider two plants with identical availability: one experiences frequent, low-cost failures; the other experiences rare but catastrophic ones. Their average uptime figures converge while their financial risk profiles diverge dramatically. A plant evaluated solely by average uptime obscures the tail risk that average performance metrics were never designed to detect.

This distinction between average performance and loss distribution has direct implications for capitalization, insurance, and management. When a financial officer asks whether the maintenance budget is adequate, the honest answer requires not just a cost average but a statement of exposure: what is the realistic range of outcomes, what does the worst credible year look like, and how does proposed mitigation shift that distribution?

III. Sources of Uncertainty

Operational uncertainty in industrial plants emerges from five interacting dimensions that compound in ways deterministic planning cannot anticipate.

A. Failure Frequency Variability

Historical maintenance records may suggest a stable average failure rate, but equipment does not fail on a fixed schedule. Asset age introduces non-stationarity: as components enter wear-out phases, hazard rates increase in ways that average-based planning will detect only after exposure has already risen materially [4]. Even with a stable long-run mean, short-term clustering of failures can create significant cash-flow exposure within compressed timeframes.

This phenomenon is formalized here as the *Average Rate Illusion*. When reliability dashboards report a stable λ, they are reporting a time-averaged expectation, not a description of the underlying failure process. A constant mean is a property of the estimator, not a guarantee of process stationarity. In practice, a non-homogeneous Poisson process (NHPP) with an increasing intensity function $\mu(t)$ can produce an annual expected failure count that appears unremarkable while the local, short-term hazard is rising materially. The average masks the drift.

The financial consequences of this illusion are asymmetric and compounding. During the useful-life phase, failures are approximately independent and temporally dispersed, allowing maintenance budgets to absorb costs incrementally. As assets transition into the wear-out phase, the independence assumption breaks down. Correlated component aging, shared operating stressors, and accelerating hazard rates combine to produce failure clustering within compressed intervals. When multiple failures occur in the same planning window, the resulting cash-flow demand does not simply sum linearly. It interacts with procurement lead times, contractor availability constraints, and replacement-cost inflation to produce what this paper terms a *compressed cash-flow shock*: a concentrated financial exposure that standard MTBF-based budgeting frameworks are structurally unable to anticipate.

The implication is direct: reliability metrics describe distributional averages under stationarity assumptions, while financial risk management requires distributional characterization across the full range of plausible outcomes, including the tail. Bridging this gap requires moving beyond point estimates of failure frequency toward frequency-severity modeling capable of capturing variance, skewness, and tail behavior under non-stationary failure conditions.

B. Severity Dispersion

The range between best and worst outcomes for a given failure mode can span orders of magnitude, from routine corrective repair to extended downtime, cascading secondary damage, emergency contractor mobilization, regulatory consequence, and environmental remediation. The expected value of this dispersion is an incomplete description of exposure; the variance, skewness, and tail behavior matter enormously, as tail events drive disproportionate financial impact over multi-year horizons [2].

C. Cost Escalation

Labor rates, engineered component prices, and logistics expenses are subject to inflationary pressure that both amplifies the severity of every failure event and erodes the real purchasing power of fixed maintenance budgets. Plants that do not model cost escalation will systematically understate future financial risk.

D. Supply Chain and Logistical Uncertainty

A technically identical failure can produce dramatically different economic outcomes depending on lead times for critical spares, contractor availability, and transportation constraints. Supply chain uncertainty effectively stretches the upper tail of the loss distribution, converting moderate technical failures into severe financial consequences. Post-pandemic experience demonstrated that supply chain assumptions embedded in maintenance planning can become obsolete rapidly and without clear warning.

E. Organizational and Decision Uncertainty

Budget cycles that do not correspond to failure event timing create systematic pressure to defer maintenance expenditures. Organizational silos between reliability, operations, and finance mean that those best positioned to quantify risk are often not the people making capital allocation decisions. Historical bias, anchoring on recent experience rather than the full distribution of possible outcomes, leaves plants structurally underprepared for conditions that have not occurred recently but remain within the plausible range.

IV. Simulation Design: The Collective Risk Model

A. Event Frequency Modeling

The number of cost-bearing events in a planning year is modeled as:

$$N \sim \text{Poisson}(\lambda) \tag{1}$$

where λ is the expected annual event rate. The Poisson distribution explicitly captures year-to-year frequency variability rather than suppressing it behind a single expected value. For a Poisson random variable, variance equals the mean, so frequency variability grows proportionally with the expected event rate. Where aging effects are present, a non-homogeneous extension using a time-varying rate $\lambda(t)$ better reflects increasing hazard rates in aging asset populations [4], and exposure projections based on a static λ will systematically understate future frequency risk in such cases.

B. Event Severity Modeling

Each event carries a financial consequence modeled as:

$$X_i \sim \text{Lognormal}(\mu, \sigma^2) \tag{2}$$

The lognormal distribution is selected because costs are strictly positive, probability density is concentrated around moderate outcomes, and the right-skewed tail represents the rare catastrophic event that drives disproportionate financial exposure [1]. While alternative heavy-tailed distributions such as Pareto may be appropriate in specific contexts, the lognormal provides a practical balance between empirical realism and parameter interpretability in industrial datasets.

Rather than parameterizing with μ and σ^2 directly, iARM uses two operationally intuitive inputs: the expected cost per event $E[X]$ and the coefficient of variation $\text{CV} = \sigma_X/\mu_X$. The log-space variance is recovered as:

$$\sigma^2 = \ln(1 + \text{CV}^2) \tag{3}$$

and the log-space mean as $\mu = \ln(E[X]) - 0.5\sigma^2$. When $\mathrm{CV} = 0$, severity dispersion vanishes and each event carries a fixed deterministic cost equal to $E[X]$; the lognormal reduces to a point mass and no severity sampling is performed. This allows a complete severity distribution to be specified from inputs that practitioners can estimate from experience or historical records, without requiring statistical fitting to large datasets that may not exist for low-frequency, high-severity events. iARM treats the plant's ERP system as the primary source of record for calibrating λ, $E[X]$, and CV. Output quality is therefore directly conditioned on ERP data integrity: incomplete work order cost capture, inconsistent labor and materials coding, short or discontinuous history windows, and asset hierarchy changes that break failure continuity will each introduce bias into parameter estimates and should be assessed before model results are used for capital decisions.

C. Severity Decomposition

In industrial applications, the severity variable X_i represents total economic impact and is decomposed as:

$$X_i = C_{\text{repair}} + C_{\text{labor}} + C_{\text{downtime}} + C_{\text{collateral}} + C_{\text{regulatory}} \quad (4)$$

where downtime cost is modeled as:

$$C_{\text{downtime}} = T_{\text{outage}} \times R_{\text{prod}} \quad (5)$$

with T_{outage} denoting outage duration and R_{prod} representing production value per unit time. Additional components may include safety consequence costs, environmental remediation, regulatory penalties, and expedited logistics. Supply chain constraints that extend T_{outage} beyond the repair interval itself are a primary driver of tail severity, as a moderate technical failure can produce a disproportionate financial consequence when critical spares carry extended lead times. Aggregation of these components produces the strictly positive, right-skewed severity distribution modeled by the lognormal above.

D. Aggregate Annual Cost

Total annual financial exposure is the compound sum:

$$S = \sum_{i=1}^{N} X_i \quad (6)$$

This compound frequency-severity structure is the standard collective risk model used across actuarial science [1], and its application to industrial cost modeling follows from the recognition that plant financial exposure is driven by the same mathematical structure as insurance loss portfolios [5].

Independence between event frequency and severity is assumed as a baseline modeling condition. However, in industrial contexts positive correlation between the two is plausible: a common-cause event such as a process upset or extreme weather period can simultaneously elevate both the number of failures and the individual cost of each. Where operational data support dependence structures, correlated extensions may be incorporated.

The expected value of S is:

$$E[S] = \lambda \cdot \mu_X \quad (7)$$

This is the quantity traditional maintenance budgeting targets. It is necessary but insufficient as a description of exposure. The variance of S for a compound Poisson process follows from substituting $\mathrm{Var}(X) = \mu_X^2 \cdot \mathrm{CV}^2$ into the standard compound Poisson variance formula:

$$\mathrm{Var}(S) = \lambda \cdot \left[\mathrm{Var}(X) + (E[X])^2\right] = \lambda \cdot \mu_X^2 \cdot (1 + \mathrm{CV}^2) \quad (8)$$

Equation (8) reveals a key insight that average-based planning obscures entirely: aggregate cost volatility is driven not only by the expected event rate λ but by the second moment of the severity distribution, which grows rapidly as CV increases. A plant with high severity dispersion, even with a moderate event rate, can exhibit cost volatility that renders budget-based planning unreliable [1]. This is the formal mathematical basis for the argument that severity distribution matters as much as, and often more than, expected cost per event when characterizing true financial exposure.

E. Monte Carlo Simulation and Risk Metrics

The compound model is evaluated via Monte Carlo simulation over 10,000 or more simulated years. For each iteration, a random event count is drawn from the Poisson distribution with constant rate λ; individual event costs are independently sampled from the lognormal severity distribution when $\mathrm{CV} > 0$, or set to the fixed deterministic value $E[X]$ when $\mathrm{CV} = 0$; and total annual loss is computed as the aggregate sum. Events are assumed independent of one another within each simulated year. From the resulting empirical distribution, three primary risk metrics are derived:

- **Expected Annual Loss (EAL):** The long-run average cost, appropriate for budget planning.
- **Value at Risk (VaR):** The loss threshold not exceeded in a specified proportion of years, evaluated at the 90th, 95th, and 99th percentiles [2].
- **Tail Value at Risk (TVaR):** The expected loss conditional on exceeding the VaR threshold, characterizing the severity, not merely the probability, of adverse outcomes [3].

The distinction between EAL and TVaR is critical: a plant may carry a manageable expected annual loss while its TVaR represents a material threat to liquidity in adverse years. Tail risk in the aggregate distribution emerges from two sources: the occurrence of a single high-severity event in the upper tail of the individual loss distribution, and the accumulation of multiple moderate events within a single year when realized event counts exceed the expected rate. Both pathways are captured within the compound model and are invisible to point-estimate budgeting. Because iARM calibrates λ, $E[X]$, and CV from ERP work order and cost records, all outputs are conditioned on the completeness and consistency of that data source. Short history windows, missing cost transactions, inconsistent asset coding, and ERP system migrations represent the primary sources of parameter uncertainty and should be documented alongside model results.

V. The iARM Framework: Four Pillars

A. Pillar One: Quantification

Quantification converts operational variability into measurable financial exposure. iARM derives EAL, VaR, and TVaR from the compound loss distribution and decomposes aggregate exposure to the asset level, identifying each component's contribution to both expected loss and tail risk.

These two contributions are not always proportional. An asset generating frequent, low-severity events may contribute substantially to EAL while contributing minimally to TVaR. An asset generating rare, catastrophic events may contribute modestly to EAL while dominating the tail. Standard maintenance prioritization based on work order frequency or average repair cost will mis-rank these assets systematically. Quantification makes the distinction visible and directs capital toward the sources of true financial exposure rather than operational visibility.

Fig. 1 illustrates Pillar One output from a prototype implementation applied to an oil refinery. With $\lambda = 9.006$ events per year, a per-event mean cost of \$885,488, and CV $= 0.50$, Monte Carlo simulation over 10,000 trials yields EAL $=$ \$7,933,966, $\text{TVaR}_{80} =$ \$12,308,503, and $\text{TVaR}_{95} =$ \$14,778,629. The histogram shows the full simulated annual loss distribution with percentile thresholds marked, confirming that tail exposure substantially exceeds the expected value and that point-estimate budgeting would systematically understate financial risk.

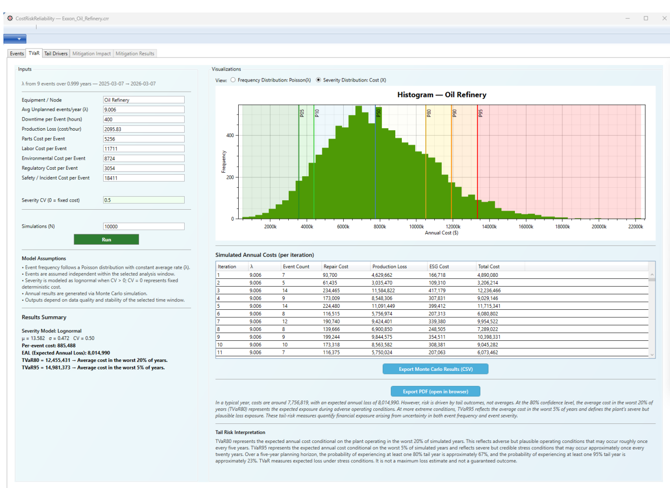

Fig. 1. Pillar One: Quantification. Simulated annual loss distribution for an oil refinery showing EAL, VaR, and TVaR percentile thresholds derived from 10,000 Monte Carlo iterations ($\lambda = 9.006$, CV $= 0.50$, mean severity $=$ \$885,488).

B. Pillar Two: Prioritization

Prioritization ranks assets and failure modes by financial risk contribution rather than historical spend, subjective urgency, or organizational advocacy. Assets are evaluated along four dimensions: (1) contribution to EAL; (2) contribution to tail exposure via TVaR decomposition; (3) sensitivity to parameter changes through scenario analysis on λ and CV; and (4) economic leverage of potential interventions.

This reframes maintenance planning as risk-adjusted portfolio management. Just as a financial portfolio manager distinguishes between assets that contribute to expected return versus tail downside risk, iARM enables plant managers to distinguish between expenditures that reduce average cost and those that reduce catastrophic exposure. A practical consequence is the systematic identification of assets that are under-resourced relative to their financial risk contribution: high-severity, low-frequency assets frequently receive less attention than their risk profile warrants precisely because they have not recently produced a visible failure event.

Financial prioritization is adopted because capital allocation decisions ultimately operate in monetary terms. Safety, environmental, regulatory, and reputational risks are not excluded from the framework; they may be incorporated through cost proxies, scenario modeling, or constraints on the optimization. The goal is decision compatibility, not reductionism.

Fig. 2 shows Pillar Two output from the same refinery. The tail driver analysis identifies that three assets account for 100% of cumulative tail contribution at the 80th percentile, with tail hit rates of 99.7%, 82.1%, and 99.6% respectively. This concentration would not be apparent from average maintenance spend data alone.

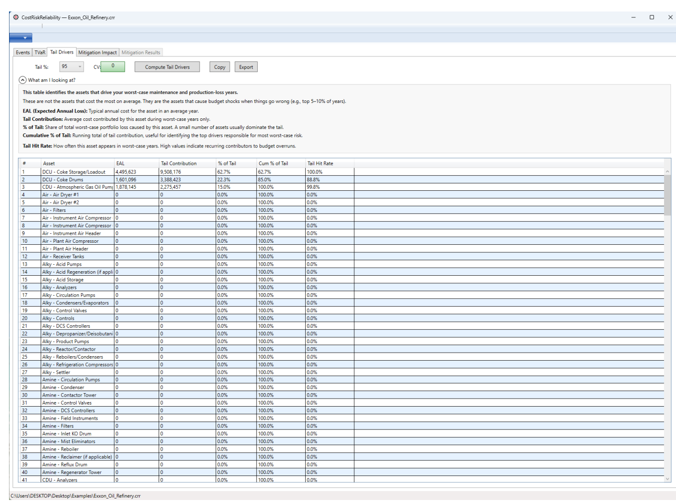

Fig. 2. Pillar Two: Prioritization. Asset-level tail driver analysis for an oil refinery identifying the equipment classes that dominate worst-case annual loss years, ranked by tail contribution, cumulative tail percentage, and tail hit rate.

C. Pillar Three: Mitigation

Mitigation evaluates how proposed interventions reshape the loss distribution and translates that reshaping into financially comparable terms. Interventions may reduce λ through inspection programs or condition monitoring, reduce CV through redundancy investment or consequence mitigation, or reduce downtime duration through spare parts inventory or contractor agreements. The decomposition in equation (4) clarifies which cost components each intervention pathway addresses and how parameter changes propagate through to the aggregate distribution.

Frequency and severity reduction multipliers assume independent mitigation effects. Where multiple controls address the same failure mode, their combined effectiveness is modeled multiplicatively under the assumption of independence. If controls overlap or interact, composite effectiveness may be lower than predicted and dependence structures should be assessed before mitigation ROI estimates are used for capital decisions.

For each intervention, iARM re-runs simulation with adjusted parameters, producing a quantified comparison of pre- and post-intervention distributions. The risk-adjusted return metrics are:

$$\text{ROI} = \frac{\Delta E[S]}{\text{Mitigation Cost}} \tag{9}$$

$$\text{Tail ROI} = \frac{\Delta \text{TVaR}_\alpha}{\text{Mitigation Cost}} \tag{10}$$

where α should be selected to reflect the organization's risk tolerance and, where relevant, align with insurance attachment points or financial covenant thresholds. Tail ROI measures conditional tail exposure reduction per dollar invested rather than realized accounting profit, and frequency reduction strategies will tend to produce broad proportional reductions in EAL while severity reduction strategies tend to produce disproportionate improvements in tail metrics relative to their EAL impact.

For multi-year capital planning, the net present value of exposure reduction is:

$$\text{NPV}_{\text{mit}} = \sum_{t=1}^{T} \frac{\Delta E[S_t]}{(1+r)^t} - C_0 \tag{11}$$

where r is the organization's cost of capital and C_0 is the initial investment. This structure allows mitigation to be evaluated using the same financial discipline applied to production expansions or strategic capital projects, enabling reliability investment to compete on equal terms in any capital allocation process.

Fig. 3 shows the batch mitigation configuration for the same refinery. Two interventions are evaluated for the CDU Atmospheric Gas Oil Pumps: a pump upgrade reducing frequency by 25% and severity by 56% at a cost of $250,000 per year, and a vibration monitoring route reducing frequency by 13% and severity by 44% at $50,000 per year.

D. Pillar Four: Communication

Communication translates probabilistic output into the vocabulary of enterprise risk management. EAL maps to budget planning. VaR maps to capital reserve and contingency planning. TVaR maps to insurance limit selection and self-insured retention decisions. Loss volatility maps to earnings stability analysis and financial covenant management.

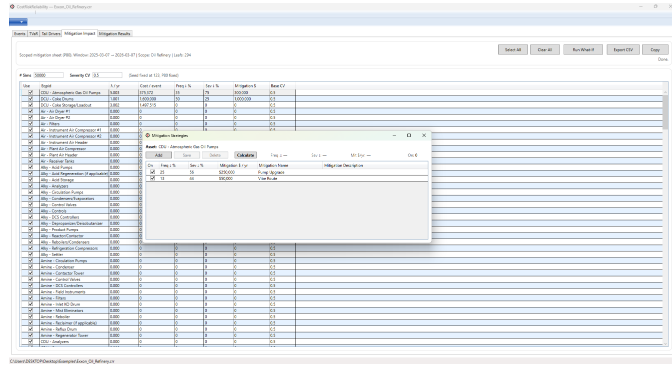

Fig. 3. Pillar Three: Mitigation. Batch mitigation configuration showing asset-level frequency and severity reduction parameters and associated mitigation costs for the refinery portfolio.

When reliability engineers express recommendations in terms of quantified exposure reduction rather than operational preference, those proposals compete on equal terms with any capital investment in the organization's portfolio. The alignment also creates accountability in both directions: reliability engineers are accountable for quantifying the financial exposure their recommendations address, and financial decision-makers are accountable for acknowledging the risk they accept when they decline to fund a recommended intervention.

Communication also supports insurance structuring: plants that can present their aggregate loss distribution with documented frequency and severity assumptions are better positioned to obtain coverage reflecting their actual risk profile, since the actuarial vocabulary of iARM is precisely the vocabulary insurance underwriters use.

Fig. 4 presents the Pillar Four output for the refinery mitigation plan. The proposed interventions reduce mean annual cost by $7,964,585 VaR_{80} by $1,246,186, and TVaR_{80} by $1,517,137, at a total mitigation cost of $1,00,000, yielding a Tail ROI of 1.52. These metrics are expressed in the financial vocabulary of enterprise risk governance and are directly actionable by capital committees without further translation.

VI. Strategic Implications

A. Reframing Maintenance as Risk Management

iARM reframes maintenance not as a cost-control function but as portfolio risk management within the plant microeconomy. The central question shifts from *how often does this asset fail* to *what is this asset's contribution to aggregate financial volatility and tail exposure*. Assets that appear stable under average performance metrics may dominate tail risk when their severity distribution is examined carefully. Conversely, assets that generate frequent but low-consequence failures may consume disproportionate operational attention and maintenance budget without materially affecting the plant's capital-at-risk or tail exposure profile.

B. Capital Allocation Discipline

Rather than allocating capital based on recent failure history, organizational advocacy, or historical spend inertia,

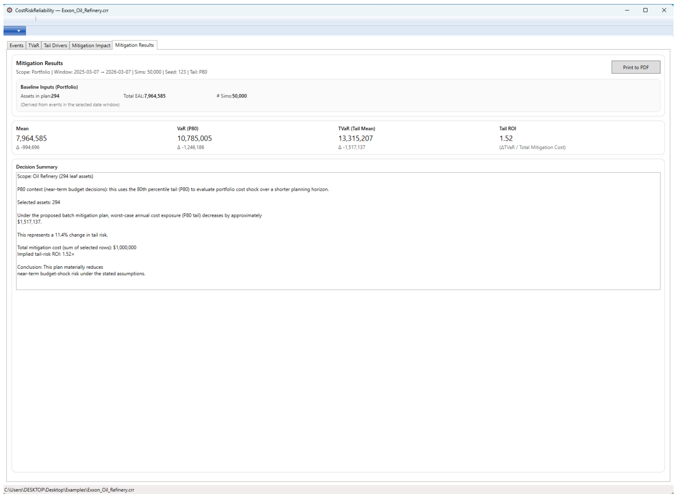

Fig. 4. Pillar Four: Communication. Portfolio-level mitigation impact showing reductions in mean cost, VaR, and TVaR under the proposed plan, with an implied Tail ROI of 1.52 on a total mitigation investment of $1,000,000.

iARM introduces a defensible alternative: allocate according to marginal reduction in financial exposure per dollar invested, across the full distribution rather than at the expected value alone. The capital efficiency of a mitigation investment is measured by the ratio of exposure reduction it produces per dollar spent. This creates a genuine portfolio perspective enabling capital committees to compare reliability investments against competing uses of capital with analytical rigor.

C. Volatility-Aware Budget Planning

The compound model demonstrates that actual annual costs can deviate substantially from expected values, by 50%, 75%, or even 100%, not due to management failure but because volatility is a mathematical property of the loss structure as shown in equation (8). By quantifying VaR and TVaR alongside EAL, iARM provides the analytical basis to align contingency reserves with modeled exposure rather than historical convention, reducing the frequency and magnitude of surprise-driven capital expenditure.

D. Aligning Operational and Financial Governance

The translation gap between engineering analysis and financial decision-making is primarily a gap in analytical language, not intelligence or expertise. iARM establishes shared language by expressing operational risk in terms native to enterprise risk governance. When reliability recommendations are expressed in quantified exposure reduction and financial decisions explicitly acknowledge the operational risk being accepted or rejected, the conversation becomes analytical rather than political, resolvable through shared evidence rather than organizational authority.

E. Institutionalizing Probabilistic Thinking

The most consequential long-term implication of iARM adoption may be cultural. Industrial organizations that implement the framework consistently develop a fundamentally different relationship with uncertainty. Reliance on anecdotal or recent experience declines, replaced by distributional analysis. Tail risk receives explicit recognition rather than being dismissed as unlikely until it occurs. And declining a mitigation investment becomes an explicit acknowledgment of a quantified exposure level rather than an abstract budget choice with no stated risk consequence.

VII. Conclusion

iARM does not eliminate operational risk; no analytical framework can. What it does is replace implicit, unquantified risk acceptance with explicit, measured risk management grounded in a rigorous characterization of exposure. By treating total annual plant cost as a compound stochastic process, decomposing severity into its operational cost components, and deriving actionable financial metrics from the resulting distribution, iARM gives reliability engineers a financial vocabulary for risks they already understand technically, gives financial officers a quantitative basis for decisions currently made on instinct, and gives plant managers a coherent framework for holding both disciplines accountable to the same picture of economic reality.

The transition is not from managing cost to eliminating risk, but from reacting to failures toward governing quantified exposure with deliberate capital discipline.

Acknowledgment

The author thanks the practitioners and engineers whose field experience informed the development of the iARM methodology.

References

[1] S. A. Klugman, H. H. Panjer, and G. E. Willmot, *Loss Models: From Data to Decisions*, 4th ed. Hoboken, NJ: Wiley, 2012.
[2] A. J. McNeil, R. Frey, and P. Embrechts, *Quantitative Risk Management: Concepts, Techniques and Tools*, Revised ed. Princeton, NJ: Princeton University Press, 2015.
[3] P. Artzner, F. Delbaen, J. Eber, and D. Heath, "Coherent measures of risk," *Mathematical Finance*, vol. 9, no. 3, pp. 203–228, 1999.
[4] M. Rausand and A. Høyland, *System Reliability Theory: Models, Statistical Methods, and Applications*, 2nd ed. Hoboken, NJ: Wiley, 2004.
[5] D. F. Mango, "Applying actuarial techniques in operational risk modeling," presented at the Enterprise Risk Management Symposium, Society of Actuaries, Chicago, IL, Apr. 2006.