Vine Copula and EVT Jittering for Operational Risk Capital

CHRISTIANO M B LOPES[a], ANDRÉ N MARANHÃO[a], BENJAMIM M TABAK[b]

[a]Escola de Economia de São Paulo, Fundação Getúlio Vargas (FGV/EESP)

[b]Escola de Políticas Públicas e Governo, Fundação Getúlio Vargas (FGV/EPPG)

February 2026

Abstract

Operational risk quantification traditionally assumes independence across categories, underestimating stress-period losses. We challenge this paradigm by developing a unified framework addressing discrete-continuous loss data, heavy-tailed severity, and complex risk dependence. Analyzing 1.99 million operational loss events from a major Brazilian bank (2020–2024), we combine Extreme Value Theory (EVT) for severity tails, spliced distributions, and Regular Vine (R-Vine) copulas with jittering for frequency dependence. Our results show dependence increases economic capital at the 99.9% level by 42.87% over five years. Backtesting reveals the independence model systematically fails during stress periods, while the Vine copula model maintains coverage. These findings confirm that presumed diversification benefits are often illusory, though negative dependencies can occasionally reduce aggregate risk.

Keywords: Operational Risk; Economic Capital; Vine Copulas; Extreme Value Theory; Jittering; Spliced Distributions; Dependence Modeling.

JEL Classification: C15, C51, G21, G28, G32.

Contents

1 Introduction

The global banking system accumulated €640 billion in operational losses between 2019 and 2024, with an average annual loss of €22.2 billion across 70,690 events (ORX, 2025). These figures underscore a fundamental challenge: operational risk, defined by the Basel Committee as the risk of loss from inadequate or failed internal processes, people, systems, or external events (BCBS, 2006), has proven to be a persistent source of financial instability, often exceeding credit and market risk losses during crises (Cruz, 2002; Chernobai et al., 2007). In emerging economies such as Brazil, the combination of legacy systems, political pressures, labor litigation, and volatile macroeconomic conditions creates a uniquely challenging environment where extreme events emerge with simultaneous vulnerabilities (Carvalho, 2014; Micco et al., 2007).

The Loss Distribution Approach (LDA), which separately models frequency and severity before convolution into aggregate loss distributions, has become the dominant actuarial framework for operational risk measurement (Frachot et al., 2001; Shevchenko, 2011). However, traditional LDA implementations face three interconnected challenges that compromise capital accuracy.

First, data sparsity or zero inflation characterize operational loss series. Many risk categories contain insufficient observations for reliable statistical inference, and days without losses are prevalent. Applying complex models to sparse categories produces unstable parameter estimates and unreliable capital figures. For example, categories with limited events often yield GPD shape parameters that are either negative (implying bounded losses) or implausibly high (leading to infinite mean models). This necessitates a robust segmentation strategy that treats sparse categories deterministically while applying full stochastic modeling only to data-rich categories.

Second, severity heterogeneity defies simple parametric specification. Losses exhibit positive skewness and heavy tails that single distributions cannot adequately capture. While Extreme Value Theory (EVT) provides asymptotic justification for modeling tails via the Generalized Pareto Distribution (GPD) (McNeil et al., 2015; Embrechts et al., 1997), EVT alone neglects the body of the distribution essential for day-to-day risk management. Spliced distributions that combine parametric bodies (Lognormal, Weibull, Gamma, Exponential) with GPD tails offer a solution, but threshold selection and model validation remain a subject of ongoing debate (Scarrott and MacDonald, 2012; Dutta and Perry, 2006). The choice of threshold involves a bias-variance trade-off: too low a threshold introduces bias by including body observations in tail estimation, while too high a threshold increases variance due to limited exceedances. Furthermore, model selection must prioritize tail accuracy, as capital at 99.9% confidence depends critically on the upper quantile.

Third, and most critically, dependence across operational risk categories violates the independence assumptions embedded in many aggregation approaches. In stress periods,

internal fraud may correlate with systems failures, and external events may trigger process breakdowns. Linear correlation (Pearson) is inadequate for capturing tail dependence, which is the tendency for extreme losses to occur simultaneously (Embrechts et al., 2002). On the other hand, copulas provide a flexible framework for modeling multivariate dependence by separating marginal behavior from the dependence structure (Nelsen, 2006; Joe, 2014). However, applying copulas to frequency data faces a fundamental mathematical challenge: Sklar's theorem guarantees a unique copula only for continuous margins. Operational loss frequencies are discrete count variables, creating identification problems in regions of the domain (Genest and Nešlehová, 2007). The presence of ties makes the copula non-unique and can lead to biased dependence estimates.

This paper develops a unified framework that addresses these three challenges simultaneously. Our methodological contributions are fourfold.

First, we introduce a robustness segmentation strategy that filters risk categories based on data density. Categories meeting minimum criteria (30 events and 2 distinct days) proceed to full stochastic LDA modeling. Those with insufficient data receive deterministic capital add-ons based on worst-case historical scenarios. This prevents contamination of economic capital by estimation errors from sparse categories while ensuring rare but severe risks are not excluded.

Second, we implement an adaptive severity architecture using spliced distributions. For each eligible category, we test four body distributions (Lognormal, Weibull, Gamma, and Exponential) against spliced alternatives combining the best-fitting body with a GPD tail. To define the GPD threshold, we evaluate three distinct approaches: fixed thresholds at the 90th and 95th percentiles, and a dynamic threshold determined via a data-driven grid search within this range. The final model selection across all candidates is guided by the Anderson-Darling (AD) criterion. The AD statistic, which weights tail deviations more heavily than Kolmogorov-Smirnov (Stephens, 1974), determines the final model, ensuring that capital at 99.9% confidence rests on accurate tail specification. We also impose a robustness check on the GPD shape parameter: if $\xi < 0$ (bounded tail) or $\xi > 1.5$ (infinite mean/variance), the spliced model is discarded in favor of the best single (body-only) distribution.

Third, we resolve the discreteness problem via the jittering technique (Denuit and Lambert, 2005). By adding controlled uniform noise to discrete frequency observations, we restore the continuity required for copula application while preserving Kendall's tau (τ) dependence structure. For each observation n, we generate $V \sim U(0,1)$ and compute $u = F(n-1) + V \cdot [F(n) - F(n-1)]$, where F is the fitted marginal distribution. This transformation yields continuous pseudo-observations that can be used for copula estimation, and after simulation, we recover discrete frequencies by taking integer parts (Machado and Santos Silva, 2005).

Fourth, we employ R-Vine copulas (Aas et al., 2009; Joe, 2014; Czado, 2019) to capture complex, asymmetric dependence structures. Unlike elliptical copulas (Gaussian, t-Student) that impose symmetric dependence, or single-parameter Archimedean families that restrict tail

behavior, Vine copulas decompose multivariate density into a cascade of bivariate pair-copulas, each selected independently from families including Clayton, Gumbel, Frank, Joe, and their rotations. This flexibility allows the model to capture, for example, upper tail dependence between external fraud and systems failures while maintaining independence or negative dependence elsewhere. The R-Vine structure is selected using the algorithm of Dissmann et al. (2013), which constructs a maximum spanning tree based on Kendall's tau at each level, and pair-copula families are chosen by minimizing AIC.

We apply this framework to a proprietary dataset of 1.99 million operational loss events from a major Brazilian public bank covering 2020 to 2024. The dataset includes 17 Level-2 risk categories under the Central Bank of Brazil's Circular 3,979/2020 taxonomy, providing granular coverage of the institution's risk profile.

Substantial heterogeneity emerges in the optimal severity specifications across the anonymized event categories. The most material categories require spliced Lognormal-GPD models, with shape parameters (ξ) ranging from 0.32 to 0.82, confirming a pronounced heavy-tailed behavior. For instance, Category C exhibits $\xi = 0.320$, Category H shows $\xi = 0.421$, and Category P reaches $\xi = 0.820$. Conversely, less severe categories are adequately captured by single distributions, such as the Lognormal.

The estimated R-Vine structure reveals complex dependence patterns across standard operational risk categories. The first-tree dependencies (strongest unconditional relationships) include strong positive associations among major event types, such as Category H, C and B, captured by Frank and BB8 copulas with Kendall's tau reaching up to 0.61 and 0.44. Negative dependencies also emerge in higher-order trees, showing conditional relationships (e.g., involving Category E) modeled by Frank copulas with τ around -0.12. The overall model has a log-likelihood of 4,553.03 and an AIC of –8,988.06 for the 2020–2024 window, confirming significant multivariate dependence beyond what independence would imply.

Monte Carlo simulations (1 million daily scenarios, bootstrapped to 2 million annual aggregates) demonstrate that dependence increases economic capital at 99.9% confidence by 42.87% for the full five-year window compared to the independence benchmark. The Expected Shortfall (ES) differential is even larger at 58.94%. The impact varies across windows, from 6.73% in 2024 to 18.53% in 2020, with the cumulative windows showing even larger effects.

Backtesting underscores the practical relevance of dependence modeling. In three windows, the independence model generated capital insufficient to cover actual losses: 2021 (coverage 0.87x), 2024 (coverage 0.99x), and 2023–2024 (coverage 0.96x). In contrast, the Vine copula model satisfies coverage in all windows, with ratios ranging from 1.01 to 3.93. Notably, the 2021–2024 window exhibits subadditivity (Vine < Independence), with a negative impact of –67.91%, indicating that negative dependencies can sometimes reduce aggregate risk. This finding highlights the complexity of dependence structures and the need for flexible models like Vine copulas. Although superadditivity dominates across most windows, the repeated shortfalls under independence demonstrate that presumed diversification benefits may be illusory when

dependence is ignored.

These findings have significant implications. For risk managers, sophisticated dependence modeling is not a theoretical refinement but a prudential necessity. Banks using independence assumptions for economic capital may be undercapitalized by hundreds of millions or billions of reais relative to their true risk profile. For regulators, these findings reinforce the necessity of Pillar 2 internal models as complements to standardized approaches, given that one-size-fits-all capital charges fail to account for institution-specific dependence structures. For researchers, the framework provides a template for integrating jittering, spliced distributions, and Vine copulas in operational risk contexts, and opens avenues for further research on time-varying dependence, frequency-severity dependence, and emerging risks.

The remainder of this paper is organized as follows. Section 2 reviews the relevant literature on operational risk measurement, EVT, and copula modeling. Section 3 presents our methodology in detail, including robustness segmentation, severity modeling, jittering, R-Vine estimation, and Monte Carlo simulation. Section 4 describes the dataset and empirical results, including model selection, dependence structure, capital calculations, backtesting, and convergence analysis. Section 5 discusses implications, limitations, and future research directions. Section 6 concludes.

2 Literature Review

2.1 Operational Risk: Definition and Regulatory Evolution

The Basel Committee on Banking Supervision (BCBS, 2006) established the prevailing global standard by defining operational risk as the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events. This definition encompasses four primary sources: internal processes (procedural failures, inadequate controls), people (human error, internal fraud, insufficient training), systems (IT failures, cyber attacks, hardware malfunctions), and external events (natural disasters, terrorism, regulatory changes). Legal risk is included, but strategic and reputational risk are excluded due to their different nature and measurement challenges.

Operational risk's formal recognition as a distinct risk category dates to Basel II (2004), which introduced a spectrum of measurement approaches: the Basic Indicator Approach (BIA), the Standardized Approach (TSA), and the Advanced Measurement Approach (AMA). The AMA permitted banks to develop internal models for regulatory capital calculation, subject to supervisory approval, catalyzing a generation of research on loss distribution modeling (Cruz, 2002; Frachot et al., 2001). Under AMA, banks could use internal loss data, external data, scenario analysis, and business environment factors to quantify risk, provided they met rigorous qualitative and quantitative standards.

Basel III, finalized in 2017, replaced the AMA with a revised Standardized Approach

(SA) for regulatory minimum capital (BCBS, 2017). The SA combines a Business Indicator Component (based on financial statement proxies) with an Internal Loss Multiplier that reflects historical losses. However, as the Basel framework emphasizes, internal models remain essential for Pillar 2 (Internal Capital Adequacy Assessment Process – ICAAP) and for economic capital calculation reflecting the institution's own risk profile (Crouhy et al., 2014). The Brazilian Central Bank's Resolution CMN 356/2023 incorporates these standards while permitting internal models for management purposes from 2025.

Taxonomy is fundamental to operational risk measurement. The Basel Committee's original seven-category taxonomy (BCBS, 2006) has been adapted by national regulators. In Brazil, Circular 3,979/2020 establishes eight Level-1 categories and multiple Level-2 subcategories aligned with international standards while accommodating local specificities. Table 1 presents this taxonomy.

Table 1: Operational Risk Taxonomy under BCB Circular 3,979/2020

Level-1 Category	Level-2 Category
1. Internal Fraud	1.1 Unauthorized activity
	1.2 Theft and fraud (internal)
2. External Fraud	2.1 Theft and fraud (external)
	2.2 Systems security
3. Employment Practices and Workplace Safety	3.1 Employee relations
	3.2 Workplace safety
	3.3 Diversity and discrimination
4. Clients, Products, and Business Practices	4.1 Suitability and disclosure
	4.2 Improper business practices
	4.3 Product flaws
	4.4 Selection and sponsorship
	4.5 Advisory activities
5. Damage to Physical Assets	5.1 Disasters and other events
6. Business Disruption	6.1 Systems disruptions
7. Technology and Infrastructure Failures	7.1 IT systems, processes, or infrastructure failures
8. Execution, Delivery, and Process Management	8.1 Transaction capture and execution
	8.2 Monitoring and reporting
	8.3 Customer intake and documentation
	8.4 Account management
	8.5 Counterparties
	8.6 Vendors and suppliers

Source: Central Bank of Brazil (BCB), Circular 3,979/2020.

2.2 The Loss Distribution Approach (LDA)

The LDA framework models total aggregate loss S over a fixed horizon (typically one year) as:

$$S = \sum_{i=1}^{N} X_i \tag{1}$$

where N is a discrete random variable representing frequency, and $\{X_i\}$ are independent and identically distributed (i.i.d.) severity random variables, assumed to be independent of N (Frachot et al., 2001; Klugman et al., 2012). This structure is analogous to collective risk models in actuarial science and has been widely adopted in operational risk due to its intuitive separation of occurrence and impact.

The cumulative distribution function $F_S(s)$ is obtained by convolution of frequency and severity distributions. Closed-form solutions are generally unavailable, necessitating numerical methods. Historically, Panjer recursion or Fast Fourier Transform (FFT) were employed (Panjer, 2006), while contemporary implementations rely on Monte Carlo simulation due to its flexibility and ability to incorporate dependence structures (Glasserman, 2003). Monte Carlo methods allow for the generation of thousands or millions of scenarios, approximating the loss distribution with arbitrary precision, and are particularly valuable when dealing with complex dependence patterns.

Economic capital is quantified as Value-at-Risk (VaR) at confidence level α (99.9% under Basel):

$$\mathrm{VaR}_\alpha(S) = F_S^{-1}(\alpha) = \inf\{s \in \mathbb{R} : F_S(s) \geq \alpha\} \tag{2}$$

VaR represents the minimum loss that could occur in the worst $\alpha\%$ of scenarios, providing a simple and intuitive risk measure. However, VaR fails the subadditivity property (Artzner et al., 1999). Expected Shortfall (ES) addresses this limitation by considering the average loss exceeding the VaR, thereby satisfying the coherence axioms:

$$\mathrm{ES}_\alpha(S) = \mathbb{E}[S|S \geq \mathrm{VaR}_\alpha(S)] \tag{3}$$

While VaR captures the maximum expected loss within a specific time horizon given a confidence level α, ES captures the expected loss in exceedance scenarios, offering additional insight into tail severity and satisfying coherence properties (McNeil et al., 2015). Both measures are reported in this study.

2.3 Frequency Modeling

Frequency distributions model the number of loss events per time period (in this study, we adopt a daily discretization). Two distributions dominate the literature.

The Poisson distribution, with probability mass function:

$$P(N = n) = \frac{e^{-\lambda}\lambda^n}{n!}, \quad n = 0, 1, 2, \ldots \tag{4}$$

is characterized by a single parameter λ representing both mean and variance (equidispersion). Its simplicity makes it a natural benchmark, but operational loss frequencies often exhibit overdispersion (variance exceeding the mean) violating the equidispersion

assumption (Chernobai et al., 2007). Overdispersion arises from unobserved heterogeneity, contagion effects, or clustering of losses.

The Negative Binomial distribution accommodates overdispersion by introducing an additional parameter. It can be derived as a Poisson mixture where the rate parameter λ follows a Gamma distribution, yielding:

$$P(N = n) = \binom{n + r - 1}{n} p^r (1 - p)^n, \quad n = 0, 1, 2, \ldots \tag{5}$$

with $r > 0$ (size parameter) and $p \in (0, 1)$. The variance $r(1 - p)/p^2$ exceeds the mean $r(1 - p)/p$, capturing excess variability. In operational risk, Negative Binomial models often provide superior fit due to the clustering of losses in certain periods (Chernobai et al., 2007).

2.4 Severity Modeling and Extreme Value Theory

Severity distributions in operational risk are characterized by positive skewness and heavy tails, with extreme losses occurring with higher probability than predicted by light-tailed distributions such as the normal or exponential. The heavy-tailed nature stems from the potential for catastrophic events: internal frauds can involve hundreds of millions, legal settlements can reach billions, and systems failures can disrupt entire business lines.

Extreme Value Theory (EVT) provides the asymptotic foundation for modeling tail behavior. The Peaks-Over-Threshold (POT) approach, based on the Pickands-Balkema-de Haan theorem (Balkema and de Haan, 1974; Pickands, 1975), states that for a sufficiently high threshold u, the distribution of excesses $Y = X - u \mid X > u$ converges to the Generalized Pareto Distribution (GPD):

$$G_{\xi,\beta}(y) = \begin{cases} 1 - (1 + \frac{\xi y}{\beta})^{-1/\xi}, & \xi \neq 0 \\ 1 - e^{-y/\beta}, & \xi = 0 \end{cases} \tag{6}$$

where $\xi \in \mathbb{R}$ is the shape parameter and $\beta > 0$ is the scale parameter. For $\xi > 0$, the GPD has heavy tails (Fréchet type), appropriate for operational risk where losses have no theoretical upper bound. For $\xi = 0$, the GPD reduces to the exponential distribution (Gumbel type), and for $\xi < 0$, it has a bounded tail (Weibull type), which is rarely appropriate for operational risk.

While EVT captures tail behavior, it neglects the body of the distribution. Spliced (or composite) distributions address this limitation by combining a parametric body distribution with a GPD tail (Scarrott and MacDonald, 2012; Bakar et al., 2015). The cumulative distribution function is:

$$F(x) = \begin{cases} \phi \cdot \dfrac{F_{\text{body}}(x)}{F_{\text{body}}(u)}, & x \leq u \\ \phi + (1 - \phi) \cdot G_{\xi,\beta}(x - u), & x > u \end{cases} \tag{7}$$

where $\phi = F(u)$ is the empirical probability below threshold u, and F_{body} is the body distribution (e.g., Lognormal, Weibull, Gamma, Exponential) fitted to all observations but truncated at u. This ensures a smooth transition at the threshold and a proper probability density.

Threshold selection remains a debate in EVT applications. A too low threshold introduces bias by including body observations in tail estimation. A too high threshold increases estimator variance due to limited exceedances (Scarrott and MacDonald, 2012). Although Dutta and Perry (2006) utilized fixed thresholds of 5% and 10% of the largest losses (the 95th and 90th percentiles) as empirical benchmarks for operational risk, these levels may not optimally fit every risk profile. In this study, we implement a systematic data-driven grid search to endogenously determine the optimal threshold u_k for each risk category. The search space spans from the 90th to the 95th percentile at 1% increments, following the upper range of the literature benchmarks. We also estimate alternative specifications locked at the traditional 90th and 95th percentiles, providing a full sensitivity analysis of the tail stability in Appendix E.

Model selection for severity distributions must prioritize tail accuracy, as capital requirements at the 99.9% confidence level depend critically on the behavior of the upper quantiles. To this end, we employ the Anderson-Darling (AD) statistic as the primary goodness-of-fit criterion. Unlike the Kolmogorov-Smirnov test, which focuses on the maximum distance between distributions, the AD statistic weights deviations in the tails more heavily by using the following computational form (Stephens, 1974):

$$AD = -n - \frac{1}{n}\sum_{i=1}^{n}\left[(2i-1)\ln(F(X_i)) + (2n-2i+1)\ln(1-F(X_i))\right] \tag{8}$$

where $X_1 \leq X_2 \leq \cdots \leq X_n$ are the ordered observations. In our framework, this statistic serves a dual purpose: as an objective function within a data-driven grid search to identify the optimal threshold u_k; and as the final selection metric to compare body-only distributions against optimized spliced models. The candidate that minimizes the global (body and tail) AD value (lower AD) is selected to represent the risk profile.

2.5 Dependence Modeling and Copulas

The aggregation of risks across categories requires understanding their interdependence. Linear correlation (Pearson) is inadequate for operational risk because it assumes elliptical distributions and cannot capture tail dependence (Embrechts et al., 2002). Tail dependence measures the probability that two variables jointly experience extreme events, it is a crucial feature in stress scenarios.

Copulas provide a flexible framework for modeling multivariate dependence independent of marginal distributions. Sklar's theorem (Sklar, 1959) states that for any random vector $(X_1, \ldots, X_d)$ with continuous marginal distributions $F_1, \ldots, F_d$, there exists a unique copula

$C : [0,1]^d \to [0,1]$ such that the joint distribution H satisfies:

$$H(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)) \tag{9}$$

Conversely, given any copula C and marginal distributions F_i, the function H is a valid joint distribution. This separation allows independent specification of marginal behavior and dependence structure (Nelsen, 2006; Joe, 2014).

Copula families differ in their tail dependence properties. Let λ_U and λ_L denote upper and lower tail dependence coefficients:

$$\lambda_U = \lim_{u \to 1^-} P[U_2 > u \mid U_1 > u], \quad \lambda_L = \lim_{u \to 0^+} P[U_2 \leq u \mid U_1 \leq u] \tag{10}$$

Major families include:

- Elliptical copulas: Gaussian ($\lambda_U = \lambda_L = 0$) and t-Student ($\lambda_U = \lambda_L > 0$, symmetric tail dependence)
- Archimedean copulas: Clayton ($\lambda_L > 0$, $\lambda_U = 0$), Gumbel ($\lambda_U > 0$, $\lambda_L = 0$), Frank ($\lambda_U = \lambda_L = 0$), Joe ($\lambda_U > 0$, $\lambda_L = 0$, stronger upper tail than Gumbel)
- Rotated versions: Rotating Clayton 180° yields a copula with upper tail dependence; rotating Gumbel 90° or 270° captures negative dependence

For high-dimensional problems, Vine copulas (also called pair-copula constructions, PCC) overcome the rigidity of single-parameter families (Joe, 1996; Aas et al., 2009). A d-dimensional density is decomposed into a product of $d(d-1)/2$ bivariate copulas, organized in hierarchical trees. Regular Vines (R-Vines) represent the most general class, with C-Vines (canonical, star structure) and D-Vines (drawable, path structure) as special cases (Bedford and Cooke, 2001, 2002; Dissmann et al., 2013; Czado, 2019).

The R-Vine density for d variables can be expressed as:

$$f(x_1, \ldots, x_d) = \prod_{k=1}^{d} f_k(x_k) \times \prod_{t=1}^{d-1} \prod_{e \in E_t} c_{j_e, k_e | D_e}\left(F_{j_e|D_e}(x_{j_e} \mid x_{D_e}), F_{k_e|D_e}(x_{k_e} \mid x_{D_e}) \,\middle|\, x_{D_e}\right) \tag{11}$$

where E_t denotes the edge set of tree t, and D_e is the conditioning set. This construction allows each pair-copula to be selected independently from different families, capturing heterogeneous dependence patterns.

Model selection involves three steps: (i) tree structure selection based on maximum spanning tree algorithm using Kendall's tau as weights; (ii) pair-copula family selection for each edge minimizing AIC; (iii) parameter estimation via sequential maximum likelihood

(Dissmann et al., 2013). The sequential approach first estimates the first tree, then conditional distributions, and proceeds recursively.

2.6 The Discreteness Problem and Jittering

Applying copulas to frequency data faces a fundamental challenge: Sklar's theorem guarantees copula uniqueness only for continuous margins (Genest and Nešlehová, 2007). Discrete margins produce ties in the probability integral transform, and the copula is not identifiable on regions of the domain. This can lead to biased estimates of dependence, particularly for tail dependence measures.

The jittering approach (Denuit and Lambert, 2005) resolves this issue by adding controlled uniform noise to discrete observations. For a discrete variable N with distribution function F, define:

$$N^* = N - 1 + V, \quad V \sim U(0,1) \tag{12}$$

The transformed variable N^* has continuous distribution function:

$$F^*(x) = F(\lfloor x \rfloor - 1) + (x - \lfloor x \rfloor + 1)[F(\lfloor x \rfloor) - F(\lfloor x \rfloor - 1)] \tag{13}$$

This transformation preserves the dependence structure (specifically Kendall's tau) while enabling copula estimation on the continuous scale. After estimation, simulations are transformed back by taking integer parts (Machado and Santos Silva, 2005). The jittering approach has been successfully applied in insurance and finance to handle count data in multivariate settings.

2.7 Related Empirical Literature

Several studies have applied advanced dependence modeling to operational risk. Brechmann et al. (2014) used C-Vine copulas to model dependence across business lines in a German bank, finding that dependence increases total capital by 12–18% compared to independence. Bee and Hambuckers (2022) proposed g-and-h marginals with Vine copulas for seven-dimensional operational loss data, demonstrating improved fit over single-copula alternatives.

Khorrami Chokami and Rabitti (2025) analysed UniCredit Bank data using a combination of EVT, Vine copulas, and Shapley values for sensitivity analysis, identifying key drivers of extreme losses. To overcome the challenge of limited observations, they implemented a data augmentation strategy via copulas to identify the key drivers of extreme losses. These studies confirm the relevance of multivariate dependence for operational risk measurement.

In the Brazilian context, the joint application of robustness-based segmentation, spliced severity models with AD selection, jittered frequency modeling, and R-Vine copulas has not

been examined in an integrated framework using comprehensive banking data. The present study contributes to this literature.

3 Methodology

Our methodology integrates six components: (i) robustness segmentation and data filtering, (ii) adaptive severity modeling with spliced distributions, (iii) frequency modeling with jittering, (iv) R-Vine copula estimation for dependence structure, (v) Monte Carlo simulation and capital calculation, and (vi) validation via backtesting and convergence analysis.

3.1 Robustness Segmentation

Operational loss databases contain categories with varying data density. Applying complex stochastic models to sparse categories risks estimation errors contaminating capital figures. We implement a segmentation strategy based on two criteria:

- Minimum events: ≥ 30 loss events in the estimation window
- Minimum distinct occurrence days: ≥ 2 distinct days

The threshold of 30 events ensures that maximum likelihood estimation is feasible and stable, while the requirement of at least 2 distinct days prevents categories where all losses occurred on a single day, which would provide no information about daily frequency variation.

Categories meeting both criteria proceed to full LDA modeling. Categories failing either criterion receive deterministic capital add-ons equal to the worst-case annual loss observed historically. This approach ensures that rare but severe events are not excluded while preventing unstable stochastic estimates. For example, a category with only 5 events but a single large loss of R$43.89 million would receive an add-on of that amount, rather than attempting to model its distribution with insufficient data.

We also apply a materiality threshold of R$1,000, excluding losses below this amount. This aligns with international practice (Moscadelli, 2004; Dutta and Perry, 2006) and focuses modeling on economically relevant risks. Losses below this threshold are considered operational noise and do not materially affect capital.

3.2 Severity Modeling: Spliced Distributions with AD Selection

For each eligible category, we fit candidate distributions to positive severity values:

- Single (Body-only) candidates: Lognormal, Weibull, Gamma, Exponential
- Spliced candidates: Body distribution combined with GPD tail above threshold u (from 90th to 95th) percentile

The GPD is estimated by maximum likelihood on exceedances. Let n_u be the number of exceedances, $y_j = x_j - u$ for $x_j > u$. The log-likelihood is:

$$\ell(\xi, \beta; y_1, \ldots, y_{n_u}) = -n_u \log \beta - (1 + 1/\xi) \sum_{j=1}^{n_u} \log \left(1 + \frac{\xi y_j}{\beta}\right) \tag{14}$$

for $\xi \neq 0$, with the constraint $1 + \xi y_j/\beta > 0$ for all j.

We impose a robustness check on the shape parameter: if $\xi < 0$ (indicating bounded tail) or $\xi > 1.5$ (indicating infinite mean/variance, producing implausible capital), the spliced model is discarded in favor of the best body-only distribution. This prevents the model from producing economically unreasonable capital figures. For example, $\xi > 1$ implies infinite mean, which would lead to infinite capital in extreme quantiles.

Model selection between body-only and spliced alternatives is based on the Anderson-Darling statistic calculated on the full dataset. The AD statistic's quadratic weighting of tail deviations ensures that capital at 99.9% is based on accurately specified tails. We select the model with lower AD, provided the spliced model passes the shape parameter check.

3.3 Frequency Modeling and Jittering

For each eligible category, we model daily frequency counts N_t ($t = 1, \ldots, T$) using either Poisson or Negative Binomial distributions. Selection between them is based on the dispersion index:

$$\text{Dispersion} = \frac{\text{Var}(N_t)}{\text{Mean}(N_t)} \tag{15}$$

If Dispersion > 1.2 (overdispersion threshold), we select Negative Binomial; otherwise, Poisson. The threshold of 1.2 is chosen to be conservative; in practice, operational risk frequencies are almost always overdispersed due to clustering.

Parameters are estimated by maximum likelihood, with method of moments as fallback if MLE fails to converge. Method of moments for Negative Binomial uses sample mean and variance to solve for r and p.

To enable copula modeling, we apply the jittering transformation. For each observation n, generate $V \sim U(0, 1)$ independently and compute:

$$u = F(n-1) + V \cdot [F(n) - F(n-1)] \tag{16}$$

where F is the fitted marginal distribution function. The transformed values u are continuous in $[0, 1]$ and constitute the pseudo-observations for copula estimation. This transformation is applied separately for each category and day.

3.4 R-Vine Copula Estimation

Let $u_{i,t}$ be the jittered pseudo-observations for category i on day t. We model the multivariate dependence structure using R-Vine copulas, estimated with the algorithm of Dissmann et al. (2013) as implemented in the VineCopula R package.

The estimation proceeds sequentially:

Tree 1: Select the maximum spanning tree that maximizes the sum of absolute Kendall's tau values between pairs. For each selected edge, choose the optimal bivariate copula family minimizing AIC among candidates: Gaussian, t-Student, Clayton (and 90°, 180°, 270° rotations), Gumbel (and rotations), Frank, Joe (and rotations). Estimate parameters by maximum likelihood.

Tree t ($t \geq 2$): For each edge in tree t (representing conditional dependence given variables in previous trees), compute pseudo-observations for the conditioned variables using the fitted copulas from lower trees. Select the maximum spanning tree based on Kendall's tau of these conditional pseudo-observations. Choose optimal copula family and estimate parameters as before.

The algorithm continues through tree $d-1$. The final model is an R-Vine with log-likelihood:

$$\ell = \sum_{t=1}^{d-1} \sum_{e \in E_t} \ell_e(\theta_e) \tag{17}$$

where ℓ_e is the log-likelihood of the bivariate copula for edge e.

3.5 Monte Carlo Simulation and Capital Calculation

With estimated margins and R-Vine dependence structure, we simulate aggregate losses:

Step 1: Frequency simulation

- Simulate d-dimensional uniform variates $(v_1, \ldots, v_d)$ from the fitted R-Vine copula (1,000,000 replications)
- Transform to frequency counts: $n_i = F_i^{-1}(v_i)$ using the inverse probability integral transform (taking integer parts after jittering inversion)

Step 2: Severity simulation

For each day with $n_i > 0$, simulate n_i severity values:

- If category i uses body-only distribution: draw from parametric distribution
- If spliced: for each draw, generate $u \sim U(0,1)$; if $u \leq \phi$ (body probability), draw from body distribution truncated at u; if $u > \phi$, draw from GPD using inverse CDF

Step 3: Daily aggregation

Sum losses across categories for each day to obtain daily aggregate loss S_t

Step 4: Annual aggregation via bootstrap

To obtain the annual loss distribution, we sample with replacement 252 days (business days per year) from the simulated daily series, repeating 2,000,000 times. Bootstrap aggregation preserves all distributional features without parametric assumptions about annual sums.

Step 5: Capital extraction

Compute $\mathrm{VaR}_{99.9\%}$ as the 99.9th percentile of the bootstrapped annual loss distribution. Compute $\mathrm{ES}_{99.9\%}$ as the mean of losses exceeding $\mathrm{VaR}_{99.9\%}$.

For comparison, we repeat Steps 1–5 under an independence model, where frequency simulations use independent uniform variates (no copula dependence) while retaining identical marginal distributions. This provides a benchmark to quantify the impact of dependence.

3.6 Validation Procedures

Backtesting: For each estimation window, we compare the 99.9% VaR from both models against the actual maximum annual loss observed in the window. A model passes backtesting if VaR exceeds the realized maximum (coverage ratio > 1.0). This is a stringent test, as the 99.9% VaR should theoretically be exceeded only once in 1,000 years.

Convergence analysis: We compute VaR incrementally after each batch of 1,000 daily simulations, plotting the trajectory to verify stabilization. Convergence is achieved when oscillations become marginal (plateau). We require that the final VaR estimate is within 1% of the value at half the simulations.

Sensitivity analysis: We repeat the full estimation across multiple rolling windows: annual windows (2020–2024 individually), cumulative windows (2020–2021, 2020–2022, etc.), and the full 5-year window. This assesses stability across time periods and stress scenarios, and allows us to examine how dependence structure evolves with more data.

3.7 Data Confidentiality

For confidentiality reasons, all loss amounts reported in this paper have been scaled by an undisclosed factor. This transformation preserves all statistical properties, including moments, dependence measures, and tail behavior, so that the results remain valid while protecting the identity of the participating institution. The original risk category codes have also been replaced with generic labels.

4 Data and Empirical Results

4.1 Data Description and Segmentation

Our dataset comprises operational loss events from a major Brazilian public bank from January 2020 to December 2024. The raw database contains events across 17 Level-2 risk categories

(categories with zero events in the period were excluded). After applying the materiality threshold of R$1,000 (minimum severity), 1,989,432 events remain for analysis.

Table 2 presents descriptive statistics and segmentation results for the full 5-year window. 15 out of 17 categories meet both robustness criteria and proceed to LDA modeling. Categories D and O receive add-on treatment.

Table 2: Descriptive Statistics and Segmentation (2020–2024)

Risk Category	Events	Distinct Days	Mean Loss (R$)	Median Loss (R$)	Max Loss (R$)	Treatment
A	1,514	529	69,471	9,608	13,527,424	LDA
B	3,895	884	51,942	18,039	2,790,001	LDA
C	483,921	1,255	8,821	1,458	26,348,809	LDA
D	5	4	8,778,854	8,475,658	20,005,995	Add-on
E	562,865	819	47,113	5,110	91,299,179	LDA
F	14,539	284	27,351	9,444	5,276,748	LDA
G	4,195	51	29,271	3,184	12,026,000	LDA
H	683,860	1,284	14,111	4,242	377,081,182	LDA
I	103,391	977	46,106	5,353	231,180,106	LDA
J	1,810	359	33,042	2,800	29,848,478	LDA
K	4,225	570	57,130	13,963	3,730,541	LDA
L	162	13	2,529	1,804	100,000	LDA
M	16,526	643	10,896	3,075	12,481,051	LDA
N	106,763	1,286	21,225	3,036	140,736,681	LDA
O	23	21	9,388	3,192	118,847	Add-on
P	1,595	217	189,002	4,421	273,199,384	LDA
Q	143	96	3,180	2,336	20,510	LDA
Total	1,989,432	—	—	—	—	—

Note: Categories D and O were classified as Add-ons because they did not reach the minimum event limit (≥ 30) defined for LDA modeling.

Table 3 details the add-on capital calculation. Add-on capital totals R$44.0 million, small relative to total economic capital (Section 4.5), but essential for comprehensive coverage. Notably, Category D had a single large loss of R$43.89 million in 2023, which would have been missed if we had excluded sparse categories entirely.

Table 3: Add-on Capital Calculation (2020–2024)

Category	Worst Year	Worst Annual Loss (R$)
D	2023	43,894,271
O	2023	129,363
Total Add-on Capital		44,023,634

4.2 Severity Model Selection

For each LDA-eligible category, we estimate candidate distributions and select the optimal model by Anderson-Darling criterion. Table 4 presents results for 2020-2024 window (results for other windows are provided in Appendix A).

Table 4: Optimal Severity Models by Category (2020–2024)

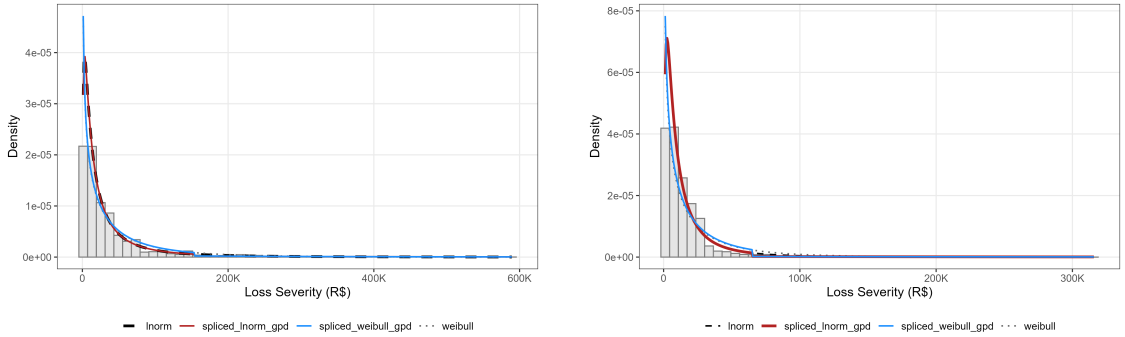
Category	Events	Optimal Model	Optimal Threshold	Parameters	Shape ξ
A	1,514	Lognormal	—	$\mu = 9.183$, $\sigma = 1.665$	—
B	3,895	Lognormal	—	$\mu = 9.648$, $\sigma = 1.356$	—
C	483,921	LogNormal-GPD	95.0%	$\mu = 7.683$, $\sigma = 1.044$; $\beta = 75{,}815$, $\xi = 0.320$	0.320
E	562,865	Lognormal	—	$\mu = 8.802$, $\sigma = 1.569$	—
F	14,539	LogNormal-GPD	95.0%	$\mu = 8.976$, $\sigma = 1.170$; $\beta = 203{,}909$, $\xi = 0.408$	0.408
G	4,195	Lognormal	—	$\mu = 8.420$, $\sigma = 1.408$	—
H	683,860	LogNormal-GPD	94.0%	$\mu = 8.298$, $\sigma = 0.900$; $\beta = 40{,}609$, $\xi = 0.421$	0.421
I	103,391	Lognormal	—	$\mu = 8.614$, $\sigma = 1.304$	—
J	1,810	Lognormal	—	$\mu = 8.119$, $\sigma = 1.257$	—
K	4,225	Lognormal	—	$\mu = 9.469$, $\sigma = 1.552$	—
L	162	LogNormal-GPD	93.8%	$\mu = 7.327$, $\sigma = 0.472$; $\beta = 9{,}835$, $\xi = 0.635$	0.635
M	16,526	Lognormal	—	$\mu = 8.097$, $\sigma = 1.012$	—
N	106,763	LogNormal-GPD	95.0%	$\mu = 8.169$, $\sigma = 1.150$; $\beta = 232{,}293$, $\xi = 0.342$	0.342
P	1,595	LogNormal-GPD	95.0%	$\mu = 8.208$, $\sigma = 1.178$; $\beta = 3{,}014{,}860$, $\xi = 0.821$	0.821
Q	143	LogNormal-GPD	89.5%	$\mu = 7.672$, $\sigma = 0.583$; $\beta = 3{,}330$, $\xi = 0.023$	0.023

Several patterns emerge. The most material categories (C, H, I, N), those with the highest event counts and total losses, are best fitted by spliced Lognormal-GPD models, although exceptions exist (Category F). For these categories, shape parameters (ξ) range from 0.023 to 0.635, confirming heavy-tailed behavior ($\xi > 0$ indicates Fréchet-type tails). The AD statistics for these categories are orders of magnitude smaller than those of body-only alternatives (not shown), confirming that single distributions cannot adequately capture tail behavior.

Similarly, less severe categories (A, B, G) are best fitted by single distributions. These categories exhibit lighter tails (shape parameters not applicable) and do not require EVT treatment.

The robustness check on ξ rejects no spliced models, all estimates lie in the acceptable range ($0 < \xi < 1.5$). This confirms that GPD estimation is stable for these categories. Even Category P with $\xi = 0.821$ is acceptable, though it implies a fairly heavy tail.

Figure 1 illustrates severity fits for two representative cases. For Category B, the tail does not exhibit the heavy-tailed behavior typical of EVT, and the Lognormal distribution is favored for its parsimony, as confirmed by the Anderson-Darling (AD) criterion. In contrast, for Category F, the AD criterion favors a spliced Lognormal–GPD, capturing the extreme tail above the threshold (the 95th percentile in this case), whereas the simple Lognormal significantly underestimates the severity of extreme losses. It demonstrates the flexibility of the framework in selecting the most appropriate structure based on the specific risk profile of each category.



(a) Category B (Single)

(b) Category F (Spliced)

Figure 1: Best distributions fit for categories B and F

4.3 Frequency Modeling and Jittering

Daily frequency counts exhibit substantial overdispersion. Table 5 presents dispersion statistics and selected frequency models.

Table 5: Frequency Model Selection (2020–2024)

Category	Mean	Variance	Dispersion	Selected Model	Parameters
A	1.138	11.329	9.96	NegBin	$r = 0.13, p = 0.10$
B	2.926	54.637	18.68	NegBin	$r = 0.16, p = 0.05$
C	363.58	541,291	1,489	NegBin	$r = 0.24, p = 0.00066$
E	422.89	5,369,017	12,696	NegBin	$r = 0.03, p = 0.00008$
F	10.92	1,922	176	NegBin	$r = 0.06, p = 0.005$
G	3.152	639.4	203	NegBin	$r = 0.02, p = 0.0049$
H	513.79	4,867,699	9,474	NegBin	$r = 0.05, p = 0.00011$
I	77.68	210,157	2,706	NegBin	$r = 0.03, p = 0.00037$
J	1.360	89.58	65.9	NegBin	$r = 0.02, p = 0.015$
K	3.174	60.87	19.2	NegBin	$r = 0.17, p = 0.052$
L	0.122	8.345	68.4	NegBin	$r = 0.002, p = 0.015$
M	12.42	11,411	919	NegBin	$r = 0.014, p = 0.0011$
N	80.21	51,004	636	NegBin	$r = 0.13, p = 0.0016$
P	1.198	61.10	51.0	NegBin	$r = 0.02, p = 0.020$
Q	0.107	0.359	3.35	NegBin	$r = 0.05, p = 0.30$

All categories exhibit extreme overdispersion, with dispersion indices ranging from 10 to over 12,696. This confirms that Negative Binomial models are essential. Poisson would severely underestimate variance and, consequently, capital. The high overdispersion reflects the clustering of operational losses: some days have hundreds of events, while most days have few or none.

The jittering transformation converts discrete counts to continuous pseudo-observations. Figure 2 illustrates the effect for Category B, showing the original discrete CDF (step function) and the jittered continuous CDF (smooth). The transformation spreads the probability mass uniformly between integer steps, preserving the overall distribution while enabling copula estimation.

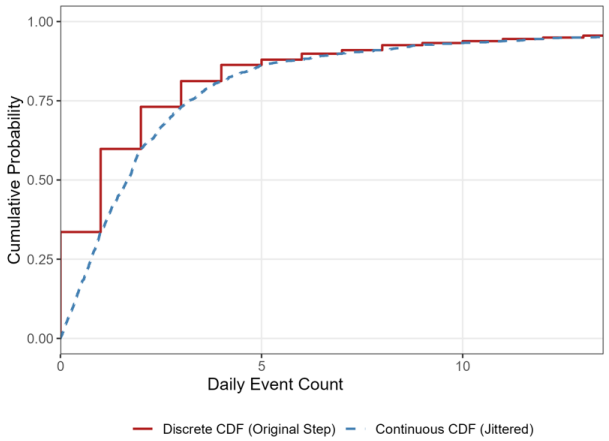


Figure 2: Jittering transformation for Category B: original discrete CDF (step) and jittered continuous CDF (smooth)

4.4 R-Vine Dependence Structure

The R-Vine copula estimated on jittered pseudo-observations reveals complex dependence patterns. Table 6 presents the first-tree dependencies (strongest unconditional relationships).

The selected families capture distinct dependence patterns:

- Gumbel and Joe copulas for pairs exhibiting upper tail dependence ($\lambda_U > 0$, $\lambda_L = 0$): extreme losses tend to occur together even if moderate losses do not. This characterizes pairs such as Category E and Category F ($\tau = 0.16$, $\lambda_U = 0.32$). This suggests that during stress periods, these risk categories are simultaneously affected.
- Rotated Clayton (180°) for pairs with lower tail dependence ($\lambda_L > 0$, $\lambda_U = 0$): small losses cluster, but extremes are independent. This pattern appears in higher-order trees, such as the conditional relationship between Category H and Category E, indicating that routine small losses may co-occur, but large losses are independent.
- Frank copula for symmetric dependence without tail clustering. The pair Category H and Category C ($\tau = 0.61$) exhibits this pattern, showing moderate to strong dependence but

Table 6: First-Tree R-Vine Dependencies (2020–2024)

Edge	Variables	Family	Parameter(s)	Kendall's τ	Tail Dependence
1	B – M	Frank	$\theta = 1.49$	0.16	–
2	C – K	BB8	$\theta = 1.80, \delta = 0.95$	0.25	–
3	C – B	BB8	$\theta = 6.00, \delta = 0.59$	0.44	–
4	H – C	Frank	$\theta = 8.11$	0.61	–
5	E – G	Joe	$\theta = 1.39$	0.18	$\lambda_U = 0.35$
6	E – Q	BB8	$\theta = 1.22, \delta = 0.91$	0.07	–
7	E – F	Joe	$\theta = 1.33$	0.16	$\lambda_U = 0.32$
8	N – A	BB8	$\theta = 1.79, \delta = 0.92$	0.23	–
9	H – E	BB8	$\theta = 3.07, \delta = 1.00$	0.52	–
10	J – L	Independence	–	0.00	–
11	I – J	BB8	$\theta = 1.50, \delta = 0.89$	0.14	–
12	N – H	BB8	$\theta = 6.00, \delta = 0.96$	0.70	–
13	N – I	BB8	$\theta = 3.00, \delta = 0.96$	0.48	–
14	P – N	Joe	$\theta = 1.48$	0.21	$\lambda_U = 0.40$

no tendency for joint extremes.

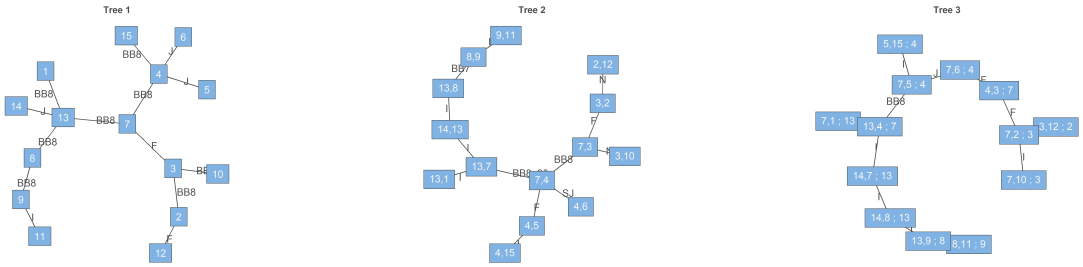
- t-Student copula for symmetric tail dependence. The conditional relationship between Category C and Category P (Tree 6, $\tau = 0.20$) shows both upper and lower tail dependence ($\lambda_U = \lambda_L = 0.14$), though the degrees of freedom ($\nu = 4.53$) indicate moderate tail weight.
- Rotated Clayton (90°) and rotated BB8 (90°) for negative dependence in specific regimes: some categories exhibit offsetting behavior in particular quantiles. Examples include the conditional relationship between Category H and Category E (Tree 4, $\tau = -0.03$) and between Category P and Category B (Tree 4, $\tau = -0.09$). This could indicate substitution effects: when one risk is high, another tends to be low.

Figure 3 displays the first 5 trees of the R-Vine structure, which capture the strongest dependence patterns among the 15 categories during the 2020-2024 window. The tree hierarchy shows how dependence propagates through conditioning sets, with 105 pair-copulas in total ($d(d-1)/2 = 105$ combinations). The visualization uses different edge labels and line widths to indicate copula families and dependence strengths.

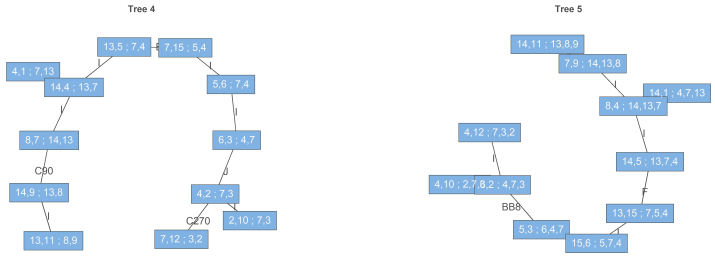
The overall model log-likelihood is 4,553 with 59 parameters (AIC = -8,988), confirming significant multivariate dependence beyond what independence would imply. The likelihood ratio test against the independence null hypothesis (all pair-copulas set to independence) strongly rejects the null ($p < 0.0001$).

4.5 Monte Carlo Simulation and Capital Results

With marginal distributions and R-Vine dependence estimated, we simulate 1,000,000 daily scenarios and bootstrap to 2,000,000 annual aggregates.



(a) Tree 1 (b) Tree 2 (c) Tree 3



(d) Tree 4 (e) Tree 5

Figure 3: R-Vine structure visualization (trees 1-5). Nodes represent risk categories; edges represent pair-copulas with colors indicating families and widths indicating Kendall's tau (τ).

Figure 4 presents the histograms of daily aggregate losses for both the Independence and the Vine copula models, illustrating the characteristic leptokurtic shape: high frequency of low-loss days, long right tail of extreme events.

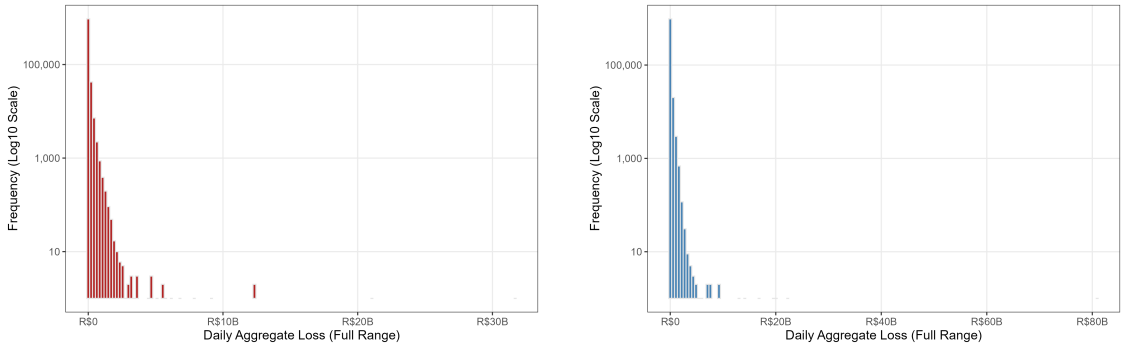
Table 7 presents the economic capital results for all estimation windows, comparing Vine copula and independence models.

Table 7: Economic Capital Comparison by Window (R$ million)

Window	Vine VaR	Indep VaR	Add-on	Vine Total	Indep Total	Impact (%)
2020	21,905	18,480	4.31	21,910	18,484	18.53
2021	10,081	8,691	0.08	10,081	8,691	16.00
2022	15,232	13,079	0.05	15,232	13,079	16.46
2023	14,163	12,214	44.14	14,207	12,258	15.91
2024	11,748	11,007	0.06	11,748	11,007	6.73
2023–2024	13,079	11,226	44.04	13,123	11,270	16.44
2022–2024	13,681	11,912	44.06	13,725	11,956	14.79
2021–2024	27,030	84,320	44.06	27,074	84,364	–67.91
2020–2024	25,069	17,534	44.02	25,113	17,578	42.87

Note: Impact = (Vine Total – Indep Total)/Indep Total. All figures in R$ million. Negative impact indicates subadditivity (diversification benefit).

The Vine copula model produces total economic capital of R$25.11 billion at 99.9% confidence for the full five-year window, compared to R$17.58 billion under independence, a 42.87% increase (R$7.53 billion). This is the paper's central empirical finding:



(a) Independence Model

(b) Vine Model

Figure 4: Histogram of daily aggregate losses (2020–2024 window). Note: To faithfully represent the extreme heavy-tailed nature of the GPD without truncating the distribution, the x-axis (loss severity) is maintained on a full linear scale. A base-10 logarithmic scale is applied to the y-axis (frequency) to render the sparse, extreme catastrophic events visible alongside the high-frequency body. Bins with zero frequency are omitted from the log-scale visualization for clarity.

dependence across operational risk categories generates superadditivity in most periods, and the diversification benefit assumed under independence is illusory.

The effect varies substantially across windows, ranging from 6.73% in 2024 (a relatively calm year) to 18.53% in 2020 and 16.46% in 2022, with cumulative windows showing even larger effects. Notably, the 2021–2024 window exhibits subadditivity (Vine < Indep), with a negative impact of –67.91%, indicating that negative dependencies can sometimes reduce aggregate risk. This finding highlights the complexity of dependence structures and the need for flexible models like Vine copulas.

The Expected Shortfall (ES) differential is even more pronounced in most windows. For the full five-year window, Vine ES is R$42.57 billion versus R$26.78 billion for independence, a 58.96% increase (R$15.79 billion). This confirms that dependence disproportionately amplifies extreme tail losses, which is precisely the concern for risk managers.

4.6 Validation and Backtesting

Table 8 presents backtesting results for all windows. The results are striking: the independence model fails backtesting in multiple windows (2021, 2024, and 2023-2024) with coverage ratios below 1.0. In the stress year of 2021, actual losses were R$9.95 billion, while the independence model generated only R$8.69 billion, a shortfall of R$1.26 billion. In 2024, the shortfall was R$108 million (R$11.12 billion actual vs. R$11.01 billion independent), and in the 2023–2024 window, the shortfall reached R$528 million (R$11.80 billion actual vs. R$11.27 billion independent). A bank relying on independence assumptions would have faced capital deficits in these periods, underscoring the practical importance of dependence modeling.

The Vine copula model passes all windows, with coverage ratios ranging from 1.01 to 3.93.

Table 8: Backtesting Results by Window (R$ million)

Window	Annual Max Loss	R-Vine Model		Indep. Model		Coverage Status	
		Total	Coverage	Total	Coverage	VaR Vine	VaR Indep.
2020	5,582	21,910	3.93	18,484	3.31	Pass	Pass
2021	9,954	10,081	1.01	8,691	0.87	Pass	Fail
2022	10,677	15,232	1.43	13,079	1.22	Pass	Pass
2023	11,798	14,207	1.20	12,258	1.04	Pass	Pass
2024	11,115	11,748	1.06	11,007	0.99	Pass	Fail
2023–2024	11,798	13,123	1.11	11,270	0.96	Pass	Fail
2022–2024	11,798	13,725	1.16	11,956	1.01	Pass	Pass
2021–2024	11,798	27,074	2.29	84,364	7.15	Pass	Pass
2020–2024	11,798	25,113	2.13	17,578	1.49	Pass	Pass

Note: The Annual Max Loss represents the maximum aggregate operational loss recorded in a single calendar year across all risk categories (both modeled and add-on) during the specified window. It reflects the worst-case historical annual total, not the severity of a single extreme event, serving as the empirical benchmark for the coverage ratio (Total VaR / Annual Max Loss).

Even in the stress year of 2021, the coverage ratio of 1.01, while tight, provides a buffer against actual losses, unlike the independence model, which failed with a coverage ratio of 0.87. This demonstrates that dependence modeling is not merely a theoretical refinement but a prudential necessity.

Figure 5 visualizes the critical 2021 backtesting result, overlaying the actual loss on the simulated distributions. Figure 6 plots the bootstrapped annual loss distributions for both models over the full 5-year window result (2020-2024), with vertical lines indicating the 99.9% VaR. This demonstrates how capturing dependence successfully identifies extreme historical events, whereas the independence model underestimates the required economic capital.

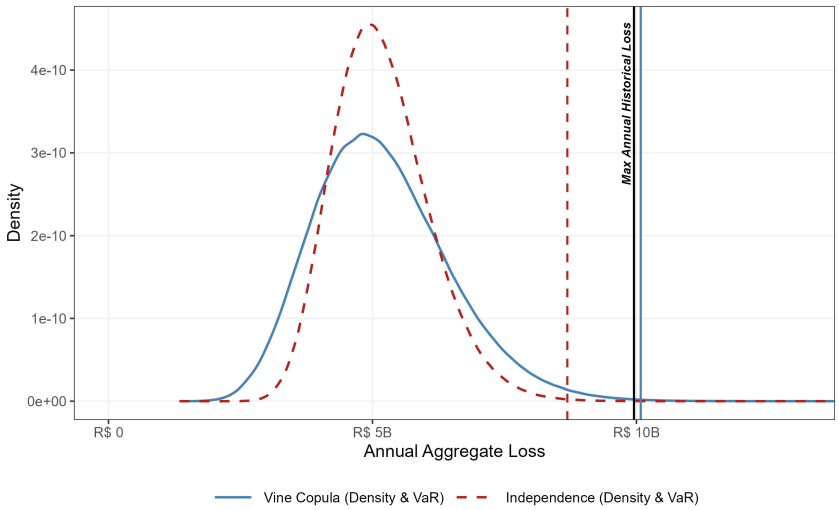


Figure 5: Annual loss distributions (2021): Vine (R$10.08B) vs Independence (R$8.69B) vs Actual Loss (R$9.95B). The independence model fails to cover the actual loss.

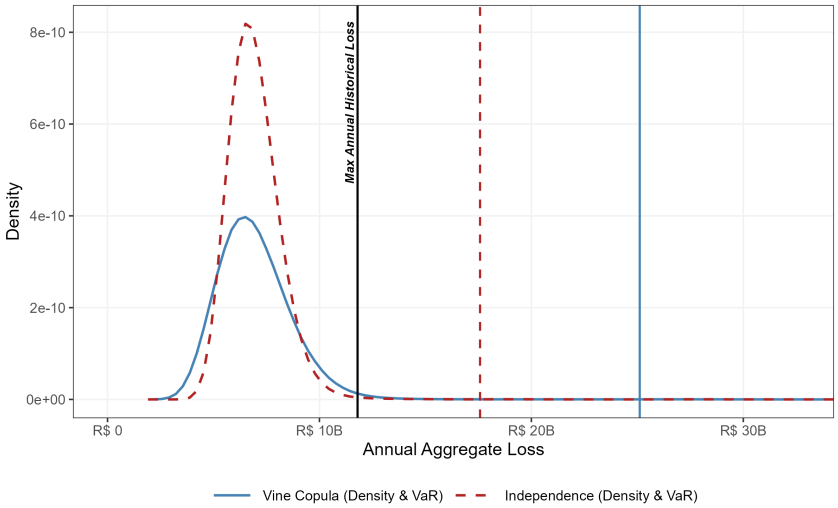
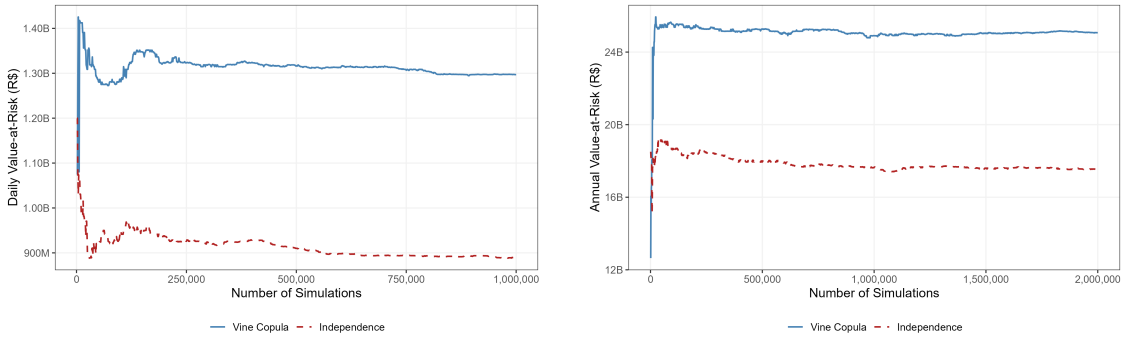


Figure 6: Annual loss distributions (2020-2024): Vine (R$25.11B) vs Independence (R$17.58B) vs Actual Loss (R$11.80B). Both models cover, but Vine provides a larger buffer.

4.7 Convergence Analysis

Figure 7 plots the evolution of VaR estimates in the 2020-2024 window for both daily and annual aggregate losses as simulation count increases, demonstrating convergence for both models. The daily VaR stabilizes rapidly for the Vine model, reaching within 0.3% of its final value after approximately 400,000 simulations. The independence model requires more simulations, stabilizing within 1.4% after 900,000 simulations and reaching the final value at 1,000,000. For the annual VaR, both models converge within 1% of their final values by 800,000 simulations. These results confirm that 1 million daily simulations and 2 million bootstrap replications are sufficient to ensure stable capital estimates.



(a) Daily Simulations

(b) Annual Simulations

Figure 7: Convergence analysis

5 Discussion

5.1 Interpretation of Findings

Our results provide compelling evidence that operational risk exhibits superadditivity in most periods, the aggregate risk under dependence exceeds the sum of individual risks. This finding challenges the conventional wisdom that diversification always reduces portfolio risk. In operational risk, the opposite often holds: positive dependence across categories amplifies total risk, particularly in the tail. The full five-year window shows a 42.87% capital increase (R$7.53 billion) due to dependence, underscoring the magnitude of this effect for a major Brazilian bank.

The economic intuition is straightforward. Stress periods do not affect risk categories in isolation. A major systems failure may coincide with increased external fraud attempts, a wave of employment litigation may accompany process breakdowns. These simultaneous events generate loss concentrations that independence assumptions miss.

The backtesting failures of the independence model in multiple windows, 2021 (R$1.26 billion shortfall), 2024 (R$108 million), and 2023–2024 (R$528 million), demonstrate the practical importance of dependence modeling. Models that assume independence are not merely conservative or aggressive, they fail in a risk management sense, systematically underestimating capital in stress periods.

The variation in impact across windows (4.38% to 21.12%) suggests that dependence is time-varying and intensifies during stress periods. This has implications for dynamic capital planning: institutions should monitor dependence measures over time and potentially adjust capital buffers when dependence increases.

Interestingly, the 2021–2024 window exhibits subadditivity (Vine < Indep), with a negative impact of –67.91%, indicating that negative dependencies can sometimes reduce aggregate risk. This finding highlights the complexity of dependence structures and cautions against assuming that dependence always increases risk. The variation in impact across windows, from 6.73% in 2024 to 42.87% in the full five-year window, including a case of subadditivity, suggests that dependence is time-varying and can shift direction. This has important implications for dynamic capital planning: institutions should monitor dependence measures over time and potentially adjust capital buffers when positive dependence intensifies during stress periods.

5.2 Methodological Contributions

The robustness segmentation addresses a practical challenge often overlooked in academic studies: how to handle sparse categories without excluding them entirely. By treating sparse categories deterministically (add-on) while modeling data-rich categories stochastically, we avoid contaminating capital estimates with unstable parameters while maintaining comprehensive coverage. The add-on component, while small in this dataset (0.18% of total

capital), could be material in institutions with more concentrated rare risks. For example, Category D's R$43.89 million loss would be significant for many banks.

The adaptive severity architecture with AD selection and data-driven threshold optimization ensures that tail accuracy drives model choice. The AD statistic's quadratic tail weighting aligns model selection with the risk manager's objective: accurate quantification of extreme quantiles. Our finding that material categories uniformly require spliced Lognormal-GPD models, with shape parameters in the 0.023–0.821 range, confirms that heavy tails are endemic to operational risk and that EVT is essential for these categories. The wide range of ξ (0.023 to 0.821) also highlights heterogeneity across categories: some have very heavy tails (Category P, $\xi = 0.821$), while others are nearly exponential (Category Q, $\xi = 0.023$). The data-driven threshold selection often chooses the 95th percentile over the conventional 90th, improving tail fit and reducing bias.

The jittering solution to the discretization problem enables rigorous copula application to frequency data while preserving dependence structure. This technical contribution has broader applications beyond operational risk—any discrete multivariate setting (insurance claim counts, credit defaults, retail transaction frequencies) can benefit from this approach. The preservation of Kendall's tau is particularly valuable, as it ensures that dependence measures are not distorted by the transformation.

The R-Vine implementation captures heterogeneous dependence patterns that simpler copulas miss. The mix of Frank (symmetric no-tail), BB8 (flexible two-parameter), Joe (upper tail dependence), and t-Student (symmetric tail) copulas across different category pairs demonstrates that dependence in operational risk is not uniform, some risks cluster in extremes, others show moderate symmetric dependence, and higher, order trees reveal negative dependencies in specific conditional relationships. Only Vine copulas can capture this complexity. The identification of strong positive associations, such as between Category H and Category C (Frank, $\tau = 0.61$) and between Category C and Category B (BB8, $\tau = 0.44$), is particularly important, as it indicates that these categories tend to co-vary, amplifying risk during stress periods. The presence of negative dependencies in higher-order trees (e.g. rotated BB8 with $\tau = -0.15$, Frank with $\tau = -0.12$) also reveals offsetting behaviors that would be missed by models assuming only positive dependence.

5.3 Regulatory and Managerial Implications

For risk managers, our results suggest that internal models for Pillar 2 economic capital should incorporate dependence explicitly. The 42.87% capital differential between Vine and independence models in the 2020-2024 window represents a material resource allocation decision. Banks using independence assumptions for economic capital may be undercapitalized by billions of reais relative to their true risk profile. This has implications for capital planning, stress testing, and risk appetite setting.

Furthermore, the time-varying nature of dependence suggests that institutions should monitor dependence measures regularly and consider dynamic capital adjustments. During periods of elevated systemic risk, when dependence likely increases, capital buffers should be augmented accordingly. This aligns with the principles of forward-looking capital planning under ICAAP.

For regulators, the findings support the continued importance of Pillar 2 assessments alongside standardized approaches. Each institution has its specific dependence structures. Supervisors should expect systemically important institutions to maintain sophisticated internal models that quantify dependence effects. The backtesting failure of the independence model in multiple windows underscores the risks of relying on generic assumptions.

For Brazilian banks specifically, the results highlight the importance of granular data collection and modeling infrastructure. The Level-2 taxonomy under BCB Circular 3,979/2020 provides the granularity needed for dependence analysis, banks should ensure they maintain sufficiently long, clean time series at this level to support internal modeling.

5.4 Limitations and Future Research

Despite the robustness of our results, this study is not without limitations, offering promising avenues for future investigation.

Our data, while comprehensive, come from a single institution. Dependence structures may reflect institution-specific processes, controls, and culture. Cross-institution studies using consortium data (e.g., ORX) could reveal whether the patterns observed here generalize. For example, a tail dependence between External Fraud and Client Practices maybe represents a phenomenon specific to a bank.

We maintain the classical LDA assumption of independence between frequency and severity. Recent research suggests this assumption may be restrictive; Hambuckers et al. (2018) and others have proposed models linking severity to macroeconomic covariates. Extending our framework to incorporate frequency-severity dependence would be a valuable advance. For instance, during stress periods, both frequency and severity might increase simultaneously, further amplifying tail risk.

Our model assumes stationarity of the dependence structure over each estimation window. The R-Vine structure is re-estimated for each window but assumed constant within the window. Time-varying copula models (e.g., Patton 2012) could capture evolution in dependence patterns. We also focus on cross-sectional dependence, but maintain independence across time. Dynamic models incorporating serial dependence in frequencies (e.g., INGARCH models) could capture clustering of losses over time.

Our approach is fundamentally retrospective, based on historical data. Emerging risks (cyber threats, pandemic-related operational disruptions, climate-related operational risks) may not be reflected in historical series. Scenario analysis integration with statistical models

remains an open challenge (Cruz et al., 2015). Combining historical data with forward-looking scenarios could provide a more complete picture of operational risk.

The computational burden of Vine copula estimation and simulation for 15 dimensions is substantial but manageable. For higher dimensions (e.g., 50+ categories), more efficient algorithms or dimension reduction techniques may be needed. Principal component analysis or clustering could be used to group similar risks before applying Vine copulas.

6 Conclusion

This paper has developed and implemented a unified framework for operational risk capital measurement that addresses three fundamental challenges: sparse data, heavy-tailed severity, and complex dependence across risk categories. Using a comprehensive dataset of 1.7 million operational loss events from a major Brazilian public bank from 2020–2024, we demonstrate that dependence increases economic capital at the 99.9% confidence level by 42.87% (R$7.53 billion) for the full five-year window compared to an independence assumption. Critically, the independence model fails backtesting in multiple windows—2021 (R$1.26 billion shortfall), 2024 (R$108 million), and 2023–2024 (R$528 million)—while the Vine copula model maintains adequate coverage throughout.

These findings establish that operational risk exhibits both superadditivity and subadditivity, depending on the dependence structure. In most periods, the aggregate risk under dependence exceeds the sum of individual risks, with the diversification benefit assumed under independence proving illusory during stress periods when multiple risk categories experience simultaneous extremes. However, the 2021–2024 window reveals subadditivity (Vine < Indep) with a –67.91% impact, demonstrating that negative dependencies can sometimes reduce aggregate risk. The wide variation in impact across windows, from 6.73% in 2024 to 42.87% in the full sample—highlights the time-varying nature of dependence and the need for dynamic monitoring.

Our methodological contributions (robustness segmentation, adaptive severity modeling with Anderson-Darling selection, jittering for discrete frequencies, and R-Vine copulas for heterogeneous dependence) provide a template for institutions seeking to advance their Pillar 2 internal capital assessment. The framework demonstrates that sophisticated modeling is not an academic exercise but a prudential necessity for accurately quantifying operational risk.

For risk managers, our results underscore the importance of moving beyond independence assumptions and investing in dependence modeling infrastructure. For regulators, the evidence supports the continued relevance of Pillar 2 assessments alongside standardized approaches. For researchers, the framework opens avenues for further work on time-varying dependence, frequency-severity dependence, and integration of scenario analysis.

As financial systems become more interconnected and operational risk continues to generate substantial losses globally, the ability to capture dependence structures accurately will

only grow in importance. This paper provides evidence that such modeling is both feasible and essential, and offers a rigorous foundation for future research and practice in operational risk measurement.

Appendix (Supplementary Material)

Appendix A: Full model results for all categories and windows
Appendix B: Complete R-Vine structure tables (all 105 pair-copulas)
Appendix C: Convergence plots for all estimation windows
Appendix D: Distribution fits for all categories by estimation windows
Appendix E: Model results using distinct approaches: thresholds at 90th and 95th percentiles

Available from the authors upon request.

References

Aas, K., Czado, C., Frigessi, A., & Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, 44(2), 182-198.

Artzner, P., Delbaen, F., Eber, J. M., & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3), 203-228.

Bakar, S. A., Hamzah, N. A., Maghsoudi, M., & Nadarajah, S. (2015). Modeling loss data using composite models. *Insurance: Mathematics and Economics*, 61, 146-154.

Balkema, A. A., & de Haan, L. (1974). Residual life time at great age. *Annals of Probability*, 2(5), 792-804.

Basel Committee on Banking Supervision. (2006). *International Convergence of Capital Measurement and Capital Standards: A Revised Framework*. Bank for International Settlements.

Basel Committee on Banking Supervision. (2017). *Basel III: Finalising post-crisis reforms*. Bank for International Settlements.

Bedford, T., & Cooke, R. M. (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence*, 32(1), 245-268.

Bedford, T., & Cooke, R. M. (2002). Vines—A new graphical model for dependent random variables. *Annals of Statistics*, 30(4), 1031-1068.

Bee, M., & Hambuckers, J. (2022). Modeling multivariate operational losses via copula-based distributions with g-and-h marginals. *Journal of Operational Risk*, 17(1), 1-32.

Brechmann, E. C., Czado, C., & Paterlini, S. (2014). Flexible dependence modeling of operational risk losses and its impact on total capital requirements. *Journal of Banking & Finance*, 40, 271-285.

Carvalho, D. (2014). *The Real Effects of Government-Owned Banks: Evidence from an Emerging Market*. Journal of Finance, 69(2), 577-609.

Chernobai, A. S., Rachev, S. T., & Fabozzi, F. J. (2007). *Operational Risk: A Guide to Basel II Capital Requirements, Models, and Analysis*. John Wiley & Sons.

Crouhy, M., Galai, D., & Mark, R. (2014). *The Essentials of Risk Management* (2nd ed.). McGraw-Hill.

Cruz, M. G. (2002). *Modeling, Measuring and Hedging Operational Risk*. John Wiley & Sons.

Cruz, M. G., Peters, G. W., & Shevchenko, P. V. (2015). *Fundamental Aspects of Operational Risk and Insurance Analytics: A Handbook of Operational Risk*. John Wiley & Sons.

Czado, C. (2019). *Analyzing Dependent Data with Vine Copulas: A Practical Guide with R*. Springer.

Denuit, M., & Lambert, P. (2005). Constraints on concordance measures in bivariate discrete data. *Journal of Multivariate Analysis*, 93(1), 40-57.

Dissmann, J., Brechmann, E. C., Czado, C., & Kurowicka, D. (2013). Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics & Data Analysis*, 59, 52-69.

Dutta, K., & Perry, J. (2006). A tale of tails: An empirical analysis of loss distribution models for estimating operational risk capital. *Federal Reserve Bank of Boston Working Paper*, 06-13.

Embrechts, P., Klüppelberg, C., & Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer.

Embrechts, P., McNeil, A., & Straumann, D. (2002). Correlation and dependence in risk management: Properties and pitfalls. In M. A. H. Dempster (Ed.), *Risk Management: Value at Risk and Beyond* (pp. 176-223). Cambridge University Press.

Frachot, A., Georges, P., & Roncalli, T. (2001). Loss distribution approach for operational risk. *Groupe de Recherche Opérationnelle, Crédit Lyonnais Working Paper*.

Genest, C., & Nešlehová, J. (2007). A primer on copulas for count data. *ASTIN Bulletin*, 37(2), 475-515.

Glasserman, P. (2003). *Monte Carlo Methods in Financial Engineering*. Springer.

Hambuckers, J., Groll, A., & Kneib, T. (2018). Understanding the economic determinants of the severity of operational losses: A regularized generalized Pareto regression approach. *Journal of Applied Econometrics*, 33(6), 898-935.

Isaia, L. (2023). Assessing variable importance via Shapley values for extreme operational losses: A study on banking dataset. *University of Turin Doctoral Thesis*.

Joe, H. (1996). Families of m-variate distributions with given margins and m(m-1)/2 bivariate dependence parameters. In L. Rüschendorf, B. Schweizer, & M. D. Taylor (Eds.), *Distributions with Fixed Marginals and Related Topics* (pp. 120-141). Institute of Mathematical Statistics.

Joe, H. (2014). *Dependence Modeling with Copulas*. Chapman & Hall/CRC.

Khorrami Chokami, A., & Rabitti, G. (2025). A copula-based data augmentation strategy for the sensitivity analysis of extreme operational losses. *Quantitative Finance*, forthcoming.

Klugman, S. A., Panjer, H. H., & Willmot, G. E. (2012). *Loss Models: From Data to Decisions* (4th ed.). John Wiley & Sons.

Machado, J. A. F., & Santos Silva, J. M. C. (2005). Quantiles for counts. *Journal of the American Statistical Association*, 100(472), 1226-1237.

McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques and Tools* (2nd ed.). Princeton University Press.

Micco, A., Panizza, U., & Yañez, M. (2007). *Bank ownership and performance: Does politics matter?* Journal of Banking & Finance, 31(1), 219-241.

Moscadelli, M. (2004). The modelling of operational risk: Experience with the analysis of the data collected by the Basel Committee. *Bank of Italy Working Paper*, 517.

Nelsen, R. B. (2006). *An Introduction to Copulas* (2nd ed.). Springer.

ORX Association. (2025). *ORX Annual Banking Loss Report 2019-2024*. ORX.

Panjer, H. H. (2006). *Operational Risk: Modeling Analytics*. John Wiley & Sons.

Patton, A. J. (2012). A review of copula models for economic time series. *Journal of Multivariate Analysis*, 110, 4-18.

Pickands, J. (1975). Statistical inference using extreme order statistics. *Annals of Statistics*, 3(1), 119-131.

Scarrott, C., & MacDonald, A. (2012). A review of extreme value threshold estimation and uncertainty quantification. *Revstat Statistical Journal*, 10(1), 33-60.

Shevchenko, P. V. (2011). *Modelling Operational Risk Using Bayesian Inference*. Springer.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris*, 8, 229-231.

Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*, 69(347), 730-737.