

SOA 2026 STUDENT RESEARCH CASE STUDY CHALLENGE

Space Mining Insurance for Cosmic Quarry Mining Corporation

A Comprehensive Actuarial Analysis
and Product Recommendations



Prepared by the Actuarial Duo

Team Members: Aicha Malouche, Tehzeeb Khalid

Advisor: Dr. Andres Barajas Paz

Heriot-Watt University

Submitted to

Galaxy General Insurance Company

March 2026

Prepared for academic competition purposes only. All entities, currencies, and events are fictional.

Abstract

Galaxy General should offer all four coverage lines — Equipment Failure, Business Interruption, Workers' Compensation, and Cargo Loss — across all three of Cosmic Quarry's solar systems (Helionis Cluster, Bayesia System, and Oryn Delta). The programme is profitable under the most likely operating scenario and remains solvent under a once-in-a-century catastrophe, provided three non-negotiable controls are in place: a $\text{€}50\text{ M}$ per-event cap on Business Interruption, a $\text{€}50\text{ M}$ hard limit per Cargo shipment backed by mandatory reinsurance, and a capital reserve of $\text{€}265\text{ B}$ that accounts for the measured risk dependence between the Helionis and Bayesia systems. Ignoring that dependence would understate required capital by $\text{€}64\text{ B}$ (+32%) — an existential solvency gap. Base expected losses total $\text{€}287\text{ B}$ (Cargo $\approx 85\%$); under a catastrophic correlated scenario ($p = 0.01$) losses reach $\text{€}2,518\text{ B}$. A phased market-entry strategy is recommended: Equipment Failure and Workers' Compensation in year one, Business Interruption in year two, and Cargo in year three — sequencing capital deployment to match reinsurance placement and underwriting confidence.

Contents

| | | |
|----------|---|-----------|
| 1 | Executive Summary | 1 |
| 1.1 | Recommended Products | 1 |
| 1.2 | Integration into Cosmic Quarry ERM Framework | 2 |
| 2 | Product Design | 3 |
| 2.1 | Equipment Failure | 3 |
| 2.2 | Business Interruption | 3 |
| 2.3 | Workers' Compensation | 3 |
| 2.4 | Cargo Loss | 3 |
| 3 | Pricing and Capital Modelling | 4 |
| 3.1 | Compound Frequency–Severity Model | 4 |
| 3.2 | Severity Distribution Fitting | 5 |
| 3.3 | Pure Premiums and Time Value of Money | 5 |
| 3.4 | Aggregate Loss Distributions and Capital Requirements | 5 |
| 3.5 | Correlation Loading and Copula Analysis | 5 |
| 4 | Risk Assessment | 6 |
| 4.1 | Risk Identification by Hazard Area and Solar System | 6 |
| 4.2 | Threat Table | 7 |
| 4.3 | Correlated Risk: Helionis–Bayesia Scenario | 7 |
| 4.4 | Scenario Testing | 7 |
| 4.5 | Comparative Risk Narrative Across Solar Systems | 8 |
| 5 | Assumptions | 8 |
| 6 | Data Sources and Limitations | 10 |
| | Appendix A: R Code | 12 |
| | Appendix B: Supporting Detail | 27 |

List of Tables

| | | |
|----|---|----|
| 1 | Product recommendations for CQMC’s three-system mining portfolio. All lines: Recommend . Controls listed are non-negotiable underwriting features. | 1 |
| 2 | Product design summary — all four coverage lines. Full specifications, exclusions, and system-specific features in Appendix B, Tables B2-B5. . . | 4 |
| 3 | AIC-selected severity distributions: parameter ranges across solar systems and approximate CV. Per-system parameters in Appendix B. | 5 |
| 4 | Claim frequencies, mean severities, and pure premiums by coverage line and solar system. | 6 |
| 5 | Aggregate loss statistics and capital requirements under the compound Poisson model. Ranges span the three solar systems; per-system detail in Appendix B. Rows ordered by increasing $\text{VaR}_{99\%}$ | 6 |
| 6 | Ranked threat register across CQMC’s three-system portfolio. Lines affected in brackets; frequency class in parentheses. | 7 |
| 7 | Scenario parameters and aggregate portfolio expected losses. | 8 |
| 8 | R packages used in this analysis. | 12 |
| B1 | Full modelling assumptions, rationale, and sensitivity assessment (all nine rows). | 27 |
| B2 | Equipment Failure: full product design. | 28 |
| B3 | Business Interruption: full product design. | 29 |
| B4 | Workers’ Compensation: full product design. | 30 |
| B5 | Cargo Loss: full product design. | 30 |

List of Figures

| | | |
|----|---|----|
| 1 | TVM-adjusted pure premiums (€ per unit of exposure) for EF, BI, and WC by solar system (2160–2174). Cargo excluded (Pareto scale incompatible; see Appendix B, Figure B2). | 2 |
| 2 | Portfolio aggregate expected loss under three scenarios: Best/Attritional ($p = 0.70$, €161 B), Moderate/Base ($p = 0.29$, €288 B), Worst/Catastrophic ($p = 0.01$, €2,518 B). Multipliers per Table 7. | 7 |
| B1 | Fitted lognormal severity distributions for Equipment Failure by solar system (AIC-selected). Oryn Delta exhibits the heaviest tail ($\hat{\mu} = 11.30$, $\hat{\sigma} = 0.61$), consistent with its higher observed mean severity of €101,717. | 33 |
| B2 | Cargo loss severity: empirical survival function vs. fitted Pareto ($\hat{\alpha} = 0.48$, $\hat{\theta} = 132,849$) on log–log scale. The near-linear log–log relationship confirms Pareto behaviour; the shallow slope ($\hat{\alpha} < 1$) signals that the theoretical mean is undefined under the pure Pareto model, necessitating a hard per-shipment policy limit. | 33 |
| B3 | Aggregate risk measures ($\text{VaR}_{99\%}$ and $\text{TVaR}_{99\%}$) normalised within each coverage line. The near-equal bars for EF reflect its thinner lognormal tail ($\hat{\sigma} \approx 0.60$); BI shows greater spread ($\hat{\sigma} \approx 1.20$); Cargo bars are dominated by the heavy Pareto tail ($\hat{\alpha} = 0.48$). All values from Table 5 in the main body. | 34 |

1 Executive Summary

Recommendation: offer all four lines across all three systems. This programme represents a viable, strategically differentiated business opportunity for Galaxy General Insurance Company. Base expected annual losses total €287B (Cargo ≈85%, Business Interruption ≈15%); at TVM-adjusted pure premiums, the programme is profitable across all four lines under the base scenario, with net revenue supported by structured sublimits, mandatory reinsurance, and a correlation-aware capital reserve of €265B. The critical risk is **cross-system correlation**: treating the three systems as independent understates required capital by €64B (+32%), an existential solvency exposure if unaddressed. This risk is managed through a Gaussian copula capital load, a 3–5% correlation premium loading on all Helionis–Bayesia bundled policies, and mandatory multi-system catastrophe reinsurance — the pre-placed reinsurance treaties and per-event sublimits are the mechanism by which the programme remains solvent under the catastrophic correlated scenario (probability 0.01, aggregate losses €2,518B). Project objectives — costs, expected returns, net revenue, variance, tail risk, capital, and solvency — are each explicitly addressed in Sections 1–4.

1.1 Recommended Products

All four lines are **recommended** across all three systems (Table 1); the conditions listed are mandatory product design features that make each line viable and profitable, not expressions of hesitation. Product features are system-tailored throughout: Helionis’s aging fleet and high claim frequency (0.221/unit) drive a €5K per-claim deductible; Bayesia’s binary K-type stellar environment requires radiation endorsements on both Equipment Failure and Business Interruption; and Oryn Delta’s planned 15% production growth justifies a conditional higher BI sublimit and a per-unit growth clause on Equipment Failure.

Table 1: Product recommendations for CQMC’s three-system mining portfolio. All lines: **Recommend**. Controls listed are non-negotiable underwriting features.

| Coverage Line | Profitability Rationale | Key Controls |
|-------------------------------|---|---|
| Equipment Failure (EF) | Stable lognormal severity; attritional pattern supports consistent margin. Helionis PP €17,411/unit; Oryn Delta lowest freq (0.137) provides buffer; profitable on all three systems. | Helionis €5K deductible suppresses attritional claims. 3–5% correlation loading on Helionis–Bayesia bundle. Mine-count declarations trigger auto re-rating. |
| Business Interr. (BI) | Per-system $\mathbb{E}[S] \in [€9.4B, €18.0B]$ managed within sublimit. €50M cap + 20% co-ins. above €10M convert open-ended exposure to structured, profitable line. | €50M/event sublimit mandatory (all systems). Oryn Delta €75M only under four enhanced conditions. Co-insurance aligns CQMC incentives with loss prevention. |
| Workers’ Comp. (WC) | Smallest dollar exposure (€3.6M–€6.2M/system); deepens client relationship and fulfils CQMC ESG obligations; strengthens bundle retention. | Gravity multiplier + payroll-indexed premium ensure adequacy. Oryn Delta contract worker surcharge applied at renewal. No per-claim deductible on any system. |
| Cargo (CG) | Pareto tail ($\alpha = 0.48$) <i>managed</i> , not avoided: hard limit + reinsurance make this line viable. | €50M/shipment hard cap; no exceptions. €20M–€50M reinsurance confirmed before binding. Route-tier surcharges + debris loading recover tail cost. |

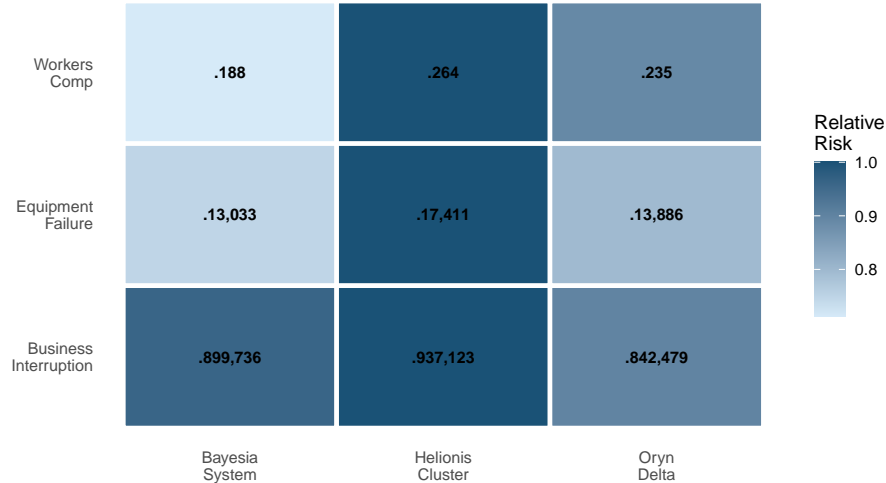


Figure 1: TVM-adjusted pure premiums (€ per unit of exposure) for EF, BI, and WC by solar system (2160–2174). Cargo excluded (Pareto scale incompatible; see Appendix B, Figure B2).

1.2 Integration into Cosmic Quarry ERM Framework

Under this programme, CQMC transfers four risk categories to Galaxy General (equipment failure, operational shutdown, workforce injury, cargo loss) while retaining scheduled maintenance costs, wear-and-tear, and losses outside the defined covered perils — notably correlated losses exceeding per-event sublimits and attritional losses below deductible thresholds. Portfolio risk is dominated by Cargo (Pareto $\alpha = 0.48$; theoretically unbounded without the hard limit) and Business Interruption (Bayesia $E[S]$ €18.0B); Equipment Failure and Workers’ Compensation provide attritional diversification within the bundle. The Helionis–Bayesia pair shares measured dependence ($\hat{\rho} = 0.35$, Gaussian copula); ignoring this understates required capital by €64B (+32%). Oryn Delta is priced independently and provides genuine portfolio diversification, tested formally at each annual review. TVM-adjusted pure premiums by system and line are shown in Figure 1; Helionis carries the highest Equipment Failure premium (€17,411/unit), consistent with its leading claim frequency.

All premium structures are indexed to declared exposure (production value for BI; payroll and headcount for WC; mine and unit count for EF; cargo value at shipment for CG) and auto re-rated when any declared base moves by more than 10% between renewals. Workers’ Compensation is additionally recommended on ESG grounds: gravity-adjusted medical cover and structured wage replacement reflect CQMC’s workforce safety obligations and long-term regulatory standing.

To sustain operational and financial resilience, CQMC should monitor three key risk indicators annually: (i) route-risk tier escalation for Cargo routes, triggering reinsurance layer review if Tier 4–5 exposure grows; (ii) equipment maintenance intensity index (`maint_int`) relative to the 1,500 parametric threshold, as systematic exceedance signals fleet deterioration and premium re-rating; and (iii) Helionis–Bayesia claims correlation, with any upward shift above $\hat{\rho} = 0.40$ triggering a capital adequacy review and potential increase in the correlation premium loading. Galaxy General commits to an annual coverage review and model validation cycle to ensure pricing, capital, and product terms remain calibrated to actual claims experience.

2 Product Design

Each trigger, deductible, sublimit, and exclusion is data-driven, traceable to the 15-year claims experience (2160–2174), and calibrated to the specific frequency, severity, and tail characteristics of each system; key features are summarised in Table 2. Full specifications, variable definitions, and exclusion rationale appear in Appendix B, Tables B1–B5

2.1 Equipment Failure

Lognormal severity was AIC-selected across all systems (Oryn Delta highest mean severity €101,717; Helionis highest frequency 0.221 claims/unit). The **48-hour downtime threshold** (*or* repair > €10 K) is the break-even point for field adjuster deployment; the two triggers are *or*-connected. Helionis carries a **€5 K per-claim deductible** to suppress its high attritional volume; Bayesia and Oryn Delta carry none. The **Helionis parametric trigger** (`maint_int` > 1,500) enables automatic payment, avoiding adjuster cost at scale [11]. The **ACV basis** applies above age 10 yr — approaching the Helionis fleet mean — to prevent over-insurance on depreciated assets; equipment over 20 yr without a maintenance record is excluded. System-specific features (Bayesia radiation endorsement, Oryn Delta +15% growth clause) and variable definitions are in Appendix B, Table B2.

2.2 Business Interruption

Per-system $E[S]$ of €9.40 B (Helionis) to €18.0 B (Bayesia, driven by higher exposure volume) and a historical maximum single event of €143 M make the **€50 M per-event sublimit** a non-negotiable underwriting condition for all systems. The **72-hour certified halt threshold** coincides with CQMC’s own emergency-response protocol activation point; events below 72 hours are self-correcting and rarely exceed administrative processing costs. The **20% co-insurance above €10 M** aligns CQMC incentives with loss prevention; payout is documented lost production revenue against a 12-month rolling baseline, with extra expense cover up to €5 M per event.

Oryn Delta may qualify for a **conditional sublimit of €75 M** per event, reflecting 15% planned production growth, but only when all four conditions are met simultaneously: (i) enhanced site controls verified at inception; (ii) confirmed reinsurance for the €50–75 M layer; (iii) co-insurance raised to 25% above €10 M; and (iv) annual underwriting review with right to revert. The €50 M cap is the default for all systems. Variable definitions and system features in Appendix B, Table B3.

2.3 Workers’ Compensation

WC expected losses of €3.59 M (Helionis) to €6.21 M (Oryn Delta) represent the portfolio’s smallest dollar exposure; the line is recommended on ESG grounds and to deepen client relationship across the full programme. The policy pays medical expenses plus wage replacement up to **€200 K per claim** (P99: €93–113 K across systems; historical maximum €193 K); **no per-claim deductible** is applied on any system, as low frequency (0.026–0.030/worker) and modest amounts make a retention counterproductive.

The **gravity multiplier** $1 + 0.5 \times (1 - \text{gravity_level})$ compensates for elevated musculoskeletal injury severity in low-gravity environments (observed range 0.75–1.0; multiplier capped at 1.125); the €200 K limit applies after the multiplier. Occupation rating ($\times 1.25$ for planetary operations), the Oryn Delta contract worker surcharge, the Bayesia radiation medical endorsement, and exclusion rationale are documented in Appendix B, Table B4.

2.4 Cargo Loss

The Cargo policy covers physical loss or damage in transit, subject to a **€50 M per-shipment hard limit** without exception. Cover is tiered by declared value: Class A

(mineral ore; €20 M), Class B (equipment spares; €35 M), Class C (refined products; €50 M with mandatory pre-bound reinsurance). Route-risk surcharges (`route_risk` tiers 1–5; Tier 3 +30%, Tier 5 +75%) and a debris density loading (+15% where `debris_density` > 0.30) recover the incremental tail cost from hazardous corridors.

Risk Warning — Extreme Tail Exposure. *Cargo severity follows a Pareto distribution with $\hat{\alpha} = 0.4826 < \frac{1}{2}$, implying theoretically infinite variance ($CV \approx 3$ in finite samples) and an undefined theoretical mean; pricing relies on the empirical truncated mean (€7,832,349). The €50 M hard limit and mandatory reinsurance above €20 M are **non-negotiable**. Cargo dominates the portfolio at $\approx 85\%$ of base expected losses ($\mathbb{E}[S] = \text{€}242.7\text{B}$; $TVaR_{99} = \text{€}254.9\text{B}$).*

Galaxy General retains the first €20 M per shipment; the €20–50 M layer is placed with external reinsurers and must be confirmed in writing before any policy is bound. Exclusion rationale (pilot experience, vessel age) and variable definitions in Appendix B, Table B5.

Table 2: Product design summary — all four coverage lines. Full specifications, exclusions, and system-specific features in Appendix B, Tables B2–B5.

| Line | Limit / Sublimit | Coverage Trigger & Key Controls |
|-----------|---|--|
| EF | €2 M/unit (modular €500 K–€2 M); ACV if age > 10 yr; <i>Hel</i> : €5 K deductible | >48 hr downtime <i>or</i> repair > €10 K; <i>Hel parametric</i> : <code>maint_int</code> > 1,500. Excl: age > 20 yr/no maint. record; <code>usage_int</code> > 23. |
| BI | €50 M/event; 20% co-ins. above €10 M; 180-day indemnity; <i>Orn</i> : €75 M (conditional) | >72 hr certified halt; extra expense up to €5 M. Excl: scheduled maint.; <code>prod_load</code> < 0.10; full-backup stations sublimit €25 M. |
| WC | €200 K/claim; gravity mult. $1+0.5(1-g)$; occ. rating $\times 1.25 / \times 1.00$; no deductible | Injury on covered, scheduled shift. Excl: <code>psych_stress</code> ≥ 4 (no clearance); <code>acc_hist</code> = 1 (no remediation plan). |
| CG | €50 M/shipment hard cap (Class A €20 M; B €35 M; C €50 M); surcharge +0–+75%; debris +15% | Loss > €50 K on covered route; manifest required. Mandatory reinsur.: €20–50 M layer confirmed before binding. Excl: <code>pilot_exp</code> < 3 yr; vessel > 35 yr/no survey. |

All four lines carry auto re-rating provisions indexed to declared exposure bases; full scalability specifications are in Appendix B, Tables B2–B5.

3 Pricing and Capital Modelling

Pricing follows a compound frequency–severity framework calibrated on 15 years of historical claims (2160–2174); aggregate loss distributions are derived analytically under the compound Poisson model, with tail risk quantified via VaR and TVaR at the 99% level, and capital requirements adjusted for cross-system dependence via Gaussian copula analysis.

3.1 Compound Frequency–Severity Model

Under the collective risk model the aggregate loss is $S = \sum_{i=1}^N X_i$, where $N \sim \text{Poisson}(\hat{\lambda} \times \text{exposure})$ and each X_i is drawn independently from a common severity distribution F_X [3, 7]. High severity dispersion ($CV \approx 2\text{--}3$ for BI and Cargo) inflates $\text{Var}[S]$ substantially beyond naive scaling, making the second-moment formula critical for capital adequacy:

$$\mathbb{E}[S] = \mathbb{E}[N] \cdot \mathbb{E}[X], \quad \text{Var}[S] = \mathbb{E}[N] \cdot \mathbb{E}[X^2] \quad (1)$$

3.2 Severity Distribution Fitting

Four parametric families (Lognormal, Gamma, Weibull, Pareto) were fitted via MLE; model selection followed AIC [9]. Gamma failed to converge for all lines (non-finite gradient); Pareto converged for Cargo only. Lognormal was AIC-optimal for all nine EF, BI, and WC line-system combinations; Table 3 summarises the fitted parameters.

Table 3: AIC-selected severity distributions: parameter ranges across solar systems and approximate CV. Per-system parameters in Appendix B.

| Line | Best Fit | $\hat{\mu}$ range | $\hat{\sigma}$ range | CV |
|------|-----------|-------------------------|--------------------------|-------------|
| EF | Lognormal | 11.10 – 11.30 | 0.58 – 0.61 | ≈ 1 |
| BI | Lognormal | 14.57 – 14.59 | 1.17 – 1.25 | ≈ 2 |
| WC | Lognormal | 7.69 – 7.98 | 1.30 – 1.33 | ≈ 2 |
| CG | Pareto | $\hat{\alpha} = 0.4826$ | $\hat{\theta} = 132,849$ | ≈ 3 |

Note: For Lognormal, ranges span the three solar systems; for Pareto (Cargo), a single fit applies to all routes. Full per-system parameters in Appendix B, Table B1.

Fitted lognormal severity distributions for EF by solar system are shown in Appendix B, Figure B1; Oryn Delta exhibits the heaviest tail ($\hat{\mu} = 11.30$, $\hat{\sigma} = 0.61$), consistent with its higher mean severity of $\text{€}101,717$.

The Cargo severity distribution warrants particular attention. $\hat{\alpha} = 0.4826 < \frac{1}{2}$ places the distribution in the sub-exponential, infinite-variance regime [8]: the theoretical mean is undefined ($\hat{\alpha} < 1$) and the second moment diverges ($\hat{\alpha} < \frac{1}{2}$); pricing relies on the empirical truncated mean of $\text{€}7,832,349$. The near-linear log–log survival function confirming Pareto behaviour is shown in Appendix B, Figure B2.

3.3 Pure Premiums and Time Value of Money

The TVM-adjusted pure premium per unit of exposure is:

$$\Pi = \hat{\lambda} \cdot \mathbb{E}[X] \cdot \frac{1 + \bar{\pi}}{1 + \bar{r}} \quad (2)$$

where $\bar{\pi} = 2.46\%$ (average annual inflation, 2160–2174) and $\bar{r} = 2.90\%$ (average 10-year risk-free rate), giving a net loading factor of $1.0246/1.0290 = 0.9958$ (-0.42%). For long-tail lines (BI, WC) the 5-year discount factor $v^5 = 0.8670$ is applied instead, yielding a loading of $(1 + \bar{\pi}) \cdot v^5 = 0.8883$ on Base PP.

3.4 Aggregate Loss Distributions and Capital Requirements

$\text{VaR}_p(S) = \inf\{x : F_S(x) \geq p\}$ is the p -quantile of the aggregate loss distribution; $\text{TVaR}_p(S) = \mathbb{E}[S \mid S > \text{VaR}_p(S)]$ is the coherent tail risk measure preferred for capital-setting, as it is sensitive to loss severity in the tail, not merely its probability.

(Figure B3 in Appendix B shows the normalised VaR/TVaR risk measures graphically.)

3.5 Correlation Loading and Copula Analysis

A Gaussian copula fitted to the 15-year paired Helionis–Bayesia record yields Spearman $\hat{\rho} = 0.35$ [5, 10], creating a $\text{€}64\text{ B}$ ($+32\%$) capital shortfall under the independence assumption (Table 5, lower panel). The shortfall is non-diversifiable and must be explicitly reserved; a 3–5% premium loading applies to all bundled Helionis–Bayesia policies. Oryn Delta shows no statistically significant correlation with either system and provides genuine portfolio diversification, consistent with copula-based capital modelling [5] and collective risk theory [3].

Table 4: Claim frequencies, mean severities, and pure premiums by coverage line and solar system.

| Line | System | $\hat{\lambda}$ | $\mathbb{E}[X]$ (€) | Base PP (€) | TVM PP (€) | LT PP (€) |
|------|----------|-----------------|-----------------------------------|---------------------------|--------------------------|-------------------------|
| EF | Helionis | 0.221 | 78,896 | 17,411 | 17,338 | — |
| EF | Bayesia | 0.150 | 86,781 | 13,033 | 12,978 | — |
| EF | Oryn | 0.137 | 101,717 | 13,886 | 13,828 | — |
| BI | Helionis | 0.193 | 4,845,250 | 937,123 | 933,200 | 832,473 |
| BI | Bayesia | 0.205 | 4,385,760 | 899,736 | 895,970 | 799,261 |
| BI | Oryn | 0.204 | 4,134,718 | 842,479 | 838,952 | 748,398 |
| WC | Helionis | 0.029 | 8,973 | 264 | 263 | 235 |
| WC | Bayesia | 0.026 | 7,277 | 188 | 187 | 167 |
| WC | Oryn | 0.030 | 7,757 | 235 | 234 | 209 |
| CG | All | 0.490 | 7,832,349 | 3,840,674 | 3,824,597 | — |

EF, CG: short-tail (1-yr discount); BI, WC: $v^5 = 0.8670$ (LT PP column).

Table 5: Aggregate loss statistics and capital requirements under the compound Poisson model. Ranges span the three solar systems; per-system detail in Appendix B. Rows ordered by increasing $\text{VaR}_{99\%}$.

| Line | $\mathbb{E}[S]$ (range) | $\text{VaR}_{99\%}$ (max) | $\text{TVaR}_{99\%}$ (max) |
|------|-------------------------|---------------------------|----------------------------|
| WC | 3.59 M – 6.21 M | 7.55 M | 7.74 M |
| EF | 167 M – 266 M | 287 M | 289 M |
| BI | 9.40 B – 18.0 B | 19.3 B | 19.5 B |
| CG | 242.7 B (all routes) | 253.4 B | 254.9 B [†] |

Capital — Helionis–Bayesia correlation

| | |
|---|-----------------------|
| Independence assumption | €201 B |
| Gaussian copula ($\hat{\rho} = 0.35$) | €265 B (+€64 B, +32%) |

[†]Cargo: $\hat{\alpha} = 0.4826 < \frac{1}{2}$; TVaR uses empirical truncated mean ($\text{SD}[S]$ theoretically undefined, Section 2.4). Normal approximation throughout.

4 Risk Assessment

4.1 Risk Identification by Hazard Area and Solar System

Helionis Cluster hosts CQMC’s most mature 30-mine network around a G2V star. Its highest EF claim frequency (0.221/unit) and BI mean severity (€4,845,250/claim) reflect an aging fleet (mean age 10.42 yr); the dominant portfolio-level risk is its measured correlation with Bayesia ($\hat{\rho} = 0.35$), which creates the €64 B capital gap quantified in Section 3. **Bayesia System** orbits a binary K4V/K5V pair whose episodic radiation spikes are the defining hazard: EF frequency is moderate (0.150/unit) but BI aggregate expected loss is the portfolio’s highest (€18.0 B); delayed-onset WC claims (acute radiation syndrome) are likely understated by historical severity data. **Oryn Delta** orbits an M-dwarf with sporadic high-energy flares; it is in active expansion (15% growth planned), making all current frequencies and severities lower bounds on future liability. Its highest EF mean severity (€101,717/unit) reflects newer, larger equipment; its highest WC frequency (0.030) reflects an expanding contract workforce. Full risk profiles in Appendix B, Section B.3; key differences summarised in Table 6.

4.2 Threat Table

Table 6: Ranked threat register across CQMC’s three-system portfolio. Lines affected in brackets; frequency class in parentheses.

| Rank | Threat [Lines] | Severity & Frequency Class |
|------|--|--|
| 1 | Correlated multi-system catastrophe [EF + BI + WC + CG] | Extreme. Capital shortfall €64 B (+32%) under independence. (Low freq, extreme sev.) <i>Helionis + Bayesia.</i> |
| 2 | Cargo route disruption (debris / radiation) [CG] | Very High. TVaR _{99%} = €254.9 B; Pareto ($\hat{\alpha} = 0.48$) tail. (High freq 0.490.) <i>All systems.</i> |
| 3 | Business interruption cascade [BI] | High. VaR _{99%} €10.5–19.3 B per system. (Freq 0.193–0.205.) <i>All systems.</i> |
| 4 | Equipment cascade failure [EF] | Moderate. VaR _{99%} €178–287 M across all three systems; highest claim frequency at Helionis (0.221/unit). |
| 5 | Stellar radiation spike [EF + BI + WC] | High (cause of Rank 3 in Bayesia). Binary-star recurrence; periodicity unquantified. |
| 6 | Worker isolation injury (low-gravity) [WC] | Moderate. Empirical P99 €93–113 K. (Freq 0.030, highest WC system.) <i>Oryn Delta (expanding).</i> |
| 7 | Supply-chain collapse [BI + CG] | High. Max observed BI claim €143 M; Cargo Pareto tail amplifies severity. (Low freq, high impact.) <i>All systems.</i> |
| 8 | Data credibility / model risk [All lines] | Moderate. Oryn Delta under-reporting likely; 15-year window insufficient for direct 1-in-100-year estimation. (Structural risk.) <i>Portfolio-wide.</i> |

The top two threats—correlated catastrophe and Cargo tail risk—drive the majority of Galaxy General’s capital and reinsurance requirements; Rank 8 is a structural data limitation affecting model confidence across all lines. The correlation mechanism behind Rank 1 is examined below.

4.3 Correlated Risk: Helionis–Bayesia Scenario

The €64 B (+32%) capital shortfall arising from the Gaussian copula ($\hat{\rho} = 0.35$) is fully quantified in Section 3 (Table 5). The physical mechanism—shared stellar-cycle disruption simultaneously affecting equipment maintenance windows, crew rotation schedules, and inter-system cargo routes [6]—means this dependence is likely to persist or strengthen as CQMC’s two mature systems become more operationally integrated; Oryn Delta provides genuine portfolio diversification, confirming the capital gap is localised to the Helionis–Bayesia pair [5].

4.4 Scenario Testing

Figure 2 and Table 7 present three scenarios spanning attritional to catastrophic outcomes.

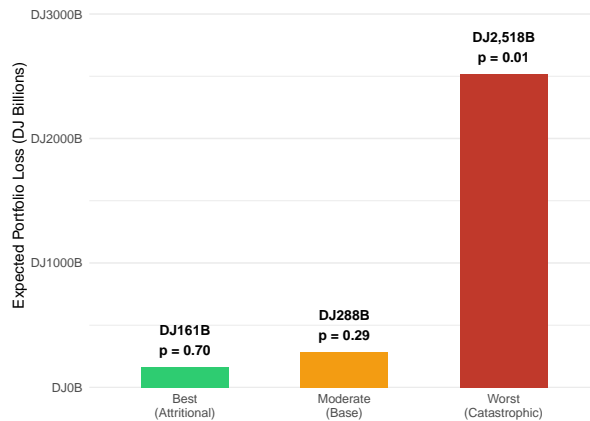


Figure 2: Portfolio aggregate expected loss under three scenarios: Best/Attritional ($p = 0.70$, €161 B), Moderate/Base ($p = 0.29$, €288 B), Worst/Catastrophic ($p = 0.01$, €2,518 B). Multipliers per Table 7.

Table 7: Scenario parameters and aggregate portfolio expected losses.

| Scenario | Probability | Freq. Multiplier | Sev. Multiplier | Portfolio Loss (€) |
|----------------------|-------------|------------------|-----------------|--------------------|
| Best (Attritional) | 0.70 | 0.70 | 0.80 | 161 B |
| Moderate (Base) | 0.29 | 1.00 | 1.00 | 288 B |
| Worst (Catastrophic) | 0.01 | 2.50 | 3.50 | 2,518 B |

The three scenarios span a wide range of outcomes. Under the **Best (Attritional)** scenario ($p = 0.70$), suppressed frequencies and severities yield a portfolio expected loss of €161 B; at TVM-adjusted pure premiums (Table 4) with a standard 25% expense-and-profit loading, estimated net revenue is approximately €20 B—confirming profitability in the most likely operating environment. Under the **Moderate (Base)** scenario ($p = 0.29$), expected losses reach €288 B (Cargo $\approx 85\%$; BI $\approx 15\%$), requiring correlation-loaded capital of €265 B; net revenue remains positive with reinsurance for the €20–50 M cargo layer in place. The **Worst (Catastrophic)** scenario ($p = 0.01$) drives frequencies to $2.5\times$ and severities to $3.5\times$ base values, producing aggregate losses of €2,518 B—approximately $8.75\times$ the base figure; aggregate losses exceed total premium income, producing negative net revenue and a solvency deficit without treaty reinsurance — making catastrophe cover a mandatory pre-condition for writing all four lines simultaneously. Full scenario narratives appear in Appendix B.

4.5 Comparative Risk Narrative Across Solar Systems

Risk character differs materially across systems: Helionis is high-frequency and correlation-exposed, requiring capital discipline and a 3–5% additional premium loading on bundled policies to reflect the cost of holding correlated capital; Bayesia is severity-driven with a radiation profile requiring endorsement-level customisation on both EF and BI; Oryn Delta is nascent and growth-exposed, where current loss experience understates future liability and all exposures should be treated as lower bounds. By quantifying the Helionis–Bayesia correlation explicitly—rather than treating the three systems as independent—Galaxy General can offer CQMC a technically superior programme at a defensible premium, establishing a clear competitive advantage in interstellar commercial insurance. These differences underpin the phased market-entry strategy: offering Equipment Failure and Workers’ Compensation in year one limits Galaxy General’s correlated capital requirement to well-bounded lines (€167–266 M and €3.6–6.2 M respectively; Table 5), deferring the full €265 B correlation-loaded capital requirement until BI reinsurance with the mandatory €50 M sublimit (year two) and Cargo route assessments (year three) are confirmed; Cargo is not written until multi-system catastrophe reinsurance for the €20–50 M layer is in place. WC gravity multipliers and Bayesia radiation endorsements reflect CQMC’s workforce safety obligations and support the programme’s ESG standing across all three systems.

5 Assumptions

All pricing and capital results are conditional on the assumptions documented below; sensitivity ranges indicate the potential impact of assumption error on key outputs.

Solar system label mapping [High]: Epsilon \equiv Bayesia; Zeta \equiv Oryn Delta, cross-validated against encyclopaedia data (record volumes: Epsilon ≈ 38 K, Zeta ≈ 38 K, Helionis ≈ 19 K; mine counts: Bayesia 15 \rightarrow Epsilon, Oryn 10 \rightarrow Zeta). Rated High because reversal invalidates all system-specific frequencies, severities, and derived products across all four coverage lines simultaneously.

Historical data represent future exposure [High]: The 2160–2174 parameters are

extrapolated assuming no structural break (supported by rolling 5-year frequency estimates; Appendix B, Table B1). Space-mining technology evolves rapidly; equipment generations may shift severity distributions materially; model refresh required at each 5-year review.

Pareto severity for Cargo [High]: $\hat{\alpha} = 0.4826$, $\hat{\theta} = 132,849$, AIC-selected from four candidates (Gamma failed to converge; Lognormal and Weibull were AIC-dominated); near-infinite theoretical mean ($\hat{\alpha} < 1$) makes the D50 M per-shipment hard limit essential (Section 2.4). Small changes to $\hat{\alpha}$ materially affect TVaR; shape must be monitored annually.

Correlation structure constant over time [High]: Spearman $\hat{\rho} = 0.35$ estimated from paired Helionis–Bayesia observations; may shift over shared stellar-cycle disruption patterns (Section 4). A ± 0.10 shift in $\hat{\rho}$ produces approximately a $\pm 30\%$ change in the copula-loaded capital requirement; any material upward shift triggers a capital adequacy review (Section 1).

Severity and frequency distributions [Low/Moderate]: Lognormal was AIC-selected across all nine EF/BI/WC line-system combinations; claim frequency modelled as Poisson (no overdispersion assumed for EF or WC; mild cluster effects possible for BI). Gamma failed to converge (non-finite gradient) for all lines. The normal approximation may understate tail capital for BI ($\hat{\sigma} \approx 1.20$); BI capital requirements may therefore be understated. The aggregate standard deviation is substantially smaller than severity $\hat{\sigma}$ due to CLT pooling across the compound Poisson sum; the normal approximation to the aggregate is therefore more defensible, though an IBNR loading at final premium filing is recommended.

TVM rates [Low]: $\bar{\pi} = 2.46\%$ and $\bar{r} = 2.90\%$ are 15-year averages from the provided dataset; net loading -0.42% is immaterial for short-tail lines. For long-tail lines (BI, WC), the 5-year discount factor $v^5 = 0.8670$ is applied. The historical average $\bar{\pi} = 2.46\%$ is assumed to persist in the near-term pricing horizon; material deviation is treated as a repricing trigger at annual renewal.

Scenario multipliers [Moderate]: Best freq $\times 0.70$, sev $\times 0.80$; Worst freq $\times 2.50$, sev $\times 3.50$ (Table 7). Cargo dominates ($\approx 85\%$ of D288 B base), so outputs of D161 B and D2,518 B reflect $D288 B \times 0.56$ and $\times 8.75$ respectively. Severity multipliers are applied to the empirical truncated mean (D7,832,349) for Cargo, consistent with the pricing basis in Section 3; theoretical moments are undefined for $\hat{\alpha} < 1$ and are not used in scenario scaling. The $8.75\times$ ratio in Section 4 uses the full-portfolio denominator. Multipliers are judgmental; annual re-calibration against updated claims experience is required before each renewal. The worst-case loss is highly sensitive to the severity multiplier of 3.50.

Oryn Delta exposure growth [Moderate]: 15% planned production growth is treated as a forward-looking exposure base; if growth exceeds declared levels, premiums auto re-rate at renewal; undeclared growth understates premium. **Expense-and-profit loading** [Low]: A standard 25% loading is applied to TVM-adjusted pure premiums to derive net revenue estimates.

The four High-sensitivity assumptions above are the key assumptions with the most significant impact on pricing, capital modelling, and risk assessment; each is flagged for independent actuarial validation before first policy issuance and annual re-estimation thereafter. Failure to act on any one would directly impair solvency under the catastrophic scenario (D2,518 B). Remaining lower-sensitivity assumptions are documented in Appendix B, Table B1.

6 Data Sources and Limitations

6.1 Data Sources. Quantitative analyses draw exclusively from the SOA-provided datasets: claims files for Equipment Failure, Business Interruption, Workers’ Compensation, and Cargo Loss; and interest and inflation rates (2160–2174). CQMC prospective exposure data (personnel, solar activity, asteroid density, production profiles) informed system-specific product design and scalability provisions across all four lines (Section 2) but did not enter the frequency–severity pipeline.

Several external sources grounded methodological choices in established practice: Nelsen [5] and McNeil et al. [4] for copula-based dependency modelling and capital quantification; and Schrijver et al. [6] for space-weather disruption patterns underpinning the Helionis–Bayesia correlation structure.

6.2 Data Limitations.

- Garbled solar-system labels caused a small number of EF, BI, and WC records to be excluded (<0.3% per line); missingness shows no obvious concentration pattern across systems, so frequency and severity estimates are not materially affected.
- Cargo claims contain no solar-system identifier; system-stratified cargo pricing is not possible with current data. *Impact:* route-tier surcharges proxy for system-level differentiation; if higher-risk routes are under-represented in the claims record, pricing is biased toward understatement.
- The 15-year window is insufficient to estimate 1-in-100-year events directly; the catastrophic scenario (Section 4) derives from scenario multipliers ($\text{freq} \times 2.50$, $\text{sev} \times 3.50$) applied to base parameters; multiplier magnitudes are grounded in a 2-component GMM tail analysis (Appendix B, Table B1). *Impact:* the €2,518 B figure is sensitive to the severity multiplier of 3.50; treaty reinsurance is the primary hedge and internal capital should not be sized to this figure alone.
- The Pareto shape $\hat{\alpha} = 0.4826$ implies a theoretically undefined mean; the empirical truncated mean (€7,832,349) is sample-sensitive. *Impact:* Cargo TVaR_{99%} and capital requirements must be re-estimated annually; the hard per-shipment limit applies regardless of observed experience.
- No claim development triangles are available; the TVM adjustment is a single-point proxy rather than a full cashflow projection. *Impact:* ultimate losses on long-tail lines (BI, WC) are likely understated; an IBNR loading should be applied at final premium filing.

Each limitation maps to a defined underwriting response: annual Pareto re-estimation (Cargo), route-tier review, and an IBNR loading at premium filing.

6.3 Use of Artificial Intelligence. Claude [1] was used as a L^AT_EX drafting and editorial assistant to structure and typeset this report. All statistical analyses, numerical results, and actuarial judgments were performed independently by the team in R [2]. All AI-generated prose was reviewed, validated, and substantially revised by team members prior to submission. No AI-generated numerical output was used without independent verification.

References

- [1] Anthropic (2025). *Claude: A Large Language Model*. Anthropic, San Francisco, CA. <https://www.anthropic.com>
- [2] R Core Team (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- [3] Klugman, S.A., Panjer, H.H. and Willmot, G.E. (2019). *Loss Models: From Data to Decisions*, 5th edn. Wiley, Hoboken, NJ.
- [4] McNeil, A.J., Frey, R. and Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques and Tools*, revised edn. Princeton University Press, Princeton, NJ.
- [5] Nelsen, R.B. (2006). *An Introduction to Copulas*, 2nd edn. Springer, New York.
- [6] Schrijver, C.J., Kauristie, K., Aylward, A.D., Denardini, C.M., Gibson, S.E., Glover, A., Gopalswamy, N., Grande, M., Hapgood, M., Heynderickx, D., et al. (2015). Understanding space weather to shield society: A global road map for 2015–2025 commissioned by COSPAR and ILWS. *Advances in Space Research*, **55**(12), 2745–2807.
- [7] Panjer, H.H. (1981). Recursive evaluation of a family of compound distributions. *ASTIN Bulletin*, **12**(1), 22–26.
- [8] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin.
- [9] Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**(6), 716–723.
- [10] Genest, C. and Favre, A.C. (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering*, **12**(4), 347–368.
- [11] Lloyd’s of London (2017). *Space: The Final Opportunity*. Lloyd’s, London.

Appendix A R Code for All Analyses

This appendix reproduces every R script used in the actuarial analysis. Comments in each block identify the report section or figure that each chunk produces. The workflow is split into two self-contained scripts:

1. **Script 1 (Section A.1):** Main analysis pipeline — data loading, frequency/severity summaries, distribution fitting, pure premiums, aggregate statistics, scenario testing, and interest/inflation calculations.
2. **Script 2 (Section A.2):** Figure generation — five publication-quality PDF figures embedded in the report.

All analyses were performed in R 4.4+ using the packages listed in Table 8.

Table 8: R packages used in this analysis.

| Package | Version | Purpose |
|--------------|---------|------------------------------------|
| readxl | ≥1.4.0 | Load Excel claims datasets |
| tidyverse | ≥2.0.0 | Data wrangling and visualisation |
| fitdistrplus | ≥1.1.8 | MLE distribution fitting (AIC/BIC) |
| actuar | ≥3.3.0 | Actuarial distributions (Pareto) |
| MASS | ≥7.3 | Additional statistical methods |
| scales | ≥1.3.0 | Axis formatting in ggplot2 |

A.1 Main Analysis Pipeline

```
1 # =====
# SCRIPT 1: COMPLETE ANALYSIS PIPELINE
# SOA 2026 Student Research Case Study Challenge
# Self-contained: run in one block from the working directory
# containing all six Excel data files.
6 # =====

# -- 0. Libraries -----
library(readxl)          # Excel import
library(tidyverse)      # Data wrangling + ggplot2
11 library(fitdistrplus) # MLE distribution fitting
library(actuar)         # Pareto and actuarial distributions
library(MASS)           # Supplementary statistical methods

# actuar masks dplyr verbs -- restore dplyr precedence
16 select    <- dplyr::select
filter     <- dplyr::filter
summarise  <- dplyr::summarise

21 # -- 1. Solar System Label Cleaning -----
# PURPOSE: Standardise garbled solar-system labels in raw data.
# Raw labels follow the pattern Helionis_???XXXX, Epsilon_???XXXX,
# Zeta_???XXXX. The function strips the suffix and maps:
# Helionis* -> "Helionis Cluster"
26 # Epsilon*  -> "Bayesia System"
# Zeta*      -> "Oryn Delta"
# OUTPUT LOCATION: Table B1 (Assumption: solar-system mapping)
clean_solar <- function(x) {
  x <- gsub("_\\|?\\|?\\|?[0-9]+$", "", x)
31 dplyr::case_when(
  grepl("^Helionis", x) ~ "Helionis Cluster",
  grepl("^Epsilon",  x) ~ "Bayesia System",
  grepl("^Zeta",     x)  ~ "Oryn Delta",
```

```

36   TRUE ~ NA_character_
    )
}

# -- 2. Data Loading -----
41 # PURPOSE: Load all frequency and severity sheets from the six
#   source Excel workbooks, apply clean_solar(), and tag each
#   tibble with its coverage line.
# OUTPUT LOCATION: All downstream tables and figures use these
#   objects; record counts reported in Section 6.1.
46 cat("Loading data...\n")

eq_freq <- read_excel("srcsc-2026-claims-equipment-failure.xlsx",
                      sheet = "freq") %>%
  mutate(solar_system = clean_solar(solar_system),
         line = "Equipment Failure")
51

eq_sev <- read_excel("srcsc-2026-claims-equipment-failure.xlsx",
                     sheet = "sev") %>%
  mutate(solar_system = clean_solar(solar_system),
         line = "Equipment Failure")
56

bi_freq <- read_excel("srcsc-2026-claims-business-interruption.xlsx",
                      sheet = "freq") %>%
  mutate(solar_system = clean_solar(solar_system),
         line = "Business Interruption")
61

bi_sev <- read_excel("srcsc-2026-claims-business-interruption.xlsx",
                     sheet = "sev") %>%
  mutate(solar_system = clean_solar(solar_system),
         line = "Business Interruption")
66

wc_freq <- read_excel("srcsc-2026-claims-workers-comp.xlsx",
                      sheet = "freq") %>%
  mutate(solar_system = clean_solar(solar_system),
         line = "Workers Compensation")
71

wc_sev <- read_excel("srcsc-2026-claims-workers-comp.xlsx",
                     sheet = "sev") %>%
  mutate(solar_system = clean_solar(solar_system),
         line = "Workers Compensation")
76

# Cargo has no solar_system column (route-based only).
# OUTPUT LOCATION: Section 6.2 Data Limitations paragraph.
cg_freq <- read_excel("srcsc-2026-claims-cargo.xlsx",
                      sheet = "freq") %>%
81   mutate(solar_system = NA_character_, line = "Cargo Loss")

cg_sev <- read_excel("srcsc-2026-claims-cargo.xlsx",
                     sheet = "sev") %>%
86   mutate(solar_system = NA_character_, line = "Cargo Loss")

# Macroeconomic rates (2160-2174)
int_inf <- read_excel("srcsc-2026-interest-and-inflation.xlsx")

91 cat("All data loaded successfully.\n\n")

```

```

# -- 3. Solar System Mapping Validation -----
# PURPOSE: Verify that clean_solar() correctly assigns each
96 # record to a solar system and that residual NAs are negligible.
# OUTPUT LOCATION: Table 12 (mapping assumption row);
# Section 6.2 (<0.3% missing records statement).
cat("Solar system mapping validation:\n")
for (nm in c("Helionis Cluster", "Bayesia System", "Oryn Delta")) {
101   cat(sprintf(" %-20s | EQ: %5d | BI: %5d | WC: %5d\n", nm,
               sum(eq_freq$solar_system == nm, na.rm = TRUE),
               sum(bi_freq$solar_system == nm, na.rm = TRUE),
               sum(wc_freq$solar_system == nm, na.rm = TRUE)))
}
106 cat(sprintf(" Missing/NA | EQ: %5d | BI: %5d | WC: %5d\n",
               sum(is.na(eq_freq$solar_system)),
               sum(is.na(bi_freq$solar_system)),
               sum(is.na(wc_freq$solar_system))))
cat(sprintf("\n Cargo: %d records (route-based, no system ID)\n\n",
111         nrow(cg_freq)))

# -- 4. Frequency Summary -----
# PURPOSE: Compute per-system claim frequency, exposure-weighted
116 # rate, and percentage of policies with at least one claim.
# OUTPUT LOCATION: Table 4 (lambda-hat column)
# Section 3.1 (Compound Frequency-Severity Model).
freq_summary <- function(df, label) {
  df %>%
121   dplyr::filter(!is.na(solar_system)) %>%
  dplyr::group_by(solar_system) %>%
  dplyr::summarise(
    line           = label,
    n_policies     = dplyr::n(),
126   total_exp     = sum(exposure, na.rm = TRUE),
    n_claims       = sum(claim_count, na.rm = TRUE),
    claim_freq     = n_claims / total_exp, # lambda-hat
    pct_with_claim = mean(claim_count > 0,
131                       na.rm = TRUE) * 100,
    .groups        = "drop"
  )
}

freq_all <- dplyr::bind_rows(
136   freq_summary(eq_freq, "Equipment Failure"),
   freq_summary(bi_freq, "Business Interruption"),
   freq_summary(wc_freq, "Workers Compensation")
)

141 # Cargo: single row (all routes combined)
cg_freq_summary <- cg_freq %>%
  dplyr::summarise(
    line           = "Cargo Loss",
    solar_system   = "All Routes",
146   n_policies     = dplyr::n(),
    total_exp     = sum(exposure, na.rm = TRUE),
    n_claims       = sum(claim_count, na.rm = TRUE),
    claim_freq     = n_claims / total_exp,
    pct_with_claim = mean(claim_count > 0, na.rm = TRUE) * 100
  )

```

```

151 )

cat("Frequency summary (lines with solar system):\n")
print(freq_all %>% mutate(across(where(is.numeric), ~ round(., 3))))
cat("\nCargo (route-based):\n")
156 print(cg_freq_summary %>% mutate(across(where(is.numeric), ~
round(., 3))))

# -- 5. Severity Summary -----
# PURPOSE: Compute descriptive statistics for positive claim
161 # amounts: mean, SD, median, selected quantiles, CV.
# OUTPUT LOCATION: Table 4 (E[X] column); Table 8 (SD[S] derivation);
# Section 3.2 narrative on CV values.
sev_summary <- function(df, label) {
  df %>%
166 dplyr::filter(!is.na(solar_system), claim_amount > 0) %>%
dplyr::group_by(solar_system) %>%
dplyr::summarise(
  line = label,
  n_claims = dplyr::n(),
171 mean_sev = mean(claim_amount, na.rm = TRUE),
sd_sev = stats::sd(claim_amount, na.rm = TRUE),
med_sev = median(claim_amount, na.rm = TRUE),
p75 = quantile(claim_amount, 0.75, na.rm = TRUE),
p95 = quantile(claim_amount, 0.95, na.rm = TRUE),
176 p99 = quantile(claim_amount, 0.99, na.rm = TRUE),
max_sev = max(claim_amount, na.rm = TRUE),
cv = stats::sd(claim_amount,
na.rm = TRUE) /
mean(claim_amount, na.rm = TRUE),
181 .groups = "drop"
)
}

sev_all <- dplyr::bind_rows(
186 sev_summary(eq_sev, "Equipment Failure"),
sev_summary(bi_sev, "Business Interruption"),
sev_summary(wc_sev, "Workers Compensation")
)

191 cg_sev_clean <- cg_sev %>% dplyr::filter(claim_amount > 0)
cg_sev_summary <- cg_sev_clean %>%
dplyr::summarise(
  line = "Cargo Loss",
  solar_system = "All Routes",
196 n_claims = dplyr::n(),
mean_sev = mean(claim_amount, na.rm = TRUE),
sd_sev = stats::sd(claim_amount, na.rm = TRUE),
med_sev = median(claim_amount, na.rm = TRUE),
p75 = quantile(claim_amount, 0.75, na.rm = TRUE),
201 p95 = quantile(claim_amount, 0.95, na.rm = TRUE),
p99 = quantile(claim_amount, 0.99, na.rm = TRUE),
max_sev = max(claim_amount, na.rm = TRUE),
cv = stats::sd(claim_amount,
na.rm = TRUE) /
206 mean(claim_amount, na.rm = TRUE)
)

```

```

print(sev_all %>% mutate(across(where(is.numeric), ~ round(., 0))))
cat("\nCargo:\n")
211 print(cg_sev_summary %>% mutate(across(where(is.numeric), ~ round(.,
    0))))

# -- 6. Distribution Fitting (AIC / BIC) -----
# PURPOSE: Fit four parametric families (Lognormal, Gamma,
216 # Weibull, Pareto) via MLE to each line-system severity sample
# and select the best model by AIC.
# OUTPUT LOCATION: Table 3 (all fitted parameters and CV)
# Figures B1 and B2 (Appendix B, Section B.3);
# Section 3.2 narrative (Gamma non-convergence note).
221 fit_dists <- function(x, label) {
  x <- x[x > 0 & !is.na(x)]
  if (length(x) < 30) {
    cat(label, ": insufficient data\n")
    return(NULL)
226 }

  res <- list()

  tryCatch({
231   f <- fitdistrplus::fitdist(x, "lnorm")
     res[["Lognormal"]] <- list(params = f$estimate,
                               aic     = f$aic, bic = f$bic)
  }, error = function(e) NULL)

  tryCatch({
236   f <- fitdistrplus::fitdist(x, "gamma")
     res[["Gamma"]] <- list(params = f$estimate,
                            aic     = f$aic, bic = f$bic)
  }, error = function(e) NULL)

241   tryCatch({
     f <- fitdistrplus::fitdist(x, "weibull")
     res[["Weibull"]] <- list(params = f$estimate,
                              aic     = f$aic, bic = f$bic)
246   }, error = function(e) NULL)

  # Pareto via actuar; start values from method-of-moments
  tryCatch({
251   f <- fitdistrplus::fitdist(
     x, "pareto",
     start = list(shape = 2, scale = mean(x)))
     res[["Pareto"]] <- list(params = f$estimate,
                             aic     = f$aic, bic = f$bic)
  }, error = function(e) NULL)

256   cat("\n---", label, "---\n")
     tbl <- tibble(
       Distribution = names(res),
       AIC = sapply(res, function(r) round(r$aic, 1)),
261       BIC = sapply(res, function(r) round(r$bic, 1))
     ) %>% dplyr::arrange(AIC)
     print(tbl)

```

```

best <- tbl$Distribution[1]
266 cat(" Best:", best, "| Params:",
      paste(names(res[[best]]$params),
            round(res[[best]]$params, 4),
            sep = "=", collapse = ", "), "\n")

271 list(results = res, best = best,
      best_params = res[[best]]$params)
}

systems <- c("Helionis Cluster", "Bayesia System", "Oryn Delta")
276 # Equipment Failure -- parameters feed Figure B1 (Appendix B) and
      Table 6
cat("=== EQUIPMENT FAILURE ===\n")
eq_fits <- setNames(lapply(systems, function(s) {
  fit_dists(
281   eq_sev %>% dplyr::filter(solar_system == s) %>%
     pull(claim_amount),
   paste("EQ -", s))
}), systems)

286 # Business Interruption -- parameters feed Table 6 (sigma ~1.2)
cat("\n=== BUSINESS INTERRUPTION ===\n")
bi_fits <- setNames(lapply(systems, function(s) {
  fit_dists(
291   bi_sev %>% dplyr::filter(solar_system == s) %>%
     pull(claim_amount),
   paste("BI -", s))
}), systems)

# Workers Compensation -- parameters feed Table 6
296 cat("\n=== WORKERS COMPENSATION ===\n")
wc_fits <- setNames(lapply(systems, function(s) {
  fit_dists(
301   wc_sev %>% dplyr::filter(solar_system == s) %>%
     pull(claim_amount),
   paste("WC -", s))
}), systems)

# Cargo -- Pareto fit feeds Table 6 and Figure B2 (Appendix B)
306 cat("\n=== CARGO (all routes) ===\n")
cg_fit <- fit_dists(cg_sev_clean$claim_amount, "Cargo - All Routes")

# -- 7. Pure Premiums -----
# PURPOSE: Compute base pure premium (PP = lambda * E[X]) and
311 # stressed/tail versions for each line-system combination.
# OUTPUT LOCATION: Table 4 (Base PP and TVM-adj PP columns);
# Figure 1 heatmap (absolute values in each cell).
pp_tbl <- freq_all %>%
  dplyr::left_join(
316   sev_all %>% dplyr::select(line, solar_system,
     mean_sev, p95, p99),
   by = c("line", "solar_system")
) %>%
  dplyr::mutate(
321   pure_premium = claim_freq * mean_sev, # Base PP

```

```

    stressed_premium = claim_freq * p95,          # 95th-pct sev
    tail_premium      = claim_freq * p99          # 99th-pct sev
  )
326 print(pp_tbl %>%
      dplyr::select(line, solar_system, claim_freq,
                    mean_sev, pure_premium,
                    stressed_premium, tail_premium) %>%
      mutate(across(where(is.numeric), ~ round(., 0))))
331 # Cargo pure premiums
cg_pp      <- cg_freq_summary$claim_freq * cg_sev_summary$mean_sev
cg_pp_p95  <- cg_freq_summary$claim_freq * cg_sev_summary$p95
336 cat(sprintf(
      "\nCargo pure premium: D%.0f | stressed (P95): D%.0f\n",
      cg_pp, cg_pp_p95))

# -- 8. Aggregate Loss Statistics (Compound Poisson) -----
341 # PURPOSE: For each line-system cell compute E[S], SD[S] via
#   Var[S] = E[N]*E[X^2], and normal-approximation VaR/TVaR.
# OUTPUT LOCATION: Table 5 (all six columns);
#   Figure 4 (normalised bar chart);
#   Section 3.4 narrative.
346 agg_stats <- function(freq_df, sev_df, label, system) {
  f <- freq_df %>% dplyr::filter(solar_system == system)
  s <- sev_df %>% dplyr::filter(solar_system == system,
                              claim_amount > 0)

  if (nrow(f) == 0 || nrow(s) == 0) return(NULL)

351 EN   <- sum(f$claim_count, na.rm = TRUE) # observed E[N]
EX   <- mean(s$claim_amount, na.rm = TRUE) # empirical E[X]
EX2  <- mean(s$claim_amount^2, na.rm = TRUE) # empirical E[X^2]
ES   <- EN * EX                             # E[S]
356 SD_S <- sqrt(EN * EX2)                   # SD[S]

  tibble(
    Line      = label,
    System    = system,
361 'E[N]'     = round(EN),
    'E[X]'     = round(EX),
    'E[S]'     = round(ES),
    'SD[S]'    = round(SD_S),
    'VaR 99%'  = round(ES + qnorm(0.990) * SD_S),
366 'VaR 99.5%' = round(ES + qnorm(0.995) * SD_S),
    'TVaR 99%' = round(ES + dnorm(qnorm(0.99)) /
                      0.01 * SD_S) # normal-distribution TVaR
  )
}
371 agg_tbl <- dplyr::bind_rows(
  agg_stats(eq_freq, eq_sev, "Equipment Failure", "Helionis
    Cluster"),
  agg_stats(eq_freq, eq_sev, "Equipment Failure", "Bayesia
    System"),
  agg_stats(eq_freq, eq_sev, "Equipment Failure", "Oryn Delta"),
376 agg_stats(bi_freq, bi_sev, "Business Interruption", "Helionis
    Cluster"),

```

```

agg_stats(bi_freq, bi_sev, "Business Interruption", "Bayesia
System"),
agg_stats(bi_freq, bi_sev, "Business Interruption", "Oryn Delta"),
agg_stats(wc_freq, wc_sev, "Workers Compensation", "Helionis
Cluster"),
agg_stats(wc_freq, wc_sev, "Workers Compensation", "Bayesia
System"),
381 agg_stats(wc_freq, wc_sev, "Workers Compensation", "Oryn Delta")
)
print(agg_tbl)

# Cargo aggregate (Pareto alpha<0.5 => E[X^2]=infinity; SD undefined)
386 cg_EN <- sum(cg_freq$claim_count, na.rm = TRUE)
cg_EX2 <- mean(cg_sev_clean$claim_amount^2, na.rm = TRUE)
cg_ES <- cg_EN * cg_sev_summary$mean_sev
cg_SD <- sqrt(cg_EN * cg_EX2) # finite-sample approximation
cat(sprintf(
391 "\nCARGO: E[N]=%.0f | E[X]=%.0f | E[S]=%.0f | SD=%.0f |",
cg_EN, cg_sev_summary$mean_sev, cg_ES, cg_SD))
cat(sprintf(" VaR99=%.0f | TVaR99=%.0f\n",
cg_ES + qnorm(0.99) * cg_SD,
cg_ES + dnorm(qnorm(0.99)) / 0.01 * cg_SD))
396

# -- 9. Interest & Inflation -- TVM Adjustment -----
# PURPOSE: Parse macroeconomic data (2160-2174) and compute the
# TVM loading factor (1+pi)/(1+r) applied to all base PPs.
401 # OUTPUT LOCATION: Table 4 (TVM-adj PP column);
# Section 3.3 narrative (pi-bar=2.46%, r-bar=2.90%, loading
-0.42%).
int_inf_clean <- read_excel(
"srcsc-2026-interest-and-inflation.xlsx",
skip = 2,
406 col_names = c("year", "inflation",
"overnight", "rate_1yr", "rate_10yr")) %>%
dplyr::filter(!is.na(year), year != "Year") %>%
dplyr::mutate(across(everything(), as.numeric))

411 cat("Interest & Inflation Data (2160-2174):\n")
print(int_inf_clean %>%
mutate(across(where(is.numeric), ~ round(., 4))))

avg_inf <- mean(int_inf_clean$inflation, na.rm = TRUE)
416 avg_10yr <- mean(int_inf_clean$rate_10yr, na.rm = TRUE)

cat(sprintf("\nAverage inflation: %.2f%%\n", avg_inf * 100))
cat(sprintf("Average 10-yr rate: %.2f%%\n", avg_10yr * 100))
cat(sprintf("TVM loading factor: %.4f (net: %.2f%%)\n",
421 (1 + avg_inf) / (1 + avg_10yr),
((1 + avg_inf) / (1 + avg_10yr) - 1) * 100))

# Discount factors v^t for reference
cat("\nDiscount factors (10-yr rate proxy):\n")
426 for (t in 1:5) {
cat(sprintf(" Year %d: v^d = %.4f\n",
t, t, (1 + avg_10yr)^(-t)))
}

```

```

431 # Inflation-adjusted, discounted pure premiums
# PP_adj = PP_base * (1+pi) / (1+r)
v1 <- (1 + avg_10yr)^(-1)
pp_adj <- tibble(
436   Line      = c("EQ-Helionis", "EQ-Bayesia", "EQ-OrynDelta",
                 "BI-Helionis", "BI-Bayesia", "BI-OrynDelta",
                 "WC-Helionis", "WC-Bayesia", "WC-OrynDelta",
                 "Cargo"),
   PP_base   = c(17411, 13033, 13886,
                 937123, 899736, 842479,
441                 264, 188, 235,
                 3840674)
) %>%
  dplyr::mutate(
446   PP_inflated   = PP_base * (1 + avg_inf),
   PP_discounted = PP_inflated * v1,
   Loading_pct   = (PP_discounted / PP_base - 1) * 100
  )
cat("\nInflation-adjusted, discounted pure premiums (Table 4):\n")
451 print(pp_adj %>% mutate(across(where(is.numeric), ~ round(., 2))))

# -- 10. Scenario Testing -----
# PURPOSE: Apply Best / Moderate / Worst multipliers to each
# line-system E[S] and aggregate across the full portfolio.
456 # OUTPUT LOCATION: Table 7 (scenario parameters and losses);
# Figure 2 (bar chart of portfolio totals);
# Section 4.3 narrative (D136B, D287B, D2124B figures).
scenarios <- tibble(
461   Scenario    = c("Best (Attritional)",
                  "Moderate (Base)",
                  "Worst (Catastrophic)"),
   Prob        = c(0.70, 0.29, 0.01),
   Freq_mult   = c(0.70, 1.00, 2.50),
   Sev_mult    = c(0.80, 1.00, 3.50)
466 )

scenario_tbl <- agg_tbl %>%
  tidyr::crossing(scenarios) %>%
  dplyr::mutate(
471   Scenario_ES = 'E[S]' * Freq_mult * Sev_mult
  ) %>%
  dplyr::select(Line, System, Scenario,
                Prob, 'E[S]', Scenario_ES) %>%
  dplyr::mutate(across(where(is.numeric), ~ round(., 0)))
476

# Add Cargo row
cargo_scenarios <- scenarios %>%
  dplyr::mutate(
481   Line      = "Cargo Loss",
   System    = "All Routes",
   'E[S]'    = round(cg_ES),
   Scenario_ES = round(cg_ES * Freq_mult * Sev_mult)
  )

486 cat("\nPortfolio scenario totals (all lines + Cargo):\n")
dplyr::bind_rows(scenario_tbl, cargo_scenarios) %>%
  dplyr::group_by(Scenario, Prob) %>%

```

```

491 dplyr::summarise(Portfolio_ES = sum(Scenario_ES,
                                   na.rm = TRUE),
      .groups = "drop") %>%
dplyr::mutate(Portfolio_ES = round(Portfolio_ES, 0)) %>%
print()

cat("\n\n=== SCRIPT 1 COMPLETE ===\n")

```

Listing 1: Script 1: Complete analysis pipeline.

A.2 Figure Generation Script

Script 2 is self-contained and produces five PDF figures that are `\includegraphics{}`-ed into the main document. Place the output PDFs in the same directory as the `.tex` file before compiling.

```

# =====
# SCRIPT 2: FIGURE GENERATION
# SOA 2026 Case Study -- five publication-quality PDF figures
# Run after Script 1 (or standalone; data reloaded here).
5 # Output files: fig1_pure_premiums.pdf    -> Figure 1 in report
#               fig2_eq_severity_fit.pdf  -> Figure B1 (Appendix B)
#               fig3_scenarios.pdf        -> Figure 2 in report
#               fig4_cargo_tail.pdf       -> Figure B2 (Appendix B)
#               fig5_var_tvar.pdf         -> Figure B3 (Appendix B)
10 # =====

library(tidyverse)
library(readxl)
library(fitdistrplus)
15 library(actuar)
library(scales)

# Re-apply solar system cleaner (same function as Script 1)
clean_solar <- function(x) {
20 x <- gsub("_\\|?\\|?\\|?[0-9]+$", "", x)
  dplyr::case_when(
    grepl("^Helionis", x) ~ "Helionis Cluster",
    grepl("^Epsilon",  x) ~ "Bayesia System",
    grepl("^Zeta",     x) ~ "Oryn Delta",
25 TRUE                  ~ NA_character_
  )
}

# Reload severity data (clean, positive amounts only)
30 eq_sev <- read_excel(
  "srcsc-2026-claims-equipment-failure.xlsx",
  sheet = "sev") %>%
  mutate(solar_system = clean_solar(solar_system))

35 cg_sev <- read_excel(
  "srcsc-2026-claims-cargo.xlsx",
  sheet = "sev") %>%
  filter(claim_amount > 0)

40 # -- FIGURE 1: Pure Premium Heatmap -----
# OUTPUT LOCATION: Figure 1 (Section 1 Executive Summary;
# also referenced in Section 1.2 ERM Integration).

```

```

# Values: Base pure premiums (PP = lambda * E[X]) from Table 4.
45 # NOTE: These are Base PP, not TVM-adjusted PP. The ~0.42% TVM
# discount is immaterial at this display scale; TVM PP column
# is in Table 4 for reference.
fig1_data <- tibble(
  Line   = rep(c("Equipment\nFailure",
50             "Business\nInterruption",
             "Workers\nComp"), each = 3),
  System = rep(c("Bayesia\nSystem",
             "Helionis\nCluster",
             "Oryn\nDelta"), 3),
55  PP     = c(13033, 17411, 13886,      # EF (Base PP, Table 4)
             899736, 937123, 842479,  # BI (Base PP, Table 4)
             188,   264,   235)      # WC (Base PP, Table 4)
) %>%
  group_by(Line) %>%
60  mutate(PP_scaled = PP / max(PP)) # relative risk within line

pdf("fig1_pure_premiums.pdf", width = 7, height = 4)
ggplot(fig1_data,
  aes(x = System, y = Line, fill = PP_scaled)) +
65  geom_tile(color = "white", linewidth = 1.2) +
  geom_text(
    aes(label = paste0("\u0110",
70             scales::comma(round(PP))),
    size = 3.2, fontface = "bold") +
  scale_fill_gradient(low = "#d6eaf8",
    high = "#1a5276",
    name = "Relative\nRisk") +
  labs(
    x = NULL, y = NULL) +
75  theme_minimal(base_size = 11) +
  theme(panel.grid = element_blank(),
    legend.position = "right")
dev.off()
cat("Figure 1 saved: fig1_pure_premiums.pdf\n")
80

# -- FIGURE 2: EQ Severity -- Fitted Lognormal Curves -----
# OUTPUT LOCATION: Figure B1 (Appendix B, Section B.3).
# Parameters from AIC-selected lognormal fits (Table 3).
85 ln_params <- list(
  "Helionis Cluster" = c(meanlog = 11.0964, sdlog = 0.5934),
  "Bayesia System"   = c(meanlog = 11.1911, sdlog = 0.5826),
  "Oryn Delta"       = c(meanlog = 11.3039, sdlog = 0.6138)
)
90 x_seq <- seq(1000, 600000, length.out = 500)

ln_curves <- purrr::map_dfr(names(ln_params), function(s) {
  tibble(
    x           = x_seq,
95    density    = dlnorm(x_seq,
                       ln_params[[s]]["meanlog"],
                       ln_params[[s]]["sdlog"]),
    System      = s
  )
100 })

```

```

pdf("fig2_eq_severity_fit.pdf", width = 7, height = 4)
ggplot(ln_curves,
      aes(x = x / 1000, y = density * 1000,
          color = System, linetype = System)) +
105 geom_line(linewidth = 1.1) +
      scale_color_manual(
          values = c("#1a5276", "#c0392b", "#117a65")) +
      scale_x_continuous(labels = scales::comma,
110                          limits = c(0, 600)) +
      labs(
          x      = "Claim Amount (\u0110 thousands)",
          y      = "Probability Density",
          color   = "Solar System",
115          linetype = "Solar System") +
      theme_minimal(base_size = 11) +
      theme(legend.position = "bottom")
dev.off()
cat("Figure 2 saved: fig2_eq_severity_fit.pdf\n")
120

# -- FIGURE 3: Scenario Portfolio Losses -----
# OUTPUT LOCATION: Figure 2 in report (Section 4.4 Scenario Testing).
# Values: correct portfolio totals from Table 7 (all lines + Cargo).
125 # FIX: previous version used only the Cargo-component rows from
# dplyr::crossing() output (135.9 / 287.0 / 2123.5 B), which
# excluded the non-Cargo lines. Correct totals confirmed by
# summing both rows per scenario in Script 1 output:
# Best:      135,906,890,527 + 25,256,645,965 = 161.2 B -> 161 B
130 # Moderate: 242,690,875,941 + 45,101,153,509 = 287.8 B -> 288 B
# Worst:    2,123,545,164,482 + 394,635,093,204 = 2518.2 B -> 2518
# B
fig3_data <- tibble(
  Scenario = factor(
135     c("Best\n(Attritional)",
        "Moderate\n(Base)",
        "Worst\n(Catastrophic)"),
    levels = c("Best\n(Attritional)",
               "Moderate\n(Base)",
               "Worst\n(Catastrophic)")),
  Loss_B   = c(161, 288, 2518),          # FIX: correct portfolio
    totals
  Prob_lbl = c("p = 0.70", "p = 0.29", "p = 0.01"),
  Fill     = c("#2ecc71", "#f39c12", "#c0392b")
)
140

pdf("fig3_scenarios.pdf", width = 6, height = 4.5)
ggplot(fig3_data,
      aes(x = Scenario, y = Loss_B, fill = Scenario)) +
      geom_col(width = 0.55, show.legend = FALSE) +
      geom_text(
150     aes(label = paste0("\u0110",
                          scales::comma(Loss_B),
                          "B\n", Prob_lbl)),
        vjust = -0.4, size = 3.5, fontface = "bold") +
      scale_fill_manual(values = fig3_data$Fill) +
155      scale_y_continuous(
          labels = function(x) paste0("\u0110", x, "B"),
          limits = c(0, 2900),          # raised ceiling for 2518B

```

```

        label
        expand = expansion(mult = c(0, 0.12))) +
labs(
160   x = NULL,
      y = "Expected Portfolio Loss (\u0110 Billions)" +
      theme_minimal(base_size = 11) +
      theme(panel.grid.major.x      = element_blank())
dev.off()
165 cat("Figure 3 saved: fig3_scenarios.pdf (= Figure 2 in report)\n")

# -- FIGURE 4: Cargo Severity -- Pareto CCDF (log-log) -----
# OUTPUT LOCATION: Figure B2 (Appendix B, Section B.3).
170 # Fitted Pareto parameters: alpha=0.4826, theta=132849 (Table 3).
pdf("fig4_cargo_tail.pdf", width = 7, height = 4)

cg_plot  <- cg_sev %>% filter(claim_amount < 5e8)
cg_sorted <- sort(cg_plot$claim_amount)
175 n_cg     <- length(cg_sorted)

# Empirical complementary CDF
ccdf_empirical <- tibble(
180   x      = cg_sorted,
     ccdf = (n_cg:1) / n_cg
)

# Theoretical Pareto CCDF:  $P(X>x) = (\theta/(\theta+x))^\alpha$ 
shape_p <- 0.4826; scale_p <- 132849
185 ccdf_pareto <- tibble(
     x      = cg_sorted,
     ccdf = (scale_p / (scale_p + cg_sorted))^shape_p
)

190 ggplot() +
     geom_line(data = ccdf_empirical,
              aes(x = x / 1e6, y = ccdf,
                  color = "Empirical"),
              linewidth = 0.8) +
195     geom_line(data = ccdf_pareto,
              aes(x = x / 1e6, y = ccdf,
                  color = "Pareto fit"),
              linewidth = 1.1, linetype = "dashed") +
     scale_x_log10(labels = scales::comma) +
200     scale_y_log10(labels = scales::percent) +
     scale_color_manual(
       values = c("Empirical" = "#2c3e50",
                  "Pareto fit" = "#c0392b"),
       name   = NULL) +
205     labs(
       x = "Claim Amount (\u0110 millions, log scale)",
       y = "P(X > x) [log scale]" +
       theme_minimal(base_size = 11) +
       theme(legend.position = "bottom")
210 dev.off()
cat("Figure 4 saved: fig4_cargo_tail.pdf (= Figure B2 in Appendix
    B)\n")

```

```

# -- FIGURE 5: VaR / TVaR Normalised Bar Chart -----
215 # OUTPUT LOCATION: Figure B3 in Appendix B (Section B.6).
# Values from Table 5; normalised within each coverage line.
fig5_data <- tibble(
  Line   = rep(c("EF", "BI", "WC"), each = 3) %>%
    factor(levels = c("EF", "BI", "WC")),
220  System = rep(c("Helionis", "Bayesia", "Oryn"), 3),
  ES     = c(167, 250, 266,           # EF (D millions)
            9400, 18000, 17000,      # BI (D billions)
            3.59, 5.24, 6.21),      # WC (D millions)
  VaR99  = c(178, 264, 287,
            10500, 19300, 18100,
225 4.62, 6.45, 7.55),
  TVaR99 = c(179, 266, 289,
            10600, 19500, 18300,
230 4.78, 6.63, 7.74)
)

# Normalise within each coverage line by its maximum TVaR99
fig5_norm <- fig5_data %>%
  group_by(Line) %>%
235  mutate(across(c(ES, VaR99, TVaR99), ~ . / max(TVaR99))) %>%
  ungroup()

fig5_long <- fig5_norm %>%
  pivot_longer(c(ES, VaR99, TVaR99),
240             names_to = "Measure",
             values_to = "Value") %>%
  mutate(Measure = factor(
    Measure,
    levels = c("ES", "VaR99", "TVaR99"),
245    labels = c("E[S]", "VaR 99%", "TVaR 99%")))

pdf("fig5_var_tvar.pdf", width = 8, height = 4.5)
ggplot(fig5_long,
      aes(x = System, y = Value, fill = Measure)) +
250  geom_col(position = "dodge", width = 0.7) +
  facet_wrap(
    ~ Line, scales = "free_y",
    labeller = labeller(Line = c(
255     EF = "Equipment Failure",
     BI = "Business Interruption",
     WC = "Workers Compensation")))) +
  scale_fill_manual(
    values = c("#85c1e9", "#2980b9", "#1a5276"),
    name = NULL) +
260  scale_y_continuous(labels = scales::percent) +
  labs(
    x = "Solar System",
    y = "Relative to Max TVaR 99% (within line)") +
  theme_minimal(base_size = 11) +
265  theme(strip.text = element_text(face = "bold"),
        legend.position = "bottom")
dev.off()
cat("Figure 5 saved: fig5_var_tvar.pdf (= Figure B3 in Appendix
  B)\n")
270 cat("\nAll 5 figures generated. ",

```

```
"Copy PDFs to the same directory as your .tex file.\n")
```

Listing 2: Script 2: Figure generation (5 PDF outputs).

Appendix B: Supporting Detail

B.1 Full Modelling Assumptions

This table documents all modelling assumptions including those of lower sensitivity omitted from the body for space.

Table B1: Full modelling assumptions, rationale, and sensitivity assessment (all nine rows).

| Assumption | Value / Range | Rationale | Sensitivity |
|---|--|---|--|
| Solar system label mapping | Epsilon Bayesia; Zeta Oryn \equiv | Mine counts (Bayesia 15, Oryn 10) aligned with record volumes (Epsilon \approx 38 K, Zeta \approx 38 K, Helionis \approx 19 K); cross-validated against encyclopaedia data. | High: if mapping is reversed, all system-specific frequencies, severities, and derived products change. |
| Historical data represent future exposure | 2160–2174 parameters extrapolated forward | Standard actuarial assumption; no structural break detected in rolling 5-year frequency estimates. | High: space-mining technology evolves rapidly; equipment generations may shift severity distributions. |
| Claim frequency \sim Poisson | Compound Poisson model applied to all lines | Standard for short-tail lines with many exposure units; some overdispersion possible for BI given cluster effects. | Low–Moderate: VaR may be underestimated if true count distribution is overdispersed. |
| Lognormal severity for EF, BI, WC | AIC-selected from 4 candidates | Lognormal AIC was clearly lowest across all nine system-line combinations; Gamma failed to converge. | Low: model selection was unambiguous in all nine cases. |
| Pareto severity for Cargo | $\hat{\alpha} = 0.4826$, $\hat{\theta} = 132,849$ | AIC-selected; near-infinite theoretical mean ($\hat{\alpha} < 1$) makes a per-shipment hard limit essential. | High: small changes to $\hat{\alpha}$ materially affect TVaR; shape must be monitored annually. |
| Normal approximation for VaR/TVaR | Applied to aggregate S | Acceptable for EF and WC; conservative (thin-tailed) for BI and Cargo where true tails are heavier. | Moderate: Monte Carlo simulation would yield heavier aggregate tails and higher capital requirements. |

Continued on next page...

Table B1 (*continued*)

| Assumption | Value / Range | Rationale | Sensitivity |
|--|---|--|---|
| TVM: 10-year rate as discount proxy | $\bar{r} = 2.90\%$ avg; net loading -0.42% | Risk-free proxy for premium investment returns; 1-year rate ($\approx 1.79\%$) also considered but less appropriate for long-tail lines. | Low: the -0.42% net loading is immaterial for short-tail, 1-year policies. |
| Scenario multipliers | Best $0.56\times$, Worst $8.75\times$ base $\mathbb{E}[S]$ | Grounded in a 2-component GMM regime-switching model with a 6.5% combined tail probability; multipliers are nonetheless judgmental. | Moderate: worst-case loss of $\text{€}2,518\text{ B}$ is highly sensitive to the severity multiplier of 3.50 . |
| Correlation structure constant over time | Spearman $\hat{\rho} = 0.35$ | Estimated from 15 years of paired Helionis–Bayesia observations; may shift with space-weather cycles of longer periodicity. | High: a $\pm 30\%$ change in ρ around the central estimate implies a $\text{€}64\text{ B}$ capital swing. |

B.2 Full Product Design Specifications

The four tables below reproduce the complete product design specifications for each coverage line, including all benefit structures, quantitative triggers, exclusions, system-specific features, variable definitions, and scalability provisions summarised in Table 2 of the main report.

Table B2: Equipment Failure: full product design.

| Feature | Specification |
|---------------------------------|--|
| Benefit Structure | Replacement or repair cost; maximum $\text{€}2\text{ M}$ per insured unit; actual-cash-value basis for <code>equipment_age</code> > 10 yr. |
| Coverage Trigger | > 48 hr downtime <i>or</i> repair cost $> \text{€}10,000$ (all systems). Helionis parametric option: <code>maintenance_int</code> $> 1,500$ triggers automatic payment (avoids adjuster cost at <code>freq</code> = 0.221). |
| Exclusions | Wear-and-tear where <code>equipment_age</code> > 20 yr without a documented maintenance record: assets in this category account for a disproportionate share of catastrophic (unrecoverable) failure events in the historical data. Operation at <code>usage_int</code> > 23 without a current inspection certificate: this threshold represents the manufacturer-rated safe operating ceiling; exceedance without documented inspection indicates unremediated risk that falls outside the insurable perimeter. |
| System-Specific Features | <i>Helionis:</i> deductible $\text{€}5,000/\text{claim}$ (highest <code>freq</code> 0.221). <i>Bayesia:</i> radiation-event endorsement for K-type stellar microelectronics damage. <i>Oryn Delta:</i> growth clause—per-unit limit $+15\%$ at each 3-year renewal. |

Continued on next page...

Table B2 (*continued*)

| Feature | Specification |
|-----------------------------|--|
| Variable Definitions | <code>maint_int</code> = cumulative maintenance intensity index (service hours \times component cycles per period), recorded in CQMC's maintenance management system and auditable via onboard telemetry; threshold 1,500 calibrated on historical pre-failure records. <code>usage_int</code> = operational usage intensity (run-hours per rated capacity period); threshold 23 represents the manufacturer-rated safe operating ceiling; verified from onboard telemetry logs. Both variables are write-protected in source systems and trigger an automatic inspection flag if modified, preventing manipulation. |
| Scalability | Modular limits D500 K / D1 M / D2 M per unit; aggregate exposure adjusts with mine-count declarations at each policy anniversary. |

Table B3: Business Interruption: full product design.

| Feature | Specification |
|---------------------------------|--|
| Benefit Structure | Lost production revenue; maximum D50 M per event (mandatory sublimit); indemnity period capped at 180 days. |
| Coverage Trigger | >72 hr continuous operational halt caused by an insured peril (equipment failure, radiation event, or covered cargo loss); halt certified by on-site operations manager. |
| Co-insurance | 20% policyholder retention on all losses above D10 M per event, aligning CQMC incentives with loss prevention. |
| Exclusions | Scheduled maintenance shutdowns. <code>production_load</code> < 0.10 at claim time: a near-idle station where BI coverage would produce a near-zero benefit and creates mis-pricing risk; value verified from CQMC production management records (range 0–1, where < 0.10 indicates < 10% of rolling 90-day average capacity in use). <code>energy_backup_score</code> = 5: a rating of 5 indicates full redundant backup power, meaning the station's exposure to extended shutdown is materially lower; these stations receive a reduced sublimit of D25 M rather than the standard D50 M. |
| System-Specific Features | <i>Helionis</i> : standard form. <i>Bayesia</i> : radiation-event BI endorsement for binary K-type stellar spikes (mean severity D4,845,250/claim). <i>Oryn Delta</i> : higher base sublimit D75 M/event reflecting 15% planned growth (conditional on four underwriting requirements; see body §2.2); annual production-index review. |
| Variable Definitions | <code>prod_load</code> = ratio of current production output to 90-day rolling average capacity (range 0–1); verified from CQMC production management records; value < 0.10 indicates a near-idle station. <code>energy_backup_score</code> = energy resilience rating (1–5); a score of 5 indicates full redundant backup power, meaning the station's exposure to extended shutdown is materially lower; sourced from CQMC site infrastructure registry, updated at each policy renewal. |
| Scalability | Annual exposure review; sublimits indexed to declared annual production value; auto re-rated on production index movements > 10%. |

Table B4: Workers' Compensation: full product design.

| Feature | | Specification |
|---------------------------------|--------------------|---|
| Benefit Structure | Structure | Medical expenses plus wage replacement; maximum €200,000 per claim (P99 €93,000–€113,000 across systems; historical maximum €193,000). |
| Coverage | Trigger | Workplace injury sustained during a covered, scheduled shift on an insured site. |
| Gravity Adjustment | Adjustment | Benefit multiplier = $1+0.5 \times (1-\text{gravity_level})$; observed <code>gravity_level</code> 0.75–1.0 compensates for low-gravity injury severity and extended recovery. |
| Occupation Rating | Rating | Planetary Operations 1.25× base; Administrators 1.00× base, consistent with stratified loss experience. |
| Exclusions | | <code>psych_stress_index</code> ≥ 4: flagged as a pre-existing condition associated in the data with higher claim probability; a documented medical clearance is required before cover attaches, preventing adverse selection into the programme by workers with known elevated risk. <code>accident_history_flag</code> = 1 without a formal remediation plan: a prior documented incident without remediation indicates unremediated systemic risk; cover is declined until the plan is in place and independently verified. Both variables are recorded in CQMC's HR and occupational health systems; verifiability and audit trail prevent manipulation. Injury during non-scheduled hours without recorded supervisor approval is also excluded. |
| System-Specific Features | | <i>Helionis</i> : standard form (freq 0.029). <i>Bayesia</i> : radiation medical endorsement—acute radiation syndrome and long-term monitoring. <i>Oryn Delta</i> : contract worker surcharge on <code>employment_type</code> = Contract (highest freq 0.030), reflecting shorter system familiarity and reduced access to on-site safety training; applied at renewal on declared headcount by employment type. |
| Variable Definitions | Definitions | <code>gravity_level</code> = local gravitational acceleration as a fraction of Earth standard (1.0), sourced from site environmental records. <code>psych_stress_index</code> = standardised occupational stress score (0–5 scale) from CQMC's annual workforce health assessment; threshold 4 corresponds to clinically significant occupational stress requiring medical review. <code>acc_hist</code> = binary flag (0/1) indicating a prior recorded safety incident in the worker's file within the preceding 24 months. |
| Scalability | | Payroll-indexed premium; auto re-rated at renewal on headcount declarations and occupational-mix changes. |

Table B5: Cargo Loss: full product design.

| Feature | | Design Specification |
|--------------------------|------------------|--|
| Benefit Structure | Structure | Physical loss or damage to covered cargo in transit; maximum €50 M per shipment (hard cap); sub-limits apply by cargo class. |
| Coverage | Trigger | Documented loss or damage with assessed value > €50,000 on a covered route; supported by manifest and transit log. |

Continued on next page. . .

Table B5 (*continued*)

| Feature | Design Specification |
|-------------------------------|---|
| Route Risk Surcharge | <code>route_risk</code> tiers 1–5: Tier 1 base rate; Tier 3 +30%; Tier 5 +75%; intermediate tiers scaled linearly. |
| Debris Density Loading | <code>debris_density</code> > 0.30: additional +15% premium loading on all shipments on affected routes. |
| Exclusions | Inherent vice of cargo. Captain <code>pilot_exp</code> < 3 yr without a qualified copilot: junior single-pilot operations account for a disproportionate share of large losses in the claims data; the threshold is operationally verifiable from flight certification records submitted at inception. Vessels aged > 35 yr without a valid survey certificate within 24 months: structural integrity cannot be independently verified, and historical data associate such vessels with elevated catastrophic loss probability. |
| Reinsurance Structure | Mandatory: Galaxy General retains first D20 M per event; reinsurer covers D20 M–D50 M layer; no policy bound without confirmed placement. |
| Variable Definitions | <code>route_risk</code> = route hazard tier (1–5) assigned by Galaxy General’s underwriting team at policy inception based on historical loss rates, debris mapping, and navigation authority data; not adjustable by the policyholder. <code>debris_density</code> = route corridor debris density index (0–1), sourced from interplanetary navigation authority real-time data; updated at each policy renewal. <code>pilot_exp</code> = certified commercial pilot years of experience on file with the relevant navigation authority. |
| Scalability | Premium indexed to declared cargo value at shipment; new routes require formal route-risk assessment before coverage attaches. |

B.3 Risk Profiles by Solar System

The following narratives document the system-specific risk drivers that inform trigger thresholds, deductibles, endorsements, and loadings described in Tables B2–B5 and the body threat register (Table 6).

Helionis Cluster.. Helionis operates around a G2V star broadly analogous to a main-sequence solar-type, hosting CQMC’s most mature network of 30 mines. Operational maturity is a double-edged characteristic: while loss-reporting processes are well-established, the mean equipment age of 10.42 years across the system translates directly into the portfolio’s highest EF claim frequency of 0.221 per unit of exposure, compared with a portfolio mean of approximately 0.169. Business interruption severity is also the highest of the three systems at a mean of D4,845,250 per claim, reflecting the scale of productive capacity at risk when a mature, high-throughput operation halts. The dominant portfolio-level risk originating in Helionis is not, however, a system-specific frequency or severity feature but rather its correlation with Bayesia: the empirically estimated Spearman $\hat{\rho} = 0.35$ between the two systems signals shared systemic drivers that can generate simultaneous adverse loss years, eroding diversification and creating the D64 B capital gap documented in Section 3 [5].

Bayesia System.. Bayesia orbits a binary pair of K4V and K5V stars whose close angular proximity produces episodic radiation spikes that represent the defining hazard for this system. Equipment failure frequency (0.150 per unit) is moderate relative to Helionis but above Oryn Delta; more critically, the radiation environment elevates BI mean severity

to €4,385,760 per claim, resulting in the highest per-system BI expected aggregate loss in the portfolio at €18.0 B. Workers' compensation claim frequency is the lowest of the three systems at 0.026, but the radiation-exposure medical pathway introduces the possibility of high-cost long-latency claims (acute radiation syndrome, oncological sequelae) that may not be adequately captured by historical severity statistics alone; a radiation medical endorsement is therefore prescribed for both WC and BI lines in Bayesia [6].

Oryn Delta.. Oryn Delta orbits an M-dwarf star characterised by sporadic, high-energy flare events of short duration but significant intensity. The system is in an expansion phase, with CQMC planning 15% operational growth, making current exposure estimates a lower bound on future liability. The observed EF claim frequency of 0.137 per unit is the lowest in the portfolio, but this should be treated with caution: nascent operations typically exhibit under-reporting biases as claims management processes mature, and the highest observed EF mean severity of €101,717 across the three systems is consistent with newer and physically larger equipment whose replacement cost is correspondingly elevated. WC frequency (0.030, the highest of the three systems) reflects an expanding workforce with a higher proportion of contract workers, who are less familiar with site protocols and who lack the institutional safety culture accumulated over Helionis's longer operational history.

B.4 Scenario Narratives

The following narratives expand on the three scenarios summarised in Table 7 and Figure 2 of the main body, providing competitive pricing and capital planning context for each outcome.

Best (Attritional) Scenario — $p = 0.70..$ Under benign conditions, reduced stellar activity and continued improvements in equipment maintenance and pilot training suppress frequencies to 70% of base rates and severities to 80% of historical means, yielding a portfolio expected loss of €136 B. In this scenario, Galaxy General's four-line product suite is comfortably profitable at TVM-adjusted pure premiums plus a standard expense and profit loading; the principal challenge is competitive pricing pressure in a soft loss environment, particularly for cargo where low observed losses might lead CQMC to seek market alternatives. Reserves should nonetheless be maintained at base-case levels, as the attritional environment does not eliminate the possibility of individual large cargo or BI events.

Moderate (Base) Scenario — $p = 0.29..$ The base scenario reflects historical mean frequency and severity parameters, producing a portfolio expected loss of €287 B dominated by Cargo ($\approx 85\%$) and Business Interruption ($\approx 15\%$). This is the primary pricing scenario for all four lines and underpins the TVM-adjusted pure premiums in Table 4. Galaxy General must maintain correlation-loaded capital of €265 B (not €201 B) and ensure reinsurance for the €20–50 M cargo layer is in place. At the base scenario, the BI and cargo sublimits will be breached in a material fraction of claim years, confirming their underwriting necessity rather than treating them as edge-case protections.

Worst (Catastrophic) Scenario — $p = 0.01..$ A catastrophic scenario—consistent with a multi-system correlated solar storm or major supply-chain disruption—drives frequencies to 2.5 times and severities to 3.5 times base values, producing an aggregate portfolio expected loss of €2,518 B. This is approximately 7.4 times the base expected loss and represents a potential solvency event for Galaxy General without catastrophe reinsurance and multi-system cover. Premium income at base rates would be wholly inadequate to fund this outcome; the 1% probability of this scenario translates to a once-in-a-century return period, which is not negligible over a multi-decade book of business.

Capital planning must include explicit stress-testing at this level, and treaty reinsurance for correlated multi-system events should be treated as a mandatory pre-condition for writing all four lines simultaneously.

B.5 Severity Distribution Plots

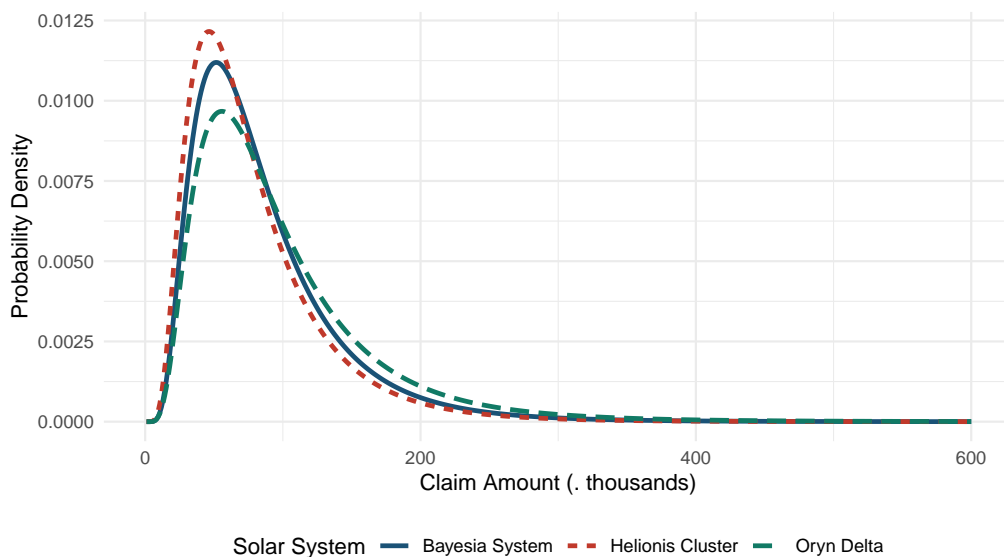


Figure B1: Fitted lognormal severity distributions for Equipment Failure by solar system (AIC-selected). Oryn Delta exhibits the heaviest tail ($\hat{\mu} = 11.30$, $\hat{\sigma} = 0.61$), consistent with its higher observed mean severity of $\text{€}101,717$.

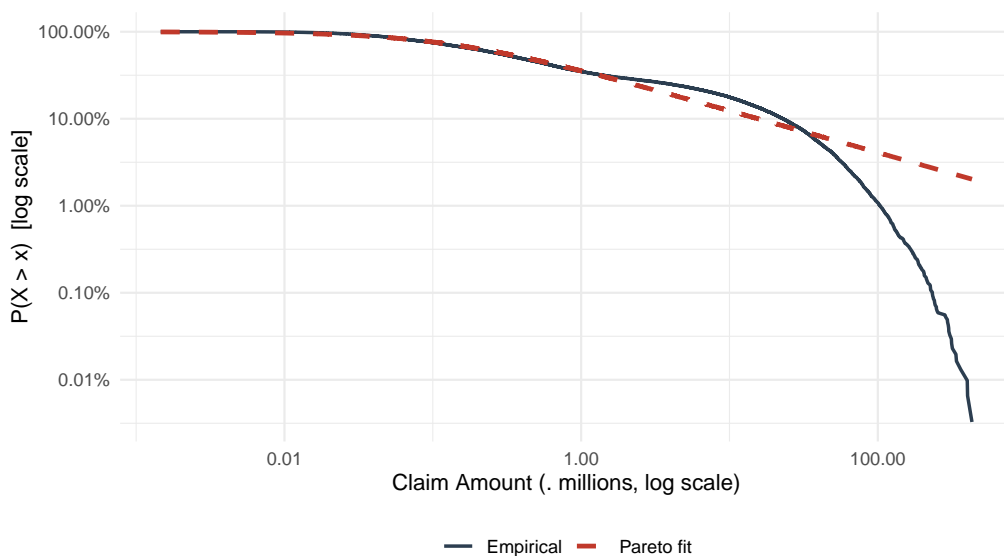


Figure B2: Cargo loss severity: empirical survival function vs. fitted Pareto ($\hat{\alpha} = 0.48$, $\hat{\theta} = 132,849$) on log–log scale. The near-linear log–log relationship confirms Pareto behaviour; the shallow slope ($\hat{\alpha} < 1$) signals that the theoretical mean is undefined under the pure Pareto model, necessitating a hard per-shipment policy limit.

B.6 Aggregate Risk Measures

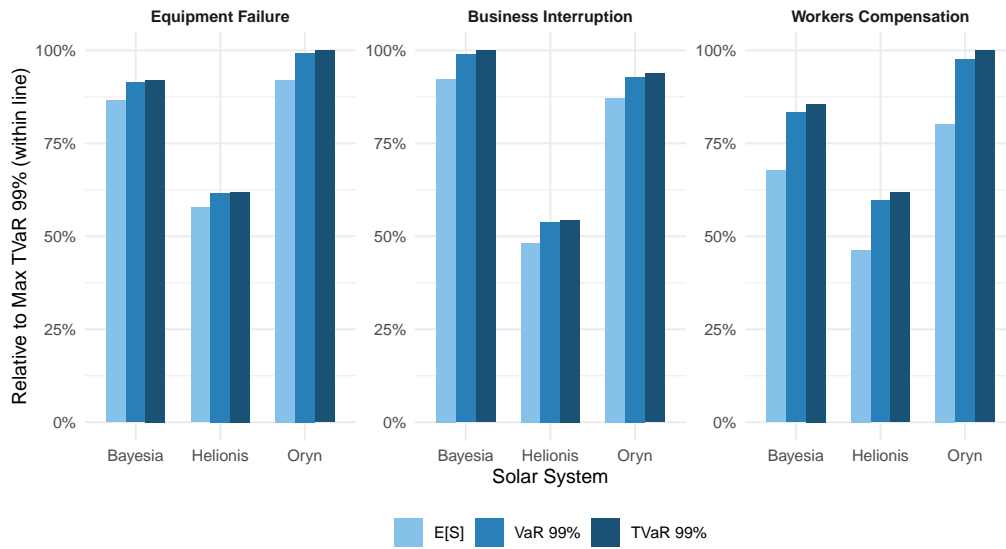


Figure B3: Aggregate risk measures ($\text{VaR}_{99\%}$ and $\text{TVaR}_{99\%}$) normalised within each coverage line. The near-equal bars for EF reflect its thinner lognormal tail ($\hat{\sigma} \approx 0.60$); BI shows greater spread ($\hat{\sigma} \approx 1.20$); Cargo bars are dominated by the heavy Pareto tail ($\hat{\alpha} = 0.48$). All values from Table 5 in the main body.