Model Development Documentation

ICAAP — Credit Risk Capital Model (Year 2025)

Bank: European Fictive Bank (EFB)
Version: 2025.0.1
Date: March 2025
Author: Credit Risk Modelling Team
Reviewers: Risk Methodology (2nd Line)

PART I — Executive Summary

The objective of this document is to describe the Year 2025 version of the ICAAP Credit Risk Capital Model used by EFB to estimate internal economic capital at a 99.9% confidence level. The model is designed to complement regulatory capital requirements by reflecting the Bank's internal view of risk, portfolio specifics, and long-term risk appetite.

Compared to previous years, the YEAR 2025 version introduces several limited changes: a refined segmentation for SME exposures, a small adjustment to downturn LGD inputs, and a revised aggregation file in Excel intended to improve transparency. However, several simplifications remain in place, particularly regarding macroeconomic dynamics, dependency structures, and the implementation approach (which still relies heavily on Excel sheets and manually executed Python scripts).

The model estimates the Bank's credit economic capital for YEAR 2025 at **€1.83bn**, representing a **6.4% increase** relative to YEAR 2024. The increase is primarily driven by a deterioration of the SME portfolio and a small upward revision of LGD assumptions.

While the core methodology remains broadly appropriate for ICAAP purposes, the reliance on manually maintained files, incomplete documentation of data lineage, and the lack of detailed validation evidence on parameter stability remain important limitations to be addressed in subsequent versions.

PART II — Model Overview

The ICAAP Credit Risk Model estimates unexpected losses (UL) at the portfolio level over a one-year horizon and at a 99.9% confidence level. The model uses a Monte-Carlo–based simulation framework combined with segment-level loss distributions derived from internal historical data.

The methodology is conceptually aligned with industry practice: exposures are grouped into segments, each segment is characterised by point-in-time PD and downturn LGD, and a loss distribution is simulated using a one-factor Gaussian copula approximation. Economic capital is calculated as the difference between the simulated 99.9% quantile and the expected loss (EL).

Despite relying on well-established quantitative principles, the model makes several structural simplifications to remain operationally manageable for the modelling team, notably:

- limited reliance on macroeconomic drivers,
- simplified treatment of correlation effects,
- partial manual processing of data,
- and an implementation that mixes Excel spreadsheets and Python scripts without an automated workflow.

The model is intended to serve as the Bank's internal credit risk capital measure rather than to replace regulatory IRB capital requirements. It informs risk appetite reporting, capital planning discussions, reverse stress testing, and the ICAAP final report.

The main outputs of the model are:

- Expected Loss (EL)
- Unexpected Loss (UL)
- Economic Capital (EC = UL – EL)
- Segment-level capital contributions
- Stress scenario impacts

The YEAR 2025 model maintains continuity with the YEAR 2024 version, as required by governance, but includes incremental updates to ensure that the inputs reflect the latest observed risk profile of the portfolio.

PART III — Data Sources

This section describes all data used in the YEAR 2025 ICAAP Credit Risk Model, their origin, structure, and known limitations. The Bank does not maintain a fully automated data warehouse for ICAAP purposes; instead, inputs are extracted manually from internal systems and reconciled by the modelling team.

3.1 Internal Risk Systems

The model uses five main internal data sources:

(1) RISK360_PD

Type: CSV monthly extract
Content: Point-in-time PD estimates based on the Bank's internal rating system
Period covered: 7 years
Size: ~150 MB per extract
Variables: counterparty_id, rating_grade, pd_12m, pd_date, segment, sector_code

Known issues:

- Several missing values for retail exposures in years N-5 and N-6.
- Rating overrides not always logged properly.
- Counterparty IDs sometimes inconsistent with Portfolio_Master.

(2) LGD_CORE

Type: Monthly CSV extract
Content: Downturn LGD estimates used for ICAAP
Size: ~80 MB
Variables: counterparty_id, lgd_downturn, collateral_flag, product_type

Known issues:

- Haircut assumptions changed over time; documentation incomplete.
- Some exposure classes (leasing) use proxy LGDs.

(3) EAD_ENGINE

Type: 40 Excel files generated by the EAD engine
Content: Exposure at Default values for all banking book exposures
Size: Approx. 1.2 GB total
Variables: counterparty_id, product_code, ead, maturity_bucket

Known issues:

- Heavy manual manipulation by operations teams.
- Version control not systematically applied.

(4) PORTFOLIO_MASTER

Type: SQL → CSV export
Content: Master table with segmentation: sector, geography, SME flag, product type
Size: ~25 MB
Known issues:

- SME flag inconsistent for approximately 3.4% of exposures.

(5) DEFAULT_HISTORY

Type: SQL table
Content: Default observations (7-year history)
Volume: ~500,000 observations
Variables: counterparty_id, default_date, recovery_amount, colead_team, write-off_date

Known issues:

- Recovery amounts sometimes missing (due to legacy system migration).
- No direct link to LGD_CORE table for older defaults.

3.2 External or Public Sources

No external data used in YEAR 2025 model, except limited cross-checks with publicly available sector default statistics from the European Banking Authority (EBA) Risk Dashboard (not directly integrated into the model).

3.3 Data Extract Summary Table

Data Source	Format	Size	Frequency	Main Use	Issues
RISK360_PD	CSV	150 MB/mo	Monthly	PDs	Missing PD / override issues
LGD_CORE	CSV	80 MB	Monthly	LGDs	Proxy LGD for some classes
EAD_ENGINE	Excel	1.2 GB	Monthly	EAD	Manual manipulation
PORTFOLIO_MASTER	CSV	25 MB	Monthly	Segmentation	SME inconsistencies
DEFAULT_HISTORY	SQL	500k rows	Daily	Backtesting	Missing recoveries

PART IV — Data Preparation & Quality Checks

Data preparation is performed using a combination of Excel workbooks and Python scripts, without an automated pipeline. This approach remains a key operational limitation but has been maintained “as-is” due to resource constraints and the need for continuity in ICAAP calculations.

4.1 Data Integration Workflow

The integration process is performed using two Python scripts:

- **`01_merge_data.py`** merges PD, LGD, EAD, segmentation, and default history into a single dataset.
- **`02_prepare_dataset.py`** applies transformations, cleans missing values, and produces the final modelling dataset.

The scripts rely on hardcoded file paths in the control workbook `Data_Control_N.xlsx`, which introduces fragility if source files are moved.

4.2 Data Lineage (Narrative)

Below is the simplified lineage description (sufficiently realistic but imperfect, comme tu veux) :

1. **Raw extracts** are stored in:
   `S:\ICAAP\YearN\Data\Raw\`
2. **Intermediate files** (corrected Excel files) are stored in:
   `S:\ICAAP\YearN\Data\Intermediate\`
3. Python scripts export the modelling dataset to:
   `S:\ICAAP\YearN\Data\Processed\dataset_modelling_N.parquet`
4. The final aggregated dataset used for simulations is stored in:
   `S:\ICAAP\YearN\Model\Input\segment_exposures_N.csv`

There is **no automated version control**; changes are tracked manually via a sheet called "Change_Log" in `Data_Control_N.xlsx`.

4.3 Data Quality Checks

Quality checks are performed through a mix of:

- Excel conditional formatting (for missing or abnormal values)
- Python assertions
- Manual spot checks

Summary of QC Checks

QC Check	Description	Result YEAR 2025
Missing PD	Flag PD < 0 or PD > 1 or PD null	1.2% flagged (manual imputation applied)
LGD consistency	LGD within [0, 1]	OK except 0.4% outliers corrected
Duplicate exposures	counterparty_id + product	0.7% duplicates removed
EAD negative values	Should be ≥ 0	14 cases corrected to 0
Sector mapping	Valid sector codes	96.6% mapped successfully
Recovery values	Missing recovery for defaults	17% missing (inherited issue)

4.4 Figure 1 — Distribution of PD Values (Described)

Description:
A unimodal histogram centered around PD = 1.8%, with a long right tail extending to 12%. SME exposures show slightly higher dispersion.

Synthetic Data (to generate your own figure):

PD bucket	Count
0–1%	420,000
1–2%	380,000
2–3%	220,000
3–5%	150,000
5–10%	45,000
>10%	8,000

4.5 Figure 2 — LGD Distribution by Portfolio

Description:
Boxplots showing higher LGD variance in SME and Retail Unsecured segments, with median downturn LGD around 48–55%.

Synthetic Data:

Segment	Median LGD	IQR	Min	Max
SME	48 %	15 %	22 %	82 %
Corporate	43 %	12 %	18 %	71 %
Retail secured	32 %	8 %	10 %	55 %
Retail unsecured	55 %	17 %	28 %	88 %

PART V — Model Assumptions

This section describes the main assumptions underlying the YEAR 2025 ICAAP Credit Risk Model. Some assumptions are inherited from previous versions and remain imperfectly documented, which represents an area for improvement noted in prior audits.

5.1 Time Horizon and Confidence Level

The model estimates unexpected losses over a **one-year horizon**, consistent with EBA ICAAP guidelines.
A **99.9% confidence level** is used to define the economic capital requirement. This level reflects internal risk appetite and aims to be broadly aligned with regulatory IRB calibration philosophy (though the model is not an IRB model).

5.2 Probability of Default (PD) Assumptions

Point-in-time PDs are used directly from the internal rating system for each counterparty.
Key assumptions:

- PDs remain **constant over the one-year horizon** (no migration simulation).
- PD overrides are included but not always transparently documented.
- For exposure classes lacking sufficient history (e.g., leasing), PDs are proxied from similar segments.

This last point has been flagged previously as a methodological gap.

5.3 Loss Given Default (LGD) Assumptions

Downturn LGD estimates from LGD_CORE are used.
Assumptions include:

- LGD is **constant across scenarios** (i.e., no scenario-specific LGD modelling).
- LGD adjustments applied by the recovery teams are included but poorly documented.
- Secured retail exposures use conservative LGDs due to incomplete collateral data.

5.4 Exposure at Default (EAD) Assumptions

EAD values are imported from EAD_ENGINE without additional floors or add-ons.
Limitations include:

- No modelling of credit conversion factors (CCF) for certain off-balance exposures.
- Some product types (overdrafts) rely on outdated EAD assumptions.

5.5 Correlation Structure

A **single-factor Gaussian copula** is used to model systematic risk.
Assumptions:

- Sector-level correlations are not differentiated; a single correlation parameter $\varrho = 0.23$ is applied.
- No geographic dependency is included.
- Correlation estimation is based on internal judgment due to limited historical depth.

5.6 Macroeconomic Effects

The model does not incorporate explicit macroeconomic drivers.
Instead, it assumes:

- PDs implicitly embed macro conditions,
- downturn LGDs already reflect adverse scenarios,
- macro shocks are addressed through separate ICAAP stress testing exercises.

This simplification remains a major limitation for ICAAP modelling.

5.7 Simplifications

- No multi-year dynamics.
- No feedback effects between segments.
- Loss distribution assumed log-normal at segment level for tractability.

These simplifications impact precision but remain acceptable for internal capital planning, given resource constraints.

PART VI — Methodology

This section provides a detailed description of the modelling methodology. It includes segmentation, loss distribution modelling, Monte-Carlo simulation, aggregation, and calculation of economic capital.

6.1 Portfolio Segmentation

The portfolio is segmented into 42 homogeneous groups according to:

- Business line (Corporate, SME, Retail)
- Product type (term loans, revolving credit, leasing)
- Geography (only partially used due to data gaps)
- Internal rating grade buckets

Segmentation aims to balance granularity and computational feasibility.
Some segments contain fewer observations than ideal (notably Retail Auto), which limits calibration robustness.

6.2 Loss Distribution Modelling

For each segment *s*:

- Expected Loss (EL_s) is computed as
  $\mathbf{EL_s = PD_s \times LGD_s \times EAD_s}$

- Unexpected Loss is modelled via a log-normal distribution calibrated on historical segment-level losses.

Let:

- μ_s = mean of log-losses
- σ_s = standard deviation of log-losses

The simulated loss for segment *s* in scenario *i* is:

$$\mathbf{L_{si} = \exp(\mu_s + \sigma_s \times Z_i)}$$

where Z_i is a standard normal random variable correlated through the one-factor structure.

This formulation is chosen for simplicity and computational speed, though it does not fully capture tail dependencies.

6.3 Copula Framework and Correlation

A Gaussian one-factor copula is used:

For each segment *s*:

$$\mathbf{Z_s = \sqrt{\varrho} \times F + \sqrt{(1 - \varrho)} \times \varepsilon_s}$$

where:

- F is the common systematic factor (standard normal),
- ε_s are idiosyncratic shocks (independent normals),
- $\varrho = 0.23$ (fixed parameter).

Segments within the same economic sector would ideally use different correlations, but historical data does not support reliable multivariate estimation. This remains an identified limitation.

6.4 Monte-Carlo Simulation

A total of **20,000 simulations** are performed for YEAR 2025.

Each simulation generates:

1. A value of the systematic factor F
2. For each segment:
    - correlated shock Z_s
    - simulated loss L_{si}

3. Portfolio loss = sum of all segment losses

Implementation

Simulations are run via script:

`02_mc_simulation.py`

- Runtime: approx. 8–10 minutes on standard laptop
- Random seed fixed for reproducibility
- Output stored in:
  `S:\ICAAP\YearN\Model\Output\loss_distribution_N.csv`

No parallelisation is used, which slows the process but remains acceptable for ICAAP frequency.

6.5 Economic Capital Calculation

Let:

- $LL_{99.9}$ be the 99.9th percentile of simulated losses
- EL be total expected loss

Then:

Economic Capital = $LL_{99.9}$ − EL

For YEAR 2025:

- 99.9% loss: €**2.11bn**
- Expected loss: €**0.28bn**
- **Economic Capital = €1.83bn**

6.6 Model Validation Steps (Internal)

Although a dedicated validation framework exists, not all steps were fully documented this year:

- Sensitivity tests were performed on PD and LGD (but not EAD).
- Backtesting was partly incomplete for Retail portfolios.
- No challenger model available for comparison.

These weaknesses explain several recurring recommendations from Audit and Risk Methodology.

PART VII — Implementation Details

The ICAAP Credit Risk Model is implemented through a combination of Excel workbooks and Python scripts. While functional, this approach is not fully industrialised and presents operational risks that have been highlighted in previous reviews.

7.1 Folder Structure

Implementation files are stored under:

```
S:\ICAAP\YearN\Model\
     ├── Input\
     ├── Intermediate\
     ├── Output\
     ├── Scripts\
     └── Documentation\
```

This structure was introduced in YEAR 2024 to improve organisation, but it remains only partially enforced. Several users continue to store temporary files directly under `Model\` or on local machines.

7.2 Excel Workbooks

The model relies heavily on five Excel files:

1. **Aggregation_N.xlsx**
    - Contains segment-level aggregation tables
    - Includes manual macros (undocumented)
2. **Data_Control_N.xlsx**
    - Hosts QC indicators
    - Contains a "Change_Log" tab updated inconsistently
3. **EAD_Imports_N.xlsx**
    - Used to consolidate the 40 EAD_ENGINE Excel files
    - Contains manual formulas (some broken links detected)
4. **PD_LGD_Review_N.xlsx**
    - SME and Corporate reconciliation
    - Reviewer comments partly missing
5. **Exposure_Summary_N.xlsx**

- Final exposure table used by Python simulation
- Generated manually through copy/paste from Aggregation_N.xlsx

Identified Weaknesses (intentional for audit purposes)

- No passwords or access restrictions on any workbook.
- Several formulas refer to local user paths (`C:\Users\Admin\Downloads\…`).
- At least 3 pivot tables have outdated filters.
- The "Review" column in PD_LGD_Review_N.xlsx has missing entries for Retail portfolios.

These are perfect seeds for future audit recommendations.

7.3 Python Scripts

Two main scripts are used:

`01_merge_data.py`

- Merges PD, LGD, EAD, and segmentation
- Performs basic QC checks
- Hardcoded file paths (not ideal)
- No unit testing

`02_mc_simulation.py`

- Performs the 20,000-run Monte Carlo simulation
- Single-threaded
- Random seed documented only in code comments
- Outputs are overwritten without archiving previous runs

There is **no Git repository**, and changes to scripts are tracked only in the Excel “Change_Log” tab, which is often incomplete.

7.4 Figure 3 — Structure of the One-Factor Simulation (Described)

Description:
A diagram representing:

- a central “Systematic factor F”,
- arrows pointing to segment-level shocks Z_s,
- segment losses L_s aggregated into portfolio loss.

Since images cannot be embedded here, you can recreate it easily in Python, PowerPoint, or Mermaid.

PART VIII — Results

This section presents the quantitative outputs of the YEAR 2025 model, including EL, UL, economic capital, and sensitivity results. Results are realistic but include small inconsistencies so that your RAG+audit prototype can detect and comment on them later.

8.1 Expected Loss (EL)

EL is computed as:

$EL = \Sigma\ PD_s \times LGD_s \times EAD_s$

For YEAR 2025:

- Total portfolio EL = **€0.28bn** (rounded)
- By segment (synthetic):

Segment	EL (€m)
Corporate	112
SME	84
Retail secured	26
Retail unsecured	58
Total	**280**

Note: The sum above is €280m; the model documentation reports €0.28bn (consistent). However, some segment-level EL values still carry 1–2% rounding inconsistencies inherited from Excel exports. Leave them — they are realistic imperfections.

8.2 Loss Distribution — YEAR 2025

The 20,000-loss simulation produces a right-skewed distribution with a heavy tail.

Key statistics:

- Mean simulated loss: €**0.29bn**
- 95% percentile: €**0.61bn**
- 99% percentile: €**1.04bn**
- 99.9% percentile: €**2.11bn**

The slight difference between EL (€0.28bn) and mean loss (€0.29bn) is caused by the log-normal calibration — this is normal.

8.3 Figure 4 — Loss Distribution Histogram (Synthetic)

Description:
A histogram with 40 bins showing:

- high mass near €0.1–0.4bn,
- long right tail up to €2.5bn,
- slight multi-modality (due to SME volatility).

Synthetic Table for Graphing:

Loss Range (€bn)	Frequency
0.0–0.2	4,820
0.2–0.4	7,540
0.4–0.6	4,130
0.6–0.8	1,970
0.8–1.0	788
1.0–1.5	550
1.5–2.0	160
2.0–2.5	42

8.4 Economic Capital

Compute:

EC = $LL_{99.9}$ − EL = €2.11bn − €0.28bn = €1.83bn

This is the YEAR 2025 economic capital figure.

8.5 Segment-Level Contributions (Important for future audit insights)

Synthetic, but realistic:

Segment	Contribution to EC
Corporate	42 %
SME	33 %
Retail secured	10 %
Retail unsecured	15 %

You will use these numbers later to audit whether the SME deterioration was appropriately reflected.

8.6 Sensitivity Analysis

PD +20%

- EC increases to **€2.08bn**
  (+13.7%)

LGD +10%

- EC increases to **€1.96bn**
  (+7.1%)

Combined shock (PD +20% & LGD +10%)

- EC increases to **€2.27bn**
  (+24%)

These are intentionally asymmetric — real models often show slightly non-linear behaviour.

8.7 Figure 5 — Sensitivity Tornado Chart (Described)

A large horizontal bar plot showing:

- PD +20% → +13.7%
- LGD +10% → +7.1%
- Combined → +24.0%

You can recreate the bars easily from the numbers above.

PART IX — Limitations

This section summarises the main weaknesses and limitations of the YEAR 2025 ICAAP Credit Risk Model. Some of them were identified in previous audits and remain partially unaddressed.

9.1 Conceptual Limitations

Absence of multi-year dynamics

The model uses a one-year horizon without modelling rating migrations or macroeconomic trajectories. While acceptable for capital planning, this simplification reduces the model's ability to capture long-term risk build-up.

Simplified correlation structure

A single correlation parameter ($\varrho = 0.23$) is applied across all segments.
This ignores sectoral and geographic dependencies, which may underestimate concentration risk.

No explicit macroeconomic drivers

PDs and LGDs implicitly include macroeconomic effects, but no macro model is used.
This makes the model less sensitive to economic downturns and limits consistency with ICAAP stress-testing frameworks.

9.2 Data & Input Limitations

Data lineage incomplete

Although YEAR 2025 introduced a more formal folder structure, lineage remains described manually in Excel.
No automated logging or metadata checks are performed.

Missing and inconsistent data

Several anomalies persist:

- PD missing for ~1.2% of exposures
- LGD outliers corrected manually
- SME flag inconsistent for ~3% of exposures
- Recovery data missing for 17% of historical defaults

These issues are corrected on an ad-hoc basis rather than through a structured data quality framework.

Manual consolidation

40 EAD spreadsheets must be manually merged using Excel.
This is error-prone and has been repeatedly flagged in past audits.

9.3 Implementation Limitations

Excel-heavy process

Critical steps rely on manual copy/paste operations.
No automated checks ensure formulas or pivot tables remain up to date.

Python scripts lack governance

Scripts have:

- hardcoded paths,
- no documentation header,
- no versioning,
- no peer review,
- output overwriting without archiving.

No reproducibility framework

Re-running the model does not guarantee identical results unless the user manually checks all Excel intermediary files.

9.4 Validation & Testing Limitations

Incomplete backtesting

Retail portfolios were not fully backtested this year due to missing recovery data.

No challenger model

Only one model configuration is available.
There is no benchmark (e.g., simpler parametric loss model or Basel formula proxy) despite recommendations from YEAR 2024 audit.

Documentation gaps

Reviewer comments are incomplete in PD_LGD_Review_N.xlsx.
Several model change rationales are not fully documented.

PART X — Governance

The ICAAP Credit Risk Model is subject to the Bank's model risk governance framework. However, in practice, governance remains partly informal and several responsibilities are not consistently enforced.

10.1 Roles and Responsibilities

Model Owner (1st line)

Credit Risk Modelling Team
Responsibilities include:

- Maintaining the model,
- Preparing documentation,
- Performing annual recalibration,
- Running the ICAAP model for YEAR 2025.

In practice, documentation is often updated close to submission deadlines, and only partially reviewed internally.

Independent Review (2nd line)

Risk Methodology Department
Responsibilities:

- Reviewing conceptual soundness,
- Checking data quality and assumptions,
- Producing an annual model opinion.

This year, the 2nd line review was issued late (mid-February), limiting time for corrections.

Audit (3rd line)

Audit performs periodic deep-dive reviews every 2–3 years.
The last deep audit was during YEAR 2023, with a follow-up in YEAR 2024.

Several YEAR 2024 recommendations remain open, particularly on data lineage and script governance.

10.2 Documentation Governance

EFB's Model Governance Policy requires:

- Annual documentation update,
- Version control,
- Complete change logs,
- Validation sign-off.

However, in practice:

- Change_Log tab is inconsistently updated,
- Some sections of the documentation are copy-pasted from YEAR 2024,
- Reviewer comments are missing for several sections.

These gaps will be exploitable in your audit generation later.

10.3 Periodic Review Cycle

Year	Review	Result
N-2	Full audit	Significant findings
N-1	Follow-up audit	Partial remediation
N	2nd line review	Several items still open
N+1	Next audit planned	TBD

This timeline is perfect for your prototype's document chain.

PART XI — Annexes (A–F)

(All annexes are written, not just listed, with deliberate imperfections.)

Annex A — Data Dictionary (Excerpt)

Only a subset of variables has complete definitions. Some entries are intentionally vague to allow audit findings later.

Variable	Description	Type	Comment
counterparty_id	Internal unique identifier	String	Some missing consistency checks
pd_12m	12-month point-in-time PD	Float	Overrides not fully documented

lgd_downturn	Downturn LGD	Float	Several proxy values
ead	Exposure at Default	Float	Imported from EAD_ENGINE
product_code	Internal product mapping	Categorical	Mapping table incomplete
sme_flag	SME indicator	Boolean	3% mismatches detected
sector_code	NACE code	String	Occasionally missing

A full dictionary would normally contain 40 variables, but only 14 are fully documented in YEAR 2025 (intentional).

Annex B — Calibration Tables

Segment-level calibration parameters (synthetic):

Segment	μ (log-mean)	σ (log-std)	Notes
Corporate	-1.85	0.92	Stable
SME	-1.60	1.20	Increased volatility
Retail secured	-2.40	0.70	Stable
Retail unsecured	-1.25	1.35	High tail risk

Documentation does **not** fully explain how μ and σ were derived (intentionally).

Annex C — Figures (Descriptions + Data)

Figure C1 — PD Trend (7 years)

Year	Mean PD
N-6	1.4%
N-5	1.6%
N-4	1.9%
N-3	2.3%
N-2	2.0%
N-1	1.9%
N	2.1%

The documentation does not analyse the spike in N-3 (intentional gap).

Annex D — Model Change Log

Version	Date	Change	Comment
N.0.1	Mar N	Initial release	Several items inherited from N-1
N-1.1	Mar N-1	Added SME segmentation	Rationale incomplete
N-2.0	Feb N-2	Introduced Python simulation	Documentation minimal

Notice the inconsistencies in version numbering — suitable for audit critique.

Annex E — Scripts Overview

01_merge_data.py

- Merges datasets
- Hardcoded paths
- No header documentation
- Missing try/except blocks

02_mc_simulation.py

- Runs simulation
- No logging
- Overwrites output files

These weaknesses are perfect for future audit recommendations.

Annex F — Open Issues & Recommendations

Issue	Status
Lack of automated lineage	Open
Incomplete backtesting	Open
Script versioning	Open
SME flag inconsistencies	Partially resolved
Missing documentation sections	Open

This gives excellent material for the reports d’audit N-1 et N-2.