MILLIMAN WHITE PAPER

Reducing the number of scenarios used for stochastic ALM valuation

Pierre-Edouard Arrouy
Jérémy Beaudet
Mohammed Bennouna
Steven Francois
Alison Tonin

The valuation of an insurance balance sheet is a complex exercise that requires the simulation of various stochastic variables, such as risk-neutral economic scenarios. In practice, given typical run-time constraints, the number of economic scenarios to be considered is limited and it is necessary to develop techniques to ensure that the stochastic valuation of the Best Estimate of Liabilities (BEL) and the Present Value of Future Profits (PVFP) converge to their true values.

In today's landscape, numerous insurance companies are actively working to reduce the computational time associated with stochastic asset-liability management (ALM) valuation, particularly under frameworks such as International Financial Reporting Standard (IFRS) 17 or within Solvency II internal models, specifically for least squares Monte Carlo (LSMC) calibration.

This paper focusses on exploring the reduction of the number of simulations as a means to achieve a reasonable objective in terms of valuation accuracy while significantly decreasing the computational time required for standard valuations. Additionally, efficiency, encompassing cost and energy considerations, has emerged as a pressing concern for businesses of all magnitudes.

The starting point of this study is the Prudent Harmonised Reduced Set of Scenarios (PHRSS) framework recently introduced by the European Insurance and Occupational Pensions Authority (EIOPA) that aims to assist companies that rely on deterministic calculations for their technical provisions by enabling them to evaluate the time value of their options and guarantees using a reduced set of economic scenarios. Nevertheless, scenario reduction based on weighted Monte Carlo techniques have enjoyed enduring popularity for internal model applications thanks to ESG Rebase.[1]

This research proposes a refined simulation reduction technique, centring on the optimisation of critical choices. A key focus lies on enhancing the method for generating and selecting trajectories, along with the reweighting optimisation function. Thanks to cross-validation, we showcase the performance of our proposed methodology using both the Milliman Economic Scenario Generator[2] and an ALM benchmark model. Notably, we demonstrate that even with fewer than 200 simulations, our approach achieves a remarkably accurate replication, equivalent to employing 3,000 simulations through a best-in-class approach.

The paper discusses the following topics:

1. Analysis of the PHRSS framework and the reduced scenario set published by EIOPA.
2. Implementation of scenario reduction techniques and study of trajectory selection approaches along with scenario adjustments.
3. Towards adaptative numbers of scenarios in view of determining an optimised scenario reduction approach.

[1] Howell, C., Leitschkis, M. & Ward, R. (July 2019). ESG Rebase. Milliman White Paper. Retrieved 1 August 2023 from https://www.milliman.com/en/insight/esg-rebase.

[2] See https://www.milliman.com/en/products/economic-scenario-generator.

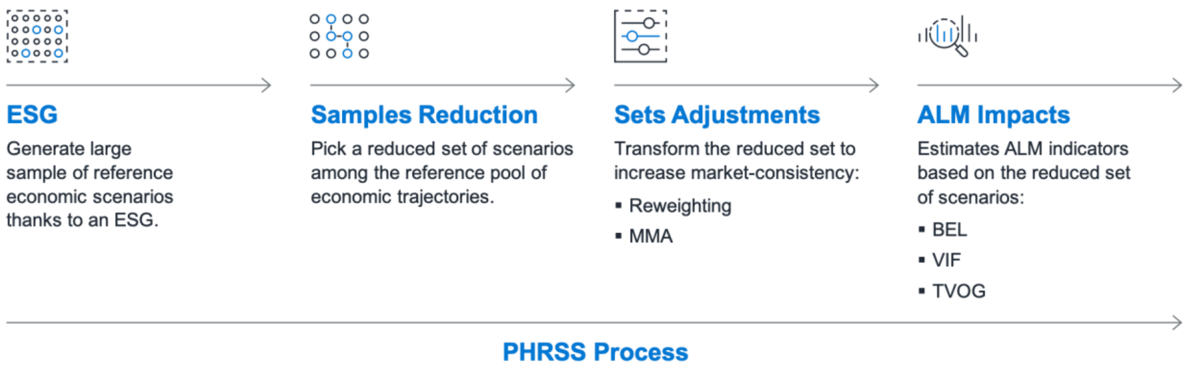
Analysis of the EIOPA PHRSS

Under Solvency II, insurers are required to carry out an exhaustive stochastic assessment of their Best Estimates of Liabilities (BELs) and their Values of In-Force (VIFs) using Monte Carlo techniques. To this extent they use a stochastic cash flow model fed with thousands of risk-neutral economic scenarios generated through an economic scenario generator (ESG). Nevertheless, many European insurance companies currently do not perform stochastic calculations for several reasons. Firstly, they might lack the required computational resources or expertise to perform complex stochastic simulations. Secondly, some companies might have a business model or risk profile that doesn't require stochastic valuation. Moreover, regulatory requirements or guidance might vary across different jurisdictions, leading some companies to opt for simpler, non-stochastic approaches to valuation.

For these reasons, EIOPA has introduced a new valuation method, called the Prudent Harmonised Reduced Set of Scenarios[3] (PHRSS) and aimed for ”low-risk profile undertaking” companies that are not equipped to assess the materiality of their time values of options and guarantees (TVOGs) and which solely perform deterministic calculations of their technical provisions. The objective of this approach is to provide insurers with a set of stochastic economic scenarios containing only a few trajectories, so that the TVOG (as well as the VIF and the BEL) can be operationally estimated by running a deterministic model several times. EIOPA has set the implementation date for this framework in 2024.

In January 2023, EIOPA disclosed six scenario sets[4] for the 31 December 2021 valuation date, each of them featuring nine scenarios. The general methodology employed by EIOPA for creating these six scenario sets involves the steps described in the diagram in Figure 1.

FIGURE 1: EIOPA METHODOLOGY



The first step of the approach consists in generating several thousands of stochastic risk-neutral trajectories incorporating various financial risk factors such as interest rates, equity indices[5] etc. This initial pool of scenarios is referred to as the “reference scenario set.” Secondly, based on this reference scenario set, EIOPA proposes three methodologies for selecting a sample of nine economic scenarios, called the “reduced scenario set." Eventually, two adjustments are made to each trajectory within the reduced scenario set to maintain acceptable properties related to the martingality and market consistency, which have been substantially altered by the reduction step:

- **First step: Weighted Monte Carlo (WMC):** WMC, also known as reweighting, is a procedure that assigns a nonuniform weight to each trajectory to improve the martingale and market consistency (swaption or equity volatilities) of Monte Carlo estimators. In practice, these weights are chosen as the solution of an optimisation problem and do not alter the trajectories but only their weights in the computation of mean estimators.

[3] EIOPA. Prudent Harmonised Reduced Set of Scenarios – First information request. Retrieved 1 August 2023 from https://www.eiopa.europa.eu/prudent-harmonized-reduced-set-scenarios-first-information-request_en.

[4] Ibid.

[5] The PHRSS doesn't include the modelling of stochastic spreads and default risk.

- **Second step: Moment Matching Adjustment (MMA):** The MMA aims at reallocating the sampling bias of martingale estimators by applying a multiplicative coefficient to each risk factor in such a way that they perfectly meet the martingality conditions expected from a so-called risk-neutral set. The MMA can significantly modify the trajectories, in particular in the case of a severe reduction of the number of simulations.

These two methods are combined to improve the robustness and the accuracy of the results, and their impact strongly depends on the order in which they are applied. EIOPA has chosen to perform the MMA after reweighting to ensure that the weighted martingale estimators across the reduced set of nine scenarios are improved.

Finally, the resulting nine economic scenarios can be used as an input of an ALM model to project insurance liability cash flows. The main virtue of this approach lies in the fact that it provides a framework operationally manageable for insurers using a deterministic ALM model to assess the TVOG based on weighted mean estimators.

The following sections present EIOPA's approach in more detail and discuss a study of the six published scenarios sets.

REVIEW OF THE MODELLING CONFIGURATION USED BY EIOPA

Regarding the initial step of generating the reference set, EIOPA prioritises operationality, by proposing an easy-to-follow and implementable methodology. In this section, we provide an overview of the ESG by which the reference set was generated by EIOPA.

In the PHRSS information request, EIOPA outlines its primary objectives as follows:

- Provide simple models that are easily implementable in practice for all market players, with the potential to add complexity.
- Detach as much as possible from market data, to provide a timeless approach to follow for all undertakings.
- On a financial perspective, to not utilise market data, which necessitates expenses for external data acquisition and the creation of substitute hypotheses for specific risk factors and currencies that lack data providers.

For these reasons, EIOPA has chosen a simple modelling framework. The table in Figure 2 summarises the risk factors and the models used to generate the EIOPA reference scenario set.

FIGURE 2: RISK FACTORS AND MODELS USED IN THE PHRSS REFERENCE SET AND ITS CALIBRATIONS

RISK FACTORS		MODEL	CALIBRATION
Interest rate	$r_m(t)$	1-factor Gaussian dynamic $dr_m(t) = \sigma_{IR} dW_t^1$	Standard Formula equivalent shock $\sigma_{IR} = 0{,}39\ \%$
Inflation		Inferred from the projected. determinist scenario	-
Real Interest rate	$r_r(t)$	Inferred from the projected. determinist scenario	-
Real estate index	$S^{RE}(t)$	Black & Scholes dynamic $\frac{dS_t^{EQ}}{S_t^{EQ}} = \sigma_{EQ}\, dW_t^2$	Standard Formula equivalent shock $\sigma_{EQ} = 0{,}19\ \%$
Equity index	$S^{EQ}(t)$	Black & Scholes dynamic $\frac{dS_t^{RE}}{S_t^{RE}} = \sigma_{RE} dW_t^3$	Standard Formula equivalent shock $\sigma_{RE} = 0{,}11\ \%$

Note that W_t^1, W_t^2, W_t^3 denote independent standard Brownian motions. The simplification of not correlating the risk factors is worth noting as it differs from standard market practices that aim at replicating historical correlation targets.

The one-factor interest rates model relies on a single normal volatility; as such it does not reproduce non-flat market swaption volatility structures, which are generally considered for the calibration of benchmark interest rate models in the industry. Likewise, the equity model assumes a constant volatility and does not account for term-dependency of implied volatilities as per most of the models employed by insurers. Moreover, in contrast to risk-neutral best practices, EIOPA calibrates the model parameters $\boldsymbol{\sigma_{EQ}}, \boldsymbol{\sigma_{RE}}$ to match the Standard Formula stresses, prioritising independence from external market data.

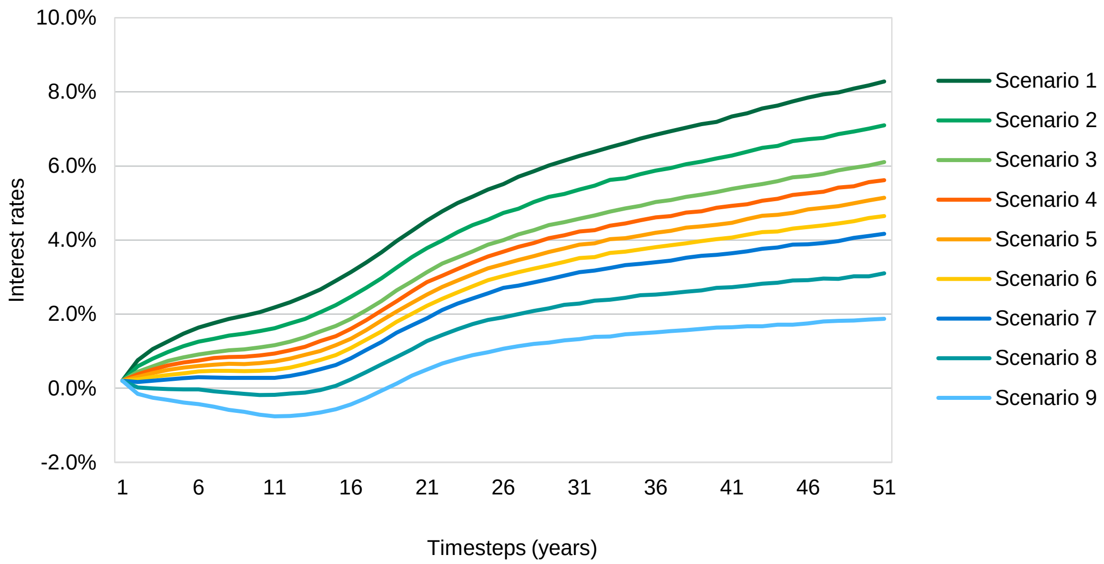
These models are then sampled to produce a large set of (e.g., 5,000) reference scenarios including the above-mentioned risk factors projected over 120 years.

DIFFERENTS APPROACHES OF SCENARIO SELECTION

This section details the key features of the different methodologies proposed by EIOPA for selecting a subset of scenarios (9) from the reference scenario set.

- **Stochastic approach (referred as Method 1 by EIOPA):** This straightforward approach invokes the Central Limit Theorem and randomly selects scenarios from the reference scenario set. This method is also very easy to implement. However, its primary limitation lies in its high sensitivity to the random number generator (RNG), as well as its seed. Indeed, the Central Limit Theorem presumes a strong independence of our trajectories. This assumption could potentially be challenged by a RNG that is not robust enough and incapable of generating sufficiently diversified samples.
- **Quantile approach (referred as Method 2 by EIOPA):** The objective of this approach is to achieve a high level of independence from the underlying RNG by constructing "quantile" scenarios. The concept involves choosing n quantile levels, typically equal to the desired number of scenarios to be extracted, (e.g., $10\%, 20\%, \dots 90\%$) and, for each time step and each risk factor, selecting the value corresponding to the empirical quantile of interest within the reference scenario set. The resulting trajectories exhibit a high degree of smoothness. Consequently, it generally fails in capturing the inherent volatility of the equity indices. From a market consistency perspective, justifying this approach becomes relatively challenging.

FIGURE 3: ILLUSTRATION OF THE METHOD “PERCENTILE LEVEL LINES” ON INTEREST RATE 10-YEAR MATURITY (SET N°3)



- **Nearest neighbour approach (referred as Methods 3 by EIOPA):** This last method combines the two previously mentioned approaches. The idea is to use a nearest-neighbour clustering approach to maintain both diversity in the chosen scenarios and a certain independence from the RNG. Additionally, this method better preserves the original trajectories, enabling the capture of consistent volatility in the indices and thereby ensuring the market consistency of the reduced set of scenarios.

  The selection procedure is defined as follows:

  1. Construct a portfolio index comprising each of the risk factors with a certain weight $\omega_{bond}, \omega_{EQ}, \omega_{RE}$ that is representative of a typical insurer portfolio.
  2. Based on this portfolio index (which would be calculated for each scenario in practice), define a date T that could be the duration of the portfolio in practice (e.g., 10 years).
  3. As in the previous method, define several quantile levels.
  4. Define a new scenario that would be equivalent to the average of the scenarios closest to the value corresponding to the portfolio's quantile.
  5. Lastly, select the "nearest neighbour" scenario that is, on average, closest to the averaged portfolio for each risk factor.

MMA AND WMC SCENARIO ADJUSTMENTS

As previously mentioned, regardless of the methodology considered, it can be necessary to adjust the retained trajectories when extracting a reduced set of scenarios. Indeed:

- The sampling error being significant on a very limited number of trajectories, the martingale properties and the market consistency (e.g., volatility) of the scenarios can be severely altered.
- Despite not being immediately apparent after the transformation, these adjustments have a crucial role in maintaining the stability of ALM impacts at a portfolio level, and effectively limiting model leakage in practice.

In this section, we will delve into more detail about the two procedures used by EIOPA for reprocessing the reduced set.

Firstly, let us introduce the reweighting (WMC) in more detail. This method assigns nonuniform weight to each trajectory to best improve the martingality and the market-consistency characteristics of the scenarios.

The primary objective of the WMC approach is to determine the optimal (nonuniform) weights for solving the following dual problem:

- Minimising the difference between the weight-adjusted trajectories and the expected theoretical values for martingality and market consistency
- Maintaining the weights as close to a uniform distribution as possible and controlling the distortion induced by the previous point. This latter aspect can be quantified using a so-called entropy function and is deeply reviewed in the PHRSS Scenario Analysis section below.

The table in Figure 4 further introduces WMC notations:

FIGURE 4: TABLES OF NOTATIONS AND REFERENCES FOR OPTIMISATION PROBLEM

WMC NOTATIONS	
Number of reference instruments	N
Price of the reference instruments (martingale and market-consistency constraints)	$(C_1, C_2, \dots, C_N)$
Number of reduced economic scenarios	ν
For the fixed reference instrument j, simulated cash flows associated with the various scenarios	$(g_{1j}, g_{2j}, \dots, g_{\nu j})$
Probability weights attributed to each scenario when computing mean estimates	$p = (p_1, p_2, \dots, p_\nu)$
Martingale and market-consistency constraint function to minimise	$\chi_w^2(p) = \frac{1}{2} \sum_{\substack{1<i<N \\ 1<j<\nu}} \left(p_i g_{ij} - C_i\right)^2$
Negative entropy function (Kullback-Leibler) to minimise. Such a function measures the distortion against equi-distributed probability weights of $\frac{1}{\nu}$	$D(p\|u) = \sum_{i=1}^{\nu} p_i \log\left(\frac{p_i}{u_i}\right) = \log(\nu) + \sum_{i=1}^{\nu} p_i \log(p_i)$

The minimisation problem of the WMC approach can be described as follows:

$$\min_p \{w_{mc} \times \chi_w^2(p) + w_e \times D(p|u)\}$$

where $\boldsymbol{w_{mc}}$ governs the importance attributed to the replication of martingale and market-consistency constraints and w_e the importance given to the entropy function (i.e., minimising the distortion of probability weights). The latter parameter w_e will be further discussed in the Selection of an Optimised Reduction Technique Based on an Iterative Algorithm section below.

According to the paper "Weighted Monte Carlo: A new Technique for Calibrating Asset-Pricings Models,"[6] this problem can be reformulated into its dual form, simplifying its implementation on any computational software. Subsequently, for each of the extracted scenarios, a respective weight is obtained requiring transforming the classic estimators such as the mean into weighted estimators, particularly to assess the ALM indicators such as the BEL, VIF or TVOG.

Let us now delve into the Moment Matching Adjustment (MMA) approach. This method alters the trajectories with the goal of meeting the martingale conditions by proportionally reallocating the empirical martingale test's sampling error. Hence, by construction, the martingale tests derived from the post-MMA scenarios precisely match their theoretical expected values.

Formally, let us consider an equity price process denoted $\boldsymbol{S(t)}$ and the associated discount factor $\boldsymbol{D(t)}$. Assuming that we have a sample of ν simulations $\left(\boldsymbol{S^i(t)}, \boldsymbol{D^i(t)}\right)$ the equity martingale test weighted estimator is given by:

$$\widehat{m_t^S} = \sum_{i=1}^{\nu} p_i D^i(t) S^i(t)$$

The MMA approach, consists in modifying the equity price simulations as follows:

$$S_{MMA}^i(t) = S^i(t) \frac{S(0)}{\widehat{m_t^S}}$$

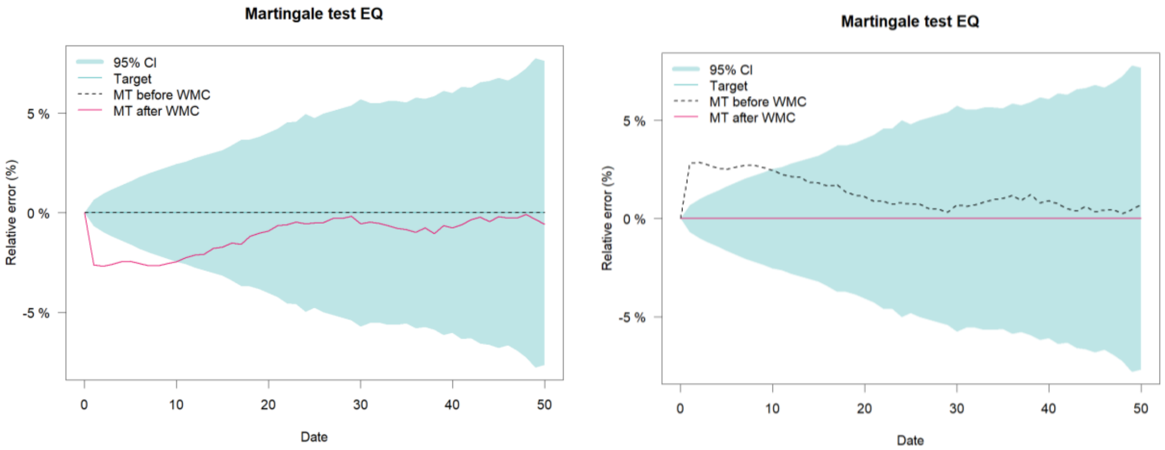
This adjusted sample now perfectly matches the martingale test, because:

$$\widehat{m_t^{S_{MMA}}} = \sum_{i=1}^{\nu} p_i D^i(t) S_{MMA}^i(t) = S(0)$$

The primary advantage of this method lies in its absolute guaranty of martingality, independently of the number of reduced scenarios used. Nevertheless, a significant downside is the drastic modification it can impose on the trajectories, often compromising other properties such as market consistency and correlations. As a result, the use of MMA is generally discouraged for Solvency II applications by regulators and auditors, and the application of such techniques must be carefully justified, as prescribed for instance by the ACPR, the supervisor of French banking and insurance undertakings.[7]

It is important to mention that the order in which the MMA and WMC are applied can lead to variations in the results of martingale tests. To illustrate this, the MMA technique is applied in Figure 5 to a reduced set of scenarios.

FIGURE 5: MARTINGALE TESTS ON TABLE WITH MMA BEFORE WMC ADJUSTMENT (LEFT) AND AFTER WMC ADJUSTMENT (RIGHT)



[6] Avellaneda, M., Buff, R., Friedman, C., Grandechamp, N., Kruk, L., & Newman, J. (2001). Weighted Monte Carlo: A new technique for calibrating asset-pricing models. International Journal of Theoretical and Applied Finance, 4(01), 91-119.

[7] ACPR (Décembre 2020). Générateurs de scénarios économiques: points d'attention et bonnes pratiques – Economic Scenario Generators: Points of Attention and Best Practices.

The application of MMA before WMC (left graph) can substantially distort the martingale tests encompassing the WMC weights. On the contrary, the application of MMA after WMC (right graph) leads, by construction, to perfectly weighted martingale tests.

PHRSS SCENARIO ANALYSIS

As previously mentioned, as part of the PHRSS information request, EIOPA has published six sets of scenarios to perform a stochastic valuation on nine simulations at 31 December 2021. The characteristics of each set are described in the table in Figure 6 (see the MMA and WMC Scenario Adjustments section above for the detail of each underlying method and adjustment).

FIGURE 6: OVERVIEW OF PHRSS SETS

	SELECTION METHOD	ADJUSTMENTS	SPECIFICITY
SET 1	Method 1: Stochastic	Reweighting & MMA	Seed 1
SET 2	Method 1: Stochastic		Seed 2
SET 3	Method 2: Quantile		Percentile list 1
SET 4	Method 2: Quantile		Percentile list 2
SET 5	Method 3: Nearest Neighbour		Seed 1
SET 6	Method 3: Nearest Neighbour		Seed 2

In this section, our goal is to compare the six sets of scenarios provided by EIOPA to a reference set generated with the Milliman Economic Scenario Generator (ESG) and representative of insurance best practices for Solvency II, calibrated to market data at 31 December 2021. This comparison is based on the analysis of martingale and market-consistency tests. Then the PHRSS sets are used in the cash flow model of a fictitious entity that groups together the main characteristics of life insurance companies operating on the French market:

- The model refers to the Standard Formula (no internal model effects, no transitional measures etc.).
- The actuarial valuation methodology used is “standard”; it might not reflect all the specificities of the various insurance companies in the market.
- The assumptions are mostly derived from market data but also based on our own knowledge of the French market.

Considering that the presence of TVOG is predominantly explained by savings contacts, the results presented in this paper focus on this type of undertaking.

Martingale test

The aim of martingale tests is to verify the martingale property of the risk-neutral trajectories of the PHRSS sets. As introduced in the previous section, PHRSS scenarios have the MMA applied. Therefore, all risk factors in the PHRSS sets have perfect martingale tests.

Market consistency

EIOPA aims for stability and being as detached as possible from market data. To this extent, the volatility of the one-factor interest rate model and the volatility parameters of the equity and real estate models are obtained by inverting the associated Standard Formula stresses. For interest rates and equity this approach can lead to discrepancies with the market volatilities generally used by insurers to calibrate their ESG models, for stressed periods during which market-implied volatilities can be significantly higher. On the other hand, the real estate volatility retained by EIOPA is in line with market practices.

Figure 7 presents the repricing of the swaption volatilities performed on the PHRSS scenarios with a Monte Carlo technique and compares them to the at-the-money (ATM) swaptions market volatilities at 31 December 2021 used to calibrate the reference ESG.

FIGURE 7: TABLE OF ABSOLUTE ERROR MONTE CARLO VOLATILITY VS. MARKET VOLATILITY

ABSOLUTE ERROR	METHOD 1		METHOD 2		METHOD 3	
$\sigma^{MC} - \sigma^{Market}$	SET 1	SET 2	SET 3	SET 4	SET 5	SET 6
Maturity / Tenor	$[\mathbf{1};\mathbf{30}]$	$[\mathbf{1};\mathbf{30}]$	$[\mathbf{1};\mathbf{30}]$	$[\mathbf{1};\mathbf{30}]$	$[\mathbf{1};\mathbf{30}]$	$[\mathbf{1};\mathbf{30}]$
$[\mathbf{1};\mathbf{7}]$	-0.19%	0.10%	-0.26%	-0.14%	-0.31%	-0.16%
$[\mathbf{10};\mathbf{30}]$	-0.17%	-0.08%	-0.15%	-0.03%	-0.08%	-0.10%

The table in Figure 7 illustrates the average absolute differences between Monte Carlo volatilities and market volatilities at 31 December 2021. To ease the analysis, the averages have been split into two zones, namely swaptions of short-term maturities (between one and seven years) and swaptions of long-term maturities (greater than 10 years). The results demonstrate that the Monte Carlo volatilities of the PHRSS scenarios underestimate the market volatilities used in the reference ESG by about 20 basis points (bps). We can also see that the difference is quite sensitive to the seed used. Indeed, on Method 1, a change of seed creates a +29bps variation between set 1 and set 2.

We now detail the repricing of the equity and real estate implied volatilities estimated with a Monte Carlo technique on the PHRSS scenarios. For Methods 1 and 3 (focussing on Sets 1 and 5), results are close given the limited number of simulations. We obtain, respectively, an average volatility of 19.42% and 17.62% for equity, as shown in Figure 8. It is interesting to note that the results are also close to the 12 December 2021 ATM market implied volatilities used in the reference ESG. Furthermore, the results remain relatively stable independently of the random generator seed used (Sets 2 and 6).

Method 2 (Sets 3 and 4), however, leads to highly underestimated volatilities of the equity compared to the target volatilities. Indeed, the paths are smooth, as specified by EIOPA in the PHRSS methodology note *“The scenarios are continuous increase or decrease in the value […] There is therefore no ’internal volatility’ in one given scenario.”*

FIGURE 8: ESTIMATED EQUITY VOLATILITY

METRIC	MARKET	SET 1	SET 2	SET 3	SET 4	SET 5	SET 6
$\overline{\sigma_{EQ}}$	19.04%	19.42%	18.22%	3.27%	4.56%	17.62%	18.22%

Note that we obtain similar results for the real estate index.

ALM indicators

PHRSS scenarios are now used in the cash flow model of our fictitious French entity, instead of using the reference ESG previously mentioned. The goal is to verify whether the PHRSS sets produce low leakage[8], which are good estimations of the VIF and allow the TVOG to be approximated, as presented in the table in Figure 9.

FIGURE 9: OUTCOMES OF THE PHRSS SCENARIOS

METRIC	SET 1	SET 2	SET 3	SET 4	SET 5	SET 6
Leakage[8]	-0.33%	0.42%	1.29%	2.34%	0.59%	-1.02%
% TVOG est.[9]	111%	104%	145%	131%	117%	130%
%VIF est.	172%	126%	395%	299%	210%	294%

[8] Leakage = Market value - BEL - VIF, presented in this table as a percentage of the market value.

[9] %TVOG est. is defined as $\frac{TVOG_{Reduced}}{TVOG_{Ref}}$.

We can see that all the PHRSS scenario sets tend to overestimate the VIF, which can be explained by the fact that the PHRSS scenarios underestimate the volatility of swaptions compared with the reference ESG (as a reminder, the economic scenarios are based on market models calibrated on market data). Related to this observation, the EIOPA points out in the PHRSS information request methodology note: *"As the PHRSS is intended to provide a materiality assessment of the TVOG, it might not be necessary to perfectly match the criterions of a fully economic valuation of the balance sheet."*

Method 1 produces a lower leakage than the other approaches and a better estimation of TVOG. In particular, Set 2 provides the closest estimates of ALM indicators; this is due to the volatilities of the swaptions being relatively close to the market volatilities observed on 31 December 2021. However, these volatilities are the furthest from the swaption target volatility used by the EIOPA to calibrate the PHRSS interest rate model. As a consequence, the good results obtained for this set emphasise the importance of using market data to calibrate the reference ESG to replicate more closely the ALM results obtained with the reference ESG.

Sets 3 and 4, corresponding to Method 2, produce the worst results in terms of VIF (395% for Set 3 and 299% for Set 4), caused by an additional underestimation of equity and real estate volatilities. Therefore, Method 2 will be excluded from the next section study.

Method 3 leads to significant distortions of the VIF and the TVOG. Nevertheless, it provides decent martingale and market-consistency results. Hereinafter, we would like to further investigate this trajectory selection method, in the framework of reducing the number of simulations.

In the next sections, we will implement similar reduction techniques on our reference ESG using market volatility targets, discuss the number of scenarios, propose a few alternative approaches and benchmark the results in the light of the above study on the PHRSS scenarios.

Implementation of scenario reduction techniques

Drawing from the EIOPA PHRSS framework analysis and its impacts on ALM metrics, we implement similar reduction techniques now using market volatility targets. A new method, inspired by the Society of Actuaries (SOA), has also been proposed to streamline and condense the number of economic scenarios. Furthermore, the concept of rescaling is introduced as an additional adjustment to the PHRSS-based sets. This proposed adaptation requires the alteration of the condensed scenario set to ideally synchronise with the target volatilities supplied. To end, the sequence of adjustments and their measurable influence on standard metrics like the root mean square relative error (RMSRE) and ALM impacts have been studied.

As a cornerstone of this analysis, the Milliman ESG solution is used to generate a preliminary set of scenarios (3,000 simulations) leveraging a variety of best-in-class models focussed on the dispersion of key risk factors and integrating multiple random number generators (RNGs) such as the Milliman hybrid-RNG and pseudo-RNG.

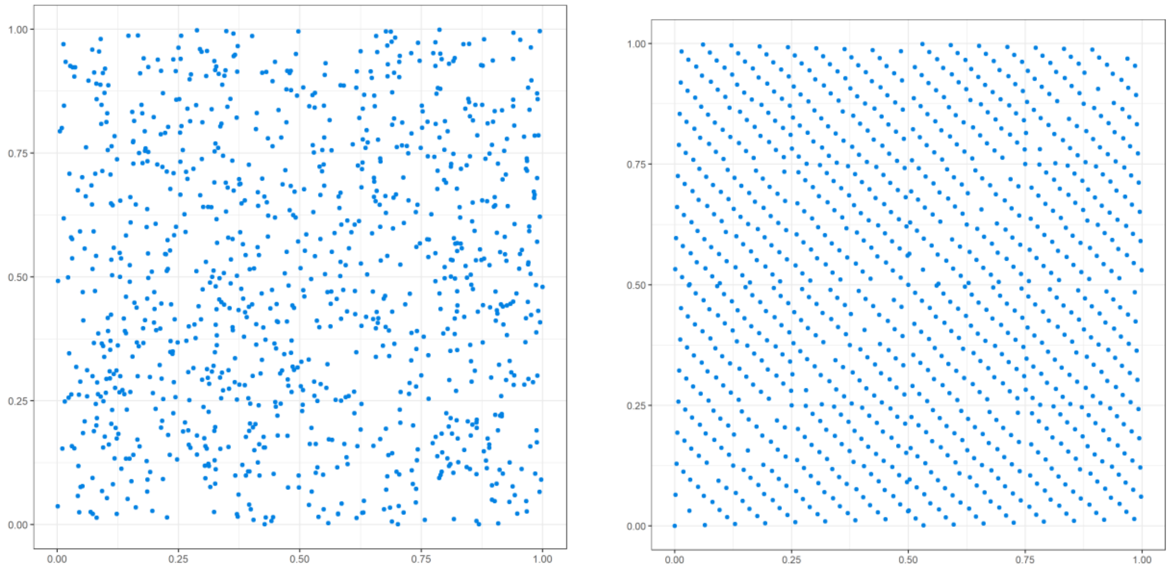
RNG SELECTION

A preliminary adjustment considered consists of modifying the random number generator used in the creation of the reference set. In fact, the role of the RNG is critical in the generation of ESGs. The quality of pseudo-RNG and its seed value can significantly impact validation tests and subsequently insurers' balance sheet valuations and leakage. Quasi-RNG as the Sobol generator is briefly discussed in the following to introduce the Milliman ESG hybrid-RNG.

The Sobol quasi-RNG falls within a broader category termed "low-discrepancy sequences." They are primarily designed to accomplish a swifter rate of convergence compared to the $\sqrt{\nu}$ given by the Central Limit Theorem. Moreover, these sequences display superior evenly distributed space-filling properties compared to standard random number generators. This methodology proves particularly meaningful in our context, as it bolsters the independence of each selected scenario, even when the number of such scenarios is relatively small.

However, traditional Sobol sequences can present challenges, particularly concerning potential lower convergence rates when used in higher-dimensional problems. To mitigate these limitations, the Milliman ESG scenarios employ a hybrid-RNG approach derived from Sobol RNG.

FIGURE 10: REPRESENTATION OF THE 9,999TH (X-AXIS) AND 10,000TH (Y-AXIS) COORDINATES OF THE FIRST 1,000 POINTS OF THE MERSENNE (LEFT), AND SOBOL (RIGHT) SEQUENCES IN DIMENSION 10,000



The hybrid Sobol sequence strategy is developed in two primary steps:

- **Formulation of a low-discrepancy sample:** Initially, a low-discrepancy (deterministic) sample is created using the Sobol sequences. This generates a sequence that fills the space in an evenly distributed manner.
- **Disruption through randomisation:** Following this, the sequence is disrupted with a randomisation process. The aim of this step is to further improve the discrepancy attributes of our sequence, while ensuring a uniform distribution across the entirety of the space.

This dual approach not only maintains the benefits of the Sobol sequences, such as evenly distributed space-filling, but also allows us to improve the convergence rate in higher dimensional problems.

A comparison between hybrid-Sobol and Mersenne-Twister RNGs is illustrated in the following. Let us first introduce the formula of the root mean square relative error (RMSRE) as follows:

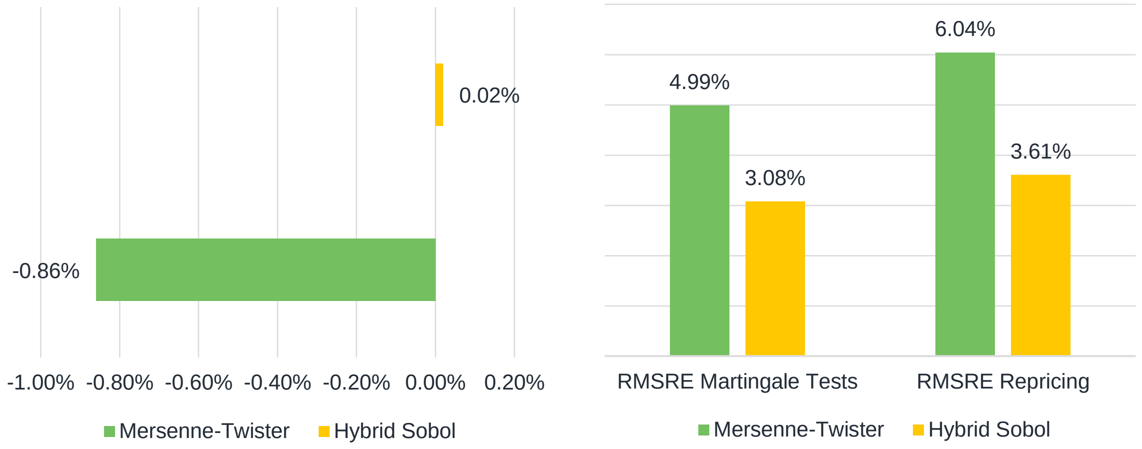
$$RMSRE = \sqrt{\frac{1}{|\mathcal{T}|}\sum_{i\in\mathcal{T}}\left(\frac{\hat{E}_i - E_i}{E_i}\right)^2},$$

where $\mathcal{T}$ is the set of all martingale tests and/or market-consistency tests, $\hat{E}_i$ is the estimated value over the simulations for the test indexed by i and E_i is the corresponding expected target value.

Note that the WMC algorithm only embeds a small subset of the RMSRE constituents to avoid convergence problems. As such the RMSRE also constitutes a relevant metric to control, a posteriori, that the WMC algorithm efficiently improves martingale and market-consistency properties of the scenarios. This metric will be largely used over the next sections to assess the performance of the scenarios from a pure ESG perspective.

The RMSREs and the convergence gaps of the cash flow model are shown in Figure 11 for 200 simulations sampled from both hybrid-Sobol and Mersenne-Twister RNGs thanks to the Milliman ESG capabilities.

FIGURE 11: IMPACT OF RNG ON RMSRE LEVELS (LEFT) AND IMPACT OF RNG ON LEAKAGE (RIGHT)



A distinct improvement in global metrics is discerned with the hybrid-Sobol RNG. Indeed, a considerable reduction in both martingale and market-consistency tests is witnessed. Additionally, the cash flow model convergence leakage is also much lower using the hybrid-Sobol RNG. It is noteworthy that this finding is also corroborated when using the scenario selection approaches introduced in this paper.

The use of the hybrid-Sobol can thus serve as a primary step for efficiently reducing the number of scenarios, which significantly mitigates the sampling error associated with traditional pseudorandom RNGs. Further intricate details about the methodology, results and benefits of the hybrid-Sobol RNG are developed in the Milliman white paper “A New Hybrid Random Number Generator for More Accurate Valuation of Insurance Liabilities.”[10]

RESCALING SCENARIO ADJUSTMENT

The rescaling adjustment method aims to adjust the equity-like trajectories, such that the volatility levels of the equity-like indices are adjusted to new target levels. In practice, the WMC algorithm is often coupled with the rescaling adjustment in view of efficiently improving its validity domain when performing equity volatility stresses for applications like internal models. As a result, in the following, we propose to study the benefits of the rescaling method in the context of reducing the sampling error by limiting volatility distortions rather than applying stresses. We formally summarise its principle below.

Under the assumption that indices follow a lognormal dynamic:

$$S_{ref}(t) = D(t)S(0)exp\left(-\frac{\sigma_{reference}^2(t)}{2} + \sigma_{reference}(t)\, W_t\right)$$

where W_t denotes the standard Brownian Motion and $D(t)$ denotes the discount factor, with the objective being to alter the volatility value to a new target volatility σ_{target}.

The rescaling factors can be calculated accordingly,

$$\lambda(t) = \frac{\sigma_{target}(t)}{\sigma_{reference}(t)}$$

[10] Andres, H., Arrouy, P.-E., Bonnefoy, P. & Boumezoued, A. (December 2022). A New Hybrid Random Number Generator for More Accurate Valuation of Insurance Liabilities. Milliman White Paper.Retrieved 1 August 2023 from https://www.milliman.com/en/insight/hybrid-rng-for-accurate-valuation-of-insurance-liabilities.

[11] American Academy of Actuaries & Society of Actuaries (May 2022). Academy Interest Rate Generator: Frequently Asked Questions (FAQ). Retrieved 2 August 2023 from https://www.soa.org/4a29df/globalassets/assets/files/static-pages/research/2022-academy-interest-rate-generator-faq.pdf.

Subsequently, the value of the rescaled index is derived by the following approach:

$$S^i_{adjust}(t) = S^i_{ref}(0)\frac{\exp\left(\left(LR^i_{ref}(t) - \overline{LR_{ref}}(t)\right) \times \lambda(t) \; + \; \overline{LR_{ref}}(t) \times \lambda^2(t)\right)}{D^i_{ref}(t)}$$

where:

- $\overline{LR_{ref}}(t) = \frac{1}{\nu}\sum_{i=1}^{\nu} LR^i_{ref}(t)$.
- $LR^i_{ref}(t) = \ln\left(\frac{S^i_{ref}(t)}{S^i_{ref}(0)} \times D^i_{ref}(t)\right)$

Additionally, it is crucial to point out that $\overline{LR_{ref}}(t)$ is an unweighted estimator of $\mathbb{E}[LR_{ref}]$. In fact, rescaling is typically conducted prior to the weighted Monte Carlo, which would slightly alter the weights assigned to the trajectories. This decision is driven by the consideration that, with a very limited number of simulations, the volatility of indices could be exceedingly distorted and not counterbalanced by an extreme choice of weights. Nevertheless, the weighted estimator is strongly dependent on the scenario selection methodology employed. A significant asymmetry in the chosen volatilities could substantially skew the estimator.

AN ALTERNATIVE APPROACH OF SCENARIO SELECTION

Attention is now shifted towards a new scenario selection methodology, initially explored by Yvonne Chueh and then employed by the Society of Actuaries (SOA).[11] This method, theoretically, should ensure good homogeneity in scenario choices. The following section explores this approach, covering:

- The main analytical properties
- Practical implementation
- Advantages and disadvantages

The fundamental idea of Yvonne Chueh's paper[12] is to assume that the present value of the insurance portfolio can be written as a continuous function of the discount factor. Specifically:

$$PV(t) = f\big(D(t)\big)$$

with f a continuous function and D the discount factor viewed at time t.

Consequently, if trajectories are close in terms of a distance on the discount factor, a relatively minor variation in the present value of the insurance portfolio should be observed. This property is analytically demonstrated by Chueh.

The goal would, therefore, be to select the trajectories based exclusively on the discount factor, as it would influence all the present values of the other risk factors through the discounting of future cash flows.

Different methods are available to implement the scenario selection in the SOA paper. The choice is made for the Second Sampling Algorithm, as it is the most frequently used in practice by various entities and is summarised as follows:

- A significance of each trajectory is defined as the L^2 norm of the discount factor vector computed over the trajectory.
- Scenarios from the reference set are sorted according to this significance.
- Quantile levels are chosen, in practice as many quantiles as numbers of scenarios to be extracted.
- The trajectories associated with the previous step quantiles of the computed significances are selected.

Beyond its theoretical validity, one key distinction between this methodology and the nearest-neighbour EIOPA approach, is its independence from a reference portfolio. As such, this approach avoids any potential statistical inferences that might be introduced by incorporating weights $w_{Bonds}, w_{EQ}, w_{RE}$ from insurers' varying portfolios.

The focus is now redirected towards the results-related ESG (martingale and market-consistency tests) and ALM impacts of the various selection methods and scenario adjustments introduced in the paper.

---

[12] Chueh, Y.C. (2002). Efficient Stochastic Modelling for Large and Consolidated Insurance Business: Interest Rate Sampling Algorithms. North American Actuarial Journal, 6(3), 88-103.

ALM RESULT ANALYSIS ON 200 SCENARIOS

In this section, the various selection methods and scenario adjustments introduced in the paper are scrutinised to provide comprehensive insights into their effects on RMSREs and ALM metrics.

In the following, we study the following trajectory selection methods:

- **Method A:** Stochastic approach, which simply involves randomly selecting trajectories, namely the ν first trajectories of the reference pool of scenarios. This approach corresponds to Method 1 defined by the EIOPA for the PHRSS impact study.
- **Method B:** Method B is an approach analogous to EIOPA's Method 3, based on a portfolio index calculated as before. Nevertheless, instead of calculating the nearest neighbour of a moving average, we then perform a quantile selection of our scenarios based on the portfolio index value at a fixed horizon date.
- **Method C:** This corresponds to the alternative selection approach based on discount factors and inspired from Chueh's work.

Other variants of Method C have been tested (e.g., taking a distance on the portfolio index instead of the discount factor), but no significant benefits were noted compared to the discount factor approach, which tends to exhibit the highest degree of robustness.

Focus has been given to 200 simulations to mitigate the effects of sampling errors and more clearly decompose the various effects. The results below are split into:

- **ESG indicators:** The martingale tests, the market-consistency tests (swaptions and equity like implied volatilities) and RMSREs.
- **ALM indicators:** The cash flow model leakage as well as the VIF and TVOG relative deviations compared to the reference values obtained with 3,000 simulations.

The tables in Figures 12 and 13 compare the impacts of the three selection approaches of 200 scenarios, considering a reference pool of 3,000 scenarios based on both the hybrid-Sobol RNG and on the Mersenne-Twister RNG.

FIGURE 12: ALM AND RMSRES OUTPUTS BASED ON A MILLIMAN ESG SET IN SOBOL, 200 SIMULATIONS

METRIC	METHOD A	METHOD B	METHOD C
Leakage	0.02%	-0.28%	0.11%
% TVOG est.	102%	99%	100%
%VIF est.	115%	91%	99%
RMSRE – Martingale tests	3.08%	3.54%	5.58%
RMSRE – Market-consistency	3.61%	6.03%	5.74%

FIGURE 13: ALM AND RMSRES OUTPUTS BASED ON A MILLIMAN ESG SET IN MERSENNE, 200 SIMULATIONS

METRIC	METHOD A	METHOD B	METHOD C
Leakage	-0.86%	0.05%	0.46%
% TVOG est.	91%	135%	99%
%VIF est.	136%	-40%	104%
RMSRE – Martingale tests	4.99%	9.46%	4.30%
RMSRE – Market-consistency	6.04%	6.27%	4.88%

The impact of the scenario selection approach is notable in particular on the VIF estimation. Independently from the RNG, Method C provides the best convergence in all cases and seems to be highly resilient to the sampling of the RNG. Moreover, Method B exhibits high discrepancies when coupled to the Mersenne-Twister RNG. These observations can be explained by the high exposure of Method B in the sampling error due to the fact it is based on a single horizon date when classifying the portfolio index. On the contrary, Method C uses the whole trajectory to compute the significance, which provides a higher degree of robustness.

Additionally, and as illustrated previously, the discrepancies in cash flow model leakage as well as VIF and TVOG estimations are significantly higher when considering the Mersenne-Twister RNG, thereby advocating for the hybrid-Sobol RNG even more convincingly. As a result, in the analysis that follows, the hybrid-Sobol RNG will be used for the generation of the reference scenario.

It is noteworthy that the performance gain of Method C over Methods B and A appears to be more pronounced with the Mersenne-Twister RNG than with the Sobol RNG. While Method C still outperforms Method A in Sobol RNG settings, the difference is not as stark as in the Mersenne-Twister scenario. While these results suggest a clear preference for Method C, they do not preclude the potential suitability of other methods under different conditions, such as the number of simulations. In the third part of this paper below, we examine these interactions more closely, with an eye towards gaining a nuanced understanding of the performance across the different selection methods.

The use of scenario adjustments, such as the MMA and the WMC, can be a way to further improve the results. To this extent, the table in Figure 14 examines the impacts of successive application of adjustments (rescaling, reweighting and the MMA) for Method A (the same conclusions apply to other methods).

FIGURE 14: ALM AND RMSRE OUTPUT BASED ON A MILLIMAN ESG SET IN SOBOL (200 SIMULATIONS), METHOD A WITH SUCCESSIVE ADJUSTMENTS

METRIC	METHOD A (NO ADJUSTMENT)	METHOD A (WITH MMA)	METHOD A (WMC)	METHOD A (WMC + MMA)	METHOD A (RESCALING + WMC + MMA)
Leakage	0.02%	-0.17%	0.11%	-0.05%	-0.34%
% TVOG est.	102%	101%	101%	101%	102%
%VIF est.	115%	109%	104%	107%	110%
RMSRE – Martingale tests	3.08%	0.00%	2.41%	0.00%	0.00%
RMSRE – Market consistency	3.61%	3.20%	5.63%	5.67%	5.03%

Acceptable results are achievable without adjustments, yet application of the MMA reduces the RMSRE of martingale tests, which is expected because by construction martingale tests are perfect, and substantially improves the estimation of the VIF. The sole use of weights allows for the best approximation of the VIF while maintaining a reasonable model leakage. Combining the weights and the MMA, as per the EIOPA recommendation, seems to deteriorate the estimation of the VIF compared to the sole application of the WMC. A key point to consider is that, while the MMA effectively ensures the martingale property of the ESG trajectories, its impact can be so substantial on a limited number of scenarios that it may introduce serious bias and deteriorate ALM results. Our various tests tend to indicate that this behaviour is even magnified when the number of scenarios decreases substantially below 200. Given this potential for adverse outcomes, caution should be exercised in the systematic application of the MMA.

The additive application of the rescaling does not bring any benefit in a 200-simulation application. Such an approach is especially relevant when performing stresses on the implied volatilities compared to the reference scenario set (e.g., in an LSMC context), which is not the case in this context because it is used to compensate the sampling error on volatilities.

This section demonstrated that the use of a hybrid random number generator, such as Sobol, markedly improves the precision while reducing the number of simulations. Additionally, we have illustrated that the WMC algorithm can efficiently enhance the results, particularly when sampling error is high (RNG, scenario selection method). Besides, the instability in the impacts for this case study reduction of 200 scenarios illustrates that a reduction towards a very limited number of scenarios (e.g., nine) can be challenging even with adjustments applied. Yet no clear benefits of the MMA and the rescaling adjustments have been identified. In the next section we focus rather on the trajectory selection methodologies coupled with the sole WMC algorithm; we study the effect of entropy in the WMC algorithm and discuss the choice of an optimal number of reduced scenarios, to enhance adaptability and robustness over various situations.

Towards adaptative numbers of scenarios

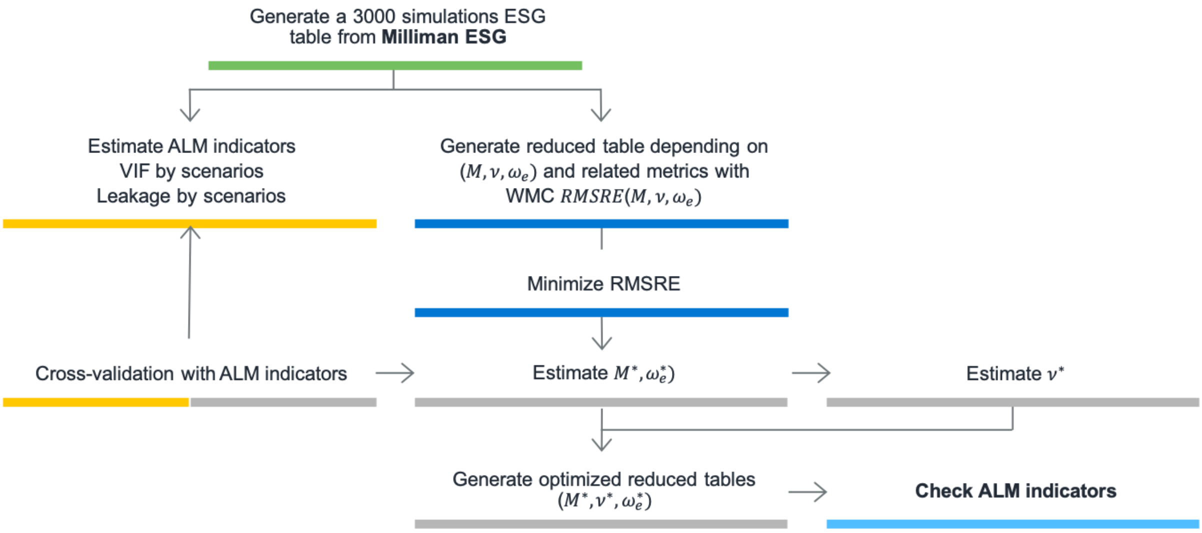
In the light of the previous analysis on 200 scenarios, the objective of this section is to determine optimal triplets (M^*, ν^*, ω_e^*), respectively denoting the trajectory selection methodology, the numbers of scenarios and the entropy weight parameters. In conducting this study, consideration will be given to three distinct metrics:

- **VIF metric:** This measure checks that the VIF obtained for a reduced set $VIF_{Reduced}$ is close to the VIF calculated for the reference set VIF_{Ref}. The metric is defined by $\left|\frac{VIF_{Reduced}}{VIF_{Ref}} - 1\right|$.
- **Leakage metric:** This measure verifies that the convergence deviation is low, and that the BEL is not degraded too much. The metric is defined by $\left|\frac{Market\ Value - (BEL_{Reduced} + VIF_{Reduced})}{BEL_{Ref}}\right|$.
- **RMSRE metric:** This measure checks that repricing and martingale tests remain at the correct level on reduced sets.

CROSS-VALIDATION OF THE SCENARIO REDUCTION TECHNIQUES

These three metrics will be used to determine good triplets for building reduced sets relying on a cross-validation approach presented in the diagram in Figure 15.

FIGURE 15: REDUCED SET CONSTRUCTION PROCESS



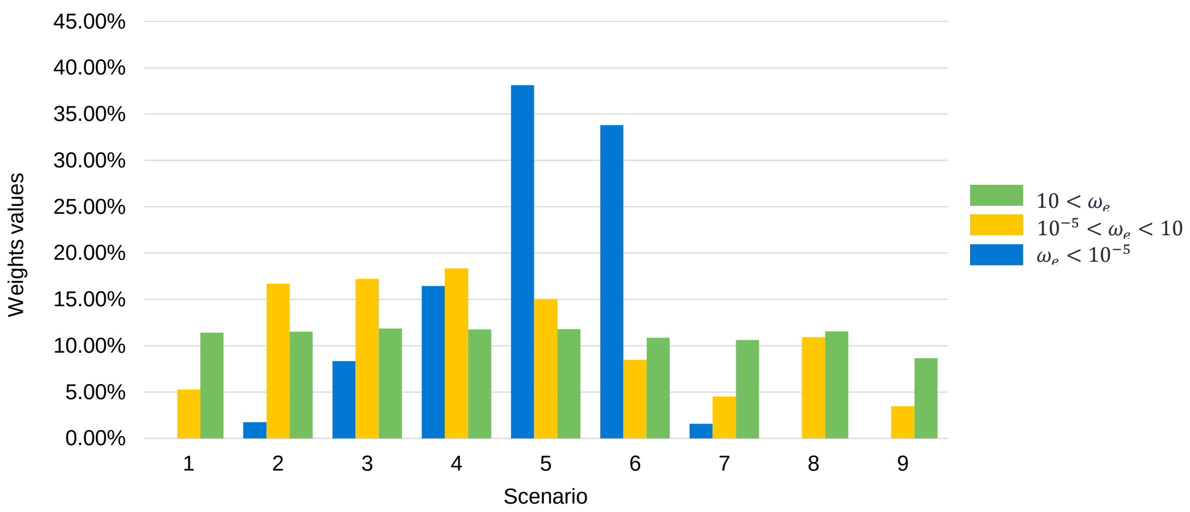
The steps are as follows:

- **Step 1:** Generation of a reference scenario set of 3,000 simulations with Milliman ESG. Throughout the remainder of this section, the reference scenario has been produced utilising the Sobol quasi-RNG.
- **Step 2:** Estimation of reference ALM indicators VIF_{Ref} and BEL_{Ref} associated with the reference set of 3,000 simulations.

- **Step 3:** Creation of reduced scenario sets by activating the WMC algorithm, according to trajectory selection methodology M, number of simulations ν and entropy weight ω_e. The WMC procedure does not distort the ESG trajectories but only their probability weights. As a result, the ALM cash flows per trajectory are not altered either but only the weights of the mean estimators involved in the computation of the ALM indicators. A cross-validation of ALM indicators can then be efficiently performed to analyse the quality of each set as a function of (M, ν, ω_e), without requiring any further ALM runs than the one associated with the reference set of 3,000 scenarios.
- **Step 4:** Determination of the optimal selection method M^* and entropy weight ω_e^* by minimising the RMSRE.
- **Step 5:** Determination of the optimal number of simulations ν^* from the optimal pair (M^*, ω_e^*).
- **Step 6:** Generation of optimised reduced sets using the WMC.
- **Step 7:** Check that the optimised reduced sets satisfy the different requirements in terms of RMSRE, VIF and leakage.

The initiation of this process involves establishing an objective measure for the entropy weight intended for simulation reduction. Initially, an examination of the role of the entropy parameter is reviewed in relation to the objective of achieving an optimally reduced set. The entropy measures the deviation of the probability weights in comparison to the uniform distribution; in practice it is measured by the Kullback-Leibler entropy function $D(p|u)$. This measure is introduced as a constraint in the WMC algorithm and its importance in the overall function to minimise is controlled by a parameter ω_e. A high (resp. low) value of ω_e will induce a limited (resp. significant) distortion of probability weights. It is noteworthy that a too low value of ω_e can result in nil probability weights being assigned to certain scenarios. While this effect may seem less impactful with many scenarios, it is questionable when the number of scenarios is very low because it contributes to reducing it even further. This effect is illustrated in the graph in Figure 15.

FIGURE 16: DISTRIBUTIONS OF WEIGHTS BY ENTROPY PARAMETER VALUES



This suggests that a precise assessment of the entropy weight ω_e is crucial. It should also be noted that the entropy weight and its impact on the WMC depend on ν. As ν increases, we can afford to lower the entropy weight to allow the WMC to assign nonuniform weights to the simulations. On the other hand, if both ν and ω_e are low then the WMC constraints can induce a large bias so that the algorithm can have difficulties to converge, leading to nil weights values.

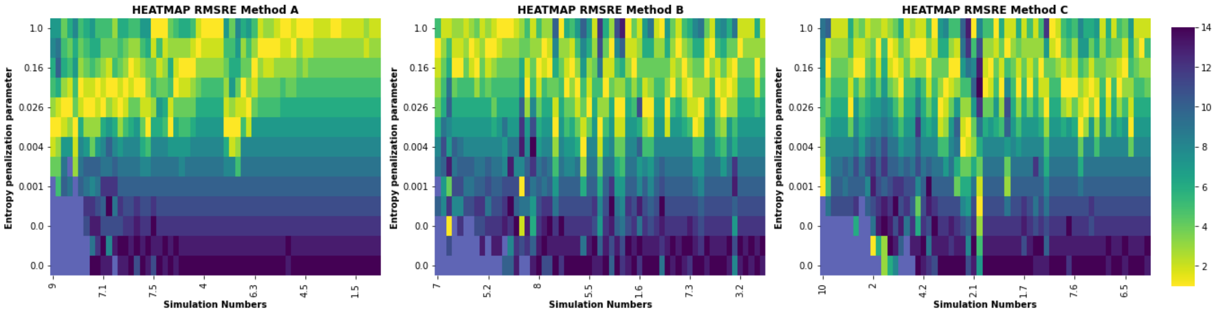
As explained above in Step 3, the knowledge of ALM cash flows by trajectory enables us to set up a cross-validation algorithm to test different methodologies M, numbers of simulations ν and entropy weights ω_e, limiting the need for run time.

There is an intention to put forth a methodology “a priori” for constructing a reduced scenario set, devoid of the ALM indicators. A natural metric that only depends on the ESG characteristics is the RMSRE. The latter serves the purpose of controlling the martingale and market-consistency properties of the reduced set. Therefore, as a preliminary step we look at the link between RMSRE and ALM metrics to determine the relevance of using RMSRE to predict ALM impacts.

CORRELATION ANALYSIS

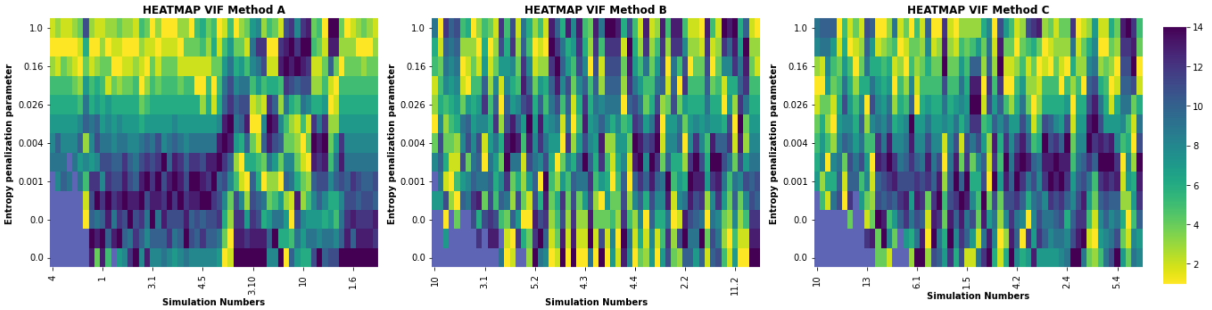
To facilitate this, a heat map of RMSRE relative to ν and ω_e is constructed in Figure 17. Following this, each ω_e value is classified according to ν and the RMSRE value estimated on the reduced set of scenarios. This analysis has been performed for each of the three trajectory selection methods discussed previously.

FIGURE 17: HEAT MAPS – RANKING EACH RMSRE VALUE (%) AS A FUNCTION OF SIMULATION NUMBERS AND ENTROPY (LOWEST VALUE OF RMSRE IS RANKED AS 1)



We also build the heat map of the VIF metric, i.e., we rank each ω_e value as a function of ν and the value of the VIF metric estimated on the reduced set of scenarios.

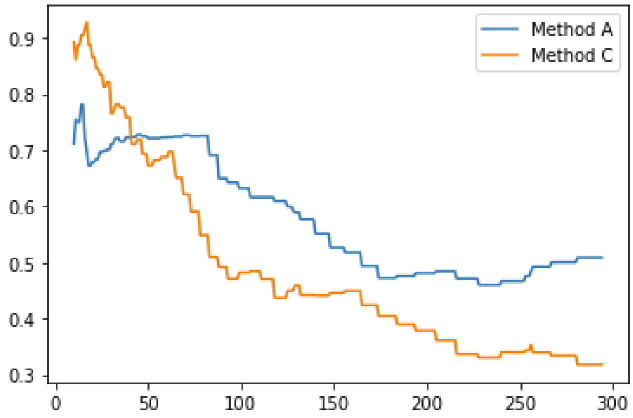
FIGURE 18: HEAT MAPS – RANKING EACH VIF METRIC VALUE (%) AS A FUNCTION OF THE SIMULATION NUMBER AND THE ENTROPY WEIGHT (LOWEST RELATIVE VALUE OF VIF IS RANKED AS 1)



Upon comparison of these heat maps, establishing a definitive connection between RMSRE and VIF proves challenging. However, some common patterns can be discerned for Method A and Method C. To assist in determining the existence of a relation between RMSRE and VIF metrics, the correlation coefficient[13] between the two ranks is computed. We obtain a correlation of 51.6% for Method A, 25% for Method B and 42.7% for Method C. Thus, restricting the analysis to Methods A and C, we seek to further characterise this link, computing the correlation depending on the number of simulations.

[13] We estimate correlation by using Pearson correlations ranked between VIF and RMSREs for each methodology, that is: $\rho = \frac{cov(X,Y)}{\sigma_x \sigma_y}$ where X, Y are respectively our VIF and RMSREs rankings provided on the heat maps.

FIGURE 19: CORRELATION BETWEEN VIF AND RMSRE FOR EACH METHOD IN TERMS OF THE NUMBER OF SIMULATIONS



This preliminary analysis confirms the existence of a link between the RMSRE and the VIF metrics. We observe a pivot value at 40 simulations before which Method A exhibits a stronger correlation than Method C. The decrease in the correlation when the number of simulations increases can be explained by the improvement in the RMSRE and VIF metrics whatever the entropy weight parameter. As such, it becomes more likely that the rank of the best RMSRE does not correspond to that of the best VIF, also explaining the “chaotic” patterns of the heat maps.

SELECTION OF AN OPTMISED REDUCTION TECHNIQUE BASED ON AN ITERATIVE ALGORITHM

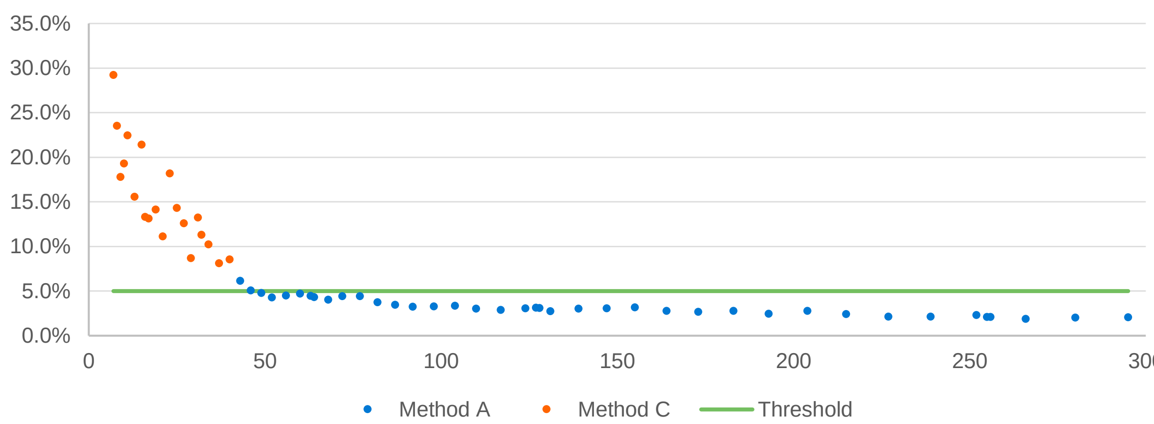
We now rely on this link to build an iterative algorithm to choose an optimised value of ω_e with regard to the ALM metrics we ultimately seek to estimate. For a given methodology M and a given number of simulations ν, the iterative algorithm we propose works as follows:

- **Step 1:** Apply the WMC algorithm considering several values of entropy weight ω_e.
- **Step 2:** For each entropy weight, determine the associated RMSRE.
- **Step 3:** Determine the value of the entropy weight such that the RMSRE is minimal.
- **Step 4:** Verify that the selected ω_e provides relatively accurate results on ALM indicators.

Based on the above study of correlations, and under the conditions of our analysis (in particular use of hybrid Sobol RNG), we can further refine this algorithm by considering Method A when the number of simulations is less than 40 and Method C otherwise as a choice for the optimal selection methodology M^*. An optimal number of simulations can be chosen to reach a precision threshold on the RMSRE and ALM metrics.

The graph in Figure 20 shows the RMSRE results obtained with the optimal entropy weight obtained with the iterative algorithm presented above as a function of the number of simulations.

FIGURE 20: RMSRE IN TERMS OF THE NUMBER OF SIMULATIONS FOR OPTIMAL ENTROPY PARAMETER



We observe a clear reduction of the RMSRE, increasing the number of simulations. An acceptance threshold has been arbitrarily set to 5% as a common tolerance level used in the industry for martingale and market-consistency tests. Applying this threshold, only simulation numbers above 50 verify this condition. Therefore, at this stage, it would seem difficult to build a good proxy for several simulations below 50. Indeed, a certain instability can be observed when a low number of simulations is used. High RMSRE levels probably indicate a problem of market consistency or risk-neutrality.

We then seek to address the impact of the iterative algorithm on the ALM indicators in the two graphs in Figures 21 and 22. In this regard, we have determined thresholds to ensure a sufficient quality of the VIF and leakage metrics. To determine these metrics we considered the example approach provided by the ACPR in its note on ESG good practices:

"It is therefore advisable to define thresholds on the basis of the maximum error they imply, either in terms of BEL valuation, or in terms of the market value (MV) of assets. As an example, an overall threshold of 0.5% for the relative error on martingality tests is equivalent to accepting an error of up to 0.5% of the MV of assets. For a life insurer whose own funds represents 10% of the market value, this represents an uncertainty of 5% on own funds ".

The ACPR also indicates that a 95% uncertainty of less than 0.2% of the BEL corresponds to generally observed market practice. Considering such a threshold of 0.2% on market value and applying the above-mentioned methodology to our fictitious entity, we obtain that the relative error on VIF must be less than 21.55%. For the purposes of prudence, we also define an arbitrary and more prudent threshold of 10%. Besides regarding the leakage metric, we are also considering the two thresholds of 0.5% and 0.2%.

FIGURE 21: VIF METRIC IN TERMS OF THE NUMBER OF SIMULATION FOR OPTIMAL ENTROPY PARAMETER

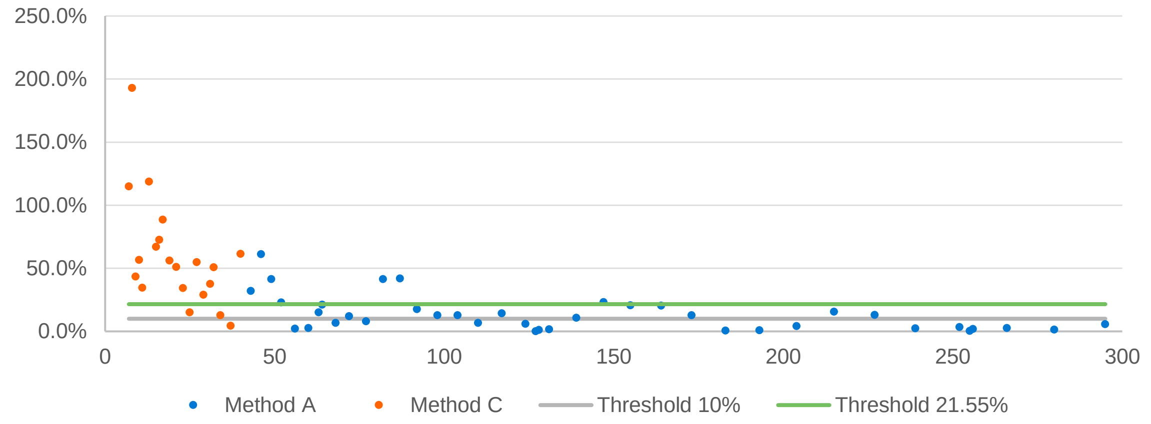
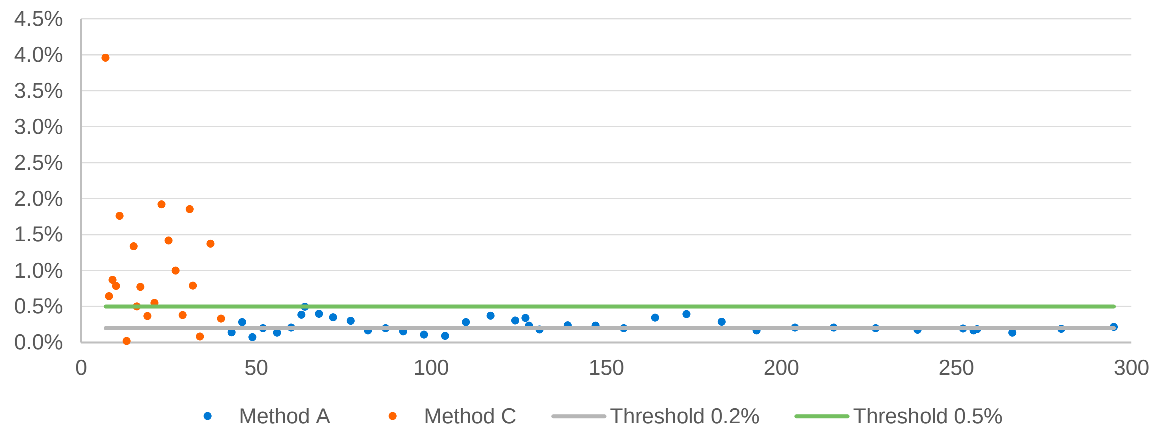


FIGURE 22: LEAKAGE METRIC IN TERMS OF NUMBER OF SIMULATIONS FOR OPTIMAL ENTROPY PARAMETER



The results seem relatively consistent with the RMSRE, in particular a significant improvement is observed as soon as the number of simulations is greater than 50, for both the VIF metric and the leakage. We note that, when the number of simulations is low, the VIF and the leakage metrics are highly unstable, although well below the thresholds for some values of ν. This observation seems to be aligned with the mixed results obtained in the first section above using the six PHRSS sets of nine scenarios provided by EIOPA; for some low values of ν, the iterative algorithm proposed above performs better than the best PHRSS set (Set 2) while for other values it is outperformed.

It is important to stress that the primary goal of EIOPA is not to replicate ALM indicators, but to detect the existence of TVOG. The PHRSS framework served as an inspiration and starting point for our research on scenario reduction. Thus, the main objective of our study is not to surpass the results obtained in the first section above, but rather to propose a robust methodology for constructing reduced sets that can yield satisfactory results when assessing ALM indicators. By reducing the set, we aim to strike a balance between computational efficiency and accuracy, enabling us to obtain accurate ALM estimates while minimising the computational burden associated with running the cash flow model on large scenario sets.

It was demonstrated that results start stabilising for a number of simulations ν greater than 50. Because we determined, that, for ν > 40, Method A (selection of the ν first trajectories in the reference scenario set) was optimal, the selection is entirely driven by the quasi-RNG (Sobol). Such a generator is ideally equi-distributed when the number of simulations[14] has the form of 2^n - 1. Thus, the simulation numbers 7, 15, 31, 63, 127 and 255 are good candidates for the optimal number of scenarios. By cross-referencing all the information we have so far, we propose to further focus on 63, 127 and 255.

The table in Figure 23 summarises the metrics obtained for these three points.

FIGURE 23: SET OF OPTIMAL NUMBER OF SIMULATIONS ON 2021 ALM MODEL

M^*	ν	ω_e^*	RMSRE	VIF METRIC	LEAKAGE METRIC	TVOG15
Method A	63	0.4	4.43%	15.1%	0.38%	2.33%
Method A	127	0.16	3.10%	0.25%	0.34%	0.04%
Method A	255	2.5	2.06%	0.33%	0.16%	0.05%

We note that all these simulation numbers pass the thresholds we have defined and, in the end, $\nu^* = 127$ seems to be the best compromise between not too many simulations and good results on validation and ALM indicators.

In order to back-test the proposed iterative algorithm, we are now looking forward to 31 December 2022. To this extent, we first generate a reference set of scenarios calibrated on the 31 December 2022 market conditions. We then apply the approach presented above to produce reduced sets of 63, 127 and 255 simulations. Finally, we run an updated version of the cash flow model (assets, liabilities, reinvestment rules, …) as of 31 December 2022. The table in Figure 24 details the associated results.

FIGURE 23: BACK-TESTING ON 2022 ALM MODEL

M^*	ν	ω_e^*	RMSRE	VIF METRIC	LEAKAGE METRIC	TVOG
Method A	63	2.5	4.99%	3.13%	0.21%	5.63%
Method A	127	0.026	3.54%	3.14%	0.55%	5.65%
Method A	255	2.5	1.96%	0.35%	0.08%	0.64%

We note that the approach successfully passes the back-test because all the tested simulation numbers provide precise ALM estimates. The proposed iterative algorithm allows us to build reduced sets, so as to obtain accurate and stable enough results while efficiently reducing the number of simulations and as such.

[14] Jäckel, P. (2002). Monte Carlo Methods in Finance (Vol. 5). John Wiley & Sons.

[15] TVOG metric is defined as follow: $\frac{TVOG_{Reduced}}{TVOG_{Ref}} - 1$.

Conclusion

In conclusion, this paper introduced the PHRSS framework that aims to assist companies that rely on deterministic calculations for their technical provisions. The framework enables these companies to evaluate the time value of their options and guarantees using a reduced set of economic scenarios. The analysis of the six reduced sets, each consisting of nine scenarios, published by EIOPA as part of an information request, revealed mixed results, with some sets performing particularly well and others being less relevant.

Subsequently, we revisited the PHRSS framework by implementing similar scenario reduction techniques and proposing an alternative scenario selection method inspired by the SOA. Our findings demonstrated that the use of a hybrid random number generator (RNG), such as Sobol, significantly improved precision while reducing the number of simulations compared to pseudorandom RNGs. Moreover, we showed that the WMC algorithm effectively enhanced the precision of ALM indicators, especially in situations with high sampling error. However, adjustments such as the MMA and rescaling did not exhibit clear benefits.

We then focussed on identifying an optimal trajectory selection methodology coupled with only the WMC algorithm. To achieve this, we introduced an iterative algorithm that allowed us to choose an optimised entropy weight value based on the RMSRE because of its high correlation with the ALM indicators when the number of simulations is low. Through cross-validation, we determined a reasonable number of simulations to obtain accurate and stable results while efficiently reducing the number of simulations.

Overall, this research presented valuable insights into the scenario reduction techniques, and trajectory selection methodologies. By using reduced sets of scenarios and implementing innovative algorithms, companies can achieve precise and reliable results while minimising computational efforts.

Milliman

Milliman is among the world's largest providers of actuarial, risk management, and technology solutions. Our consulting and advanced analytics capabilities encompass healthcare, property & casualty insurance, life insurance and financial services, and employee benefits. Founded in 1947, Milliman is an independent firm with offices in major cities around the globe.

milliman.com

CONTACT

Pierre-Edouard Arrouy
pierre-edouard.arrouy@milliman.com

Jérémy Beaudet
jeremy.beaudet@milliman.com

Mohammed Bennouna
Mohammed.Bennouna@milliman.com

Alison Tonin
Alison.Tonin@milliman.com

For more information, please visit
milliman.com/en/products/economic-scenario-generator