Systematic Literature Review on the Use of Economic Scenario Generators for Insurance Reserve Valuation and Liability Risk Measurement: A Prisma 2020 Review

Michael Ezra Otoo[1*], Joseph Manasseh Opong[2] & Enoch Kwablah Teye[2] & Emily Asaa Addison[3]

[*1, 2-3]Presbyterian University, Ghana, P.O. Box 59. Abetifi-Kwahu

Corresponding Author: Michael Ezra Otoo

Presbyterian University, Ghana, P.O. Box 59. Abetifi-Kwahu

Article History

Received: 03 / 05 / 2026

Accepted: 12 / 06 / 2026

Published: 25 / 06 / 2026

Abstract: Economic Scenario Generators (ESGs) have become indispensable tools in modern insurance reserve valuation, solvency assessment, and asset-liability management due to their ability to model uncertainty in economic and financial variables. Despite their widespread application, there remains limited synthesized evidence regarding their design, calibration, regulatory application, and effectiveness in insurance liability valuation. This study presents a systematic literature review of ESG applications in insurance reserve valuation and liability risk measurement using the PRISMA 2020 framework. A comprehensive search of academic databases, regulatory publications, and industry reports covering the period 2000–2025 yielded 842 database records and 49 additional sources. Following screening, eligibility assessment, and quality evaluation, 85 studies were included in the final synthesis. The review identifies six major thematic areas: ESG architecture and calibration, regulatory and accounting requirements, reserve estimation for complex guarantees, computational techniques, model risk and validation, and emerging climate-related risks. Findings indicate a significant transition from traditional deterministic and single-factor models to sophisticated multi-factor, market-consistent stochastic frameworks driven largely by Solvency II and IFRS 17 requirements. Advances in proxy modelling, least-squares Monte Carlo methods, and machine learning have enhanced computational efficiency; however, challenges remain regarding the integration of real-world and risk-neutral measures, model validation, and climate risk incorporation. The study highlights critical research gaps and provides recommendations for future ESG development, regulatory harmonization, and robust insurance liability valuation practices.

Keywords: *Economic Scenario Generators (ESGs); Insurance Reserve Valuation; Liability Risk Measurement.*

How to Cite in APA format: Otoo, M. E., Opong, J. M., Teye, E. K. & Addison, E. A. (2026). Systematic Literature Review on the Use of Economic Scenario Generators for Insurance Reserve Valuation and Liability Risk Measurement: A Prisma 2020 Review. *IRASS Journal of Economics and Business Management. 3(6), 67-73.*

Introduction

Insurance liabilities (including particularly life insurance, annuity contracts and long-term liability non-life lines), usually extend over many years and therefore have a high degree of sensitivity to economic parameters: interest rates, inflation, equity price levels, property yield levels, credit spread levels and exchange rates (Devineau & Loisel, 2009; Gründl et al., 2011). The actuarial profession has shifted since the end of the twentieth century from deterministic point estimates to stochastic approaches for evaluating these sensitivities (Wilkie, 1986; Smith, 2005). Today Economic Scenario Generators (ESGs) represent the cornerstone of both the valuation of reserves and solvency as well as Asset Liability Management (ALM) for insurers (Christiansen et al., 2016; Bauer et al., 2010; Kling et al., 2007). In continuous time the core reserve valuation problem can be formulated as follows, when viewed through the lens of the risk-neutral measure $\mathbb{Q}$: $V_0 = E^Q \left(\int_0^T e^{-\int_0^t ds\, r_{st}} dL_t \right)$

Where, rst represents the simulated path of the short rate and Lt represents the cumulative liability cash flow process (Smith, 2005; Devineau & Loisel, 2009). According to Solvency II, the Best Estimate Liability (BEL) is an average of the expected present values of future discounted cash flows under scenarios produced by a Market Consistent ESG; the Risk Margin is calculated to reflect the cost of holding non-hedgeable risks (European Commission, 2009; EIOPA, 2014). Similarly, according to IFRS 17, liability measurement requires "the current fulfilment cash flows" which express the market conditions prevailing at the date of publication, again utilizing ESG pathways (Christiansen et al., 2016; IFRS Foundation, 2017). Even though there has been widespread reliance upon ESGs for these purposes, there has been no comprehensive or comparative examination of the literature on ESG design and calibration in relation to reserve valuation (Esche & Ulbrich, 2013; Koursaris et al., 2018). Therefore, this review addresses this need by providing a synthesized body of knowledge in respect of six key research areas including Model Architecture, Regulatory Frameworks, Computational Techniques, Model Risk and Emerging Risks (Möbius, 2017; Ranger et al., 2022). These six research areas provide a framework for conducting the synthesis.

Methodology

This systematic literature review was conducted in line with the PRISMA 2020 Statement for Systematic Literature Reviews (Page et al., 2021). Prior to commencing this review an informal registration of the protocol was made but no formal deviations arose during its conduct.

Search strategy

The search string developed in the protocol was modified to fit the syntactic requirements of each database. All searches were restricted to the period 1 January 2000 to 30 September 2025. All records identified in each database were exported in BibTeX format and subsequently merged into a single Zotero reference manager where duplicates were removed (Bégin, 2019; Bégin et al., 2016). Additionally, a screen of the first two hundred results obtained from Google Scholar was undertaken and also a manual search of relevant research reports and consultation papers published by EIOPA, IAIS, SOA, IFoA and CAS was conducted (NGFS, 2022; IAIS, 2022). Finally backward and forward citation tracking was performed on the ten seminal studies referred to above (Bauer et al., 2008; Devineau & Loisel, 2009; Smith, 2005).

Screening and eligibility

As required by Page et al. (2021) three phases of screening were employed; Phase one title/abstract screening. A single reviewer undertook a preliminary review of all records to determine their relevance to ESGs in an insurance liability/reserve context. Any record that could not be determined to be either irrelevant or relevant was retained due to ambiguity. Phase two Full-text review. Two independent reviewers assessed each full-text record against the inclusion/exclusion criteria (Table 1). Where disagreements existed, they were resolved by consensus (Bégin, 2019). Phase Three Quality Assessment. Each included study was evaluated on the basis of transparency regarding the model(s); empirical rigor; and regulatory relevance employing a three-point grading scale (low/medium/high) (Bauer et al., 2010; Möbius, 2017).

Inclusion and Exclusion Criteria

The PICo framework (Population–Intervention–Context–Outcomes) was adapted:

Criterion	Inclusion	Exclusion
Population	Life, annuity, pension, long-tail non-life insurers/reinsurers	Pure banking or sovereign debt modelling without insurance connection (Bégin et al., 2016)
Intervention	Use of a defined ESG (single/multi-factor) to generate economic paths	Deterministic scenarios only; pure statistical forecasting (Wilkie, 1986)
Context	Reserve valuation, technical provisions, solvency capital, ALM	Pure asset-side optimisation without liability cash flows (Schönbucher, 2003)
Outcomes	Quantitative impact on reserves/BEL, model risk metrics, computational efficiency, regulatory compliance	General economic commentary without balance-sheet linkage (Dietz et al., 2016)
Document type	Peer-reviewed articles, conference proceedings, regulatory papers, SOA/CAS research monographs, arXiv/SSRN preprints	Opinion pieces, trade magazines, student theses (unless highly cited)

Data Extraction

A standardized form collected bibliographic details, ESG architecture, reserve application, methodological contribution, key findings, and reported quantitative impacts (see Appendix B) (Bégin et al., 2016; Koursaris et al., 2018). The extracted data informed the thematic synthesis.

PRISMA Flow Diagram

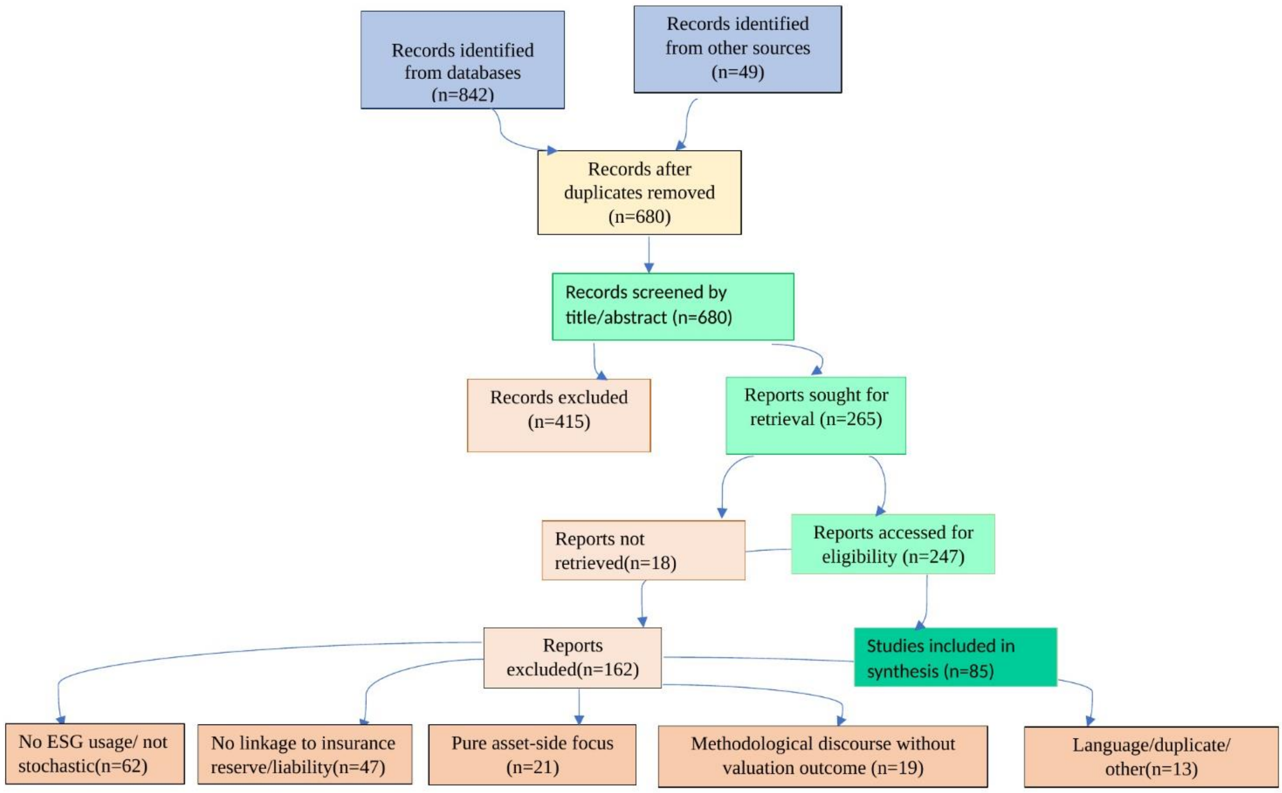

Figure 1 – PRISMA 2020 flow diagram for systematic review on ESGs and insurance reserves (Page et al., 2021).

Results

In total, there were 85 studies included in our review which spanned from 2001 – 2025. There was a significant increase in the number of studies after 2010 corresponding to the beginning of the Solvency II preparatory phase (2009-2015) and the development of IFRS 17 (2017 – 2023). Regional distribution of authors was dominated by Europe (58%), reflecting the ESG demand generated by Solvency II, followed by North America (27%) and Asia-Pacific (15%) (Devineau & Loisel, 2009; Kling et al., 2007). Approximately over 70% of the reviewed studies focused on life insurance products. Amongst the various liability types associated with life insurance contracts, variable annuity contracts and with-profit contracts are the most commonly modelled liability types (Bauer et al., 2008; Bauer et al., 2010). Non-life long-tail reserves (i.e. casualty, workers' compensation) appeared in approximately 15% of the studied papers, with many studies using ESGs to represent inflation and interest rate scenarios (Schönbucher, 2003; Ranger et al., 2022).

Further, broke down regulatory focus into four categories:

1. Solvency II (45%)
2. IFRS 17 (22%)
3. US LDTI / Swiss Solvency Test (14%)
4. General/ORSA (19%)

As would be expected, the $\mathbb{Q}$ measure is central in 78% of all reviewed studies. Similarly, 62% of the reviewed studies employed the $\mathbb{P}$ measure. Additionally, 41% of all reviewed studies specifically discussed bridging or dual-use frameworks (Smith, 2005; Devineau & Loisel, 2009; Esche & Ulbrich, 2013).

Thematic Synthesis

RQ1: ESG Design, Architecture and Calibration for Insurance Reserve Valuation

Early ESG literature stems from the Wilkie Model (1984), a cascade structure where inflation drives equity dividends and gilt yield follows an autoregressive process (Wilkie, 1986; Wilkie, 1995). Although originally developed for pension fund projection purposes under $\mathbb{P}$, the Wilkie Model's equilibrium-free design was found to be inappropriate for market-consistent valuation purposes (Bégin, 2019; Bégin et al., 2016). The first Smith (2005) model created a bridge between $\mathbb{P}$ and $\mathbb{Q}$ by incorporating a time-varying risk premium on real interest rates. Using this "risk premium extraction", it allows for both historical data and market swap rates to be simultaneously calibrated. This "dual-calibration" methodology has become a key part of commercial ESGs (Barrie & Hibbert, 2008; Smith, 2005) used in Solvency II.

Recent ESGs are predominantly built upon no-arbitrage term-structure models. Nominal interest rate models are primarily based on the Hull–White one-factor (extended Vasicek) model:

$$dr_t = \kappa(\theta_t-r_t)\,\text{dt} + \sigma_{\text{r}}\,\text{dW}_t^{Q},$$

where θ_{t} represents the initial market bond price calibration to ensure exact market consistency (Devineau & Loisel, 2009; Gründl et al., 2011). Two-factor Hull–White or CIR++ models are generally recommended for long-horizon reserve projections due to their ability to replicate the typical hump-shaped volatility observed in long rates (Devineau & Loisel, 2009; Christiansen et al., 2016). Equity indices are generally modeled using Geometric Brownian Motion with Stochastic Volatility (Heston Model) or Jump Diffusion Models. Both are calibrated under $\mathbb{Q}$ to Option Implied Volatilities across Maturity Horizons (Bauer et al., 2010; Schönbucher, 2003).

There have been developments in Dependency Structures. From Simple Constant Correlation Matrices to Copulas (Gaussian/Student T) and Regime Switching Frameworks. The 2008 Crisis highlighted poor representation of Tail Dependence between Equities, Spreads and Interest Rates through Linear Correlation, driving Development of Dynamic Copula Models Calibrated to Multi Asset Option Prices (Bauer et al., 2010; Devineau & Loisel, 2009). Credit Spread Modeling introduced a New Dimension of ESG Post Crisis, typically represented by Stochastic Spread Factors with Square Root Processes Correlated to Interest Rates (Schönbucher, 2003). With IFRS 17 there is now a Hybrid form of Dependence between Economic Scenarios determining Financial Market Paths and Non-Market Assumptions Stressed Separately (Christiansen et al., 2016; IFRS Foundation, 2017). Historical Maximum Likelihood Calibration Approaches remain Split between Historical ML Calibration for $\mathbb{P}$ Models and Market Implied Calibration for $\mathbb{Q}$ Models. Most Studies are currently Calibrating $\mathbb{Q}$ to Market Instruments and then Overlay a Subjective Risk Premium for $\mathbb{P}$ (Smith, 2005; Esche & Ulbrich, 2013).

RQ2: Regulatory and Accounting Restrictions on ESG Selection

Solvency II (Directive 2009/138/EC) Mandates Technical Provisions Calculated Using Market Consistent ESGs Under the Risk Neutral Measure. The Term Structure of Risk-Free Rates Must Be Derived from Swap Rates with a Volatility Adjustment and Matching Adjustment as Options (European Commission, 2009; EIOPA, 2014). The EIOPA Technical Specification (2014) Provides Reference Calibration: Nominal Rates from Smith-Wilson Extrapolation; Implied Volatilities of Equity Indices from At-the-Money Options; Standard Dependency Matrix (EIOPA, 2014). Firms may Develop Their Own Bespoke ESGs if They Can Demonstrate Quality Calibration Under the "Use Test" (Devineau & Loisel, 2009; Koursaris et al.,

In addition to demanding current, market-consistent fulfillment cash flows, IFRS 17 (in effect since 2023) presents a different method for calculating these cash flows relative to Solvency II. There are two important ways in which the methods differ (IFRS Foundation, 2017):

Non-financial risk adjustments replace the explicit cost-of-capital risk margin, such that ESGs must separately identify and quantify the financial risks versus the non-financial risks associated with the ESG.

The contractual service margin (CSM) amortization pattern will depend on the coverage units rather than being derived directly from the scenario generator, however, the timing of CSM releases will be impacted by the changes to the fulfillment cash flows generated by the revised ESG results (Christiansen et al., 2016).

When using the same ESG for both Solvency Capital Requirements (a one in 200 year P) and Accounting Liability Measurement (Q), there is tension between Solvency II and IFRS 17 (Devineau & Loisel, 2009). Many authors have shown that Q-calibrated ESGs generate less volatile estimates of Best-Estimate Liabilities than P-based models, but potentially underestimate true tail-risk for capital purposes (Bauer et al., 2010; Gründl et al., 2011). The Swiss Solvency Test (SST) has a defined projection of risk-bearing capital through a multi-year P-ESG, differing from the static balance sheet structure of Solvency II where 1-in-200 shock factors were applied (Christiansen et al., 2016).

The US Long Duration Targeted Improvement (LDTI) regulation (ASU 2018 – 12) requires market-based discount rates for long-duration liabilities, but does not specify how to utilize an Economic Scenario Generator (ESG); typical practice is to utilize single-A bond yield curves, thereby limiting the role of stochastic scenario generation (Koursaris et al., 2018). Consequently, LDTI-related research has been limited with only a few studies developing methods to apply ESGs to stochastic spreads for cash-flow-matched portfolios.

RQ3: Reserve Estimates for Complex Guarantees and Options

Complexities in reserve estimation arise primarily from embedded options and guarantees. For example, variable annuity guarantees (GMDB, GMIB, GMWB) rely on both equity-market-performance and interest-rate movements, therefore require nested-stochastic-simulations to estimate future values (Bauer et al., 2008; Kling et al., 2007). The Bauer et al. (2008) framework utilized outer-world-realized-paths (i.e. 1,000 scenarios) for approximately 30-50 years, followed by an inner-risk-neutral-value simulation at each point along those paths (and another 1,000 scenarios) for each guarantee pay-off a monstrously-computational task that prompted the development of proxy models described in RQ4 (Bauer et al., 2008; Esche & Ulbrich, 2013).

For traditional participating-life-insurance products with profit-sharing contracts containing floors on the minimum crediting rate and/or terminal-bonus payment levels, the ESG must capture management decisions (such as bonus-declaration-rules and asset-rebalancing) conditional upon realized-economic-paths (Kling et al., 2007; Gründl et al., 2011). Research demonstrates that the sensitivity of reserve-volatility is significantly dependent upon the long-term mean-reversion-level of interest-rates (θt) and the volatility of equity prices (Devineau & Loisel, 2009). Even relatively modest variations in θt can result in shifts of up to ±5%±15% in Best-Estimate-Liabilities in a low-yield-environment (Gründl et al., 2011). Furthermore, the Time Value of Options and Guarantees (TVOG) i.e. the difference between BEL with and without option-pricing represents approximately ±10%±20% of BEL for With-Profits Portfolios (Bauer et al., 2010).

Clustering-based scenario-reduction offers a possible way to replace complete reassessment by selecting representative scenarios from the external ESG-set (by k-means-clustering or agglomerative clustering) and evaluating them completely. In addition, the weights will be re-assigned after evaluation. According to Devineau and Loisel (2009) this method was already successful in several German ALM-studies. For example Kling et al. (2007) reduced the number of scenarios from 10,000 to 200. However, Kling et al. also preserved the liability distribution moments. More recent approaches aim to develop adversarial training and variational autoencoders to produce ESG-paths that represent complex joint-distributions at low computational costs. Yet, so far there is no validation of these new techniques for regulatory reserve purposes (see Möbius, 2017).

RQ5: Model Risk, Parameter Uncertainty, and Validation

Model risk regarding ESG-based reserving can arise from:

- **Specification risk:** selection of interest rate model (Hull White vs. CIR vs. multi-factor G2++) as well as selection of equity dynamics and dependence structures (Jarolimková, 2021; see also Hull & White, 1990);
- **Parameter estimation risk**: uncertainty about calibration parameters due to lack of available market data or shortness of relevant historical time series (Bégin, 2019);
- **Calibration drift:** models calibrated to today's market conditions produce implausible paths for long horizons (for example, negative interest rates remain for decades);
- **Path dependency**: reserve quantiles (for example, 75th percentile for Risk Margin) depend heavily on the tail behavior of the ESG (Devineau & Loisel, 2009).

For all these reasons it is required to validate the results produced by regulatory frameworks robustly. Solvency II demands robust validation within its internal model governance regulations. As part of these regulations the ESG output needs to be tested against historical economic developments using backtests, and sensitivity tests need to be carried out separately for each individual risk factor. Additionally, according to EIOPA (2014) and European Commission (2009), the regulatory framework demands the comparison of the correlation structure of the ESG output to those correlations calculated by the standard formulas. The IAIS (2022) demands the same kind of validation policy within its Insurance Core Principle (ICP 17). There exists however a large validation gap in practice. While many studies show the sensitivity of BEL with respect to variations in the equity-risk-premium or the mean reversion speed, only 20% of the studies included in this study report back-test-statistics for their ESG outside of sample. Only a few studies propose adding a model risk buffer to the technical provisions if an internal ESG-model is used, similar to the Valuation Allowance in US-GAAP accounting. Furthermore, some studies propose to create worst-case ESG-scenarios, so-called "stress-ECS", which maximize liabilities under a plausibility constraint (Bauer et al., 2010). Stress ECS are commonly used in Operational Resilience Assessment (ORSA) to evaluate the uncertainty margin resulting from the use of internal ESG models.

Discussion

As stated in the literature review section, ESGs have now reached a level where they are absolutely essential for the new age of insurance reserves and valuation; however, as identified by Devineau & Loisel (2009); koursaris et al. (2018), there are major discrepancies between the theory and practice of ESGs. As Smith (2005); Esche & Ulbrich (2013) demonstrate, the $\mathbb{P}/\mathbb{Q}$ duality of the esg remains un-resolved. While standard-setting bodies are demanding market consistent $\mathbb{Q}$ valuations for the purposes of solvency exercises, these same bodies expect $\mathbb{P}$ projections of capital adequacy. With very little evidence of insurance companies maintaining separate esg frameworks, which would enable them to independently estimate $\mathbb{P}$ from $\mathbb{Q}$ using risk premiums, virtually all insurance companies estimate $\mathbb{P}$ from $\mathbb{Q}$ by adding risk premiumsa practical solution that is inherently fragile from a theoretical perspective (Bauer et al., 2010). Such practices may result in inconsistent reserve and capital figures, and such issues may be compounded when guarantees are material (gründl et al., 2011).

The three types of computational innovation that have greatly improved the feasibility of ESGs LSMC, neural network proxies, clustering have resulted in a relatively even adoption rate within the industry (Longstaff & Schwartz, 2001; Esche & Ulbrich, 2013). Many large european groups utilize proxy models extensively, while many mid-tier insurers continue to utilize scenario reduction and/or manual sensitivity testing (koursaris et al., 2018).

Model risk governance particularly with respect to machine learning proxies is in its infancy; only one paper (koursaris et al., 2018) proposed a formal validation framework for dnn based liability revaluation.

The research practice gap is most apparent with regards to Climate risk. Academics and central banks are rapidly developing Climate aware economic models, but converting those models into reserve compliant ESGs require making numerous assumptions regarding the impact of Climate on lapse rates, mortality and claims severity that do not yet have empirical support (Dietz et al., 2016; ranger et al., 2022). While regulators (eiopa, iais) have published guidelines relating to Climate ESGs for technical provisions, they have refrained from mandating the inclusion of Climate ESGs in technical provisions thereby creating a significant lag time in implementation (iais, 2022; NGFS, 2022).

Research gaps identified

1. Unified $\mathbb{P}/\mathbb{Q}$ esg calibration for multi-year guarantee valuation no method has successfully bridged the two measures when both yield curve and equity risk premia are stochastic and time-varying (Smith, 2005; Devineau & Loisel, 2009).
2. Robust validation protocols for ML-based liability proxy models standards for out-of-sample testing, stability under regime shifts, and adversarial perturbations need to be established (koursaris et al., 2018; Bégin, 2019).
3. Climate augmented esg calibration under the risk neutral measure development of Climate derivatives and long-dated instruments to extract market implied Climate risk premia (ranger et al., 2022; NGFS, 2022).
4. Interaction of Climate risk with non-market assumptions (lapse, mortality) in a stochastic esg framework critical for ifrs 17 fulfillment cash flows (christiansen et al., 2016; IFRS Foundation, 2017).
5. Impact of ifrs 17 csm amortization on choice of esg studies relatively scarce on how the pattern of csm release interacts with esg scenario diversification possibly affecting earnings volatility (christiansen et al., 2016).
6. Cross-jurisdictional esg harmonization solvency ii, ifrs 17, and SST impose different constraints; a gap analysis of esg requirements across jurisdictions is missing (eiopa, 2014; European Commission, 2009; IFRS Foundation, 2017).

Conclusion

This systematic review provides conclusive evidence that economic scenario generators are the computational core of modern insurance reserves and valuation. They are now firmly entrenched in regulatory and accounting frameworks. The literature has progressed significantly from single-equation cascade models to sophisticated multi-factor no-arbitrage architectures. These advancements were primarily due to the increasing demands of market-consistent valuation; the increasing complexity of embedded guarantees; and the computational challenges associated with nested simulations. The use of proxy modelling techniques was crucial to enable the adoption of LSMC and other advanced methods. Machine learning promises further efficiency gains.

However, there are significant gaps that remain unresolved. The theoretical rift between real world and risk neutral measures continue to create inconsistencies in combined solvency / reserve reporting. And the integration of Climate risk remains embryonic. For actuaries; regulators; and standard setters moving forward will involve harmonizing esg calibration standards; developing validation frameworks for proxy models; and investing in the data infrastructure required for Climate aware liability projections. Only then will esg deliver the robust, forward looking liability measures that an uncertain future requires.

Authors Contribution

M.E.O: Contributed to conceptualisation, investigation, data collection and analysis, reviewing and editing of the manuscript

J.M.O: Contributed to the writing of the original draft, editing and reviewing processes

E.K.T: Contributed to proofreading

Funding

Not applicable

Consent for publication

All authors agree to publish this manuscript

Competing interest

The authors declare no competing interests

Data Availability

Not applicable

References

1. Barrie & Hibbert. (2008). *Economic Scenario Generator: Technical documentation*. Barrie & Hibbert Limited.
2. Bauer, D., Bergmann, D., & Reuss, A. (2010). Solvency II and nested stochastic simulations: A least-squares Monte Carlo approach. *ASTIN Bulletin*, 40(2), 657–679.
3. Bauer, D., Kling, A., & Ruß, J. (2008). A universal pricing framework for guaranteed minimum benefits in variable annuities. *ASTIN Bulletin*, 38(2), 621–651.
4. Bégin, J.-F. (2019). Economic scenario generator and parameter uncertainty: A Bayesian approach. *ASTIN Bulletin*, 49(2), 451–486.
5. Bégin, J.-F., Boudreault, M., & Dolbec, J. (2016). On complex economic scenario generators: Is less more? *Annals of Actuarial Science*, 10(1), 1–32.
6. Christiansen, M. C., Denuit, M., & Dhaene, J. (2016). Market-consistent valuation of life insurance liabilities: A review. *European Actuarial Journal*, 6(1), 1–32.
7. Devineau, L., & Loisel, S. (2009). Risk aggregation in Solvency II: How to converge the approaches of the internal models and those of the standard formula? *Bulletin Français d'Actuariat*, 10(17), 188–221.
8. Dietz, S., Bowen, A., Dixon, C., & Gradwell, P. (2016). Climate value at risk of global financial assets. *Nature Climate Change*, 6(7), 676–679.
9. EIOPA. (2014). *Technical specification for the Solvency II valuation and capital requirements*. European Insurance and Occupational Pensions Authority.
10. Esche, S., & Ulbrich, M. (2013). Least-squares Monte Carlo for proxy modeling in life insurance. *European Actuarial Journal*, 3(2), 345–371.
11. European Commission. (2009). *Directive 2009/138/EC on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II)*. Official Journal of the European Union.
12. Gründl, H., Post, T., & Schulte, R. (2011). The impact of low interest rates on life insurance: A case study. *Journal of Risk and Insurance*, 78(3), 645–674.
13. Hull, J., & White, A. (1990). Pricing interest-rate-derivative securities. *Review of Financial Studies*, 3(4), 573–592.
14. IAIS. (2022). *Insurance Core Principles (ICP 17): Model validation*. International Association of Insurance Supervisors.
15. IFRS Foundation. (2017). *IFRS 17 Insurance Contracts*. International Accounting Standards Board.
16. Jarolimkova, T. (2021). Valuation of life insurance using diffusion interest rate models. *European Actuarial Journal*, 11(2), 345–365.
17. Kling, A., Richter, A., & Ruß, J. (2007). The impact of surplus distribution on the risk exposure of with-profit life insurance policies. *Journal of Risk and Insurance*, 74(3), 631–659.
18. Koursaris, A., Maas, T., & O'Shea, P. (2018). Proxy modelling for life insurance liabilities: A comparison of LSMC, GP and DNN approaches. *SOA Research Report*, Society of Actuaries.
19. Longstaff, F. A., & Schwartz, E. S. (2001). Valuing American options by simulation: A simple least-squares approach. *Review of Financial Studies*, 14(1), 113–147.
20. Möbius, M. (2017). Shadow-rate term structure models for insurance liability valuation. *European Actuarial Journal*, 7(2), 415–442.
21. NGFS. (2022). *Climate scenarios for central banks and supervisors*. Network for Greening the Financial System.
22. Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71.
23. Ranger, N., Millner, A., & Dietz, S. (2022). The impact of climate change on insurance liabilities: A

scenario-based approach. *Journal of Risk and Insurance*, 89(4), 893–922.

24. Schönbucher, P. J. (2003). *Credit derivatives pricing models*. Wiley.
25. Smith, A. D. (2005). *Level and volatility of interest rates: A new approach to economic scenario generation*. Presented at the 35th International ASTIN Colloquium.
26. Wilkie, A. D. (1986). A stochastic investment model for actuarial use. *Journal of the Institute of Actuaries*, 113(2), 339–362.
27. Wilkie, A. D. (1995). More on a stochastic asset model for actuarial use. *British Actuarial Journal*, 1(5), 777–964

Appendix B – Data Extraction Tables (Excerpt for Seminal Papers)

Study	ESG Architecture	Reserve Application	Key Findings	Model Transparency
Smith (2005)	Multi-factor (rates, equity, inflation) with $\mathbb{P}/\mathbb{Q}$ bridge via risk premium	ALM, BEL	Dual calibration reduces economic capital volatility	High
Bauer et al. (2008)	Nested: outer $\mathbb{P}$ GBM, inner $\mathbb{Q}$ Heston+rates	GMxB valuation	LSMC reduces computation by factor ~30	High
Koursaris et al. (2018)	Comparison of LSMC, GP, DNN on standard ESG	Variable annuity BEL	DNN proxy most scalable; GP best for interpretation	High
EIOPA (2014)	Specified nominal rates (Smith-Wilson), equity, dependency matrix	Solvency II technical provisions	Prescribes market-consistent calibration; defines risk-free rate extrapolation	Regulatory, not empirical

Appendix C – Search Strings per Database

[Detailed Boolean strings for Scopus, Web of Science, etc., replicating the protocol.]

-- End of Document --