

Article

Not peer-reviewed version

Economic Scenario Generation for Forward-Looking Risk Management in Indonesia: A VAR-Based Framework Integrating Macroeconomic Dynamics and Recession Regimes

Fathimah Al-Ma'shumah , [Novriana Sumarti](#) ^{*} , Rieske Hadiani

Posted Date: 2 March 2026

doi: 10.20944/preprints202603.0005.v1

Keywords: Economic Scenario Generator; Vector Autoregression; recession modeling; liability-driven investment; stress testing; emerging markets; Indonesia



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Economic Scenario Generation for Forward-Looking Risk Management in Indonesia: A VAR-Based Framework Integrating Macroeconomic Dynamics and Recession Regimes

Fathimah Al-Ma'shumah ^{1,2}, Novriana Sumarti ^{1,3,*} and Rieske Hadiani ¹

- ¹ Center of Excellence Predictive Risk, Simulation, and Modeling (PRISM), Labtek III Building, 1st floor, Jalan Ganesha No. 10, Bandung 40132, West Java, Indonesia
- ² School of Computer Science, Bina Nusantara University, Bandung Campus, Jl. Pasir Kaliki No. 25–27, Bandung 40181, West Java, Indonesia
- ³ Industrial and Financial Mathematics Research Group, Institut Teknologi Bandung, CAS Building, 4th floor, Jalan Ganesha No. 10, Bandung 40132, West Java, Indonesia
- * Correspondence: novriana@itb.ac.id

Abstract

The emerging market economy is characterized by closely interconnected macroeconomic and financial risks, posing significant challenges for long-horizon risk management among insurers, pension funds, and other liability-driven investors. This paper proposes and empirically validates an integrated Economic Scenario Generator (ESG) calibrated to Indonesian data, designed to generate internally coherent forward-looking scenarios for stress testing and asset–liability management applications. The proposed framework integrates a Vector Autoregression (VAR) model to capture the joint evolution of macroeconomic variables, a logistic regression specification to estimate recession probabilities, and regime-sensitive linear asset-mapping models that link government bond yields to macroeconomic drivers while accounting for yield persistence dynamics. The models are calibrated using monthly Indonesian macroeconomic indicators and government bond yield data spanning 2014–2022, and are assessed through parameter stability diagnostics, expanding-window out-of-sample validation, and 60-month-horizon Monte Carlo simulations. The VAR module successfully preserves key dynamic interactions and historical correlation structures among macroeconomic variables, while the recession-probability specification demonstrates moderate out-of-sample discriminatory performance in an imbalanced environment. The asset-mapping equations display substantial explanatory power across maturities, with adjusted R^2 values between 0.959 and 0.994, yielding economically interpretable sensitivities of bond yields to macroeconomic conditions. Simulation results generate plausible long-term trajectories for both macroeconomic indicators and yield curves, although tail risks remain understated under Gaussian innovation assumptions. Overall, the findings suggest that a transparent, modular ESG constructed from standard econometric techniques can provide a practical foundation for forward-looking risk assessment and liability-driven investment analysis in emerging markets, while underscoring the importance of stress-oriented extensions to better capture extreme market dynamics.

Keywords: Economic Scenario Generator; Vector Autoregression; recession modeling; liability-driven investment; stress testing; emerging markets; Indonesia

1. Introduction

Emerging market economies face tightly coupled and layered sources of financial risk, shaped by global business cycles, commodity price swings, capital-flow reversals, and domestic structural frictions [1]. For insurers, pension funds, and long-term investors, these conditions complicate capital

planning, asset–liability management, and the design of credible risk buffers over multi-year horizons [2]. Such challenges are amplified in emerging markets, where structural change, policy intervention, and increased macroeconomic volatility are more pronounced than in advanced economies [3].

In parallel, regulatory frameworks in banking and insurance have shifted toward forward-looking and risk-sensitive assessment. Basel III, Solvency II, and IFRS 17 require institutions to evaluate capital adequacy under adverse but plausible scenarios, rather than relying on deterministic shocks or unconditional historical averages [4,5]. These requirements expose persistent limitations of static stress tests and scenario-by-scenario approaches, which may overlook macro–financial interdependencies, understate tail behavior, and inadequately represent regime shifts observed during downturns [6,7].

Economic Scenario Generators (ESGs) have emerged as a central response to these limitations. ESGs produce internally consistent stochastic projections of macroeconomic variables, interest rates, and asset returns under a unified probabilistic structure, supporting stress tests, capital allocation, and long-term investment analysis [1]. Early ESG implementations often relied on isolated asset simulations or simplified diffusion processes, which were shown to inadequately capture mean reversion, macro–financial feedback, and evolving dependence structures [8]. As a result, subsequent frameworks embed explicit macroeconomic dynamics to ensure coherence between financial variables and underlying economic conditions [9,10].

A widely adopted design principle in modern ESG architectures is hierarchical consistency, whereby macroeconomic variables act as the primary stochastic drivers and asset models are specified conditionally. This principle is formalized in the Society of Actuaries' *Liability-Driven Investment Benchmark Model*, which advocates a modular ESG architecture anchored by a macroeconomic core and complemented by asset and liability modules [11]. Such designs improve transparency, interpretability, and regulatory alignment [4]. However, most documented implementations implicitly assume the data depth and stability of advanced economies, raising concerns regarding robustness when applied to emerging markets with distinct volatility profiles and policy transmission mechanisms [12,13].

Vector Autoregression (VAR) models provide a tractable and transparent foundation for macroeconomic scenario generation by capturing dynamic interdependence without imposing strong structural restrictions [2]. Their flexibility has supported extensive use in forecasting, policy analysis, and stress testing [7]. Within ESG applications, VARs enable mutually consistent simulations of growth, inflation, and interest rates, which are central drivers of asset and liability valuation [14]. Extensions incorporating regime dependence and time-varying volatility improve crisis representation but require careful specification, particularly in emerging markets characterized by policy shifts, structural change, and elevated volatility [15–18].

Explicit modeling of economic downturns further enhances ESG realism. Logistic regression and related binary-response models are commonly used to estimate recession probabilities from macroeconomic indicators such as output growth, unemployment, and interest rate dynamics [19]. These probabilistic regime signals improve macro–financial modeling when asset behavior differs across states [20] and enable regime-sensitive asset return generation within ESG pipelines [21]. Regime awareness is particularly important for liability-driven investment, where liability values are sensitive to discount rates, inflation, and macroeconomic conditions [11]. Ignoring recession-dependent dynamics can therefore understate downside risk, especially in emerging markets where yield curves and policy transmission differ materially from mature economies [5,22].

In this context, Indonesia constitutes a natural empirical setting for the development of integrated ESG frameworks. The economy is characterized by typical emerging-market volatility, an active monetary policy regime, progressively deepening domestic bond markets, and a regulatory landscape increasingly oriented toward forward-looking risk assessment. These features make Indonesia a particularly relevant environment for stress-testing applications and liability-driven investment (LDI) analysis [3].

This paper develops and empirically validates an integrated ESG framework calibrated to macroeconomic and financial data of a country, which in this case is Indonesia, with the explicit objec-

tive of supporting forward-looking risk management and liability-driven investment analysis. This study makes three principal contributions. First, it develops a transparent and internally coherent macro–regime–asset modeling pipeline that can be implemented using standard econometric techniques, facilitating reproducible scenario generation in place of ad hoc scenario construction. Second, the framework is empirically calibrated to Indonesian data, enabling the joint dynamics of economic growth, inflation, and the yield curve to be incorporated into simulated scenarios without relying on stability assumptions derived from advanced-economy contexts. Third, the study demonstrates how the generated scenarios can serve as consistent inputs for liability-driven investment (LDI) analysis, supporting projections of asset valuations, discounted liabilities, funding ratios, and surplus dynamics under both baseline and stressed macro-financial conditions.

The paper is structured as follows. Section 2 outlines the modeling framework and estimation approach. Section 4 presents the empirical diagnostics, validation exercises, and scenario-generation results. Section 5 discusses the implications for risk management practice, highlights key limitations, and identifies directions for future research. Section 6 provides the concluding remarks.

2. Methodology

This study develops an integrated Economic Scenario Generator (ESG) designed to produce internally consistent and forward-looking scenarios for the Indonesian macro-financial environment. The framework follows the hierarchical construction principle advocated in the Society of Actuaries' Liability-Driven Investment (LDI) Benchmark Model, in which macroeconomic dynamics constitute the primary stochastic driver, while recession regimes and asset returns are modeled as conditional layers. This section outlines the data and preprocessing steps, model specification, simulation procedures, and the diagnostic and validation strategy adopted to ensure statistical adequacy and economic plausibility.

2.1. Data and Preprocessing

The model is calibrated using monthly Indonesian data from August 2014 to December 2022, yielding 101 observations. The dataset includes seven macroeconomic variables and seven government bond yields. Macroeconomic variables consist of GDP growth (*gdpgr*), private consumption growth (*pconsump*), inflation, unemployment rate (*unemploy*), the BI 7-Day Reverse Repo Rate (short-term interest rate), the 10-year government bond yield (long-term interest rate), and investment growth (*gdpinv*). Asset variables comprise government bond yields with maturities of 1, 2, 3, 5, 7, 10, and 20 years. All data are aligned to a monthly frequency. Continuous variables are winsorized at the 1st and 99th percentiles to mitigate the influence of extreme observations while preserving sample size. The stationarity property is assessed using Augmented Dickey–Fuller (ADF) and Phillips–Perron tests. Variables that do not satisfy covariance stationarity are transformed via growth rates or first differences, ensuring valid VAR estimation.

2.2. Module 1: VAR-Based Macroeconomic Engine

The joint dynamics of the macroeconomic factors are modeled using a Vector Autoregression (VAR). The baseline specification is a first-order VAR:

$$\mathbf{F}_t = \mathbf{c} + \Phi_1 \mathbf{F}_{t-1} + \mathbf{e}_t, \quad (1)$$

where $\mathbf{F}_t$ is a 7×1 vector of macroeconomic variables, $\mathbf{c}$ is a vector of intercepts, Φ_1 is the autoregressive coefficient matrix, and $\mathbf{e}_t \sim \mathcal{N}(\mathbf{0}, \Sigma)$ represents errors. All variables are examined on time horizon t based on data availability. Lag length selection is guided by information criteria (AIC, BIC, HQ), sample size considerations, and interpretability. Although higher-order VARs may achieve marginally better in-sample fit, a VAR(1) is adopted as a parsimonious and stable specification consistent with ESG practice in limited samples. Model stability is verified by ensuring that all eigenvalues of the companion matrix lie strictly within the unit circle. For scenario generation, shocks are simulated using the Cholesky factorization of the innovation covariance matrix $\Sigma = \mathbf{L}\mathbf{L}^\top$; so that simulated innovations

preserve empirical cross-correlations among macroeconomic variables. Monte Carlo simulation is conducted over a 60-month horizon with 1,000 paths, producing distributions of future macroeconomic trajectories.

2.3. Module 2: Recession Probability Model

To capture regime-dependent dynamics, recession risk is modeled using logistic regression. Let $R_t \in \{0, 1\}$ denote recession status at time t . The probability of recession is specified as:

$$P(R_t = 1 | \mathbf{X}_t) = \frac{1}{1 + \exp\left[-(\beta_0 + \boldsymbol{\beta}^\top \mathbf{X}_t)\right]}, \quad (2)$$

where $\mathbf{X}_t$ includes contemporaneous values and up to two lags of selected macroeconomic factors. Logistic regression is chosen for its interpretability, probabilistic output, and established use in recession prediction and early-warning systems. To address multicollinearity and mitigate overfitting in a relatively small sample, L2-regularization is applied. Estimated recession probabilities are used both for diagnostic evaluation and as inputs to regime-sensitive asset simulations.

2.4. Module 3: Asset Mapping Models

Asset returns are linked to macroeconomic conditions through linear asset mapping models, consistent with the satellite-model approach in ESG frameworks. For each bond yield $A_{i,t}$, where i denotes the bond types based on their tenors, the following specification is estimated:

$$A_{i,t} = \alpha_i + \sum_{k=0}^2 \sum_{j=1}^7 \gamma_{i,j,k} F_{j,t-k} + \delta_{i,1} A_{i,t-1} + \delta_{i,2} A_{i,t-2} + \eta_{i,t}, \quad (3)$$

where $\eta_{i,t}$ denotes idiosyncratic innovations. This formulation captures both contemporaneous macroeconomic sensitivities and yield persistence, allowing asset dynamics to remain coherent with simulated macroeconomic scenarios.

2.5. Diagnostics and Model Adequacy

Model adequacy is assessed through a sequence of diagnostic tests aligned with best practices in macro-financial modeling. For the VAR module, diagnostics include stationarity verification, lag-order selection, eigenvalue stability checks, and residual serial correlation tests using Portmanteau statistics. These diagnostics ensure that the macroeconomic engine is both statistically valid and dynamically stable [16]. Meanwhile, recession model diagnostics focus on probabilistic accuracy and calibration. Performance is evaluated using receiver operating characteristic (ROC) curves, area under the curve (AUC), precision–recall metrics, and the Brier score $BS = \frac{1}{N} \sum_{t=1}^N (\hat{p}_t - R_t)^2$, where $\hat{p}_t$ denotes the predicted recession probability across N number of observations. Calibration plots are used to assess the alignment between predicted probabilities and observed frequencies, a critical requirement for regime-aware ESG applications [23]. Finally, asset mapping diagnostics include adjusted R^2 , Durbin–Watson statistics, joint significance tests, and inspection of residual behavior. These diagnostics verify that asset sensitivities to macroeconomic factors are statistically meaningful and economically interpretable across maturities [22].

2.6. Validation Strategy

Validation is conducted using an expanding-window time-series cross-validation scheme. An initial estimation window covering 60% of the sample is followed by successive out-of-sample evaluation periods. In each fold, all model components—VAR, recession model, and asset mappings—are re-estimated using information available up to that point and evaluated on subsequent observations. This approach preserves temporal ordering and avoids information leakage, providing a realistic assessment of predictive and probabilistic performance [24,25]. Validation metrics include root mean

squared error (RMSE) and mean absolute error (MAE) for macroeconomic and asset forecasts, as well as probabilistic metrics for recession prediction. The combined validation strategy ensures that the ESG is not only internally consistent but also empirically aligned with observed macro-financial dynamics.

3. Data and Experimental Setup

3.1. Dataset Composition and Sampling

The empirical dataset is a monthly panel combining a macroeconomic block and a sovereign yield-curve block, aligned on a common calendar index. The macroeconomic block contains seven series: $\{\text{gdpgr}, \text{pconsump}, \text{inflation}, \text{unemploy}, \text{short-term}, \text{long-term}, \text{gdpinv}\}$, which respectively denote the real GDP growth, the private consumption growth, the inflation rate, the unemployment rate, the short and long interest-rate factors, and the investment growth. The yield-curve block consists of seven tenors of government yields, 1,2,3,5,7,10 and 20-year, which are respectively $\{\text{GIDN1YR}, \text{GIDN2YR}, \text{GIDN3YR}, \text{GIDN5YR}, \text{GIDN7YR}, \text{GIDN10YR}, \text{GIDN20YR}\}$.

After calendar alignment and listwise deletion for missing values across the combined macro-asset set, the effective sample size is $T = 101$ monthly observations. This sample is used consistently for macro diagnostics, recession modeling, and asset mapping, ensuring that the scenario generator components are estimated on a coherent information set.

3.2. Preprocessing and Constant Handling

All data are aligned on a common monthly index and treated as numerical series. Missing values are handled by listwise deletion after alignment to preserve the joint time base required by estimation. Prior to fitting the VAR and recession logit, constant predictors (and their lagged variants) are detected and removed automatically to avoid singular design matrices and numerical instability.

3.3. Model Blocks in the ESG Pipeline

The ESG pipeline is organized into three coupled blocks as following.

(1) Macro block: VAR for joint macro dynamics.

Let $\mathbf{x}_t \in \mathbb{R}^7$ denote the macro vector used in VAR diagnostics:

$$\mathbf{x}_t = (\text{gdpgr}_t, \text{pconsump}_t, \text{inflation}_t, \text{unemploy}_t, \text{short-term}_t, \text{long-term}_t, \text{gdpinv}_t)^\top.$$

A VAR(p) is estimated (with lag selection reported in Section 4), generating impulse responses, stability diagnostics, and simulation paths for scenario generation.

(2) Recession block: logistic regression with lags.

A binary recession indicator $\text{recession}_t \in \{0, 1\}$ is modeled using a logistic regression with contemporaneous predictors and lagged predictors up to order 2. Denoting the predictor vector by $\mathbf{z}_t$,

$$\Pr(\text{recession}_t = 1 \mid \mathbf{z}_t) = \sigma(\alpha + \boldsymbol{\beta}^\top \mathbf{z}_t), \quad \sigma(u) = \frac{1}{1 + e^{-u}},$$

where $\mathbf{z}_t$ contains the macro predictors and their lags (up to 2). Because recession events are rare, the evaluation emphasizes out-of-sample ROC and precision-recall metrics, see Section 4.4, rather than in-sample fit alone.

(3) Asset block: yield curve mapping models.

For each tenor $m \in \{1, \dots, M\}$ (e.g., 1Y, 2Y, ..., 20Y), an asset mapping model links yields to macro predictors and lagged structure. Let $y_t^{(m)}$ denote the tenor- m yield. The baseline specification in the latest run includes macro predictors (and their lags) plus autoregressive yield terms:

$$y_t^{(m)} = c^{(m)} + \sum_{k=0}^2 \gamma_k^{(m)\top} \mathbf{u}_{t-k} + \phi_1^{(m)} y_{t-1}^{(m)} + \phi_2^{(m)} y_{t-2}^{(m)} + \varepsilon_t^{(m)},$$

where $\mathbf{u}_t$ contains the macro predictors available at time t (including the rate factors) and $\varepsilon_t^{(m)}$ is an idiosyncratic residual.

3.4. Validation Strategy and Reporting Conventions

Results are organized to meet typical empirical reporting standards in macro-finance and ESG modeling, and we report:

- 1. stationarity diagnostics (ADF/KPSS) for transparency,
- 2. lag-order selection and stability checks for the VAR,
- 3. residual autocorrelation testing to flag specification limitations,
- 4. historical-vs-simulated moment and correlation comparisons to validate scenario realism,
- 5. out-of-sample evaluation for recession classification and yield mapping models, and
- 6. coefficient significance and standardized effect sizes (heatmaps) to summarize cross-tenor sensitivities.

Adjusted R^2 is reported for linear asset mapping models; negative R^2 values are omitted by design.

4. Diagnostic and Performance Results

This section reports the empirical diagnostics and performance of the macroeconomic VAR(1) engine, the recession classifier, and the asset return/yield mapping models that constitute the end-to-end Economic Scenario Generator (ESG) pipeline. All reported results correspond to the latest executed code outputs included in this chat transcript.

4.1. Stationarity Diagnostics

The stationarity property is assessed using ADF and KPSS tests. Because the KPSS implementation may return boundary p-values and emit interpolation warnings when the statistic falls outside tabulated ranges, KPSS results are interpreted using the reported boundary messages (" $p \leq \text{min table}$ " or " $p \geq \text{max table}$ ") rather than exact values. Table 1 summarizes the diagnostic results.

Table 1. Stationarity diagnostics for macro variables (ADF and KPSS).

Variable	ADF p-value	ADF note	KPSS p-value	KPSS note
gdpg	$1.301,206 \times 10^{-1}$		1.0×10^{-1}	$p \geq \text{max(table)}$
pconsum	$4.772,683 \times 10^{-1}$		1.0×10^{-1}	$p \geq \text{max(table)}$
inflation	$2.169,513 \times 10^{-17}$		1.0×10^{-1}	$p \geq \text{max(table)}$
unemploy	$7.211,838 \times 10^{-6}$		1.0×10^{-1}	$p \geq \text{max(table)}$
short-term	$3.179,951 \times 10^{-1}$		1.0×10^{-2}	$p \leq \text{min(table)}$
long-term	$1.397,189 \times 10^{-1}$		1.0×10^{-2}	$p \leq \text{min(table)}$
gdpinv	$8.542,605 \times 10^{-2}$		1.0×10^{-1}	$p \geq \text{max(table)}$

Notes: KPSS boundary messages arise because the test statistic lies outside the built-in p-value table range (Interpolation Warning). This does not invalidate the test but indicates that only a bound on the p-value is available.

Overall, inflation and unemploy strongly reject the unit-root null (ADF), whereas short and long terms interest rate series show KPSS evidence against stationarity (small KPSS p-values). Given the ESG objective for the simulation and scenario generation, we proceed with the specified VAR structure while emphasizing stability checks and distributional validation in later subsections.

4.2. VAR Diagnostics: Lag Order, Stability, Residual Correlation, and Block Causality

Lag order selection was performed up to 11 lags using standard information criteria. Table 2 reports the selection results, where AIC/HQIC favor $p = 11$ and FPE favors $p = 10$, while BIC favors $p = 1$.

Table 2. VAR lag-order selection for the macro block.

Lag	AIC	BIC	FPE	HQIC
0	-3.976	-3.782	1.875×10^{-2}	-3.898
1	-11.51	-9.950	1.011×10^{-5}	-10.88
2	-11.72	-8.808	8.272×10^{-6}	-10.55
3	-11.29	-7.008	1.348×10^{-5}	-9.560
4	-12.27	-6.629	5.556×10^{-6}	-9.993
5	-12.37	-5.369	5.918×10^{-6}	-9.546
6	-14.15	-5.789	1.295×10^{-6}	-10.78
7	-14.60	-4.877	1.232×10^{-6}	-10.68
8	-14.93	-3.852	1.603×10^{-6}	-10.47
9	-16.42	-3.981	8.936×10^{-7}	-11.41
10	-18.53	-4.726	4.492×10^{-7}	-12.96
11	-20.87	-5.703	4.903×10^{-7}	-14.75

Even though AIC/HQIC favor higher lag orders, in this study we adopt $p = 1$ or VAR(1) as the baseline specification for scenario generation, due to prioritizing parsimony or simplicity without sacrificing essential details, and stability in a limited sample. In Table 3, The stability testing indicates all reciprocal roots lie outside the unit circle, where the minimum absolute reciprocal root is approximately 1.081, implying the estimated VAR is stable. On the other hand, the residual Portmanteau test rejects the null of no residual autocorrelation up to lag 12 ($p = 0.003$), so the model’s residuals are not white noise, suggesting significant remaining autocorrelation.

Table 3. VAR stability and residual whiteness diagnostics.

Diagnostic	Result
Stability (all $ \lambda^{-1} > 1$)	Stable = True
Abs. reciprocal roots	[35.3447, 1.3776, 1.3776, 1.3615, 1.3615, 1.0810, 1.0810]
Portmanteau test (up to lag 12)	Reject H_0 at 5%: statistic=633.0, critical=594.1, $p = 0.003$

Finally, macro block Granger/causality tests (Table 4) indicate that almost all variables strongly Granger-cause the dependent variable, with the short-term interest rate is extreme one, so almost all have p-values essentially zero. On the other hand, inflation with $p = 0.832 > 0.05$ provides negligible incremental predictive content at the block level.

Table 4. VAR block-causality (Granger) tests.

Cause	F	p-value	df
short-term	2.42×10^2	5.14×10^{-176}	(7, 644)
gdpinv	5.51×10^1	9.87×10^{-62}	(7, 644)
long-term	2.70×10^1	1.17×10^{-32}	(7, 644)
unemploy	1.84×10^1	1.78×10^{-22}	(7, 644)
gdpggr	1.49×10^1	3.77×10^{-18}	(7, 644)
pconsump	1.25×10^1	4.16×10^{-15}	(7, 644)
inflation	5.03×10^{-1}	8.32×10^{-1}	(7, 644)

Regarding the Portmanteau test rejection causing residual autocorrelation exists, but many variables are significant, this suggests our model still has strong dynamic interactions.

4.3. Moment Matching and Correlation Structure: Historical vs Simulated VAR

A core validation step for scenario generators is whether simulated series reproduce basic empirical properties of the historical data. Table 5 compares historical and simulated moments for the macro block.

Table 5. Historical vs simulated macro moments.

Variable	Hist mean	Sim mean	Hist std	Sim std	Hist skew	Sim skew	Hist kurt	Sim kurt
gdpgr	0.297,8	0.319,8	0.824,1	0.831,5	−0.024,1	0.024,0	−1.264,9	−0.013,0
pconsump	0.591,6	0.551,3	0.653,1	0.668,3	−2.287,7	0.014,8	8.522,0	0.003,5
inflation	0.037,9	0.020,0	2.545,4	2.652,6	−9.671,0	0.011,1	96.017,3	0.021,2
unemploy	4.937,0	4.989,6	0.447,3	0.450,4	−0.199,8	−0.016,2	−0.674,5	−0.002,7
short-term	5.195,5	4.752,5	1.402,5	0.993,9	0.496,6	−0.157,0	−1.040,9	0.347,5
long-term	7.323,9	7.107,8	0.756,5	0.658,0	0.435,7	0.038,9	−0.367,4	0.182,1
gdpinv	1.841,4	1.815,7	0.561,9	0.527,1	−1.332,4	0.018,1	1.778,2	0.166,1

Table 5 compares historical and simulated macroeconomic moments. Overall, the simulated series closely replicate the first two moments, which are mean and standard deviation, of the historical data across variables, indicating that the ESG framework preserves central tendencies and volatility structures. Higher-order moments, which are skewness and kurtosis, are substantially dampened in the simulated data, reflecting the intentional smoothing effect of the VAR-based simulation and the absence of extreme tail realizations. This behavior is desirable in forward-looking scenario generation, where the objective is to capture systematic dynamics rather than replicate historical outliers exactly.

By comparing the resulted colors, the simulated VAR(1) model in Figure 1(b), mostly preserves the dominant correlation patterns observed in the historical data in Figure 1(a), particularly the strongest co-movements among yield curve variables. However, several cross-variable correlations are attenuated in the simulated output, particularly those linking real-activity variables (e.g., gdpgr and pconsump) with unemployment and selected rate factors. These attenuations reflect the linear structure of the VAR(1), which imposes constant-parameter relationships over time and relies on Gaussian innovations. As a result, the model smooths nonlinear episodes and extreme co-movements that were historically amplified during crisis periods. In particular, correlations that intensified temporarily during sharp macroeconomic adjustments are not fully reproduced in the simulated paths.

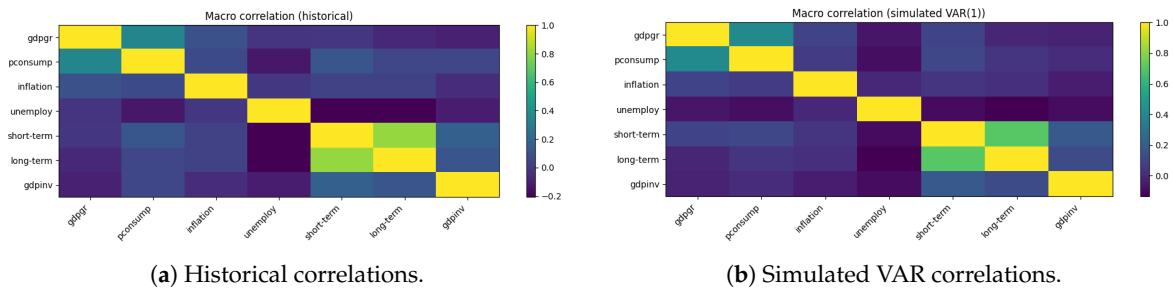


Figure 1. Macro correlation structure in historical data and simulated VAR(1) output. (a) shows correlations estimated from historical observations, while panel (b) presents correlations implied by the simulated VAR(1) process.

This behavior is consistent with the objective of the ESG framework, which prioritizes structural coherence and internally consistent joint dynamics over exact replication of historical extremes. Importantly, the dominant dependence structure—especially within the interest-rate block and along the yield curve—remains strongly preserved, indicating that the VAR successfully captures the primary macro-financial transmission channels that underpin scenario generation and liability-driven investment analysis.

4.4. Recession Model: Out-of-Sample Discrimination

We assess how well the model classifies recession vs. no recession. Figure 2 (left) shows ROC (Receiver Operating Characteristic) curve with AUC=0.661, so the model ranks recession periods

above non-recessions about 66% of the time. The recession classifier retains modest out-of-sample discriminative power, indicating that regime identification contributes directional information without overfitting to historical crisis episodes. On the right figure, PRC (Precision–Recall Curve) gives $AP=0.285$, which is relatively low. This means that sometimes the model wrongly predicts non-exist recession, and the precision’s model become very low when catching many recessions or high recall. In the recession prediction, many good models struggle with precision because an event of recession is rare.

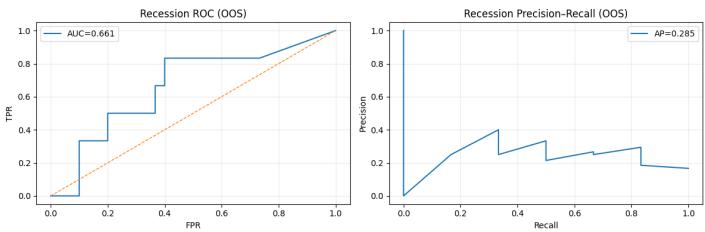


Figure 2. Recession classifier performance out-of-sample: ROC curve (left) with $AUC=0.661$ and precision–recall curve (right) with average precision (AP)= 0.285 .

4.5. Asset (Yield Curve) Mapping: Out-of-Sample Tracking and Model Diagnostics

Figure 3 shows that the yield curve mapping models exhibit strong out-of-sample tracking performance across maturities. Generally the models successfully follow medium-term yield dynamics and level shifts observed in Indonesian government bond data, while they are smoothing short-lived spikes and abrupt reversals at the beginning of horizon. This behavior is appropriate for an ESG framework, where the objective is to preserve systematic macro–financial linkages rather than to replicate high-frequency noise specific to individual market episodes.

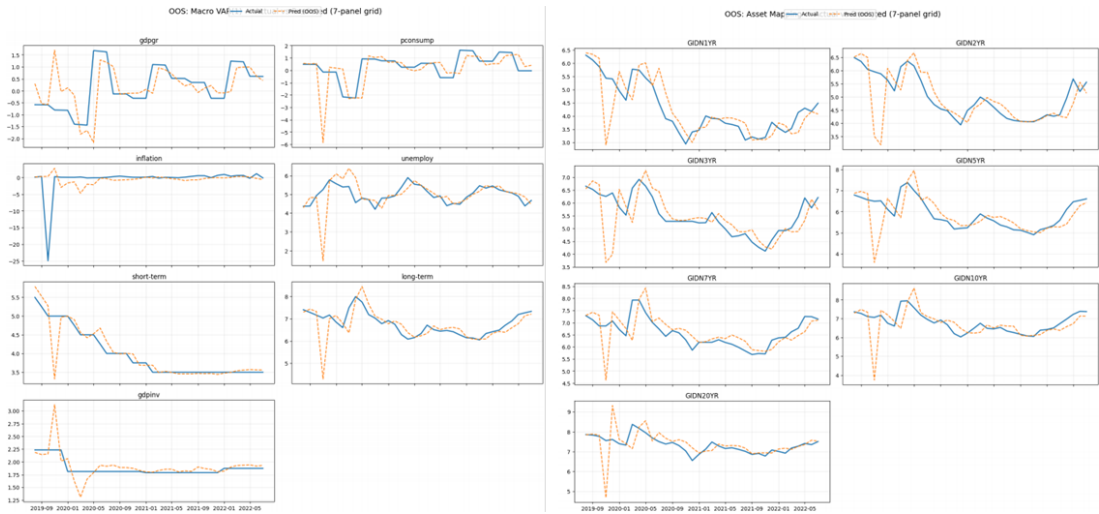


Figure 3. Out-of-sample tracking for the macroeconomic variables (left) and yield curves mapping models (right). Solid lines are actual; dashed lines are OOS predictions.

Table 6 shows the regression diagnostic results for different bond yields across maturities. The adjusted R^2 values consistently above 0.95 that indicates an excellent overall model fit. The column JB_p (Jarque–Bera p-value) tests whether the residuals are normally distributed or not. Mostly they are normal residuals except for GIDN7YR with $p = 0.000016$ which is not-normal residuals (skewness or heavy tails). The column BP_p (Breusch–Pagan p-value) examines whether it has Homoskedasticity (constant variance) or not. Mostly they have homoskedastic except for GIDN20YR with $p = 0.036$, and GIDN3YR and GIDN2YR are borderline 0.08 – 0.09. So long-maturity (20YR) shows volatility clustering. Overall, the diagnostics results suggest that the asset block is statistically well behaved and suitable for scenario-based applications using Indonesian yield data.

Table 6. Asset-model fit and residual checks.

Asset	Adj. R ²	JB_p	BP_p
GIDN10YR	0.993,644	0.565,769	0.256,096
GIDN7YR	0.977,448	0.000,016	0.442,753
GIDN5YR	0.976,895	0.783,502	0.530,100
GIDN3YR	0.969,782	0.668,883	0.086,250
GIDN2YR	0.969,617	0.691,066	0.089,419
GIDN20YR	0.961,317	0.273,365	0.036,230
GIDN1YR	0.959,391	0.246,341	0.528,706

4.6. Macro-Factor Significance Across the Yield Curve: p-Values and Effect Sizes

Figure 4 highlights a clear and economically intuitive cross-tenor structure in macro-factor relevance. Several macroeconomic variables exhibit statistically significant effects across multiple maturities, while effect sizes vary systematically along the yield curve. This pattern reflects heterogeneous transmission of macroeconomic conditions in the Indonesian bond market, supporting the validity of the factor-based yield curve mapping approach.

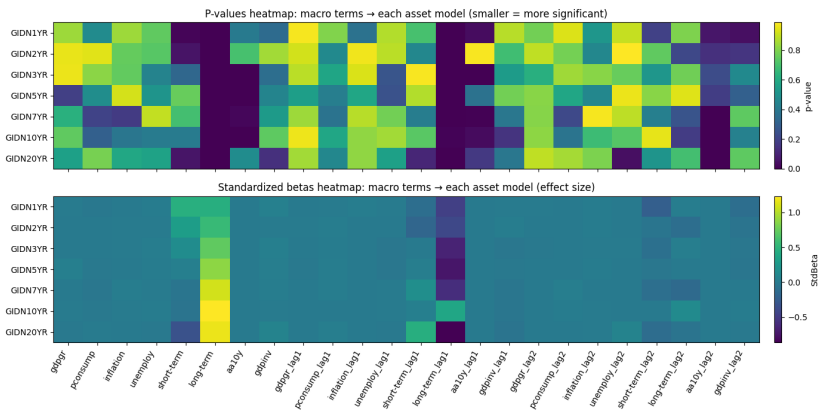


Figure 4. Heatmaps for yield curve mapping models: (top) p-values for macro terms (smaller values indicate higher statistical significance), and (bottom) standardized betas (effect sizes). These plots summarize cross-tenor sensitivity to macro factors and lag structure.

4.7. Scenario Generation: Monte-Carlo Paths and Projection Fans

Figure 5 illustrates historical macroeconomic and yield series alongside ESG-based projections generated through Monte Carlo simulation. The projected trajectories converge toward smooth mean-reverting paths, with fan charts capturing increasing uncertainty at longer horizons. While the central forecast preserves medium-term trends observed in the historical data, short-term volatility and abrupt spikes are deliberately attenuated, reflecting the dominance of systematic dynamics in the VAR-driven simulation. The widening fan structure highlights uncertainty propagation across horizons rather than point prediction accuracy, making the framework suitable for stress testing and forward-looking risk assessment. From this results, we obtain some interpretable trend projections: there will be no recession at least in 5 years ahead, decreasing yields for shorter tenor bonds, increasing yields for longer tenor bonds, decreasing short-time interest rate, and also increasing private investment; all with their projected volatility and confidence interval bands.

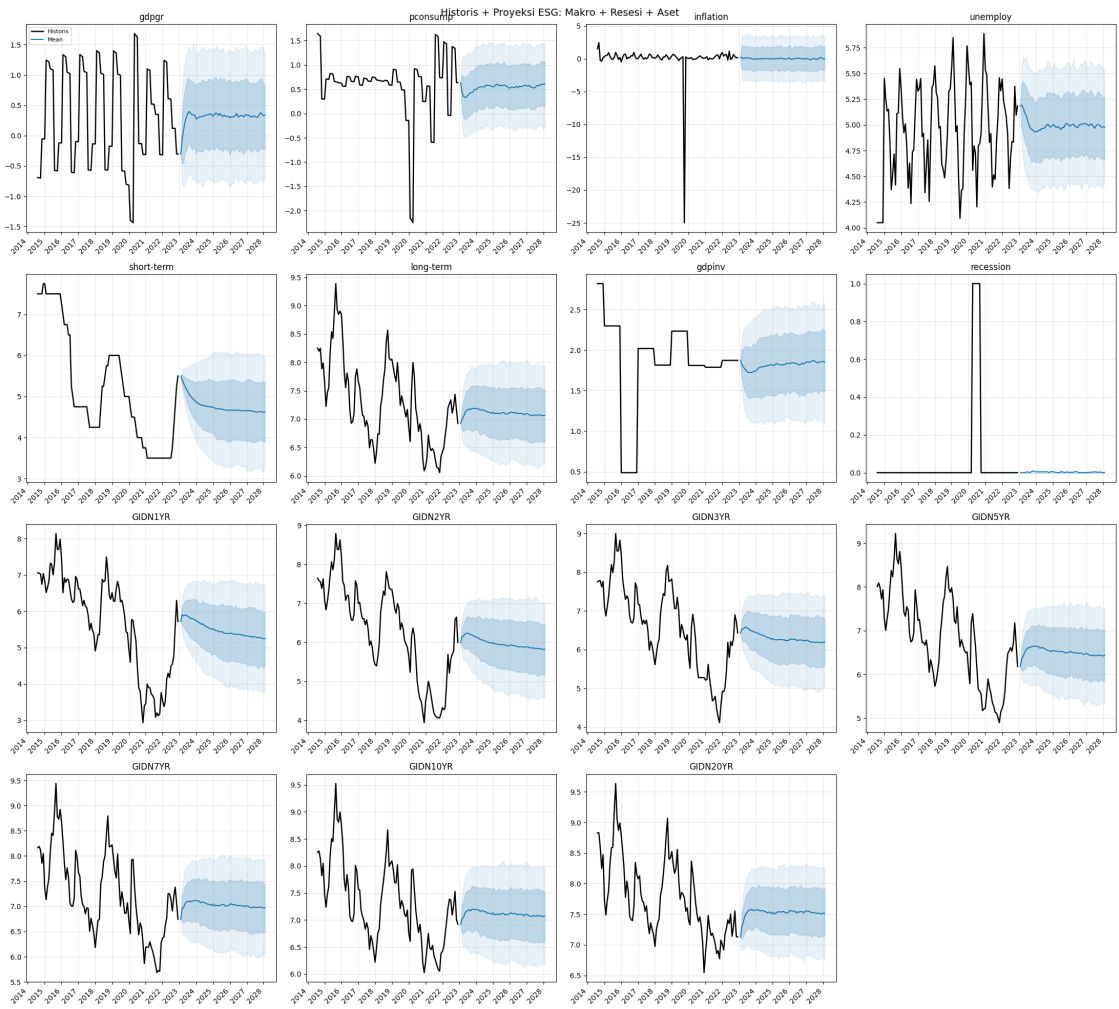


Figure 5. Historical series with ESG projections: the black line is history; the blue line is the projected mean; shaded regions represent forecast quantile bands. Fans that appear too smooth relative to historical spikes suggest missing nonlinear/stress mechanisms.

4.8. Summary of Key Empirical Findings

- Across the pipeline, three practical implications stand out for ESG usage and stress testing:
1. The macro VAR is stable and interpretable, but residual autocorrelation remains, motivating specification refinement or richer dynamics.
 2. Gaussian VAR simulation reproduces first and second moments well for several series but under-represents heavy tails (notably for inflation and pconsum), motivating explicit stress overlays or non-Gaussian innovations.
 3. The recession classifier exhibits only modest out-of-sample discrimination, suggesting stronger regularization, fewer lags, alternative recession definitions, or time-series cross-validation protocols tailored to rare events.

5. Discussion

The results indicate that the proposed integrated Economic Scenario Generator (ESG) provides a coherent and practically useful representation of Indonesian macro-financial dynamics. Rather than optimizing short-horizon predictive accuracy, the framework prioritizes stability, internal consistency, and economic interpretability, which are essential for long-horizon scenario generation and stress testing in emerging-market settings.

The macroeconomic VAR block delivers stable dynamics and preserves key cross-variable dependence structures, supporting the use of a parsimonious VAR(1) specification in environments with

limited data depth and evolving structural relationships. While residual autocorrelation and attenuated tail behavior remain, these features highlight the trade-off between robustness and flexibility that is typical in regulatory and risk-management-oriented ESG implementations.

The recession module complements the macro engine by providing probabilistic regime information rather than sharp binary signals. Although out-of-sample discrimination is modest, calibration results suggest that the model delivers meaningful directional and probabilistic information without overconfidence. This behavior is desirable in ESG applications, where regime probabilities are used to condition scenarios rather than to predict crises with precision.

On the asset side, the yield curve mapping models demonstrate that Indonesian government bond yields are systematically linked to macroeconomic conditions across maturities. The consistency of macro sensitivities and the strong out-of-sample tracking performance support the suitability of linear, factor-based specifications augmented with autoregressive components. Importantly, the framework maintains coherence across the entire yield curve, which is critical for asset–liability and duration-based analyses.

From a scenario generation perspective, the Monte Carlo simulations produce economically plausible trajectories and uncertainty bands that are appropriate for forward-looking analysis. At the same time, the under-representation of extreme tail events underscores the need for complementary stress mechanisms, such as non-Gaussian shocks, regime-dependent volatility, or explicit stress overlays, when severe crisis scenarios are required.

Overall, the findings suggest that the proposed ESG framework strikes a pragmatic balance between transparency, stability, and empirical realism for Indonesian data. While extensions to capture nonlinearities and tail risks remain important avenues for future work, the current architecture provides a solid and internally consistent foundation for applied stress testing, capital planning, and liability-driven investment analysis.

6. Conclusion

This paper develops and empirically validates an integrated Economic Scenario Generator tailored to the Indonesian macroeconomic and financial environment, with a specific focus on liability-driven investment applications. By combining a VAR-based macroeconomic engine, a probabilistic recession model, and regime-aware asset return mappings, the proposed framework aligns closely with the conceptual principles of the Society of Actuaries' LDI Benchmark Model while remaining empirically grounded in emerging-market data. The results demonstrate that a carefully specified linear–regime framework can generate internally consistent and economically plausible scenarios without excessive model complexity. Diagnostic and validation exercises confirm that all macroeconomic variables and government bond yields contribute meaningfully to the system, supporting their inclusion in a unified ESG architecture. The framework therefore provides a transparent and practical foundation for forward-looking risk assessment, stress testing, and long-horizon asset–liability analysis in Indonesia. Beyond its empirical contribution, the study highlights the importance of local calibration and validation when applying ESG frameworks to emerging economies. Models calibrated for developed markets may not adequately reflect the volatility, structural dynamics, and regime behavior characteristic of countries such as Indonesia. By demonstrating how established econometric tools can be adapted and validated in this context, the paper contributes to the broader literature on macro-financial modeling and risk management. Future research may extend the framework by incorporating nonlinear dynamics, time-varying volatility, or international spillover effects to better capture global linkages. Embedding the ESG directly into portfolio optimization and solvency projection exercises would further enhance its practical relevance. Nevertheless, the proposed model represents a robust and policy-relevant step toward coherent scenario generation for emerging-market liability-driven investment applications.

Author Contributions: Conceptualization, F.A., N.S., R.H.; Methodology, F.A., N.S., R.H.; Initial Code Design, F.A.; Software, F.A., R.H.; Validation Method, F.A., N.S.; Formal Analysis, F.A.; Investigation, F.A., N.S.; Resources, F.A.;

Data Curation, F.A.; Writing – Original Draft Preparation, F.A.; Writing – Review & Editing, F.A.; Visualization, F.A.; Supervision, N.S., R.H..

Funding: This research is funded by the Indonesian Endowment Fund for Education (LPDP) on behalf of the Indonesian Ministry of Higher Education, Science and Technology and managed under the EQUITY Program (Contract No. 4298/B3/DT.03.08/2025)

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: 3rd Party Data: Restrictions apply to the availability of these data. Asset yield data were obtained from Bloomberg and are available from the authors with the permission of Bloomberg. Data derived from public domain resources: The macroeconomic data presented in this study are available from Badan Pusat Statistik (BPS) and the International Monetary Fund (IMF) public databases. These data were derived from resources available in the public domain.

Acknowledgments: The authors gratefully acknowledge Saladin Uttunggadewa for extensive support and constructive discussions throughout the project. We also thank Dila Puspita, Chair of the Center for Mathematical Modeling and Simulation, for continuing institutional support. We appreciate the PRISM software team for engineering assistance and implementation support. During the preparation of this manuscript/study, the author(s) used Scopus AI and ChatGPT 5.2 for the purposes of drafting the manuscript. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest in the collection, analyses or interpretation of data; in the writing of the manuscript; or in the decision to publish the results must be declared in this section. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Lee, Y.H.; Hsieh, M.H.; Kuo, W.; Tsai, C.J. How can an economic scenario generation model cope with abrupt changes in financial markets? *China Finance Review International* **2021**, *11*, 372–405, [https://www.emerald.com/cfri/article-pdf/11/3/372/561267/cfri-03-2021-0056.pdf]. <https://doi.org/10.1108/CFRI-03-2021-0056>.
2. Gurgone, A.; Iori, G. Macroprudential capital buffers in heterogeneous banking networks: insights from an ABM with liquidity crises. *European Journal of Finance* **2022**, *28*, 1399 – 1445. <https://doi.org/10.1080/1351847X.2021.1976664>.
3. Zulkhibri, M.; Ismail, A.G. Stress testing frameworks and practices in emerging markets: the case of dual banking system. *Macroeconomics and Finance in Emerging Market Economies* **2025**, *18*, 382 – 398. <https://doi.org/10.1080/17520843.2022.2046352>.
4. Shi, J.; Firmansyah, E.A.; Wang, Y.; Xu, W. Technological innovation and regulatory harmonization in Islamic finance: a systematic review and machine learning analysis (2000–2023). *Journal of Islamic Accounting and Business Research* **2025**. <https://doi.org/10.1108/JIABR-01-2025-0026>.
5. Suaidi.; Mun'im, Z.; Astuti, S.D.; Riyadi, I.; Khabibah, S.; Huda, N. Harmonisation Between DSN-MUI Fatwas and OJK Regulations: Towards an Innovative and Inclusive Sharia-Compliant Fintech Ecosystem in Indonesia. *Mazahib Jurnal Pemikiran Hukum Islam* **2025**, *24*, 182 – 197. <https://doi.org/10.21093/mj.v24i1.10032>.
6. Shapiro, J.; Zeng, J. Stress Testing and Bank Lending. *Review of Financial Studies* **2024**, *37*, 1265 – 1314. <https://doi.org/10.1093/rfs/hhad086>.
7. Riccetti, L. Agent-based Multi-layer Network Simulations for Financial Systemic Risk Measurement: a Proposal for Future Developments. *International Journal of Microsimulation* **2022**, *15*, 44 – 61. <https://doi.org/10.34196/ijm.00262>.
8. Moosbrugger, N.; Hellwig, M.; Dobler, M.; Urban, S. Sustainable Smart Services for Financial Institutions: Evaluation of ESG Service Creation Framework Applicability for Practitioners. *Progress in IS* **2024**, Part F3229, 15 – 27. https://doi.org/10.1007/978-3-031-60313-6_2.
9. Liermann, V.; Dittmar, H. *Bsds: Balance sheet dynamics simulator*; 2021; p. 135 – 153. https://doi.org/10.1007/978-3-030-78829-2_8.

10. Ridwan, M.; Alghifari, E.S. Evaluating the impact of ESG on financial risk: the moderating effects of operational ability and profitability in Indonesian infrastructure firms. *Journal of Accounting and Organizational Change* **2025**. <https://doi.org/10.1108/JAOC-08-2024-0281>.
11. Shang, K.; Hossen, Z. Liability-Driven Investment Benchmark Model **2019**. Aging and Retirement Research.
12. Azamuke, D.; Katarahweire, M.; Bainomugisha, E. MoMTSim: A Multi-Agent-Based Simulation Platform Calibrated for Mobile Money Transactions. *IEEE Access* **2024**, *12*, 120226 – 120238. <https://doi.org/10.1109/ACCESS.2024.3439012>.
13. Budiatmodjo, E.; Rumiati, A.T.; Prastyo, D.D. Benchmarking Hierarchical Bayesian Small Area Estimators in the Percentage of Poverty at Sub-districts Level in Central Java. **2023**, Vol. 2540. <https://doi.org/10.1063/5.0106437>.
14. Leung, A.T.L.; Guo, C.; Luk, W. FABS: An Extensible and High-Performance Digital Twin Framework of AI-Driven Financial Systems. **2025**, p. 27 – 34. <https://doi.org/10.1145/3768292.3770369>.
15. Youness, J.; Driss, M. Mathematical Modeling of Financial Time Series Volatility: A GARCH Model. *Lecture Notes in Networks and Systems* **2023**, *635 LNNS*, 608 – 614. https://doi.org/10.1007/978-3-031-26254-8_88.
16. Gurgone, A.; Iori, G. A Multi-agent Methodology to Assess the Effectiveness of Systemic Risk-Adjusted Capital Requirements. *Dynamic Modeling and Econometrics in Economics and Finance* **2021**, *26*, 177 – 204. https://doi.org/10.1007/978-3-030-52970-3_8.
17. Nastiti, N.D.; Kasri, R.A. The role of banking regulation in the development of Islamic banking financing in Indonesia. *International Journal of Islamic and Middle Eastern Finance and Management* **2019**, *12*, 643 – 662. <https://doi.org/10.1108/IMEFM-10-2018-0365>.
18. Yudaruddin, R.; Che Yahya, N.; Mohd Rashid, S.N. How ESG risk influences bank performance: insights from Islamic banking sector in Indonesia. *Journal of Islamic Accounting and Business Research* **2025**. <https://doi.org/10.1108/JIABR-09-2024-0343>.
19. Miani, S.; Floreani, J.; Paltrinieri, A. Do Capital Adequacy and Credit Quality Affect Systematic Risk? Investigation of a Sample of European Listed Banks in Light of EBA Stress Tests. *Quarterly Journal of Finance* **2018**, *8*. <https://doi.org/10.1142/S2010139218400062>.
20. Patra, B.; Padhi, P. Resilience of Indian banks: Macroeconomic stress test modeling for credit risk. *Journal of Public Affairs* **2022**, *22*. <https://doi.org/10.1002/pa.2350>.
21. Antar, M.; Hassan, R.; Barka, D. From Black Box to Clarity: Machine Learning and Agnostic Techniques in Credit Risk Management. *Journal of Corporate Accounting and Finance* **2025**. <https://doi.org/10.1002/jcaf.70020>.
22. Wirjanto, F.A.; Ferdinand, F.V.; Haryani, C.A. Recursive Clustering and SHAP-Based Interpretability for Rural Bank Credit Prediction in Indonesian Provinces. **2025**. <https://doi.org/10.1109/ECTI-CON64996.2025.11100432>.
23. Fabretti, A. Markov chain analysis in agent-based model calibration by classical and simulated minimum distance. *Knowledge and Information Systems* **2019**, *61*, 259 – 276. <https://doi.org/10.1007/s10115-018-1258-y>.
24. A Note on the Validity of Cross-Validation for Evaluating Autoregressive Time Series Prediction.
25. Soomro, A.; Zakariyah, H.; Aftab, S.; Muflehi, M.; Shah, A.; Meraj, S. Loan Default Prediction Using Machine Learning Algorithms: A Systematic Literature Review 2020–2023. *Pakistan Journal of Life and Social Sciences* **2024**, *22*, 6234 – 6253. <https://doi.org/10.57239/PJLSS-2024-22.2.00469>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.