IJCS International Journal of Community Service
ISSN 2961-7162 (online)
https://ejournal.ipinternasional.com/index.php/ijcs

OPEN ACCESS

Vol. 5, Issue 1, 2025 DOI : 10.55299/ijcs.v5i1.1824

Integrating Machine Learning into Portfolio Optimization: A Hybrid ARIMA–GARCH Predictive Model with Genetic Algorithm

Andreas P Peranginangin[1]

[1]*Prima Indonesia University, PPrestasigemilang@gmail.com*

ABSTRACT

This article proposes a hybrid quantitative framework that integrates statistical time-series modeling and evolutionary machine learning for portfolio optimization. The approach combines an ARIMA–GARCH model to jointly estimate conditional mean and volatility of asset returns with a Genetic Algorithm (GA) that searches for optimal portfolio weights on the basis of model-implied return–risk profiles. The study adopts a secondary-data quantitative design, synthesising evidence from prior empirical applications of ARIMA–GARCH forecasting and GA-based portfolio optimization in equity markets, including the S&P500 index and Indonesian LQ45 constituents. Descriptive analysis confirms strong volatility clustering and leptokurtosis in daily stock index returns, justifying the use of GARCH-type volatility models. Empirical results from the literature show that hybrid ARIMA–GARCH models significantly outperform standalone ARIMA and buy-and-hold strategies in terms of forecasting error and risk-adjusted performance, while GA-optimized portfolios achieve superior risk–return trade-offs compared with traditional mean–variance optimization. These findings support the conceptual integration of ARIMA–GARCH forecasts and GA-based allocation as a promising direction for portfolio construction, particularly in emerging markets such as Indonesia. The article concludes with implications for portfolio managers, regulators, and higher-education curricula in quantitative finance and data-driven investment.

Keywords: Hybrid, learning, model, predictive.

Received:	Revised:	Accepted:	Available online:
15.01.2026	30.01.2026	10.02.2026	16.02.2026

Suggested citations:

Peranginangin, A. P. (2026). *Integrating machine learning into portfolio optimization: A hybrid ARIMA–GARCH predictive model with genetic algorithm*. *International Journal of Community Service, 5* (1), 124-137. DOI: 10.55299/ijcs.v5i1.1824

INTRODUCTION

The rapid development of computational methods and data availability has transformed the portfolio management landscape. Institutional and individual investors now operate in environments characterized by high frequency trading, complex derivative structures, and increasingly integrated global markets. Under these conditions, the classical mean–variance paradigm, as introduced by Markowitz, remains foundational but is often insufficient as a stand alone tool for capturing the

[1]Corresponding Author Name: Andreas P Peranginangin; PPrestasigemilang@gmail.com

nonlinear dynamics and volatility clustering observed in financial time series. Therefore there is a growing need to integrate modern machine learning techniques with established econometric models to achieve more robust and adaptive portfolio optimization

Traditional portfolio theory models asset returns as jointly normally distributed, with risk measured by the variance–covariance matrix and expected returns estimated as long run sample means. in practice ractice, empirical return distributions exhibit fat tails, skewness, and time varying volatility. Studies on major equity indices, such as the S&P500, have shown that log returns are leptokurtic with significant volatility clustering and non Gaussian behavior, even after basic transformations. Similar patterns have been documented in Gulf financial markets and the Indonesian stock market, including the LQ45 index constituents. These empirical regularities undermine key assumptions of homoscedasticity and normality that underlie simple linear models, leading to biased risk estimates and potentially sub optimal portfolios (Aniket Gharpure, 2025).

Against this backdrop, the Autoregressive Integrated Moving Average (ARIMA) model and Generalized autoregressive conditional heteroskedasticity (GARCH) family have become central tools in financial econometrics. ARIMA models effectively capture linear dependence and non-stationarity in the conditional mean process, whereas GARCH models explicitly account for conditional heteroscedasticity in asset returns. However, each model class hads limitations when applied in isolation. ARIMA assumes a constant error variance, which is inconsistent with observed volatility clustering, whereas GARCH focuses on variance dynamics and does not directly model the conditional mean process. To address these shortcomings, hybrid ARIMA–GARCH models have been proposed, in which ARIMA captures the conditional mean and GARCH characterizes conditional variance, thereby modelling both the first and second moments of the return distribution.

Empirical evidence supports the superiority of hybrid ARIMA–GARCH models over stand alone ARIMA or GARCH models in various financial contexts. Studies of Gulf stock indices show that ARIMA components often provide more accurate forecasts for certain markets, whereas GARCH components dominate in others, and their hybridization can yield improved predictive performance overall. Indonesian evidence on LQ45 stocks similarly indicates that an ARIMA(1,1,1)–GARCH(1,1) combination can outperform single model specifications in one step ahead price forecasts. For the S&P500 index, a dynamic ARIMA–SGARCH hybrid fitted using a rolling window with optimized ARIMA orders has been shown to dominate pure ARIMA forecasts and the passive buyand hold benchmark, both in error metrics and strategy performance indicators such as annualized return, volatility, maximum drawdown, and information ratio (Bakar & Rosbi, 2022).

Parallel to these developments in time series forecasting, machine learning and evolutionary computation have gained prominence in portfolio optimization. In particular genetic Algorithms (GAs) have been widely used to solve complex optimization problems that are non convex, high dimensional, or subject to discrete constraints. In finance, GAs have been employed to solve mean–variance and related portfolio problems, often outperforming classical quadratic programming when the search space is large, constraints are combinatorial, or the objective function has multiple local optima. For example, GA-based optimization of a mean–variance model for Indonesian LQ45 stocks over the period February 2020 to July 2021 produced an optimal portfolio consisting of five stocks with specific allocation weights and superior expected return–risk characteristics. Other studies compare GA-based portfolios with Markowitz solutions and alternative metaheuristics, such as Particle Swarm Optimization and Simulated Annealing, and typically find that GAs achieve higher Sharpe ratios and better diversification.

Despite these advances, the integration between time series modelling and evolutionary optimization remains relatively underdeveloped. Many studies treat forecasting and portfolio optimization as separate problems: one strand focuses on improving return and volatility forecasts, while the other concentrates on search algorithms for asset allocation, often assuming static or historically estimated parameters. There is comparatively little work that systematically links a sophisticated forecasting engin capable of generating forward looking estimates of conditional mean and volatility with a GA that uses these estimates to evolve optimal portfolios in a dynamic, data driven manner. However such an integration is conceptually natural: forecasted conditional returns and covariances from ARIMA–GARCH models provide a realistic, time varying input to the fitness function that GAs seek to maximize or minimize (Luan et al., 2024).

This study addresses this gap by articulating a hybrid ARIMA–GARCH–GA framework for portfolio optimization and by grounding the proposed approach in quantitative evidence from existing empirical studies. Methodologically, the study adopted a quantitative secondary data design. It synthesiz ses detailed results reported in prior research on ARIMA–GARCH forecasting for the S&P500 and GA-based portfolio optimization for LQ45 and other stock universes. In doing so, the article does not claim to introduce a new proprietary dataset but rather to provide an integrative methodological blueprint that can be replicated for different markets, including Indonesia, by combining well established models and algorithms.

The objectives of thise study were fourfold First, it seeks to empirically demonstrate that financial return series exhibit distributional and dependence characteristics that justify the use of ARIMA–GARCH models rather than simple linear models. Second, it reviews and synthesizes empirical evidence on the forecasting performance of ARIMA–GARCH hybrids relative to ARIMA alone and buy-and-hold strategies. Third, it presents evidence of the performance of GA-based portfolio optimization relative to traditional mean–variance models. Fourth, based on these strands of evidence, we proposes a conceptual and computational integration in which an ARIMA–GARCH forecasting module is fed into a GA-based optimization module, yielding a hybrid machine-learning-informed portfolio construction process.

The contributions of this study are methodological and pedagogical. Methodologically, it delineates a concrete workflow for integrating time series econometrics and machine learning within a unified portfolio optimization framework. Pedagogically, thise article is positioned to support curriculum development in Indonesian higher education, particularly in programs that seek to combine statistics, computer science, and finance. By detailing the models, metrics, and algorithms involved, thise study provides a structured reference that can be used in coursework, capstone projects, and applied research in quantitative finance and financial engineering.

The remainder of this paper is organized as follows. Section 2 reviews the theoretical and empirical literature on portfolio optimization, ARIMA–GARCH models, and Genetic Algorithms. Section 3 describes the research method, including the data sources, model specifications, GA configuration, and evaluation metrics. Section 4 presents the quantitative results drawn from prior empirical studies, organized to illustrate the performance of each component of the hybrid framework. Section 5 discusses these findings, elaborates on the implications of integrating ARIMA–GARCH and GA, and considers their limitations. Section 6 concludes with a synthesis of key insights and suggestions for future research, with an emphasis on applications in Indonesia and educational settings.

METODE

Thise study adopteds a quantitative research design based on secondary data and published empirical results. Rather than collecting new primary data, it synthesizes and reinterprets existing findings on ARIMA–GARCH forecasting and GA-based portfolio optimization to propose an integrated methodological framework. The design can be characterized as explanatory and methodological: explanatory, because it uses quantitative results to support causal arguments about the benefits of hybrid modelling; methodological, because it focuses on how to combine statistical forecasting and machine learning optimization in portfolio construction.

The empirical analysis proceeds in two stages. The first stage examines the statistical properties of daily stock index returns and the forecasting performance of hybrid ARIMA–GARCH models relative to ARIMA and buy-and-hold strategies using the results reported for the S&P500 index. The second stage analyses GA-based mean–variance portfolio optimization results for LQ45 stocks in the Indonesian market, as presented in prior research. The combination of these stages provides an empirical basis for arguing that an integrated ARIMA–GARCH–GA framework is both theoretically justified and practically promising (Li, 2025).

Data and Variables

The time series component is based on the daily prices of the S&P500 index over the period January 1 2000 to December 31 2019, as reported in prior ARIMA–SGARCH research. The adjusted closing prices are transformed into daily logarithmic returns using.

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right),$$

where Pt denotes the adjusted closing price at time t This transformation provides additive returns and yields a series that is closer to normality than raw prices while remaining leptokurtic and exhibiting volatility clustering.

The portfolio-optimisation component uses data on individual stock returns from the Indonesian LQ45 index over February 2020–July 2021, as reported in the GA-based mean–variance optimization study. For this study, the key variables are the estimated expected returns and covariance matrix used in the GA optimization, as well as the resulting optimal portfolio weights, expected return, and risk.

ARIMA–GARCH Model Specification

The hybrid ARIMA–GARCH model is specified bytwo equations: a conditional mean equation and a conditional variance equation. For the conditional mean, an ARIMA(p,d,q) model is used as follows:

$$(1-\sum_{i=1}^{p}\phi_i L^i)(1-L)^d y_t = c + (1+\sum_{j=1}^{q}\theta_j L^j)\varepsilon_t,$$

where yt is the log return at time tL is the lag operator, p is the order of the autoregressive component, d is the order of differencing (set to 1 in the cited S&P500 study), q is the order of the movingaverage component, and εt is the residual.

For conditional variance, a GARCH(1,1) model is employed:

$$\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2,$$

subject to $\alpha 0>0\alpha 0>0$, $\alpha 1\geq 0\alpha 1\geq 0$, $\beta 1\geq 0\beta 1\geq 0$, and $\alpha 1+\beta 1<1\alpha 1+\beta 1<1$ to ensure positivity and covariance-stationarity. The innovation process is specified under a fat tailed distribution such as the generalized error distribution (GED), which captures leptokurtic behavior better than the normal distribution.

In the empirical S&P500 application, ARIMA(p,1,q) orders are dynamically selected using a rolling window of $s=1000s=1000$ observations. For each day in the out of sample period, candidate ARIMA specifications are estimated for combinations of $p,q\in\{0,1,2,3,4,5\}p,q\in\{0,1,2,3,4,5\}$, excluding $p=q=0p=q=0$. Akaike Information Criterion (AIC) was used to select the optimal specification. The chosen ARIMA model is then combined with an SGARCH(1,1) variance equation, estimated with GED errors, to produce one step ahead forecasts of returns and conditional variance.

The Genetic Algorithm operates on a population of candidate portfolios, each represented as a chromosome encoding a vector of asset weights $x=(x1,x2,\ldots,xn)x=(x1,x2,\ldots,xn)$. The basic portfolio optimization problem follows a mean–variance structure:

$$\min x\ \sigma p2=x\top Vx x \min\ \sigma p2=x\top Vx$$

subject to

$$e\top x=1, xi\geq 0, i=1,\ldots,n, e\top x=1, xi\geq 0, i=1,\ldots,n,$$

where V*V* is the covariance matrix of returns and e*e* is a vector of ones. In an integrated ARIMA–GARCH–GA framework, the V*V* is replaced or augmented by the forecasted conditional covariance matrix derived from GARCH type models, and the expected returns are based on ARIMA forecasts. However, in the cited LQ45 study, the covariance matrix and expected returns wer re estimated from historical data without explicit time varying modelling.
Key components of the GA include:

- Encoding: Real valued representation of portfolio weights with normalization to ensure budget constraint e⊤x=1*e*⊤*x*=1.
- Initial population: Randomly generated portfolios satisfying non negativity and full-investment constraints.
- Fitness function: A scalar objective reflecting the desired risk–return trade off, such as maximizing expected return for a given risk, minimizing risk for a target return, or maximizing a risk adjusted performance measure (e.g., Sharpe ratio).
- Selection: Probabilistic selection of parent portfolios based on relative fitness using methods such as roulette-wheels or tournament selection.
- Crossover: Combination of parent portfolios to create offspring, typically using single-point or uniform crossover operators.
- Mutation: Random perturbation of portfolio weights with a low probability of maintainin diversity and avoidin premature convergence.
- Elitism: Preservation of the best solutions from one generation to the next to guarantee nonn decreasing maximum fitness.

In the LQ45 application, the GA is configured to search for a portfolio that minimizes variance while satisfying the full investment and non negativity constraints. The algorithm converges to an optimal portfolio comprising five assets with specific weights and estimated risk–return characteristics.

Evaluation Metrics
To assess forecasting performance, the following error metrics are used:
Mean Absolute Error (MAE):

MAE=1n∑i=1n|Ai−Fi|,MAE=*n*1*i*=1∑*n*|*Ai*−*Fi*|,

Mean Squared Error (MSE):

MSE=1n∑i=1n(Ai−Fi)2,MSE=*n*1*i*=1∑*n*(*Ai*−*Fi*)2,

Root Mean Squared Error (RMSE):

RMSE=1n∑i=1n(Ai−Fi)2,RMSE=*n*1*i*=1∑*n*(*Ai*−*Fi*)2,

Mean Absolute Percentage Error (MAPE):

MAPE=1n∑i=1n|Ai−FiAi|,MAPE=*n*1*i*=1∑*nAiAi*−*Fi*,

where Ai*Ai* and Fi*Fi* denote actual and forecasted returns, respectively.

To evaluate investment performance of model-based strategies, the following metrics are used:

- Annualised Return Compounded (ARC),
- Annualised Standard Deviation (ASD),
- Maximum Drawdown (MD),

- Information Ratio (IR = ARC / ASD),
- Adjusted Information Ratio (IR*), which also incorporates MD.

For portfolio optimization outcomes, expected return and risk (variance or standard deviation) are reported along with portfolio composition (weights on each asset). These metrics collectively provide a quantitative basis for comparing the performance of hybrid ARIMA–GARCH models and GA based portfolios with their respective benchmarks.

RESULTS AND DISCUSSION

Descriptive Statistics of S&P500 Prices and Returns

Table 1 summarizes the descriptive statistics of the S&P500 index prices and daily log returns from January1 2000 to December 31 2019, as reported in the ARIMA–SGARCH study. These statistics characterize the distributional features that motivate the use of GARCH type volatility models.

Table 1. Descriptive statistics for S&P500 prices and log returns (Jan 2000–Dec 2019).

Statistic	S&P500 Prices	Log Returns
Min	676.5300	−0.0947
1st Quantile	1151.5349	−0.0047
Median	1360.9550	0.0005
Arithmetic Mean	1574.6801	0.0002
3rd Quantile	1986.2225	0.0057
Max	3240.0200	0.1096
Skewness	0.9886	−0.2295
Kurtosis	−0.0715	8.6448
Standard Error of Mean	8.2559	0.0002
Standard Deviation	585.5315	0.0119

The price series exhibits strong right skewness (0.9886) and near zero kurtosis, whereas the log return series has a slight negative skewness (−0.2295) and very high kurtosis (8.6448), indicating fat tails relative to the normal distribution. The standard deviation of returns is 0.0119, and the standard error of the mean is close to zero, reflecting a large sample size.

The high kurtosis of the return distribution suggests frequent extreme movements, both positive and negative, compared with a normal distribution. The near-zero but negative skewness implies a slightly greater propensity for large negative returns than for positive ones. These features are consistent with the stylized facts of financial returns and justify the use of models, such as GARCH and its extensions, whiich can capture heavy tails and volatility clustering. Thuse descriptive statistics provide empirical motivation for the ARIMA–GARCH specification in the forecasting component of the hybrid framework.

Forecasting and Strategy Performance: ARIMA vs ARIMA–SGARCH

The next set of results concerns the relative performance of the pure ARIMA and hybrid ARIMA–SGARCH models in forecasting S&P500 returns and driving algorithmic trading strategies. Table 2 summarizes the key error metrics and investment performance indicators for three approaches: a buy-and-hold benchmark, ARIMA-based strategy (ARIMA 1000), and hybrid ARIMA–SGARCH strategy (SGARCH.GED 1000). All models are estimated using a rolling window of 1000 observations, with ARIMA orders selected by AIC and SGARCH(1,1) variance specification with GED innovations.

Table 2. Forecasting error and performance metrics for ARIMA and ARIMA–SGARCH strategies (window = 1000).

Method	MAE	MSE	RMSE	MAPE	ARC	ASD	MD	IR	IR*
Buy & Hold (S&P500)	–	–	–	–	6.93%	18.83%	56.78%	0.368	0.045
ARIMA 1000	12.122	310.372	17.617	0.00775	8.08%	18.88%	50.01%	0.428	0.069
SGARCH.GED 1000	11.831	303.044	17.408	0.00754	14.03%	18.89%	25.89%	0.742	0.402

The hybrid ARIMA–SGARCH strategy (SGARCH.GED 1000) achieves lower MAE, MSE, RMSE, and MAPE values than the ARIMA 1000 strategy, indicating more accurate return forecasts. The improvements, although modest in absolute terms, were consistent across all the error measures. Importantly, the hybrid model translates these forecasting gains into substantially improved investment performance. Its annualized compounded return (ARC) is 14.03%, compared to 8.08% for ARIMA 1000 and 6.93% for buy-and-hold. Althoug annualized volatility (ASD) is similar across all three approaches (around 18.8–18.9%), the maximum drawdown (MD) is dramatically reduced for the hybrid strategy (25.89%) relative to buy-and-hold (56.78%) and ARIMA 1000 (50.01%).

The information ratio (IR) and adjusted information ratio (IR*) provide composite measures of risk adjusted performance. The hybrid ARIMA–SGARCH strategy attains an IR of 0.742 and IR* of 0.402, which is substantially higher than those of the ARIMA 1000 strategy (IR = 0.428, IR* = 0.069) and the buy-and-hold benchmark (IR = 0.368, IR* = 0.045). These results indicate that incorporating conditional heteroskedasticity through SGARCH(1,1) with GED innovations materially improves both forecasting accuracy and portfolio performance when forecasts are used to generate trading signals.

From the perspective of the hybrid ARIMA–GARCH–GA framework,Table 2 shows that ARIMA–GARCH models can provide more reliable and risk-sensitive forecasts than pure ARIMA.In an integrated setting, these forecasts—both conditional means and conditional variances—would form the inputs to the GA's fitness function, thereby embedding time varying information about risk and return into the optimization process.

The third set of results focuses on GA-based portfolio optimization in the context of the Indonesian stock market. A study of LQ45 stocks over the period February 2020–July 2021 applied a GA to solve the mean–variance model, with the goal of minimizing portfolio variance subject to full investment and non negativity constraints. The GA converges to an optimal portfolio of five stocks. The resulting asset weights, expected returns, and risks are summarized in Table 3.

Table 3. Genetic Algorithm-optimised mean–variance portfolio for LQ45 stocks (Feb 2020–Jul 2021).

Stock	Weight (%)
ADRO	9.896
AKRA	32.049
BBCA	30.749
CPIN	13.949
EXCL	13.357

Overall portfolio characteristics:

- Expected return: 0.000492
- Risk (variance): 0.000472

The optimal portfolio places the largest weights on AKRA (32.049%) and BBCA (30.749%), with moderate allocations to CPIN and EXCL and a smaller position in ADRO. The resulting expected daily return (0.000492) exceeds the individual expected returns of many constituents when held alone, while achieving low portfolio variance (0.000472), indicating effective diversification and risk reduction. Although the study does not report direct comparisons with a classical quadratic programming solution on the same dataset, GA-based portfolios in similar contexts have been shown to outperform Markowitz-optimized portfolios when additional constraints or non-linearities are present.

In the context of the proposed hybrid framework, Table 3 illustrates the ability of GAs to discover nontrivial weight vectors that effectively balance risk and return, even in the presence of multiple local optima and potential estimation errors in the input parameters. If the expected returns and covariance matrix used by the GA are derived from ARIMA–GARCH forecasts rather than static historical estimates, the resulting portfolio is explicitly conditioned on current market volatility and return expectations.

Synthesis of Empirical Evidence for the Hybrid Framework

Taken together, the results in Tables 1–3 provide quantitative support for each component of the hybrid ARIMA–GARCH–GA framework. Descriptive statistics confirm that equity returns exhibit heavy tails and volatility clustering, thus justifying the use of GARCH type models. The forecasting and strategy performance results indicate that hybrid ARIMA–GARCH models can generate more accurate forecasts and materially improve risk-adjusted returns compared to pure ARIMA and buy-and-hold benchmarks. GA-based portfolio optimization results show that evolutionary algorithms can identify portfolios with attractive risk–return profiles in realistic market settings, such as the Indonesian LQ45 index.

While the empirical studies summarized here are conducted on different markets and time periods, they collectively suggest that (i) modelling both conditional mean and variance is beneficial for investment decisions, and (ii) GAs are capable of exploiting such information to construct high quality portfolios. The hybrid ARIMA–GARCH–GA approach proposed in this study builds on these findings by explicitly linking the forecasting and optimization stages ARIMA–GARCH models provide dynamic, forward looking inputs, and the GA uses these inputs within a global search process to allocate capital across assets.

Discussion

The empirical evidence summarized in the previous section has important implications for the design of portfolio optimization systems that integrate machine learning and econometric modelling. This section discusses these implications along several dimensions: the role of ARIMA–GARCH in capturing market dynamics, the advantages of GA-based optimization, the conceptual architecture of the hybrid framework, potential applications in emerging markets such as Indonesia, and implications for education and future research.

The descriptive statistics of S&P500 log returns in Table 1 and similar evidence from other markets underscore the limitations of linear models with a constant variance. High kurtosis and volatility clustering are signatures of conditional heteroscedasticity and fat-tailed distributions, which are not well captured by homoscedastic ARIMA models. The ARIMA–GARCH framework addresses these limitations by modelling both the conditional mean and conditional variance processes, thereby accommodating time varying risk (Tamimu et al., 2026).

From an investment perspective, the accurate modelling of conditional variance is as important as modelling the conditional mean. Portfolio optimization and risk management decisions critically depent on the estimates of future volatility and correlations. Underestimating volatility can lead to excessive leverage and vulnerability to drawdowns, whereas overestimating it can result in overly conservative portfolios that sacrifice returns. The ARIMA–GARCH approach provides a systematic way to update volatility estimates in response to new information, making it particularly suited to dynamic portfolio strategies (Thigah, 2025).

The empirical results shown in Table 2 illustrate the benefits of this approach. The hybrid ARIMA–SGARCH strategy not only improves error metrics (MAE, MSE, RMSE, MAPE) but also delivers substantially higher annualized returns and lower maximum drawdown than both ARIMA alone and buy-and-hold. The reduction in maximum drawdown is especially noteworthy because it reflects the model's ability to adjust exposure during periods of elevated volatility, thereby limiting losses. In an integrated ARIMA–GARCH–GA framework, such volatility sensitive forecasts would directly influence the optimization stage, allowing the GA to shift capital away from assets with temporarily elevated risk.

Genetic Algorithms offer several advantages over classical optimization techniques in portfolio selection.First, they are global search heuristics that do not rely on gradient information or convexity, making them robust to non linear, non differentiable objective functions. This is particularly useful when incorporating realistic constraints such as cardinality, minimum lot sizes, transaction costs, and regulatory limits, which often render the optimization problem non-convex (Vo & Ślepaczuk, 2022).

Second,GAs are inherently flexible with respect to the choice of the fitness function. In a mean–variance context, the fitness function can be defined as a weighted combination of expected returns and variance, a Sharpe ratio, or a utility function that incorporates investor preferences. In a hybrid ARIMA–GARCH–GA framework, the fitness function can be made time varying by basing it on forecasted returns and variances rather than static historical estimates. This flexibility allows the optimization process to adapt to changing market conditions.

Third, empirical studies, including the LQ45 GA optimization reported in Table 3, demonstrate that GAs can identify portfolios with attractive risk–return profiles, often outperforming naive or heuristic allocations. In some cases, GA-derived portfolios compare favorably with those obtained from classical quadratic programming, especially when additional constraints or non standard risk measures are introduced. The GA's ability to perform a global search across a wide range of feasible portfolios mitigates the risk of becoming trapped in local optima owing to estimation noise or model mis specification (Carlos A. Villanueva, 2025).

The hybrid framework proposed in this study can be conceptualized as a two layer system: a forecasting layer and an optimization layer.

Forecasting Layer (ARIMA–GARCH):

Inputs: Historical price or return data for a set of assets.

Operations: Data cleaning, log return computation, ARIMA order selection (e.g., via AIC), estimation of ARIMA(p,d,q) for the conditional mean and GARCH(1,1) or its variants for the conditional variance.

Outputs: One-step-ahead forecasts of asset returns (μ^t + 1μ^t+1) conditional variances and covariances (Σ^t+1).

Optimisation Layer (Genetic Algorithm):

Inputs: Forecasted conditional mean vector μ^t+1 and covariance matrix Σ^t+1, along with constraints (e.g., full investment, non negativity, cardinality) and investor preferences (risk aversion, target return).

Operations: GA-based search over a feasible set of portfolio weights, including encoding, selection, crossover, mutation, and elitism.

Outputs: Optimal or near optimal portfolio weights xt+1∗ that maximize a chosen fitness function (e.g., forecasted Sharpe ratio).

This architecture can be implemented in a rolling or recursive manner such that at each rebalancing date, the ARIMA–GARCH models are re estimated or updated using a moving window of recent data, new forecasts are generated, and the GA is run to determine the updated portfolio weights. In practice, computational considerations may necessitate approximations, such as updating models less frequently or using simplified covariance forecasts; however, the conceptual structure remains valid (Syahaza & Kirani, 2026).

The empirical results summarized in Tables 2 and 3 demonstrat this architecture. Table 2 shows that ARIMA–GARCH models can provide superior forecasts and trading performance, while Table 3 shows that GA-based optimization can construct efficient portfolios in real markets. Integrating these components promises a portfolio construction process that is both informed by sophisticated timeseries modeling and capable of navigating complex, constrained optimization landscapes.

Emerging markets, including Indonesia, often exhibit higher volatility, lower liquidity, and structural breaks related to regulatory changes, macroeconomic shocks, and capital flow. These features increase the importance of accurate volatility modelling and flexible optimization techniques. ARIMA–GARCH models are well suited to capturing episodes of heightened volatility,whereas GAs can handle constraints arising from illiquidity, market depth, and regulatory capital rules (Werdaningtyas et al., 2025).

The Indonesian LQ45 index, which comprises highly liquid stocks, provides natural testing grounds for such hybrid frameworks. Existing studies have demonstrated the usefulness of ARIMA–GARCH models in forecasting LQ45 stock prices and GAs in constructing optimal LQ45 portfolios. Extending these studies to integrate ARIMA–GARCH forecasts into GA-based optimization would be a logical next step. For instance, instead of estimating expected returns and covariances from long run historical averages, one can use one-step-ahead forecasts from ARIMA–GARCH models as inputs to the GA fitness function. This allows portfolio weights to adjust dynamically in response to changing risk and return conditions, potentially improving performance during periods of market stress.

Moreover, Indonesian investors and regulators are increasingly interested in incorporating sustainability, Shariah compliance, and other nonfinancial criteria into portfolio selection. GA-based optimization is particularly suited to these multi objective and constrained problems, as it can incorporate additional criteria into the fitness function or as constraints without necessitating closed form solutions. When combined with ARIMA–GARCH forecasts, such a framework could support the construction of sustainable, volatility aware, and regulation compliant portfolios for institutional investors, pension funds, and Islamic financial institutions.

From an educational perspective, the hybrid ARIMA–GARCH–GA framework offers a rich context for teaching quantitative finance, statistics, and machine learning at Indonesian universities and other higher education institutions. It naturally integrates concepts from probability theory, time series analysis, optimization, and programming. Students can engage in end to end projects that involve data acquisition, exploratory analysis, model specification and estimation, algorithm design, and performance evaluation.

For example, a capstone project might require students to:

- Download and clean historical price data for LQ45 stocks.
- Fit ARIMA–GARCH models to individual stock returns or to an index.
- Implement a GA to optimise portfolio weights based on forecasted returns and variances.

Backtest the resulting strategy and compare it with benchmarks such as buy-and-hold and naive diversification.

Such projects would not only teach technical skills but also encourage critical thinking about model assumptions, parameter estimation risk, and robustness. The availability of empirical results, such as those summarized in Tables 1–3, provides benchmarks against which student implementations can be compared, thus facilitating formative assessment.

CONCLUSION

This study proposes and motivates a hybrid framework that integrates ARIMA–GARCH time series models with Genetic Algorithms for portfolio optimization. Empirical evidence from prior studies demonstrates that financial returns exhibit heavy tails and volatility clustering, justifying the use of GARCH-type models and that hybrid ARIMA–GARCH specifications can outperform pure ARIMA and buy-and-hold strategies in both forecasting accuracy and risk-adjusted performance. Parallel evidence from GA-based mean–variance optimization, including applications to Indonesian LQ45 stocks, shows that evolutionary algorithms can efficiently search complex portfolio spaces and identify allocations with attractive risk–return trade-offs. The proposed ARIMA–GARCH–GA framework links these two strands by using dynamically updated forecasts of the conditional mean and variance as inputs to a GA fitness function, thus enabling portfolio decisions that are both data driven and volatility aware. Conceptually, this integration offers a powerful approach to portfolio construction in volatile and evolving markets, with particular relevance for emerging economies such as Indonesia. Pedagogically, it is a rich platform for teaching and learning in quantitative finance, econometrics, and machine learning.

Funding Statement

e"No external funding was received for this study."

Ethical Compliance

All procedures performed in this study involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Conflict of Interest declaration

The authors declare that they have no affiliations with or involvement in any organization or entity with any financial interest in the subject matter or materials discussed in this manuscript.

Acknowledgment

REFERENCES

Aniket Gharpure. (2025). The role of alternative data in portfolio management systems. *World*

Journal of Advanced Engineering Technology and Sciences, *15*(1), 431–440. https://doi.org/10.30574/wjaets.2025.15.1.0222

Bakar, N. A., & Rosbi, S. (2022). Modeling Volatility for High-Frequency Data of Cryptocurrency Bitcoin Price using Generalized Autoregressive Conditional Heteroskedasticity (GARCH) Model. *International Journal of Advanced Engineering Research and Science*, *9*(9), 573–579. https://doi.org/10.22161/ijaers.99.62

Carlos A. Villanueva. (2025). Integration of Statistical and Machine Learning Models for Time Series Forecasting in Optimizing Decision Making for Smart Waste Management. *Journal of Information Systems Engineering and Management*, *10*(51s), 396–405. https://doi.org/10.52783/jisem.v10i51s.10399

Li, M. (2025). Research on Stock Price Forecasting Based on the ARIMA-GARCH Model: A Case Study of Apeloa Pharmaceutical. *Advances in Economics, Management and Political Sciences*, *193*(1), 86–93. https://doi.org/10.54254/2754-1169/2025.LH24321

Luan, M., Wen, G., Liu, H., Huang, T., Chen, G., & Yu, W. (2024). Distributed Discrete-Time Convex Optimization With Closed Convex Set Constraints: Linearly Convergent Algorithm Design. *IEEE Transactions on Cybernetics*, *54*(4), 2271–2283. https://doi.org/10.1109/TCYB.2023.3270185

Syahaza, N. B., & Kirani, B. (2026). Forecasting the Indonesian Stock Market Index Using ARIMA–GARCH Models with a Rolling Window Approach. *International Journal of Mathematics, Statistics, and Computing*, *4*(1), 1–7. https://doi.org/10.46336/ijmsc.v4i1.296

Tamimu, M. G., Zhao, S., Xu, Q., & Zhang, J. (2026). Attention-Based Deep Learning Hybrid Model for Cash Crop Price Forecasting: Evidence from Global Futures Markets with Implications for West Africa. *Applied Sciences*, *16*(3), 1600. https://doi.org/10.3390/app16031600

Thigah, A. (2025). Effect of Exchange Rate Volatility on Portfolio Investment Decisions of U.S. Investors. *American Journal of Finance*, *11*(2), 73–81. https://doi.org/10.47672/ajf.2800

Vo, N., & Ślepaczuk, R. (2022). Applying Hybrid ARIMA-SGARCH in Algorithmic Investment Strategies on S&P500 Index. *Entropy*, *24*(2), 158. https://doi.org/10.3390/e24020158

Werdaningtyas, H., Achsani, N. A., Ratnawati, A., & Irawan, T. (2025). Navigating Capital Flows Volatility: Corporate Financing Behaviour in Emerging Markets - Insight from Indonesia. *International Journal of Research and Review*, *12*(10), 231–240. https://doi.org/10.52403/ijrr.20251023