BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 1, December 2025 | Hal 502-510
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i1.884

Hybrid EGARCH-LSTM for Price Forecasting and Risk Estimation of Daily USD/IDR (2015–2025)

Alfaiz Arifin Setia Budi*, Siti Salsabila Maryanto, Muhammad Althafino, Salsabila Sekar Nadia, Zulvania Armiana, Shaifudin Zuhdi

Faculty of Information Technology and Data Science, Data Science Study Program, Universitas Sebelas Maret, Surakarta, Indonesia
Email: [1,*]alfaiz.asb1204@student.uns.ac.id, [2]sitisalsamar.03@student.uns.ac.id, [3]malthafino36@student.uns.ac.id, [4]salsabilasnad@student.uns.ac.id, [5]zulvaniaarmiana@student.uns.ac.id, [6]szuhdi@staff.uns.ac.id
Correspondence Author Email: alfaiz.asb1204@student.uns.ac.id

Abstract−Forecasting volatility in the USD/IDR exchange rate poses a critical challenge for Indonesia's financial stability, especially given how the data tends to display non-linear characteristics along with extreme outliers—what we commonly call fat-tails. This research develops a hybrid EGARCH-LSTM architecture to address these challenges by bringing together the precision of econometric modeling with deep learning's adaptability. Our dataset consists of 2,605 daily observations of the USD to IDR exchange rate ranging from 2015 to 2025. We extract volatility features using an EGARCH(1, 1) model with a t-distribution, which will then be entered as exogenous input into a Long-Short-Term Memory (LSTM) network. Our analysis shows a strong contrast between price predictions and risk estimates. In terms of price forecasting, the market demonstrates remarkable efficiency. The simple Naive Forecast, with an RMSE of 100.47, proved extremely difficult to outperform, lending support to the Random Walk Hypothesis. However, the Hybrid EGARCH-LSTM demonstrated superior volatility prediction capabilities, achieving the lowest Out-of-Sample Volatility MAE of 0.0075 compared to 0.0083 for the standalone EGARCH model. The EGARCH-LSTM hybrid model achieved the best performance, passed the Kupiec backtest (P-value = 0.4373), and significantly outperformed the pure econometric model, which failed due to excessive conservatism in its estimation. This study concludes that although accurate price predictions are still unattainable in efficient markets, the EGARCH-LSTM hybrid architecture provides a powerful and reliable tool for risk estimation. These results offer significant implications for hedging and risk management practices in the Indonesian foreign exchange market.

Keywords: Time Series; EGARCH; LSTM; Hybrid Model; Value at Risk (VaR)

1. INTRODUCTION

Exchange rates are one of the most crucial macroeconomic indicators that determine the stability of a country's economic foundation, particularly for developing countries that adopt an open financial system, such as Indonesia. In the context of global trade and finance, exchange rates act as a key price variable connecting domestic economic conditions with global market dynamics. Fluctuations in exchange rates, particularly the USD to Rupiah (USD to IDR), have a direct impact on various strategic economic sectors, ranging from changes in import costs, export earnings, the value of foreign debt payments along with interest, to inflation rates, which ultimately determine the purchasing power of the public. Small changes in global economic variables can lead to significant instability in exchange rates, making the foreign exchange (forex) market a highly dynamic, complex, and unpredictable system. Thus, the ability to predict the direction of movement as well as the risk of exchange rate fluctuations becomes an important aspect. This is necessary not only by monetary authorities in determining the right intervention strategies, but also by investors and financial market participants in their efforts to manage and mitigate risk.

Empirical findings show that exchange rate volatility acts as a significant moderating variable in the relationship between currency depreciation and trade balance performance in developing countries [1]. These findings confirm that volatility plays a role in accelerating market adjustments to exchange rate changes, resulting in price movement patterns that are difficult to explain with linear models. This complexity in dynamics is reinforced by research conducted by Kartono et al. [2], which found that the movement of the USD against the IDR has characteristics resembling turbulence in fluid flow, where particles move randomly at high speeds. This indicates that the USD-to-IDR exchange rate movement is non-linear and complex. These characteristics render traditional linear models inadequate for price forecasting or risk estimation. Thus, an approach is needed that can capture the non-linear nature and heteroscedasticity of exchange rate data.

Various previous studies have identified two approaches to modeling exchange rates and their volatility. The first approach is the use of econometric models, particularly the GARCH (Generalized Autoregressive Conditional Heteroskedasticity) family of models and its variants, such as EGARCH (Exponential GARCH). The EGARCH model has the advantage of capturing the phenomenon of volatility clustering while also accounting for asymmetric or leverage effects, which is the tendency for volatility to increase more sharply in response to negative information compared to positive information at the same intensity level [3]. This advantage makes the EGARCH model more adaptive in capturing market turbulence. Nevertheless, econometric models have limitations in capturing the dynamics of average prices because they rely on assumptions of linearity or specific distributions that are often not met in economic data. The second approach is the deep learning model, particularly the Long Short-Term Memory (LSTM) architecture, which is an advancement of the Recurrent Neural Network (RNN). This model is effective in learning long-term dependencies and non-linear patterns in financial time series data [4]. As a result, LSTM proves to be superior to linear regression models in handling complex time series data. The reliability of LSTM in processing financial data is supported by comparative research on the USD to IDR exchange rate, which shows that LSTM has more consistent and stable performance compared to other architectures such as transformers [5].

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 1, December 2025 | Hal 502-510
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i1.884

Nevertheless, both approaches have limitations when used separately. A pure LSTM model, while effective for price forecasting, often overlooks structural information regarding current market volatility conditions [6]. Conversely, a pure EGARCH model, although designed to model volatility, often fails to produce robust risk estimates such as Value at Risk (VaR) in backtesting, and cannot capture complex, non-linear temporal dependencies. These limitations suggest that a single model is not sufficient to explain the complex and volatile dynamics of the foreign exchange market.

With the development of computational methods, an increasing number of studies propose the use of hybrid models that combine EGARCH and LSTM. This approach leverages EGARCH's ability to extract conditional volatility and market sensitivity, which are then used as additional features in LSTM networks to learn non-linear price movement patterns. Several studies have shown that hybrid models, namely GARCH-LSTM and EGARCH-LSTM, can produce more accurate predictive performance compared to single models, whether in the context of the stock market [7], cryptocurrencies [3], commodities [8], or equity markets [9]. In the foreign exchange market, the hybrid model approach has proven effective in estimating risk based on the Value at Risk (VaR) measure [6].

Prior studies have explored hybrid methodologies to mitigate the limitations of single models. For instance, Nsengiyumva et al. [6] demonstrated that GARCH-LSTM significantly improves VaR estimation for African currencies. Similarly, Koo and Kim [7] found that combining GARCH with LSTM enhances volatility prediction in stock markets. However, most existing studies utilize sequential input methods. This research differentiates itself by employing a MIMO architecture where daily volatility features are treated as exogenous inputs to capture the 2015–2025 USD/IDR dynamics.

Based on this, this study proposes the application of a hybrid EGARCH-LSTM model to model and predict the daily USD to IDR exchange rate. The study aims to compare the hybrid model's performance against ARIMA, Naive Forecast, and Standard LSTM benchmarks, as well as performing risk estimation based on Value at Risk (VaR 95%) evaluated using the Kupiec backtesting test. With this approach, the study is expected to make an empirical contribution to the analysis of the Indonesian foreign exchange market and offer a more robust approach in price forecasting and risk estimation compared to using the EGARCH or LSTM models separately.

2. RESEARCH METHODOLOGY

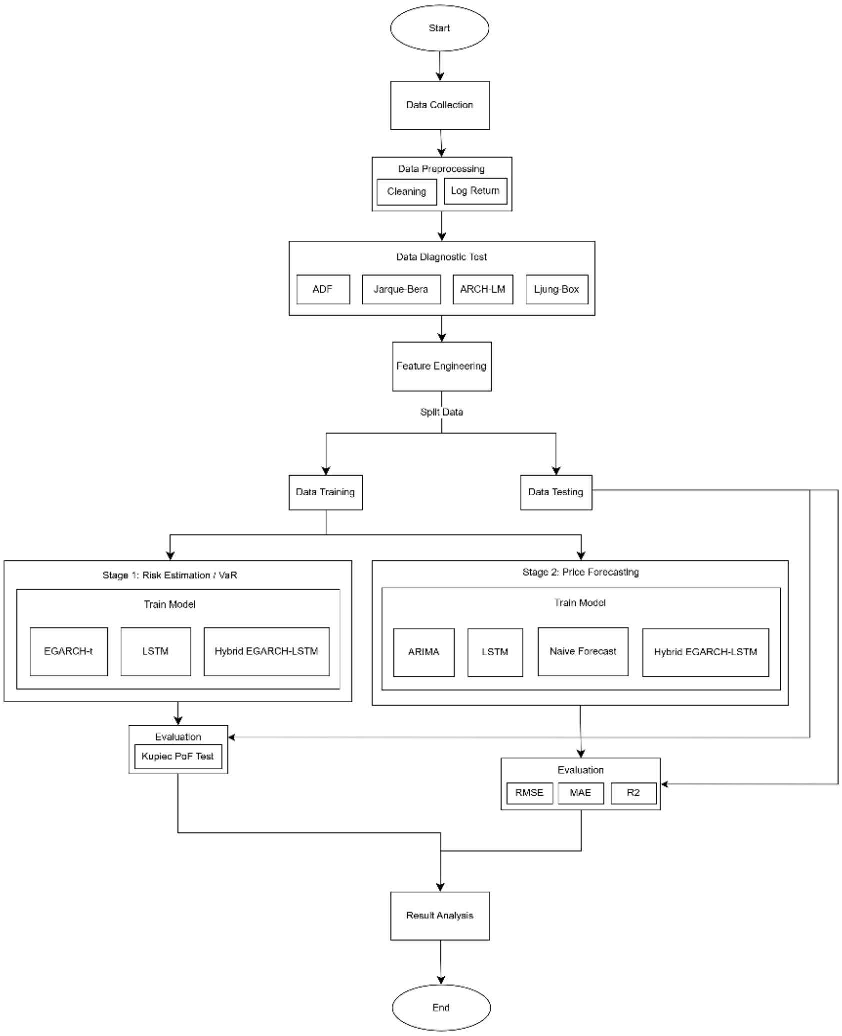

Figure 1. Research Flowchart for Hybrid EGARCH-LSTM Modeling

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 1, December 2025 | Hal 502-510
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i1.884

Figure 1 outlines the structured progression of this research. The workflow begins with the data collection phase, during which daily USD/IDR exchange rate observations are obtained. These observations then move to the data preprocessing stage, where missing entries are addressed and closing prices are transformed into log returns to enhance statistical reliability. After preprocessing, a set of diagnostic tests is applied to identify essential time series characteristics such as stationarity, distributional properties, and volatility clustering. These tests include the Augmented Dickey-Fuller (ADF), Jarque-Berra, ARCH-LM, and Ljung-Box tests.

Armed with the diagnostic findings, the study advances to the Feature Engineering stage, where volatility patterns are extracted as inputs for the deep learning architecture. The dataset is then partitioned using the Split Data procedure to create distinct training and testing subsets, ensuring that data leakage is fully mitigated. The modelling framework proceeds along two parallel stages. The first stage involves VaR estimation, in which the EGARCH-t model, the standard LSTM, and the Hybrid EGARCH-LSTM are trained and then evaluated using the Kupiec Probability of Failure (PoF) test to assess their validity. The second stage addresses price forecasting, benchmarking the hybrid model against ARIMA, standard LSTM, and naive forecasting approaches, with performance measured through RMSE, MAE, and R^2. The study concludes with the result analysis phase, which integrates outcomes from both modeling streams to formulate empirical insights on model robustness.

2.1 Basic Research Framework

This study employs a quantitative and experimental research design to understand the dynamics of the USD/IDR exchange rate. Daily secondary data covering the 2015–2025 period is obtained from Yahoo Finance, with logarithmic returns used to ensure statistical stability. The logarithmic return rt at time t is computed using the formulation shown in Equation (1).

$$rt = ln(\frac{Pt}{Pt-1}) \tag{1}$$

Where Pt represents the closing price on day t, and $Pt - 1$ represents the closing price of the previous day.

The methodological rationale rests on a fundamental duality in financial time series, while price levels behave stochastically, volatility exhibits structural clustering. This characteristic motivates the integration of an econometric volatility model (EGARCH) with a deep learning network (LSTM), forming a hybrid framework intended to enhance risk calibration. The hybrid architecture is designed around a separation of tasks principle. EGARCH serves as the volatility extractor, capturing asymmetric and clustered variance patterns, whereas LSTM handles temporal price dependencies. The extracted volatility series is then incorporated as an exogenous input to the neural network, enabling it to learn price dynamics within the appropriate risk regime. Model validation considers both forecasting accuracy and risk estimation reliability, combining conventional error metrics with formal backtesting through the Kupiec Probability of Failure test.

2.2 Data Characteristics and Diagnostic Tests

After removing observations with missing values, the dataset comprises 2,605 valid observations. Given the sequential nature of time-series data, the train-test split is performed in chronological order. An 80:20 split is applied, allocating data from 2015 to early 2023 for model training and parameter estimation, while the period from 2023 to 2025 is reserved exclusively for model evaluation. To characterize the underlying statistical properties of the series, several diagnostics tests are conducted prior to model development.

2.2.1 Stationarity Test using Augmented Dickey Fuller (ADF)

Stationarity test is assessed on the Close data using ADF which identifies the presence of unit roots. This test is widely employed in financial econometrics to determine whether differencing or transformation is required [1][2].

The ADF test is an extension of the simple Dickey Fuller test to determine whether financial time series data is stationary or not [10]. The ADF test is widely employed to assess whether a time series exhibits unit root behavior [11]. The presence of a unit root is indicative of non-stationarity, implying that the series exhibits time-varying mean and variance rather than remaining constant over time. The mathematical model of ADF test shown in Equation (2)

$$\Delta yt = \alpha + \beta t + \gamma yt - 1 + \delta 1 \Delta yt - 1 + \delta 2 \Delta yt - 2 + \cdots + \delta p \Delta yt - p + \mathcal{E}t \tag{2}$$

Where Δyt is the first difference of time series y, α is the constant term, and βt represents the linear trend. The term $yt - 1$ refers to the lagged level of y, while $\Delta yt - 1$ through $\Delta yt - p$ correspond to its lagged first difference up to p lags. The $\mathcal{E}t$ is an error term. The null hypothesis asserts that the series possesses a unit root, implying non-stationarity, while the alternative hypothesis states that the series is stationary.

2.2.2 Normality Test Using Jarque Bera

Distributional characteristics of the log-return series are examined through the Jarque Bera test to identify departures from normality, particularly the presence of heavy tails, a feature extensively documented in prior literature [3] [6] [12]. The Jarque–Bera test, developed by Jarque and Bera in 1980, is a commonly used method for checking whether a variable follows a normal distribution. The test statistic is constructed from the sample's skewness (S) and kurtosis (K). Under the assumption of normality, the expected values of S and K are 0 and 3, respectively. The test statistic JB of Jarque Bera is defined by Equation (3)

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 1, December 2025 | Hal 502-510
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i1.884

$$JB = \frac{n}{6}\left(S^2 + \frac{(K-3)^2}{4}\right) \quad (3)$$

Where n denotes the sample size, S^2 represents the sample skewness, and K is the sample kurtosis. Under normality, skewness equals zero and kurtosis equals three, so S^2 and $(K-3)^2$ measure the deviation of the sample distribution from the normal benchmark [13]. Insights from this diagnostic justify the adoption of a Student's t-distribution within the volatility specification, as financial returns, especially in emerging markets, commonly display excess kurtosis relative to a Gaussian distribution.

2.2.3 ARCH-LM Test

Conditional heteroskedasticity is examined using the ARCH–LM test [14], a Lagrange Multiplier procedure that evaluates autocorrelation in the squared residuals to identify ARCH behavior. This diagnostic is essential for determining whether the residual variance is non-constant and dependent on past disturbances, thereby guiding the selection of appropriate volatility models such as ARCH or GARCH. These properties provide methodological support for employing GARCH-family models during the feature-engineering stage [15] [16].

2.2.4 Ljung Box Test

The Ljung–Box test is a diagnostic procedure used to determine whether a time series exhibits significant autocorrelation. This test evaluates autocorrelation at several lags, making it suitable for assessing randomness or checking whether a fitted model has adequately captured serial dependence. The null hypothesis of the Ljung–Box test states that the series behaves like white noise, meaning that there is no significant autocorrelation at any of the lags examined. Conversely, the alternative hypothesis states that at least one lag contains statistically significant autocorrelation, indicating that the series is not white noise. This combination of diagnostics ensures that the selected modeling framework is consistent with the facts of stylized financial time series and capable of capturing their main statistical regularities [17].

2.3 Hybrid Model Architecture

2.3.1 Feature Engineering via EGARCH

This study trains the EGARCH(1,1)-t model exclusively on training data to extract volatility patterns. The EGARCH model was chosen for its ability to capture asymmetric leverage effects, whereby markets typically exhibit stronger volatility reactions to bad news than to good news [3]. The conditional volatility generated by this model is not discarded but is retained and included as an additional exogenous feature, referred to as Volatility_Feature.

2.3.2 MIMO LSTM Implementation

Once the engineering features are ready, we normalize all variables using MinMaxScaler and apply a 30-day sliding window for processing. Fjellström [18] and other researchers have shown how this technique helps improve stationarity in financial time series data. Our LSTM architecture follows a MIMO (Multi-Input Multi-Output) design that takes two simultaneous data streams: historical price information and volatility features that we extracted earlier. The network itself is structured with an initial LSTM layer containing 128 units, followed by a Dropout layer set at 0.2 to prevent overfitting, followed by another LSTM layer with 64 units. The model is compiled using the Adam optimizer with a learning rate of 0.001 and Mean Squared Error (MSE) loss function. Training is conducted over 50 epochs with a batch size of 32 to ensure convergence. Recent comparative research by Islamy and Wahabi [19] and Liu [20] emphasizes the importance of obtaining the right LSTM layer configuration to effectively capture nonlinear temporal patterns. The comparative studies by Islamy and Wahabi [19] and Liu [20] highlight the importance of carefully selecting the LSTM layer configuration to effectively capture nonlinear temporal dependencies. Consistent with the findings of Xu et al. [21], this multi-input design allows the model to learn price dynamics while obtaining information about prevailing market risks. This additional source of information provides a clear advantage over conventional single-input LSTM architectures.

2.4 Evaluation Scenarios

To obtain a comprehensive assessment of model reliability, performance is evaluated in two stages using a test data set. The first stage is Risk Value estimation. This stage tests each model's ability to generate Value at Risk (VaR) estimates at a 95% confidence level. The Kupiec Proportion of Failures (POF) test is used as the primary validation tool, as it is widely used to assess whether the observed frequency of VaR exceedances is consistent with the expected level. This procedure allows for an objective assessment of whether the Hybrid model improves VaR calibration without exhibiting excessive conservatism a problem often associated with pure EGARCH specifications in emerging markets [15].

The second stage is price This stage evaluates the predictive accuracy of the models in forecasting closing prices. Performance is quantified using the Root Mean Squared Error (RMSE). Consistent with standard practice in markets that may display random-walk behavior, the Hybrid and Standard LSTM models are benchmarked against a Naive Forecast and ARIMA. For the ARIMA model, the optimal order (p, d, q) is determined using the auto_arima algorithm which minimizes the Akaike Information Criterion (AIC) to select the best-fitting model automatically. As noted by Stevenson [22] and Zhang [23], including a Naive baseline is essential for determining whether more complex forecasting models offer improvements that are practically and statistically meaningful.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 1, December 2025 | Hal 502-510
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i1.884

3. RESULT AND DISCUSSION

3.1 Diagnostic Test Results and Modeling Implications

3.1.1 Stationarity Analysis and the Random Walk Hypothesis

The analysis began with the application of standard stationarity tests to the closing price data of the USD/IDR exchange rate. The results showed that both the Augmented Dickey-Fuller (ADF) test (P-value 0.0502) and the KPSS test (P-value 0.0213) indicated the non-stationarity of the price series. In financial econometrics, this finding has significant theoretical weight because it is consistent with the Random Walk Hypothesis and is a characteristic of efficient financial markets [23] [24]. The implication is that the mean and variance of exchange rates are not constant but constantly change in response to new macroeconomic information. The absence of a constant mean or memory indicates that the Naive Forecasting (Random Walk) model serves as a theoretical benchmark that is unexpectedly difficult, often challenging the predictive capacity of even sophisticated non-linear models.

3.1.2 Distributional Properties and Leptokurtosis

The next experiment focused on log-return distribution. The Jarque-Bera normality test produced a P-value of 0.0000, which led to a strong rejection of the null hypothesis of normality. These statistical results provide convincing empirical evidence of fat tails or leptokurtosis in the return distribution. This is recognized as a classic feature in financial time series, especially in emerging markets such as Indonesia [25] [26]. Practically, this indicates that extreme market movements such as currency crises or geopolitical shocks occur more frequently than predicted by the standard Gaussian curve. Recognizing this reality requires a critical methodological decision, namely the adoption of the t-Student distribution in the proposed EGARCH framework. Failure to account for fat tails using the standard normal distribution will result in underestimating the probability of these extreme tail risks, thereby reducing the resilience of risk management components.

3.1.3 Volatility Clustering and Serial Dependence

The final stage involves analyzing the Log_Return series for temporal dependence. Both the ARCH-LM test (P-value < 0.00001) and the Ljung-Box test (P-value < 0.00001) confirm the presence of significant volatility clustering and nonlinear temporal dependence. This indicates that while the direction of price movements may follow a random walk, the magnitude of price changes exhibits dependence, with periods of high volatility tending to cluster together. According to Liu [20] and Mualifah et al. [27], these statistical characteristics provide strong justification that GARCH-type models are capable of capturing the observed heteroscedasticity (volatility clustering) and LSTM performance in modeling nonlinear serial dependence in time series.

The diagnostic phase therefore confirmed that a hybrid architecture is not merely a sophisticated design choice but a statistical necessity to adequately address the diverse properties observed in the USD/IDR exchange rate, which is random price movement, non-normal (fat-tailed) errors, and clustered volatility.

3.1.4 Synthesis and Architectural Justification

Collectively, these diagnostic results paint a clear picture of a complex financial asset. The price follows a random walk, yet the returns exhibit leptokurtosis, significant volatility clustering, and non-linear serial dependencies. This confirms the central premise that no single, simple econometric or machine learning model can optimally capture all these competing phenomena. The hybrid EGARCH-LSTM architecture is thus theoretically justified as the most sound solution, specifically engineered to simultaneously manage the volatility dynamics (via EGARCH) and the non-linear temporal patterns (via LSTM).

3.2 Risk Modeling (VaR) Performance Evaluation

Prior to risk estimation, we evaluated the volatility forecasting performance. The Hybrid EGARCH-LSTM achieved the lowest Out-of-Sample Volatility MAE (0.0075), outperforming the pure EGARCH model (0.0083). This superior volatility tracking provides a solid foundation for VaR calculation. The initial stage of model validation focused on the accuracy of the 95% Value at Risk (VaR) estimate using Kupiec's Proportion of Failures (POF) test. This test serves as a critical stress assessment, which requires a valid model to accurately predict the frequency of actual losses exceeding the VaR threshold. To achieve calibration, the model must produce a P-value greater than 0.05. The results of the 515-day backtest are summarized in Table 1.

Table 1. Backtesting Results (Kupiec POF Test)

Model	Violations (N)	Rate	P-Value (Kupiec)	Status (Target: > 0.05)
Hybrid EGARCH-LSTM	22	4.27%	0.4373	Accepted (Valid)
LSTM (VaR)	18	3.50%	0.0984	Accepted (Valid)
EGARCH(1,1)-t	16	3.11%	0.0346	Rejected (Failed)

Table 1 highlights a clear contrast in model adaptability. The EGARCH-LSTM hybrid model emerged as the clear winner, being the only model that passed the calibration test with a robust P-value of 0.4373. The observed breach rate

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 1, December 2025 | Hal 502-510
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i1.884

was 4.27%, which is very close to the theoretical VaR target of 5%. This result places the hybrid model in the ideal "Goldilocks" zone, neither too aggressive nor too conservative. This success strongly supports the premise that combining GARCH structural volatility insights with LSTM non-linear learning capabilities significantly sharpens risk precision [6] [21]. By incorporating EGARCH derived volatility as a feature, the LSTM component appears to have developed the capacity to dynamically adjust risk thresholds as market conditions evolve, effectively leveraging the strengths of econometric modeling and deep learning.

Although the Standard LSTM passed the Kupiec test with a lower violation rate (3.50%) than the Hybrid model (4.27%), the Hybrid model is considered superior in terms of capital efficiency. A 3.50% violation rate implies the LSTM is overly conservative, holding unnecessary idle capital. The Hybrid model's rate (4.27%) is closer to the theoretical target of 5%, striking a better balance between safety and efficiency.

Conversely, the standalone EGARCH(1,1)-t model failed the POF test with a P value of 0.0346. This failure is very thought-provoking: not because it underestimates risk, but because of excessive caution. This model only recorded 16 violations (3.11%), far below the expected 26 violations. Although it appears safe, this excessive risk calculation forces companies to allocate "idle capital," which is excess reserves that could otherwise be used productively. The theoretical explanation lies in the rigidity of the model; static econometric models, once trained on periods of high historical volatility (e.g., the COVID-19 era in our training data), become "haunted" by past risks [15] [16]. They continue to predict high-risk scenarios even after the market has stabilized, resulting in excess capital reserves that fail standard backtests.

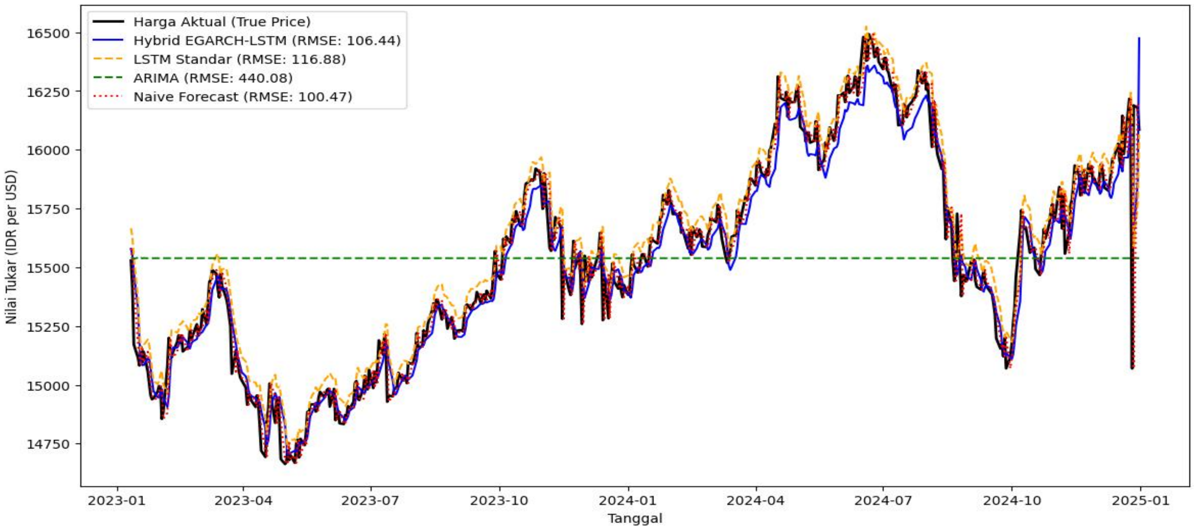

Figure 2. 95% VaR Validation (USD/IDR)

This statistical narrative is visually confirmed in Figure 2. The visual comparison highlights different behavioral patterns during market transitions. The 'Best VaR' trajectory (Hybrid EGARCH-LSTM, blue) demonstrates superior adaptability, closely tracking the Actual Logarithm (gray). When market volatility spikes, the Hybrid model's VaR bounds immediately widen to capture risk; conversely, they quickly narrow when the market calms down. The 'VaR Fail' line (EGARCH(1,1)-t, red dotted) maintains a consistently lower number of violations, illustrating this "ghost effect"; its rigidity prevents it from dynamically adapting to the new, calmer market reality. This confirms that the LSTM memory gate successfully leverages volatility information from the previous EGARCH layer, enabling smarter and more context-sensitive risk decision-making.

3.3 Price Forecasting Evaluation

The final phase of model evaluation involves Out-of-Sample (OOS) forecasting for both USD/IDR closing prices and their associated volatility, with a particular focus on the RMSE metric for prices and the predictive power of the model during unprecedented market turbulence. Table 2 presents the comparative performance across models during this validation period.

Table 2. Comparison of Price Prediction Accuracy Metrics

Model	RMSE	MAE	R^2
Naive Forecast	100.4722	60.6633	0.9479
Hybrid EGARCH-LSTM	106.4385	73.9233	0.9415
LSTM Standar	116.8772	82.9251	0.9295
ARIMA	440.0777	364.7442	0.0000

As shown in Table 2, Out-of-sample data strongly reinforces the main findings. The ARIMA model performed poorly with an RMSE of 440.08 and an R^2 of 0.0000. This indicates that the ARIMA model failed to capture the stochastic trend of the exchange rate and essentially predicted the mean value (flat line), which is a common limitation of linear

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 1, December 2025 | Hal 502-510
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i1.884

models in random walk data. While Naive Forecast remains the best, the Hybrid EGARCH-LSTM (RMSE 106.44) demonstrated an improvement of approximately 8.9% over the Standard LSTM (RMSE 116.88), indicating that incorporating volatility features helps reduce forecasting errors during turbulent periods. This reaffirms the market efficiency hypothesis and the intrinsic random walk behavior of price series [23]. Attempts by even the most sophisticated models to outperform this short-term random walk are fundamentally constrained by market information efficiency.

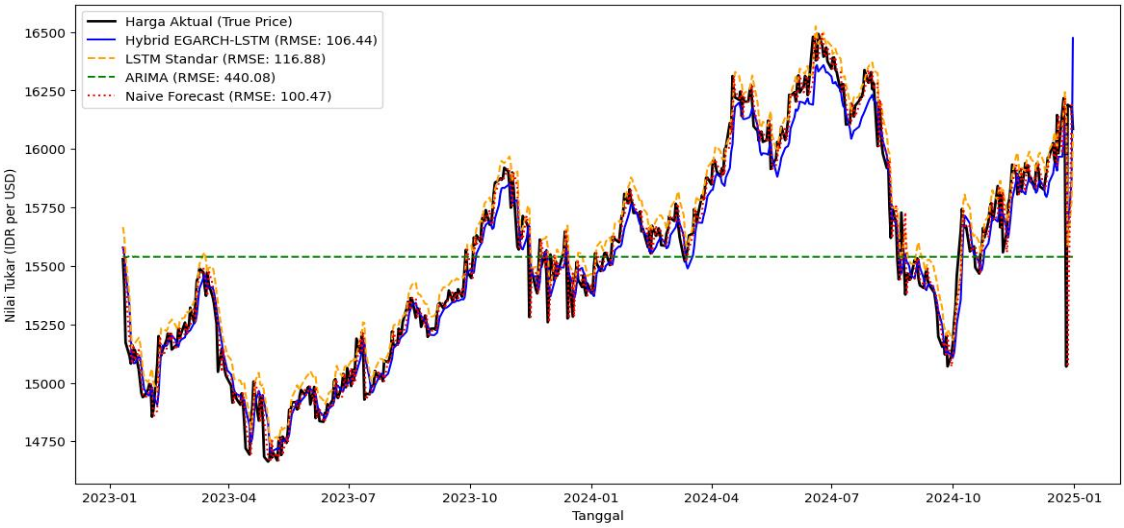

Figure 3. Comparison of Price Predictions (USD/IDR)

However, Hybrid EGARCH-LSTM shows clear superiority in the field of volatility prediction. This model achieves the lowest Out-of-Sample Volatility MAE (0.0075), significantly outperforming both Standard LSTM (0.0098) and standalone EGARCH(1,1)-t (0.0083). The 10.6% reduction in MAE compared to the standalone EGARCH model confirms that integrating EGARCH features with the non-linear time-series learning capabilities of LSTM provides demonstrably better conditional volatility estimates.

These results confirm the superiority of the architecture established in Stage 1 (VaR accuracy) and are visually illustrated in Figure 3. The graph shows that the Hybrid model's volatility estimates are much more responsive to sharp peaks and troughs in real market variance than other models, highlighting its superior ability to capture the dynamics of volatility clustering in out-of-sample periods. The Standard LSTM, which does not explicitly model heteroscedasticity, consistently underestimates or lags in responding to these critical volatility spikes.

Finally, the results of Phase 3 confirm the dual nature of the USD/IDR time series and validate the choice of a hybrid model: while the Price Level is best estimated using the Naive model due to market efficiency, Risk Dynamics (Volatility) are best captured and predicted by the EGARCH-LSTM Hybrid architecture. This confirms that the hybrid approach provides maximum predictive power in the areas most important for risk management and capital allocation.

3.4 Discussion

3.4.1 Comparative Analysis with Hybrid Models

The validation results confirm the main hypothesis of this study: the superior risk estimation capabilities of our EGARCH-LSTM P-value 0.4373 Multi-Input Multi-Output (MIMO) model compared to single models underscores the importance of hybrid architecture. This finding is not isolated, but rather aligns significantly with recent advances in financial literature. For example, AlMadany et al. [3] show that a similar GARCH-LSTM hybrid model performs better in highly volatile cryptocurrency markets, while Nsengiyumva et al. [6] report a significant improvement in Value at Risk (VaR) estimation for African currencies (USD/RWF) using the GARCH-LSTM approach.

Our research expands on these global findings by confirming that this hybrid advantage applies to the highly liquid USD/IDR pair. Furthermore, we provide a crucial architectural distinction. Unlike studies that use a sequential input approach [3], our MIMO design treats volatility as an exogenous feature. This fusion method successfully captures the increased information from the GARCH component without reducing the raw price signal. This shows that the model combination mechanism is as important as the selection of individual components for performance.

3.4.2 Structural Breaks and the Failure of Static Parameters

The main contribution of this study is an empirical diagnosis of why traditional econometric models fail in dynamic environments. As shown in Table 1, the EGARCH(1,1)-t model was rejected with a P-value (0.0346) because it was too conservative. The limitations of standard GARCH models in handling structural changes and dynamic shifts in volatility regimes have been well documented [15] [16].

The exact mechanism of this failure is rooted in parameter rigidity. The training period (2015–2023) included a significant spike in volatility, particularly during the COVID-19 pandemic. The EGARCH model “locked in” these high-

ISSN 2774-3659 (Media Online)
Vol 6, No 1, December 2025 | Hal 502-510
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i1.884

risk parameters. When the market regime shifted to a relatively calmer phase during the testing period (2023–2025), the static model failed to adapt, continuously overestimating risk. In contrast, the LSTM component of the hybrid model, by utilizing its forget gate and dynamic weight updates, successfully “forgot” the high volatility regime and adjusted to the new market reality. This provides strong empirical support for the argument that hybrid models offer a robust solution to the limitations of traditional econometrics in emerging markets [20].

3.4.3 The Dichotomy of Market Efficiency and Risk Predictability

The results presented in Table 2 establish a significant theoretical dichotomy. While some studies have found success in predicting commodity prices using hybrid models [8], our analysis confirms that the Random Walk Hypothesis remains fundamentally valid for the USD/IDR exchange rate, as evidenced by the clear dominance of the Naive Forecast (RMSE 100.47). This is very much in line with Zhang's [23] statement that the Naive baseline is often unbeatable in financial time series due to non-stationarity and information efficiency.

However, this finding does not negate the value of the hybrid model. On the contrary, it explains its true purpose. The significant RMSE reduction (8.9%) achieved by MIMO EGARCH-LSTM compared to Standard LSTM confirms that “distillation facts” such as volatility clustering contain important predictive information that is often overlooked by uninformed deep learning models. Therefore, the main implication is that there is a clear separation. Although consistent price predictions remain difficult to achieve due to market efficiency, risk is highly predictable.

This difference has very important practical value for the real sector. For Indonesian companies exposed to currency fluctuations, the main objective is not short-term speculation, but rather accurate estimation of the maximum potential loss VaR to ensure adequate hedging and capital reserves. In this particular capacity, it functions as a “Risk Radar” rather than a “Crystal Ball.” The proposed EGARCH-LSTM architecture has proven to be indispensable, offering a level of precision and dynamic adaptability that is unmatched by pure traditional statistics or basic AI models.

4. CONCLUSION

This study finds that the Hybrid EGARCH–LSTM model provides improved performance in Value-at-Risk (VaR) estimation for the USD/IDR exchange rate. The model meets the 95% VaR requirement by passing the Kupiec test (P-value 0.4373) and producing a violation rate of 4.27% (22 violations), which aligns with the expected threshold. In comparison, the standard EGARCH(1,1)-t model does not satisfy the backtest criteria (P-value 0.0346) due to its overly conservative forecasts. These results indicate that combining EGARCH-based volatility dynamics with the temporal learning capabilities of LSTM yields a more responsive framework for risk measurement. This is quantitatively supported by the Hybrid model achieving the lowest Volatility MAE (0.0075), confirming its ability to track market turbulence more accurately than single models. For price forecasting, the findings show that no model consistently outperforms the random walk benchmark. The Naive forecast achieves the lowest RMSE (100.47), surpassing both the Hybrid EGARCH–LSTM (106.43) and the standard LSTM (116.87). This outcome reflects the inherent non-stationarity and limited predictability of USD/IDR price levels rather than a shortcoming of the modelling approach. Overall, the hybrid model is more suitable for risk-related applications than for directional price forecasting. Future work may consider incorporating macroeconomic exogenous variables or applying the hybrid architecture to assets with different market efficiencies.

REFERENCES

[1] M. Bosupeng, A. Naranpanawa, and J. J. Su, “Does exchange rate volatility affect the impact of appreciation and depreciation on the trade balance? A nonlinear bivariate approach,” *Econ Model*, vol. 130, Jan. 2024, doi: 10.1016/j.econmod.2023.106592.

[2] A. Kartono, M. Febriyanti, S. T. Wahyudi, and Irmansyah, “Predicting foreign currency exchange rates using the numerical solution of the incompressible Navier–Stokes equations,” *Physica A: Statistical Mechanics and its Applications*, vol. 560, Dec. 2020, doi: 10.1016/j.physa.2020.125191.

[3] N. N. AlMadany, O. Hujran, G. Al Naymat, and A. Maghyereh, “Forecasting cryptocurrency returns using classical statistical and deep learning techniques,” *International Journal of Information Management Data Insights*, vol. 4, no. 2, Nov. 2024, doi: 10.1016/j.jjimei.2024.100251.

[4] N. Nengah *et al.*, “(print) JSIKTI: Jurnal Sistem Informasi dan Komputer terapan Indonesia Deep Learning Approach for USD to IDR Forecasting with LSTM Deep Learning Approach for USD to IDR Forecasting with LSTM "Deep Learning Approach for USD to IDR Forecasting with LSTM,” *JSIKTI: Jurnal Sistem Informasi dan Komputer Terapan Indonesia*, vol. 8, no. 1, pp. 91–100, 2025, [Online]. Available: www.infoteks.orgJournalPageisavailabletohttps://infoteks.org/journals/index.php/jsikti

[5] E. Verianto and A. F. Shimbun, “TRANSFORMER WITH LAGGED FEATURES FOR HANDLING LONG-TERM DATA DEPENDENCY IN TIME SERIES FORECASTING,” *JIKO (Jurnal Informatika dan Komputer)*, vol. 7, no. 3, pp. 232–241, Dec. 2024, doi: 10.33387/jiko.v7i3.9247.

[6] E. Nsengiyumva, J. K. Mung’atu, and C. Ruranga, “Hybrid GARCH-LSTM Forecasting for Foreign Exchange Risk,” *FinTech*, vol. 4, no. 2, Jun. 2025, doi: 10.3390/fintech4020022.

[7] E. Koo and G. Kim, “A Hybrid Prediction Model Integrating GARCH Models With a Distribution Manipulation Strategy Based on LSTM Networks for Stock Market Volatility,” *IEEE Access*, vol. 10, pp. 34743–34754, 2022, doi: 10.1109/ACCESS.2022.3163723.

[8] C. S. Huang and A. Sayed, “Novel forecasting of white maize futures volatility: a hybrid GARCH-based bi-directional LSTM model,” *Cogent Economics and Finance*, vol. 13, no. 1, 2025, doi: 10.1080/23322039.2025.2484422.

ISSN 2774-3659 (Media Online)
Vol 6, No 1, December 2025 | Hal 502-510
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i1.884

[9] W. Nhlapho, M. Atemkeng, and J. C. Ndogmo, "An attention-guided hybrid statistical and deep learning modeling for enhanced time series forecasting: A case study of South African telecommunication companies," *Sci Afr*, vol. 30, Dec. 2025, doi: 10.1016/j.sciaf.2025.e02950.

[10] Z. Guo, "Research on the Augmented Dickey-Fuller Test for Predicting Stock Prices and Returns," *Advances in Economics, Management and Political Sciences*, vol. 44, no. 1, pp. 101–106, Nov. 2023, doi: 10.54254/2754-1169/44/20232198.

[11] J. R. Porto de Carvalho, E. D. Assad, A. F. de Oliveira, and H. Silveira Pinto, "Annual maximum daily rainfall trends in the midwest, southeast and southern Brazil in the last 71 years," 2014, *Elsevier B.V.* doi: 10.1016/j.wace.2014.10.001.

[12] H. Malmsten and T. Terasvirta, "Stylized Facts of Financial Time Series and Three Popular Models of Volatility," vol. 3, Sep. 2004.

[13] T. Thadewald and H. Büning, "Jarque–Bera Test and its Competitors for Testing Normality – A Power Comparison," *J Appl Stat*, vol. 34, no. 1, pp. 87–105, Jan. 2007, doi: 10.1080/02664760600994539.

[14] R. F. Engle, "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, vol. 50, no. 4, pp. 987–1007, 1982, doi: 10.2307/1912773.

[15] J. C. Smolović, M. L. Božović, and S. Vujošević, "GARCH models in value at risk estimation: Empirical evidence from the montenegrin stock exchange," *Economic Research-Ekonomska Istrazivanja* , vol. 30, no. 1, pp. 477–498, Apr. 2017, doi: 10.1080/1331677X.2017.1305773.

[16] N. Godfrey Emenogu, "Selecting superior GARCH model with backtesting approach in First Bank of Nigeria stock returns," 2023. [Online]. Available: https://www.researchgate.net/publication/378139201

[17] H. Hassani and M. Yeganegi, "Selecting optimal lag order in Ljung–Box test," *Physica A: Statistical Mechanics and its Applications*, vol. 541, p. 123700, Dec. 2019, doi: 10.1016/j.physa.2019.123700.

[18] C. Fjellström, "Long Short-Term Memory Neural Network for Financial Time Series," Jan. 2022, [Online]. Available: http://arxiv.org/abs/2201.08218

[19] C. C. Islamy and A. Wahabi, "Sistemasi: Jurnal Sistem Informasi Perbandingan Model LSTM dan GRU untuk Peramalan Angin A Comparative Study of LSTM and GRU Models for Wind Forecasting." [Online]. Available: http://sistemasi.ftik.unisi.ac.id

[20] J. Liu, "A Hybrid Model Integrating LSTM with Multiple GARCH-Type Models for Volatility and Var Forecast," European Alliance for Innovation n.o., Jun. 2023. doi: 10.4108/eai.6-1-2023.2330313.

[21] K. Xu, Y. Wu, M. Jiang, W. Sun, and Z. Yang, "Hybrid LSTM-GARCH Framework for Financial Market Volatility Risk Prediction," 2024. [Online]. Available: https://www.mfacademia.org/index.php/jcssa

[22] S. Stevenson, "A Comparison of the Forecasting Ability of ARIMA Models," *Journal of Property Investment & Finance*, vol. 25, pp. 223–240, May 2007, doi: 10.1108/14635780710746902.

[23] C. Zhang, "Movement Prediction-Adjusted Naive Forecast: Is the Naive Baseline Unbeatable in Financial Time Series Forecasting?," Oct. 2025, [Online]. Available: http://arxiv.org/abs/2406.14469

[24] K. Kladívko and P. Österholm, "Do market participants' forecasts of financial variables outperform the random-walk benchmark?," *Financ Res Lett*, vol. 40, p. 101712, 2021, doi: https://doi.org/10.1016/j.frl.2020.101712.

[25] J. He, "Capturing Four Stylized Facts of Financial Time Series in GARCH and Stochastic Volatility Models," 2020. doi: 10.23977/ICEMGD2020.051.

[26] D. Kim and M. Shin, "Volatility models for stylized facts of high-frequency financial data," *J Time Ser Anal*, vol. 44, no. 3, pp. 262–279, May 2023, doi: https://doi.org/10.1111/jtsa.12666.

[27] L. N. A. Mualifah, A. M. Soleh, and K. A. Notodiputro, "Comparison of GARCH, LSTM, and Hybrid GARCH-LSTM Models for Analyzing Data Volatility," *International Journal of Advances in Soft Computing and its Applications*, vol. 16, no. 2, pp. 150–165, 2024, doi: 10.15849/IJASCA.240730.10.