Proceedings of ICMRED 2026 Symposium: The Future of Work: Strategy, Workforce Transformation, and Organizational Renewal
DOI: 10.54254/2754-1169/2026.LD32298

A Comparative Study of ARIMA and Hybrid STL-ETS-GARCH(-SNAIVE) Models in Forecasting the Growth Enterprise Market

Zhiyan Yang

Hunan University, Changsha, China
202421060125@hnu.edu.cn

Abstract. The close price of the Growth Enterprise Market reflects the economic environment and market risk appetite. Since it is composed of growing companies such as technology and biopharmaceuticals, a rising closing price implies higher profit expectations in these startups, looser market liquidity, and a greater preference for market risk. In order to solve the issues mentioned above, this paper will apply the traditional ARIMA model and STL decomposition, respectively, processing ChiNext data from 2015 to 2025 after the transformation of Box-Cox. Employ ETS model in the trend and the ARIMA-GARCH model in the remainder which has heteroscedasticity. Compare difference before and after using the SNAIVE method in the seasonal portion in terms of weak seasonality. After calculating the economic effect, the hybrid STL-ETS-GARCH model achieved the highest strategic Sharpe ratio (0.876) and positive excess Sharpe ratio (0.212) in the out-of-sample testing, with its directional accuracy (49.2%) and profit-loss ratio (1.22) demonstrating robust performance. Although the STL-ETS-GARCH-SNAIVE model recorded the highest directional accuracy (51.3%), its Sharpe ratio (0.018) was significantly deteriorated due to the inclusion of redundant seasonal components. The ARIMA model delivered moderate performance. The paper raises a more elaborate method on the condition that seasonal components are handled prudently.

Keywords: Time Series Forecasting, ARIMA, STL Decompose.

1. Introduction

Close price of Growth Enterprise Market is highly volatile due to the unstable profit ability of enterprises in the growth stage [1]. The performance of ChiNext Index reflects the capital market's valuation in innovative enterprises. Besides, regulatory authorities can adjust policies and evaluate the effectiveness of capital market reforms, judging from the fluctuations and trend of ChiNext Index. Past studies have applied traditional ARIMA models or ARIMA-GARCHA models to forecast the closing price of Growth Enterprise Market [2,3]. The ARIMA model captures the linear trend in the short terms and has high interpretability. However, the ChiNext Index is influenced by countless factors and possesses significant nonlinear characteristics. The traditional ARIMA model fixes the variance and ignores changes in fluctuations. Besides, it is not suitable for long-term data

DOI: 10.54254/2754-1169/2026.LD32298

from 2015 to 2025. The ARIMA-GARCH model focuses on volatility clusters and overcomes the shortcomings of the traditional ARIMA models. STL has been implemented in financial research to model separately [4]. This method improves flexibility and extracts key characteristics of the data. The article intends to explore a comprehensive predictive model that simultaneously captures ChiNext Index's linear trend, limited seasonality and volatility aggression. It proposes a framework of exerting hybrid ETS, ARIMA-GARCH, and SNAIVE models. It evaluates not only the prediction accuracy (RMSE) but also economic performance (Sharpe ratio, direction accuracy, profit-to-loss ratio) and model stability (CVaR). Empirical evidence shows that improper handling of seasonal components can harm strategy performance under weak seasonal patterns.

2. Methodology

The closing price of Growth Enterprise Market data in China is sourced from China Stock Market & Accounting Research Database (CSMAR) through the purchased database in Hunan University(data.csmar.com).

2.1. Data derivation

Select the stock with code "399006" from the daily market data of domestic indices and extract the closing index from 1/5/2015 to 12/12/2025. Set frequency to 252 while transferring into time series data because it is the annual trading days in China's A-share market [5]. Divide 80% of the data into a training set. To stabilise variance, do a Box-Cox transformation to the data [6].

$$\mathrm{Y}\,(\lambda) = \begin{cases} \frac{y^{\lambda}-1}{\lambda}, & \lambda \neq 0 \\ \ln(y), & \lambda = 0 \end{cases} \tag{1}$$

where y denotes the original non-negative data to be transformed, λ denotes the transformation parameter that determines the transformation type. The Box-Cox lambda function in R suggests lambda=-0.1380812.

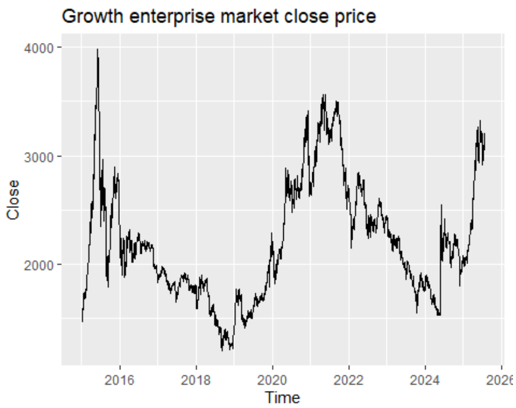

Figure 1. Growth market price

Figure 1 shows a plot of the Growth Enterprise Market. Figure 2 shows a seasonal graph of the data, which indicates that the data lacks seasonality.

DOI: 10.54254/2754-1169/2026.LD32298

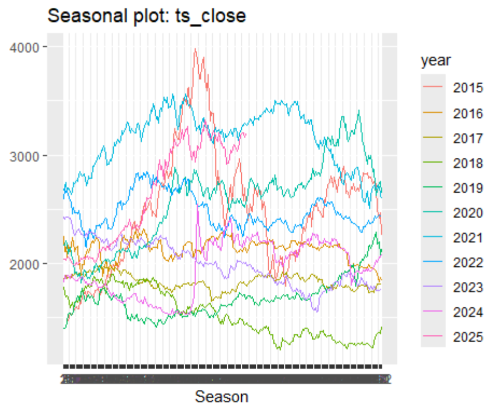

Figure 2. Growth market price by seasons

2.2. Traditional ARIMA

The traditional ARIMA model converts a non-stationary series into a stationary series after the d-th order difference. The formula controls auto-regressive order p and moving average order [7]:

$$\Delta^d \cdot X_t = c + \varphi^1 \cdot \Delta^d \cdot X_{t-1} + \varphi^2 \cdot \Delta^d \cdot X_{t-2} + \ldots + \varphi_p \cdot \Delta^d \cdot X_{t-p} + \varepsilon_t + \theta^1 \cdot \varepsilon_{t-1} + \theta^2 \cdot \varepsilon_{t-2} + \theta_q \cdot \varepsilon_{t-q} \quad (2)$$

X_t is time series at time t.c is constant term. $\varphi_i(i = 1, \ldots, p)$ are auto − regressive coefficients. $\Delta^d \cdot X_{t-i}(i = 1, \ldots, p)$ is the d-th difference series at time t-i. ε_t is error at time t. $\theta_j(j = 1, \ldots, q)$ are moving average coefficients. ARIMA (2,1,2) is recommended based on AICc value with significant coefficients ar2=0.8383, ma2=0.0625. Ljung-Box test results in p value of 0.4141, demonstrating that residuals are approximately white noise.

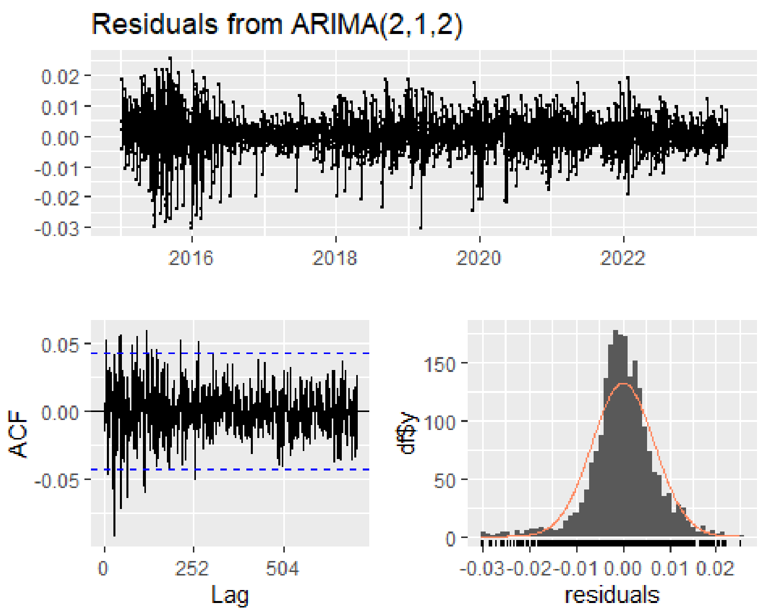

Figure 3. Residuals from ARIMA

DOI: 10.54254/2754-1169/2026.LD32298

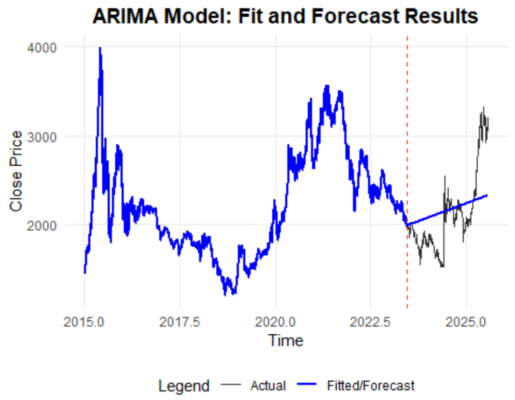

Figure 4. Residuals from ARIMA results

Figure 3 shows residuals of the fit model from ARIMA (2,1,2). Figure 4 shows a plot of actual value, fitted value and forecast value.

Table 1. Output of test data accuracy measurement

	ME	RMSE	MAE	MPE	MAPE
Test set	-47.00182	380.428	289.0071	-5.514229	13.64226

Table 1 shows the output of test data accuracy measurement.

2.3. STL-ETS-GARCH (-SNAIVE)

Decompose original data into trend, seasonal and remainder components.

2.3.1. STL decompose

It is computed that the seasonal ratio is 11.09%.

$$\text{Seasonal ratio} = \frac{\text{sd(seasonal component)}}{\text{sd(train component)}} \tag{3}$$

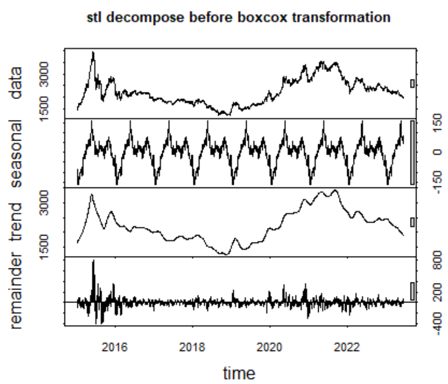

Figure 5. Boxcox transformation

Figure 5 shows a graph of trend, seasonal and remainder components after decomposition.

DOI: 10.54254/2754-1169/2026.LD32298

2.3.2. Trend modeling

ETS model literate error, trend and seasonal parameters through exponential smoothing [8]. The trend component is revealed to be slowly fluctuating and stable, which is compatible with ETS framework. The forecast package in R supports selecting ETS model automatically [9,10]. Box-Cox transformation is necessary for each portion before modelling independently. The model selected by the function is ETS(A,Ad,N) with smoothing parameters alpha=0.9999, beta=0.8561, phi=0.98. The model can be reflected by the following formula after the substitution of parameters:

$$\begin{gathered}\text{Observation Equation}: \quad y_t = l_{t-1} + 0.98b_{t-1} + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}\left(0{,}0.0016^2\right) \\ \text{Level}: l_t = 0.9999y_t + 0.0001\left(l_{t-1} + 0.98b_{t-1}\right) \\ \text{Damped Trend}: b_t = 0.8561\left(l_t - l_{t-1}\right) + 0.141022 \\ \text{Initial States}: l_0 = 12.3319, \quad b_0 = 0.0056 \\ h - \text{step Forecast}: \widehat{y}_{t+h|t} = l_t + 50b_t\left(1 - 0.98^h\right)\end{gathered} \tag{4}$$

Where y_t is the observed value at time t, l_t is the level component at time t, b_t is the damped trend component at time t. ϵ_t is the random error of the observation equation. h represents forward steps of the forecast.

2.3.3. Remainder modelling

The ARCH-LM test constructs the Lagrange multiplier statistic to determine whether autocorrelation exists in the conditional variance of the residuals, thereby verifying the significance of volatility clustering [11]. The residual is first tested for stationarity for ADF test, resulting a p value of 0.01, which confirms no difference is needed [12]. The Ljung-Box test(lag=10) on the squared residuals yields a p-value of 0, indicating significant ARCH effects. Thus, adopting the ARIMA-GARCH model captures both the mean trend and volatility clustering of the series [13]. Build ARIMA on the remainder segment. ARIMA (2,0,1) is selected with significant coefficients.

Construct sGARCH (1,1) – ARMA (2,1) model with specifications:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \epsilon_t + \theta_1 \epsilon_{t-1} \tag{5}$$

where y_t denotes the residual attamed, μ is the constant mean term ϕ_1, ϕ_2 are AR coefficients ϕ_1 is the MA coefficient and ϵ_t " is the innovation term.

$$\sigma_t^2 = \omega + \alpha_1 \epsilon_{t\text{-}1}^2 + \beta_1 \sigma_{t\text{-}1}^2 \tag{6}$$

where σ_t^2 is the conditional variance attimet, ω is the constant term, α_1 captures the" short-term impact on volatility and β_1 reflects the long-term persistence of volatility. The rug arch package in R supports opting GARCH model automatically [14]. The model is tested via the Ljung-

DOI: 10.54254/2754-1169/2026.LD32298

Box test(lag=10) on standardised residuals and squared standardised residuals. Diagnostic tests show no residual autocorrelation (Ljung-Box, lag=10, p=0.466) "

2.3.4. Seasonal term processing

For weak seasonality, two methods are applicable. The following will explain the two methods. Ignore seasonality by means of merging the seasonal component into the remainder. Build the model the same way as mentioned above while modeling remainder part. New remainder series are stationary via ADF test (p=0.01). The R function results in sGARCH (1,1) – ARIMA (1,1) model. Diagnostic tests show no residual autocorrelation (Ljung-Box, lag = 10, p = 0.7167468) or remaining ARCH effects (Ljung-Box on squared residuals=0.9846622). Assume that the future value of the seasonal term is equivalent to the most recent historical observation corresponding to the same seasonal period, which is the seasonal naive method [15].

2.4. Economic performance evaluation

To evaluate economic performance, calculate Directional Accuracy, Strategy Sharpe Ratio, Benchmark Sharpe Ratio, Excess Sharpe Profit/Loss Ratio, and Outperforms Benchmark of each model [16,17]. To assess model stability, introduce Directional Accuracy Volatility, Directional Accuracy Conditional Value at Risk (10%), Sharpe Ratio Volatility, and Sharpe Ratio Conditional Value at Risk (10%) of each model [18,19].

3. Results

For STL-ETS-GARCH(-SNAIVE) method, compare the real value with the fitted value and forecast value after integration and calculate the accuracy of the training set and test set separately [20,21].

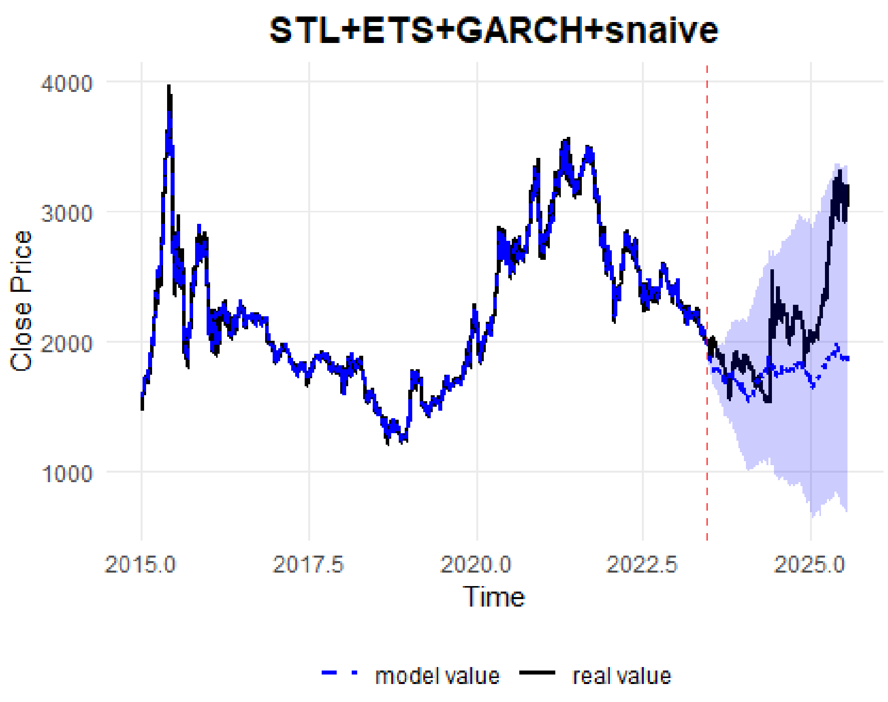

Figure 6. Model value

DOI: 10.54254/2754-1169/2026.LD32298

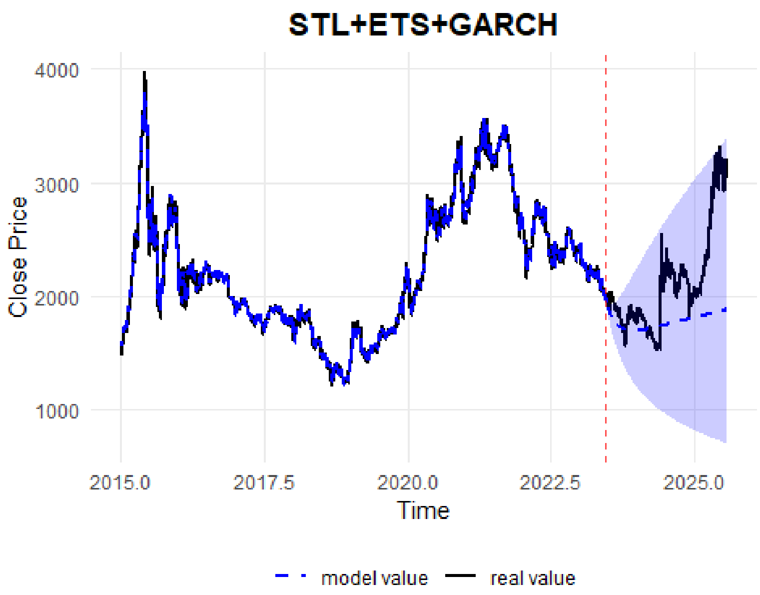

Figure 7. Real value

Figure 6 shows graph of fitted, forecast and actual value from STL-ETS-GARCH-SNAIVE model. Figure 7 shows graph of fitted, forecast and actual value from STL-ETS-GARCH model. All predicted values fall within the 80% confidence interval.

Table 2. Economic performance and stability comparison of competing models

Model	Directional Accuracy (%)	Sharpe Ratio	Profit–Loss Ratio	Max Drawdown (%)	Break-even Cost (bps)	Benchmark Sharpe	Excess Sharpe	Outperforms BH
ARIMA	48.7	0.775	1.22	-31.9	-1.8	0.664	0.111	Yes
STL–ETS–GARCH–SNAIVE	51.3	0.018	0.95	-53.7	1.8	0.664	-0.646	No
STL–ETS–GARCH	49.2	0.876	1.22	-32.5	-1.0	0.664	0.212	Yes

Table 3. Stability evaluation of forecasting models

Model	DA Volatility	DA CVaR (10%)	Sharpe Volatility	Sharpe CVaR (10%)
ARIMA	0.0984	0.2740	3.9328	-6.1529
STL–ETS–GARCH–SNAIVE	0.1189	0.2727	3.6411	-6.3675
STL–ETS–GARCH	0.0999	0.2759	3.8348	-5.5318

Table 2 and Table 3 shows the result of the STL-ETS-GARCH model performs the best in the Sharpe ratio and Excess Sharpe ratio. STL-ETS-GARCH-SNAIVE improves direction accuracy but drops sharply in Sharpe ratio, indicating that the seasonal naive method in weak seasonality introduces noise, which damages economic performance. Results of the traditional ARIMA model are acceptable as a benchmark, but not as good as the hybrid model. The STL-ETS-GARCH model

DOI: 10.54254/2754-1169/2026.LD32298

with optimal performance does not have the worst stability and performs well in CVaR value(Sharpe Ratio Conditional Value at Risk (10%)).

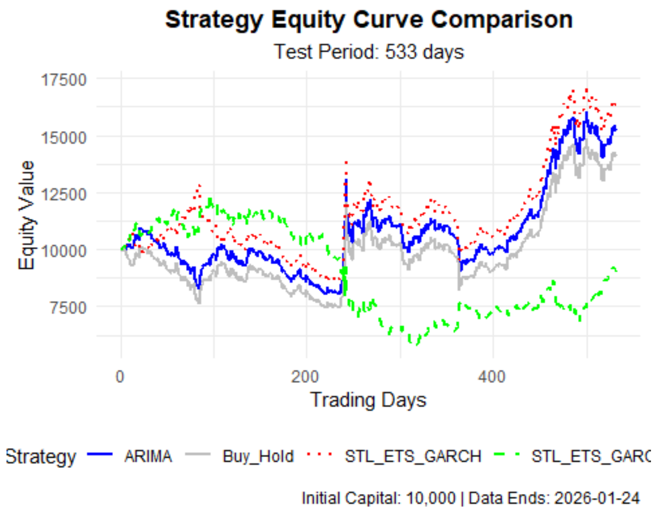

Figure 8. Strategy equity curve comparison

Figure 8 shows a graph of the strategy equity curve comparison.

4. Discussion

STL decomposition allows selecting the most suitable model separately based on the characteristics of each component. GARCH model controls residual volatility, which enhances the risk-adjusted return of strategy. From the STL-ETS-GARCH-SNAIVE model that can conclude that ignoring seasonal patterns is superior for weak seasonality. Separately modelling accumulates errors due to the addition of residuals from each model. This explains the high RMSE of both train and test data compared to traditional ARIMA model. However, the model is flexible and explainable. It captures more traits of the data and increases economic performance. Future studies can seek ways of reducing cumulative errors and combine LSTM models in forecasting the trend component. For close price of Growth Enterprise Market, STL-ETS-GARCH(-SNAIVE) hybrid model framework can generate trading strategies that significantly outperform the traditional ARIMA model in terms of economic performance. Meanwhile, research practice provides empirical warnings for dealing with weak seasonal components in the financial sequence.

5. Conclusion

This study compares a traditional ARIMA model with a hybrid STL–ETS–GARCH framework (with and without SNAIVE adjustment) in forecasting the closing price of the Growth Enterprise Market (ChiNext Index) from 2015 to 2025. Beyond conventional statistical accuracy evaluation, this paper incorporates economic performance and stability measures, including directional accuracy, Sharpe ratio, excess Sharpe ratio, profit–loss ratio, drawdown, and Conditional Value at Risk (CVaR), to assess the practical value of forecasting models. The results show that while the ARIMA (2,1,2) model provides reasonable statistical performance and well-behaved residual diagnostics, its economic performance remains moderate. Although ARIMA captures linear dependencies effectively, it fails to model volatility clustering and structural characteristics embedded in financial time series, limiting its risk-adjusted returns. In contrast, the STL–ETS–

DOI: 10.54254/2754-1169/2026.LD32298

GARCH hybrid model demonstrates superior economic effectiveness. By decomposing the original series and modeling the trend component via ETS and the remainder via ARIMA–GARCH, the framework captures both structural trend dynamics and volatility clustering. The model achieves the highest Sharpe ratio and excess Sharpe ratio in out-of-sample testing, indicating that volatility modeling significantly enhances risk-adjusted performance. The comparison between STL–ETS–GARCH and STL–ETS–GARCH–SNAIVE further highlights the importance of appropriate seasonal treatment. Although incorporating the seasonal naive method slightly improves directional accuracy, it substantially deteriorates the Sharpe ratio under weak seasonality. This finding suggests that imposing redundant seasonal structures may introduce noise and undermine economic outcomes, even when predictive accuracy appears marginally improved. Overall, the study contributes by integrating decomposition techniques with volatility modeling and by extending evaluation criteria from statistical metrics to economic and stability performance. The findings suggest that for financial series with weak seasonal patterns, careful handling—or even exclusion—of seasonal components is preferable. The STL–ETS–GARCH framework provides a more economically meaningful forecasting approach than the traditional ARIMA model, offering practical implications for quantitative trading and risk management in emerging equity markets.

References

[1] Li, S. (2024) Stock Closing Price Prediction Based on the ARIMA-GARCH Model. Frontiers in Business, Economics and Management.
[2] Rao, A., Sharma, G.D., Tiwari, A.K., Hossain, M.R. and Dev, D. (2025) Crude Oil Price Forecasting: Leveraging Machine Learning for Global Economic Stability. Technological Forecasting and Social Change, 216, 124133. https: //doi.org/10.1016/j.techfore.2025.124133
[3] Rubio, L., Palacio Pinedo, A., Mejía Castaño, A. and Others (2023) Forecasting Volatility by Using Wavelet Transform, ARIMA and GARCH Models. Eurasian Economic Review, 13, 803-830. https: //doi.org/10.1007/s40822-023-00243-x
[4] Zhang, Y., Li, J. and Wang, H. (2023) Hybrid Time Series Decomposition and Deep Learning Model for Stock Market Forecasting. Expert Systems with Applications, 213, 119095.
[5] Kim, S. and Kim, H. (2022) Forecast Evaluation and Economic Value of Volatility Models in Equity Markets. Journal of Empirical Finance, 66, 1-18.
[6] Apostolou, N. and Charalambous, C. (2022) Volatility Modeling in Emerging Stock Markets Using GARCH-Type Models. Finance Research Letters, 48, 102921.
[7] Patton, A.J. and Weller, B.M. (2020) What You See Is Not What You Get: The Costs of Trading Market Anomalies. Journal of Financial Economics, 137, 515-549. https: //doi.org/10.1016/j.jfineco.2020.02.012
[8] Zahrani, A., Rohmawati, A. and Saadah, S. (2021) STL Decomposition and SARIMA Model: The Case for Estimating Value-at-Risk of Covid-19 Increment Rate. International Journal on Information and Communication Technology, 7, 1-10.
[9] Hyndman, R.J., Koehler, A.B., Snyder, R.D. and Grose, S. (2002) A State Space Framework for Automatic Forecasting Using Exponential Smoothing Methods. International Journal of Forecasting, 18, 439-454.
[10] Hyndman, R.J. and Khandakar, Y. (2008) Automatic Time Series Forecasting: The Forecast Package for R. Journal of Statistical Software, 27, 1-22.
[11] Hyndman, R.J. and Athanasopoulos, G. (2021) Forecasting: Principles and Practice. 3rd Edition, OTexts.
[12] Box, G.E.P. and Jenkins, G.M. (2010) Time Series Analysis: Forecasting and Control. Wiley.
[13] Liu, H. and Shi, J. (2013) Applying ARMA–GARCH Approaches to Forecasting Short-Term Electricity Prices. Energy Economics, 37, 152-166.
[14] Pesaran, M.H. and Timmermann, A. (1992) A Simple Nonparametric Test of Predictive Performance. Journal of Business & Economic Statistics, 10, 461-465.
[15] Kourtis, A. (2016) The Sharpe Ratio of Estimated Efficient Portfolios. Finance Research Letters, 17, 72-78.
[16] Rockafellar, R.T. and Uryasev, S. (2002) Conditional Value-at-Risk for General Loss Distributions. Journal of Banking & Finance, 26, 1443-1471.

DOI: 10.54254/2754-1169/2026.LD32298

[17] Park, C.H. and Irwin, S.H. (2007) What Do We Know about the Profitability of Technical Analysis? Journal of Economic Surveys, 21, 786-826.
[18] Bouri, E., Shahzad, S.J.H. and Roubaud, D. (2021) Co-Movement and Risk Spillovers in Stock Markets: Evidence from Emerging Economies. Research in International Business and Finance, 58, 101453.
[19] Lahmiri, S. and Bekiros, S. (2022) Hybrid ARIMA and Machine Learning Models for Financial Forecasting: A Comparative Study. Chaos, Solitons & Fractals, 154, 111644.
[20] Fang, Y., Huang, W. and Wu, Y. (2023) Volatility Forecasting in Emerging Markets Using Hybrid Decomposition Models. Finance Research Letters, 52, 103575.
[21] Ahmed, W.M.A., Bouri, E. and Tiwari, A.K. (2024) Forecasting Financial Time Series Using Hybrid Volatility Models. International Review of Financial Analysis, 91, 102857.