

DO GARCH MODEL IMPROVE FORECASTS OF MALAYSIAN LOG PRODUCTION? A COMPARATIVE ARIMA AND GARCH ANALYSIS

Tengku Mardhiah Tengku Jalal^{1*}

Norshaieda Abdullah²

Noor Ilanie Nordin³

Mohd Faiez Suhaimin⁴

Muhammad Irfan Haqiem Hairan⁵

Wabeel Luqman Shahmiral⁶

¹ Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Cawangan Kelantan, Machang, Kelantan, Malaysia (E-mail: tengkum908@uitm.edu.my)

² Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Cawangan Kelantan, Machang, Kelantan, Malaysia (E-mail: norsha052@uitm.edu.my)

³ Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Cawangan Kelantan, Machang, Kelantan, Malaysia (E-mail: ilanie@uitm.edu.my)

⁴ Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Cawangan Kelantan, Machang, Kelantan, Malaysia (E-mail: mdfaiez821@uitm.edu.my)

⁵ Human Capital Development Division, Public Service Department, Government Complex C, Putrajaya, Malaysia (E-mail: irfanhaqiem42@gmail.com)

⁶ Zikay Factoring Sdn. Bhd., Chow Kit, Kuala Lumpur, Malaysia (E-mail: wabeelluqman@gmail.com)

Article history

Received date : 10-9-2026

Revised date : 11-9-2026

Accepted date : 17-10-2026

Published date : 25-10-2026

To cite this document:

Tengku Jalal, T. M., Abdullah, N., Nordin, N. I., Suhaimin, M. F., Hairan, M. I. H., & Shahmiral, W. L. (2026). Do GARCH model improve forecasts of Malaysian log production? A comparative ARIMA and GARCH analysis. *Journal of Islamic, Social, Economics and Development (JISED)*, 11 (87), 1143-1151.

Abstract: *Reliable forecasts of log production are important for timber-supply planning, sustainable forest management. This study evaluates the forecasting performance of conventional conditional-mean models against conditional-variance models for Malaysian logs production, a series characterized by non-stationarity and distinct structural changes. Utilizing annual production data from 1947 to 2021, the research methodologically compares the out-of-sample accuracy of various autoregressive integrated moving average (ARIMA) specifications with generalized autoregressive conditional heteroskedasticity (GARCH) type models. The models were estimated using an in-sample period from 1947 to 2002, reserving 2003 to 2021 as a chronological holdout sample for evaluation based on RMSE, MAE, and MAPE metrics. The results demonstrate that the ARIMA(0,2,1) model outperforms GARCH(1,1) specification. These findings provide a substantial forecasting advantage, offering forestry stakeholders a robust analytical tool for sustainable resource planning.*

Keywords: *ARIMA; GARCH; log production; Malaysia; time-series forecasting; volatility; forestry*

Introduction

The Malaysian timber industry depends on a reliable and sustainable supply of logs for downstream activities such as sawmilling, plywood manufacturing, furniture production, and other wood-based industries. The Malaysia Policy on Forestry recognises the importance of controlling timber production from natural and planted forests in accordance with sustainable production capacity and industrial requirements (Ministry of Energy and Natural Resources, 2022). Reliable forecasts of log production are therefore necessary to support production planning, raw-material management, policy formulation, and environmental conservation.

Malaysia's log production has experienced a substantial long-term transformation. The annual volume increased from 138,000 m³ in 1947 to a peak of approximately 43.51 million m³ in 1992 before declining to approximately 5.58 million m³ in 2021. Contemporary Malaysian timber-production data are collected from establishments registered with the relevant forestry authorities. The official definition of timber logs excludes rubberwood logs, poles, coconut trunks, and oil-palm trunks (Department of Statistics Malaysia [DOSM], 2024). Earlier research also documented a declining trend in log production and exports in Sarawak between 1997 and 2006 (Ismail & Zaki, 2010).

The substantial increase followed by a prolonged decline creates a challenging forecasting problem. The series is non-stationary, contains large changes in magnitude, and may exhibit periods of different volatility. Conventional autoregressive integrated moving average (ARIMA) models are designed primarily to model the conditional mean of a series. By contrast, generalised autoregressive conditional heteroskedasticity (GARCH) models capture changes in conditional variance after an appropriate mean equation has been specified (Bollerslev, 1986; Engle, 1982). Therefore, ARIMA and GARCH should not be treated as statistically identical or interchangeable forecasting techniques.

Previous studies have demonstrated the applicability of ARIMA models to forestry-related data. Acquah et al. (2014) applied a seasonal ARIMA model to monthly wood-product exports in Ghana, whereas Huda et al. (2023) identified ARIMA(0,2,1) as an appropriate model for annual Indonesian log production. For Malaysia, Tengku Jalal et al. (2025) compared several Box–Jenkins specifications and concluded that ARIMA(0,2,1) produced the most accurate forecasts. However, the earlier Malaysian study focused only on ARIMA models and did not examine whether conditional-variance modelling could improve forecast performance.

The present study extends the previous ARIMA analysis by comparing ARIMA with GARCH models. This study aims to identify the best-performing ARIMA and GARCH models. The study contributes by providing a methodologically consistent comparison between conventional conditional-mean modelling and models that jointly represent conditional mean and variance.

Literature Review

Trends in Logs Production

Log production represents an important upstream component of the timber industry. Changes in log supply may influence processing capacity, exports, employment, and the availability of raw materials for downstream industries. At the same time, timber production must be balanced against forest conservation and sustainable harvesting requirements. Malaysia's forestry policy consequently emphasises the monitoring of natural- and planted-forest production and the development of alternative timber supplies (Ministry of Energy and Natural Resources, 2022).

Ismail and Zaki (2010) examined log production and exports in the major producing divisions of Sarawak. Their findings indicated a declining trend in production from both hill and peat-swamp forests between 1997 and 2006. At the national level, the production series similarly increased until the early 1990s before entering a prolonged downward phase. However, a univariate production series cannot identify the causal effects of forest policy, timber demand, economic conditions, harvesting restrictions, or changes in forest area. These factors can only be offered as possible contextual explanations.

The long-term pattern should also not be automatically classified as a cycle. A cyclical component refers to recurrent increases and decreases occurring over non-fixed periods. The Malaysian series contains one dominant expansion phase followed by an extended contraction. It may therefore represent a structural change in trend rather than a repeatedly observed economic cycle.

ARIMA models

The Box–Jenkins approach combines autoregressive, differencing, and moving-average components to represent the serial dependence of a time series (Box et al., 2015). An ARIMA(p, d, q) model may be written as

$$\phi(B)(1 - B)^d y_t = c + \theta(B)\varepsilon_t$$

where p is the autoregressive order, d is the differencing order, q is the moving-average order, B is the backshift operator, and ε_t is a white-noise innovation.

ARIMA models have been applied successfully to several types of forestry data. Acquah et al. (2014) used a seasonal ARIMA model for monthly wood-product exports from Ghana. Huda et al. (2023) found that ARIMA(0,2,1) adequately represented annual Indonesian log production. Tengku Jalal et al. (2025) subsequently identified the same ARIMA order for Malaysian log production based on out-of-sample RMSE, MAE, and MAPE.

Model identification should not depend solely on ACF and PACF plots. Stationarity tests, information criteria, residual diagnostics, and out-of-sample forecast accuracy should be considered together. Excessive differencing can generate unstable extrapolations and unnecessarily increase forecast uncertainty (Hyndman & Athanasopoulos, 2021). This concern is particularly relevant when an ARIMA model fitted to a sharply declining series eventually produces negative forecasts.

GARCH

Engle (1982) developed the ARCH model to capture conditional variance that depends on earlier squared innovations. Bollerslev (1986) generalised the model by allowing variance to depend on both past squared innovations and earlier conditional variances. A GARCH(1,1) model is expressed as

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}$$

where σ_t^2 denotes conditional variance, $\omega > 0$, $\alpha \geq 0$, and $\beta \geq 0$. The condition $\alpha + \beta < 1$ is commonly imposed to ensure covariance stationarity.

GARCH models do not directly forecast the level of production. They model the variance of innovations around a conditional-mean equation. Therefore, a GARCH comparison requires a

suitable mean specification, such as an ARMA model for logarithmic production changes. The predicted changes must then be accumulated and transformed back to the original production scale.

Forecast evaluation

Models should be compared using data that were not used during estimation. A chronological holdout sample maintains the temporal ordering of observations, while rolling-origin evaluation can provide additional evidence across multiple forecast origins (Bergmeir et al., 2018). Forecast accuracy may be evaluated using MAE, RMSE, and MAPE. RMSE places greater weight on large errors, MAE describes the average absolute error in the original measurement unit, and MAPE expresses the average error relative to the actual observations.

Hyndman and Koehler (2006) highlighted several limitations of percentage-error measures and recommended scaled measures when comparisons involve different series. Nevertheless, MAPE remains interpretable in the current study because all actual production observations are positive and all models are evaluated using the same series and period.

Methodology

Data

The study used 75 annual observations of Malaysian log production from 1947 to 2021, measured in cubic metres. The first 56 observations, covering 1947–2002, were used for model estimation. The remaining 19 observations from 2003 to 2021 were retained as an out-of-sample evaluation set. All models were estimated using information available only up to 2002. Nineteen-step-ahead forecasts were then generated and compared with the actual production recorded from 2003 to 2021. This ensured that all models were evaluated over the same years and on the same production scale.

Exploratory analysis and stationarity

The temporal pattern was initially examined using a time plot, ACF, and PACF. Stationarity was assessed using the augmented Dickey–Fuller test. The test examines null hypothesis (the series contains a unit root) against alternative hypothesis (The series is stationary).

For the GARCH analysis, proportional changes in production were represented through annual simple returns:

$$r_t = \frac{y_t - y_{t-1}}{y_{t-1}}$$

This transformation prevented differences in the magnitude of production from dominating the variance equation. It also allowed GARCH parameters to represent changes in proportional volatility.

ARIMA modelling

Based on the order of differencing and the observations of the ACF and PACF plot, Seven candidate models were evaluated: ARIMA(2,2,2), ARIMA(2,2,1), ARIMA(1,2,1), ARIMA(0,2,2), ARIMA(0,2,1), ARIMA(1,2,0), and ARIMA(2,2,0).

AIC and BIC were used to assess in-sample parsimony. Residual independence was evaluated using the Ljung–Box test (Ljung & Box, 1978). Final selection emphasised by out-sample RMSE, MAE, and MAPE.

GARCH modelling

For GARCH(1,1), the return forecasts were accumulated and transformed back into forecasts of actual log-production values:

$$y_t = y_{t-1} + r_t \times y_{t-1}$$

Then, the out-sample MAE, RMSE and MAPE were assessed and compared with the best ARIMA model.

The application of GARCH to annual production data also requires caution. GARCH models are normally most informative when a sufficiently long series contains clear volatility clustering. Annual data from 1947 to 2021 provide only 75 level observations and 74 logarithmic changes. The relatively small sample may limit the stability of asymmetric GARCH parameters.

Accuracy measures

Forecast performance was evaluated using:

$$\text{MAE} = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}$$

and

$$\text{MAPE} = \frac{100}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right|$$

The model with consistently lower out-of-sample errors was regarded as the preferred forecasting model.

Analysis and Discussion

Temporal behaviour

Figure 1 shows the time plot of Malaysia's logs production over 1947 to 2021. Malaysian log production increased from 138,000 m³ in 1947 to a peak of 43.51 million m³ in 1992. Production subsequently declined to 5.58 million m³ in 2021, representing an approximate reduction of 87.2% from its peak.

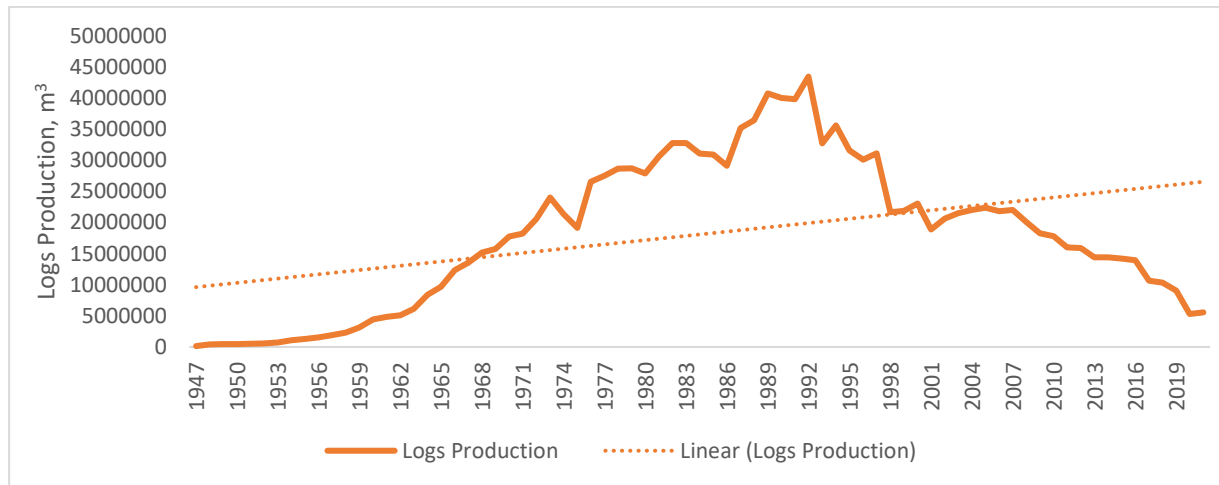


Figure 1: Time plot of Malaysia's Logs Production with trend line

Therefore, the series should be characterised as having an initial expansion followed by a prolonged contraction. It is inaccurate to describe the complete series as having an overall increasing trend. Moreover, because the data are annual, a seasonal component cannot be estimated.

Selection of the ARIMA model

The first-differenced production series produced an ADF statistic of -3.4508 and $p = 0.0571$. At the 5% significance level, the null hypothesis could not be rejected. After second differencing, the ADF statistic was -7.4625, with $p = 0.010$, supporting stationarity. Consequently, the candidate ARIMA models used $d = 2$.

Table 1 presents the out-of-sample results for the seven ARIMA models.

Table 1: Performance of Candidate ARIMA Models

Model	AIC	BIC	Ljung–Box p	RMSE	MAE	MAPE (%)
ARIMA(2,2,2)	1760.59	1770.53	0.9736	9,260,943	3,118,941	73.48
ARIMA(2,2,1)	1764.92	1772.87	0.8310	9,830,336	8,941,282	78.36
ARIMA(1,2,1)	1763.57	1769.54	0.7347	8,675,013	8,412,103	68.32
ARIMA(0,2,2)	1763.54	1769.50	0.7377	10,511,800	10,117,772	84.23
ARIMA(0,2,1)	1767.02	1770.99	0.1373	2,479,976	2,146,219	16.57
ARIMA(1,2,0)	1783.14	1787.12	0.0113	2,870,900	2,510,802	19.19
ARIMA(2,2,0)	1767.54	1773.50	0.4601	3,076,462	2,790,404	18.68

ARIMA(2,2,2) had the lowest AIC, whereas ARIMA(0,2,2) had the lowest BIC. However, ARIMA(0,2,1) produced the lowest out-of-sample RMSE, MAE, and MAPE. Given the forecasting objective, ARIMA(0,2,1) was selected as the best ARIMA specification.

ARIMA(1,2,0) failed the Ljung–Box diagnostic at the 5% significance level because $p = 0.0113$. It should therefore not be classified as having white-noise residuals.

ARIMA versus GARCH

Figure 2 shows that the stationary data of logs production are relatively skewed to the left. Furthermore, kurtosis value for the stationary data was also calculated where the value obtained

is 5.600948. Since the stationary data of logs production are skewed and the kurtosis value is more than three, the GARCH models can be applied for the data series.

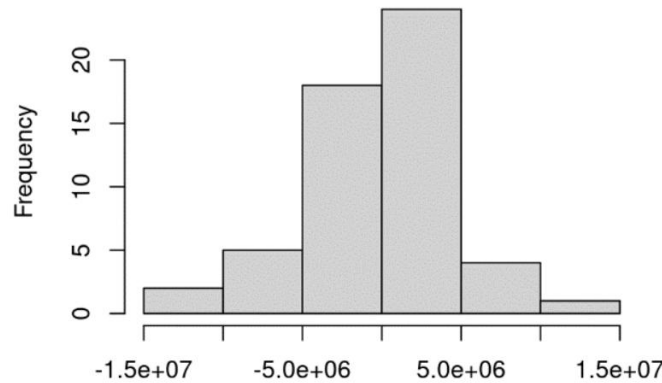


Figure 2: Histogram for stationary data of Malaysia Logs Production

Table 2 compares the best ARIMA specification with the GARCH(1,1) specification. The results clearly indicate that the ARIMA(0,2,1) model outperformed the GARCH(1,1) model across all three accuracy metrics. The introduction of the GARCH(1,1) specification substantially increased the forecasting errors, yielding a much higher RMSE (6,723,894.80), MAE (5,269,540.10), and MAPE (25.45%) compared to the ARIMA model. Therefore, incorporating a time-varying conditional variance did not improve, but rather deteriorated, the point forecasts of Malaysian log production.

Table 2: Comparison of the Best ARIMA and GARCH-Type Models

Model	RMSE	MAE	MAPE (%)
ARIMA(0,2,1)	2,479,976.49	2,146,219.11	16.57
GARCH(1,1)	6,723,894.80	5,269,540.10	25.45

The superiority of the ARIMA(0,2,1) model can be explained by the primary statistical characteristics of the production series. Malaysian log production is heavily dominated by major shifts in its conditional mean and a strong long-term trend (an initial expansion followed by a prolonged contraction). The poor out-of-sample performance of the GARCH model suggests that persistent short-term volatility clustering, or ARCH effects, are either weak or entirely absent in this annual series. GARCH models are designed principally to capture dynamic conditional variance; when the data is relatively homoskedastic, estimating additional variance parameters unnecessarily increases model complexity (violating the principle of parsimony) without providing any predictive.

However, this result does not imply that ARIMA models are universally superior to GARCH models for all forms of production data. Rather, it demonstrates that for this particular annual production series and evaluation period, conventional conditional-mean modelling was highly adequate, and accommodating conditional heteroskedasticity provided no forecasting benefit.

Conclusion

This study evaluated the forecasting performance of conventional conditional-mean models against conditional-variance models using Malaysian log production data spanning from 1947 to 2021. Exploratory analysis revealed that the production series is non-stationary,

characterising a dominant expansion phase that peaked at approximately 43.51 million cubic metres in 1992, followed by a prolonged contraction phase down to 5.58 million cubic metres in 2021.

To address the non-stationarity, the series required second differencing, which informed the evaluation of seven candidate autoregressive integrated moving average (ARIMA) models. Through a methodologically consistent out-of-sample evaluation utilizing a holdout period from 2003 to 2021, the ARIMA(0,2,1) specification was identified as the most accurate. This model minimized forecasting errors, producing a Root Mean Square Error (RMSE) of 2,479,976.49, a Mean Absolute Error (MAE) of 2,146,219.11, and a Mean Absolute Percentage Error (MAPE) of 16.57%.

While the stationary data exhibited a left skew and a kurtosis value of 5.600948—which theoretically justified the application of generalized autoregressive conditional heteroskedasticity (GARCH) models—the GARCH(1,1) specification substantially deteriorated forecast accuracy. The ARIMA(0,2,1) model outperformed the GARCH(1,1) model across all evaluated accuracy metrics. Incorporating a time-varying conditional variance increased the forecasting errors considerably, yielding a much higher RMSE of 6,723,894.80, a MAE of 5,269,540.10, and a MAPE of 25.45% compared to the ARIMA model.

The superiority of the ARIMA(0,2,1) model indicates that Malaysian log production is primarily dominated by major shifts in its conditional mean and a strong long-term trend, rather than by persistent short-term volatility clustering or ARCH effects. Applying GARCH models to relatively homoskedastic data unnecessarily increases model complexity, which violates the principle of parsimony without providing any predictive advantage. Therefore, for this specific annual production series and evaluation period, conventional conditional-mean modelling proved highly adequate, demonstrating that accommodating conditional heteroskedasticity offers no forecasting benefit for this particular forestry. Furthermore, Hwang and Valls Pereira (2006) provide a foundational Monte Carlo study demonstrating how MLE estimators perform in small samples. They found that standard GARCH(1,1) estimations tend to be negatively biased with limited data, consistently underestimating volatility persistence.

References

- Acquah, S. B., Avuglah, R. K., & Harris, E. (2014). Time series behavior of the volume of wood products export in Ghana. *International Journal of Statistics and Probability*, 3(4), 12–24. <https://doi.org/10.5539/ijsp.v3n4p12>
- Bergmeir, C., Hyndman, R. J., & Koo, B. (2018). A note on the validity of cross-validation for evaluating autoregressive time series prediction. *Computational Statistics & Data Analysis*, 120, 70–83. <https://doi.org/10.1016/j.csda.2017.11.003>
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327. [https://doi.org/10.1016/0304-4076\(86\)90063-1](https://doi.org/10.1016/0304-4076(86)90063-1)
- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5th ed.). Wiley.
- Department of Statistics Malaysia. (2024, September 1). Production of major timber products by state. OpenDOSM. https://open.dosm.gov.my/data-catalogue/timber_production
- Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366), 427–431. <https://doi.org/10.1080/01621459.1979.10482531>
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), 987–1007. <https://doi.org/10.2307/1912773>
- Huda, K., Wibowo, C. E., & Gunawan, V. (2023). Implementation of time series forecasting with Box–Jenkins ARIMA method on wood production of Indonesian forests. *AIP Conference Proceedings*, 2738(1), Article 060004. <https://doi.org/10.1063/5.0141187>
- Hwang, S., & Valls Pereira, P. L. (2006). Small sample properties of GARCH estimates and persistence. *The European Journal of Finance*, 12, 473–494. <https://doi.org/10.1080/13518470500039436>
- Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts. <https://otexts.com/fpp3/>
- Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679–688. <https://doi.org/10.1016/j.ijforecast.2006.03.001>
- Ismail, M. H., & Zaki, P. H. (2010). A study on trend of logs production and export in the state of Sarawak, Malaysia. *International Journal of Marketing Studies*, 2(1), 92–95. <https://doi.org/10.5539/ijms.v2n1p92>
- Ljung, G. M., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297–303. <https://doi.org/10.1093/biomet/65.2.297>
- Ministry of Energy and Natural Resources. (2022). *Malaysia policy on forestry (Version 2.0)*. Government of Malaysia. [https://www.nres.gov.my/ms-my/pustakamedia/Penerbitan/Malaysia%20Policy%20on%20Forestry%20\(Ver%202.0\).pdf](https://www.nres.gov.my/ms-my/pustakamedia/Penerbitan/Malaysia%20Policy%20on%20Forestry%20(Ver%202.0).pdf)
- Tengku Jalal, T. M. T., Abdullah, N., Hairan, M. I. H., & Shahmiral, M. W. L. (2025). Modelling Malaysia production of logs using Box–Jenkins model. *Journal of Mathematics and Computing Science*, 11(2), 151–160. <https://doi.org/10.24191/jmcs.v11i2.9529>