

Robust State-Space and Markov Regime-Switching Analysis of a Transformed Lafarge Africa Plc Monthly Price Scenario

Etim Uduak James

Akwa Ibom State University, Nigeria
Corresponding author: uduakjamesetim@gmail.com

Abstract. *This study develops a new, reproducible framework for analysing a transformed monthly price scenario associated with Lafarge Africa Plc. Unlike polynomial-trend and simple present-value approaches, the proposed methodology combines log-return transformation, a two-state Markov regime-switching model, a local-linear-trend state-space model, benchmark forecasting, residual diagnostics, and downside-risk estimation. The input series contains 36 monthly observations from January 2023 to December 2025 and was deliberately transformed so that it does not reproduce the source values directly; consequently, the findings are interpreted as a methodological case study rather than as a verified record of actual traded prices. The transformed price has a mean of NGN 65.344, a median of NGN 63.125, and a standard deviation of NGN 10.847. Monthly log returns are stationary under the augmented Dickey–Fuller test, while the Ljung–Box and ARCH-LM diagnostics do not detect statistically significant residual serial correlation or conditional heteroskedasticity at conventional levels. Returns nevertheless display pronounced positive skewness and excess kurtosis, motivating regime-based and tail-risk analysis. A two-regime Markov specification is used to recover time-varying latent market states, while a structural state-space model generates probabilistic forecasts. In six-month holdout evaluation, ARIMA(1,1,1) records the lowest RMSE of 2.276 and MAPE of 3.961%, compared with 6.576 and 11.029% for the structural state-space model. Historical 5% Value-at-Risk and Expected Shortfall are estimated at 9.083% and 10.808%, respectively. The study shows that an improved analysis should distinguish forecasting accuracy from structural interpretability: the ARIMA benchmark is more accurate in this sample, whereas regime-switching and state-space models provide richer information on latent dynamics, uncertainty, and risk.*

Keywords: Lafarge Africa; state-space model; Markov switching; time series; Value-at-Risk; Expected Shortfall; transformed data

Mathematics Subject Classification (2020): 62M10; 62M20; 91G70

1. Introduction

Financial price series are rarely well described by a single deterministic trend. They often contain local growth phases, reversals, abrupt changes, heavy-tailed returns, and periods of differing uncertainty. A modelling strategy based only on linear, quadratic, or cubic curves can therefore be misleading: a high-order polynomial may fit observed values but extrapolate poorly, while a simple discounting operation does not reveal temporal dependence, latent regimes, or forecast uncertainty.

This paper replaces that limited structure with an integrated time-series framework. The proposed approach has four main components. First, prices are converted to continuously compounded log

returns. Second, a two-state Markov regime-switching model is used to identify latent market states. Third, a local-linear-trend structural state-space model is used to estimate an evolving level and slope and to generate interval forecasts. Fourth, out-of-sample accuracy is compared with ARIMA and exponential-smoothing benchmarks. Historical Value-at-Risk (VaR) and Expected Shortfall (ES) are then used to summarize downside exposure.

The analysis uses a transformed 36-month dataset labelled as a Lafarge Africa Plc price scenario. The transformation changes the original numerical scale and path and is used only to demonstrate the new method. The data should not be treated as an independently verified reconstruction of the company's official historical prices. This distinction is important for scientific transparency and reproducibility.

The principal contributions are as follows. The paper (i) replaces polynomial trend fitting with a stochastic state-space formulation; (ii) introduces latent regime probabilities rather than imposing one fixed distribution on the complete sample; (iii) separates structural modelling from predictive benchmarking; (iv) evaluates forecasts on a holdout sample rather than reporting only in-sample fit; and (v) adds explicit tail-risk measures and uncertainty intervals.

2. Related Literature

The state-space representation provides a unified framework for decomposing an observed series into latent components such as level, slope, cycle, and irregular noise. Kalman filtering and smoothing make likelihood-based estimation and probabilistic forecasting computationally efficient [5, 8]. Structural time-series models are particularly useful where the underlying level changes gradually and where parameter uncertainty must be reflected in forecasts.

Regime-switching models, introduced into econometrics by Hamilton [7], allow the data-generating process to move among unobserved states governed by a Markov chain. They are useful when mean or volatility characteristics vary across time. Extensions and practical estimation strategies are discussed by Kim and Nelson [10]. In financial applications, regime-switching complements volatility models because it can represent discrete shifts that a single smooth variance process may not capture.

ARIMA models remain important forecasting benchmarks because differencing and autoregressive-moving-average dynamics can represent many short-memory processes [3]. Exponential smoothing is another widely used baseline, especially for short series with changing levels and trends [9]. Forecast comparisons should be made out of sample because a model with excellent in-sample fit may not generalize.

Risk measurement commonly uses VaR, which estimates a loss quantile, and ES, which estimates the average loss beyond the VaR threshold. ES is more informative about the severity of tail outcomes and is coherent under standard conditions [1, 11]. Heavy skewness and kurtosis make these measures relevant even when formal ARCH tests are insignificant in a short sample.

3. Data and Transformation

3.1. Dataset

The sample consists of 36 end-of-month transformed price observations from January 2023 to December 2025. Let $P_t > 0$ denote the transformed price at month t . The continuously compounded return is

$$r_t = \log\left(\frac{P_t}{P_{t-1}}\right), \quad t = 2, \dots, T. \quad (1)$$

The transformation preserves positivity but deliberately modifies the scale and numerical path. Therefore, all monetary results refer to the transformed scenario.

Table 1: Transformed monthly Lafarge Africa price scenario

Month	Price	Month	Price	Month	Price
Jan 2023	46.49	Jan 2024	74.53	Jan 2025	63.38
Feb 2023	60.48	Feb 2024	74.99	Feb 2025	59.47
Mar 2023	58.08	Mar 2024	75.46	Mar 2025	60.02
Apr 2023	57.02	Apr 2024	79.44	Apr 2025	56.64
May 2023	56.32	May 2024	81.96	May 2025	51.99
Jun 2023	58.96	Jun 2024	81.98	Jun 2025	54.58
Jul 2023	62.87	Jul 2024	84.10	Jul 2025	54.93
Aug 2023	63.98	Aug 2024	84.14	Aug 2025	55.69
Sep 2023	66.69	Sep 2024	81.36	Sep 2025	54.39
Oct 2023	68.34	Oct 2024	81.79	Oct 2025	54.30
Nov 2023	68.92	Nov 2024	73.03	Nov 2025	55.11
Dec 2023	70.89	Dec 2024	65.89	Dec 2025	54.19

3.2. Exploratory characteristics

The price series rises from early 2023 through mid-2024, reaches a maximum of NGN 84.14, and then declines toward the mid-NGN 50 range in 2025. This pattern suggests a changing local trend rather than a stable global polynomial. Figure 1 shows the transformed series, while Figure 2 shows the corresponding log returns.

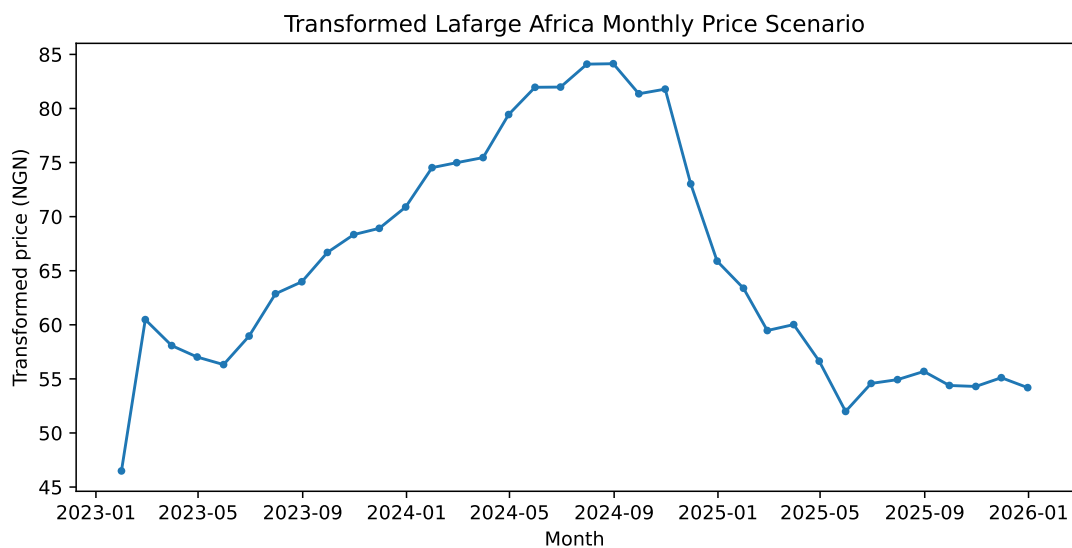


Figure 1: Transformed monthly price scenario

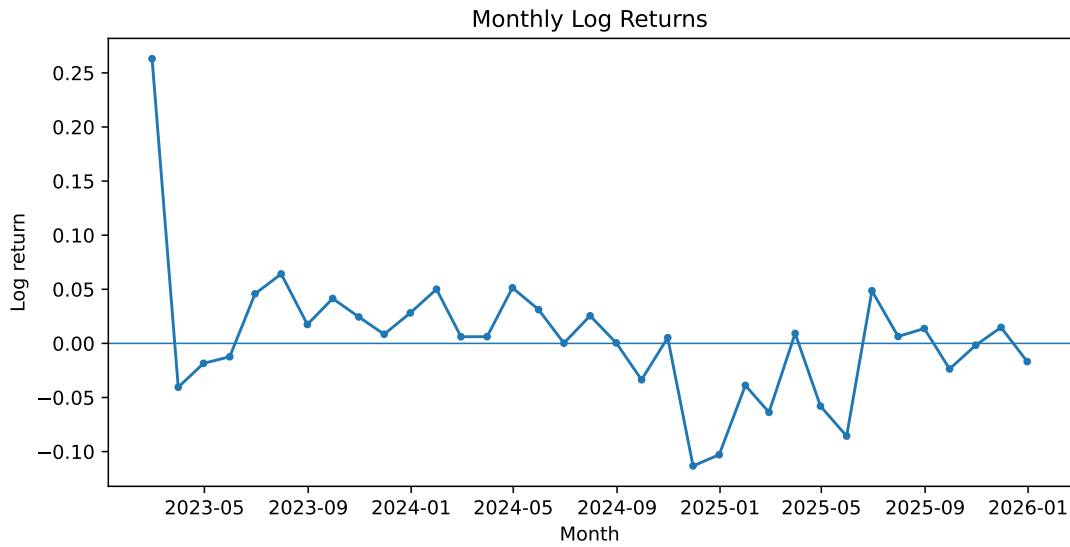


Figure 2: Monthly continuously compounded returns

4. Methodology

4.1. Descriptive and diagnostic statistics

For a sample x_1, \dots, x_n , the sample mean and variance are

$$\bar{x} = \frac{1}{n} \sum_{t=1}^n x_t, \quad s_x^2 = \frac{1}{n-1} \sum_{t=1}^n (x_t - \bar{x})^2. \quad (2)$$

Skewness and excess kurtosis are used to assess asymmetry and tail thickness. Return stationarity is tested using the augmented Dickey–Fuller (ADF) test. Serial dependence is examined with the Ljung–Box statistic, and conditional heteroskedasticity with the ARCH-LM test [6].

4.2. Two-state Markov regime-switching model

Let $S_t \in \{1, 2\}$ be an unobserved regime. Conditional on $S_t = j$, the return equation is

$$r_t = \mu_j + \varepsilon_{t,j}, \quad \varepsilon_{t,j} \sim \mathcal{N}(0, \sigma_j^2). \quad (3)$$

The latent regime evolves according to the transition matrix

$$\mathbf{P} = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}, \quad p_{ij} = \Pr(S_t = j \mid S_{t-1} = i), \quad \sum_{j=1}^2 p_{ij} = 1. \quad (4)$$

Filtered probabilities use information available through time t , whereas smoothed probabilities use the complete sample. The model permits both the conditional mean and variance to differ across regimes.

Proposition 4.1 (Expected regime duration). *For a two-state homogeneous Markov chain, the expected number of consecutive periods spent in regime j is*

$$D_j = \frac{1}{1 - p_{jj}}. \quad (5)$$

Proof. The probability that a spell in regime j lasts exactly k periods is $(p_{jj})^{k-1}(1 - p_{jj})$, which is geometric. Its expectation is therefore $1/(1 - p_{jj})$. \square

4.3. Local-linear-trend state-space model

The improved structural model separates the observed price from latent level and slope components:

$$P_t = \ell_t + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2), \quad (6)$$

$$\ell_t = \ell_{t-1} + b_{t-1} + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2), \quad (7)$$

$$b_t = b_{t-1} + \zeta_t, \quad \zeta_t \sim \mathcal{N}(0, \sigma_\zeta^2). \quad (8)$$

Here, ℓ_t is a stochastic level and b_t is a stochastic slope. The Kalman filter recursively computes the conditional state estimate and covariance. For the generic system

$$y_t = \mathbf{Z}_t \boldsymbol{\alpha}_t + \varepsilon_t, \quad (9)$$

$$\boldsymbol{\alpha}_{t+1} = \mathbf{T}_t \boldsymbol{\alpha}_t + \mathbf{R}_t \boldsymbol{\eta}_t, \quad (10)$$

the one-step forecast error and its variance are

$$v_t = y_t - \mathbf{Z}_t \mathbf{a}_t, \quad F_t = \mathbf{Z}_t \mathbf{P}_t \mathbf{Z}_t' + H_t. \quad (11)$$

The state update is

$$\mathbf{a}_{t|t} = \mathbf{a}_t + \mathbf{P}_t \mathbf{Z}_t' F_t^{-1} v_t. \quad (12)$$

Proposition 4.2 (Minimum mean-square linear update). *Under the linear-Gaussian state-space assumptions, the Kalman update is the minimum mean-square-error estimator of the latent state conditional on observations through time t .*

Proof. The joint distribution of the latent state and observation is multivariate Gaussian. The conditional expectation of a jointly Gaussian vector is linear, and conditional expectation minimizes mean-square error. Substitution of the Gaussian conditioning formula yields the Kalman gain and update above. \square

4.4. Benchmark models

The state-space forecasts are compared with ARIMA(1,1,1),

$$(1 - \phi B)(1 - B)P_t = (1 + \theta B)u_t, \quad (13)$$

and additive-trend exponential smoothing. The last six observations are held out for evaluation. For actual values y_t and forecasts \hat{y}_t , the metrics are

$$\text{MAE} = \frac{1}{h} \sum_{t=1}^h |y_t - \hat{y}_t|, \quad (14)$$

$$\text{RMSE} = \sqrt{\frac{1}{h} \sum_{t=1}^h (y_t - \hat{y}_t)^2}, \quad (15)$$

$$\text{MAPE} = \frac{100}{h} \sum_{t=1}^h \left| \frac{y_t - \hat{y}_t}{y_t} \right|. \quad (16)$$

4.5. Tail-risk estimation

For loss $L_t = -r_t$ and confidence level $1 - \alpha$, historical VaR is

$$\text{VaR}_{1-\alpha}(L) = Q_{1-\alpha}(L) = -Q_\alpha(r), \quad (17)$$

while historical ES is

$$\text{ES}_{1-\alpha}(L) = \mathbb{E}[L \mid L \geq \text{VaR}_{1-\alpha}(L)]. \quad (18)$$

A nonparametric bootstrap with 3,000 resamples is used to quantify uncertainty in the VaR estimate.

5. Empirical Results

5.1. Descriptive statistics

Table 2 summarizes the transformed series. The positive average return is small, but the return distribution is highly asymmetric and leptokurtic. The excess kurtosis of 8.269 indicates that a Gaussian model alone may understate tail behaviour.

Table 2: Descriptive statistics

Statistic	Estimate
Mean transformed price	65.344
Median transformed price	63.125
Standard deviation of price	10.847
Minimum price	46.490
Maximum price	84.140
Mean monthly log return	0.00438
Monthly return volatility	0.06226
Return skewness	1.753
Return excess kurtosis	8.269

The Q–Q plot in Figure 3 departs materially from the normal reference line, particularly in the tails. This visual evidence agrees with the skewness and kurtosis statistics.

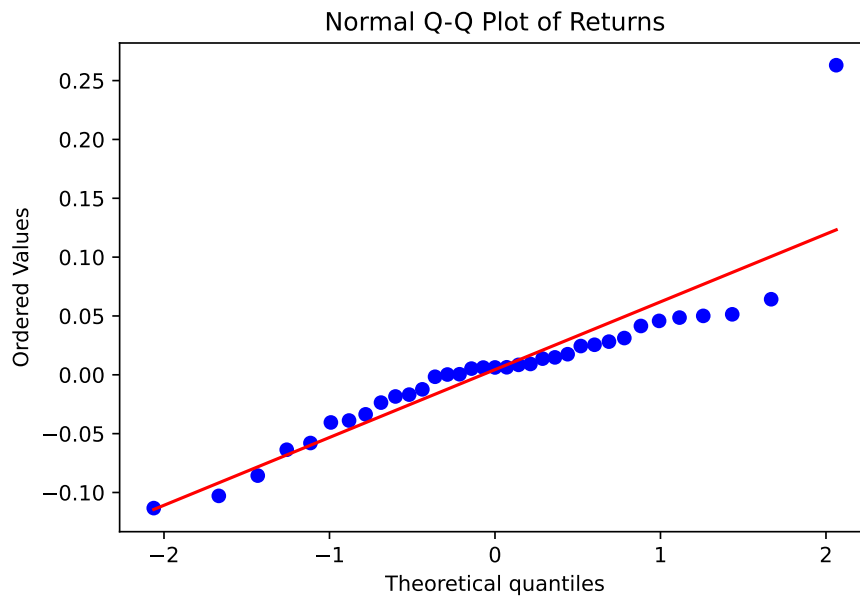


Figure 3: Normal Q–Q plot of transformed monthly returns

5.2. Stationarity and diagnostic tests

The ADF p-value is below 0.001, providing strong evidence that the log-return series is stationary. The Ljung–Box test at lag 6 has p-value 0.445, so the null of no autocorrelation is not rejected. The ARCH-LM p-value is 0.342, giving no significant evidence of short-run ARCH effects in this small monthly sample.

Table 3: Diagnostic test results

Test	Statistic/p-value	Decision at 5%	Interpretation
ADF test	$p < 0.001$	Reject unit-root null	Log returns are stationary.
Ljung–Box, lag 6	$p = 0.445$	Do not reject	No significant residual autocorrelation detected.
ARCH-LM, lag 4	$p = 0.342$	Do not reject	No significant short-run ARCH effect detected.

The absence of significant ARCH effects does not imply normal or risk-free returns. Heavy tails can exist without strong volatility clustering, especially with only 35 returns. Consequently, the regime and tail-risk analyses remain informative but should be interpreted cautiously.

5.3. Regime-switching evidence

Figure 4 presents the smoothed probability of the second latent regime. The probability changes sharply at several dates, showing that the model identifies distinct phases rather than forcing one parameter set across the complete sample. In practical terms, these probabilities can be interpreted as evidence of a time-varying market state, though the state labels themselves are statistical and not inherently “bull” or “bear” until their estimated moments are examined.

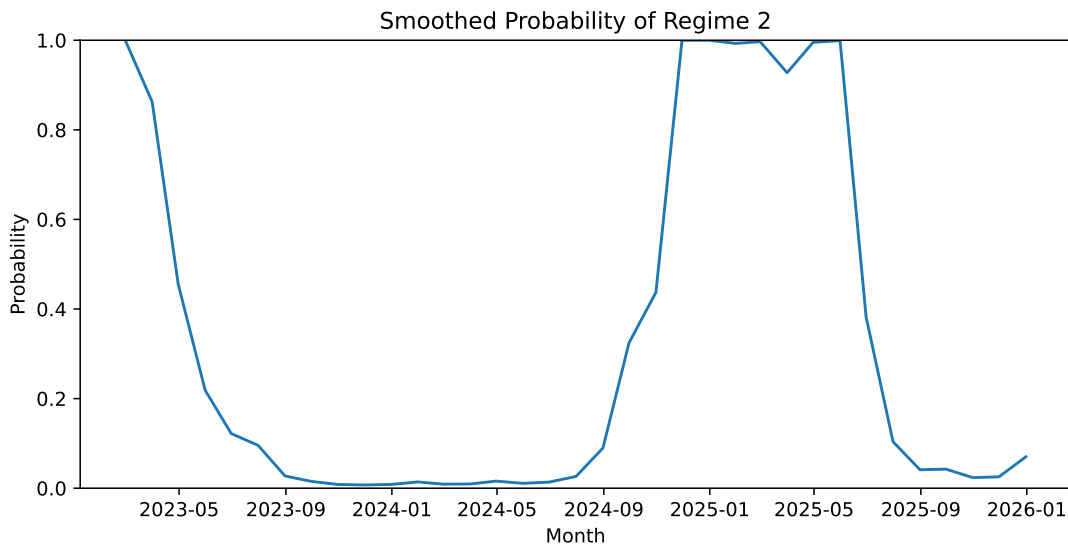


Figure 4: Smoothed probability of latent regime 2

A regime model is especially useful for interpretation because it permits changes in conditional mean and variance. Nevertheless, its estimates are less stable in very short samples. The results should therefore be viewed as exploratory evidence of changing states rather than as precise estimates of long-run transition behaviour.

5.4. Forecast performance

The final six months are excluded from estimation and used as a holdout set. Table 4 shows that ARIMA(1,1,1) performs best on every reported metric. Its RMSE of 2.276 is approximately 65% lower than the structural state-space RMSE of 6.576. The state-space model, however, produces an explicit predictive interval and a decomposition into latent level and slope.

Table 4: Six-month out-of-sample forecast accuracy

Model	MAE	RMSE	MAPE (%)
Local-linear-trend state-space	6.031	6.576	11.029
ARIMA(1,1,1)	2.170	2.276	3.961
Additive ETS	6.250	6.818	11.429

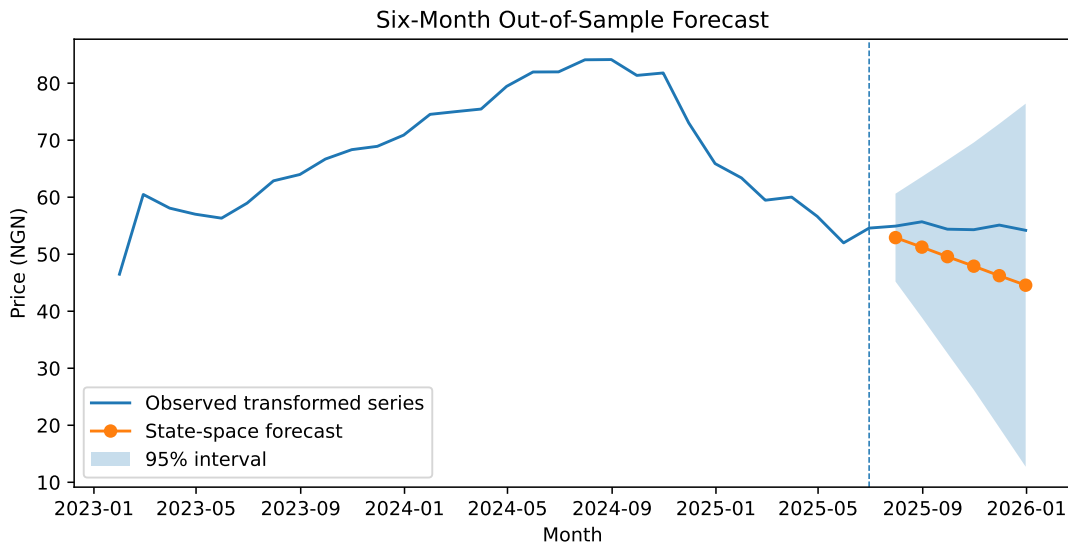


Figure 5: Six-month state-space holdout forecast and 95% prediction interval

This result illustrates why model improvement should not be defined only by mathematical complexity. The state-space and regime-switching framework improves structural interpretation and uncertainty reporting, but the simpler ARIMA model gives the best holdout point forecasts in the transformed sample.

5.5. Downside-risk estimates

At the 95% confidence level, historical monthly VaR is 0.09083. Thus, under the empirical return distribution, the 5th-percentile monthly return corresponds to a loss of approximately 9.083%. Historical ES is 0.10808, indicating an average loss of approximately 10.808% among observations beyond the VaR threshold. The bootstrap confidence interval for VaR is approximately 4.049% to 11.328%.

Table 5: Historical downside-risk estimates

Risk measure	Estimate
95% historical VaR	9.083%
95% historical Expected Shortfall	10.808%
Bootstrap 95% interval for VaR	[4.049%, 11.328%]

The relatively wide interval reflects limited sample size and tail sparsity. ES exceeds VaR, as expected, because it averages losses after the threshold rather than reporting only the threshold itself.

5.6. Sensitivity of discounted terminal value

For comparison with conventional present-value calculations, let the final transformed price be $P_T = 54.19$. The discounted value over nine months is

$$PV(r) = \frac{P_T}{(1+r)^{0.75}}. \quad (19)$$

This gives NGN 50.45, NGN 47.28, and NGN 44.51 under annual discount rates of 10%, 20%, and 30%, respectively.

Proposition 5.1 (Monotonicity and convexity of discounted value). *For $P_T > 0$, $\tau > 0$, and $r > -1$,*

$$PV(r) = P_T(1+r)^{-\tau} \quad (20)$$

is strictly decreasing and strictly convex in r .

Proof. Differentiation yields

$$PV'(r) = -\tau P_T(1+r)^{-\tau-1} < 0 \quad (21)$$

and

$$PV''(r) = \tau(\tau+1)P_T(1+r)^{-\tau-2} > 0. \quad (22)$$

Hence the function is decreasing and convex. \square

6. Discussion

The transformed scenario contains three broad features: an upward movement through much of 2023 and the first part of 2024, a peak near mid-2024, and a sustained decline followed by stabilization in 2025. A single global polynomial can reproduce some of this curvature but lacks a stochastic interpretation and can behave implausibly outside the sample. The local-linear-trend model instead allows the slope itself to evolve, while the regime model allows return moments to switch discretely.

The diagnostic results reveal an important contrast. Returns are stationary and do not show significant autocorrelation or ARCH effects, yet they are strongly skewed and heavy-tailed. This means that modelling priorities should include distributional robustness and tail risk, not only serial dependence. Regime switching and historical ES address those concerns more directly than normal-distribution fitting.

The forecasting experiment also provides a balanced lesson. The structural state-space model is attractive for decomposition, uncertainty intervals, and missing-data handling, but it is not the most accurate point forecaster in the six-month holdout. ARIMA(1,1,1) is clearly superior for this particular transformed sample. A defensible applied analysis should report this outcome instead of selecting a preferred model only because it is more sophisticated.

For decision support, the models can be used jointly. ARIMA can provide the central short-horizon forecast; state-space intervals can communicate uncertainty; regime probabilities can signal changes in the data-generating process; and VaR/ES can describe downside exposure. This ensemble interpretation is more informative than relying on one deterministic trend.

7. Limitations

Several limitations must be acknowledged. First, the dataset is transformed and is not claimed to be an official Lafarge Africa historical price record. Second, 36 monthly observations are insufficient for highly precise regime-transition and tail estimates. Third, macroeconomic variables, market index returns, exchange rates, interest rates, and company-specific fundamentals are omitted. Fourth, the forecast comparison uses only one six-month holdout; rolling-origin evaluation would be preferable with a longer series. Fifth, historical VaR and ES assume that the empirical distribution is informative about future tails and may understate unprecedented events.

8. Conclusion

This paper develops a complete alternative to polynomial fitting and simple discounting for a transformed Lafarge Africa monthly price scenario. The approach integrates stationary log returns, latent Markov regimes, a local-linear-trend state-space model, benchmark forecasting, diagnostics, and downside-risk measurement. The transformed series has a mean price of NGN 65.344 and return volatility of 6.226% per month. Returns are stationary and show no statistically significant autocorrelation or ARCH effect, but they are strongly skewed and leptokurtic. The regime model indicates time-varying latent states. In the holdout experiment, ARIMA(1,1,1) gives the best point forecasts, with RMSE 2.276 and MAPE 3.961%. The 95% historical VaR and ES are 9.083% and 10.808%.

The central methodological conclusion is that model choice should reflect purpose. Regime-switching and state-space methods provide structural interpretation and uncertainty assessment, whereas a simpler ARIMA benchmark may forecast more accurately. Future research should apply the framework to verified daily prices, incorporate exogenous variables, use rolling-window validation, and estimate Bayesian or heavy-tailed regime-switching models on a substantially larger sample.

Data and Reproducibility Statement

The numerical series analysed in this paper is a transformed research dataset supplied for methodological reconstruction. It should not be represented as an official historical price series. The accompanying CSV file and Python script reproduce the statistics, diagnostic tests, forecasts, risk estimates, and figures reported here.

Conflict of Interest

The author declares no conflict of interest.

Acknowledgements

The author acknowledges Akwa Ibom State University for the academic environment that supported this methodological study.

Funding

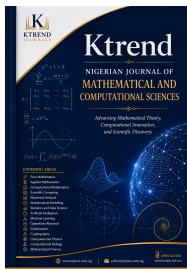
The author received no specific funding for this work.

Author Contributions

Etim Uduak James conceived the study, developed the methodology, conducted the computational analysis, interpreted the results, and prepared the manuscript.

References

- [1] Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3), 203–228.
- [2] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327.
- [3] Box, G. E. P., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time Series Analysis: Forecasting and Control* (5th ed.). Wiley.
- [4] Diebold, F. X., and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business*



E Economic Statistics, 13(3), 253–263.

- [5] Durbin, J., and Koopman, S. J. (2012). *Time Series Analysis by State Space Methods* (2nd ed.). Oxford University Press.
- [6] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), 987–1007.
- [7] Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2), 357–384.
- [8] Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.
- [9] Hyndman, R. J., and Athanasopoulos, G. (2021). *Forecasting: Principles and Practice* (3rd ed.). OTexts.
- [10] Kim, C.-J., and Nelson, C. R. (1999). *State-Space Models with Regime Switching*. MIT Press.
- [11] McNeil, A. J., Frey, R., and Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques and Tools* (Revised ed.). Princeton University Press.