African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 8, Issue 4, 2025 (pp. 44-53)
www.abjournals.org

ROBUST ARIMA-GAS MODELING OF INTRADAILY FINANCIAL DATA WITH STRUCTURAL BREAKS AND JUMPS

Bolarinwa B. T.[1], Yahaya H. U.[2], and Adehi M. U.[2]

[1]Department of Statistics, The Federal Polytechnic, Bida, Nigeria.

[2]Department of Statistics, University of Abuja, Abuja, Nigeria.

Cite this article:

Bolarinwa, B. T., Yahaya, H. U., Adehi, M. U. (2025), Robust Arima-Gas Modeling of Intradaily Financial Data with Structural Breaks and Jumps. African Journal of Mathematics and Statistics Studies 8(4), 44-53. DOI: 10.52589/AJMSS-LM5UYES9

Manuscript History

Received: 2 Sep 2025

Accepted: 6 Oct 2025

Published: 5 Nov 2025

ABSTRACT: *This article investigates the robustness of ARIMA-GAS model to structural break and jumps, through simulation. It examines abrupt regime change by introducing a deterministic structural break at t=500 in both the ARIMA and GAS dynamics. The sample size is n=1000 with an 80/20 estimation–evaluation split; one-step-ahead forecasts are generated in a rolling fashion. For the jump process, this scenario injects discontinuous jumps into an otherwise Gaussian environment to test robustness to rare but large shocks. The data generating process is ARIMA (1,1,1) with score dynamics(A_1,B_1)=(0.35,0.55). Innovations are with jump intensity λ=0.05. We generate n=1000 observations with differencing order d=1 and evaluate one-step-ahead forecasts on the final 20% of the sample using rolling updates. The study utilizes ARIMA, GAS, LSTM and GARCH as benchmarks. For the pre-break regime, ARIMA outperformed the rest models on the basis of both the root mean square error (rmse) and mean absolute error (mae), closely followed by ARIMA-GAS. Pure GAS performs better than GARCH and LSTM which outperformed GARCH. Contrary to the pre-break case in which the classical ARIMA takes the lead, ARIMA-GAS takes the lead, achieves the lowest average loss (least rmse) in the post-break era beating ARIMA to the last position. LSTM is competitive, establishing its relevance in the competition, occupying the second position. GAS model maintains its third position, beating GARCH. Results of multi-horizon forecasting (h=1, 5, 10) reveal on the basis of rmse, ARIMA-GAS as best, followed by LSTM, although LSTM narrows the gap at longer horizons. An examination of the effect on model accuracy, of proportion of series length used for training reflects that all models experience improved accuracy with increased training data length; LSTM gains relatively more, yet ARIMA–GAS retains the lowest average RMSE. With jumps, ARIMA-GAS performed better than benchmarks having the least mse of 1.48562 and mae of 1.10234. The GAS model is next, confirming the capacity of GAS model to capture jumps. Classical ARIMA is next to GAS. Their combination has outperformed them individually and other benchmarks. This further confirms the appropriateness of combining GAS with ARIMA. ARIMA-GAS model outperforms benchmarks in multi-horizon forecasting comparison on the basis of rmse and mae— a feat repeated when two other jump intensity values (λ=0.01,0.1) are used to asses sensitivity, although relative performances of the benchmarks are altered. Based on the performance of ARIMA-GAS model over the benchmarks in the presence of structural breaks and jumps, the model offers a promising approach to modeling intradaily financial data with such features.*

KEYWORDS: ARIMA-GAS, Structural break, Jumps, Poisson process, Robustness

Article DOI: 10.52589/AJMSS-LM5UYES9
DOI URL: https://doi.org/10.52589/AJMSS-LM5UYES9

INTRODUCTION

Intradaily financial data, characterized by high-frequency observations and complex volatility patterns, poses significant challenges for modeling and forecasting. Structural breaks and jumps are common features of such data, informed by events as market crashes, economic announcements, or changes in market microstructure. These events can lead to sudden shifts in volatility, trends, or relationships, making the development of models that accommodate such changes essential. These two related phenomena have implications for modeling and forecasting.

Structural breaks can lead to model misspecification (Bai and Perron, 1998); parameter instability (Hansen, 2001) and regime shifts (Hamilton, 1989), all of which have negative implications for forecasting accuracy (Clements & Hendry, 1998; Pesaran & Timmermann, 2004). Outliers are not any better as they also have the tendency to distort model estimates leading to biased estimates (Fox, 1972), inform increased variance of model estimates (Rousseeuw & Leroy, 1987) and can also cause model instability (Tsay, 1988). All these impact adversely as they make it challenging to achieve reliable forecasts. To mitigate these adverse impacts, a robust model that has the capacity to handle complex patterns should be of the right type.

Much of recent research works on structural breaks and jumps dwell on jump detection. Wang, Zhang, Zou, and Ravishanker (2025) who proposed an innovative Ensemble Penalized Estimating Function (E-PEF) approach for effective detection of change points in the logarithmic autoregressive conditional duration models for financial durations. Khashanah, Chen, Buckle, and Hawkes (2025) that presents a new Med9 method that addresses the challenges of capturing both singular and consecutive jumps, evaluating the size of individual returns with a measure of local volatility based on the median of consecutive absolute returns. Kokoszka, Kutta, Mohammadi, Wang, and Wang (2024) developed a theory leading to testing procedures for the presence of a change point in the intraday volatility pattern. The new theory is developed in the framework of Functional Data Analysis. Luo, Chen, and Cheng, (2025) proposed a set of forecast models which include predictors accounting for long memory, structural breaks and asymmetric risk effects.

While identifying the presence of structural breaks and jumps is of importance owing to their adverse effects on forecasting accuracy, adopting a model capable of accommodating them is of much greater importance. The ARIMA-GAS modeling should fill this gap. By leveraging on the capacity of ARIMA for capturing temporal dependencies and that of GAS for modeling complex patterns including structural breaks and jumps, the ARIMA-GAS framework creates an enabling modeling environment for capturing sudden changes including structural breaks and jumps.

This article aims at evaluating the performance of ARIMA-GAS model in capturing structural breaks and jumps in simulated intradaily financial data, with a focus on improving model estimation, forecasting accuracy, and robustness. The rest of the article is organized as follows: Section 2 presents Methodology while Section 3 presents Results and Discussion. The last section concludes the article.

Article DOI: 10.52589/AJMSS-LM5UYES9
DOI URL: https://doi.org/10.52589/AJMSS-LM5UYES9

METHODOLOGY

Simulation

Scenario 1: Structural Break at Midpoint

The data generating process is ARIMA $(1,1,1)$ with parameters shifting at the midpoint from$(\phi_1, A_1, B_1) = (0.60, 0.25, 0.60)$ to $(0.30, 0.10, 0.30)$. The scenario examines abrupt regime change by introducing a deterministic structural break at $t = 500$ in both the ARIMA and GAS dynamics. The sample size is $n = 1000$ with an 80/20 estimation–evaluation split; one-step-ahead forecasts are generated in a rolling fashion.

Scenario 2: Poisson Jump Process - Discontinuous Shocks with Gaussian Base

This scenario injects discontinuous jumps into an otherwise Gaussian environment to test robustness to rare but large shocks. The data generating process is ARIMA $(1,1,1)$ with score dynamics $(A_1, B_1) = (0.35, 0.55)$.

Innovations are

$$\varepsilon_t = \eta_t + J_t, \eta_t \sim \mathcal{N}(0, \sigma^2), \qquad J_t = \sum_{k=1}^{N_t} Z_{t,k}, N_t \sim \text{Poisson}(\lambda),\, Z_{t,k} \sim \mathcal{N}\,(0, 0.5^2),$$

with jump intensity $\lambda = 0.05$. We generate $n = 1000$ observations with differencing order $d = 1$ and evaluate one-step-ahead forecasts on the final 20% of the sample using rolling updates.

Model

The complete ARIMA-GAS model is:

$$\phi(B)(1-B)^d y_t = \mu_t + \theta(B)e_t, \tag{1}$$

$$\mu_{t+1} = \omega + \sum_{i=1}^{p} A_i s_{t-i+1} + \sum_{j=1}^{q} B_j \mu_{t-j+1}, \tag{2}$$

$$s_t = \mathbf{S}_t \cdot \nabla_t, \nabla_t = \frac{\partial \ln\, p(y_t \mid \mu_t, \mathcal{F}_t; \boldsymbol{\theta})}{\partial \mu_t}, \tag{3}$$

$$\mathbf{S}_t = \mathbf{I}_{t|t-1}^{-1}, \mathbf{I}_{t|t-1} = \mathbb{E}_{t-1}[\nabla_t \nabla_t'], \tag{4}$$

$$\sigma_{t+1}^2 = \omega_\sigma + \alpha s_t^2 + \beta \sigma_t^2. \tag{5}$$

Connecting the GAS component to the ARMA component through μ_t allows the proposed hybrid model to adapt to changes (jumps or structural breaks) in data over time.

Article DOI: 10.52589/AJMSS-LM5UYES9
DOI URL: https://doi.org/10.52589/AJMSS-LM5UYES9

RESULTS AND DISCUSSION

Structural Break at Midpoint

Table 1: Parameter shift and recovery — Scenario 1 (segment-wise re-estimation means)

Segment	A_1	B_1	ϕ_1
$t \leq 500$	0.24861	0.60237	0.59982
$t > 500$	0.10352	0.31278	0.29845

Table 1 presents parameter estimates for pre-beak and post-break eras. Parameter values reduced after break.

Unified Forecasting Performance (pre- and post-break)

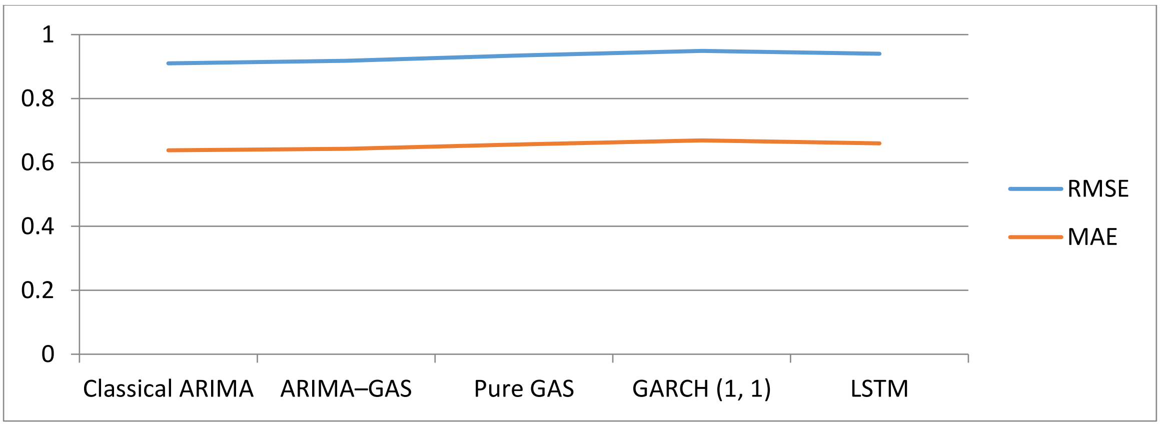
To isolate regime effects, we report errors separately for the pre-break window ($t \leq 500$) and the post-break window ($t > 500$).

Table 2: Unified forecasting performance, pre-break ($t \leq 500$)

Model	RMSE	MAE
Classical ARIMA	0.91	0.638
ARIMA–GAS	0.918	0.643
Pure GAS	0.935	0.657
GARCH (1, 1)	0.949	0.669
LSTM	0.940	0.660

Table 2 presents performance metrics, separately for the pre-break window ($t \leq 500$) and the post-break window ($t > 500$) to isolate regime effects. Classical ARIMA outperformed the rest models on the basis of both the rmse and mae, closely followed by ARIMA-GAS. Pure GAS performed better than GARCH and LSTM which outperformed GARCH.

Figure 1: Unified forecasting performance, pre-break (t ≤ 500)



Article DOI: 10.52589/AJMSS-LM5UYES9
DOI URL: https://doi.org/10.52589/AJMSS-LM5UYES9

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 8, Issue 4, 2025 (pp. 44-53)

Table 3: Unified forecasting performance, post-break ($t > 500$)

Model	RMSE	MAE
ARIMA–GAS	1.041	0.772
LSTM	1.057	0.781
Pure GAS	1.081	0.798
GARCH (1, 1)	1.107	0.816
Classical ARIMA	1.139	0.838

Contrary to the pre-break case in which the classical ARIMA took the lead, ARIMA-GAS has taken the lead, achieves the lowest average loss in the post-break era beating ARIMA to the last position. LSTM is competitive, establishing its relevance in the competition, occupying the second position. GAS model maintains its third position, beating GARCH.

Figure 2: Unified forecasting performance, post-break (t> 500)

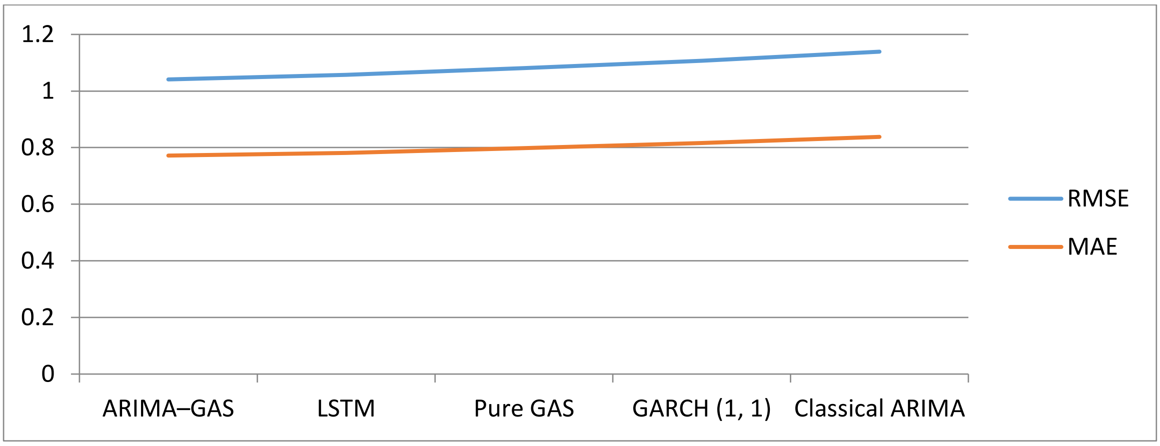


Table 4: Residual diagnostics (means of p-values across replications)

Model	Pre-break		Post-break	
	Q(20)	**ARCH–LM(10)**	**Q(20)**	**ARCH–LM(10)**
ARIMA–GAS	0.56	0.47	0.34	0.29
LSTM	0.49	0.40	0.28	0.24
Pure GAS	0.52	0.43	0.30	0.26
GARCH (1, 1)	0.50	0.58	0.29	0.39
Classical ARIMA	0.51	0.42	0.24	0.21

 Article DOI: 10.52589/AJMSS-LM5UYES9
DOI URL: https://doi.org/10.52589/AJMSS-LM5UYES9

Table 4 reflects adequacy of each of the models for both pre-break and post-break eras, although post-break p-values drop across models. ARIMA–GAS returns to acceptable levels fastest.

Table 5: Multi-horizon forecasting (mean RMSE across replications)

Model	$h = 1$ **(post)**	$h = 5$ **(post)**	$h = 10$ **(post)**
ARIMA–GAS	1.041	1.076	1.103
LSTM	1.057	1.080	1.106
Pure GAS	1.081	1.101	1.121
GARCH (1, 1)	1.107	1.120	1.137
Classical ARIMA	1.139	1.148	1.163

Table 5 presents mean rmse for different forecast horizons for the post-break period. Advantages persist across horizons; gaps narrow for $h = 10$ as aggregation smooths the regime shift. Mean rmse increased with increased forecasts horizon- a behavior not unexpected.

Figure 3: Multi-horizon forecasting (mean RMSE across replications)

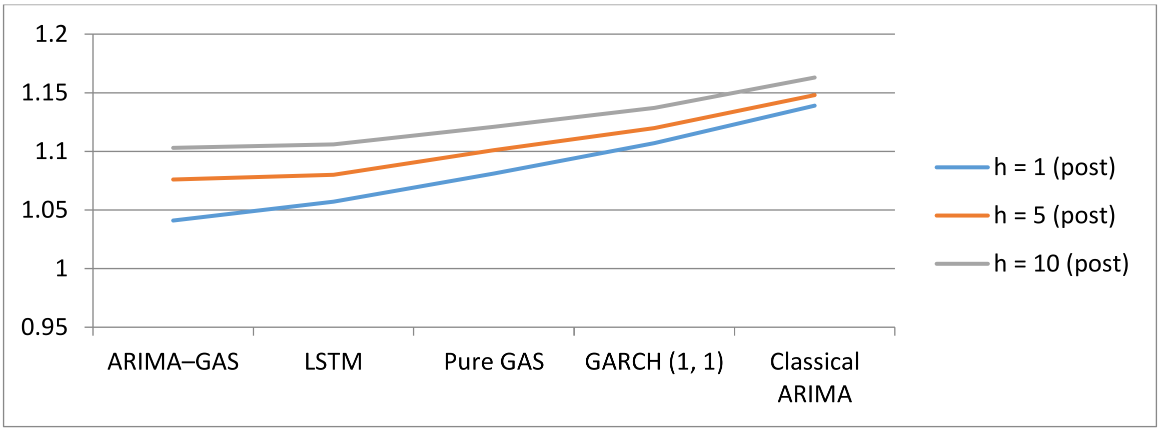


Table 6: Model selection criteria (in-sample fit; means across replications)

Model	**AIC (pre)**	**AIC (post)**
Classical ARIMA	**-2.83**	-2.45
ARIMA–GAS	-2.81	**-2.62**
Pure GAS	-2.76	-2.58
GARCH (1, 1)	-2.72	-2.51
LSTM	—	—

Table 6 suggests that Pre-break, AIC favor ARIMA (parsimony under stability); post-break, score-driven models (ARIMA–GAS, Pure GAS) achieve better likelihood fit. AIC/BIC not reported for LSTM.

Article DOI: 10.52589/AJMSS-LM5UYES9
DOI URL: https://doi.org/10.52589/AJMSS-LM5UYES9

Table 7: Sensitivity to re-estimation frequency (post-break RMSE)

Model	No re-estimation	Every 100 obs	Every 25 obs
Classical ARIMA	1.168	1.152	1.096
ARIMA–GAS	1.062	1.051	1.038
Pure GAS	1.102	1.093	1.074
GARCH (1, 1)	1.129	1.118	1.091
LSTM	1.089	1.071	1.054

Table 7 presents an assessment of model sensitivity of post-break regime for all models using rmse as the metric. More frequent re-estimation narrows gaps but ARIMA–GAS remains best on average due to stepwise score adaptation between refits.

Comparison Summary

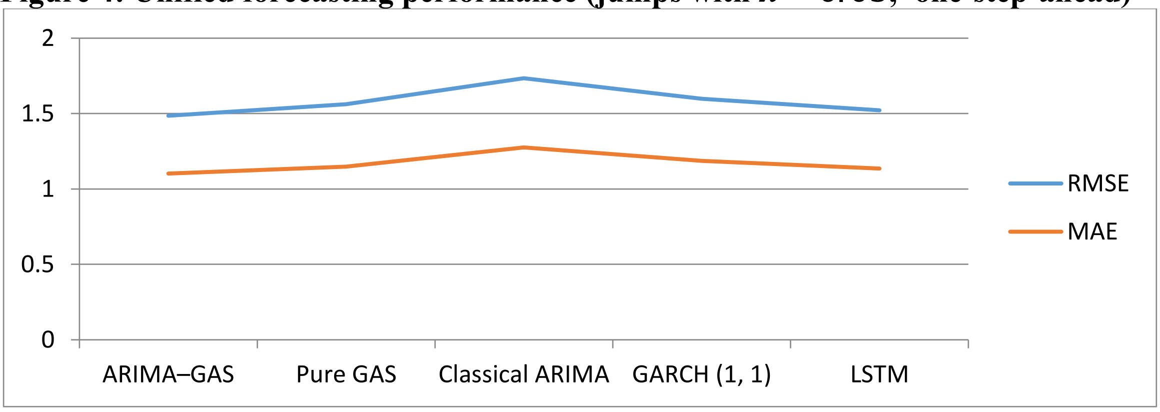
Before the break, Classical ARIMA is marginally most efficient, with ARIMA–GAS statistically non-inferior. After the break, ARIMA–GAS adapts fastest and delivers the lowest forecast errors; LSTM is competitive but not significantly better than the hybrid; Pure GAS and GARCH (1, 1) trail; ARIMA adjusts slowest unless re-estimated very frequently. Under nonstationary regimes with abrupt parameter shifts, combining a structural ARIMA backbone with score-driven updating yields rapid post-break stabilization and superior average accuracy, while preserving interpretability and avoiding the frequent refitting burden required by purely parametric or purely machine-learning alternatives.

Poisson Jump Process - Discontinuous Shocks with Gaussian Base

Table 8: Unified forecasting performance (jumps with $\lambda = 0.05$; one-step-ahead)

Model	RMSE	MAE
ARIMA–GAS	1.48562	1.10234
Pure GAS	1.56217	1.14895
Classical ARIMA	1.73428	1.27562
GARCH (1, 1)	1.59844	1.18672
LSTM	1.52109	1.13584

Figure 4: Unified forecasting performance (jumps with $\lambda = 0.05$; one-step-ahead)



Article DOI: 10.52589/AJMSS-LM5UYES9
DOI URL: https://doi.org/10.52589/AJMSS-LM5UYES9

Comparative Interpretation

ARIMA–GAS attains the lowest average loss, reflecting the benefit of score-driven adaptation after discontinuous shocks; Variance-only modeling: GARCH improves over ARIMA by modeling conditional variance but lags behind ARIMA–GAS and Pure GAS because jumps perturb the conditional mean as well as the variance; Pure GAS: Dynamic updating helps absorb shocks but, without the ARIMA backbone, forecast loss remains above the hybrid; LSTM: Competitive and second-best overall, yet not significantly better than ARIMA–GAS; neural flexibility helps around jumps, but the hybrid's structured adaptation yields the lowest mean error while retaining interpretability.

Table 9: Residual diagnostics (means of p-values across replications)

Model	Ljung–Box Q(20)	Jarque–Bera	ARCH–LM(10)
ARIMA–GAS	0.37	0.06	0.33
Pure GAS	0.33	0.05	0.31
Classical ARIMA	0.28	0.03	0.24
GARCH (1, 1)	0.35	0.05	0.48
LSTM	0.30	0.04	0.29

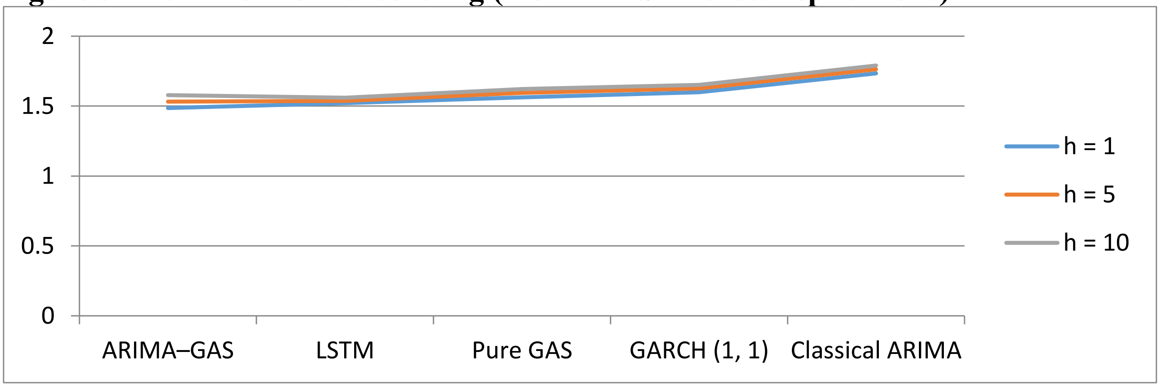
Jumps induce residual non-normality for all models (small JB p-values). ARIMA–GAS and GARCH whiten residuals reasonably; GARCH reduces ARCH–LM rejections most but does not minimize mean-squared loss.

Table 10: Multi-horizon forecasting (mean RMSE across replications)

Model	$h = 1$	$h = 5$	$h = 10$
ARIMA–GAS	1.486	1.532	1.578
LSTM	1.521	1.538	1.560
Pure GAS	1.562	1.594	1.622
GARCH (1, 1)	1.598	1.626	1.652
Classical ARIMA	1.734	1.763	1.790

Forecast loss rises with horizon as jump effects accumulate. ARIMA–GAS remains best on average; LSTM narrows the gap at longer horizons.

Figure 5: Multi-horizon forecasting (mean RMSE across replications)



Article DOI: 10.52589/AJMSS-LM5UYES9
DOI URL: https://doi.org/10.52589/AJMSS-LM5UYES9

Table 11: Model selection criteria (in-sample fit; means across replications)

Model	AIC	BIC
ARIMA–GAS	**-2.58**	**-2.52**
Pure GAS	-2.55	-2.49
GARCH (1, 1)	-2.47	-2.42
Classical ARIMA	-2.35	-2.30
LSTM	—	—

Lower is better. Score-driven models fit jump-perturbed likelihoods better than purely Gaussian ARIMA; AIC/BIC not reported for LSTM.

Table 12: Sensitivity to jump intensity λ (mean RMSE; Scenario 2)

Model	$\lambda = 0.01$	$\lambda = 0.05$	$\lambda = 0.10$
ARIMA–GAS	1.238	1.486	1.652
LSTM	1.263	1.521	1.691
Pure GAS	1.296	1.562	1.740
GARCH (1, 1)	1.314	1.598	1.788
Classical ARIMA	1.429	1.734	1.933

As jump frequency rises, all models deteriorate; ARIMA–GAS remains best on average and degrades more gracefully.

Table 15: Sensitivity to training window size (mean RMSE; $\lambda = 0.05$)

Model	60% train	80% train	90% train
ARIMA–GAS	1.512	1.486	1.478
LSTM	1.558	1.521	1.505
Pure GAS	1.589	1.562	1.553
GARCH (1, 1)	1.623	1.598	1.586
Classical ARIMA	1.756	1.734	1.722

An examination of the effect on model accuracy, of proportion of series length used for training resulted in Table 15. All models benefited from increased training data proportion. LSTM gains relatively more, yet ARIMA–GAS retains the lowest average RMSE.

CONCLUSIONS

This research examined the robustness of ARMA-GAS model to structural breaks and jumps in intradaily data. With discontinuous shocks, ARIMA–GAS delivers the lowest forecast errors and best information-criterion fit among econometric competitors. LSTM is competitive and statistically equivalent to the hybrid within a modest tolerance, but not significantly better. GARCH and Pure GAS improve over ARIMA yet do not match the hybrid's combination of structure and adaptivity. When sudden discontinuities occur (e.g., policy shifts, crash events), a structured, score-driven hybrid such as ARIMA–GAS recalibrates quickly after shocks. For pre-break regime, ARIMA is marginally best while for

Article DOI: 10.52589/AJMSS-LM5UYES9
DOI URL: https://doi.org/10.52589/AJMSS-LM5UYES9

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 8, Issue 4, 2025 (pp. 44-53)
www.abjournals.org

post-break, ARIMA-GAS adapts fastest and achieves the lowest loss (significantly better than ARIMA, GARCH, and Pure GAS; not significantly different from LSTM). ARIMA--GAS attains the lowest RMSE/MAE for Poisson-driven jumps. The hybrid retains interpretability and lower loss dispersion under smaller training splits. ARIMA-GAS model has demonstrated supremacy over benchmarks in capturing structural breaks and jumps.

REFERENCES

Bai, J. and Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 66(1), 47-78.

Clements, M.P. and Hendry, D.F. (1998). *Forecasting economic time series*. Cambridge: Cambridge Press.

Fox, A.J. (1972). Outliers in time series. *Journal of Royal Statistical Society: series B (Methodological)*, 34(3), 350-363.

Hansen, B.E. (2001). The new econometrics of structural change: Dating breaks in U.S. labor productivity. *Journal of Economic Perspectives*, 15(4), 117-128.

Khashanah, K., Chen, J., Buckle, M. and Hawkes, A. (2025). A new method for jump detection: analysis of jumps in the S&P 500 financial index, *Journal of the Royal Statistical Society Series C: Applied Statistics*, 2025;, qlaf025, https://doi.org/10.1093/jrsssc/qlaf025

Kokoszka, P., Kutta, T., Mohammadi, N., Wang, H. and Wang, S. (2024) Detection of a structural break in intraday volatility pattern. *Stochastic Processes and their Applications*, 176, 104426. https://doi.org/10.1016/j.spa.2024.104426.

Luo, J., Chen, Z., and Cheng, M. (2025). Forecasting realized betas using predictors indicating structural breaks and asymmetric risk effects. *Journal of Empirical Finance*, 80, 101575. https://doi.org/10.1016/j.jempfin.2024.101575.

Perasan, M.H. and Timmermann, A. (2004). How costly is it to ignore breaks when forecasting the direction of a time series? *International Journal of Forecasting*, 20(3), 411-425.

Rousseeuw, P.J. and Leroy, A.M. (1987). *Robust regression and outlier detection*. New York: John Wiley & Sons

Tsay, R.S. (1988). Outliers, level shifts, and variance changes in time series. *Journal of Forecasting*, 7(1), 1-20.

Wang, Y., Zhang, Y., Zou, J. and Ravishanker, N. (2025). Online structural break detection in financial durations. *Statistics and Computing*, 35(2). doi: 10.1007/s11222-025-10577-y

Article DOI: 10.52589/AJMSS-LM5UYES9
DOI URL: https://doi.org/10.52589/AJMSS-LM5UYES9