

Data Article

Title: Volatility Modelling of the JSE Top40 Index: Assessing the GAS Framework Against GARCH and Hybrid GARCH–XGBoost

Authors: Israel Maingo, Thakhani Ravele and Caston Sigauke

Affiliations: Department of Mathematical and Computational Sciences, University of Venda, Private Bag X5050, Thohoyandou, 0950, South Africa.

Contact email: caston.sigauke@univen.ac.za

Abstract

This paper studies the volatility dynamics of the JSE Top40 Index by estimating a univariate GAS model with time-varying location, scale, and shape parameters (identity score scaling) and comparing its density and point-forecast performance against a standalone ARMA(3,2)-EGARCH(1,1) model and a hybrid ARMA(3,2)-EGARCH(1,1)-XGBoost framework. The GAS model is estimated on 3,515 daily observations, and several conditional densities are examined. The Student-t GAS model (GAS-STD) obtains the lowest information criteria within the GAS family (AIC = 10,188.142; BIC = 10,243.626) and exhibits statistically significant persistence in location and scale dynamics. Statistical diagnostics provide evidence of correct density calibration (Normalised Log Score = 1.1932; Uniform score = 0.4417), although residual skewness remains (IID-Test skewness p = 0.0134). Out-of-sample analysis shows that GAS-STD performs strongly in density and risk forecasting, producing accurate 5% VaR and ES paths and passing coverage backtests (Kupiec LRuc 13 p = 0.8414; DQ p = 0.2281). However, short-horizon point forecasts are best delivered 14 by the hybrid ARMA-EGARCH-XGBoost model (RMSE = 0.1386), with Diebold-Mariano tests confirming a transitive ranking: Hybrid > ARMA-EGARCH > GAS-STD. Simulation experiments highlight the sensitivity of tail behaviour to degrees-of-freedom (e.g., kurtosis v = 5 ≈ 7.32). Overall, GAS-STD is a strong density and risk model for the JSE Top40, while the hybrid framework excels in short-term volatility forecasting. The data are stored in an Excel file.

Specifications Table

Subject area	<i>Volatility Modelling and Short-Term Forecasting</i>
More specific subject area	<i>Financial Market Modelling and Forecasting</i>
Type of data	<i>Excel file</i>
How data was acquired	<i>Provided and from the internet</i> https://za.investing.com/indices/89_ftse-jse-top-40-historical-dat
Data format	<i>Filtered and analysed.</i>
Experimental factors	<i>N/A</i>

Experimental features	<i>N/A</i>
Data source location	<i>FTSE/JSE Top 40 index Data portal webpage</i>
Data accessibility	<i>The data is hosted on GitHub https://github.com/csiqauke</i>
Related research article	Volatility Modelling of the JSE Top40 Index: Assessing the GAS Framework Against GARCH and Hybrid GARCH–XGBoost

Value of the Data

The data can be used for financial time series forecasting in cross-sectional, temporal, or cross-temporal frameworks.

Data

The data used in this study is from Wall Street Journal Markets, FTSE/JSE Top 40 index, and it can be accessed from <https://za.investing.com/indices/ 89 ftse-jse-top-40-historical-dat>. The data comprises Date and Close price.

Experimental Design, Materials, and Methods

Data used in the study is from FTSE/JSE Top 40 index.

Acknowledgements

The authors acknowledge FTSE/JSE Top 40 index for providing the data.

Funding sources

The 2025-2026 NRF Masters Postgraduate Scholarship support towards this research is hereby acknowledged.

References

The FTSE/JSE Top 40 index from <https://za.investing.com/indices/ 89 ftse-jse-top-40-historical-dat> (accessed on 25 April 2025).