Interdisciplinary Joint project report

OptiverGroup13Report[aff-1]

This project aims to develop a sophisticated volatility prediction model tailored for high-frequency trading environments, a critical need in the financial sector. The primary focus was on leveraging ultra-high-frequency data to accurately forecast the volatility of over 100 stocks, enhancing trading strategies and risk management.
Our comprehensive analysis led to the selection of the ARIMA-GARCH model, which demonstrated superior ability to manage the complexities of high volatility stocks effectively. This model was chosen for its exceptional balance of predictive performance and computational efficiency, standing out for its robust handling of non-stationary data and volatility clustering.

optiver | stock | report

Aim of Project

The primary goal of this project is to develop an advanced volatility prediction model specifically designed for high-frequency trading environments. This project addresses the critical need for accurate volatility prediction in the financial sector, especially for firms like Optiver, where the value of options is directly influenced by the volatility of the underlying asset. The project aims to utilise ultra-high-frequency data to create a predictive tool that accurately forecasts volatility (Bannouh, 2013)

Consider a day trader who specialises in high-volatility stocks and depends on precise and immediate data to make informed trading decisions. The volatility prediction tool developed in this project is designed to meet this trader's needs - our persona - by providing rapid and reliable forecasts of stock volatility within minute-long intervals. The use of advanced time-series models and the integration of comprehensive performance measures ensures that the tool is not only accurate and stable but also interpretable. This tool can be used to assess risk, determine entry and exit points, and tailor his trading strategies to maximise returns. This tool can also serve as an evaluation resource for after trading hours, allowing users to analyse patterns and behaviours of different stocks, enhancing their understanding and improving their future trading strategies.

Data Description

The dataset provided by Optiver includes high-frequency trading data for 126 individual stocks, captured in CSV files. Each file contains order book snapshots taken at one-second intervals during multiple 10-minute time buckets throughout a trading day. For each snapshot, the dataset records details about the two highest bid prices and the two lowest ask prices, along with their corresponding sizes. By understanding the metrics used, users can proceed with strategising how to predict future volatility.

Methodology

We began by developing a business understanding and building a persona to provide a clear path for the project. The persona highlights that our focus was to predict high-volatility stocks to enhance decision-making using

Significance Statement

A key outcome from this project was the integration of the ARIMA-GARCH model into a user-friendly Shiny app. This application allows traders to adjust and interact with the model's inputs, providing real-time forecasts of stock volatility. This capability is vital for traders who specialise in high-volatility stocks and relies on precise data for trading decisions.
The practical relevance of this tool is significant; it not only enhances the ability of traders to anticipate and react to market conditions but also serves as an educational and strategic resource outside of trading hours. By enabling detailed analysis and fostering a deeper understanding of market dynamics, this model and its application helps with strategic decision-making in volatile trading scenarios.

Author affiliations: [aff-1]The University of Sydney

Anabel Geraldine, Jeremy Gebrael, Jiaqi Fan, Pranav Sundaram, Royston Quek and Thomas Lunn

predictive models. We proceeded with Exploratory Data Analysis (EDA) to ascertain data quality and identify key trends and volatility patterns across low and high volatility stocks. Our process continued with selection of stocks based on volatility and a cross-validation of the results, ensuring a focus on high-volatility trading scenarios.

For model development, we employed different time series predictive models where each model was evaluated using chosen statistical metrics. Moreover, the model's robustness was validated through residual analysis. After comparing these metrics across different models, a final model was chosen. This model then underwent testing with different training sizes and window configurations to evaluate their stability and impact on prediction.

The comprehensive approach culminated in deploying the final model through a Shiny app designed for use by traders, allowing real-time volatility forecasting based on the latest market data. This methodology not only supported the development of an accurate and efficient predictive model but also provided a scalable solution adaptable to the dynamic nature of financial markets.

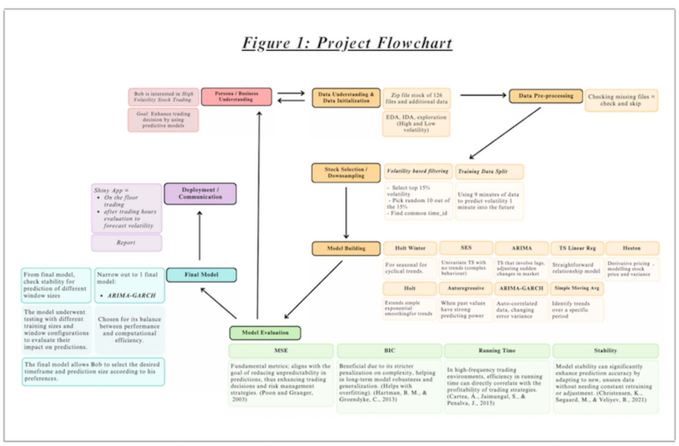

Figure 1: Project Flowchart

Approach

Data pre-processing. This phase involved a comprehensive data preparation and pre-processing strategy to ensure the integrity and usability of the data for subsequent modelling. We began by downloading the dataset, which consisted of 126 stock files accompanied by additional data. A crucial step in the pre-processing phase was the identification and exclusion of missing stock files. Specifically, we identified files 12, 24, 25, 45, 49, 54, 57, 65, 71, 79, 91, 92, 106, 117, and 121 as missing, and they were all skipped to maintain the robustness of our model.

EDA and IDA were then conducted to gain a deeper understanding of the characteristics and dynamics within the dataset. Our primary focus was to distinguish between stocks with high and low volatility.

Exploratory Data Analysis

During the EDA, the stock files were loaded, and the transaction records per second were examined. These records contained different layers of information on bid and ask prices. To analyse the data effectively, we aggregated the data to calculate the average bid and ask prices per 10 seconds. These averages were then plotted over time to illustrate the price fluctuations in 10-second intervals.

Next, we analysed price trends across different time ID segments. The average change in bid and ask prices was plotted to demonstrate how prices fluctuated over different time periods. This analysis provided insights into the temporal dynamics of price movements.

Additionally, spread analysis was conducted to evaluate market liquidity and stability. We calculated the bid-ask spreads and plotted the change in average spreads over time. This analysis helped us observe trends in the bid-ask spreads, which are indicative of market liquidity and stability.

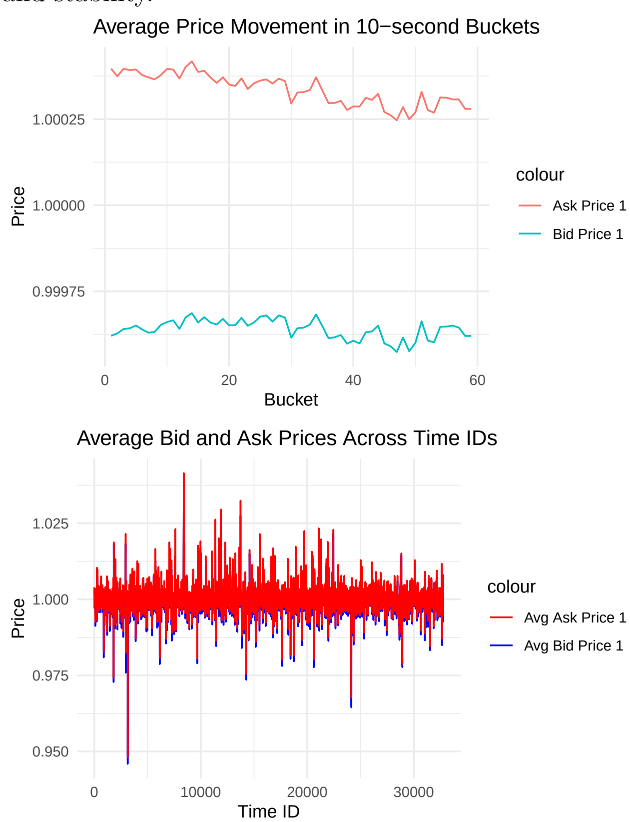

Figure 2: EDA stuff output

These analyses helped us understand market price behaviour and liquidity, providing a foundation for further research.

Stock Selection. After identifying high-volatility stocks as our main interest, we proceeded to the stock selection or classification phase. This step was crucial as it provided essential data regarding high volatility. In the context of financial markets, where the precision and applicability of predictions can significantly influence trading outcomes, this reduction is particularly valuable.

From the stocks ordered based on their mean volatility, we selected the top 15%. To ensure robustness and prevent selection bias, ten stocks were chosen at random from this subset. This method provided a diverse yet focused group of stocks, representative of typical conditions in volatile trading environments. Additionally, to standardise analysis and model testing, 10 common time IDs were randomly selected across these stocks.

We then proceeded with cross-validation of our data to ensure the model's robustness. From the stocks selected, we partitioned the time_ids to facilitate cross-validation, conducting the training phase across four iterations. To validate our findings, we then employed the Kruskal-Wallis test. This process was foundational in building a predictive model tailored to effectively support strategic trading decisions in highly volatile market conditions.

Model Building. In developing our volatility prediction model, we carefully selected a suite of diverse yet complementary models, each chosen for their unique strengths and suitability for handling the complexities of time series data, particularly in the context of high-volatility stock trading. The models included in this analysis are Holt-Winters, Simple Exponential Smoothing (SES), ARIMA, Time Series Linear Regression, and Heston Model variants. Each of these models offers specific benefits for predicting financial market volatility, making them well-suited for the aim of the project.

Holt-Winters Model: This model accounts for seasonal variations, which can be particularly useful for seasonal or cyclical stock patterns. Its ability to integrate seasonality directly into the forecasting model makes it particularly valuable for our dataset, where such patterns can significantly impact volatility predictions.

Holt Model: Extending simple exponential smoothing, the Holt model captures trends in data, which is beneficial for forecasting trends in volatile stocks. This model is suitable for data with a systematic trend, providing a more nuanced approach to forecasting compared to SES.

Simple Exponential Smoothing (SES): Chosen for its simplicity and effectiveness in forecasting data with little to no trend or patterns, SES is useful when dealing with high-frequency trading data that may not exhibit long-term trends. It is also computationally efficient, making it ideal for real-time trading applications.

Autoregression Model: Effective for data where past values have strong predictive power on future values, which is common in financial markets. This model is promising in stable financial conditions where past values can accurately predict future trends.

ARIMA Model:Provides a robust approach for modelling time series that involve lags of both the forecast variable and the forecast errors, helping to adjust for sudden changes in the market. ARIMA can model a wide range of data behaviours, offering a foundational model for forecasting trends in volatile stocks.

ARIMA-GARCH Model: A hybrid model combining ARIMA's capabilities in handling non-stationary data with GARCH's strength in modelling time-varying volatility. This model is effective for stock market data where volatility tends to be auto-correlated and the error variance changes over time.

Time Series Linear Regression: Applies a linear approach to model relationships within the data, particularly effective for identifying and quantifying straightforward relationships and trends over time.

Simple Moving Average Model: Identifies trends over a specified period, smoothing out short-term fluctuations and highlighting longer-term trends or cycles. Its simplicity and effectiveness in providing a clear view of the underlying trends in volatile data make it a valuable tool.

Heston Model: Typically used for derivative pricing, this model observes intricate dynamics in financial markets by modelling stock prices and their variance. It is particularly useful for capturing the complex behaviour of financial instruments.

While these models vary in complexity and approach, they share a common goal of accurately forecasting financial market behaviour. By evaluating and comparing each of these time series models, we aim to determine which model best supports our overall objective.

Model Evaluation. In developing our volatility prediction model, a crucial step was the selection and evaluation of appropriate statistical metrics to assess each model's performance effectively. These metrics are essential for making informed comparisons and refining our models to ensure optimal performance.

To achieve this, we initially considered a broad array of evaluation metrics which were selected for their relevance to our project's goals, offering insights into various aspects of model performance. After thorough consideration, we decided to focus on MSE, BIC, and training time as our key metrics:

Mean Squared Error (MSE): This metric is fundamental for reducing unpredictability in predictions, enhancing trading decisions and risk management strategies. MSE measures the average of the squares of the errors—that is, the average squared difference between the estimated values and the actual value. Lower MSE values indicate better model performance (Poon & Granger, 2003).

Bayesian Information Criterion (BIC): BIC is particularly beneficial for its stringent penalisation of model complexity, which is crucial for maintaining performance in a changing market. It helps in model selection by balancing the fit of the model and its complexity, thereby avoiding overfitting (Hartman & Groendyke, 2013).

Running Time: This metric is assessed to ensure that the efficiency of our model aligns with the operational demands of real-time financial trading. Fast training times are essential for high-frequency trading environments where timely decision-making is critical (Cartea, Jaimungal, & Penalva, 2015).

Focusing on these key metrics allows us to develop a robust volatility prediction model that meets the high standards required for effective trading strategies.

Model/Metric	MSE	BIC	Time (s)
Holt-Winter	5.2207e-07	-232.4981	2.303
Holt	9.0324e-07	-225.8566	2.229
SES	5.2698e-07	-236.3798	0.845
Autoregressive	5.1354e-06	-215.0550	1.151
ARIMA	5.7514e-07	-216.0940	23.77
ARIMA-GARCH (N-steps)	1.6427e-08	-7565.4846	209.66
Time Series Linear Regression	2.2409e-08	-101376.7750	12.86
Simple Moving Average	5.387e-07	-212.33	16.78
HAR-RV Model	1.67146e-40	-6219.484078	30.91
Heston	0.032212835593	20000000019.19	54.47

Figure 3. Tabulated Performance metrics for 9 models

After evaluating various models, the ARIMA-GARCH model was chosen for predicting market volatility due to its excellent performance in key metrics. It achieved a MSE of 1.6427e-08 and an exceptionally low BIC of -7565.4846, although it had a slightly longer training time of 209.66 seconds. This model effectively balances accuracy with the ability to manage non-stationarity and volatility clustering, which are crucial for financial markets.

Other models, while faster, did not match ARIMA-GARCH in overall performance. While models like HAR-RV, despite showing extremely low MSE and BIC values, raised concerns about potential overfitting. Overfitting can significantly affect the practical deployment of these models in trading scenarios, where generalisability to new data is essential.

Final Model. From our evaluation of various models, we have identified the ARIMA-GARCH model as the best performing, and most suitable for our goals. The ARIMA-GARCH model stands out due to its optimal balance between predictive performance and computational efficiency. The model integrates ARIMA's proficiency in capturing non-stationary aspects of time series data with GARCH's ability to model conditional volatility effectively. This combination is ideal for predicting the erratic movements typical of high-volatility stocks.

```
{'whiskers': [<matplotlib.lines.means': []}
(-5e-09, 9.5e-08)
{'whiskers': [<matplotlib.lines.
```

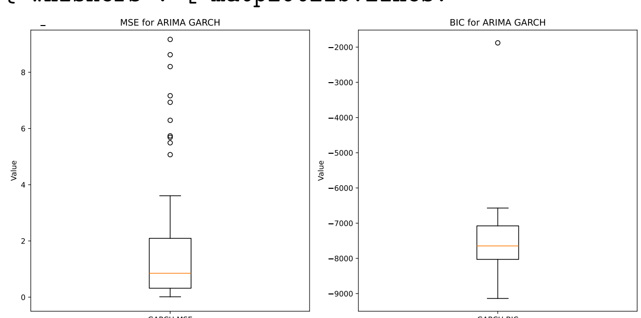

Figure 4: Boxplot of MSE, BIC & Residual analysis

```
<matplotlib.collections.PathCollection o
<matplotlib.lines.Line2D object at 0x1609ffed0>
Text(0.5, 0, 'Fitted Values')
Text(0, 0.5, 'Standardized Residuals')
Text(0.5, 1.0, 'Residuals vs Fitted Values')
```

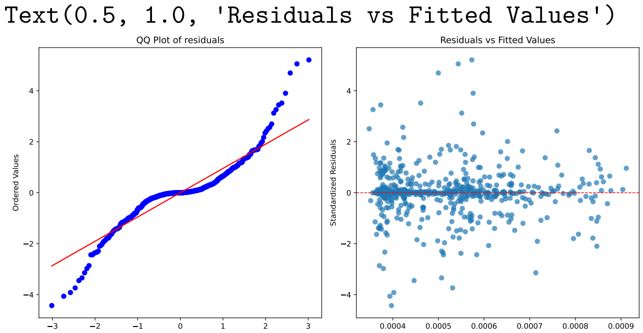

Figure 5 : QQ plot of residuals and residual vs fitted values

The plots presented in the image are diagnostic tools used to evaluate the performance of an ARIMA-GARCH model. The first plot, a QQ plot of residuals, compares the quantiles of the standardised residuals from the ARIMA-GARCH model to the quantiles of a standard normal distribution. Ideally, if the residuals follow a normal distribution, the points should align along the 45-degree line. In this plot, while the central points adhere closely to this line, deviations at the tails suggest the presence of heavy tails or outliers. However, the variation in the residual is expected and acceptable, as volatility and time series data often do not follow normal distribution, typically exhibiting extreme values due to market dynamics.

The second plot, which shows the residuals versus fitted values, is used to check for patterns that might indicate model inadequacies. A good model will have residuals scattered randomly around zero, without any discernible patterns. In this plot, the residuals appear to be randomly dispersed around the horizontal axis, suggesting that there are no obvious patterns or systematic errors in the model. However, the presence of some residuals with high magnitudes could indicate periods of volatility not fully captured by the model.

```
<Figure size 1000x600 with 0 Axes>
{'whiskers': [<matplotlib.lines.Line2D object
Text(0.5, 1.0, 'Comparison of MSEs Across Each
Text(0, 0.5, 'MSE')
Text(0.5, 0, 'Model')
(-5e-09, 1.2e-07)
```

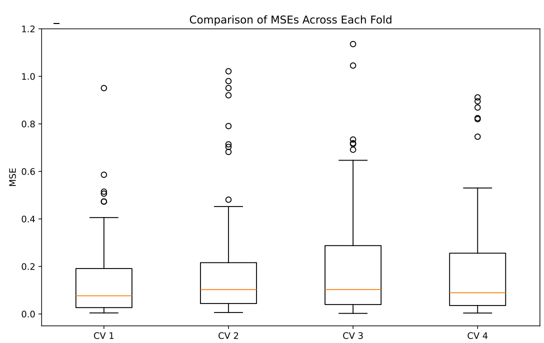

Figure 6: Comparsion of MSEs Across Each Fold

```
Text(0.5, 1.0, 'MSEs across diff train sizes')
Text(0, 0.5, 'MSE')
(-5e-09, 1.5e-07)
Text(0.5, 0, 'Minutes of train data')
```

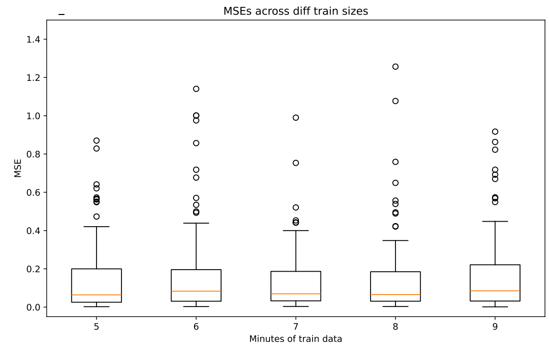

Figure 7: MSEs Across Diff train sizes

```
{'whiskers': [<matplotlib.lines.
Text(0.5, 1.0, 'MSEs across different prediction
Text(0, 0.5, 'MSE')
(-5e-09, 1.25e-07)
Text(0.5, 0, 'Time (s)')
```

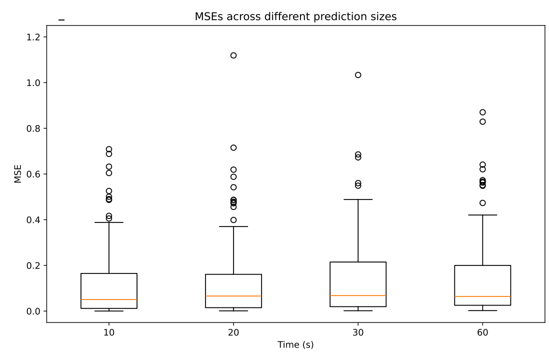

Figure 8: MSEs Across Different prediction sizes

Furthermore, we tested the stability of the ARIMA-GARCH model by varying training and prediction sizes. This analysis was crucial to determine the model's reliability across different scenarios. The results confirmed that the ARIMA-GARCH model maintains consistent performance regardless of these variations, demonstrating no significant differences in the outcomes. These diagnostic plots are essential in validating the assumptions of the ARIMA-GARCH model, ensuring its reliability and stability in capturing the underlying data dynamics.

Deployment (Shiny App). Shiny provides an excellent framework that allows us to get started quickly without needing additional knowledge of HTML or CSS. It enables users to interact with data through a simple, interactive interface. We designed the application to allow users to upload a stock CSV file to predict volatility. The app reads the data from the file, and allows users to select the training window and prediction size through drop-down menus, making the data analysis process more intuitive and user-friendly.

Furthermore, the Shiny app has the ability to filter stock data by time ID, enhancing the precision of the analysis. The app responds to user inputs in real-time and updates the results instantly, which is very crucial for strategies that depend on quick responses to market volatility. In the final export phase, we can easily deploy the Shiny app to Shiny Server or shinyapps.io for access by targeted users, simplifying the sharing of data analysis results and interactive applications. By providing a direct and interactive approach to volatility prediction, our Shiny app serves as a valuable tool for traders, offering both advantages and strategic depth in navigating the complexities of volatile markets.

Conclusion and Findings

Throughout this project, we successfully developed and refined a volatility prediction model tailored for high-frequency trading environments. The ARIMA-GARCH model emerged as the most effective mode, providing an optimal balance between accuracy and computational efficiency. Our analysis confirmed that the ARIMA-GARCH model is particularly adept at managing the unpredictable and dynamic nature of high volatility stocks. The robustness of this model was systematically tested across various training sizes and window configurations. With the results indicating consistent performance, thereby confirming the model's reliability and stability.

The integration of our model to our Shiny app exemplifies practical innovation by offering real-time data processing capabilities. Ths interactive and user-friendly interface of the app enables users to customise their analysis parameters easily, with this app, traders can adapt quickly to market changes, enhancing decision making processes and risk management strategies.

Limitation and Future Research

The primary limitation is the model's dependency on historical data, which might not represent and capture future market conditions, particularly during unforeseen market events. Additionally, the ARIMA-GARCH model, while based on our robustness evaluation, requires considerable computational resources for real-time data processing, which could be optimised further.

Future enhancement could include the integration of deep learning algorithms to improve accuracy and potentially uncovering non-linear relationships. Another area for development is the inclusion of other external financial instruments such as the market news, macroeconomic indicators, or liquidity, as some form of exogenous data. These enhancements aim to provide a more holistic view of the market, allowing traders to make more informed decisions based on a comprehensive analysis of both quantitative and qualitative data.

Student Contributions

Royston - Holt-Winter, Holt, SES, Autoregressive, ARIMA, ARIMA-GARCH, Slides, Shiny

Pranav - TS Linear Regression, Shiny

Jeremy - Simple Moving Average, Time Analysis, Slide Visual, Script

JiaQi - HAV-RV Model, EDA stuff, Script, Qmd report generate

Anabel - Extra features analysis, Figure1, Script, Report written, Referencing

Thomas - Heston, Presentation Slides, Script

Appendix

Our shiny app can be see here:
https://oriole.shinyapps.io/data3888/

References

Aljez, R. M. 2016. "Comparative Performance of ARIMA and ARCH/GARCH Models on Time Series of Traffic Accidents in Gaza Strip." https://www.semanticscholar.org/paper/Comparative-Performance-of-ARIMA-and-ARCH-GARCH-on-Aljez/5ff9c599dc36fa9996b32e459a88c4a7a8a8f9f4#citing-papers.

Amendola, A., and V. Candila. 2016. "Evaluation of Volatility Predictions in a VaR Framework." *Quantitative Finance* 16 (5): 695–709. https://www.tandfonline.com/doi/epdf/10.1080/14697688.2015.1062122?needAccess=true.

Bannouh, K. 2013. "Measuring and Forecasting Financial Market Volatility Using High-Frequency Data." PhD thesis, Erasmus University Rotterdam. https://www.erim.eur.nl/doctoral-programme/phd-in-management/phd-tracks/detail/314-measuing-and-forecasting-financial-market-volatility-using-high-frequency-data/.

Box, G. E., and G. M. Jenkins. 1976. *Time Series Analysis: Forecasting and Control*. Holden-Day. http://repo.darmajaya.ac.id/4781/1/Time%20Series%20Analysis_%20Forecasting%20and%20Control%20%28%20PDFDrive%20%29.pdf.

Cartea, Á., S. Jaimungal, and J. Penalva. 2015. *Algorithmic and High-Frequency Trading*. Cambridge University Press. https://assets.cambridge.org/97811070/91146/frontmatter/9781107091146_frontmatter.pdf.

Christensen, K., M. Søgaard, and B. Veliyev. 2021. "A Machine Learning Approach to Volatility Forecasting." CREATES Research Paper 2021-03. https://pure.au.dk/ws/portalfiles/portal/208284743/rp21_03.pdf.

Dokuchaev, N. 2014. "Volatility Estimation from a Short Time Series of Stock Prices." *Journal of Nonparametric Statistics* 26 (2): 373–84. https://ideas.repec.org/a/taf/gnstxx/v26y2014i2p373-384.html.

Hartman, B. M., and C. Groendyke. 2013. "Model Selection and Averaging in Financial Risk Management." University of Connecticut; Robert Morris University. https://hartman.byu.edu/docs/files/HartmanGroendyke_ModelSelection_p.pdf.

Hyndman, R. J., and G. Athanasopoulos. 2013. *Forecasting: Principles and Practice*. 2nd ed. Hoboken, NJ: OTexts. https://otexts.com/fpp2/.

Poon, S.-H., and C. W. J. Granger. 2003. "Forecasting Volatility in Financial Markets: A Review." *Journal of Economic Literature* 41 (2): 478–539. https://www.aeaweb.org/articles?id=10.1257/002205103765762743.

Sørbø, S., and M. Ruocco. 2024. "Navigating the Metric Maze: A Taxonomy of Evaluation Metrics for Anomaly Detection in Time Series." *Data Mining and Knowledge Discovery* 38 (6): 1027–68. https://doi.org/10.1007 /s10618-023-00988-8.

Sparks, J. J., and Y. V. Yurova. 2006. "Comparative Performance of ARIMA and ARCH/GARCH Models on Time Series of Daily Equity Prices for Large Companies." Department of Information; Decision Sciences, University of Illinois at Chicago. https://www.researchgate.net/publication/268199171 performance_of_ARIMA_and_companies.

Zhang, Y. 2021. "Average Half-Hour Trading Volume During Day and Night Trading Hours for Each Half-Hour Fig 1." https://www.researchgate.net/figure/Average-half-hour-trading-volume-during-day-and-night -trading-hours-For-each-half-hour_fig1_348458659.