

Article

Enhancing Prediction by Incorporating Entropy Loss in Volatility Forecasting

Renaldas Urniezius ^{1,*}, Rytis Petrauskas ¹, Vygasdas Vaitkus ¹, Javid Karimov ¹, Kestutis Brazauskas ¹, Jolanta Repsyte ¹, Egle Kacerauskiene ², Torsten Harms ³, Jovita Dargiene ⁴ and Darius Ezerskis ¹

¹ Department of Automation, Kaunas University of Technology, Studentu St. 48, 51367 Kaunas, Lithuania

² Faculty of Electrical and Electronics Engineering, Kaunas University of Technology, Studentu St. 48, 51367 Kaunas, Lithuania; egle.savickaite-kacerauskiene@ktu.lt

³ Fakultät Wirtschaft, Duale Hochschule Baden-Württemberg, Erzbergerstraße 121, 76133 Karlsruhe, Germany

⁴ Higher Education Institution, Kauno Kolegija, Pramones pr. 20, 50468 Kaunas, Lithuania

* Correspondence: renaldas.urniezius@ktu.lt

Abstract

In this paper, we propose examining Heterogeneous Autoregressive (HAR) models using five different estimation techniques and four different estimation horizons to decide which performs better in terms of forecasting accuracy. Several different estimators are used to determine the coefficients of three selected HAR-type models. Furthermore, model lags, calculated using 5 min intraday data from the Standard & Poor's 500 (SPX) index and the Chicago Board Options Exchange Volatility (VIX) index as the sole exogenous variable, enrich the models. For comparison and evaluation of the experimental results, we use three metrics: Quasi-Likelihood (QLIKE), Mean Absolute Error (MAE), and Mean Squared Error (MSE). An empirical study reveals that the Entropy Loss Function consistently achieves the best QLIKE results in all the horizons, especially in the weekly horizon. On the other hand, the performance of the Robust Linear Model implies that it can provide an alternative to the Entropy Loss Function when considering the results of the MAE and MSE metrics. Moreover, research shows that adding more informative lags, such as Realized Quarticity for the Heterogeneous Autoregressive model yielding the Realized Quarticity (HARQ) model, and incorporating the VIX index further improve the general results of the models. The results of the proposed Entropy Loss Function and Robust Linear Model suggest that they successfully achieve significant forecasting accuracy for HAR models across multiple forecasting horizons.

Keywords: volatility forecasting; entropy loss function; realized volatility; HAR; HAR-RV; HARQ



Academic Editors: Marcin Makowski and Jan Śladowski

Received: 9 June 2025

Revised: 22 July 2025

Accepted: 25 July 2025

Published: 28 July 2025

Citation: Urniezius, R.; Petrauskas, R.; Vaitkus, V.; Karimov, J.; Brazauskas, K.; Repsyte, J.; Kacerauskiene, E.; Harms, T.; Dargiene, J.; Ezerskis, D. Enhancing Prediction by Incorporating Entropy Loss in Volatility Forecasting. *Entropy* **2025**, *27*, 806. <https://doi.org/10.3390/e27080806>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Volatility is the quality or state of being likely to change suddenly [1]. In finance, it is often referred to as the variance or standard deviation of an asset's returns over a given period. As markets have become more complex and information-driven, accurate forecasting of volatility has become one of the main focuses of academics and practitioners [2]. Since the volatility depends on the market, macroeconomic conditions, financial leverage, and trading activity, it is often used to make financial decisions, such as asset allocation, market timing, derivative pricing, and risk evaluation over time [3]. As a result, researchers have developed models to forecast volatility and explore alternative ways to measure it, including range-based estimators that use high–low price ranges rather than returns [4].

In the second half of the 20th century, scholars began developing models that better captured the dynamic behavior of returns and volatility. Early theoretical works emphasized the importance of accurate volatility estimation, as seen in the Black–Scholes model [5], where volatility is critical in pricing derivative securities and managing financial risk. The introduction of Black–Scholes stimulated interest in alternative approaches to volatility modeling. A foundational breakthrough came with the development of Engle’s Autoregressive Conditional Heteroskedasticity (ARCH) model, which became the basis for many subsequent studies on volatility forecasting [6]. Further, Bollerslev introduced the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model, which extended the ARCH framework by incorporating a more flexible lag structure for past variances [7]. The later introduced Exponential Generalized Autoregressive Conditional Heteroskedasticity (EGARCH) approach [8] has a few advantages over GARCH. The first is that EGARCH measures the log returns, resulting in positive conditional variance. The second is that this model’s asymmetry allows for the capture of the leverage effect [9]. Glosten et al. introduced the Glosten–Jagannathan–Runkle Generalized Autoregressive Conditional Heteroskedasticity (GJR-GARCH) model, which improved the GARCH version by evaluating an additional variable, allowing for the capture of possible asymmetries [9,10]. Despite many claims that ARCH and stochastic models provide poor forecasting, the results of [11] contradict such a statement. The availability of intraday data has enabled the use of realized volatility to examine the stochastic properties of asset returns [12]. Building on stochastic volatility frameworks, research has revealed that volatility exhibits rough, fractal-like behavior rather than smooth diffusion. A later study, completed by Poon and Granger, concluded that GARCH models were dominating ARCH models in volatility forecasting [13]. Recently, Gatheral et al. proposed the Rough Volatility model, in which volatility followed a fractional Brownian motion with a Hurst exponent below 0.5, offering improved empirical fit and theoretical insights compared to classical stochastic volatility (SV) models [14]. The following study by Ulugbode and Shittu proposed a Transition EGARCH model that outperformed standard GARCH-type models in forecasting the conditional volatility of the Nigerian Stock Exchange. However, the authors also emphasized that more complex non-linear models were not inferior to simpler alternatives and may provide greater flexibility in capturing market dynamics [15].

Over time, it became apparent that standard volatility models were insufficient to estimate daily volatility, especially over more extended periods, despite the available data. The increased quantities of available data allowed for a reduction in prediction horizons, shifting from quarterly or monthly to weekly or daily forecasts. Andersen et al. developed a method for modeling and forecasting the realized volatility and correlation. Researchers specified and estimated a long-memory Gaussian Value-at-Risk (VaR) model for daily realized volatilities. The proposed model successfully forecasted volatility and surpassed the GARCH model [16]. Later, Corsi proposed the Heterogeneous Autoregressive model of Realized Volatility (HAR-RV), which effectively captured long-memory properties in volatility. The model consistently outperformed short-memory alternatives across various forecasting horizons, including one day, one week, and two weeks [17]. Further, Andersen et al. proposed a new Heterogeneous Autoregressive model of Realized Variance with Continuous and Jump components (HAR-RV-CJ) forecasting model that supplemented the HAR-RV model, incorporating the discontinuous jumps to better capture the whole structure of asset return variability [18]. Afterward, research indicated that jump variation was important in predicting future volatility—negative jumps resulted in higher volatility and vice versa [19]. However, Prokopczuk et al. [20] found that explicit inclusion of jumps did not improve forecasting accuracy after comparing various HAR-RV model extensions over multiple forecasting periods. Despite the widespread application of GARCH-type

models, recent studies suggest that realized and implied volatility measures contain valuable forward-looking information and can significantly enhance the forecasting accuracy of GARCH-based models [21]. Therefore, researchers have proposed simple yet effective forecasting models based on realized volatility (RV), which explicitly account for temporal variation in forecast errors. One such model is the HARQ model, an extension of the HAR framework that incorporates error-based adjustments to enhance forecast accuracy [22]. To better account for regime-switching behavior and nonlinear dynamics in electricity markets, recent extensions of the HAR-RV model, such as the Logistic Smooth Transition HAR (LST-HAR) framework, have been proposed, showing improved forecasting accuracy over traditional linear HAR specifications [23]. To address cross-asset volatility dynamics, Cubadda and Hecq [24] proposed a Vector HAR (V-HAR) index model, which captured standard volatility components across multiple assets and improved forecast performance in multivariate settings. Later, Clements and Preve suggested an approach to enhance volatility forecasting by supplementing HAR models with Ridge Regression (RR) and simple Weighted Least Squares (WLS) techniques [25]. Furthermore, Li et al. extended the HAR model by incorporating over 200 cross-market predictors into a shrinkage framework using Least Absolute Shrinkage and Selection Operator (LASSO) and Elastic Net techniques. Investigation results showed that integrating global realized volatility components, jump measures, and uncertainty indices significantly improved forecasting accuracy [26]. Recently, Michael et al. augmented the HAR model with a range of volatility estimators by applying dimensionality reduction techniques to implied volatility surfaces and calibrating stochastic volatility model parameters, including Heston and Bates, to extract implied volatility estimators. This analysis indicated that these augmentations significantly enhanced daily realized volatility forecasting and could effectively improve VIX prediction [27].

The need for improved volatility forecasting and the development of computing capabilities led to the investigation and development of advanced methods, such as machine learning (ML) and Artificial Neural Networks (ANNs). Therefore, Luong and Dokuchaev introduced the Random Forest (RF) algorithm for forecasting realized volatility, suggesting that using purified implied volatility as an input and applying the RF technique significantly enhanced the predictive performance of the traditional HAR model [28]. Later on, Bucci demonstrated that feedforward neural networks could outperform traditional volatility forecasting models by effectively capturing complex nonlinear dynamics and structural breaks in financial time series. Among the models evaluated, Long Short-Term Memory (LSTM) and Nonlinear Autoregressive with Exogenous Input (NARX) delivered the most accurate realized volatility forecasts [29]. Zhang et al. proposed the use of Temporal Convolutional Networks (TCNs) for forecasting stock volatility and Value-at-Risk (VaR), with the empirical results indicating superiority of this model over the traditional GARCH-type models in forecasting both volatility and downside risk [30]. However, ML approaches often encounter challenges, as expressed by Ge et al., who systematically reviewed studies published since 2015 on neural network-based approaches for financial volatility forecasting, describing the difficulty in directly comparing model performance across studies and the gap between modern and standard ML techniques [31]. Further, Zahid and Saleem investigated the performance of Support Vector Machine (SVM) models under high-volatility conditions during the COVID-19 pandemic, revealing that the Radial Basis Function (RBF) kernel outperformed both linear and polynomial kernels [32]. As discussed by Christensen et al., neural networks are superior to detecting nonlinearity in market dynamics and extracting valuable information related to implied volatility, especially over longer forecasting horizons [33]. Therefore, Souto and Moradi introduced the Neural Basis Expansion Analysis for Time Series with Exogenous Variables (NBEATSx)

model for realized volatility forecasting, demonstrating its superior predictive accuracy across multiple stock indices when compared to both traditional econometric models, such as HAR and GARCH, and advanced deep learning architectures, including LSTM and Temporal Convolutional Networks (TCNs) [34]. Additionally, Zhang et al. proposed a machine learning approach that leveraged intraday volatility commonality across multiple assets to enhance forecasting accuracy, and another approach to forecast the coming-day RV by using past intraday RVs as a predictor. The latter proposed approach yielded a superior out-of-sample forecast compared with traditional models [35]. Recently, multiple researchers have noticed that the application of machine learning (ML) methods to capture nonlinear dynamics and complex structures often delivers far better volatility forecasting results than traditional econometric models [36]. The study by Lolic indicated that Random Forest and Gradient Boosting models consistently outperformed traditional econometric models [37]. Beg analyzed machine learning models for volatility forecasting and revealed that Random Forest (RF) consistently delivered the most accurate predictions and high correlation with the observed data. On the other hand, neural network or Gradient Boosting models performed moderately, requiring extensive tuning for optimal results. Linear regression and Support Vector Regression (SVR) yielded the weakest forecasts due to their limited ability to capture nonlinear dependencies in financial data [38]. The study by Mansilla-Lopez et al. concluded that machine learning models consistently outperformed traditional econometric approaches in forecasting financial market volatility. The authors proposed a unifying definition of volatility by comparing various statements in different articles: “volatility is an indicator of market risk, which measures the variation in the returns of a financial asset over a period of time” [39].

Recent research has introduced hybrid models that integrate GARCH-family structures with neural networks, leveraging the strengths of both econometric and machine learning approaches. Following the desire to develop hybrid models, Monfared and Enke applied different ML methods to volatility forecasting. The results showed improvement for extreme event forecasting, but the authors did not suggest using a hybrid approach for low-volatility periods due to the complexity of the model [40]. Implementing Artificial Neural Networks (ANNs) alongside GARCH models improved forecast accuracy by over 10% when applied to Latin American financial markets, demonstrating the effectiveness of hybrid methodologies in emerging market contexts [41]. Yang et al. introduced another approach, employing hybrid modeling by integrating the Support Vector Machine (SVM) algorithm within a big data framework. Their results demonstrated that effective feature selection and dimensionality reduction significantly enhanced prediction accuracy and computational efficiency in large-scale volatility forecasting tasks [42]. Later, Trierweiler Ribeiro et al. integrated the Heterogeneous Autoregressive (HAR) framework with an Echo State Neural Network (ESN) and Particle Swarm Optimization (PSO), leveraging different strengths from each method. The HAR-ESN-PSO model demonstrated an increased predictive accuracy across multiple forecasting horizons [43]. Liu et al. combined Bidirectional Recurrent Neural Networks (Bi-RNNs), Gated Recurrent Units (GRUs), and Particle Swarm Optimization (PSO). The proposed GBP (GRU–BiRNN–PSO) model showed increased forecasting performance across several datasets, thus demonstrating improved learning capacity and generalization ability compared to traditional methods [44]. The same year, Mishra et al. introduced a more accurate hybrid model integrating GARCH and LSTM components within a bagged attention mechanism. The Multi-Task Generalized Autoregressive Conditional Heteroskedasticity (MT-GARCH) and Multi-Task Learning Generalized (MTL-GARCH) models demonstrated superior predictive accuracy and robust risk estimates across diverse markets during elevated volatility windows, such as the COVID-19 pandemic [45]. Additionally, Brini and Toscano introduced SpotV2Net, an NN

with a Graph Attention Network (GAT) architecture for multivariate intraday spot volatility forecasting. The model captured dynamic cross-asset dependencies and spillover effects, delivering significant gains in prediction accuracy compared to traditional econometric and alternative machine learning models [46]. Most recently, Hu et al. combined Convolutional Neural Networks (CNNs) with the Heterogeneous Autoregressive–Kitchen Sink (HAR-KS) to forecast the direction of stock market volatility. Empirical results from the Chinese stock market showed that the CNN-HAR-KS model outperformed traditional econometric and standalone deep learning models in forecasting performance [47]. Furthermore, Li et al. augmented the classical HAR-RV and LSTM models with additional influencing factors. The results showed that while factor integration improved the performance of both models, the LSTM consistently outperformed HAR-RV. Furthermore, incorporating Principal Component Analysis (PCA) into the LSTM architecture yielded the highest forecasting accuracy across all experimental configurations [48]. Finally, Kumar et al. combined Variational Mode Decomposition (VMD) with deep learning models, including ANN, LSTM, and GRU. The proposed Q-VMD-ANN-LSTM-GRU approach achieved satisfactory forecasting results across multiple stock indices, demonstrating strong potential for improving financial risk management, stress testing, and investment strategy formulation [49].

Researchers have increasingly recognized that effective volatility forecasting requires more than just historical price data or directional predictions; it also demands attention to the magnitude of price movements and the integration of public sentiment with macroeconomic indicators. To address this need, Wang et al. proposed the Economics (ECON) framework, which combined filtered tweet sentiments, government-derived macroeconomic statistics, and historical price data to forecast stock movement and volatility together. The ECON model significantly improved predictive accuracy by modeling various relationships between individual stocks, industry sectors, and macroeconomic conditions [50]. Consequentially, Shi et al. introduced a novel framework for volatility forecasting by developing NumBERT, a pre-trained language model specifically designed to enhance the interpretation of numerical information in earnings call transcripts. Their approach improves contextual understanding of financial text and significantly increases the accuracy of 3-day volatility predictions [51]. The following study by Li et al. developed a Hierarchical Transformer-based model to forecast asset volatility by extracting risk information from annual reports. The authors constructed investment portfolios using Natural Language Processing (NLP), which successfully predicted beta values that outperformed the SPX by an average of 21 %. The results of these studies demonstrated the potential of deep language models in volatility-aware asset management strategies [52].

This analysis of the literature shows the progression and shift from traditional econometrics to advanced forecasting methods, such as employing machine learning or using hybrid approaches (see Figure 1). This research paper proposes to take an additional step and presents a novel methodology for enhancing volatility forecasting by implementing entropy loss. The proposed method is based on the HAR-RV model, which is widely used in the field of volatility forecasting. The HAR-RV model captures the long-memory properties of volatility and has been shown to outperform other models in various studies. By incorporating entropy loss into the HAR-RV framework, we aim to improve the accuracy of volatility predictions. To verify the proposed method, a large dataset of high-frequency SPX data covering the period from 2 January 2008 to 15 April 2025 is used. The results demonstrate that the HAR-RV model with entropy loss significantly outperforms traditional HAR-RV models regarding forecasting accuracy. This research contributes to the existing literature by providing a novel approach to volatility forecasting that combines the strengths of HAR-RV models with the benefits of entropy loss.

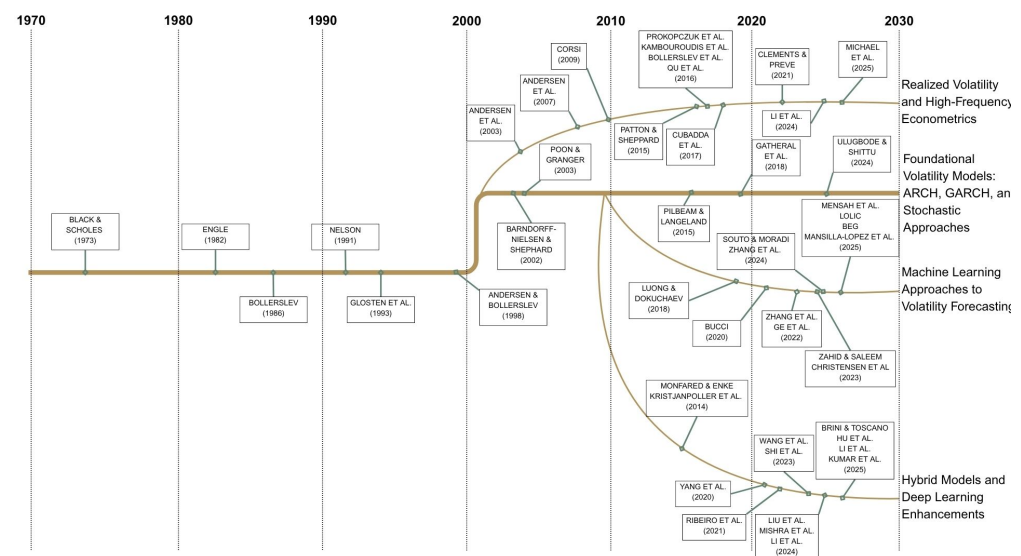


Figure 1. Development of volatility forecasting models [5–52].

This paper is organized as follows: Section 2 introduces the methodology and development of the proposed method. Section 3 presents the investigation’s results and discussions. Finally, Section 4 concludes the whole article.

2. Materials and Methods

2.1. Data Description

Table 1 provides a sample of the high-frequency dataset used in this study. The data, obtained under a licensed agreement from “First-rate Data”, consisted of five variables: Datetime, Open, High, Low, and Close prices. Each observation corresponded to a distinct 5 min interval within standard U.S. stock market trading hours. For every time interval, the dataset recorded the opening price, the highest and lowest prices observed during the interval, and the closing price. The following list provides definitions for each variable.

- Datetime: indicating the timestamp of the 5 min subinterval of the values (e.g., “2008-01-02 09:35:00”);
- Open: the opening price at the beginning of a 5 min subinterval (e.g., “1470.17”);
- High: the highest price reached within a 5 min subinterval (e.g., “1470.17”);
- Low: the lowest price reached within a 5 min subinterval (e.g., “1467.88”);
- Close: the closing price at the end of the 5 min subinterval (e.g., “1469.49”).

The intraday (5 min interval) SPX dataset used in this study covered the period from 2 January 2008 to 15 April 2025 and contained a total of 3,467,665 5 min intervals, which was equal to approximately 4350 trading days.

Table 1. SPX 5 min raw data snapshot.

Datetime	Open	High	Low	Close
2008-01-02 09:35:00	1470.17	1470.17	1467.88	1469.49
2008-01-02 09:40:00	1469.78	1471.71	1469.39	1471.22
2008-01-02 09:45:00	1471.56	1471.77	1470.69	1470.78
2008-01-02 09:50:00	1470.28	1471.06	1470.1	1470.74
2008-01-02 09:55:00	1470.81	1470.81	1468.42	1469.37

Figure 2 illustrates the evolution of the SPX 5 min closing price from 2008 to 2025, highlighting significant financial crisis periods and prolonged bull markets, as observed through high-frequency market data.

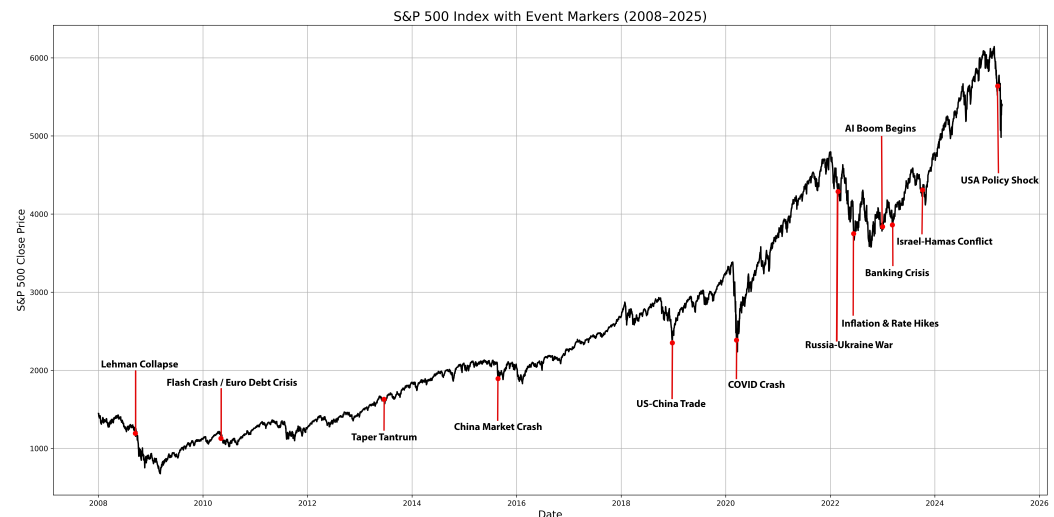


Figure 2. The evolution of the SPX index based on 5 min closing prices (P) since January 2008.

Several factors contributed to these changes, including supply and demand differences in the market. Overall, the primary factors were as follows:

- Last-minute updates or new information, such as macroeconomic announcements and earnings reports;
- Herding behavior or panic selling, plausibly caused by changes in investor sentiment [53];
- Changes in market liquidity and trading intensity reinforce price volatility, particularly during periods of stress [54].

Such factors make the modeling of time-dependent variance (volatility) a critical component in financial econometrics.

One of the primary reasons for using such extensive data is that to perform a detailed analysis, model evaluation should extend over longer periods that include periods of high and low volatility. The data used in this study contained the following major events that had a significant impact on the stock market and the SPX index between 2008 and 2025:

- The great financial crisis in 2008–2009.
- The European sovereign debt crisis in 2010–2012.
- The “Taper Tantrum” in 2013.
- The Chinese stock market turmoil and oil price crash in 2015–2016.
- The U.S.–China trade war in 2018.
- The COVID-19 pandemic in 2020.
- Inflation, aggressive rate hikes, and the Russia–Ukraine war in 2022. A combination of factors created a bear market for most of 2022:
 - High inflation;
 - Monetary tightening;
 - Russia–Ukraine conflict.
- The U.S. regional banking crisis and Israel–Hamas conflict in 2023.
- The Artificial Intelligence (AI) boom in 2023–2025.
- Weakening corporate outlook, political uncertainty, and Israel–Iran 12-day war in 2025.

Using intraday frequency data provided temporal granularity, a key component for constructing realized variance. Additionally, it provided detailed observations on long-term price evolution. As HAR-type models capture nonlinear, clustered, and shock-sensitive features of financial markets, the data frame was well suited for examining realized variance through a multi-horizon lag structure.

2.2. Return and Realized Measure Construction

Theoretically, volatility is an instantaneous variable. Hence, it is possible to study it in a continuous-time diffusive environment. Therefore, the focus was on a single financial asset and the process modeling of its logarithmic price ($\log P_t$), assuming it continuously evolved in the market during trading hours. Furthermore, in this article, the closing price of a single asset at time t was denoted as P .

If $P_{t-1+i\Delta}$ is the closing price at the end of the i th subinterval, $P_{t-1+(i-1)\Delta}$ the price at the beginning of the i th subinterval, $t-1$ the start of the current day, t the end of the current day, Δ the time interval in fractions of a day, and M the total number of subintervals [25], then $r_{t,i}$, the logarithmic return for the i th subinterval during trading day t , can be expressed as Equation (1):

$$r_{t,i} = \log P_{t-1+i\Delta} - \log P_{t-1+(i-1)\Delta}, \quad i = 1, 2, \dots, M, \quad (1)$$

where the sampling frequency, $M = 1/\Delta$, is the number of subintervals in a day. For instance, with 5 min of data in a 6.5 h (390 min) trading day, this equals $M = 390/5 = 78$. The choice of interval length affects the bias–variance trade-off in estimating the realized variance, as presented in Table 2. Empirical studies show that 5 min intervals generally minimize the Mean Squared Error when estimating daily variance using intraday returns [55].

Table 2. Intraday frequencies and subinterval counts (M).

Interval Length (Δ)	Subintervals per Day	Notes
1 min	390	Ultra-high frequency
2 min	195	
3 min	130	
5 min	78	
10 min	39	Standard for RV computation
15 min	26	
30 min	13	
60 min (1 h)	6.5	
195 min	2	Usually rounded to 6 or 7
390 min	1	Half-day intervals
		One daily observation (EOD)

Classical models such as Ordinary Least Squares (OLS) and many robust regressions often rely on assumptions of normally distributed errors. However, in high-frequency financial data, returns typically exhibit fat tails and volatility clustering, violating these assumptions and motivating the use of alternative or robust modeling techniques. While the classical HAR model estimated via OLS relies on assumptions that are violated in high-frequency financial data (e.g., normality, homoskedasticity), it remains widely used due to its relative success in forecasting realized variance. Nevertheless, robust and distribution-flexible extensions of HAR are adopted to address fat-tailed and heteroskedastic error structures in practice.

Researchers have discovered that logarithmic returns exhibit approximately Gaussian properties at intraday levels, making them preferable for volatility modeling [11,56]. One of the advantages of taking the logarithm of price differences is that it reduces skewness in the return distribution and helps to stabilize variance, especially in volatile market conditions. Logarithmic return transformation enhances the statistical properties of the data, making it a suitable input feature for regression-based models, such as OLS, WLS, and robust HAR variants. Thus, these estimation techniques become more reliable and less

sensitive to outliers and sudden price jumps. Considering this, Equation (2) describes the daily logarithmic return for the active part of trading day t .

$$r_t = \sum_{i=1}^M r_{t,i}, \quad (2)$$

where r_t is the total return for the day t , and M is the number of subintervals in a day.

Figure 3 illustrates how the SPX 5 min log returns change over the sample period. The series captures intraday return fluctuations over approximately 3.4 million intervals. The plot highlights several important empirical features that are commonly observed, such as volatility clustering in high-frequency financial data. The return values typically change around zero. The visible spikes reflect the crisis periods, such as the 2008 global financial crisis and the 2020 pandemic, among others. These spikes highlight the heavy-tailed and heteroskedastic character of return distributions by reflecting times of high uncertainty and abrupt price changes.

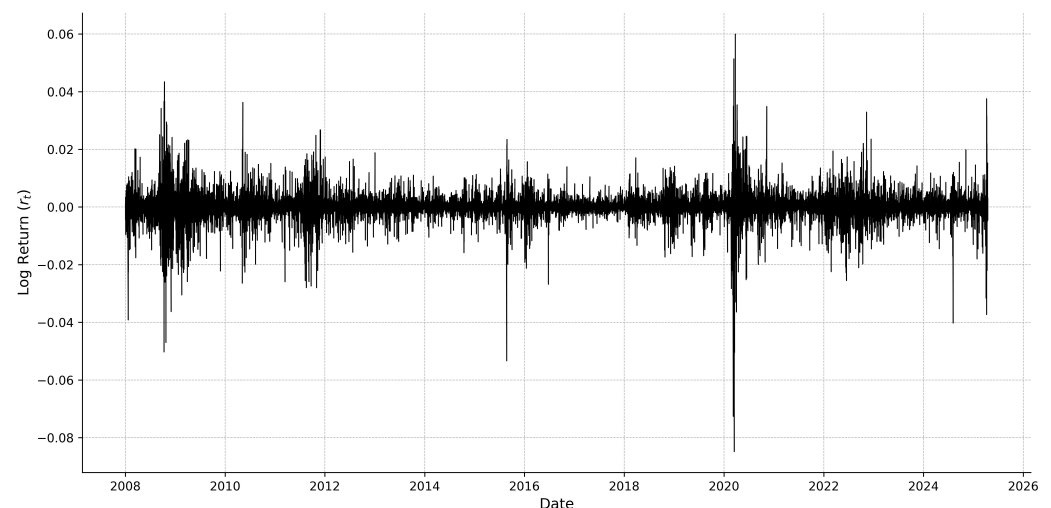


Figure 3. The 5 min log returns (r_t) of the SPX index since January 2008.

2.3. Stochastic Modeling of Log Price Dynamics

Before analyzing the dynamics of the logarithmic price process, it is essential to understand the underlying stochastic process that explains its evolution, namely, Brownian motion. Brownian motion, also known as the Wiener process, forms the foundation of continuous-time modeling in financial economics. In stochastic calculus, it is essential to the formulation of asset price dynamics. A standard Brownian motion is a continuous-time stochastic process W_t with the following properties [57]:

- $W_0 = 0$ —the starting point is always equal to zero.
- With probability 1, the function $t \rightarrow W_t$ is continuous over the period t .
- The process $\{W_t\}_{t \geq 0}$ has stationary, independent increments.
- The increment $W_{t+s} - W_t$ follows a NORMAL (0, t) distribution.

A standard d -dimensional Wiener process is a vector-valued stochastic process that can be defined as Equation (3):

$$W_t = (W_t^{(1)}, W_t^{(2)}, W_t^{(3)}, \dots, W_t^{(d)}), \quad (3)$$

where components $W_t^{(i)}$ are independent, standard one-dimensional Wiener processes.

Determining how the logarithm of the asset price changes over time is one of the fundamental problems in financial asset price modeling. The majority of forecasting and

volatility estimation techniques, including the HAR family of models, make forecasts based on hypotheses of a continuous-time stochastic process that underlies returns. The conventional constant-volatility model and the more realistic stochastic volatility (SV) framework are the two main modeling frameworks. These models differ in their handling of return variance, which has important ramifications for both the empirical validity of financial models and the precision of volatility predictions.

2.3.1. Classical Approach: Geometric Brownian Motion (GBM)

Equation (4) describes how the log of the price (P) of a single asset changes continuously over time during the trading day:

$$d \log(P_t) = \mu_t dt + \sigma_t dW_t, \quad (4)$$

where μdt is the instantaneous drift term, which represents the expected direction of the logarithmic price. $\sigma_t = \sqrt{\sum r_t^2}$ is the volatility process that inflates the price change concerning W_t , and W is a standard Brownian motion. $\sigma_t dW_t$ is the random movement of prices where volatility comes from. We can see the change in the log price ($\log(P_t)$) in Figure 4.

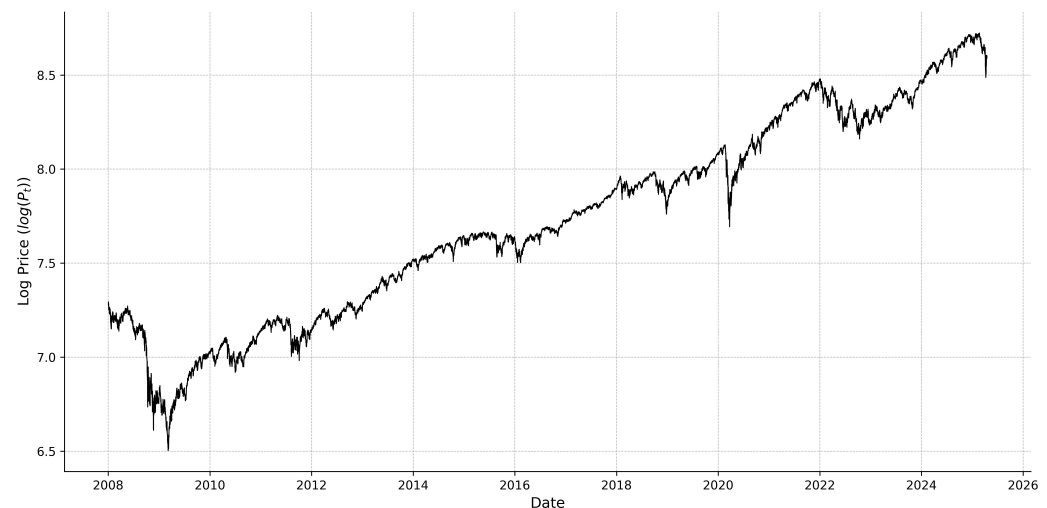


Figure 4. SPX intraday 5 min ($\log P_t$) over time.

Due to its closed form and ease of analysis, this formulation has long been a cornerstone of financial theory and frequently linked to the Black–Scholes option pricing model. Nevertheless, it assumes that volatility is always constant, with normally distributed returns, which do not accurately reflect practical scenarios [58]. Thus, in empirical finance, these assumptions are considered oversimplified. The classic GBM framework poorly captures such behaviors in actual markets that exhibit volatility clustering, leptokurtosis (fat tails), and time-varying risk [59]. Following this, Equation (5) defines W_t for the current application:

$$dW_t = \frac{d \log(P_t) - \mu_t dt}{\sigma_t}, \quad (5)$$

where dW_t is the increment in the Brownian motion process, $d \log(P_t)$ is the change in the logarithm of the asset return price, μ_t is the instantaneous drift term, and σ_t is the instantaneous volatility. Given the use of high-frequency intraday data, the instantaneous drift component $\mu_t dt$ in Equation (4) becomes negligible in comparison to the variance-driven component of price evolution. Both the theoretical and empirical literature support this assumption [11], resulting in the reduction of Equation (5) to Equation (6), which

emphasizes that the underlying Brownian motion W_t , scaled by instantaneous volatility σ_t , can be represented by normalized log price increments:

$$dW_t = \frac{d \log(P_t)}{\sigma_t} \quad (6)$$

2.3.2. Generalized Approach: Stochastic Volatility Models

To overcome the limitations of the classic GBM approach, researchers have developed stochastic volatility models that simulate variance as a stochastic process rather than a fixed constant. Barndorff-Nielsen and Shephard notably proposed Equation (7) in [12]:

$$dy^*(t) = \{\mu + \beta\sigma^2(t)\}dt + \sigma(t)dw(t), \quad (7)$$

where $y^*(t) = \log(P_t)$ is the logarithmic price; μ is the drift term, representing the expected rate of change in the logarithmic price; $\sigma^2(t)$ is the instantaneous variance that evolves stochastically over time; β reflects a risk premium component that increases the drift in proportion to the current level of variance; and $dw(t)$ is a standard Brownian motion. $\sigma(t)$ is the instantaneous (spot) volatility, modeled as a latent stochastic process that evolves continuously over time. Unlike constant-volatility models, the SV framework allows $\sigma^2(t)$ to change in response to shocks, mean-revert, or persistent behavior, which reflects actual market dynamics even more accurately. The stochastic volatility approach addresses some key empirical features that the classic GBM approach is unable to capture, such as volatility clustering. Additionally, the inclusion of $\beta\sigma^2(t)$ in the drift term introduces a risk–return trade-off, which allows expected returns to change proportionally to volatility. By allowing stochastically variant changes, the model is now capable of replicating heavy tails. Barndorff-Nielsen and Shephard defined the SV model over small intervals as Equation (8) [12]:

$$y_i = y^*(i\Delta) - y^*((i-1)\Delta), \quad (8)$$

where y_i is the change in the logarithmic price over the i th subinterval of length Δ during the trading day. This approach is equivalent to Equation (9), assuming that each return is conditionally normally distributed, as per the SV model:

$$r_{t,i} | \sigma_{t,i}^2 \sim \mathcal{N}(\mu_i, \sigma_{t,i}^2), \quad (9)$$

where $r_{t,i}$ is the return for the i th subinterval of day t , $\sigma_{t,i}^2$ is the instantaneous variance at that subinterval, and μ_i is the drift term for that subinterval. The actual latent volatility of the asset at time t is not directly observable. Instead, the objective is to estimate the total variation in volatility accumulated over a day, known as the integrated variance (IV) (see Equation (10)).

$$IV_t = \int_{t-1}^1 \sigma^2(t_s) dt_s, \quad (10)$$

where IV_t is the integrated variance for trading day t , and $\sigma^2(t_s)$ is the instantaneous variance at time t_s within that day. This quantity captures the total uncertainty or risk in asset returns for day t . It is the theoretical limit of the sum of intraday return variances as the sampling frequency increases. In volatility models such as HAR or HARQ, the instantaneous drift (μ_t) is not the focus—the randomness or volatility part is, because the drift is slight compared to volatility at high frequencies. A key insight from stochastic volatility (SV) theory is that the integrated variance over a given time interval can be recovered directly from the sample path of the log price process.

In reality, the instantaneous variance is unobservable due to the discrete and noisy nature of market data, analogous to measuring continuous stochastic fields with limited-

resolution instruments in statistical physics. This idea parallels projection methods in nonequilibrium systems and has been formalized in financial econometrics by [12,60], echoing techniques developed in statistical physics by [61]. Moreover, observed high-frequency returns nonparametrically estimate the integrated variance without requiring knowledge of the drift (μ) or the precise form of the stochastic process governing $\sigma(t)$. It is model-free regarding the volatility dynamics and is consistent as long as the sampling frequency is sufficiently high. One way to estimate it is by using realized variance (RV), as shown in Equation (11):

$$RV_t = \sum_{i=1}^M r_{t,i}^2 \approx IV_t = \int_{t-\Delta t}^t \sigma^2(t_s) dt_s, \quad \text{as } M \rightarrow \infty \text{ or } \Delta t \rightarrow 0, \quad (11)$$

where RV_t is the realized variance for trading day t , $r_{t,i}$ is the return for the i th subinterval of day t , and Δt is the length of the subinterval in fractions of a day. Realized variance in the trading day t can be defined as the sum of the squared returns. The realized variance (RV) is a reasonable estimate of the actual volatility over a day using high-frequency data. This connection is foundational to the HAR-RV framework and the use of high-frequency data in volatility forecasting. Nevertheless, even then, there is some error in that estimate (see Equation (12)). Under certain conditions on M , the sampling frequency and estimation error η_t defined in [12] can be determined using Equation (13):

$$RV_t = IV_t + \eta_t, \quad (12)$$

$$\frac{\eta_t}{\sqrt{(2\Delta IQ_t)}} = \frac{RV_t - IV_t}{\sqrt{(2\Delta IQ_t)}}, \quad (13)$$

where η_t is the estimation error, RV_t is the realized variance for trading day t , IV_t is the integrated variance for trading day t , and IQ_t is the integrated quarticity for trading day t .

Assuming the error term is scaled by the integrated quarticity, its form follows a standard normal distribution, $\mathcal{N}(0,1)$. This normalization implies that realized quarticity (RQ) can serve as an estimator of integrated quarticity (IQ). In this context, Equations (14) and (15) formalize the definitions of RQ and IQ, respectively.

$$RQ_t = \frac{M}{3} \sum_{i=1}^M r_{t,i}^4, \quad (14)$$

where RQ_t is the realized quarticity for trading day t , and $r_{t,i}$ is the return for the i th subinterval of day t . The factor $\frac{M}{3}$ is a scaling factor.

$$IQ_t = \int_{t-1}^t \sigma_s^4 ds, \quad (15)$$

where IQ_t is the integrated quarticity for trading day t , and σ_s^4 is the fourth power of the instantaneous volatility at time s . As the sampling frequency increases, RV becomes closer to IV, but the estimation error (η_t) remains. Equation (12) scales the estimator error because it is directly uninterpretable. To statistically interpret this error meaningfully, it should be differently scaled, which leads to results with a known distribution and allows for confidence intervals.

The incorporation of higher-order moments, such as integrated quarticity, has been used for refining the statistical properties of realized measures and addressing the scaling of estimation errors. In light of this, attention is next directed toward the Heterogeneous Autoregressive (HAR) class of models, wherein realized volatility components across multiple horizons are utilized for modeling and forecasting return variance.

2.4. Heterogeneous Autoregressive-Type Model Specifications

This section presents the Heterogeneous Autoregressive (HAR) model, adopted in this study as the baseline model and originally proposed by [17]. Being a pioneering study, Corsi's work demonstrated that if realized variance was modeled on the basis of multiple heterogeneous time horizons, it would be possible to effectively account for long-memory and volatility persistence in financial markets. According to this setup, this research adds to the original HAR model specifications by including realized quarticity and the VIX index as an external variable that is forward-looking in terms of market expectations. These modifications are intended to improve the model's predictive power, particularly when uncertainty is heightened.

2.4.1. Heterogeneous Autoregressive Model of Realized Volatility

Since high-frequency intraday data are widely available, current researchers have considered using realized volatility to create forecasting models for the time-varying return volatility. Because of its ease of use and reliable prediction results, the HAR model has become one of the most popular forecasting models in practical applications. Corsi first introduced the HAR model [17], derived from a simple expansion of heterogeneous models such as ARCH or HARCH examined by Müller et al. [62]. To parameterize the conditional variance of the discretely sampled returns, this method uses the squared returns across longer and/or shorter forecasting horizons, along with the delayed squared returns over the same horizon. Assuming a typical trade week and month are 5 and 22 days, respectively, as a linear function of the daily, weekly, and monthly realized variance components, RV is specified in the original HAR model and defined as Equation (16):

$$RV_t = \beta_0 + \beta_1 RV_{t-1}^{(d)} + \beta_2 RV_{t-1}^{(w)} + \beta_3 RV_{t-1}^{(m)} + e_t \quad \begin{cases} RV_{t-1}^{(d)} \equiv RV_{t-1}, \\ RV_{t-1}^{(w)} \equiv \frac{1}{5} \sum_{i=1}^5 RV_{t-i}, \\ RV_{t-1}^{(m)} \equiv \frac{1}{22} \sum_{i=1}^{22} RV_{t-i}, \end{cases} \quad (16)$$

where RV_t is the realized variance of a trade day t , and $RV_{t-1}^{(d)}$, $RV_{t-1}^{(w)}$, and $RV_{t-1}^{(m)}$ are the daily, weekly, and monthly realized variances at time $t - 1$, respectively. The β coefficients are the parameters to be estimated, and e_t is the zero mean error term. This RV definition parsimoniously captures the strong persistence seen in most realized variance series.

2.4.2. Heterogeneous Autoregressive Model of Realized Quarticity

Realized Quarticity (RQ) provides a deeper insight into the dynamics of the volatility process. While Realized Variance (RV) measures the magnitude of price movements, RQ quantifies the "volatility of volatility." It shows how stable the overall market risk is. A high RQ suggests that the level of risk is not only elevated but also highly unstable and prone to sudden jumps. By incorporating this measure, the HARQ model is expected to gain forecasting advantages. The primary benefit is the model's enhanced ability to adapt to changing market structures, allowing it to better distinguish between periods of high but persistent volatility and periods where the variance process itself is erratic. This adaptiveness is anticipated to yield more accurate and reliable forecasts, particularly during episodes of market stress or significant structural shifts. Bollerslev et al. proposed the extended HAR by (typically) estimating it with OLS and considering the error that

occurs from RV estimation by employing Realized Quarticity, which gave the model its name, HARQ [22]. Equation (17) describes the extended HAR:

$$RV_t = \beta_0 + \left(\beta_1 + \beta_{1Q} \sqrt{RQ_{t-1}^{(d)}} \right) RV_{t-1}^{(d)} + \left(\beta_2 + \beta_{2Q} \sqrt{RQ_{t-1}^{(w)}} \right) RV_{t-1}^{(w)} + \left(\beta_3 + \beta_{3Q} \sqrt{RQ_{t-1}^{(m)}} \right) RV_{t-1}^{(m)} + e_t, \quad (17)$$

where $RQ_{t-1}^{(d)}$, $RQ_{t-1}^{(w)}$, and $RQ_{t-1}^{(m)}$ are the daily, weekly, and monthly realized quadratic lags at time $t - 1$, respectively. The β coefficients are the parameters to be estimated, and e_t is the zero mean error term. For short-term forecasting, Bollerslev proposed using the daily lag, as described by Equation (18). Such a change in the equation is helpful since the estimation error is primarily responsible for the attenuation bias in the predictions (RV is less persistent than IV). When the measurement uncertainty recorded by RQ is larger, this paradigm reduces the weight of the past RV observations:

$$RV_t = \beta_0 + \left(\beta_1 + \beta_{1Q} \sqrt{RQ_{t-1}^{(d)}} \right) RV_{t-1}^{(d)} + \beta_2 RV_{t-1}^{(w)} + \beta_3 RV_{t-1}^{(m)} + e_t, \quad (18)$$

where $RQ_{t-1}^{(d)}$ is the daily realized quadratic lag at time $t - 1$. The β coefficients are the parameters to be estimated, and e_t is the zero mean error term.

Exogenous variables can further extend the HARQ model. One such variable is VIX, which stands for the Volatility Index and gauges the market's expectation of future volatility over the coming 30 days. Sometimes, others refer to this index as the “fear index” since it rises as investors anticipate market uncertainty or stability. The Chicago Board Options Exchange (CBOE) determines this index from the S&P 500 Index option prices. It reflects the implied volatility of S&P 500 index options and does not consider the past volatility as RV . The introduction of VIX to the HARQ model results in the Heterogeneous Autoregressive Model with Realized Quarticity and Exogenous Variables (HARQ-X) model expressed in Equation (19):

$$RV_t = \beta_0 + \left(\beta_1 + \beta_{1Q} \sqrt{RQ_{t-1}^{(d)}} \right) RV_{t-1}^{(d)} + \beta_2 RV_{t-1}^{(w)} + \beta_3 RV_{t-1}^{(m)} + \beta_4 VIX_{t-1} + e_t, \quad (19)$$

where VIX_{t-1} is the VIX value at time $t - 1$. The β coefficients are the parameters to be estimated, and e_t is the zero mean error term.

With the HAR, HARQ, and HARQ-X specifications introduced, attention must next be directed to the estimation procedures used to evaluate these models empirically. Given the diverse statistical characteristics of high-frequency realized variance, the choice of estimator determines the robustness and accuracy of the forecasting results. In this study, no additional constraints were applied for the estimation. Therefore, the smoothing phenomenon was out of scope when assessing relative volatility.

2.5. HAR Model Estimation and Evaluation

The OLS estimator commonly estimates the HAR model, where realized variance is the dependent variable. While it is a standard empirical application because of some well-known statistical properties of RV , OLS is not always the best choice. The stylized facts of realized variance, such as spikes/outliers, conditional heteroskedasticity, and non-Gaussianity, violate some of the most fundamental assumptions of the OLS model [25]. Therefore, OLS can only be considered the best estimator under the Gauss–Markov assumptions [63]. Consequently, OLS estimation results are inefficient, finitely sample-biased, and overly sensitive to outliers. Alternative estimation techniques such as Weighted Least

Squares (WLS), Robust Linear Model (RLM), and Least Absolute Deviation (LAD) were studied in this work to address these limitations.

2.5.1. Ordinary Least Squares

OLS is a linear regression method that estimates the parameters of a linear model by minimizing the sum of the squared differences between the observed and predicted values. For HAR model estimations, OLS finds $\beta = (\beta_1, \beta_2 \dots \beta_n)^\top$ coefficients by minimizing the Residual Sum of Squares (RSS) defined by Equation (20):

$$\hat{\beta} = \arg \min_{\beta} \sum_{t=23}^T \left(RV_t - \left(\beta_0 + \beta_1 RV_{t-1}^{(d)} + \beta_2 RV_{t-1}^{(w)} + \beta_3 RV_{t-1}^{(m)} \right) \right)^2, \quad (20)$$

where RV_t is the realized variance at time t , and $RV_{t-1}^{(d)}$, $RV_{t-1}^{(w)}$, and $RV_{t-1}^{(m)}$ are the daily, weekly, and monthly realized variances at time $t - 1$, respectively. The β coefficients are the parameters to be estimated. When considering asymptotic efficiency, OLS is the leading estimator for β , assuming the errors in autoregressions are independent, normally (Gaussian) distributed, and homoskedastic. Unfortunately, in financial volatility modeling, residuals are often neither normally distributed nor independent, and volatility shows heteroskedasticity. Therefore, while OLS can provide unbiased estimates under milder conditions, its efficiency and robustness are questionable in practice [63].

In this work, the logarithmic realized variance ($\log RV$) served as both a proxy and a target for forecasting. There are several advantages to using $\log RV$, one of which is avoiding negative values. Another positive feature is that it reduces skewness, making the data more symmetrical and closer to a normal distribution. Therefore, Equation (21) represents the reformulated minimization problem:

$$\hat{\beta} = \arg \min_{\beta} \sum_{t=23}^T \left(\begin{array}{l} \log(RV_t) - \left(\beta_0 + \beta_1 \cdot \log(RV_{t-1}^{(d)}) + \right. \\ \left. \beta_2 \cdot \log(RV_{t-1}^{(w)}) + \beta_3 \cdot \log(RV_{t-1}^{(m)}) + \right. \\ \left. \beta_4 \cdot (\log(RV_{t-1}) \cdot \log(RQ_{t-1})) + \beta_5 \cdot \log(VIX_{t-1}) \right) \end{array} \right)^2, \quad (21)$$

This minimization problem represents the estimation of an HARQ model with the exogenous variable VIX , where the dependent variable is the logarithm of the realized variance of day t , denoted as $\log(RV_t)$. The regressors include various lagged components, such as daily, weekly, and monthly realized variance terms— $\log(RV_{t-1}^{(d)})$, $\log(RV_{t-1}^{(w)})$, and $\log(RV_{t-1}^{(m)})$, respectively—that capture the heterogeneous memory of financial variance. The inclusion of the interaction term $(\log(RV_{t-1}) \cdot \log(RQ_{t-1}))$ allows the model to adjust for measurement errors in the realized variance, as advocated by Bollerslev, Patton, and Quaedvlieg [22].

2.5.2. Weighted Least Squares

Weighted Least Squares (WLS) is a regression technique that extends OLS by allowing for heteroskedasticity in the error terms. In WLS, each observation is assigned a weight, which can be used to correct for non-constant variance in the residuals. While OLS minimizes the sum of squared residuals (see Equation (22)), WLS minimizes the weighted sum of squared residuals, where each residual is scaled by a weight w_t (see Equation (23)):

$$\hat{\beta} = \arg \min_{\beta} \sum_{t=23}^T (RV_t - \widehat{RV}_t)^2, \quad (22)$$

$$\hat{\beta} = \arg \min_{\beta} \sum_{t=23}^T w_t (RV_t - \widehat{RV}_t)^2; \quad (23)$$

where β represents the regression coefficients that this model minimizes, w_t is the weight assigned to the observation at time t , and \widehat{RV}_t is the predicted value of the realized variance at time t . Such an approach helps reduce the influence of more volatile or noisier periods. For this reason, it is well suited for the HAR model and can be defined by Equation (24):

$$\hat{\beta} = \arg \min_{\beta} \sum_{t=23}^T w_t \left(\begin{array}{c} \log(RV_t) - (\beta_0 + \beta_1 \cdot \log(RV_{t-1}^{(d)}) + \\ \beta_2 \cdot \log(RV_{t-1}^{(w)}) + \beta_3 \cdot \log(RV_{t-1}^{(m)}) + \\ \beta_4 \cdot (\log(RV_{t-1}) \cdot \log(RQ_{t-1})) + \beta_5 \cdot \log(VIX_{t-1})) \end{array} \right)^2, \quad (24)$$

where w_t is the weight assigned to the observation at time t . The WLS estimator outperforms the OLS estimator if each weight (w_t) is inversely proportional to the conditional variance of the associated error (e_t). In this approach, errors that are likely to be significant weigh less. Although Patton and Sheppard employed a straightforward WLS approach to estimate the HAR model, they did not consider other options [19].

2.5.3. Least Absolute Deviations

The Least Absolute Deviation (LAD) estimator is another robust alternative to the Ordinary Least Squares (OLS) method for linear regression analysis. Whereas OLS minimizes the sum of the squared residuals, LAD minimizes the sum of the absolute residuals and is therefore less sensitive to outliers in the dependent variable [64].

Describing explanatory variables as vector X_{t-1} (see Equation (25)), estimation coefficients as β (see Equation (27)), and the dependent variable as y_t (see Equation (26)), we have:

$$X_{t-1} = \begin{pmatrix} 1, \log(RV_{t-1}^{(d)}), \log(RV_{t-1}^{(w)}), \log(RV_{t-1}^{(m)}), \\ (\log(RV_{t-1}) \cdot \log(RQ_{t-1})), \log(VIX_{t-1}) \end{pmatrix}, \quad X_{t-1} \in \mathbb{R}^k \quad (25)$$

$$y_t = \log(RV_t), \quad y_t \in \mathbb{R}^k \quad (26)$$

$$\beta = (\beta_1, \beta_2 \dots \beta_n)^T; \quad \beta \in \mathbb{R}^k \quad (27)$$

where k is the number of explanatory variables, y_t is the dependent variable, and X_{t-1} is the vector of explanatory variables at time $t - 1$. The LAD estimator $\hat{\beta}$ is defined as the solution to the optimization problem described in Equation (28):

$$\hat{\beta} = \arg \min_{\beta} \sum_{t=23}^T |y_t - X_{t-1} \cdot \beta|, \quad (28)$$

where X_{t-1} is the vector of explanatory variables at time $t - 1$, and y_t is the dependent variable at time t . This equation can be transformed and used as Equation (29):

$$\hat{\beta} = \arg \min_{\beta} \sum_{t=23}^T \tau(y_t - X_{t-1} \cdot \beta), \quad (29)$$

where $0 < \tau < 1$ in quantile regression represents the intended quantile level estimation. Let u_t be the error; then, it can be described as Equation (30):

$$u_t = y_t - X_{t-1} \cdot \beta, \quad (30)$$

where u_t is the error term at time t . When $\tau = 0.5$, it is referred to as median regression, which is equivalent to LAD regression. Therefore, in this implementation, $\tau = 0.5$ was used for quantile regression. The mathematical representation of the absolute value function $\tau(u)$ is defined as Equation (31). This function is used to calculate the absolute error in the LAD regression.

$$\tau(u) = |u| = \begin{cases} u, & \text{if } u \geq 0 \\ -u, & \text{otherwise} \end{cases}, \quad (31)$$

where u is the error term at time t . The linear and symmetric penalty of the absolute function accounts for both overestimates and underestimates. As there are outliers and a heavy-tailed error distribution, which is common for financial time series, this feature of LAD enhances its robustness. However, due to its non-differentiability at zero, LAD estimation typically requires optimization techniques such as linear programming. Despite this, its resilience to extreme values makes it especially suitable for volatility modeling and robust econometric forecasting.

2.5.4. Robust Regression Models

Although OLS is a good option in ideal circumstances, in real life, unideal conditions make this model extremely sensitive to outliers or odd observations in the data. More reliable estimators, such as the widely used M-estimator introduced by Huber, have thus been suggested as substitutes [65]. The M-estimator of β solves the HAR model's minimization problem and is defined by Equation (32):

$$\hat{\beta} = \arg \min_{\beta} \sum_{t=23}^T \rho(y_t - X_{t-1} \cdot \beta), \quad (32)$$

where ρ is a symmetric function that has a unique minimum at zero and is predefined, y_t is the dependent variable at time t , and X_{t-1} is the vector of explanatory variables at time $t - 1$. In the presence of outliers or heteroskedastic errors, RLM is a powerful tool to estimate coefficients. It represents the idea of minimizing a robust objective function. The Minkowski loss function [66,67], defined in Equation (33), provides a flexible and powerful alternative to the traditional loss functions used in the regression analysis employed in this research. Thus, it was used to estimate the model when applied as a custom function in this study.

$$L = \sum_{t=23}^T |RV_t - \widehat{RV}_t|^p, \quad (33)$$

where L is the loss function, RV_t is the realized variance at time t , \widehat{RV}_t is the predicted value of the realized variance at time t , and p is the sensitivity parameter of the Minkowski distance. One of the advantages of the Minkowski loss function is that it generalizes both the Least Absolute Deviation (LAD) and Ordinary Least Squares (OLS) approaches by allowing the exponent p to be tuned. As seen from Equation (33), the choice of p directly influences the estimator's sensitivity to outliers and the tail behavior of the error distribution. There is a trade-off between robustness and efficiency in this process. When $p = 1$, this function is as sensitive as the LAD (Manhattan norm); when $p = 2$, this function acts like the OLS (Euclidean norm) estimator. Otherwise, when $1 < p < 2$, this function

becomes more sensitive to outliers than LAD but less sensitive than OLS, which means the Minkowski loss reduces the influence of extreme observations without completely disregarding them. Fractional values of p , such as $p = 1.2$ or $p = 1.4$, offer a balance that avoids instability or inefficiency problems. These features make it especially valuable for financial time series analysis.

The continuous nature of the Minkowski norm also enables fine-tuned control over the estimator's properties, something that is hard to achieve using discrete robust loss functions such as Huber's or Tukey's bi-weight. This tunability enables practitioners to calibrate the model more precisely to the data's structure, thereby enhancing both forecast accuracy and model reliability. For the implementation of this method in this work, the minimization problem can be described by Equation (34):

$$\hat{\beta} = \arg \min_{\beta} \sum_{t=23}^T |y_t - X_{t-1} \cdot \beta|^p, \quad (34)$$

where y_t is the dependent variable at time t , X_{t-1} is the vector of explanatory variables at time $t - 1$, and β is the vector of coefficients to be estimated.

2.5.5. Entropy Loss Function

The Entropy Loss Function (ELF) is a robust alternative to the traditional loss functions used in regression analysis. To estimate the HAR model coefficients, an alternative approach—the Entropy Loss Function [68,69]—was also applied. The Entropy Loss Function (ELF) applies both squared and relative error penalties. The ELF is defined by Equation (35), which minimizes the entropy of the residuals:

$$\hat{\beta} = \arg \min_{\beta} L = \arg \min_{\beta} \sum_{t=23}^T \left[\frac{(y_t - \hat{y}_t)^2 \cdot (1 - K_{exp})}{2 \cdot y_t^2} + \frac{(y_t - \hat{y}_t)^2 \cdot K_{exp}}{2} \right], \quad (35)$$

where L is the loss function, y_t is the actual value of the logarithmic RV at time t , \hat{y}_t is the predicted value of the dependent variable at time t , and K_{exp} is a user-defined constant weighting parameter governing the trade-off between relative and absolute error penalization. The assumption of this study is that $y_t = RV_t$ and $\hat{y}_t = \widehat{RV}_t$. Rearranging this equation results in Equation (36):

$$L_t = \frac{(y_t - \hat{y}_t)^2}{2} \cdot \left(\frac{1 - K_{exp}}{y_t^2} + K_{exp} \right), \quad (36)$$

where L_t is the loss function, y_t is the actual value of the logarithmic RV at time t , \hat{y}_t is the predicted value of the dependent variable at time t , and K_{exp} is a user-defined constant weighting parameter governing the trade-off between relative and absolute error penalization.

After rearranging the equation with the variables defined for this work, the minimization problem can be described by Equation (37):

$$\hat{\beta} = \arg \min_{\beta} L = \arg \min_{\beta} \sum_{t=23}^T \left[\frac{(y_t - X_{t-1} \cdot \beta)^2}{2} \cdot \left(\frac{1 - K_{exp}}{y_t^2} + K_{exp} \right) \right], \quad (37)$$

where y_t is the dependent variable at time t , X_{t-1} is the vector of explanatory variables at time $t - 1$, β is the vector of coefficients to be estimated, and K_{exp} is a user-defined constant weighting parameter governing the trade-off between relative and absolute error penalization. Figures 5 and 6 illustrate the pseudo-code of the definition and the implementation of the Entropy Loss Function, respectively.

```

Function custom_loss(beta, X, y, k_exp):
    y_pred = X * beta                # predicted log(RV2)
    error_sq = (y - y_pred)^2        # squared error

    epsilon = small constant         # to avoid division by
    zero

    term1 = error_sq * (1 - k_exp) / (2 * (y^2 + epsilon))
    term2 = error_sq * k_exp / 2
    L = sum(term1 + term2)           # total log-likelihood
    return L

```

Figure 5. Pseudo-code of the custom Entropy Loss Function.

```

Given:
    X = regressors (lags of log(RV2))
    y = target (future log(RV2))
    k_exp = weight parameter for error term scaling

Step 1: Initialize beta = [0, 0, 0, ...]
Step 2: Optimize: minimize (custom_loss(beta, X, y, k_exp))
        → Use scipy.minimize (L-BFGS-B)
Step 3: Return beta_hat (estimated coefficients)
Step 4: Use beta_hat for forecasting

```

Figure 6. Pseudo-code of the estimation procedure using custom loss.

The parameter $K_{exp} \in [0, 1]$ determines the trade-off between the relative loss when $K_{exp} \rightarrow 0$ and the squared error loss when $K_{exp} \rightarrow 1$. When applied to volatility modeling, the ELF allows the model to remain sensitive to variance magnitudes while avoiding excessive penalization of large-scale fluctuations. The proposed method penalizes both the absolute and proportional deviations in a unified manner. As a result, the coefficient estimation process enhances both robustness and accuracy.

2.5.6. Summary

To summarize, the estimation techniques used in Section 2.5 have been employed to fit HAR-type models, emphasizing the need for robustness in the presence of financial data irregularities, such as outliers and heteroskedasticity. While OLS is a standard technique for this problem, it performs inaccurately due to its sensitivity to non-Gaussian errors. To address this limitation, two different approaches—WLS and LAD—were applied. Furthermore, the introduced Minkowski loss function balanced out the OLS and LAD. Additionally, an entropy-based loss function was applied, which also balanced absolute and relative error penalties. These techniques collectively strengthen the predictive power and resilience of HAR models under real-world market conditions.

2.6. Forecast Design

In this section, the forecasting procedure is discussed, including rolling window forecasting and multi-horizon forecast targets. The forecasting aims to generate out-of-sample predictions of future log realized variance.

2.6.1. Rolling Window Forecasting

The applied rolling window forecasting evaluates a model's genuine out-of-sample forecasting. In this approach, the model uses the latest available data each time and also prevents possible look-ahead bias. The forecasting procedure operates by estimating the model using a fixed-size training window consisting of the most recent 1000 daily observations at each iteration. Then, the loss function obtains model parameters. Such estimated coefficients generate a single or multi-step forecast of realized variance for the corresponding target horizon. Finally, following each forecast, the window is rolled forward by one day by discarding the oldest observation and including the newest one. This process continues until the entire out-of-sample evaluation period passes for forecasting. The result

is a sequence of rolling out-of-sample predictions. Each forecast uses only past data, which makes them directly comparable to observed actual realized variance values. This framework reflects how the model would work in real-time scenarios.

2.6.2. Forecasting Targets

The designed forecasting framework predicts the future value of the logarithm of realized variance. To reflect different investment and risk management perspectives, the resulting forecasts take multiple time horizons into account. This study set four distinct forecast targets for evaluation, namely, daily ($D = 1$), weekly ($D = 5$), biweekly ($D = 10$), and monthly ($D = 22$). The logarithmic scale of cumulative realized variance expresses each of the targets described in Equation (38):

$$y_t^D = \log \sum_{j=1}^D RV_{t+j}, \quad (38)$$

where y_t^D is the target variable at time t for horizon D , and RV_{t+j} is the forecasted realized variance at time $t + j$. For the empirical practice of HAR-type models, the abovementioned targets are common. To ensure compatibility with the regression framework and preserve the positive nature of the forecast variable, both actual realized variance and forecasted variance were log-transformed. The log transformation of the model features stabilized variance and improved the distributional properties of the model residuals. The non-overlapping rolling sums composed the target variables in order to maintain clarity and reduce autocorrelation in the forecast errors across horizons.

2.7. Evaluation Metrics

This study employed two performance metrics for the assessment of the forecasting results: Mean Absolute Error (MAE) and Quasi-Likelihood (QLIKE) loss. These metrics ensured that the evaluation was fair, unbiased, and methodologically rigorous. The Mean Absolute Error (MAE) (see Equation (10)) measures the average absolute deviation between the forecasted and actual values of log-realized variance. One of the main advantages of the MAE is that it is independent of model training under some specific objective function. This advantage means that models fitted with different estimation strategies are comparable on an equal basis. Another reason for using the MAE is that it is robust to outliers due to its linear penalization of error magnitudes, which is also important given the nature of financial time series. Since the target variable is on a log scale, the MAE provides a direct error measurement within the same domain, maintaining coherence with the regression specification.

$$MAE = \frac{1}{n} \sum_{t=1}^n \left| \log RV_t - \widehat{\log RV_t} \right|, \quad (39)$$

where n is the number of observations, $\log RV_t$ is the actual value of log-realized variance at time t , and $\widehat{\log RV_t}$ is the forecasted value of log-realized variance at time t .

The Quasi-Likelihood (QLIKE) loss provides a likelihood-based evaluation criterion tailored to variance forecasting. The derivation of the Gaussian Quasi-Likelihood function is the QLIKE function, which is known to be asymptotically optimal under conditional heteroskedasticity. Its key advantage is its asymmetric penalty structure. Suppose the variance under-prediction is penalized more severely than over-prediction. That makes it well aligned with risk-sensitive financial applications. For example, QLIKE is particularly relevant when model outputs aid in estimating risk management, portfolio allocation, or Value-at-Risk (VaR) estimation. Although the models use and predict variance in

log space, QLIKE cannot be directly applied to log RV forecasts and actuals. Therefore, the regression and forecast are in log RV; thus, to compute QLIKE, they must be changed back (Equations (40) and (41)) as follows:

$$RV_{actual} = \exp(\log RV_t), \quad (40)$$

$$RV_{forecast} = \exp(\widehat{\log RV_t}); \quad (41)$$

where RV_{actual} is the actual value of realized variance at time t , and $RV_{forecast}$ is the forecasted value of realized variance at time t . After this transformation, the formula for QLIKE can be expressed as Equation (42):

$$QLIKE = \frac{RV_{actual}^2}{RV_{forecast}^2} - \log\left(\frac{RV_{actual}^2}{RV_{forecast}^2}\right) - 1, \quad (42)$$

where RV_{actual} is the actual value of realized variance at time t , and $RV_{forecast}$ is the forecasted value of realized variance at time t .

Together, MAE and QLIKE maintain a balance between robustness, interpretability, and theoretical efficiency. MAE ensures fairness across models, regardless of the estimation strategy [70], while QLIKE provides a probabilistically grounded criterion sensitive to variance estimation accuracy. Their joint application supports a comprehensive and equitable evaluation framework for comparing the predictive performance of HAR-type models across different forecast horizons and estimation methods.

3. Results and Discussion

This section presents the empirical results of the HAR-type models examined in this study. Three different extensions of the HAR model were considered, namely, HAR-RV, HARQ, and HARQ-X, each with five different estimators used for performance comparison. The out-of-sample forecasting performance evaluation, using four horizons—daily, weekly, biweekly, and monthly—was based on aggregated realized variance. The parameter p of the RV estimator was tuned via exhaustive search to demonstrate the models' ability to achieve error minimization for different error criteria with different p values.

Table 3 presents a comparative analysis of the HAR-RV model with the estimators OLS, LAD, WLS, RLM, and ELF. According to the results, the primary contenders (the cell values highlighted in **bold**) for this model were the RLM and entropy estimation techniques. While the Entropy Loss Function consistently achieved the lowest QLIKE values across all forecasting horizons, suggesting its superiority in capturing variance dynamics, RLM with a custom Minkowski loss function ($p = 1.3$) demonstrated its accuracy in the MAE metric. Although the entropy estimator did not exhibit the best MAE performance, its results were not significantly different from those of other estimators. In contrast, this was not true for the RLM ($p = 1.3$) technique. Given the consistently strong performance of the entropy-based model in minimizing the QLIKE loss and achieving significant results in the MAE metric, the entropy technique is a robust alternative to conventional estimation techniques within the HAR-RV framework. On the other hand, the WLS technique delivered the best MSE values, which is not surprising, considering that it explicitly minimizes the weighted Mean Squared Error in the optimization problem.

Table 4 reflects relatively similar results, comparing the exact applied estimators with the same conditions applied for the HARQ model. Likewise, in this instance, the ELF demonstrated the best performance in the QLIKE metric due to its alignment with logarithmic likelihood-based forecasting. Furthermore, the RLM with a custom Minkowski loss function ($p = 1.3$) estimator delivered optimal MAE results for all forecasting horizons.

The WLS approach with RQ weight performed better, as indicated by MSE values similar to the previous investigation. In this experiment, the Entropy Loss Function resulted in the same magnitude of MAE as other estimation techniques. Therefore, these results lead to the conclusion that the Entropy Loss Function is well suited for the HARQ model.

Table 3. HAR-RV model's (Equation (16)) out-of-sample forecasting results using different estimation techniques.

	OLS	LAD	WLS _{RQ}	RLM $p = 2.1$	RLM $p = 1.3$	ELF $K_{exp} = 0.1$
Daily horizon						
QLIKE	0.342028	0.345926	0.342150	0.341260	0.345827	0.321536
MAE	0.579774	0.579553	0.579665	0.579930	0.579114	0.580111
MSE	0.541348	0.541339	0.541206	0.541537	0.540675	0.541350
Weekly horizon						
QLIKE	0.166367	0.177811	0.166370	0.165485	0.173596	0.165114
MAE	0.384559	0.382425	0.384370	0.385148	0.382333	0.385893
MSE	0.251494	0.255576	0.251327	0.251533	0.253271	0.251493
Biweekly horizon						
QLIKE	0.227243	0.248310	0.227262	0.227243	0.241540	0.223707
MAE	0.423649	0.421023	0.423434	0.424573	0.420693	0.426678
MSE	0.318779	0.326519	0.318584	0.318715	0.322992	0.318777
Monthly horizon						
QLIKE	0.370091	0.436770	0.370073	0.365128	0.413238	0.359545
MAE	0.464938	0.459841	0.464795	0.467163	0.458579	0.473115
MSE	0.397698	0.419242	0.397519	0.397602	0.409113	0.397699

Table 4. HARQ model's (Equation (18)) out-of-sample forecasting results using different estimation techniques.

	OLS	LAD	WLS _{RQ}	RLM $p = 2.1$	RLM $p = 1.3$	ELF $K_{exp} = 0.1$
Daily horizon						
QLIKE	0.308979	0.323848	0.308999	0.307671	0.319208	0.306458
MAE	0.543546	0.542170	0.5434	0.543920	0.541912	0.544287
MSE	0.478849	0.481396	0.478685	0.479025	0.479668	0.479030
Weekly horizon						
QLIKE	0.155273	0.165662	0.155253	0.154416	0.162043	0.154245
MAE	0.366248	0.364350	0.366114	0.366870	0.364036	0.367386
MSE	0.231442	0.235199	0.231280	0.231505	0.233154	0.231651
Biweekly horizon						
QLIKE	0.218227	0.239794	0.218224	0.216364	0.233033	0.215052
MAE	0.543546	0.542170	0.5434	0.543920	0.541912	0.544287
MSE	0.301743	0.309658	0.301538	0.301734	0.305979	0.302221
Monthly horizon						
QLIKE	0.366199	0.431019	0.366162	0.361283	0.409767	0.356297
MAE	0.458733	0.453891	0.458593	0.460855	0.452515	0.466573
MSE	0.389303	0.409066	0.389112	0.389269	0.400158	0.390595

The evaluation of the HARQ-X model was conducted under the same conditions, with one exception—the introduction of an additional coefficient in the Entropy Loss

Function ($K_{exp} = 0.01$). The results in Table 5 exhibit the superiority of ELF in minimizing the QLIKE metric. In contrast, the RLM ($p = 1.3$) showed the lowest MAE for daily and monthly horizons, while LAD exhibited the lowest MAE for weekly and biweekly horizons. This experiment presented the lowest Entropy Loss Function results in MAE compared with other models, especially in longer forecasting horizons. Conversely, the Robust Linear Model with a custom Minkowski loss function ($p = 2.1$) yielded balanced trade-off results on both metrics. The WLS approach with RQ weight demonstrated the lowest MSE values with this model as well. Conversely, the Robust Linear Model with a custom Minkowski loss function ($p = 2.1$) yielded balanced trade-off results on all the metrics.

Table 5. HARQ-X model' (Equation (19)) out-of-sample forecasting results using different estimation techniques.

	OLS	LAD	WLS _{RQ}	RLM $p = 2.1$	RLM $p = 1.3$	ELF $K_{exp} = 0.1$	ELF $K_{exp} = 0.01$
Daily horizon							
QLIKE	0.278937	0.290429	0.278922	0.277805	0.287298	0.276995	0.268285
MAE	0.522870	0.523373	0.522898	0.523026	0.522701	0.523440	0.526647
MSE	0.446430	0.448930	0.446519	0.446398	0.447735	0.446812	0.449631
Weekly horizon							
QLIKE	0.132194	0.138245	0.132158	0.131710	0.136197	0.131713	0.129863
MAE	0.345263	0.339749	0.345213	0.346180	0.340703	0.346390	0.350214
MSE	0.204014	0.203218	0.203967	0.204422	0.202857	0.204676	0.206861
Biweekly horizon							
QLIKE	0.190002	0.206361	0.189978	0.188780	0.200254	0.188493	0.184266
MAE	0.384005	0.376994	0.383951	0.385496	0.377558	0.386880	0.396262
MSE	0.265001	0.266835	0.264918	0.265749	0.264141	0.267138	0.274253
Monthly horizon							
QLIKE	0.337472	0.381946	0.337457	0.333633	0.367668	0.331365	0.319398
MAE	0.427724	0.417575	0.427714	0.430771	0.417262	0.437475	0.469005
MSE	0.353712	0.362413	0.353673	0.354721	0.356853	0.360805	0.389078

Overall, the results in Tables 3–5 suggest that introducing more lag to HAR-type models improves their accuracy. Due to the differences in each model optimization problem for these techniques, some of them demonstrated superior forecasting performance against others. For example, the WLS technique generally outperformed other models across all horizons, as measured using the MSE criterion. Although LAD primarily minimized the MAE metric in its optimization problem, it was outperformed by RLM with a custom Minkowski loss after tuning the p -value, demonstrating the flexibility of the latter approach. The Entropy Loss Function showed the best performance in the QLIKE metric compared to other models and techniques used in this work, reflecting its strong alignment with the underlying distributional characteristics of variance forecasts.

Figures 7 and 8 illustrate the QLIKE error metric for the HARQ-X model estimated with OLS over the entire period analyzed. As the visualization in Figures 7 and 8 demonstrate, the model's forecast accuracy significantly deteriorates during periods of high market stress. These periods of poor performance, indicated by spikes in the QLIKE metric, align closely with major market events that have been labeled on the chart for comparison. One of the candidate examples is the sharp increase in forecast error during the COVID-19 shock in early 2020, but other marked events also show a clear impact on the model's predictive

power. Although a detailed causal analysis of each individual error spike is beyond the scope of this study, this visualizes the model's primary failure points, demonstrating its sensitivity to the types of structural breaks represented by the labeled events.

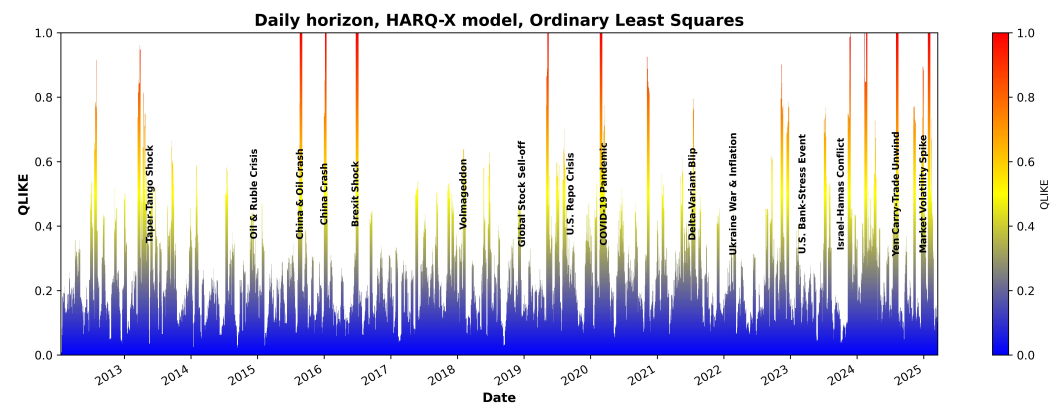


Figure 7. QLIKE estimates using the HARQX model with ordinary least squares for the daily horizon.

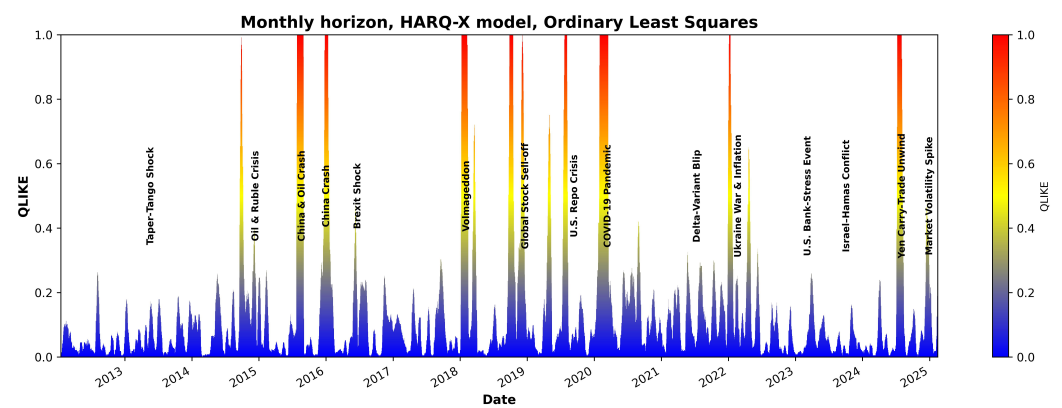


Figure 8. QLIKE estimates using the HARQX model with ordinary least squares for the monthly horizon.

Despite producing better overall average results, the effect of the entropic loss function (as shown in Figures 9 and 10) suggests that the estimation model infrastructure and constraints require further investigation.

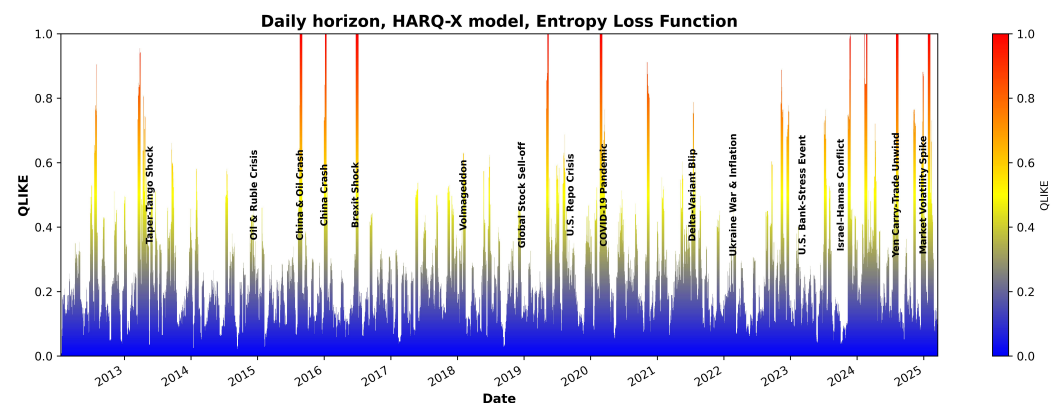


Figure 9. QLIKE estimates using the HARQX model with entropy loss for the daily horizon.

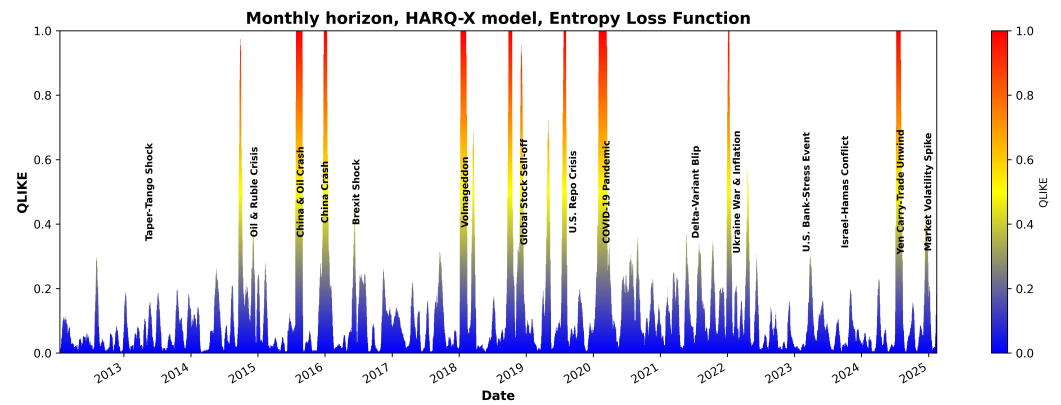
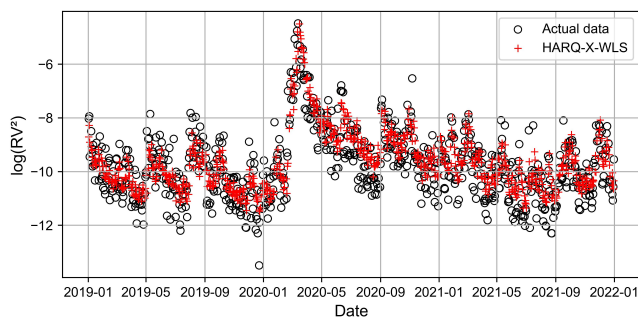
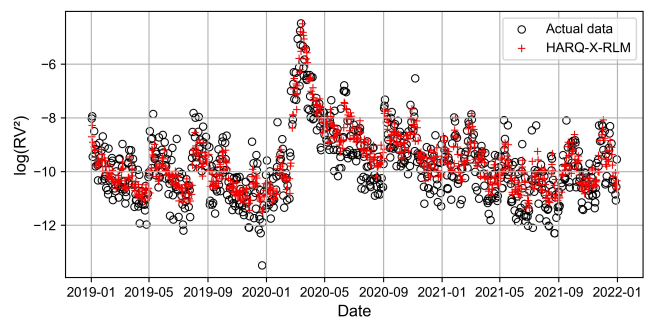


Figure 10. QLIKE estimates using the HARQX model with entropy loss for the monthly horizon.

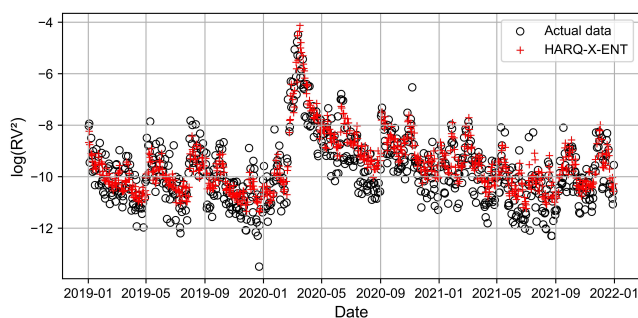
Figures 11–14 present the time series plots of the HARQ-X model with daily, weekly, biweekly, and monthly horizons, respectively. These figures show only the optimal results generated by three estimation techniques: WLS (Figures 11a–14a), ELF (Figures 11b–14b), and RLM ($p = 1.3$, Figures 11c–14c). The specially selected timeline, from January 2019 to January 2022, better represents the models' performance and actual values of the logarithmic realized variance. Additionally, during that period, stock markets crashed due to COVID-19. This period of increased volatility allowed us to observe volatility spikes and forecasting results.



(a)

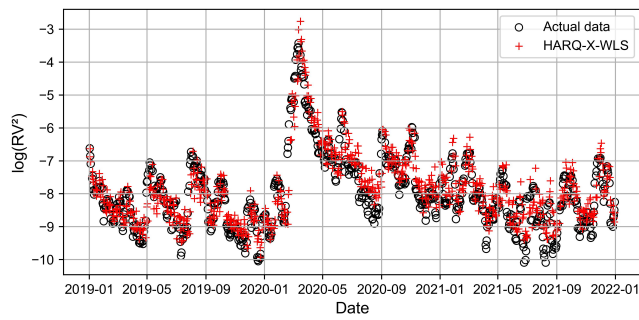


(b)

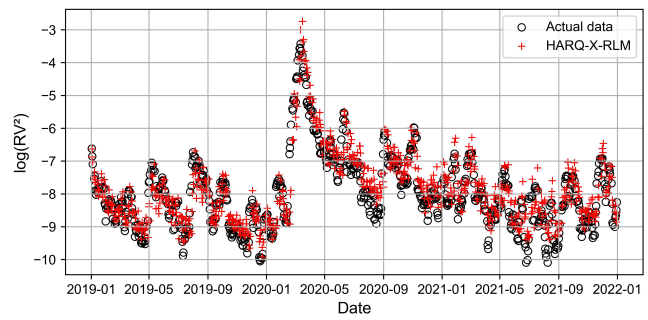


(c)

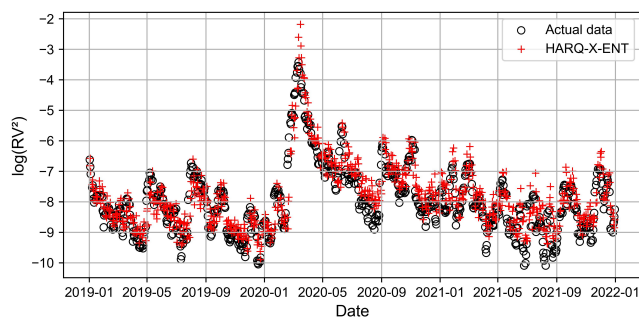
Figure 11. Best-performing HARQ-X models for the daily horizon, where (a) represents the WLS approach, (b) represents the RLM ($p = 1.3$) approach, and (c) represents the ELF approach.



(a)

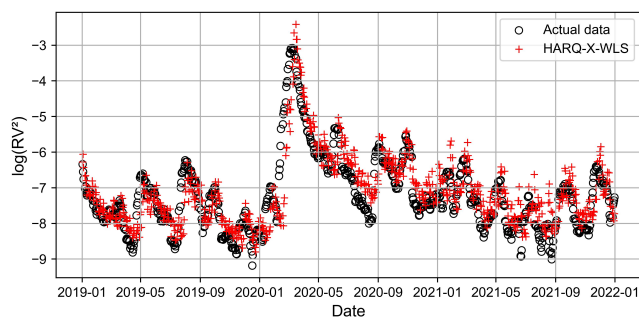


(b)

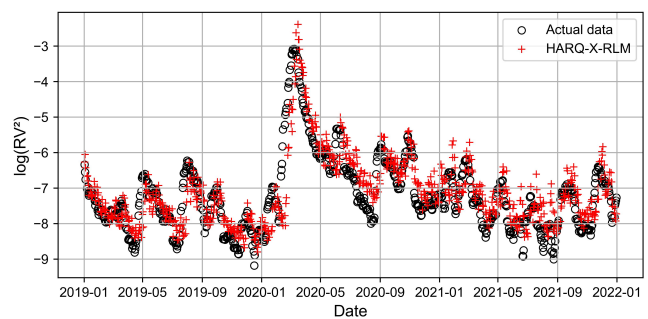


(c)

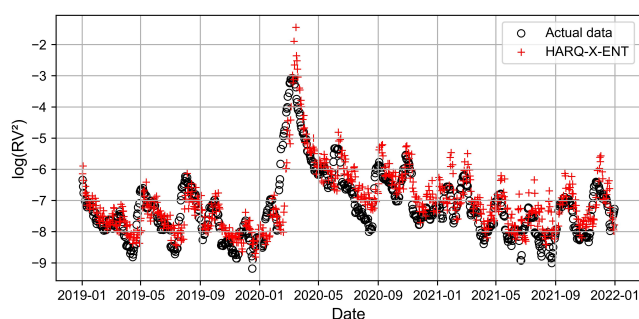
Figure 12. Best-performing HARQ-X models for the weekly horizon, where (a) represents the WLS approach, (b) represents the RLM ($p = 1.3$) approach, and (c) represents the ELF approach.



(a)



(b)



(c)

Figure 13. Best-performing HARQ-X models for the biweekly horizon, where (a) represents the WLS approach, (b) represents the RLM ($p = 1.3$) approach, and (c) represents the ELF approach.

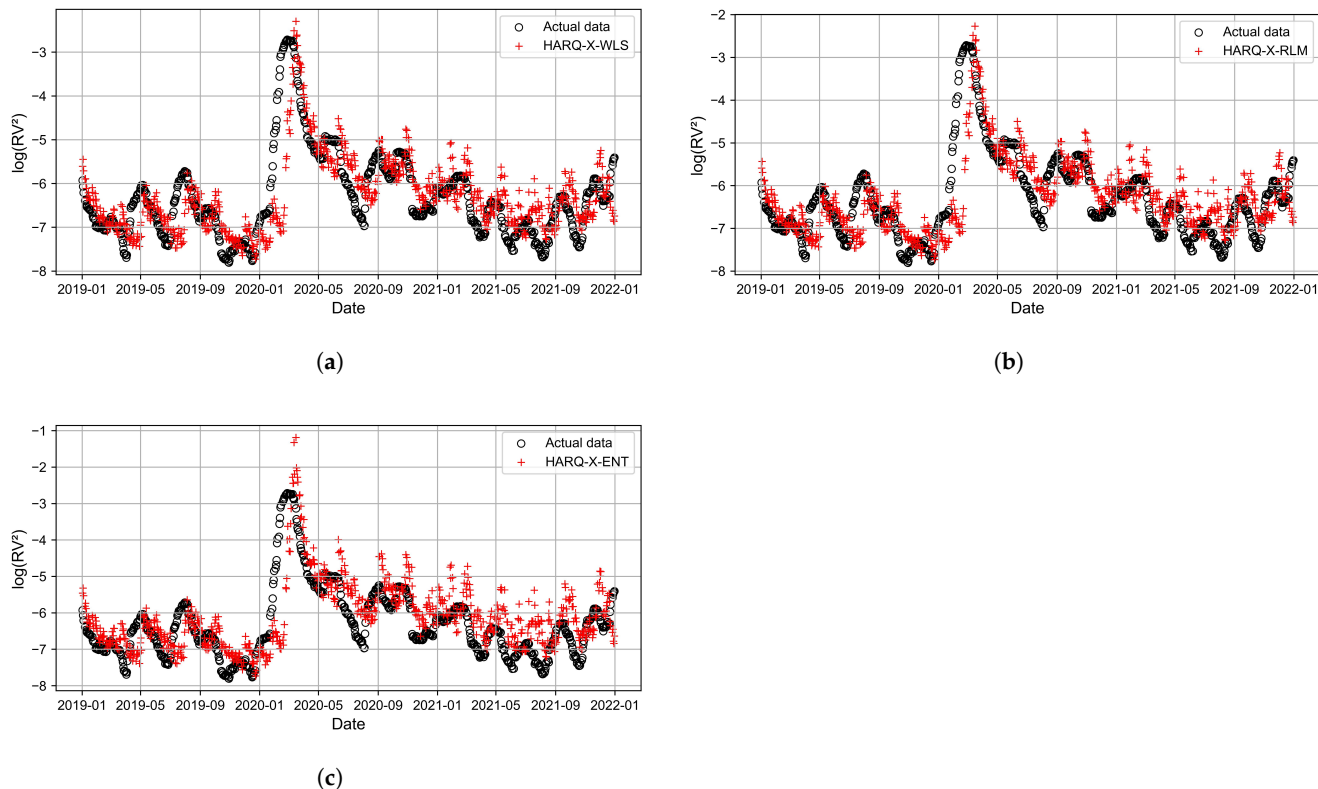


Figure 14. Best-performing HARQ-X models for the monthly horizon, where (a) represents the WLS approach, (b) represents the ELF approach, and (c) represents the RLM ($p = 1.3$) approach.

4. Conclusions

In conclusion, this study comprehensively explored HAR-type models, including the Heterogeneous Autoregressive for Realized Variance (HAR-RV), the Heterogeneous Autoregressive Model with Realized Quarticity (HARQ), and the Heterogeneous Autoregressive Model with Realized Quarticity extension incorporating an exogenous variable, VIX (HARQ-X). Generally, the observed forecasting models showed that the performance of the HAR-type models improved when extended with the realized quarticity and an exogenous variable (VIX). Moreover, this research presented five distinctive HAR-type model coefficient estimation techniques to suggest an alternative approach to traditional methods. The presented techniques were Ordinary Least Squares (OLS), Least Absolute Deviation (LAD), Weighted Least Squares (WLS), Robust Linear Model (RLM) with a custom Minkowski function, and Entropy Loss Function (ELF).

The forecasting accuracy evaluation took place over several forecasting periods, including daily (1 D), weekly (5 D), biweekly (10 D), and monthly (22 D). With the availability of high-frequency SPX stock market index data, the calculated realized variance acted as both a model feature and a target. Additionally, the coefficients estimated by the different techniques, including the Mean Absolute Error (MAE), the Mean Squared Error (MSE), and the Quasi-Likelihood (QLIKE) coefficients, were evaluated and compared to assess model performance. The empirical results of this study revealed that the ELF performed best when evaluating the coefficients of the HAR-RV and HARQ models, showing the lowest QLIKE coefficient and decent MAE and MSE results. Moreover, the RLM technique exhibited the least MAE while showing decent results for QLIKE and MSE. The HARQ-X model demonstrated the best performance in terms of QLIKE coefficients when estimated with the ELF. The RLM with a custom Minkowski loss function ($p = 1.3$) yielded the lowest MAE for daily and monthly horizons, while LAD exhibited the lowest MAE for weekly

and biweekly horizons. The Entropy Loss Function results in this experiment showed the lowest MAE compared to other models, especially in longer forecasting horizons. The WLS approach with the realized quarticity weight factor obtained the best MSE results; conversely, the Robust Linear Model with a custom Minkowski loss function ($p = 2.1$) yielded balanced trade-off results on all the metrics.

Author Contributions: Conceptualization, R.U. and V.V.; methodology, R.U., V.V., J.K. and R.P.; software, R.P. and J.K.; validation, J.K., V.V. and K.B.; formal analysis, E.K.; investigation, T.H.; data curation, K.B. and D.E.; writing—original draft preparation, R.P. and J.R.; writing—review and editing, R.P.; visualization, J.R. and J.D.; supervision, R.U. and V.V.; project administration, R.U. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ANN	Artificial Neural Network
ARCH	Autoregressive Conditional Heteroskedasticity
ARIMA	Autoregressive Integrated Moving Average
Bi-RNN	Bidirectional Recurrent Neural Network
CNN	Convolutional Neural Network
ECON	Economics
EGARCH	Exponential Generalized Autoregressive Conditional Heteroskedasticity
ELF	Entropy Loss Function
ESN	Echo State Neural Network
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
GBM	Geometric Brownian Motion
GBP	(GRU–BiRNN–PSO)
GJR-GARCH	Glosten–Jagannathan–Runkle Generalized Autoregressive Conditional Heteroskedasticity
GRU	Gated Recurrent Unit
HAR	Heterogeneous Autoregressive
HARQ	Heterogeneous Autoregressive model with Realized Quarticity
HAR-RV	Heterogeneous Autoregressive model of Realized Volatility
HAR-RV-CJ	Heterogeneous Autoregressive model of Realized Variance with Continuous and Jump components
HARQ-X	Heterogeneous Autoregressive Model with realized Quarticity and Exogenous Variables
IQ	Integrated Quarticity
LAD	Least Absolute Deviation
LASSO	Least Absolute Shrinkage and Selection Operator
LSTM	Long Short-Term Memory
LST-HAR	Logistic Smooth Transition HAR
MAE	Mean Absolute Error
ML	Machine Learning
MT-GARCH	Multi-Task Generalized Autoregressive Conditional Heteroskedasticity
MTL-GARCH	Multi-Task Learning Generalized

NARX	Nonlinear Autoregressive with Exogenous Input
NBEATSx	Neural Basis Expansion Analysis for Time Series with Exogenous Variables
NLP	Natural Language Processing
OLS	Ordinary Least Squares
PCA	Principal Component Analysis
PSO	Particle Swarm Optimization
QLIKE	Quasi-Likelihood Loss
RF	Random Forest
RLM	Robust Linear Model
RQ	Realized Quarticity
RV	Realized Variance
SPX	Standard & Poor's 500
SV	Stochastic Volatility
SVM	Support Vector Machine
TCN	Temporal Convolutional Network
V-HAR	Vector HAR
VIX	Chicago Board Options Exchange Volatility
VMD	Variational Mode Decomposition
WLS	Weighted Least Squares

References

- Cambridge Dictionary. Volatility. 2025. Available online: <https://dictionary.cambridge.org/dictionary/english/volatility> (accessed on 3 May 2025).
- Correia, M.; Kang, J.; Richardson, S. Asset volatility. *Rev. Account. Stud.* **2018**, *23*, 37–94. [\[CrossRef\]](#)
- Schwert, G.W. Why Does Stock Market Volatility Change Over Time? *J. Financ.* **1989**, *44*, 1115–1153. [\[CrossRef\]](#)
- Chou, R.Y.; Chou, H.; Liu, N. Range Volatility: A Review of Models and Empirical Studies. In *Handbook of Financial Econometrics and Statistics*; Lee, C.F., Lee, J.C., Eds.; Springer: Berlin/Heidelberg, Germany, 2015; pp. 2029–2050. [\[CrossRef\]](#)
- Black, F.; Scholes, M. The Pricing of Options and Corporate Liabilities. *J. Polit. Econ.* **1973**, *81*, 637–654. [\[CrossRef\]](#)
- Engle, R.F. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica* **1982**, *50*, 987–1007. [\[CrossRef\]](#)
- Bollerslev, T. Generalized autoregressive conditional heteroskedasticity. *J. Econom.* **1986**, *31*, 307–327. [\[CrossRef\]](#)
- Nelson, D.B. Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica* **1991**, *59*, 347–370. [\[CrossRef\]](#)
- Pilbeam, K.; Langeland, K.N. Forecasting exchange rate volatility: GARCH models versus implied volatility forecasts. *Int. Econ. Econ. Policy* **2015**, *12*, 127–142. [\[CrossRef\]](#)
- Glosten, L.R.; Jagannathan, R.; Runkle, D.E. On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *J. Financ.* **1993**, *48*, 1779–1801. [\[CrossRef\]](#)
- Andersen, T.G.; Bollerslev, T. Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts. *Int. Econ. Rev.* **1998**, *39*, 885–905. [\[CrossRef\]](#)
- Barndorff-Nielsen, O.E.; Shephard, N. Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models. *J. R. Stat. Soc. Ser. Stat. Methodol.* **2002**, *64*, 253–280. [\[CrossRef\]](#)
- Poon, S.H.; Granger, C.W.J. Forecasting Volatility in Financial Markets: A Review. *J. Econ. Lit.* **2003**, *41*, 478. [\[CrossRef\]](#)
- Gatheral, J.; Thibault, J.; Rosenbaum, M. Volatility is rough. *Quant. Financ.* **2018**, *18*, 933–949. [\[CrossRef\]](#)
- Ulugbode, M.A.; Shittu, O.I. Transition E-Garch Model for Modeling and Forecasting Volatility of Equity Stock Returns in Nigeria Stock Exchange. *Int. J. Res. Innov. Soc. Sci. (IJRISS)* **2024**, *VII*, 2120–2134. [\[CrossRef\]](#)
- Andersen, T.G.; Bollerslev, T.; Diebold, F.X.; Labys, P. Modeling and Forecasting Realized Volatility. *Econometrica* **2003**, *71*, 579–625. [\[CrossRef\]](#)
- Corsi, F. A Simple Approximate Long-Memory Model of Realized Volatility. *J. Financ. Econom.* **2009**, *7*, 174–196. [\[CrossRef\]](#)
- Andersen, T.G.; Bollerslev, T.; Diebold, F.X. Roughing It Up: Including Jump Components in the Measurement, Modeling, and Forecasting of Return Volatility. *Rev. Econ. Stat.* **2007**, *89*, 701–720. [\[CrossRef\]](#)
- Patton, A.J.; Sheppard, K. Good Volatility, Bad Volatility: Signed Jumps and the Persistence of Volatility. *Rev. Econ. Stat.* **2015**, *97*, 683–697. [\[CrossRef\]](#)
- Prokopczuk, M.; Symeonidis, L.; Wese Simen, C. Do Jumps Matter for Volatility Forecasting? Evidence from Energy Markets. *J. Futur. Mark.* **2016**, *36*, 758–792. [\[CrossRef\]](#)

21. Kambouroudis, D.S.; McMillan, D.G.; Tsakou, K. Forecasting Stock Return Volatility: A Comparison of GARCH, Implied Volatility, and Realized Volatility Models. *J. Futur. Mark.* **2016**, *36*, 1127–1163. [\[CrossRef\]](#)
22. Bollerslev, T.; Patton, A.J.; Quaedvlieg, R. Exploiting the errors: A simple approach for improved volatility forecasting. *J. Econom.* **2016**, *192*, 1–18. [\[CrossRef\]](#)
23. Qu, H.; Chen, W.; Niu, M.; Li, X. Forecasting realized volatility in electricity markets using logistic smooth transition heterogeneous autoregressive models. *Energy Econ.* **2016**, *54*, 68–76. [\[CrossRef\]](#)
24. Cubadda, G.; Guardabascio, B.; Hecq, A. A vector heterogeneous autoregressive index model for realized volatility measures. *Int. J. Forecast.* **2017**, *33*, 337–344. [\[CrossRef\]](#)
25. Clements, A.; Preve, D.P.A. A Practical Guide to harnessing the HAR volatility model. *J. Bank. Financ.* **2021**, *133*, 106285. [\[CrossRef\]](#)
26. Li, Z.C.; Xie, C.; Wang, G.J.; Zhu, Y.; Zeng, Z.J.; Gong, J. Forecasting global stock market volatilities: A shrinkage heterogeneous autoregressive (HAR) model with a large cross-market predictor set. *Int. Rev. Econ. Financ.* **2024**, *93*, 673–711. [\[CrossRef\]](#)
27. Michael, N.; Mihai, C.; Howison, S. Options-driven volatility forecasting. *Quant. Financ.* **2025**, *25*, 443–470. [\[CrossRef\]](#)
28. Luong, C.; Dokuchaev, N. Forecasting of Realised Volatility with the Random Forests Algorithm. *J. Risk Financ. Manag.* **2018**, *11*, 61. [\[CrossRef\]](#)
29. Bucci, A. Realized Volatility Forecasting with Neural Networks. *J. Financ. Econom.* **2020**, *18*, 502–531. [\[CrossRef\]](#)
30. Zhang, C.X.; Li, J.; Huang, X.F.; Zhang, J.S.; Huang, H.C. Forecasting stock volatility and value-at-risk based on temporal convolutional networks. *Expert Syst. Appl.* **2022**, *207*, 117951. [\[CrossRef\]](#)
31. Ge, W.; Lalbakhsh, P.; Isai, L.; Lenskiy, A.; Suominen, H. Neural Network-Based Financial Volatility Forecasting: A Systematic Review. *ACM Comput. Surv.* **2022**, *55*, 14:1–14:30. [\[CrossRef\]](#)
32. Zahid, S.; Saleem, H.M.N. Stock Volatility Prediction Using Machine Learning during COVID-19. *Stat. Comput. Interdiscip. Res.* **2023**, *5*, 99–119. [\[CrossRef\]](#)
33. Christensen, K.; Siggaard, M.; Veliyev, B. A Machine Learning Approach to Volatility Forecasting*. *J. Financ. Econom.* **2023**, *21*, 1680–1727. [\[CrossRef\]](#)
34. Souto, H.G.; Moradi, A. Introducing NBEATSx to realized volatility forecasting. *Expert Syst. Appl.* **2024**, *242*, 122802. [\[CrossRef\]](#)
35. Zhang, C.; Zhang, Y.; Cucuringu, M.; Qian, Z. Volatility Forecasting with Machine Learning and Intraday Commonality. *J. Financ. Econom.* **2024**, *22*, 492–530. [\[CrossRef\]](#)
36. Mensah, N.; Agbeduamenu, C.O.; Obodai, T.N.; Adukpo, T.K. Leveraging Machine Learning Techniques to Forecast Market Volatility in the U.S. *EPRA Int. J. Econ. Bus. Manag. Stud. (EBMS)* **2025**, *12*, 76–85.
37. Lolic, M. Tree-Based Methods of Volatility Prediction for the S&P 500 Index. *Computation* **2025**, *13*, 84. [\[CrossRef\]](#)
38. Beg, K. Comparative Analysis of Machine Learning Models for Sectoral Volatility Prediction in Financial Markets. *J. Inf. Syst. Eng. Manag.* **2025**, *10*, 837–846. [\[CrossRef\]](#)
39. Mansilla-Lopez, J.; Mauricio, D.; Narváez, A. Factors, Forecasts, and Simulations of Volatility in the Stock Market Using Machine Learning. *J. Risk Financ. Manag.* **2025**, *18*, 227. [\[CrossRef\]](#)
40. Monfared, S.A.; Enke, D. Volatility Forecasting Using a Hybrid GJR-GARCH Neural Network Model. *Procedia Comput. Sci.* **2014**, *36*, 246–253. [\[CrossRef\]](#)
41. Kristjanpoller, W.; Fadic, A.; Minutolo, M.C. Volatility forecast using hybrid Neural Network models. *Expert Syst. Appl.* **2014**, *41*, 2437–2442. [\[CrossRef\]](#)
42. Yang, R.; Yu, L.; Zhao, Y.; Yu, H.; Xu, G.; Wu, Y.; Liu, Z. Big data analytics for financial Market volatility forecast based on support vector machine. *Int. J. Inf. Manag.* **2020**, *50*, 452–462. [\[CrossRef\]](#)
43. Trierweiler Ribeiro, G.; Alves Portela Santos, A.; Cocco Mariani, V.; dos Santos Coelho, L. Novel hybrid model based on echo state neural network applied to the prediction of stock price return volatility. *Expert Syst. Appl.* **2021**, *184*, 115490. [\[CrossRef\]](#)
44. Liu, J.; Xu, Z.; Yang, Y.; Zhou, K.; Kumar, M. Dynamic Prediction Model of Financial Asset Volatility Based on Bidirectional Recurrent Neural Networks. *J. Organ. End User Comput. (JOEUC)* **2024**, *36*, 1–23. [\[CrossRef\]](#)
45. Mishra, A.K.; Renganathan, J.; Gupta, A. Volatility forecasting and assessing risk of financial markets using multi-transformer neural network based architecture. *Eng. Appl. Artif. Intell.* **2024**, *133*, 108–223. [\[CrossRef\]](#)
46. Brini, A.; Toscano, G. SpotV2Net: Multivariate intraday spot volatility forecasting via vol-of-vol-informed graph attention networks. *Int. J. Forecast.* **2025**, *41*, 1093–1111. [\[CrossRef\]](#)
47. Hu, N.; Yin, X.; Yao, Y. A novel HAR-type realized volatility forecasting model using graph neural network. *Int. Rev. Financ. Anal.* **2025**, *98*, 103881. [\[CrossRef\]](#)
48. Li, H.; Huang, X.; Luo, F.; Zhou, D.; Cao, A.; Guo, L. Revolutionizing agricultural stock volatility forecasting: A comparative study of machine learning and HAR-RV models. *J. Appl. Econ.* **2025**, *28*, 2454081. [\[CrossRef\]](#)
49. Kumar, S.; Rao, A.; Dhochak, M. Hybrid ML models for volatility prediction in financial risk management. *Int. Rev. Econ. Financ.* **2025**, *98*, 103915. [\[CrossRef\]](#)

50. Wang, S.; Bai, Y.; Ji, T.; Fu, K.; Wang, L.; Lu, C.T. Stock Movement and Volatility Prediction from Tweets, Macroeconomic Factors and Historical Prices. In Proceedings of the 2023 IEEE International Conference on Big Data (BigData), Sorrento, Italy, 15–18 December 2023; pp. 1863–1872. [\[CrossRef\]](#)
51. Shi, M.X.; Chen, C.C.; Huang, H.H.; Chen, H.H. Enhancing Volatility Forecasting in Financial Markets: A General Numeral Attachment Dataset for Understanding Earnings Calls. In *Proceedings of the 13th International Joint Conference on Natural Language Processing and the 3rd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (Volume 2: Short Papers)*; Park, J.C., Arase, Y., Hu, B., Lu, W., Wijaya, D., Purwarianti, A., Krisnadhi, A.A., Eds.; Association for Computational Linguistics: Stroudsburg, PA, USA, 2023; pp. 37–42. [\[CrossRef\]](#)
52. Li, X.; Xu, Y.; Yang, L.; Zhang, Y.; Dong, R. NLP-Based Analysis of Annual Reports: Asset Volatility Prediction and Portfolio Strategy Application. In Proceedings of the 32nd Irish Conference on Artificial Intelligence and Cognitive Science, CEUR Workshop Proceedings, Dublin, Ireland, 9–10 December 2024; Volume 3910, pp. 228–240.
53. Baker, M.; Wurgler, J. Investor Sentiment in the Stock Market. *J. Econ. Perspect.* **2007**, *21*, 129–151. [\[CrossRef\]](#)
54. Chordia, T.; Roll, R.; Subrahmanyam, A. Evidence on the speed of convergence to market efficiency. *J. Financ. Econ.* **2005**, *76*, 271–292. [\[CrossRef\]](#)
55. Liu, L.Y.; Patton, A.J.; Sheppard, K. Does anything beat 5-min RV? A comparison of realized measures across multiple asset classes. *J. Econom.* **2015**, *187*, 293–311. [\[CrossRef\]](#)
56. Andersen, T.G.; Tim, B.; Francis X, D.; Labys, P. The Distribution of Realized Exchange Rate Volatility. *J. Am. Stat. Assoc.* **2001**, *96*, 42–55. [\[CrossRef\]](#)
57. Karatzas, I.; Shreve, S.E. *Brownian Motion and Stochastic Calculus*, 2nd ed.; Number 113 in Graduate texts in mathematics; Springer: New York, NY, USA, 1996.
58. Szczypińska, A.; Piotrowski, E.W.; Makowski, M. Deterministic risk modelling: Newtonian dynamics in capital flow. *Phys. A Stat. Mech. Its Appl.* **2025**, *665*, 130499. [\[CrossRef\]](#)
59. Mandelbrot, B. The Variation of Certain Speculative Prices. *J. Bus.* **1963**, *36*, 394–419. [\[CrossRef\]](#)
60. Comte, F.; Renault, E. Long memory in continuous-time stochastic volatility models. *Math. Financ.* **1998**, *8*, 291–323. [\[CrossRef\]](#)
61. Zwanzig, R. Memory Effects in Irreversible Thermodynamics. *Phys. Rev.* **1961**, *124*, 983–992. [\[CrossRef\]](#)
62. Müller, U.A.; Dacorogna, M.M.; Davé, R.D.; Olsen, R.B.; Pictet, O.V.; von Weizsäcker, J.E. Volatilities of different time resolutions—Analyzing the dynamics of market components. *J. Empir. Financ.* **1997**, *4*, 213–239. [\[CrossRef\]](#)
63. Wooldridge, J.M. *Introductory Econometrics: A Modern Approach*, 6th ed.; Cengage Learning: Boston, MA, USA 2016.
64. Koenker, R.; Bassett, G. Regression Quantiles. *Econometrica* **1978**, *46*, 33–50. [\[CrossRef\]](#)
65. Huber, P.J. *Robust Statistics*; Wiley: New York, NY, USA, 1981.
66. Vera, J.F.; Heiser, W.J.; Murillo, A. Global Optimization in Any Minkowski Metric: A Permutation-Translation Simulated Annealing Algorithm for Multidimensional Scaling. *J. Classif.* **2007**, *24*, 277–301. [\[CrossRef\]](#)
67. De Nolasco Santos, F.; D’Antuono, P.; Noppe, N.; Weijtjens, W.; Devriendt, C. Minkowski logarithmic error: A physics-informed neural network approach for wind turbine lifetime assessment. In Proceedings of the ESANN 2022 Proceedings, Bruges, Belgium, 5–7 October 2022; pp. 357–362. [\[CrossRef\]](#)
68. Urniezius, R.; Survyla, A.; Paulauskas, D.; Bumelis, V.A.; Galvanauskas, V. Generic estimator of biomass concentration for *Escherichia coli* and *Saccharomyces cerevisiae* fed-batch cultures based on cumulative oxygen consumption rate. *Microb. Cell Factories* **2019**, *18*, 190. [\[CrossRef\]](#)
69. Urniezius, R.; Survyla, A. Identification of Functional Bioprocess Model for Recombinant *E. coli* Cultivation Process. *Entropy* **2019**, *21*, 1221. [\[CrossRef\]](#)
70. Hyndman, R.J.; Koehler, A.B. Another look at measures of forecast accuracy. *Int. J. Forecast.* **2006**, *22*, 679–688. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.