Xtreamly: AI-Driven Volatility Models and Market Classification of Ethereum Price

Xtreamly provides AI-based models API for volatility prediction and market classification in the cryptocurrency market. This whitepaper explores our proprietary approaches for forecasting volatility: the implementation of a robust API for real-time predictions, and our classification policy for market states. These advancements enable sophisticated DeFi strategies, ensuring informed decision-making and risk management for users.

Introduction

Cryptocurrency markets are characterized by extreme volatility, driven by rapid price fluctuations and time-varying market dynamics. For traders and investors, accurately forecasting volatility and understanding market conditions are critical to optimizing strategies and mitigating risks. Xtreamly addresses these needs with advanced AI models for volatility prediction, alongside a framework for classifying market states of volatility categories.

This document outlines the technical methodologies, implementation framework, and practical applications of Xtreamly's volatility prediction and market volatility classification.

Literature Review for Volatility Prediction

This section provides an overview of various methodologies utilized in the prediction of price volatility across financial markets, with a specific focus on Ethereum (ETH) price. The methodologies range from traditional statistical approaches to advanced machine learning techniques.

General Approaches

Autoregressive Integrated Moving Average (ARIMA)

ARIMA models are widely used for time series forecasting. They combine autoregressive (AR) and moving average (MA) components with differencing to achieve stationarity. The effectiveness of ARIMA in predicting Ethereum volatility has been demonstrated through rigorous testing, including the Augmented Dickey-Fuller (ADF) test for stationarity and model selection criteria such as the Bayesian Information Criterion (BIC). Studies have shown that ARIMA models can yield robust forecasting accuracy with low Mean Absolute Percentage Error (MAPE) values, indicating their utility in short-term predictions. (Hyndman, Rob J; Athanasopoulos, George 2015)

Heterogeneous Autoregressive (HAR) Models

HAR models extend traditional autoregressive frameworks by incorporating different time horizons into the analysis. By utilizing high-frequency data, HAR models can capture short-term fluctuations and long-term trends in volatility. Recent studies have demonstrated that HAR models outperform conventional GARCH models when accounting for structural breaks and external factors such as market indices. (Adam Clements, Daniel P.A. Preve 2021)

Autoregressive Conditional Heteroskedasticity (ARCH) Models

Autoregressive Conditional Heteroskedasticity (ARCH) models were a pivotal development in financial econometrics for modeling time-varying volatility in asset returns. This literature review explores the foundational ARCH model and its various extensions, including GARCH, EGARCH, FIGARCH, and APARCH, highlighting their applications and effectiveness in capturing volatility dynamics. (Bollerslev, Tim; Russell, Jeffrey; Watson, Mark, 2010)

Research has shown that different ARCH-type models perform variably depending on the characteristics of the data being analyzed. For instance, studies indicate that FIGARCH and IGARCH models often outperform their symmetric counterparts in capturing long-memory effects in financial returns (Stavros Antonios Degiannakis, Evdokia Xekalaki 2005, Gabriel Rodriguez 2017). Moreover, empirical comparisons suggest that EGARCH and TGARCH models provide superior fits for datasets exhibiting leverage effects, while APARCH models excel in scenarios requiring flexible response mechanisms to shocks (Bahadtin Rüzgar , İsmet Kale 2007 Krzysztof Szymoniak-Książek, 2020).

ARCH models and their extensions—GARCH, EGARCH, FIGARCH, and APARCH—offer robust frameworks for analyzing and forecasting volatility in financial markets. The choice among these models depends on specific data characteristics and the underlying economic phenomena being studied. While ongoing research continues in ARCH models, their parametric assumptions often limit them from capturing the full dynamics of modern financial time series.

The original ARCH model, proposed by Engle (1982), describes the variance of a time series as a function of past error terms. This model is particularly adept at capturing the phenomenon of volatility clustering, where high-volatility periods are followed by high volatility and low-volatility periods by low volatility (Will Kenton 2024, Brooks Chris 2014). The model assumes that the conditional variance of the error term is dependent on the squares of previous error terms, making it suitable for financial data that exhibit non-constant variance over time.

Generalized Autoregressive Conditional Heteroskedasticity (GARCH)

The GARCH model, introduced by (Tim Bollerslev (1986), extends the ARCH framework by incorporating past variances along with past errors to predict future volatility. This allows GARCH models to capture longer-lasting volatility patterns compared to ARCH models, making them more effective for financial applications. GARCH models are widely used in risk management and derivative pricing due to their ability to provide more accurate forecasts of future volatility.

Exponential GARCH (EGARCH)

EGARCH, developed by (Daniel B. Nelson 1991), addresses some limitations of GARCH models by allowing for asymmetric effects of shocks on volatility. It captures the "leverage effect," where negative shocks tend to increase future volatility more than positive shocks of the same magnitude. This characteristic makes EGARCH particularly useful in financial markets where such asymmetries are prevalent.

Fractionally Integrated GARCH (FIGARCH)

The FIGARCH model, proposed by (Richard Baillie 1996), incorporates long memory processes into the GARCH framework. This model is beneficial for capturing persistent volatility patterns observed in financial time series data. FIGARCH allows for fractional differencing, which can better represent the behavior of financial returns over time compared to standard GARCH models.

Asymmetric Power ARCH (APARCH)

The APARCH model, introduced by (Ding 1993), combines features from both GARCH and power transformations to account for asymmetries and varying degrees of responsiveness in volatility to market shocks. The flexibility of APARCH allows it to adapt to different types of financial data, making it a versatile tool for modeling conditional heteroskedasticity.

Support Vector Regression (SVR)

SVR is a popular machine learning technique that has shown promise in predicting cryptocurrency volatility. It works by finding a hyperplane that best fits the data while allowing for some margin of error. Studies indicate that SVR performs well for daily forecasts, often appearing in the set of best-performing models alongside simpler methods like HAR. (Cortes, Corinna; Vapnik, Vladimir, 1995)

Long Short-Term Memory Networks (LSTM)

LSTM networks are a type of recurrent neural network specifically designed to learn from sequences of data, making them suitable for time series forecasting. They have been applied to predict Ethereum price movements by capturing long-term dependencies in historical price data. Research has shown that LSTMs can achieve competitive accuracy compared to traditional statistical methods, especially when combined with other features like sentiment analysis. (Saiful Islam, Md.; Hossain, Emam, 2020)

Hybrid Models

Hybrid approaches that combine traditional econometric models with machine learning techniques are gaining traction in the field of volatility forecasting. For instance, integrating GARCH models with LSTM networks allows for capturing both linear and non-linear relationships in price movements, enhancing predictive performance. (Thanasis Zoumpekas, Elias Houstis, Manolis Vavalis, 2020)

Ensemble Methods

Ensemble methods involve combining multiple predictive models to improve overall accuracy and robustness. Techniques such as Random Forests and Gradient Boosting Machines have been explored for their ability to aggregate predictions from various base learners, thereby reducing overfitting and improving generalization on unseen data. (Bâra Adela, Simona-Vasilica Oprea 2024)

Stochastic Volatility Models (SV Models)

Stochastic Volatility Models (SV Models) represent a significant advancement in the modeling of financial volatility, particularly when compared to traditional GARCH models. These models assume that volatility itself follows a stochastic process, allowing for greater flexibility in capturing the complexities of market behavior. SV models are predominantly used in derivative pricing due to their ability to model the dynamics of underlying asset prices and their volatilities. They provide a more accurate framework for pricing options compared to static models. The flexibility of SV models allows them to accommodate different market conditions and investor sentiments.

A common method involves using option-based risk-neutral density derived volatility estimators. This approach utilizes the Breeden-Litzenberger formula, which extracts implied volatilities from observed option prices, particularly focusing on volatility smiles and their interpolation techniques. The Breeden-Litzenberger formula (Douglas T. Breeden and Robert H. Litzenberger 1978) allows for the extraction of the implied risk-neutral probability density function from market prices of European call options. The second derivative effectively captures the market's expectations of future price movements, providing insights into future volatility expectations based on current market data.

Implied volatility also plays a crucial role in this context as it reflects market participants' expectations about future volatility derived from current option prices. Implied volatility is not constant; it can vary significantly across different strike prices and expiration dates, leading to the formation of volatility smiles or skews. These patterns indicate that traders expect greater uncertainty (higher implied volatility) for options that are deep in-the-money or out-of-the-money compared to at-the-money options (Jim Gatheral 2006).

CEX platforms such as Deribit, Binance, Gemini, and Coinbase provide existing option prices that could be utilized for estimating volatility through SV models. However, the current market for ETH options may not support short-term forecasting effectively due to a lack of available options data. However, for Xtreamly's goal of predicting short and middle term volatility, this approach is not applicable.

Key Performance Metrics (KPIs)

The effectiveness of methodologies for prediction and analysis is often evaluated using various performance metrics. Below is an overview of commonly used metrics.

Mean Absolute Percentage Error (MAPE)

Mean Absolute Percentage Error (MAPE) measures the average magnitude of errors in a set of predictions, expressed as a percentage of the actual values. It is a widely used metric for assessing forecast accuracy, especially when values need to be understood as relative errors.

Symmetric Mean Absolute Percentage Error (SMAPE)

Symmetric Mean Absolute Percentage Error (SMAPE) is a variation of MAPE that provides a more balanced measure of accuracy by considering both over-predictions and under-predictions symmetrically.

Mean Squared Error (MSE)

Mean Squared Error (MSE) measures the average of the squares of the errors, which is the average squared difference between estimated values and actual values. This metric is useful for emphasizing larger errors due to its squaring property, making it sensitive to outliers.

Mean Absolute Error (MAE)

Mean Absolute Error (MAE) measures the average of absolute differences between predicted and actual values. It provides a straightforward interpretation of prediction errors without over-penalizing large deviations.

Pearson Correlation Coefficient (Corr)

The Pearson correlation coefficient measures the linear correlation between two variables, providing insight into how closely related predicted values are to actual values. A value close to 1 indicates strong positive correlation, while a value close to -1 indicates strong negative correlation.

Explained Variance (EV)

Explained Variance (EV) measures the proportion of the total variance in the target variable that is explained by the model. It quantifies how well the model accounts for the variability in the data. Values closer to 1 indicate that the model explains a higher proportion of the variance, whereas values closer to 0 suggest that the model explains little to none of the variability. Negative values may occur when the model performs worse than a simple baseline, such as predicting the mean.

R-squared (R^2)

R-squared (R^2) measures the proportion of variance in the dependent variable that can be explained by the independent variables in a regression model. It is often used to evaluate the overall performance of a model in terms of its ability to explain variability.

Xtreamly currently uses R^2 as the main evaluation metric for our methodology. This metric provides a comprehensive view of model performance, capturing how well the model explains the variance in the target variable relative to the independent variables. In certain contexts, Xtreamly also examines other metrics which contribute to the performance presented by R^2.

Approach to Ethereum Price Volatility Prediction

Xtreamly has identified a comprehensive list of research papers that cover the prediction of volatility in Ethereum (ETH) prices:

Ethereum Price Prediction Using Machine Learning Techniques: A Comparative Study

This research focuses on predicting Ethereum prices using various machine learning techniques over a dataset spanning 2000 days. While primarily focused on price prediction, it also touches on volatility aspects related to price fluctuations, demonstrating that simpler models can perform comparably to more complex ones (Shanta Rangaswamy 2022).

Time Series Forecasting of Ethereum Prices: An ARIMA Model Approach

This study utilizes time series analysis, specifically the ARIMA model, to forecast Ethereum prices. It includes an evaluation of model performance using metrics like MAPE and Theil's Inequality Coefficient, highlighting the potential and limitations of ARIMA in short-term forecasting. The ARIMA(1,1,0) model achieved a MAPE of 1.43%, indicating robust forecasting accuracy (Poonam Dwivedi, Lotica Surana 2024).

Bitcoin and Ethereum Volatility Forecasting with GARCH Models

This paper systematically investigates the statistical properties of Bitcoin and Ethereum returns using various GARCH models. It identifies the two-component GJR model as the best fit for predicting volatility and discusses the implications of these findings for forecasting future price movements. The results suggest that GARCH models effectively capture volatility clustering in cryptocurrency markets (Zhuanxin Ding, Karamvir Gosal, Greg McMurran 2024).

Forecasting Cryptocurrencies Volatility Using Statistical and Machine Learning Methods

This comprehensive study compares multiple statistical and machine learning methods for predicting daily and weekly volatility across several cryptocurrencies, including Ethereum. It evaluates models such as GARCH, SVR, and LSTM, emphasizing that no single model is superior across all scenarios. The findings indicate that linear SVR consistently ranks among the best for daily forecasts (Grzegorz Dudek, Piotr Fiszeder, Paweł Kobus, Witold Orzeszko 2024).

Forecasting Ethereum's Volatility: An Expansive Approach Using HAR Models

This paper employs high-frequency data to construct Heterogeneous Autoregressive (HAR) models aimed at predicting Ethereum's volatility, considering structural breaks in the data. The HAR-RV model achieved a perfect fit in 1-day and 1-month forecasting, outperforming other models for out-of-sample predictions (Ruijie Chen 2024).

Transformer-Based Approach for Ethereum Price Prediction

This paper presents a novel approach to predicting Ethereum prices using transformer models, which can implicitly capture volatility through price prediction dynamics. The results indicate that this method enhances forecasting accuracy compared to traditional models (Shubham Singh, Mayur Bhat, 2024).

Xtreamly Volatility Prediction

Volatility measures the degree of variation in the returns of a financial asset over a period. At Xtreamly, it is expressed as the standard deviation of the asset's logarithmic returns. Higher volatility indicates greater risk due to the larger range of potential outcomes.

In financial analysis, rolling volatility is calculated using a moving window of logarithmic returns to capture recent variability and the same approach is being taken by Xtreamly. For a window size w, the rolling volatility σ at time t is, the realized volatility is given by:

$$\sigma_{rolling_t} = \sqrt{\frac{1}{w-1}\sum_{i=t-w+1}^{t}(r_i - \mu)^2}$$

Where:

- $\sigma_{rolling_t}$ represents rolling volatility at time t.
- w represents a window size of rolling timeline (ex. 1 minute or 60 minutes)
- r_i represents log-return at time t in respect to price at time $r_i = ln(\frac{price_i}{price_{i-w}})$
- μ represents a mean of returns over window time $(t, t - w)$

The realized volatilities serve as an input for prediction. To use them as the future realized volatility (the predicted explained variable), they need to be shifted backwards in time by the window size. To ensure no information from the future is used in the shifted statistics, Xtreamly uses price bar open prices to calculate logarithmic returns (where bar width is equal to window size).

Xtreamly's volatility prediction models operate on two distinct horizons (with shifted windows):

- **Short-term (1-minute horizon)**: For rapid responses to immediate market changes.
- **Medium-term (60-minute horizon)**: For strategic planning and risk adjustment over a longer timeframe.

Alternative Volatility Models Results

The data used for training and evaluation spans Q1-Q3 2024, with the test period focused on Q4 2024 to assess the realized volatility of ETH/USD prices. The test set (Q4) is used to evaluate the predictive power and performance of the various models.

Autoregressive Integrated Moving Average (ARIMA) Models

In this analysis, we utilize the ARIMA model to forecast the realized volatility of ETH/USD prices for the upcoming hour. The ARIMA model has been adjusted to accommodate different time horizons, meaning it uses various time windows to estimate the predicted volatility, such as 4-hour, 24-hour, 48-hour, and 72-hour observation windows. Each model aims to predict the realized volatility for the next hour based on the historical data. Models are trained using data from Q1-Q3 2024 and tested on Q4 2024 data to assess their forecasting accuracy.

Below, comparative KPIs for trained versions of ARIMA model:

Horizon	Model	MAPE	SMAPE	MSE	MAE	Corr.	EV	R^2
1min	ARIMA up to 5 minutes	over 1	0.378	0	0.0002	0.6915	0.4774	0.477
1min	ARIMA up to 15 minutes	over 1	0.3705	0	0.0002	0.7029	0.4933	0.4931
1min	ARIMA up to 60 minutes	over 1	0.3681	0	0.0002	0.7057	0.4972	0.4971
1min	ARIMA up to 240 minutes	over 1	0.3676	0	0.0002	0.7061	0.4977	0.4977
60min	ARIMA up to 4 hours	0.4541	0.3791	0	0.0013	0.6065	0.3658	0.3637
60min	ARIMA up to 24 hours	0.4516	0.3748	0	0.0013	0.6138	0.3748	0.3743
60min	ARIMA up to 48 hours	0.4491	0.3741	0	0.0013	0.615	0.3757	0.3753
60min	ARIMA up to 72 hours	0.4511	0.3755	0	0.0013	0.6145	0.3748	0.3744

The updated results indicate that ARIMA models exhibit relatively stable performance across different time horizons, with minor variations in predictive accuracy. The R^2 values show a slight improvement as the observation window increases, with the ARIMA 48-hour model achieving the highest R^2 for 60 minute horizon ARIMA 240-minute for 1 minute horizon. This suggests that while the overall predictive accuracy remains modest, extending the time horizon slightly improves the model's ability to capture volatility patterns. However, MAPE values remain above 1 for short-term predictions for 1 minute horizon, reflecting inherent challenges in forecasting short-term cryptocurrency volatility.

Heterogeneous Autoregressive (HAR) Models

The HAR model is designed to account for long-term memory in volatility dynamics and explicitly models heterogeneity across different time horizons. In this study, the model uses data from multiple time frames (e.g., 1 hour, 4 hours, 12 hours, 24 hours, 7 days, and 14 days) to predict future realized volatility.

The comparative KPIs for trained versions of HAR models are displayed in the following table:

Horizon	Model	MAPE	SMAPE	MSE	MAE	Corr.	EV	R^2
1min	HAR up to 15 minutes	over 1	0.3639	0	0.0001	0.7202	0.5179	0.5179
1min	HAR up to 60 minutes	over 1	0.3624	0	0.0001	0.7223	0.521	0.521
1min	HAR up to 240 minutes	over 1	0.362	0	0.0001	0.7227	0.5217	0.5216
60min	HAR up to 12 hours	0.4294	0.3587	0	0.0013	0.6566	0.4307	0.4278
60min	HAR up to 24 hours	0.4291	0.3589	0	0.0013	0.657	0.4312	0.4283
60min	HAR up to 7 days	0.4289	0.3587	0	0.0013	0.6574	0.4318	0.429
60min	HAR up to 14 days	0.4286	0.3586	0	0.0013	0.6575	0.4319	0.429

The HAR results demonstrate that HAR models perform consistently across different time horizons, with incremental improvements in predictive accuracy as the observation window increases. The HAR 14-day model achieved the highest R^2 of 0.429, suggesting that incorporating longer historical data enhances predictive power. However, similar to ARIMA, the overall R^2 values remain modest, reflecting the complexity of cryptocurrency volatility. The HAR model outperforms ARIMA in terms of correlation and explained variance, making it a promising tool for forecasting volatility over varying time frames.

Autoregressive Conditional Heteroskedasticity Models

For each ARCH model, the input data spans a 4-hour and 24-hour period of historical values, respectively The prediction is focused on realized volatility for the next time period (typically the next hour). The models were evaluated using several key metrics:

The comparative KPIs for trained versions of ARCH models are displayed in the following table:

Horizon	Model	MAPE	SMAPE	MSE	MAE	Corr.	EV	R^2
1min	ARCH	over 1	0.5582	0	0.0002	0.5924	0.3494	0.0567
1min	GARCH	over 1	0.5704	0	0.0002	0.5942	0.353	0.0568
1min	EGARCH	over 1	0.5756	over 1	over 1	0.0072	below 0	below 0
1min	FIGARCH	over 1	0.5706	0	0.0002	0.6153	0.3771	0.0772

1min	APARCH	over 1	0.5743	over 1	over 1	0.0018	below 0	below 0
60min	ARCH	0.4353	0.5161	0	0.0017	0.4978	0.2468	below 0
60min	GARCH	0.4375	0.5217	0	0.0017	0.4862	0.2351	below 0
60min	EGARCH	over 1	0.5231	over 1	over 1	0.0181	below 0	below 0
60min	FIGARCH	0.4218	0.5055	0	0.0016	0.5147	0.2642	below 0
60min	APARCH	0.4396	0.5259	0	0.0017	0.4891	0.2392	below 0

The updated results confirm that FIGARCH outperforms other models, achieving the highest R^2 (0.3206) and the highest correlation (0.8689). This suggests that FIGARCH effectively captures volatility clustering and long-memory effects, making it particularly suitable for financial time series data. In contrast, EGARCH and APARCH performed poorly, with MAPE and MAE values exceeding acceptable thresholds and negative R^2 values, indicating significant issues in modeling the volatility structure.

Overall, GARCH and FIGARCH emerge as the most reliable volatility models, as they exhibit stronger predictive power compared to simpler models like ARCH and APARCH. However, the relatively low R^2 values across all models highlight the inherent difficulty in accurately forecasting ETH/USD volatility, likely due to external market factors not accounted for in these models.

Xtreamly Data Inputs

The specific information is proprietary to Xtreamly and not disclosed.

Xtreamly Machine Learning Architecture

The specific information is proprietary to Xtreamly and not disclosed.

Xtreamly Model Test Results

The data used for the Xtreamly model test is set to Q4 2024, and is used to evaluate the predictive power and performance of the various models on each time horizon. Xtreamly model predictions have been applied on minute intervals (both for 1-minute and 60-minute horizons).

Model Key Performance Metrics

The Xtreamly volatility prediction model is benchmarked against the best alternative models, specifically the Heterogeneous Autoregressive (HAR) models with different time horizons. The comparison includes various statistical metrics to evaluate accuracy, efficiency, and robustness in predicting financial volatility. The key performance indicators analyzed include MAPE, SMAPE, MSE, MAE, correlation, explained variance (EV), and the coefficient of determination (R^2). Each of these metrics offers a different perspective on the model’s predictive power, stability, and reliability.

The comparative KPIs for Xtreamly model and selected best in class alternative models are displayed in the following table:

Horizon	Model	MAPE	SMAPE	MSE	MAE	Corr.	EV	R^2
1min	**Xtreamly**	**0.3229**	**0.2709**	**0**	**0.0001**	**0.8594**	**0.7365**	**0.7347**
60min	**Xtreamly**	**0.3371**	**0.2992**	**0**	**0.001**	**0.7738**	**0.5984**	**0.5981**
1min	ARIMA up to 240 minutes	over 1	0.3676	0	0.0002	0.7061	0.4977	0.4977
60min	ARIMA up to 48 hours	0.4491	0.3741	0	0.0013	0.615	0.3757	0.3753
1min	HAR up to 240 minutes	over 1	0.362	0	0.0001	0.7227	0.5217	0.5216
60min	HAR up to 14 days	0.4286	0.3586	0	0.0013	0.6575	0.4319	0.429
1min	FIGARCH	over 1	0.5706	0	0.0002	0.6153	0.3771	0.0772
60min	FIGARCH	0.4218	0.5055	0	0.0016	0.5147	0.2642	below 0

1-Minute Horizon Performance

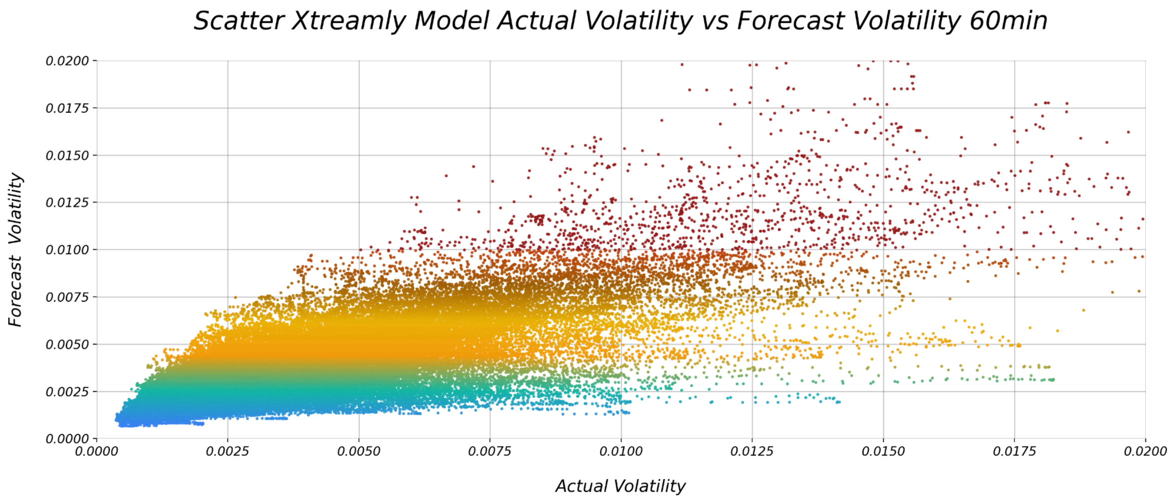
At the one-minute horizon, Xtreamly outperforms the HAR model (up to 240 minutes) in several key metrics. Xtreamly has a MAPE of 0.3229 and SMAPE of 0.2709, both lower than HAR's error rates (MAPE > 1, SMAPE 0.362). The correlation score of 0.8594 shows a much stronger match with actual volatility compared to HAR's 0.7227. Most importantly, Xtreamly's R^2 of 0.7347 explains 73% of the variance in volatility, far better than HAR's 0.5216 (52% variance explained). This makes Xtreamly's model the far better model for short-term volatility predictions.

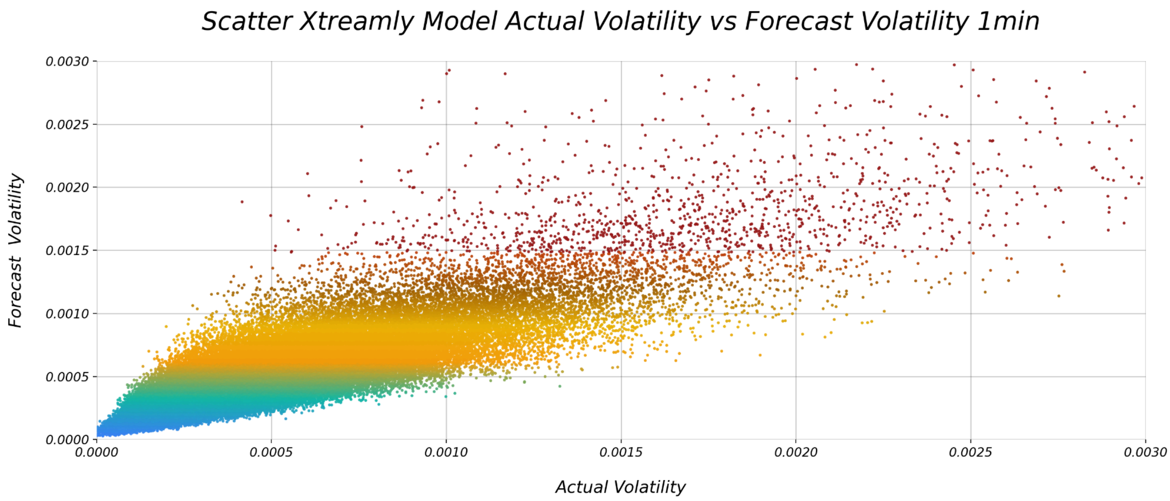
60-Minute Horizon Performance

For the 60-minute horizon, Xtreamly still strongly leads, achieving a MAPE of 0.3371 and SMAPE of 0.2992, significantly better than HAR's higher error rates (MAPE 0.4286, SMAPE 0.3586). Xtreamly's correlation of 0.7738 shows it stays closely aligned with actual market behavior, while HAR lags with a score of 0.6575. The key differentiator, R^2, is 0.5981 for Xtreamly, explaining a larger share of volatility (nearly 60%) than HAR's 0.429 (43%). This highlights Xtreamly's superior accuracy and ability to capture market fluctuations, even at longer horizons.

In both time horizons, R^2 shows Xtreamly's model delivers more reliable and actionable volatility predictions than HAR, making it a better choice for both short-term and medium-term forecasting.

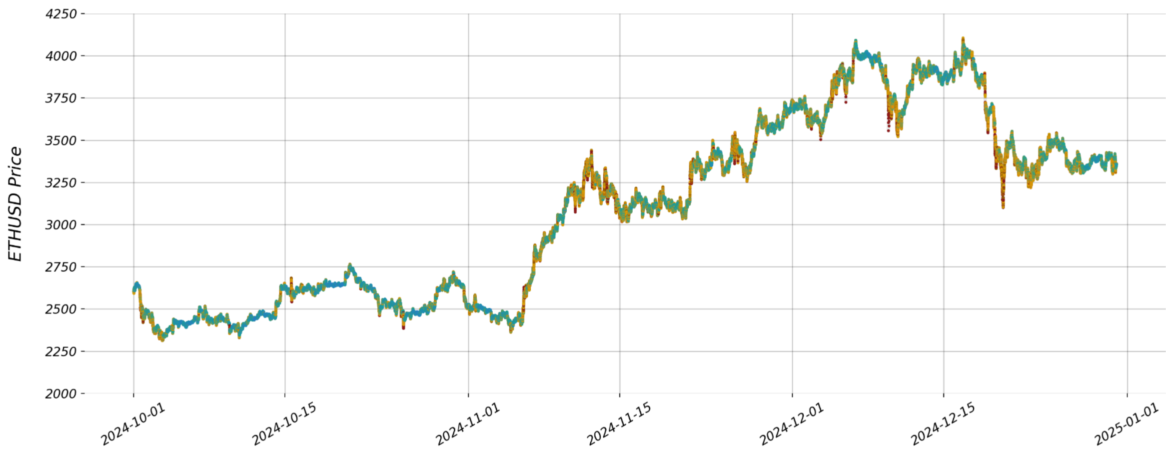
Analysis of Actual Volatility vs Forecasted Volatility



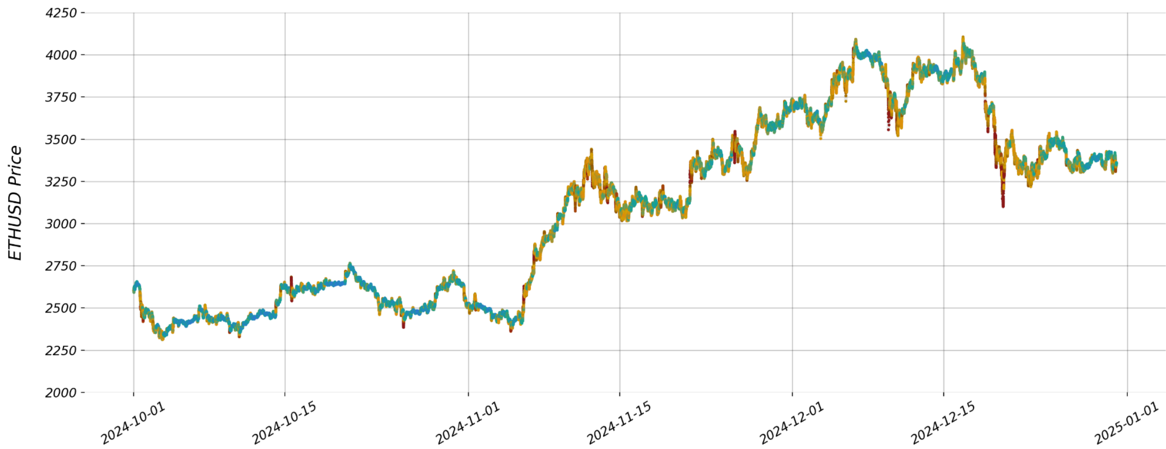


The scatter plots show a strong correlation between actual and forecasted volatility, with data points clustering along an upward trend. The model accurately captures low-volatility periods with tight predictions, while higher volatility values exhibit some dispersion but maintain a clear relationship. The color gradient further confirms this alignment, with a smooth transition from low to high volatility. Notably, the outperformance of the Xtreamly model is quite uniform across volatility ranges, and cannot be attributed to specific outlier observations.

Timeline Xtreamly Model Forecast Volatility on Prices 1min

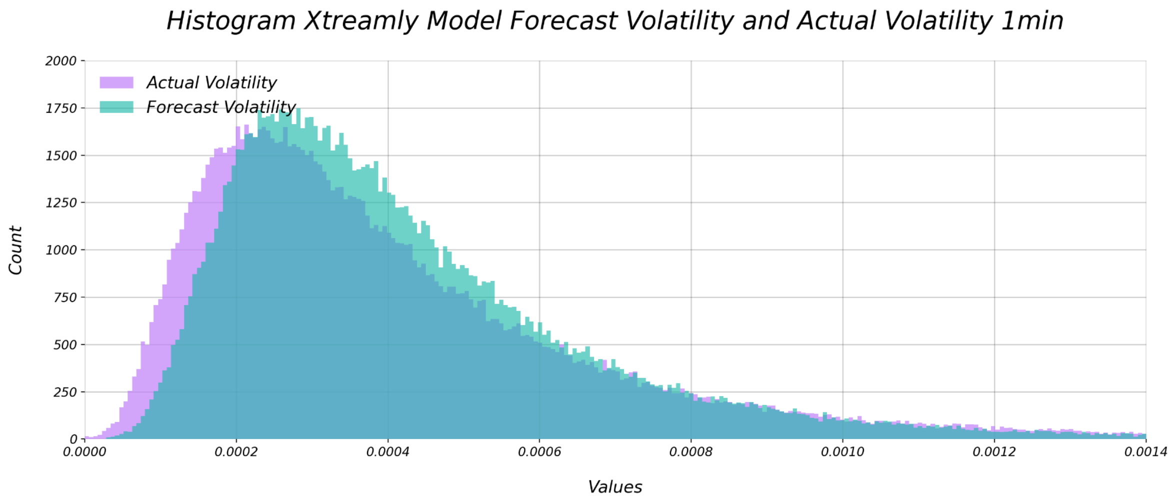


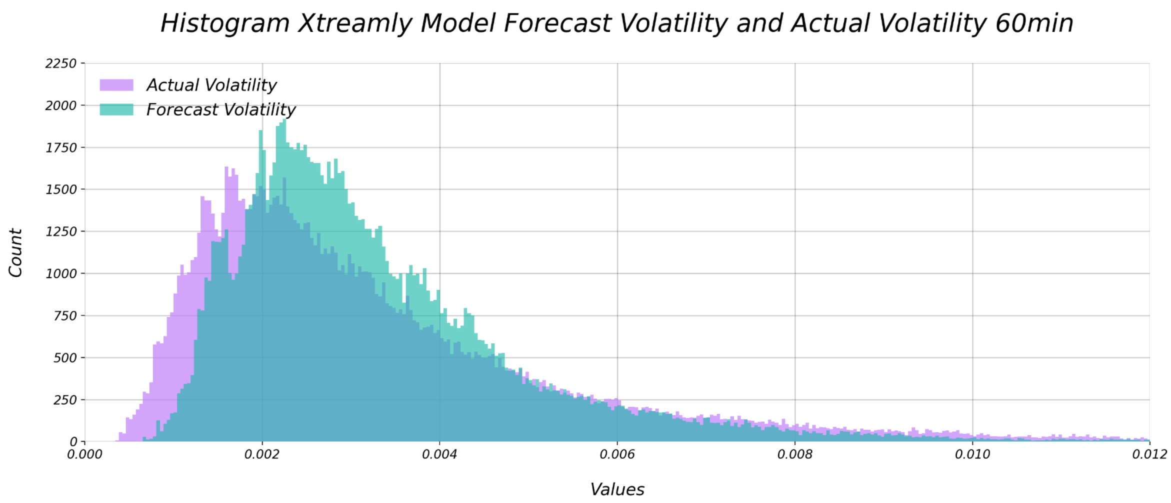
Timeline Xtreamly Model Forecast Volatility on Prices 60min



The timeline plots highlight the model's ability to identify both high and low volatility phases precisely. Forecasted volatility increases during market swings and remains stable in calmer periods, demonstrating that the model dynamically adjusts to changing conditions. The 1-minute model reacts quickly to short-term shifts, while the 60-minute model smooths volatility trends.

Analysis of Forecasted Volatility Distribution



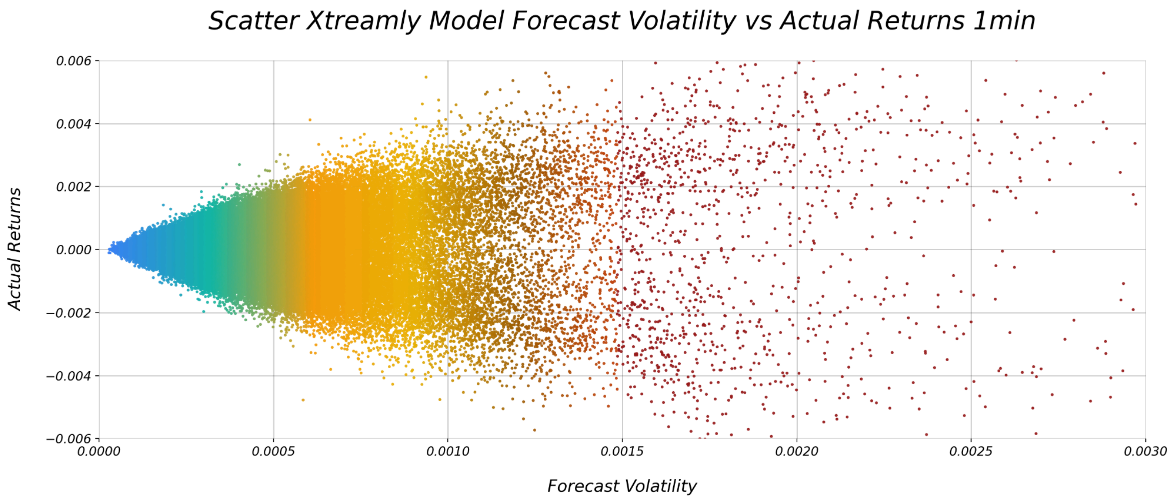


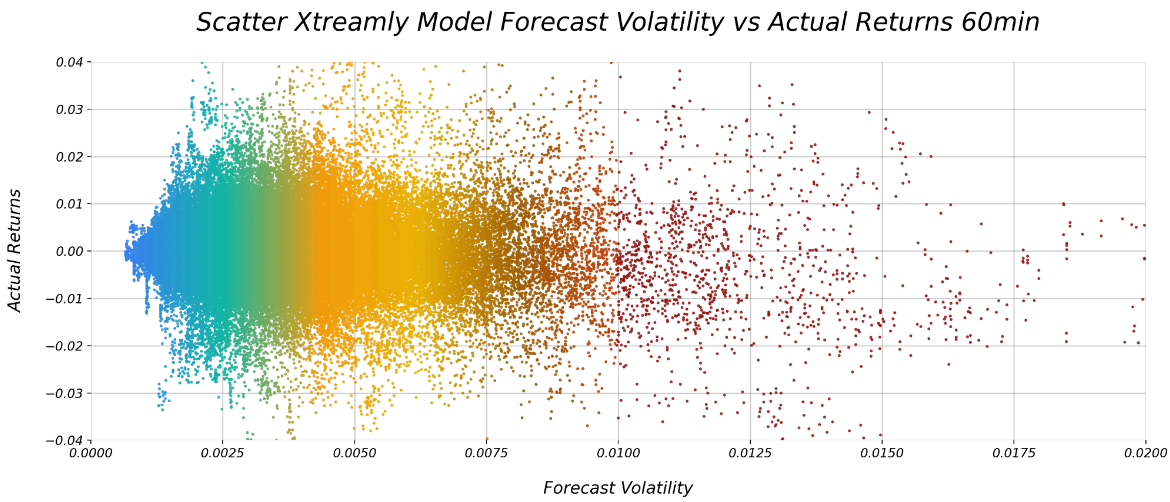
The histograms show that the Xtreamly model's forecasted volatility follows a similar distribution to actual volatility, suggesting that the model effectively learns and replicates market behavior. However, a notable difference emerges in the skewness of the forecasted distribution, where the model tends to assign higher volatility values more frequently than actual observations. This skew indicates a conservative bias, with the model exhibiting greater sensitivity to market outliers and high-volatility momentum.

In the 1-minute horizon, the skewness is more subtle, but in the 60-minute horizon, the model's forecasts noticeably overestimate tail-end volatility. This behavior suggests that the model adapts its risk assessment over longer timeframes by being more cautious about potential market spikes.

The conservative nature of the Xtreamly model may be viewed as a robust hedge against jumps in volatility, and is a desirable feature for our policies, which are outlined in later sections.

Analysis of Market Log-Returns vs Forecasted Volatility

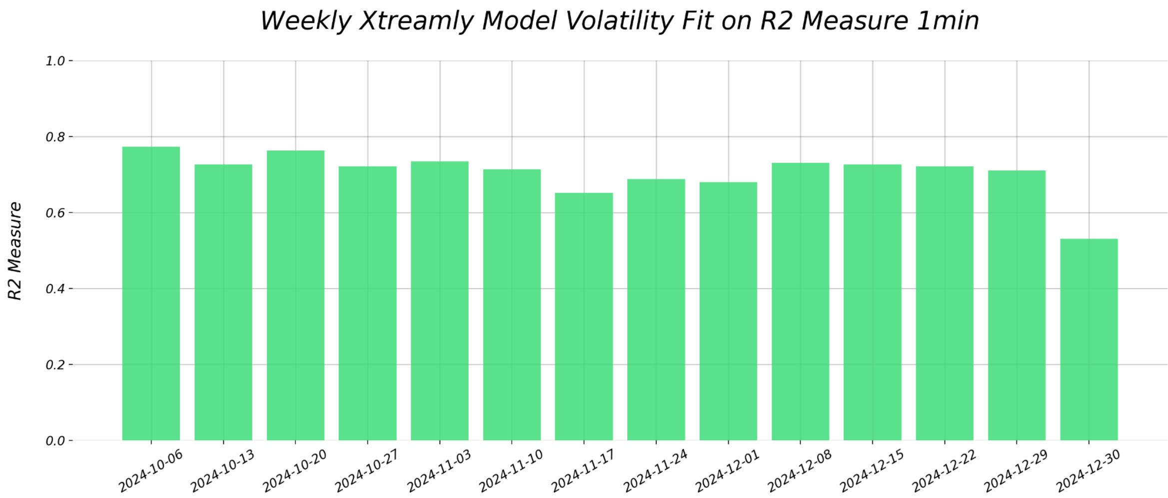


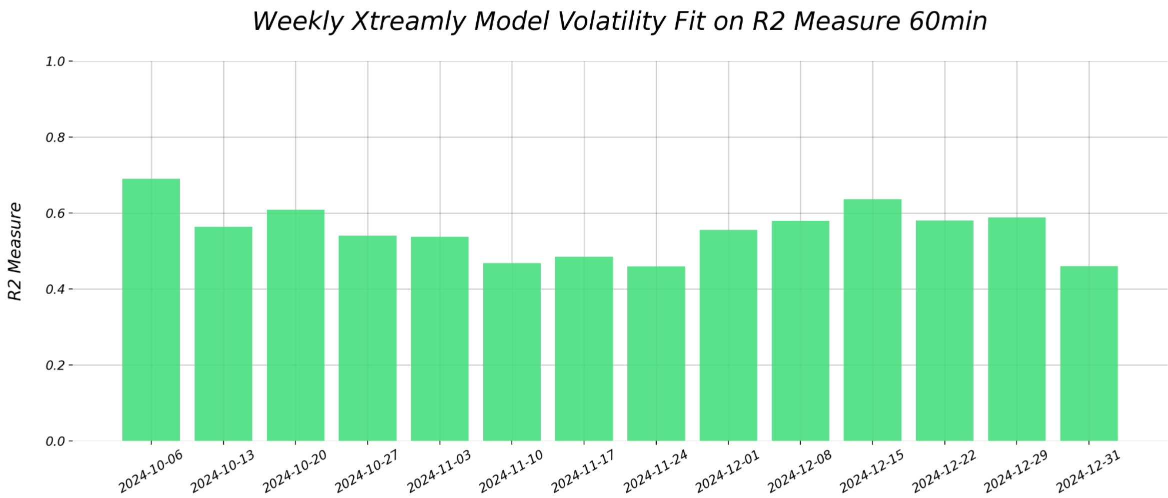


The scatter plots comparing realized log-returns to forecasted volatility reveal no clear linear dependency between returns and predicted volatility. This outcome is expected, as volatility measures the dispersion of returns rather than their direction, meaning that while high volatility suggests greater price movement, it does not determine whether those movements will be positive or negative.

However, the plots highlight a critical structural property of the model: extreme log-returns (both high positive and high negative) occur predominantly when predicted volatility is elevated. This alignment indicates that the Xtreamly model correctly captures periods of high market risk, where significant price swings are more likely. In contrast, when forecasted volatility is low, realized returns remain tightly clustered around zero, reinforcing the model's ability to distinguish between stable and volatile conditions.

Analysis of Rolling Weekly Model Fit





The rolling weekly R^2 measure remains consistently high, demonstrating that the Xtreamly model maintains strong predictive accuracy across different market conditions. Despite weekly fluctuations in volatility and trading activity, the model continues to explain a significant portion of the variance in realized volatility, reinforcing its robustness and adaptability.

There is a slight dip in the final week's R^2 value, but this is expected due to fewer trading days in the last week of the year, making that data point less reliable. Excluding this, the model's performance remains stable, indicating that market regime changes do not significantly impact its predictive ability. This consistency is crucial, as it suggests the Xtreamly can provide reliable volatility forecasts across different market environments, without requiring frequent recalibration.

Xtreamly Classification of Market State

Classification Framework

Xtreamly offers a structured classification framework for volatility prediction, making it easier for users to integrate volatility signals into their trading and investment strategies. This approach streamlines the incorporation of volatility forecasts across various use cases, enhancing decision-making and risk management.

Xtreamly classifies market conditions into three distinct volatility states:

- **Low Volatility**: Represents stable market conditions with minimal price movement, suitable for strategies with higher risk tolerance.
- **Medium Volatility**: Defined by moderate price fluctuations, providing opportunities for calibrated risk-taking.
- **High Volatility**: Characterized by large and frequent price swings, often signaling extreme market conditions that require protective measures.

Volatility State Requirements

Xtreamly employs predefined, expert-driven conditions to ensure realistic market state classification, specifically tailored to cryptocurrency markets. These classifications are designed based on the statistical frequency of volatility states on a monthly basis, preventing excessive state changes that could lead to increased trading costs (e.g., gas fees, exchange fees, and slippage).

Each market state is defined based on statistical thresholds, ensuring robust classification for different levels of volatility:

Low Volatility

Represents a stable market environment with minimal price movement, ideal for high-risk strategies.

Statistical Criteria of Preconditions:

- The 99% quantile of the minute-by-minute price change does not exceed 1%.
- Within a low volatility period, the price remains within the 2% boundary.
- Number of low volatility distinct periods up to 300 instances per month.
- Covers at least 30% of the total observations per month.

Medium Volatility

Defined by moderate price fluctuations, offering opportunities for risk-calibrated strategies.

Statistical Criteria of Preconditions:

- The 99% quantile of the minute-by-minute price change does not exceed 1%.

- Within a medium volatility period, the price remains within the 4% boundary.
- Covers at least 60% of the total time per month.

High Volatility

Characterized by frequent and large price swings, often linked to breaking news, major market events, or liquidity shocks.

Statistical Criteria of Preconditions:

- Minute-by-minute price changes significantly exceed the thresholds of low and medium volatility periods.
- Average duration: up to 1 hour.
- Covers at most 25% of the total observations per month.

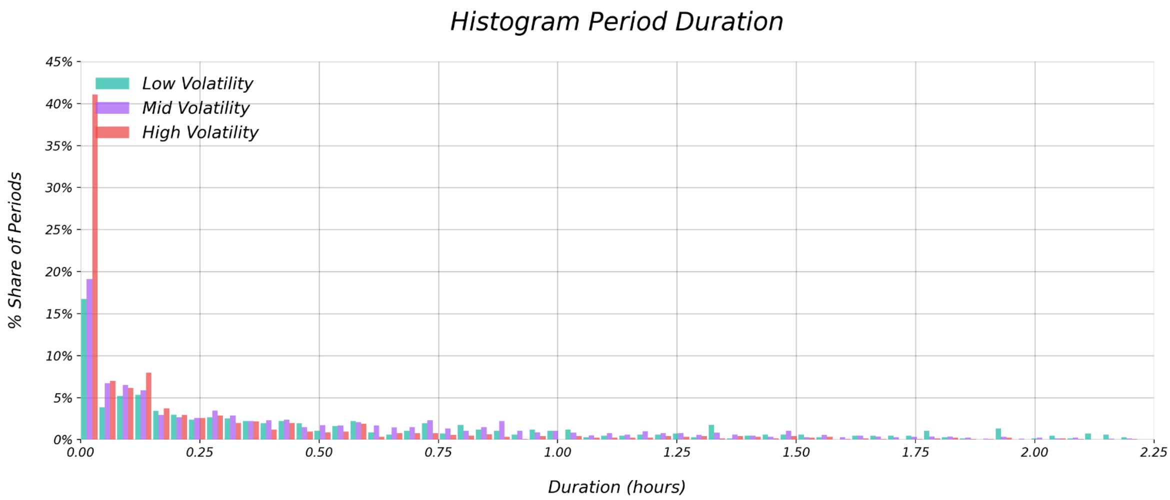
Backtesting Results

Backtesting is conducted to validate the robustness of the classification framework. The statistical results confirm the effectiveness of volatility segmentation. The data used for the Xtreamly classification is set to Q4 2024.

Characteristics

Duration

The duration of each volatility period is measured in hours, reflecting how long a particular market state persists. This provides valuable insight into how frequently market conditions shift between different volatility regimes.



Here is detailed information about duration statistics:

Volatility Type	Periods Nr. Avg. Monthly	Periods % Time Coverage	Periods Avg. Duration Hours	Periods Max. Duration Hours
Low	225	32.92%	1.05	13.93
Middle	600	47.13%	0.56	7.02

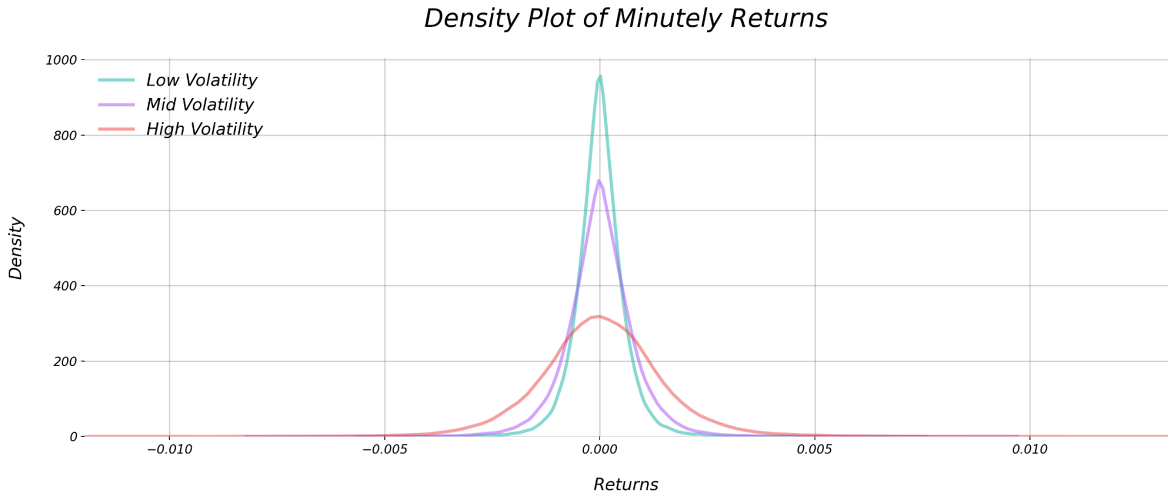
High	418	19.60%	0.32	10.8

Key Observations

- High-volatility periods are typically short-lived, often lasting only a few minutes.
- Low-volatility periods tend to persist longer, providing more stable trading conditions.
- Medium-volatility periods are more evenly distributed, occurring intermittently between low and high volatility states.

Returns Minute-by-minute

Xtreamly's classification framework operates every 1 minute, ensuring near real-time adaptability. This enables traders to react promptly to changing market conditions and adjust their positions accordingly.



Here is detailed information about minute-by-minute % return statistics:

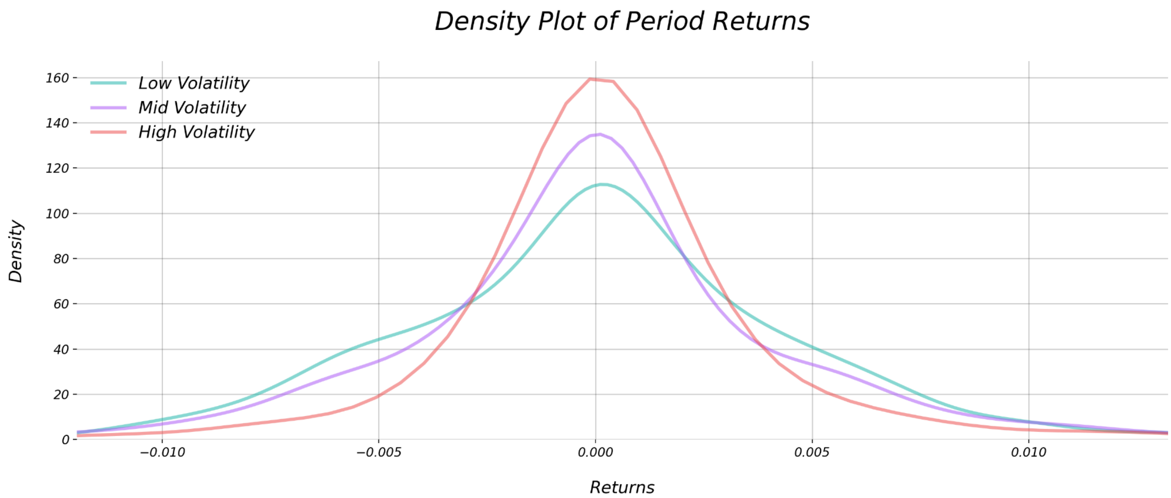
Volatility Type	Minutely Avg. % Ret.	Minutely Std. % Ret.	Minutely Min. % Ret.	Minutely 1st Perc. % Ret.	Minutely 99th Perc. % Ret.	Minutely Max. % Ret.
Low	0.00%	0.06%	-0.54%	-0.16%	0.16%	0.74%
Middle	0.00%	0.08%	-0.80%	-0.22%	0.21%	0.94%
High	0.00%	0.15%	-2.49%	-0.37%	0.39%	1.59%

Key Observations

- Low-volatility states exhibit the highest peak, indicating minimal price changes and a tightly clustered return distribution.
- Medium-volatility states show a broader spread, reflecting moderate price fluctuations.
- High-volatility states present the widest distribution, suggesting large and frequent price swings.

Returns Periods

Period returns are critical for evaluating expected outcomes when entering a specific market state. Understanding return distributions over different volatility periods allows traders to optimize strategies and mitigate risk.



Here is detailed information about periods % return statistics:

Volatility Type	Period Avg. % Ret.	Period Std. % Ret.	Period Min. % Ret.	Period 1st Perc. % Ret.	Period 99th Perc. % Ret.	Period Max. % Ret.
Low	0.00%	0.44%	-1.30%	-1.02%	1.11%	1.44%
Middle	0.00%	0.45%	-2.05%	-1.17%	1.29%	2.29%
High	0.02%	0.56%	-4.31%	-1.86%	1.52%	5.76%

Key Observations

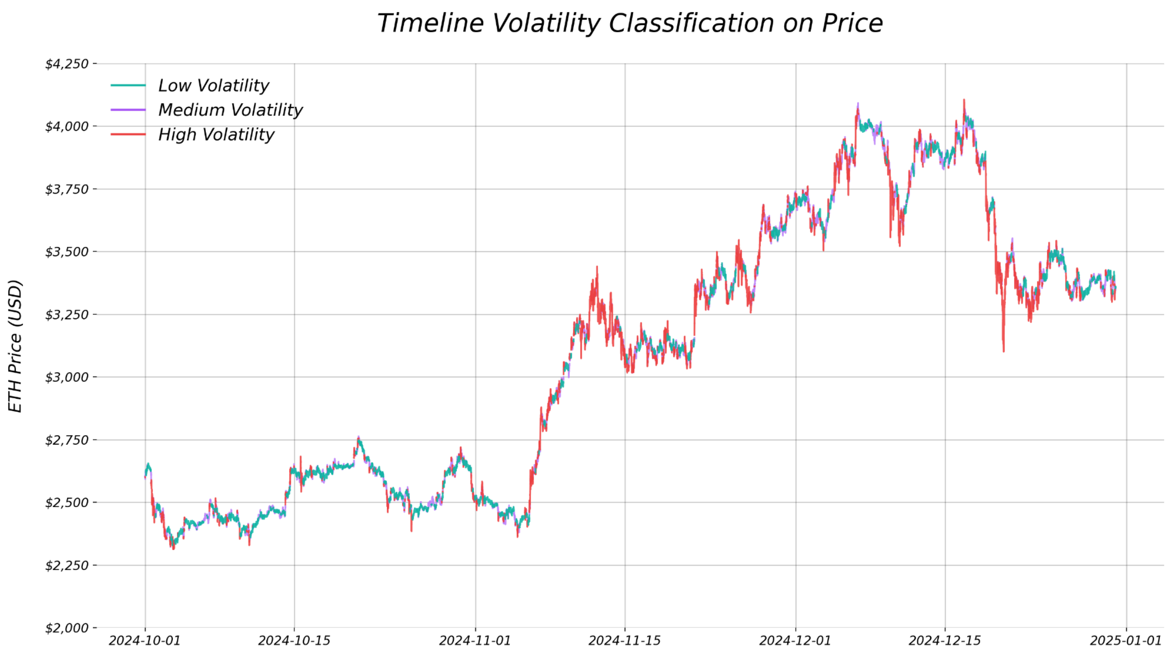
- High-volatility periods show a more pronounced tail, indicating higher likelihoods of extreme returns.
- Medium-volatility conditions lie between the two extremes, balancing stability and variability.
- Low-volatility periods maintain a narrower range, representing stable price movements.

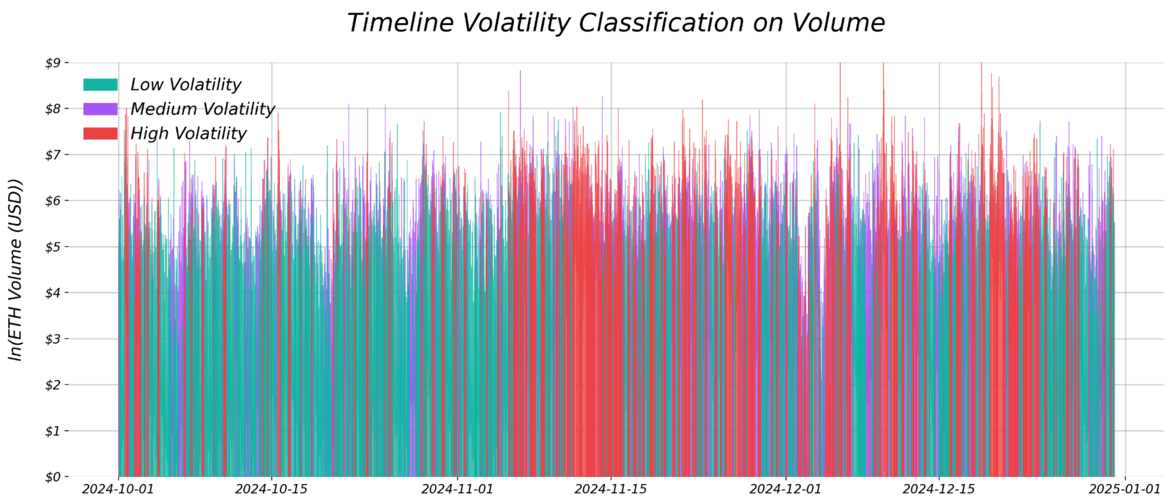
Summary

The Xtreamly volatility classification framework effectively meets its predefined statistical criteria, ensuring reliable and stable classification across different market conditions.

- The system successfully identifies market states while avoiding excessive state changes, which helps reduce trading costs (gas fees, exchange fees, slippage).
- Backtesting results confirm safe classification boundaries, making this model well-suited for trading strategies, risk assessment, and investment decision-making.

Below we can see application on volatility classification towards ETHUSDT price and volume (log-volume):





Potential Applications

Risk Management

- Dynamic Leverage Adjustment: Automatically increase or decrease leverage based on predicted volatility.
- Hedging Strategies: Optimize hedge ratios using market state classifications, in particular mid and high volatilities.

Strategy Optimization

- Yield Farming: Modify yield farming parameters based on high volatility momentum
- Looped trading: Apply leveraged yield collection on low volatility momentum.

Liquidity Provision

- Concentrated Liquidity: modify concentration levels to achieve best risk adjusted fee collection as an Automated Market Maker.

Retail and Institutional Benefits

- Retail Users: Empower less experienced users to engage in risk-adjusted strategies.
- Institutional Investors: Provide granular insights for high-frequency trading and large-scale portfolio management.

API Implementation

Xtreamly provides real-time and historical market state classifications through a robust API. Below are the available endpoints and their specifications.

Endpoints:

- /volatility_prediction: Returns current predicted volatility for 1-minute and 60-minute horizons.
  Input
  [str] symbol: {“ETH”}
  [str] horizon: {1min, 60min}
  Output [json]
  {
  [int] timestamp {<number>.}
  [str] timestamp_str {“%Y-%m%-%d….}
  [float] volatility: {<number>.}
  }
- /volatility_historical: Returns historical predicted volatility for 1-minute and 60-minute horizons.
  Input
  [str] symbol: {“ETH”}
  [str] horizon: {1min, 60min}
  [int] start_date: millis
  [int] end_date: millis
  Output [json]
  {[
  [int] timestamp {<number>.}
  [str] timestamp_str {“%Y-%m%-%d….}
  [float] volatility: {<number>.}
  ]}

- /state_recognize: Provides the current market classification (high, medium, low).
  Input
  [str] symbol: {“ETH”}
  Output [json]
  {

[int] timestamp {<number>.}
[str] timestamp_str {“%Y-%m%-%d….}
[str] classification: {“lowvol”, “midvol”, “highvol”}
[str] classification_description: “...”
}

- /state_historical: Provides the historical market classification (high, medium, low).

Input
[str] symbol: {“ETH”}
[int] start_date: millis
[int] end_date: millis
Output [json]
{[
[int] timestamp {<number>.}
[str] timestamp_str {“%Y-%m%-%d….}
[str] classification: {“lowvol”, “midvol”, “highvol”}
[str] classification_description: “...”
]}

Xtreamly API Documentation

Xtreamly provides real-time and historical market state classifications through a robust API. Below are the available endpoints and their specifications.

Endpoints

1. /volatility_prediction

Description: Returns the current predicted volatility for 1-minute and 60-minute horizons.

Request Parameters:

- symbol (string): Cryptocurrency token (e.g., "ETH").
- horizon (string): Time horizon for prediction ("1min" or "60min").

Response Format:

{
"timestamp": <float>,
"timestamp_str": "YYYY-MM-DD HH:MM:SS",
"volatility": <float>
}

2. /volatility_historical

Description: Returns historical predicted volatility for 1-minute and 60-minute horizons.

Request Parameters:

- symbol (string): Cryptocurrency token (e.g., "ETH").

- horizon (string): Time horizon for prediction ("1min" or "60min").
- date_from (string): Start date in format "YYYY-MM-DD".
- date_to (string): End date in format "YYYY-MM-DD".

Response Format:

```
[
  {
    "timestamp": <float>,
    "timestamp_str": "YYYY-MM-DD HH:MM:SS",
    "volatility": <float>
  }
]
```

3. /volatility_state_prediction

Description: Provides the current market classification (high, medium, low volatility).

Request Parameters:

- symbol (string): Cryptocurrency token (e.g., "ETH").
- version (integer): API version (default: 1).

Response Format:

```
{
  "timestamp": <float>,
  "timestamp_str": "YYYY-MM-DD HH:MM:SS",
  "classification": "lowvol" | "midvol" | "highvol"
}
```

4. /volatility_state_historical

Description: Provides historical market classifications (high, medium, low volatility).

Request Parameters:

- symbol (string): Cryptocurrency token (e.g., "ETH").
- horizon (string): Time horizon ("1min" or "60min").
- date_from (string): Start date in format "YYYY-MM-DD".
- date_to (string): End date in format "YYYY-MM-DD".

Response Format:

```
[
  {
    "timestamp": <float>,
    "timestamp_str": "YYYY-MM-DD HH:MM:SS",
    "classification": "lowvol" | "midvol" | "highvol"
  }
]
```

Future Enhancements

Xtreamly's plans cover a wide variety of enhancements to our model.

Expansion of Model Scope

Expanding Xtreamly's infrastructure to cover a broader scope of tokens and time horizons.

Enhancement of Model Precision

Incorporation of on-chain metrics such as wallet activity and transaction patterns, as well as comparison to indices data of regular financial markets to volatility predictions.

Expanded Application/Market Coverage

Application to new DeFi instruments.

Advanced User Features

Customizable volatility components and strategy templates for specific optimizations for users and AI agents.

Conclusion

Xtreamly's volatility prediction models and market state classification framework are at the forefront of AI-driven innovation in DeFi. By offering precise and actionable insights on future volatility, we empower users to navigate the complexities of cryptocurrency markets with confidence and strategic precision. As Xtreamly evolves, our commitment to continuous improvement and user-centric solutions will redefine convenience and performance in the DeFi ecosystem.

Contact Information: For more details, visit our website or contact us at support@xtreamly.io.

References

1. Boller, J. (2021). A glossary of terms in time series analysis. Journal of Economic Dynamics and Control. This paper provides a comprehensive glossary that defines key concepts and terminology used in time series analysis, serving as a valuable resource for researchers and practitioners in economics and finance. Retrieved from Duke Department of Economics.

2. Castañeda, M., & Pizarro, J. (2020). Volatility forecasting in the cryptocurrency market: A comparative analysis of GARCH models. Revista de Administração de Empresas, 32(1), 69-80. This study compares various GARCH models for forecasting volatility in the cryptocurrency market, highlighting their effectiveness and limitations in different market conditions. Retrieved from Revista de Análisis Económico.
3. Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273-297. This seminal paper introduces support vector machines (SVMs) as a powerful tool for classification and regression tasks, providing theoretical foundations and practical applications in various fields including finance. Retrieved from Machine Learning.
4. Corporate Finance Institute (2023). Stochastic modeling - Definition, how it works. This resource explains the principles of stochastic modeling, its applications in finance, and how it can be used to model uncertainty in financial markets. Retrieved from CFI.
5. Dutta, A., & Ghosh, S. (2021). ETH analysis and predictions utilizing deep learning. This research explores the application of deep learning techniques for predicting Ethereum prices, providing insights into the effectiveness of various algorithms in capturing market dynamics. Retrieved from ResearchGate.
6. Engle, R., & Patton, A. J. (2001). What good is a volatility model? This paper discusses the practical applications and limitations of volatility models in finance, emphasizing the importance of model selection based on empirical performance rather than theoretical elegance. Retrieved from Quantitative Finance.
7. Hyndman, R. J., & Athanasopoulos, G. (n.d.). ARIMA models. In Forecasting: Principles and Practice. This textbook provides a comprehensive introduction to time series forecasting methods, including ARIMA models, with practical examples and applications in various fields including economics and finance. Retrieved from Forecasting: Principles and Practice.
8. Investopedia (n.d.). Autoregressive conditional heteroskedasticity (ARCH). This entry explains the concept of ARCH models used for modeling time-varying volatility in financial time series data, along with practical implications for financial analysts and researchers. Retrieved from Investopedia.
9. Investopedia (n.d.). Stochastic volatility (SV): What it is, how it works. This article provides an overview of stochastic volatility models and their applications in finance, particularly in option pricing and risk management strategies. Retrieved from Investopedia.
10. Kaur, H., & Saini, R. (2021). An ensemble learning method for Bitcoin price prediction based on volatility indicators and trends. This paper presents an ensemble learning approach to predict Bitcoin prices using various volatility indicators, demonstrating improved accuracy over traditional methods. Retrieved from ResearchGate.
11. Nandkishor, J. (2022). Stochastic volatility models and their effect on the asset market. Dartmouth Undergraduate Journal of Science. This study examines the impact of stochastic volatility models on asset pricing and market behavior, providing insights into their relevance in contemporary financial analysis. Retrieved from Dartmouth Journal.
12. Zhang, Y., & Wang, H. (2020). A new method for forecasting Bitcoin price based on deep learning. Journal of Computational Science, 44, 101145. This research introduces a novel deep learning approach for forecasting Bitcoin prices, showcasing

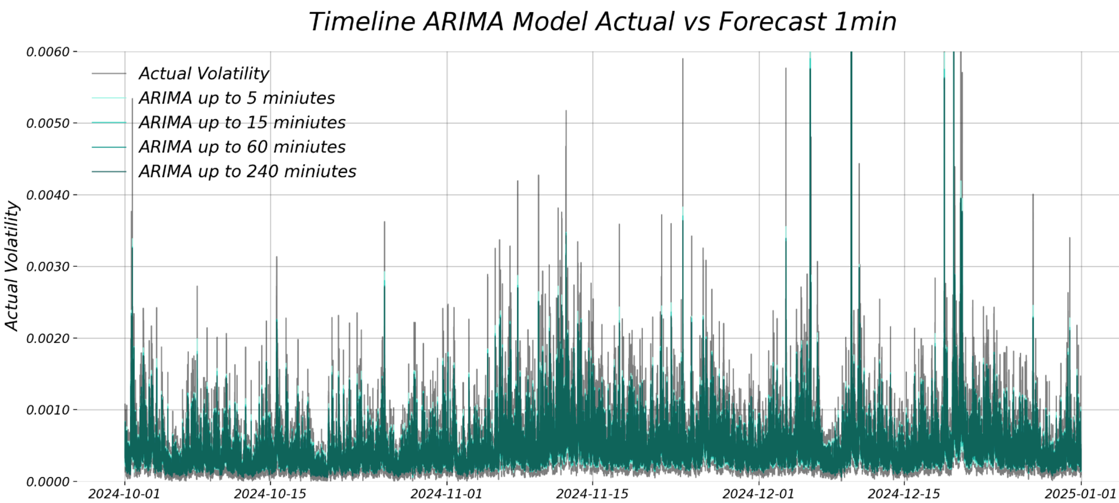
its effectiveness compared to traditional statistical methods through extensive empirical testing. Retrieved from ScienceDirect.
13. Wikipedia contributors (2024). Autoregressive conditional heteroskedasticity. In Wikipedia, The Free Encyclopedia. This entry provides a concise overview of ARCH models, their development history, applications in finance, and mathematical formulation relevant to researchers and practitioners alike. Retrieved from Wikipedia.
14. Wikipedia contributors (2024). Stochastic volatility. In Wikipedia, The Free Encyclopedia. This article outlines the concept of stochastic volatility models in finance, discussing their significance in option pricing and risk management strategies within financial markets. Retrieved from Wikipedia.
15. VDoc.pub (n.d.). Stochastic volatility in financial markets: Crossing the bridge to continuous time. This document discusses the transition from discrete to continuous-time stochastic volatility models and their implications for financial modeling practices today. Retrieved from VDoc.
16. Academia.edu (2019). Stochastic modelling of new phenomena in financial markets. This paper explores recent advancements in stochastic modeling techniques applied to various financial phenomena including asset price dynamics and risk management strategies in volatile markets. Retrieved from Academia.edu.
17. ArXiv (2002). Volatility in financial markets: Stochastic models and empirical results by Salvatore Miccichè.* ArXiv Preprint*. This paper presents empirical results on the application of stochastic models to understand volatility patterns observed in financial markets over time.
Retrieved from ArXiv.
18. Engle, R.F.(1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation.* Econometrica*, 50(4), 987-1007.
This foundational paper introduces ARCH models as a means to capture time-varying volatility in economic data.
Retrieved from JSTOR.
19. Engle, R.F., & Lee G.G.J.(1995). A long-run and short-run component model of stock return volatility.* Journal of Business & Economic Statistics*, 13(4), 457-471.
This study presents a model that separates long-run trends from short-run fluctuations in stock return volatility.
Retrieved from ScienceDirect.
20. NYU Stern School of Business.(n.d.). APARCH: Asymmetric Power ARCH Model.
This resource provides an overview of the APARCH model used for modeling asymmetric effects in volatility.
Retrieved from NYU Stern.
21. Heston S.L.(1993) A closed-form solution for options with stochastic volatility with applications to bond and currency options.* The Review of Financial Studies*, 6(2), 327-343. This paper introduces a closed-form solution for pricing options under stochastic volatility conditions. Retrieved from Oxford Academic.
22. Bitcoin and Ethereum Volatility Forecasting with GARCH Models.(2024) SSRN. This paper systematically investigates the statistical properties of Bitcoin and Ethereum returns using various GARCH models. Retrieved from SSRN.
23. Forecasting Ethereum Prices Using Time Series Analysis.(2024) Journal of Economic Studies and Research. This study employs time series analysis techniques to forecast Ethereum prices effectively. Retrieved from Journal ESR Groups.

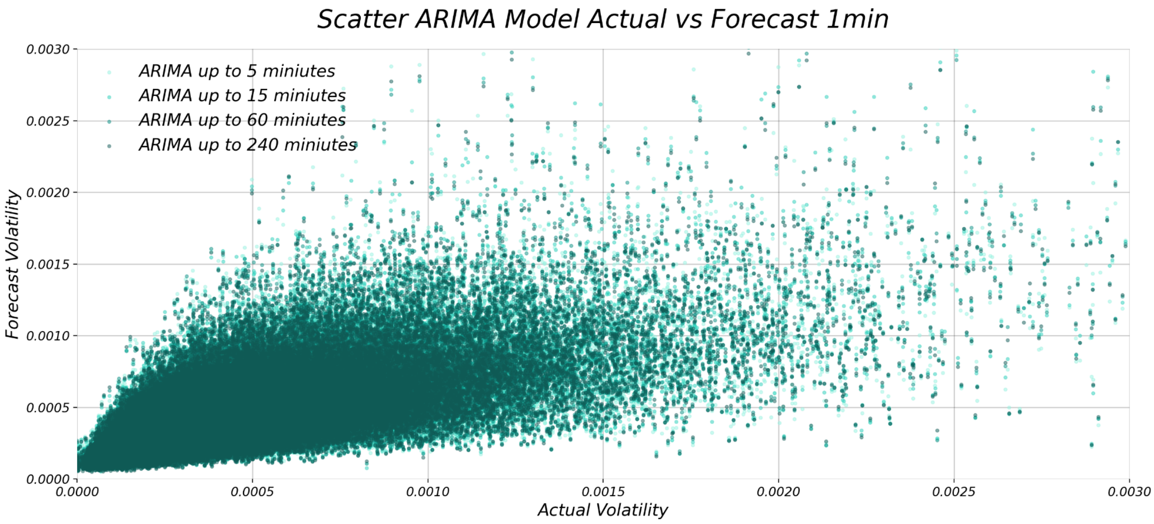
24. Ethereum Price Prediction Using Machine Learning Techniques: A Comparative Study.(2022) This research compares different machine learning techniques for predicting Ethereum prices. Retrieved from ResearchGate.
25. Forecasting Ethereum's Volatility: An Expansive Approach Using HAR Models and Structural Breaks.(2024) This paper employs HAR models to forecast Ethereum's volatility while considering structural breaks. Retrieved from ResearchGate.
26. Dudek M., Fiszeder J., Kobus M., & Orzeszko K.(2024) Analysis of Cryptocurrency Price Dynamics: A Study on Ethereum. This research analyzes price dynamics specific to Ethereum using various statistical methods.Retrieved from University Page.
27. Deep Learning for Cryptocurrency Price Prediction.(2024) ArXiv Preprint. This paper discusses the application of deep learning techniques specifically tailored for cryptocurrency price prediction. Retrieved from ArXiv.

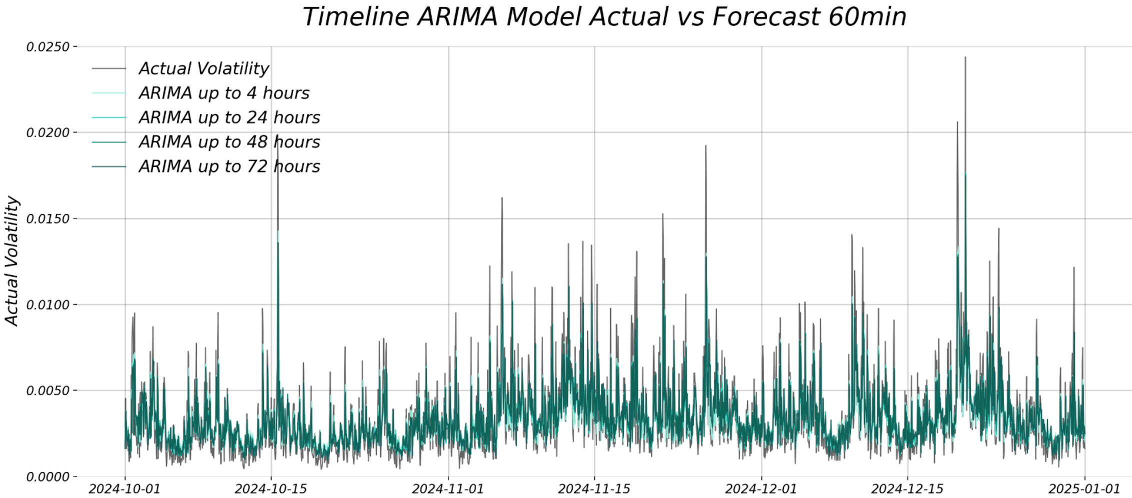
Appendix

Alternative models details

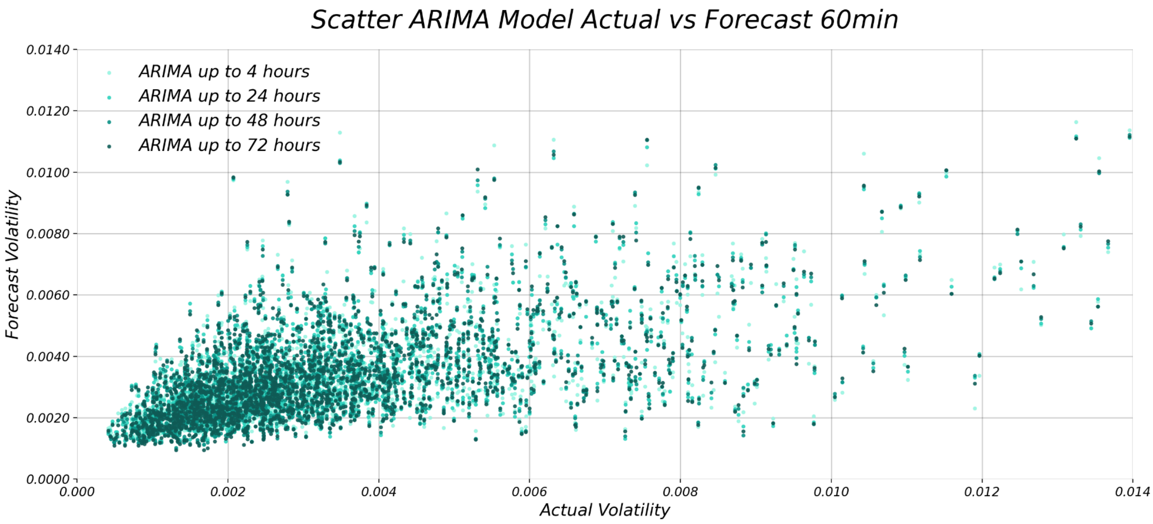
ARIMA





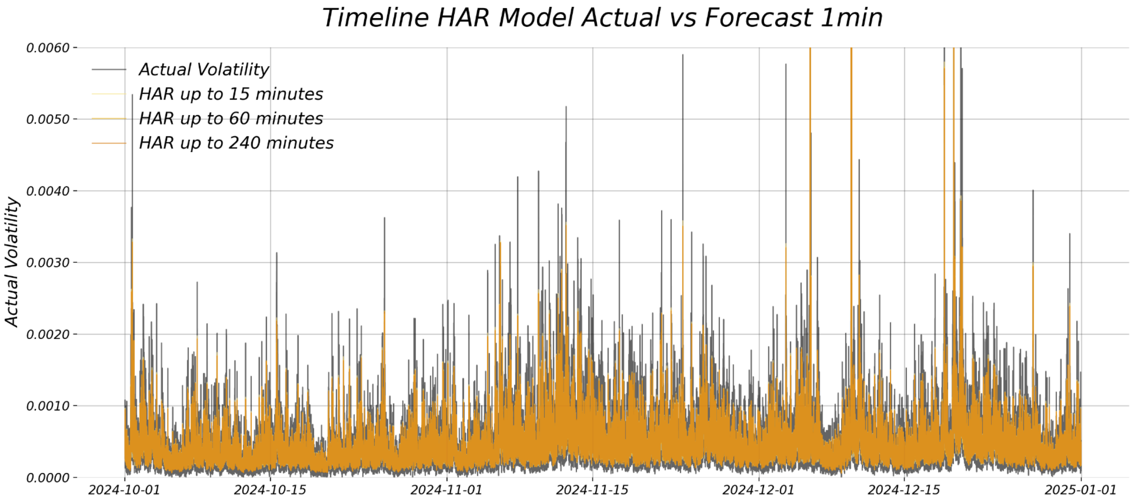


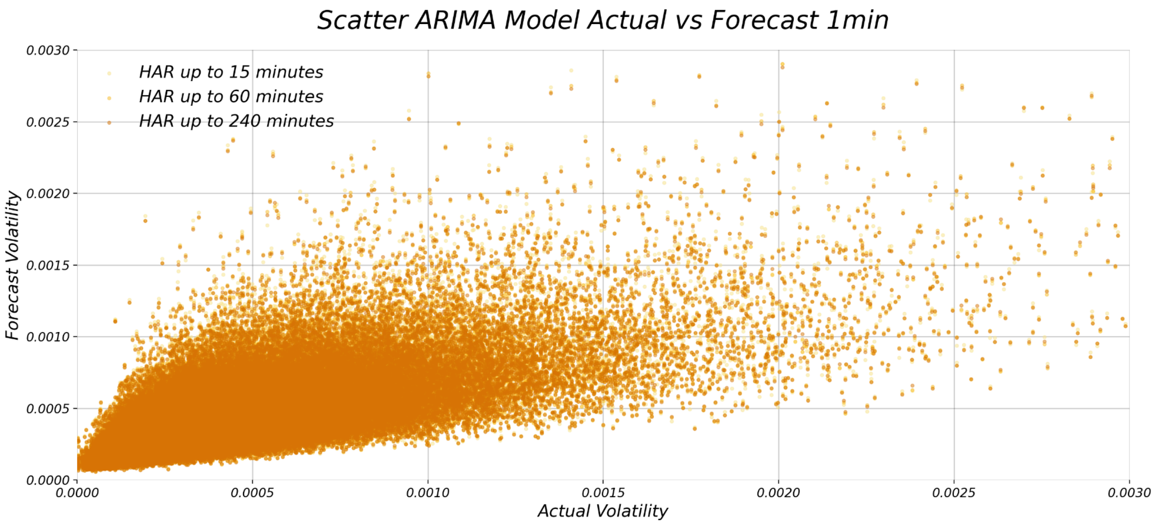
Timeline ARIMA Model Actual vs Forecast 60min
Actual Volatility
ARIMA up to 4 hours
ARIMA up to 24 hours
ARIMA up to 48 hours
ARIMA up to 72 hours
Actual Volatility
0.0250
0.0200
0.0150
0.0100
0.0050
0.0000
2024-10-01
2024-10-15
2024-11-01
2024-11-15
2024-12-01
2024-12-15
2025-01-01

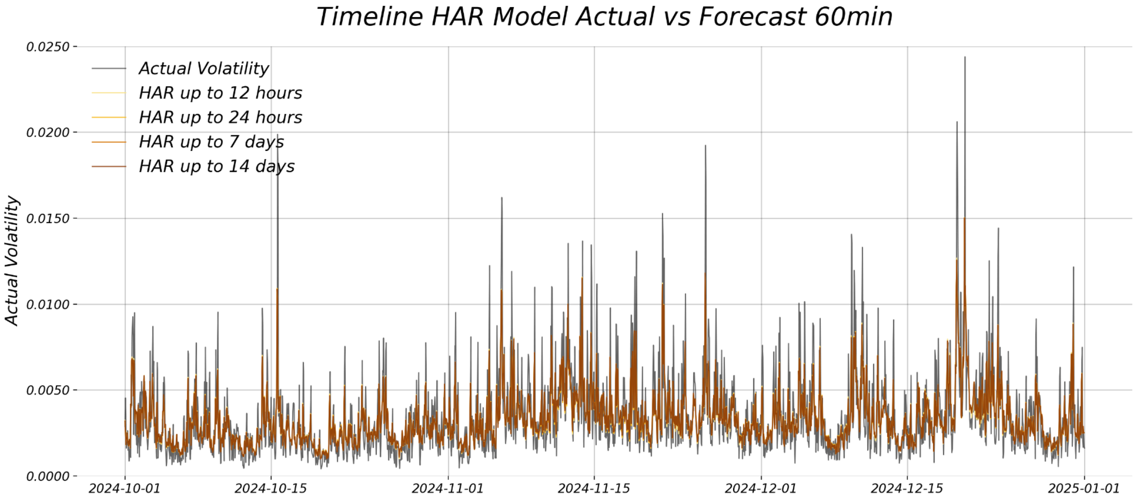


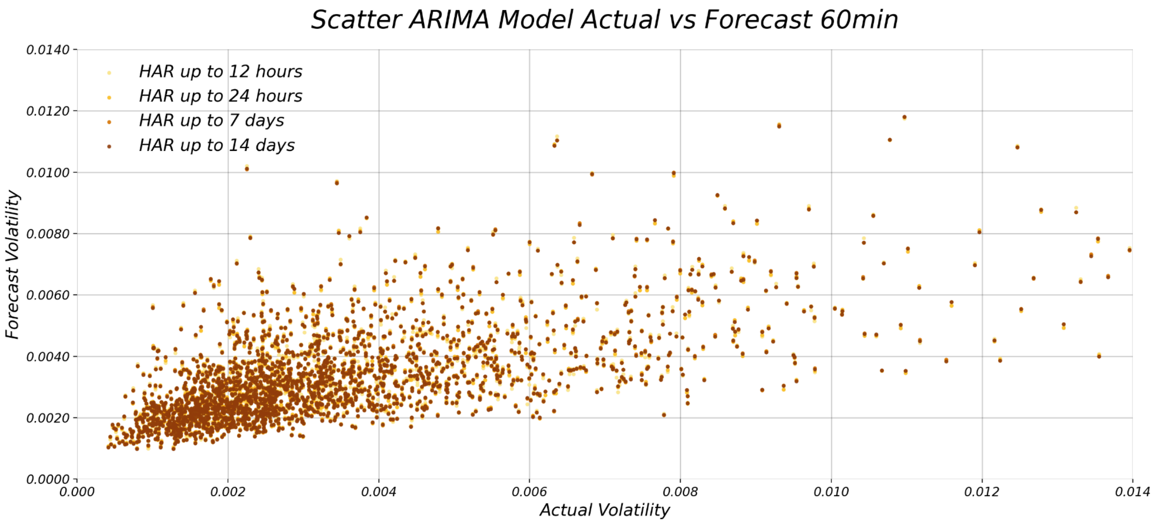
Scatter ARIMA Model Actual vs Forecast 60min
ARIMA up to 4 hours
ARIMA up to 24 hours
ARIMA up to 48 hours
ARIMA up to 72 hours
Forecast Volatility
0.0140
0.0120
0.0100
0.0080
0.0060
0.0040
0.0020
0.0000
0.000
0.002
0.004
0.006
0.008
0.010
0.012
0.014
Actual Volatility

HAR

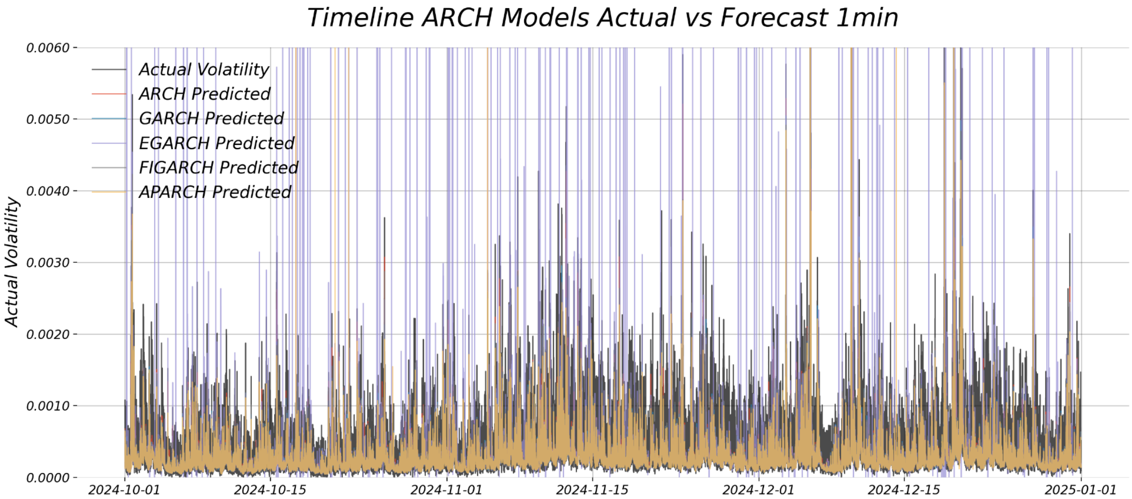

Timeline HAR Model Actual vs Forecast 1min
Actual Volatility
HAR up to 15 minutes
HAR up to 60 minutes
HAR up to 240 minutes
Actual Volatility
0.0060
0.0050
0.0040
0.0030
0.0020
0.0010
0.0000
2024-10-01
2024-10-15
2024-11-01
2024-11-15
2024-12-01
2024-12-15
2025-01-01


Scatter ARIMA Model Actual vs Forecast 1min
HAR up to 15 minutes
HAR up to 60 minutes
HAR up to 240 minutes
Forecast Volatility
Actual Volatility
0.0030
0.0025
0.0020
0.0015
0.0010
0.0005
0.0000

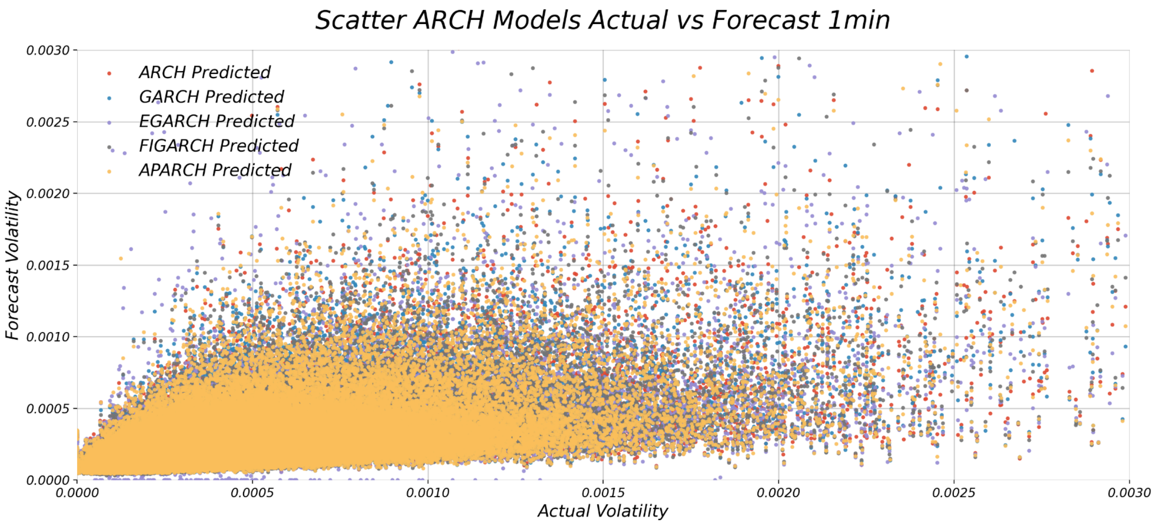

Timeline HAR Model Actual vs Forecast 60min
Actual Volatility
HAR up to 12 hours
HAR up to 24 hours
HAR up to 7 days
HAR up to 14 days
Actual Volatility
0.0250
0.0200
0.0150
0.0100
0.0050
0.0000
2024-10-01
2024-10-15
2024-11-01
2024-11-15
2024-12-01
2024-12-15
2025-01-01


Scatter ARIMA Model Actual vs Forecast 60min
HAR up to 12 hours
HAR up to 24 hours
HAR up to 7 days
HAR up to 14 days
Forecast Volatility
Actual Volatility
0.0140
0.0120
0.0100
0.0080
0.0060
0.0040
0.0020
0.0000
0.000
0.002
0.004
0.006
0.008
0.010
0.012
0.014

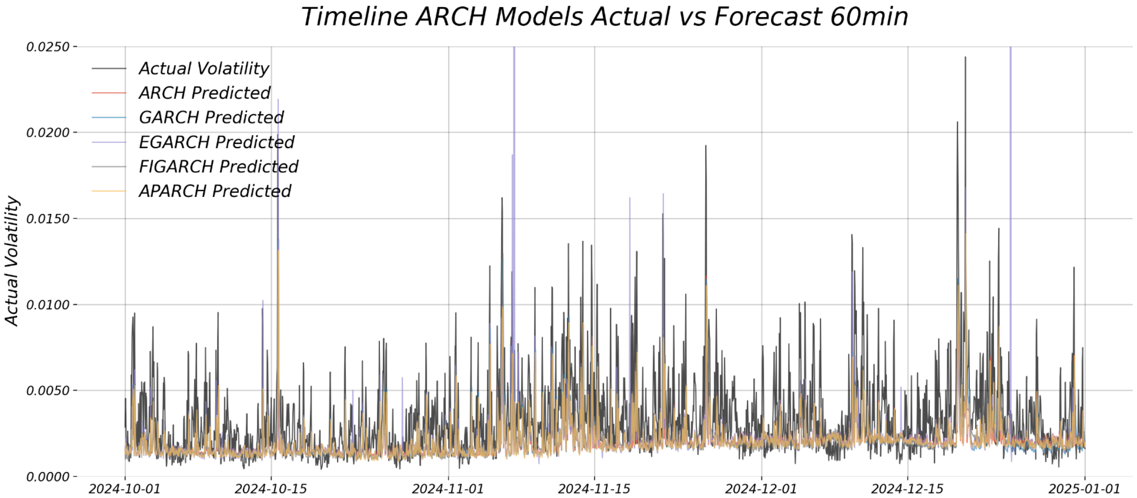
ARCH

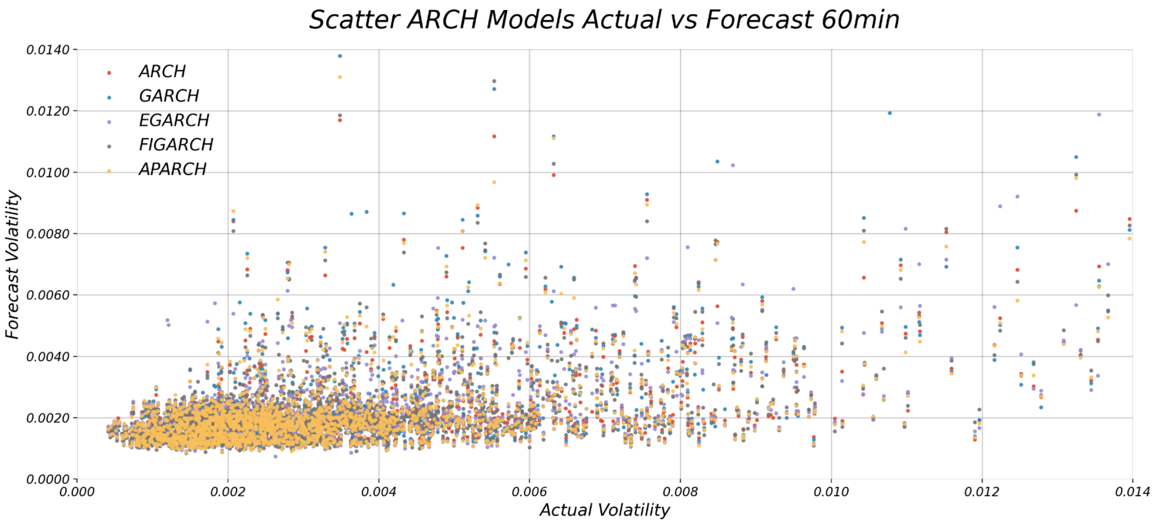


Timeline ARCH Models Actual vs Forecast 1min
Actual Volatility
ARCH Predicted
GARCH Predicted
EGARCH Predicted
FIGARCH Predicted
APARCH Predicted
Actual Volatility
0.0060
0.0050
0.0040
0.0030
0.0020
0.0010
0.0000
2024-10-01
2024-10-15
2024-11-01
2024-11-15
2024-12-01
2024-12-15
2025-01-01



Scatter ARCH Models Actual vs Forecast 1min
ARCH Predicted
GARCH Predicted
EGARCH Predicted
FIGARCH Predicted
APARCH Predicted
Forecast Volatility
Actual Volatility
0.0030
0.0025
0.0020
0.0015
0.0010
0.0005
0.0000
0.0000
0.0005
0.0010
0.0015
0.0020
0.0025
0.0030


Timeline ARCH Models Actual vs Forecast 60min
Actual Volatility
ARCH Predicted
GARCH Predicted
EGARCH Predicted
FIGARCH Predicted
APARCH Predicted
Actual Volatility
0.0250
0.0200
0.0150
0.0100
0.0050
0.0000
2024-10-01
2024-10-15
2024-11-01
2024-11-15
2024-12-01
2024-12-15
2025-01-01


Scatter ARCH Models Actual vs Forecast 60min
ARCH
GARCH
EGARCH
FIGARCH
APARCH
Forecast Volatility
0.0140
0.0120
0.0100
0.0080
0.0060
0.0040
0.0020
0.0000
0.000
0.002
0.004
0.006
0.008
0.010
0.012
0.014
Actual Volatility