# Beyond GARCH: Kernel-Based Volatility and Tail-Risk Forecasting for Ethereum

Lei Pan*

**Abstract**

This paper studies volatility prediction for Ethereum in the post-Merge era. Using daily ETH/USD returns from 15 September 2022 to 23 April 2026, we compare standard GARCH(1,1), Heston–Nandi GARCH(1,1), cross-validated and aggregated EWMA predictors, and Nadaraya–Watson kernel-regression predictors. The kernel forecasts are constructed from a rank-transformed state vector that captures recent volatility and signed-return conditions, allowing the conditional variance function to be nonlinear and state dependent. The results show that forecast performance is strongly horizon dependent. At the one-day horizon, the kernel predictor using the fitted GARCH volatility state delivers the lowest final cumulative squared prediction error, outperforming the standard GARCH benchmark and all EWMA-type competitors. At the ten-day-ahead horizon, the advantage of local nonparametric information weakens, and the mean-reverting structure of GARCH becomes more valuable. The estimated kernel surface reveals that predicted ETH volatility is highest when elevated recent volatility coincides with negative signed-return pressure. Conditional quantile results further show that kernel-based VaR improves lower-tail risk forecasts, especially at the 1% quantile. Overall, the evidence suggests that post-Merge Ethereum volatility is persistent, asymmetric, heavy-tailed, and nonlinear, and is best modelled by combining economically meaningful volatility states with flexible nonparametric forecasting maps.

**Keywords:** Ethereum; volatility forecasting; kernel regression; GARCH; value-at-risk
**JEL Codes:** C14; C22; C53; G17

## 1 Introduction

At 6:42 a.m. UTC on 15 September 2022, Ethereum completed the Merge. The event was not a routine software update. It replaced Ethereum's proof-of-work consensus mechanism with proof-of-stake and transformed the economic structure of the network. A blockchain that had previously relied on energy-intensive mining became a staking-based settlement layer for decentralised finance, stablecoins, non-fungible tokens, tokenised assets, and smart-contract applications. In financial terms, the Merge created a clean historical break. The asset still traded under the same ticker, ETH, but the mechanism producing network security, staking rewards, validator incentives, and investor narratives changed substantially. For empirical asset pricing and risk management, this raises a direct question: should post-Merge Ethereum volatility still be modelled as if it were merely another cryptocurrency price series, or does the new regime require a forecasting framework that can adapt to nonlinear and state-dependent market behaviour?

This question matters because Ethereum is no longer only a speculative digital token. It is the native asset of a large programmable settlement network, and its price dynamics are increasingly connected to institutional trading, derivatives activity, staking conditions, decentralised-finance liquidity, regulatory

*Corresponding author. School of Accounting, Economics, and Finance, Curtin University, Kent Street, Bentley, Australia. Email: lei.pan@curtin.edu.au.

news, and exchange-traded-product expectations. The approval process for spot Ether exchange-traded products in the United States further strengthened the connection between ETH and traditional financial markets. As a result, volatility forecasting for Ethereum is no longer a narrow technical exercise. It is relevant for portfolio allocation, derivative pricing, margin setting, risk limits, exchange supervision, and the measurement of downside exposure. A forecast of tomorrow's ETH variance is not simply a number produced by a statistical model; it is a summary of how the market interprets information about network fundamentals, regulatory events, liquidity conditions, and investor positioning.

The academic literature on digital assets has developed rapidly since the introduction of Bitcoin. Early work emphasised the economic novelty of cryptocurrencies and the institutional design of decentralised payment systems (Nakamoto 2008, Böhme et al. 2015, Yermack 2015, Dwyer 2015). A related literature studies whether cryptoassets behave as currencies, commodities, speculative assets, or alternative stores of value (Cheah and Fry 2015, Dyhrberg 2016, Baur et al. 2018, Bouri et al. 2017, Klein et al. 2018). Another strand investigates market efficiency, price discovery, persistence, and spillover effects in cryptocurrency markets (Urquhart 2016, Nadarajah and Chu 2017, Phillip et al. 2018, Wei 2018, Brauneis and Mestel 2018, Caporale et al. 2018, Koutmos 2018). These studies generally conclude that cryptocurrency markets exhibit high volatility, heavy tails, time variation, and changing degrees of efficiency. They also show that digital assets cannot be fully understood using the same empirical regularities that describe equities, currencies, or precious metals.

Within this broader literature, volatility modelling has become one of the central empirical problems. Cryptocurrency returns display the stylised facts familiar from financial econometrics, volatility clustering, excess kurtosis, serial dependence in squared returns, and large jumps, but these features are often more extreme than in mature asset classes. Katsiampa (2017) compares GARCH-type models for Bitcoin and shows that rich conditional-variance specifications can capture important features of crypto volatility. Gkillas and Katsiampa (2018) use extreme-value methods to study cryptocurrency tail behaviour, while Borri (2019) documents the importance of conditional tail risk. Katsiampa et al. (2019) provide evidence of volatility spillovers among leading cryptocurrencies, including Ether. Liu and Tsyvinski (2021) show that the risk-return trade-off of cryptocurrencies is distinct from that of stocks, currencies, and precious metals. Taken together, these contributions suggest that crypto volatility is persistent and forecastable, but also nonlinear, unstable, and sensitive to market-state changes.

The standard econometric starting point for modelling volatility is the ARCH–GARCH tradition. Engle (1982) introduced autoregressive conditional heteroskedasticity, and Bollerslev (1986b) generalised it into the GARCH framework that remains a benchmark in financial econometrics. Extensions such as EGARCH, GJR-GARCH, and news-impact models allow volatility to respond asymmetrically to positive and negative shocks (Nelson 1991, Glosten et al. 1993, Engle and Ng 1993b). The Heston–Nandi GARCH model provides a tractable parametric structure in which conditional variance responds to signed shocks and can be connected to option valuation (Heston and Nandi 2000b). These models are attractive because they are parsimonious, interpretable, and easy to implement. They impose a small number of parameters on a potentially complex volatility process, and this discipline is often valuable for multi-step forecasting.

However, precisely this parametric discipline can become restrictive in a market such as Ethereum. A single GARCH persistence coefficient assumes that the same volatility-updating rule applies in calm periods, crash periods, upside rallies, regulatory repricing episodes, and liquidity-driven volatility bursts. A fixed asymmetry parameter assumes that the difference between positive and negative shocks is stable over time. These assumptions are convenient, but they may be too strong for post-Merge ETH. Ethereum experienced several distinct regimes after September 2022, including the immediate adjustment to proof-

of-stake, the Shapella upgrade, the spot-ETF repricing episode, and subsequent periods of sharp market stress and recovery. If the volatility response depends on where the market is located in the state space, then a model with a constant global updating rule may either underreact in turbulent periods or overreact when shocks are transitory.

This paper therefore studies whether a flexible kernel-regression approach can improve volatility prediction for Ethereum in the post-Merge era. The idea is simple but powerful. Instead of imposing a fixed parametric law of motion for conditional variance, kernel regression estimates expected future squared returns by looking at historical observations that occurred under similar market states. In the spirit of the nonparametric regression literature (Nadaraya 1964, Watson 1964, Härdle 1990), and following the volatility-prediction framework of Klemelä (2017b) and the Bitcoin application of Sanli et al. (2025), the estimator treats volatility forecasting as a conditional-expectation problem. The target variable is the future squared return, while the conditioning variables summarise the recent volatility level and the recent direction of returns. This permits the forecast function to be nonlinear, locally adaptive, and state dependent.

The empirical design of the paper is deliberately comparative. We do not argue that kernel regression should replace parametric volatility models in every setting. Instead, we ask where and when a nonparametric method adds information beyond standard benchmarks. The paper compares standard GARCH(1,1), Heston–Nandi GARCH(1,1), cross-validated EWMA predictors, aggregated EWMA predictors, and kernel-regression predictors. Forecast accuracy is evaluated using cumulative squared prediction errors relative to the GARCH benchmark. The analysis considers one-day and ten-day prediction horizons, because the value of local information need not be the same at short and longer horizons. A model that responds quickly to the current state may be very effective for tomorrow's realised variance, but may perform less well when volatility shocks mean-revert over a longer horizon.

A key feature of the paper is the construction of the predictor space. The first predictor measures recent volatility conditions using smoothed squared returns or a fitted volatility state. The second predictor measures recent signed-return conditions. These variables are then transformed into normal scores using empirical ranks. This transformation preserves the ordering of observations while making the marginal distributions more homogeneous. For kernel regression, this step is important because distances in the predictor space determine the weights assigned to historical observations. Without transformation, a single bandwidth can imply very different degrees of smoothing in dense and sparse regions of the raw predictor space. After transformation, the bandwidth has a more stable interpretation across market states. Economically, the resulting state vector distinguishes calm states, downside-stress states, and upside-turbulence states.

The paper makes three main contributions. First, it provides a post-Merge Ethereum-specific volatility forecasting exercise. Much of the existing crypto-volatility literature focuses on Bitcoin or on multi-asset cryptocurrency panels. Ethereum is different. It is a programmable settlement asset, its supply and staking incentives changed after the Merge, and its market narrative is closely linked to smart-contract activity and institutional adoption. By focusing on daily ETH returns from the post-Merge period, this paper isolates a cleaner economic regime and asks whether volatility forecasting behaves differently after the network's structural transition.

Second, the paper brings together parametric and nonparametric volatility forecasting within a single empirical horse race. Standard GARCH and Heston–Nandi GARCH provide interpretable benchmarks for persistence and asymmetry. EWMA predictors provide adaptive smoothing rules closely related to industry risk management practice (J.P. Morgan/Reuters 1996). Kernel regression provides a flexible alternative that does not impose a constant volatility-response coefficient. The comparison is therefore

not merely between complicated and simple models. It is a comparison between global parametric structure, adaptive exponential smoothing, and local nonparametric learning.

Third, the paper gives the kernel estimator an economic interpretation rather than treating it as a black-box forecasting device. The estimated kernel surface shows how predicted ETH realised variance changes across the transformed state space. High predicted volatility appears in regions where recent absolute returns are large and recent signed returns are negative, consistent with a downside-stress interpretation. Calm states generate lower predicted variance. Upside-turbulence states are distinct from downside-stress states. This surface-based evidence links forecast performance to economic mechanisms: volatility clustering, sign dependence, and nonlinear regime variation. In this sense, the kernel estimator does not merely improve or fail to improve a loss function; it reveals how the Ethereum volatility process changes across market states.

The paper also extends the forecasting exercise to conditional quantile estimation. Volatility forecasts are useful, but risk managers often require tail-risk forecasts such as value-at-risk. Following the quantile and interval-forecasting literature (Koenker and Bassett 1978, Kupiec 1995, Christoffersen 1998, McNeil and Frey 2000, Engle and Manganelli 2004), the paper constructs lower-tail return quantiles by combining conditional volatility forecasts with residual quantile inversions. This step is especially relevant for ETH because daily returns are heavy-tailed and large losses are not well summarised by conditional variance alone. The comparison between Student-$t$ and empirical residual quantiles provides additional evidence on whether distributional assumptions matter for Ethereum tail-risk measurement.

The empirical findings support a nuanced conclusion. Kernel regression is most valuable when the forecast horizon is short and the current state contains strong information about near-term volatility. In the one-day-ahead exercise, the kernel predictor using the GARCH volatility state performs especially well relative to the standard GARCH benchmark and other EWMA-type competitors. This suggests that combining an economically meaningful parametric volatility state with a flexible nonlinear forecasting map is useful for post-Merge ETH. At longer horizons, however, the advantage of local adaptivity weakens. The ten-day evidence indicates that mean reversion embedded in the standard GARCH forecast can be valuable when recent volatility shocks are sharp but transitory. Therefore, the paper does not claim that nonparametric methods dominate uniformly. Rather, it shows that the forecasting value of nonlinear local information is horizon dependent.

The rest of the paper proceeds as follows. Section 2 presents the empirical forecasting framework, including EWMA volatility with asymmetry, GARCH and Heston–Nandi GARCH benchmarks, kernel regression, and the forecast-evaluation criterion. Section 3 describes the ETH data, constructs and transforms the predictors, reports the parametric estimates, evaluates prediction performance across horizons, interprets the estimated kernel surface, and studies conditional quantile forecasts. Section 4 concludes by discussing the implications for Ethereum volatility modelling, risk management, and the use of flexible nonparametric methods in digital-asset markets.

## 2 Empirical Framework

Let $P_t$ denote the closing price of ETH on day $t$, and let $R_t = P_t/P_{t-1} - 1$ denote the daily net return. Following Klemelä (2017a) we take squared returns $R_t^2$ as the target of volatility prediction.

## 2.1 EWMA volatility with asymmetry

Classical RiskMetrics EWMA is a special case of the asymmetric recursion (Engle and Ng 1993a):

$$\hat{f}_\delta(t) = (1-\delta)\,\hat{f}_\delta(t-1)\left(\frac{R_t}{\sqrt{\hat{f}_\delta(t-1)}} - \lambda\right)^2 + \delta\,\hat{f}_\delta(t-1), \tag{1}$$

with smoothing parameter $\delta \in [0,1]$ and asymmetry parameter $\lambda \in \mathbb{R}$; $\lambda = 0$ recovers the standard EWMA. Given a finite grid $\delta_1, \ldots, \delta_S$, a cross-validated (CV) predictor selects, at each $t$, the $\delta$ minimising the cumulative squared one-step error, while the aggregated predictor of Klemelä (2017a) forms the convex combination:

$$\hat{f}(t) = \sum_{s=1}^{S} q_{t,s}\,\hat{f}_{\delta_s}(t), \qquad q_{t,s} = \frac{\exp\{-\mathrm{CSSPE}_t(\delta_s)\}}{\sum_{\nu=1}^{S}\exp\{-\mathrm{CSSPE}_t(\delta_\nu)\}}. \tag{2}$$

## 2.2 GARCH(1,1) and Heston–Nandi GARCH(1,1)

The standard GARCH(1,1) process of Bollerslev (1986a) is:

$$\sigma_t^2 = \beta_0 + \beta_1 R_{t-1}^2 + \alpha\,\sigma_{t-1}^2, \qquad R_t = \sigma_t \varepsilon_t,\ \varepsilon_t \overset{\text{iid}}{\sim} (0,1). \tag{3}$$

The $\eta$-step-ahead forecast of $R_{t+\eta}^2$ admits the closed form

$$\mathbb{E}[R_{t+\eta}^2 \mid \mathcal{F}_t] = \bar{\sigma}^2 + (\beta_1+\alpha)^{\eta-1}(\sigma_{t+1}^2 - \bar{\sigma}^2), \qquad \bar{\sigma}^2 = \frac{\beta_0}{1-\beta_1-\alpha}. \tag{4}$$

The Heston–Nandi (HN) GARCH(1,1) model of Heston and Nandi (2000a) introduces a skewness parameter $\lambda$ that governs the asymmetric response of conditional variance to signed shocks:

$$\sigma_t^2 = \beta_0 + \beta_1 \frac{(R_{t-1} - \lambda\sigma_{t-1}^2)^2}{\sigma_{t-1}^2} + \alpha\,\sigma_{t-1}^2. \tag{5}$$

A positive $\lambda$ amplifies the effect of negative shocks on future volatility (classical leverage), while a negative $\lambda$ amplifies the effect of positive shocks (anti-leverage). Both parameter sets are estimated by conditional Gaussian maximum likelihood.

## 2.3 Nadaraya–Watson kernel regression

Define transformed predictors $Z_t = (Z_{1t}, Z_{2t})$ by:

$$Z_{1t} = \sqrt{\sum_{i=1}^{t} \omega_i^{(1)}(t)\,R_i^2}, \qquad Z_{2t} = \sum_{i=1}^{t} \omega_i^{(2)}(t)\,R_i, \qquad \omega_i^{(\nu)}(t) \propto G\left(\frac{t-i}{g_\nu}\right), \tag{6}$$

where $G$ is a smoothing kernel and $g_\nu$ a bandwidth. We then apply a rank-based transformation so that $X_t = (X_{1t}, X_{2t})$ has approximately standard-normal marginals; this restores spatial homogeneity, as is visible in Figure 3. The Nadaraya–Watson kernel regression predictor of $R_{t+\eta}^2$ is:

$$\hat{f}_h(t,\eta) = \sum_{i=k}^{t-\eta} w_i(t)\,R_{i+\eta}^2, \qquad w_i(t) = \frac{K_h(X_t - X_i)}{\sum_{j=k}^{t-\eta} K_h(X_t - X_j)}, \tag{7}$$

with Gaussian product kernel $K_h$. The bandwidth $h$ is chosen either by minimising an expanding-window squared-error criterion or via the normal-reference rule $h_t = \big(4/(d+2)\big)^{1/(d+4)} t^{-1/(d+4)}$ with $d = 2$.

### 2.4 Cumulative squared prediction errors and quantile estimation

Following Klemelä (2017a), Şanlı et al. (2025) we compare predictors $\hat{f}_1$ and $\hat{f}_2$ via the differenced cumulative sum of squared prediction errors

$$D_t = \mathrm{CSSPE}_t(\hat{f}_1) - \mathrm{CSSPE}_t(\hat{f}_2), \qquad \mathrm{CSSPE}_t(\hat{f}) = \sum_{i=t_0}^{t-\eta} (R^2_{i+\eta} - \hat{f}(i))^2. \tag{8}$$

A downward-sloping $D_t$ indicates $\hat{f}_1$ is outperforming $\hat{f}_2$; the advantage of this running comparison is that outlier-driven events change the slope of $D_t$ only locally, whereas conventional out-of-sample splits can inherit a single extreme observation into their aggregate ranking.

For conditional quantile (VaR) estimation, let $q_t(\tau) = F_t^{-1}(\tau)$ be the $\tau$-quantile of the conditional return distribution. Given a volatility forecast $\hat{\sigma}_t$ we write $\hat{q}_t = \hat{\mu}_t + \hat{\sigma}_t\, \hat{q}(\varepsilon)$ where $\hat{q}(\varepsilon)$ is either a Student-$t$ inversion or an empirical quantile of past standardised residuals. Predictor comparisons use the quantile loss $\rho_\tau(u) = u\,(\tau - \mathbf{1}\{u < 0\})$ and the associated cumulative sum difference.

## 3 Data and Application

### 3.1 Data

Our sample consists of daily ETH/USD reference rates from the Coin Metrics Community Data Archive. We use the series from 15 September 2022 (the Merge) through 23 April 2026, corresponding to $T = 1{,}316$ daily net-return observations after differencing. Table 1 reports summary statistics. The return distribution is leptokurtic (excess kurtosis 4.71) and mildly right-skewed (+0.42). The price path (Figure 1) covers three distinct regimes: a low-volatility recovery from mid-2023 into mid-2024, a sharp repricing around the ETF approval of May 2024, a renewed bull run culminating in the August 2025 all-time high near USD 4,800, and a drawdown through early 2026. The net-return path (Figure 2) exhibits the usual volatility clustering, with the largest drawdown days concentrated around the collapse of FTX (late 2022), the March 2023 regional-banking shock, and the liquidation cascades that followed the 2024 ETF approval and the 2025 peak.

Table 1: Descriptive statistics: ETH daily net returns, 15 September 2022 – 23 April 2026.

| **Statistic** | **Value** | **Statistic** | **Value** |
|---|---|---|---|
| Observations ($T$) | 1,316 | Min | −0.1793 |
| Mean | 0.00095 | Max | +0.2086 |
| Standard deviation | 0.03432 | Skewness | +0.4167 |
| Price min (USD) | 1,093 | Excess kurtosis | 4.7066 |
| Price max (USD) | 4,831 | | |

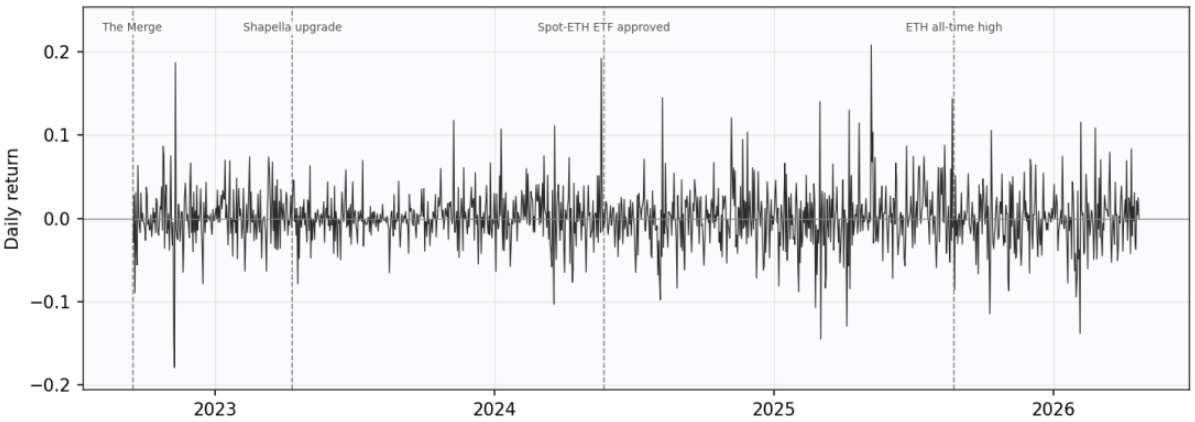
### 3.2 Predictor construction and transformation

The kernel-regression exercise requires a low-dimensional state vector that summarises the information in the recent return history without imposing a fully parametric volatility law. Following the construction

Figure 1: Ethereum daily USD price, 15 September 2022 – 23 April 2026


Note: Dashed lines mark the Merge, the Shapella upgrade, spot-ETH-ETF approval, and the August 2025 all-time high.

Figure 2: Ethereum daily net returns, 16 September 2022 – 23 April 2026




Note: The figure plots daily ETH/USD net returns, $R_t = P_t/P_{t-1} - 1$, from 16 September 2022 to 23 April 2026. The vertical dashed lines mark the Merge, the Shapella upgrade, the spot-ETH-ETF approval, and the August 2025 ETH all-time high. The return series displays pronounced volatility clustering, with large positive and negative return observations concentrated around major market events and stress episodes.

in subsection 2.3, we therefore build a two-dimensional predictor vector

$$Z_t = (Z_{1t}, Z_{2t}),$$

where the first component captures the recent magnitude of volatility and the second component captures the recent direction of signed returns. This distinction is important in the Ethereum market because large movements in ETH prices are not only clustered in magnitude, but are also often associated with directional episodes, such as liquidation cascades, ETF-related repricing, and sharp post-rally corrections. A purely magnitude-based predictor may capture volatility persistence, but it may miss the asymmetric information contained in sequences of positive or negative returns.

The first predictor is a smoothed measure of recent squared returns,

$$Z_{1t} = \left( \sum_{i=1}^{t} \omega_i^{(1)}(t) R_i^2 \right)^{1/2}, \tag{9}$$

where $R_i^2$ is the squared daily ETH return and $\omega_i^{(1)}(t)$ is a time-decaying kernel weight. The square-root transformation places $Z_{1t}$ on the same scale as a volatility measure rather than a variance measure, which improves interpretability and reduces the influence of extremely large squared-return observations. Economically, $Z_{1t}$ can be interpreted as a nonparametric volatility state variable: it is high when recent ETH returns have been large in absolute value and low when the market has been relatively calm.

The second predictor is a smoothed measure of recent signed returns,

$$Z_{2t} = \sum_{i=1}^{t} \omega_i^{(2)}(t) R_i. \tag{10}$$

This variable is included to allow the conditional volatility forecast to depend on the recent direction of market movements. In equity markets, negative returns are often followed by higher volatility because of leverage and risk-premium effects. In cryptocurrency markets, however, the sign effect may be more ambiguous. Positive return runs can also be followed by higher volatility because speculative inflows, momentum trading, and leveraged long positions increase the probability of abrupt reversals. The inclusion of $Z_{2t}$ therefore allows the kernel estimator to learn either leverage-type or anti-leverage-type patterns directly from the data, rather than imposing one asymmetry channel ex ante.

For both predictors, the weights are defined by:

$$\omega_i^{(\nu)}(t) = \frac{G\left(\frac{t-i}{g_\nu}\right)}{\sum_{j=1}^{t} G\left(\frac{t-j}{g_\nu}\right)}, \qquad \nu = 1, 2, \tag{11}$$

where $G(\cdot)$ is a smoothing kernel and $g_\nu$ is the temporal bandwidth. The normalisation in Equation (11) ensures that the weights are non-negative and sum to one at each date $t$. Hence both predictors are constructed using only information available up to date $t$, which avoids look-ahead bias in the forecasting exercise. In the baseline specification, we set

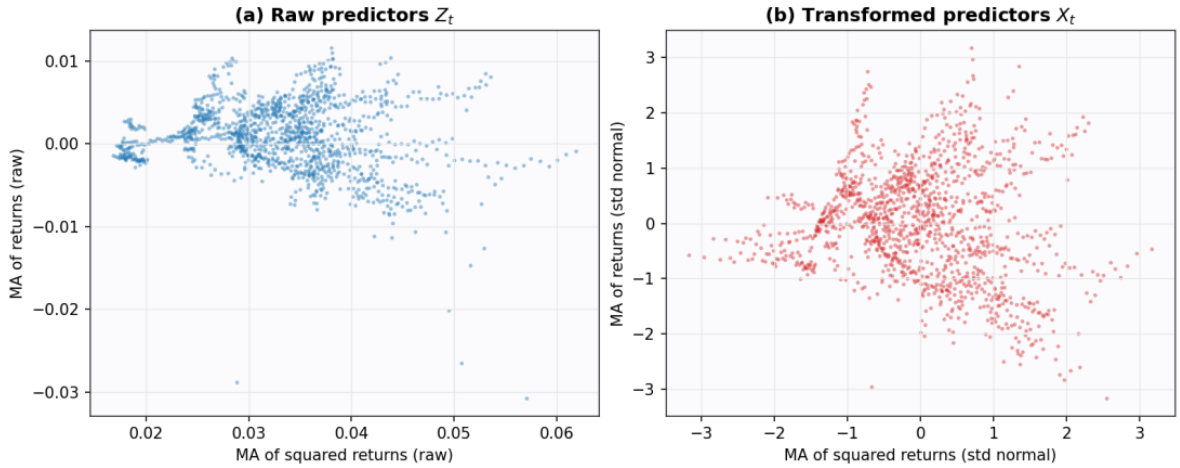
$$g_1 = g_2 = 40.$$

This choice gives the predictor vector a medium-frequency memory of approximately two trading months. The bandwidth is long enough to smooth out isolated daily jumps, but short enough to respond to volatility regime changes in the post-Merge ETH market. A much smaller bandwidth would make the

predictors excessively sensitive to one-day return outliers, whereas a much larger bandwidth would blur the transition between calm and turbulent market states.

The raw predictor cloud is shown in Panel (a) of Figure 3. Several features are worth noting. First, the horizontal axis, corresponding to $Z_{1t}$, is strictly positive and concentrated around the middle of the empirical volatility distribution. Second, the vertical axis, corresponding to $Z_{2t}$, is centred close to zero but displays a visibly asymmetric lower tail. Third, the raw cloud is highly uneven: many observations are concentrated in a dense central region, while relatively few observations occupy the high-volatility tails. This uneven geometry creates a practical difficulty for kernel regression with a global bandwidth. If the bandwidth is chosen to fit the dense centre of the distribution, the estimator may become too local in the tails. Conversely, if the bandwidth is enlarged to accommodate sparse tail observations, the estimator may oversmooth the central region where most forecasts are made.

Figure 3: Explanatory variables used by the kernel regression




Note: Panel (a) shows the raw predictors $Z_t = (Z_{1t}, Z_{2t})$; panel (b) shows the rank-transformed predictors $X_t = (X_{1t}, X_{2t})$ whose marginals are approximately standard normal.

To address this problem, we apply a rank-based marginal transformation to the two raw predictors. Let

$$\widehat{F}_{\nu,t}(z)$$

denote the empirical distribution function of the $\nu$th predictor, computed from the available sample of past predictor observations. The transformed predictor is defined as

$$X_{\nu t} = \Phi^{-1}\left[\frac{\text{rank}(Z_{\nu t}) - \frac{1}{2}}{N_t}\right], \qquad \nu = 1, 2, \tag{12}$$

where $\Phi^{-1}(\cdot)$ is the inverse standard-normal distribution function, $\text{rank}(Z_{\nu t})$ is the empirical rank of $Z_{\nu t}$ among the relevant observations, and $N_t$ is the number of observations used in the empirical transformation. The small continuity correction $1/2$ prevents the transformed values from taking infinite values at the empirical boundaries. This transformation preserves the ordinal information in each predictor while mapping the marginal distributions approximately into standard-normal variables.

The transformed state vector is therefore

$$X_t = (X_{1t}, X_{2t}), \tag{13}$$

where $X_{1t}$ is the normal-score transformation of the smoothed volatility predictor and $X_{2t}$ is the normal-

score transformation of the smoothed signed-return predictor. The transformation is monotonic in each margin, so it does not change the ranking of observations within each predictor. Instead, it changes the metric of the predictor space. Observations that are compressed in the raw scale become more evenly spread in the transformed scale, while extreme observations are moved toward the tails of the standard-normal domain.

Panel (b) of Figure 3 shows the transformed predictor cloud. Compared with Panel (a), the transformed observations fill the bivariate predictor space much more evenly. This is useful because the Nadaraya–Watson estimator relies on distances between the current state $X_t$ and historical states $X_i$. A global Gaussian product kernel is more appropriate when the predictor space is approximately homogeneous. Without the rank transformation, the same numerical bandwidth would imply very different degrees of smoothing in dense and sparse regions of the raw predictor space. After the transformation, the distance metric becomes more balanced, and the bandwidth parameter has a more stable interpretation across market states.

The final predictor construction can therefore be summarised in three steps. First, we compute daily ETH net returns from the price series and construct two smoothed historical variables: a volatility-level predictor based on squared returns and a signed-return predictor based on raw returns. Second, we transform each raw predictor into a normal score using its empirical rank. Third, we use the transformed vector $X_t = (X_{1t}, X_{2t})$ as the state variable in the kernel-regression forecast of future squared returns. This procedure produces a compact, nonparametric representation of the ETH volatility state while allowing the subsequent forecasting model to capture nonlinearities, volatility clustering, and sign-dependent effects without imposing a restrictive parametric GARCH structure.

### 3.3 GARCH(1,1) and Heston–Nandi GARCH(1,1) estimates

This subsection reports the parametric volatility estimates used as benchmark models for the nonparametric kernel-regression forecasts. The purpose of these estimates is twofold. First, the GARCH-type models provide a conventional reference point for evaluating the magnitude and persistence of Ethereum volatility in the post-Merge period. Second, the Heston–Nandi specification provides a parametric benchmark for assessing whether the response of conditional volatility differs between positive and negative ETH return shocks. This is useful because the predictor construction in subsection 3.2 explicitly allows the kernel-regression forecast to depend on both recent volatility magnitude and the recent direction of signed returns.

Let $h_t = \sigma_t^2$. The standard GARCH(1,1) benchmark is estimated as:

$$h_t = \beta_0 + \beta_1 R_{t-1}^2 + \alpha h_{t-1}, \tag{14}$$

where $\beta_0 > 0$, $\beta_1 \geq 0$, $\alpha \geq 0$, and $\alpha + \beta_1 < 1$. In this specification, $\beta_1$ measures the short-run response of conditional variance to new squared-return innovations, while $\alpha$ measures the persistence of the lagged conditional variance. Therefore, the GARCH volatility-persistence coefficient is:

$$\rho^{GARCH} = \alpha + \beta_1. \tag{15}$$

This clarification is important because $\alpha$ alone is not the total persistence of the GARCH recursion. A value of $\rho^{GARCH}$ close to one would imply near-integrated volatility dynamics, whereas a lower value indicates faster mean reversion of daily conditional volatility.

The Heston–Nandi GARCH(1,1) model is estimated as:

$$h_t = \beta_0 + \beta_1 \frac{(R_{t-1} - \lambda h_{t-1})^2}{h_{t-1}} + \alpha h_{t-1}. \tag{16}$$

The additional parameter $\lambda$ governs the asymmetric response of future volatility to signed shocks. Under the normalisation in Equation (16), the sign of $\lambda$ has a precise interpretation. For two shocks of equal absolute size, $R_{t-1} = r$ and $R_{t-1} = -r$, where $r > 0$, the difference between the Heston–Nandi innovation terms is:

$$\frac{(-r - \lambda h_{t-1})^2}{h_{t-1}} - \frac{(r - \lambda h_{t-1})^2}{h_{t-1}} = 4\lambda r. \tag{17}$$

Hence, a positive $\lambda$ implies that negative shocks raise future volatility more than positive shocks of the same magnitude. This corresponds to the classical leverage effect. Conversely, a negative $\lambda$ implies that positive shocks raise future volatility more than negative shocks of the same magnitude, which can be interpreted as an anti-leverage or speculative-upside volatility effect.

Taking expectations in Equation (16) also shows that the relevant persistence coefficient for the Heston–Nandi recursion is not $\alpha$ alone. Since

$$\mathbb{E}\left[\frac{(R_{t-1} - \lambda h_{t-1})^2}{h_{t-1}} \,\middle|\, \mathcal{F}_{t-2}\right] = 1 + \lambda^2 h_{t-1},$$

the implied persistence of the conditional-variance recursion is:

$$\rho^{HN} = \alpha + \beta_1 \lambda^2. \tag{18}$$

Thus, in the rolling analysis below, we compare $\alpha + \beta_1$ for the standard GARCH model with $\alpha + \beta_1 \lambda^2$ for the Heston–Nandi model.

Both models are estimated by conditional Gaussian maximum likelihood. Ignoring constants, the log-likelihood contribution is:

$$\ell_t(\theta) = -\frac{1}{2}\left[\log h_t(\theta) + \frac{R_t^2}{h_t(\theta)}\right], \tag{19}$$

where $\theta$ denotes the relevant parameter vector. The Gaussian likelihood is interpreted as a quasi-likelihood criterion rather than as a literal assumption that ETH returns are conditionally Gaussian. This distinction is important because Table 1 shows that ETH daily returns are strongly leptokurtic. The estimates should therefore be read as fitted volatility dynamics, not as evidence that the full return distribution is Gaussian.

Table 2 reports the full-sample maximum-likelihood estimates. For the standard GARCH(1,1) model, the estimated persistence is:

$$\widehat{\rho}^{GARCH} = \widehat{\alpha} + \widehat{\beta}_1 = 0.750 + 0.0667 = 0.8167. \tag{20}$$

This value is clearly below the near-unit persistence often found in equity-index volatility models, suggesting that post-Merge ETH volatility is persistent but not close to integrated at the daily frequency. Using the GARCH estimates, the implied unconditional variance is:

$$\widehat{\bar{h}}^{GARCH} = \frac{\widehat{\beta}_0}{1 - \widehat{\alpha} - \widehat{\beta}_1} = \frac{2.07 \times 10^{-4}}{1 - 0.750 - 0.0667} \approx 1.13 \times 10^{-3}, \tag{21}$$

which corresponds to an unconditional daily volatility of approximately 3.36%. This is close to the sample standard deviation reported in Table 1, indicating that the full-sample GARCH model captures

Table 2: Full-sample GARCH(1,1) and Heston–Nandi GARCH(1,1) estimates

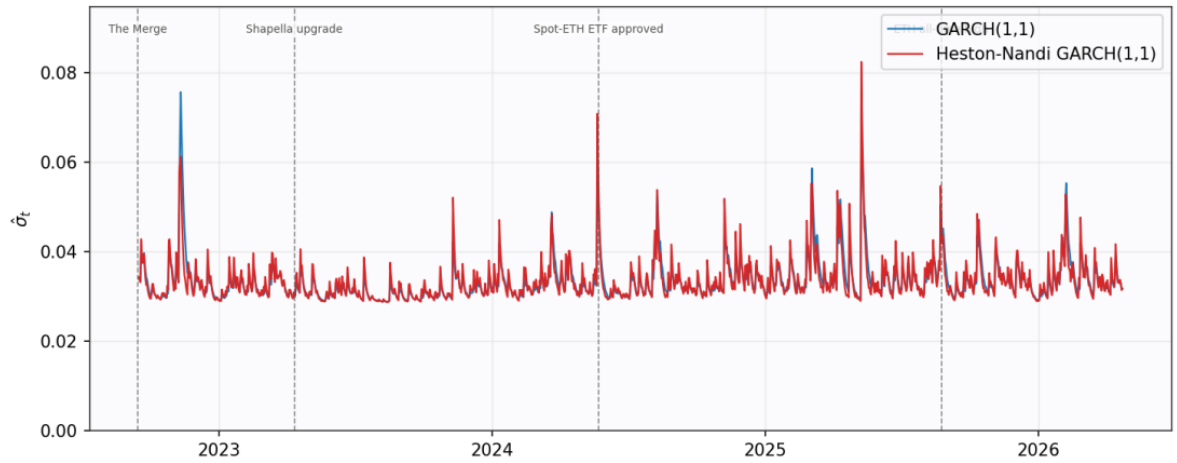
| Model | $\beta_0$ | $\beta_1$ | $\alpha$ | $\lambda$ |
|---|---|---|---|---|
| GARCH(1,1) | $2.07 \times 10^{-4}$ | 0.0667 | 0.750 | — |
| Heston–Nandi GARCH(1,1) | $2.85 \times 10^{-4}$ | $1.15 \times 10^{-4}$ | 0.653 | −0.312 |

the average volatility level of ETH reasonably well.

The full-sample Heston–Nandi estimate of $\lambda$ is negative, $\widehat{\lambda} = -0.312$. Under the sign convention in Equation (16), this should not be interpreted as classical leverage. Instead, the full-sample estimate points to a weak anti-leverage pattern, meaning that positive ETH return shocks raise subsequent conditional variance slightly more than negative shocks of the same absolute magnitude. This result is economically plausible in a cryptoasset market because volatility can increase during speculative upward repricing episodes as well as during market crashes. In particular, positive return runs may attract leveraged long positions, momentum trading, and retail inflows, all of which can make the market more fragile and raise subsequent volatility.

Figure 4 compares the full-sample fitted conditional-volatility paths from the standard GARCH and Heston–Nandi models. The two paths track each other closely over most of the sample, suggesting that the dominant empirical feature is volatility clustering rather than persistent model-specific disagreement. The main differences occur around the largest volatility spikes, including the FTX-related turbulence, the spot-ETH-ETF repricing episode, and the high-volatility period around the 2025 peak. These are precisely the episodes in which a model with signed-shock asymmetry may deviate from a symmetric GARCH benchmark. Nevertheless, the close co-movement of the two fitted paths suggests that the asymmetric channel is secondary relative to the common volatility-persistence component.

Figure 4: Full-sample conditional volatility: GARCH(1,1) vs. Heston–Nandi GARCH(1,1)

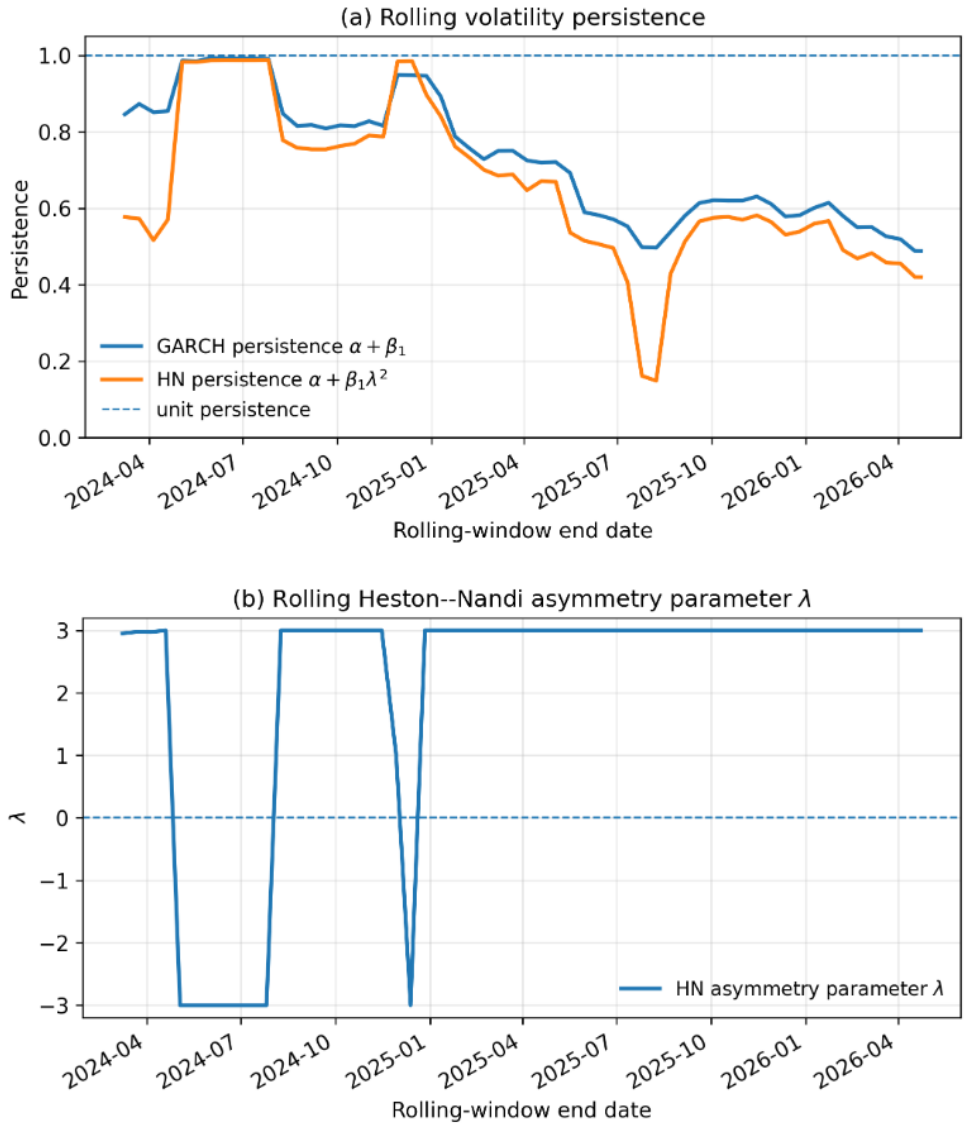



Note: The figure compares fitted full-sample conditional volatility paths from the standard GARCH(1,1) and Heston–Nandi GARCH(1,1) models. Volatility is reported as $\widehat{\sigma}_t = \sqrt{\widehat{h}_t}$. The two paths are estimated using daily ETH net returns over the post-Merge sample. The vertical dashed lines mark the Merge, the Shapella upgrade, the spot-ETH-ETF approval, and the August 2025 ETH all-time high.

Figure 5 reports rolling 540-day estimates, refitted every two weeks. Panel (a) plots the appropriate persistence measures: $\alpha + \beta_1$ for GARCH(1,1) and $\alpha + \beta_1\lambda^2$ for Heston–Nandi GARCH(1,1). The rolling estimates reveal substantial time variation in ETH volatility persistence. GARCH persistence is high in the first half of the rolling sample, rising close to unity during parts of 2024. It then declines markedly after the 2025 market peak and remains well below its earlier level through early 2026. This

pattern suggests that volatility shocks became less persistent in the later post-Merge period. One possible interpretation is that, as the ETH market matured and institutional participation increased after the ETF-related repricing period, shocks were absorbed more quickly. Another interpretation is that the market shifted from a prolonged trend-repricing regime into a more rapidly mean-reverting drawdown regime.

Figure 5: Rolling 540-day parameter estimates




Note: Panel (a) reports the rolling persistence measures. For the standard GARCH(1,1) model, persistence is defined as $\alpha + \beta_1$. For the Heston–Nandi GARCH(1,1) model under Equation (16), the implied conditional-variance persistence is $\alpha + \beta_1\lambda^2$. Panel (b) reports the Heston–Nandi asymmetry parameter $\lambda$. Under the normalisation in Equation (16), positive $\lambda$ corresponds to classical leverage, while negative $\lambda$ corresponds to anti-leverage. Estimates close to the imposed bounds should be interpreted cautiously.

The Heston–Nandi persistence path is broadly similar but more volatile. In some windows, the Heston–Nandi persistence is close to the GARCH persistence, indicating that the asymmetric specification does not substantially change the implied mean reversion of volatility. In other windows, especially around the 2025 peak and subsequent correction, the Heston–Nandi persistence falls more sharply. This suggests that allowing for signed-shock asymmetry changes the estimated persistence of the volatility process in episodes dominated by large directional price movements.

Panel (b) reports the rolling Heston–Nandi asymmetry parameter $\lambda$. Unlike the full-sample estimate, the rolling estimates are not uniformly negative. Instead, $\lambda$ switches sign across windows and is frequently estimated near the imposed positive boundary. Under the sign convention in Equation (16), positive values of $\lambda$ correspond to classical leverage, while negative values correspond to anti-leverage. The rolling evidence therefore suggests that the asymmetry mechanism is not stable over time. ETH volatility appears to alternate between regimes in which downside shocks are more volatility-amplifying and regimes in which upside shocks are more volatility-amplifying.

The boundary behaviour of $\lambda$ should be interpreted cautiously. In rolling windows, the asymmetry parameter is identified mainly by a limited number of large positive and negative return observations. When a window contains few extreme observations, or when volatility clustering can be explained almost equally well by persistence and asymmetry, the likelihood may push $\lambda$ toward the boundary. For this reason, the rolling $\lambda$ estimates are best viewed as evidence of regime-dependent asymmetry rather than precise structural estimates of a stable leverage parameter.

Overall, the parametric estimates deliver three implications for the forecasting exercise. First, ETH volatility is persistent but not uniformly near-integrated; the corrected rolling estimates show that persistence changes substantially across post-Merge market regimes. Second, the Heston–Nandi model does not materially alter the average fitted volatility path relative to GARCH, but it does change the estimated persistence and shock response during large directional episodes. Third, the rolling asymmetry parameter is unstable and often boundary-sensitive, suggesting that a fixed parametric asymmetry channel may be too restrictive for Ethereum volatility. This provides additional motivation for the kernel-regression approach: by using the transformed state vector $X_t = (X_{1t}, X_{2t})$, the nonparametric model allows future squared returns to depend flexibly on both recent volatility magnitude and recent signed-return conditions, without imposing a constant leverage or anti-leverage parameter over the entire sample.

### 3.4 Prediction horizon $\eta = 1$

We first evaluate the forecasting performance of all competing predictors at the one-day-ahead horizon, $\eta = 1$. This is the most demanding horizon in the empirical exercise because the forecast targets the next daily realised squared return, $R^2_{t+1}$, using only information available at date $t$. For each predictor $j$, define the one-step-ahead prediction error as:

$$e_{j,t+1} = R^2_{t+1} - \widehat{f}_j(t, 1), \tag{22}$$

where $\widehat{f}_j(t, 1)$ denotes the predictor-specific forecast of $R^2_{t+1}$. The cumulative squared prediction error is therefore

$$\text{CSSPE}^{(1)}_{j,t} = \sum_{i=t_0}^{t-1} \left( R^2_{i+1} - \widehat{f}_j(i, 1) \right)^2, \tag{23}$$

where $t_0$ is the first evaluation date. Consistent with subsection 2.4, the graphical comparison is based on the differenced cumulative squared prediction error relative to the standard GARCH(1,1) benchmark:

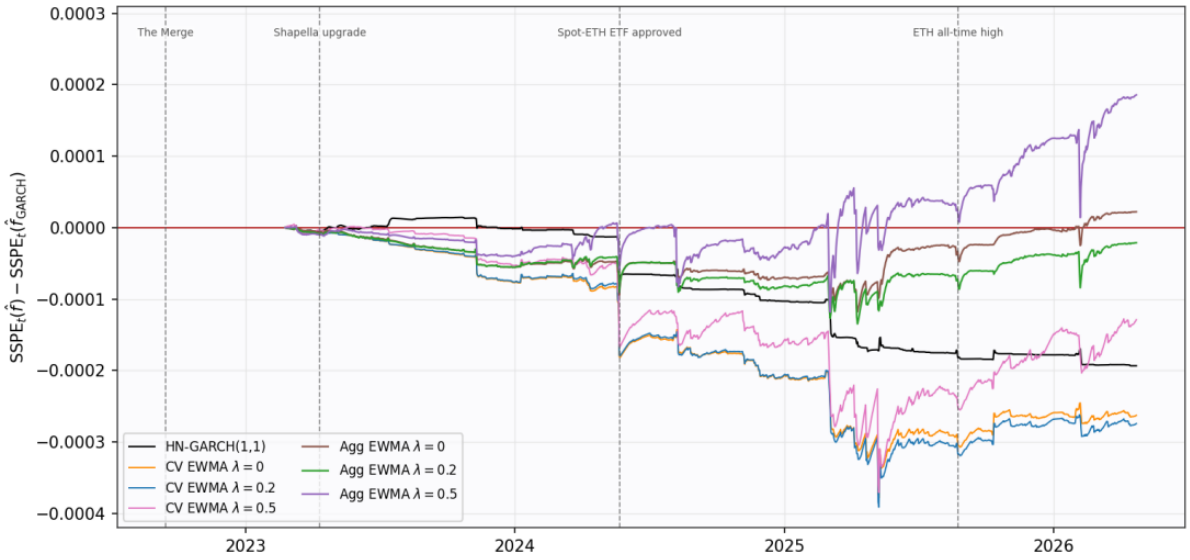
$$D^{(1)}_{j,t} = \text{CSSPE}^{(1)}_{j,t} - \text{CSSPE}^{(1)}_{GARCH,t}\,. \tag{24}$$

Hence $D^{(1)}_{j,t} < 0$ means that predictor $j$ has accumulated a smaller squared prediction error than the GARCH(1,1) benchmark up to date $t$. A downward movement in $D^{(1)}_{j,t}$ indicates that the predictor is gaining forecasting accuracy relative to GARCH, while an upward movement indicates that the predictor is losing ground.

Figure 6 plots the differenced CSSPE paths for the seven non-kernel benchmark predictors: Heston–

Nandi GARCH, three cross-validated EWMA predictors, and three aggregated EWMA predictors. The figure shows that most competing predictors improve on the GARCH(1,1) benchmark for a substantial part of the sample. The improvement becomes particularly visible after the spot-ETH-ETF repricing episode, when volatility dynamics become more difficult for a single fixed GARCH recursion to capture. The cross-validated EWMA predictors with $\lambda = 0$ and $\lambda = 0.2$ perform especially well among the EWMA-type forecasts, suggesting that adaptive smoothing is valuable at the daily horizon.

Figure 6: Differenced CSSPE for the prediction horizon $\eta = 1$




Note: The figure plots $\text{CSSPE}^{(1)}_{j,t} - \text{CSSPE}^{(1)}_{GARCH,t}$. Negative values indicate that predictor $j$ outperforms the standard GARCH(1,1) benchmark. The vertical dashed lines mark the Merge, the Shapella upgrade, the spot-ETH-ETF approval, and the August 2025 ETH all-time high.

The ranking is clearer in Table 3, which reports final-date CSSPE values and the percentage difference relative to the GARCH(1,1) benchmark:

$$\Delta^{(1)}_j = 100 \left( \frac{\text{CSSPE}^{(1)}_{j,T}}{\text{CSSPE}^{(1)}_{GARCH,T}} - 1 \right). \tag{25}$$

Negative values of $\Delta^{(1)}_j$ therefore indicate an improvement over GARCH.

Table 3: Final cumulative squared prediction errors at $\eta = 1$ and relative difference against the GARCH(1,1) benchmark

| Predictor | CSSPE (final) | % vs. GARCH |
|---|---|---|
| GARCH(1,1) | 0.009586 | 0.00% |
| HN-GARCH(1,1) | 0.009392 | −2.02% |
| CV-EWMA $\lambda = 0$ | 0.009323 | −2.74% |
| CV-EWMA $\lambda = 0.2$ | 0.009312 | −2.86% |
| CV-EWMA $\lambda = 0.5$ | 0.009457 | −1.34% |
| Agg-EWMA $\lambda = 0$ | 0.009608 | +0.23% |
| Agg-EWMA $\lambda = 0.2$ | 0.009565 | −0.22% |
| Agg-EWMA $\lambda = 0.5$ | 0.009772 | +1.94% |
| Kernel (EWMA vol) | 0.009451 | −1.40% |
| Kernel (GARCH vol) | 0.009261 | −3.39% |

Several points follow from the table. First, the standard GARCH(1,1) benchmark is not dominated

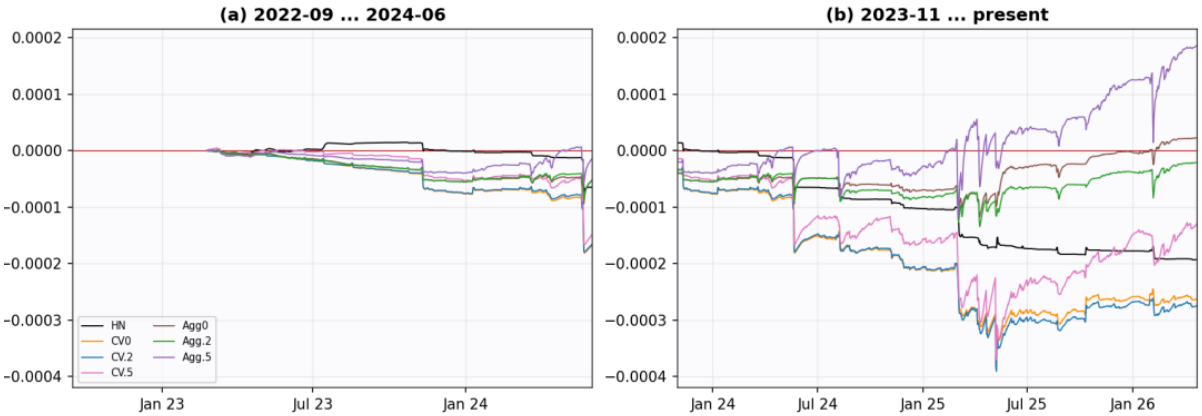
by all alternatives, but it is outperformed by most of them. The Heston–Nandi GARCH model reduces the final CSSPE by 2.02%, indicating that allowing for signed-shock asymmetry improves one-day volatility forecasting relative to the symmetric GARCH benchmark. This is consistent with the evidence in subsection 3.3 that ETH volatility reacts differently across directional market states.

Second, within the EWMA class, cross-validation performs better than aggregation. The two strongest EWMA predictors are the cross-validated models with $\lambda = 0.2$ and $\lambda = 0$, which reduce the final CSSPE by 2.86% and 2.74%, respectively. The difference between them is small, suggesting that mild asymmetry improves the forecast, but that most of the predictive gain comes from adaptive smoothing rather than from a large imposed asymmetry term. By contrast, the aggregated EWMA predictors are weaker. The aggregated predictor with $\lambda = 0.2$ only slightly improves on GARCH, while the aggregated predictors with $\lambda = 0$ and $\lambda = 0.5$ finish above the GARCH benchmark. This indicates that averaging over a grid of smoothing parameters can be costly when the evaluation period contains a relatively clear locally optimal smoothing choice.

Third, the best final performer is the kernel predictor using the GARCH-based volatility state. Its final CSSPE is 0.009261, corresponding to a 3.39% reduction relative to GARCH. This result is important because it shows that the information in the GARCH volatility state remains useful, but that imposing a linear parametric GARCH recursion is not the most effective way to exploit that information. The kernel estimator can use the GARCH volatility state as a predictor while allowing the mapping from the state variable to future squared returns to be nonlinear and local. In contrast, the kernel predictor based on the EWMA volatility state improves on GARCH by 1.40%, which is positive but weaker than the GARCH-state kernel. This suggests that the conditional-volatility state extracted from GARCH contains more useful one-day forecasting information than the simpler smoothed squared-return proxy.

Figure 7 splits the differenced-CSSPE paths into two overlapping sub-samples. This split is useful because the full-sample figure compresses several distinct market regimes into a single panel. Panel (a) covers the period from the Merge through the ETF-approval window. During this earlier phase, the differences across predictors are relatively small. Forecast rankings move gradually, and no single predictor accumulates a large advantage. This pattern is consistent with a market environment in which volatility is clustered but not yet dominated by the strong post-ETF repricing dynamics.

Figure 7: Differenced CSSPE for $\eta = 1$ across two overlapping sub-samples




Note: Panel (a) reports the early evaluation period from 2022-09 to 2024-06. Panel (b) reports the later evaluation period from 2023-11 to 2026-04. Negative values indicate that the predictor outperforms the standard GARCH(1,1) benchmark.

Panel (b) covers the later part of the sample, from late 2023 through April 2026. The divergence across

predictors is much clearer in this period. The cross-validated EWMA predictors and Heston–Nandi model accumulate most of their gains after the ETF-approval episode and around the high-volatility period surrounding the 2025 peak. In contrast, the aggregated EWMA predictor with $\lambda = 0.5$ deteriorates substantially and ends with the largest loss relative to GARCH. This is consistent with the view that a large fixed asymmetry correction is too rigid for ETH: the direction and strength of volatility asymmetry vary over time, as documented by the rolling Heston–Nandi estimates in Figure 5.

The one-day results also clarify how the ETH findings differ from the Bitcoin evidence in Şanlı et al. (2025). In their Bitcoin application, aggregation across smoothing parameters plays a central role. In the present ETH application, however, aggregated EWMAs do not dominate their cross-validated counterparts. A plausible explanation is that the post-Merge ETH sample contains a more sharply defined regime shift around ETF-related repricing and the subsequent 2025 peak. Once the regime becomes more stable, selecting a single locally successful smoothing parameter can outperform averaging over a grid that includes less suitable alternatives.

Overall, the evidence at $\eta = 1$ supports three conclusions. First, standard GARCH(1,1) is a useful benchmark, but it is not the best one-day predictor of ETH squared returns. Second, adaptive smoothing and mild asymmetry improve short-horizon volatility prediction, especially after the ETF-related regime shift. Third, the strongest performance comes from combining a parametric volatility state with a nonparametric forecasting map, as in the Kernel (GARCH vol) predictor. This finding provides direct empirical support for the main methodological argument of the paper: Ethereum volatility prediction benefits from retaining economically meaningful volatility states while allowing the conditional forecast function to be flexible, nonlinear, and regime-dependent.

### 3.5 Prediction horizon $\eta = 10$

We next evaluate forecasting performance at the ten-day-ahead horizon, $\eta = 10$. This horizon is conceptually different from the one-day exercise in subsection 3.4. At $\eta = 1$, the forecast targets the next realised squared return, $R^2_{t+1}$, and therefore mainly measures how well each model captures very short-run volatility clustering. At $\eta = 10$, the forecast target is instead the squared return ten days ahead, $R^2_{t+10}$. Thus, the exercise evaluates whether the current volatility state contains information about future volatility conditions approximately two calendar weeks later. Importantly, the target is not the cumulative ten-day realised variance, but the single future squared return at horizon $\eta = 10$, consistent with the forecasting object defined in subsection 2.3.

For predictor $j$, the ten-day-ahead forecast error is:

$$e_{j,t+10} = R^2_{t+10} - \widehat{f}_j(t, 10), \tag{26}$$

where $\widehat{f}_j(t, 10)$ denotes the forecast formed at date $t$ for the squared return realised at date $t + 10$. The corresponding cumulative squared prediction error is:

$$\mathrm{CSSPE}^{(10)}_{j,t} = \sum_{i=t_0}^{t-10} \left( R^2_{i+10} - \widehat{f}_j(i, 10) \right)^2 . \tag{27}$$

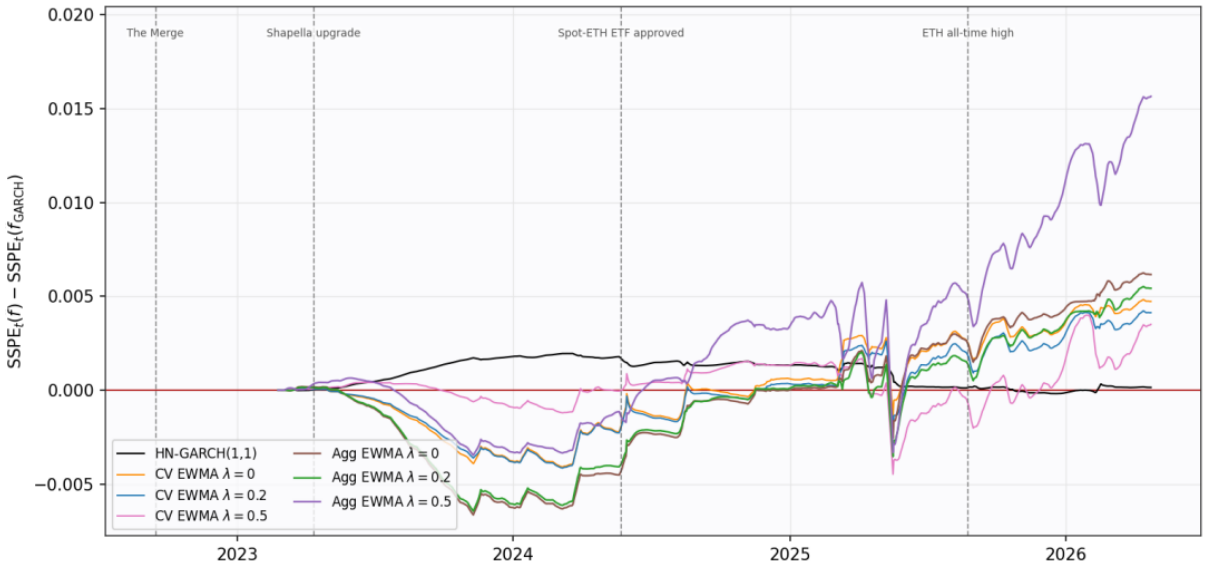
As in subsection 2.4, all predictors are compared with the standard GARCH(1,1) benchmark through the differenced cumulative squared prediction error:

$$D^{(10)}_{j,t} = \mathrm{CSSPE}^{(10)}_{j,t} - \mathrm{CSSPE}^{(10)}_{GARCH,t} . \tag{28}$$

Negative values of $D_{j,t}^{(10)}$ indicate that predictor $j$ has accumulated a smaller squared prediction error than GARCH, while positive values indicate that GARCH has performed better. Since the forecasts are produced daily but evaluated at a ten-day horizon, adjacent forecast errors are naturally overlapping. The CSSPE paths should therefore be interpreted as cumulative forecasting-performance comparisons rather than as independent daily forecast errors.

Figure 8 reports the differenced CSSPE paths for the ten-day prediction horizon. The result differs sharply from the one-day horizon. At $\eta = 1$, several EWMA-type predictors and the kernel predictor based on the GARCH volatility state improve on the GARCH benchmark. At $\eta = 10$, however, most EWMA-type predictors lose their advantage as the sample progresses. The deterioration is especially clear after the volatility disturbances around 2025, when several differenced-CSSPE paths move upward and become positive. This means that, at the ten-day horizon, the standard GARCH(1,1) benchmark is harder to beat than it is at the one-day horizon.

Figure 8: Differenced CSSPE for the prediction horizon $\eta = 10$




Note: The figure plots $\text{CSSPE}_{j,t}^{(10)} - \text{CSSPE}_{GARCH,t}^{(10)}$. Negative values indicate that predictor $j$ outperforms the standard GARCH(1,1) benchmark at the ten-day-ahead horizon. Positive values indicate that GARCH accumulates a smaller squared prediction error. The vertical dashed lines mark the Merge, the Shapella upgrade, the spot-ETH-ETF approval, and the August 2025 ETH all-time high.

The strongest underperformance is observed for the aggregated EWMA specification with $\lambda = 0.5$. This model accumulates a large positive CSSPE difference by the end of the sample, indicating that a relatively strong fixed asymmetry correction is not well suited to ten-day-ahead ETH volatility prediction. The result is consistent with the rolling Heston–Nandi evidence in Figure 5, which shows that the sign and magnitude of volatility asymmetry are unstable over time. A fixed asymmetric EWMA rule may perform well in one regime but become misaligned when the market enters a different volatility regime.

The cross-validated EWMA predictors also behave differently at $\eta = 10$ than at $\eta = 1$. At the one-day horizon, cross-validation is useful because it allows the model to choose a smoothing parameter that adapts to recent volatility conditions. At the ten-day horizon, however, this short-run adaptivity can become less valuable. A smoothing parameter that performs well for tomorrow's squared return may not provide the best signal for volatility ten days ahead. The ten-day forecast requires a model to separate temporary volatility bursts from persistent volatility states. If the cross-validated EWMA reacts too strongly to recent jumps, it may extrapolate temporary volatility spikes too far into the future.

The relative strength of GARCH at $\eta = 10$ is economically meaningful. In subsection 2.2, the multi-step GARCH forecast has the closed-form representation

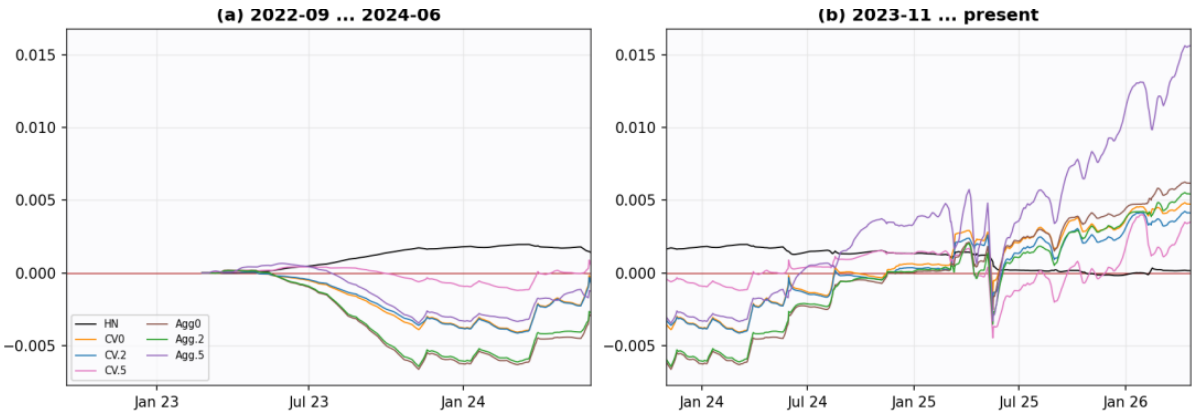
$$\mathbb{E}_t(R^2_{t+\eta}) = \bar{\sigma}^2 + (\alpha + \beta_1)^{\eta-1} \left(\sigma^2_{t+1} - \bar{\sigma}^2\right), \tag{29}$$

where $\bar{\sigma}^2$ is the unconditional variance. For $\eta = 10$, this expression automatically shrinks the one-step conditional variance forecast back toward its long-run level at a rate governed by $\alpha + \beta_1$. This mean-reversion mechanism is valuable when the underlying volatility process is persistent but not near-integrated, as documented in subsection 3.3. It prevents the model from treating short-lived volatility jumps as permanent shifts in future volatility.

This explains why the ten-day results are almost the mirror image of the one-day results. At $\eta = 1$, local adaptivity is valuable because tomorrow's volatility is strongly related to today's market state. At $\eta = 10$, excessive local adaptivity can become costly because the forecast must account for mean reversion over a longer horizon. The standard GARCH forecast benefits from its parametric multi-step structure: it uses the current volatility state but gradually discounts its influence as the horizon increases. In contrast, EWMA-type rules are more directly tied to the recent history of squared returns and may not impose enough mean reversion at longer horizons.

Figure 9 splits the ten-day CSSPE paths into two overlapping sub-samples. This decomposition is useful because the full-sample figure compresses several market regimes into a single panel. Panel (a), covering the earlier part of the evaluation period, shows that several non-GARCH predictors remain competitive before the strongest post-ETF volatility divergence. In this phase, the forecasting differences are moderate, and some EWMA variants temporarily outperform GARCH. This suggests that the poor final performance of the ten-day EWMA forecasts is not uniform across the whole sample.

Figure 9: Differenced CSSPE for $\eta = 10$ across two overlapping sub-samples




Note: Panel (a) reports the early evaluation period from 2022-09 to 2024-06. Panel (b) reports the later evaluation period from 2023-11 to 2026-04. Negative values indicate that the predictor outperforms the standard GARCH(1,1) benchmark at the ten-day-ahead horizon.

Panel (b), covering the later part of the sample, shows that the ranking changes substantially after the market enters the more turbulent 2025–2026 period. Several EWMA-type predictors accumulate positive differenced CSSPEs, while the GARCH benchmark becomes relatively more accurate. This pattern is consistent with the interpretation that the later post-Merge ETH market contains sharp but partially transitory volatility bursts. Models that react strongly to recent shocks perform well at very short horizons, but their advantage decays when the target is ten days ahead.

These results also distinguish the present ETH evidence from the Bitcoin findings of Şanlı et al.

(2025). In their Bitcoin application, aggregated EWMA predictors perform especially well at longer horizons. In the present post-Merge ETH sample, the opposite pattern emerges: aggregation does not deliver a clear long-horizon advantage, and the aggregated EWMA with stronger asymmetry performs particularly poorly. One possible explanation is that post-Merge Ethereum experienced a sequence of institutionally driven regime changes, including the ETF-related repricing episode and the 2025 peak. In such an environment, averaging over a broad grid of smoothing parameters may combine useful forecasts with poorly matched alternatives, reducing performance at longer horizons.

Overall, the ten-day evidence delivers three conclusions. First, the forecasting advantage of adaptive EWMA and asymmetric predictors is horizon-dependent: models that perform well at $\eta = 1$ do not necessarily perform well at $\eta = 10$. Second, the standard GARCH(1,1) benchmark benefits from its closed-form mean-reverting multi-step forecast, which is particularly useful when volatility shocks are persistent but ultimately transitory. Third, the underperformance of fixed asymmetric EWMA rules at $\eta = 10$ reinforces the main motivation for flexible modelling: Ethereum volatility is nonlinear and regime-dependent, but the usefulness of local information declines as the forecasting horizon lengthens.

### 3.6 Kernel regression predictors

This subsection examines the two kernel-regression predictors used in the empirical comparison. The purpose is to evaluate whether the nonparametric forecasting map developed in subsection 2.3 improves on a purely parametric volatility recursion. The kernel forecasts differ from GARCH and EWMA in one important respect: instead of imposing a fixed linear updating rule for the conditional variance, they use historical observations with similar state variables to estimate the conditional expectation of future squared returns. Hence the kernel predictor can capture nonlinearities, local regime dependence, and state-specific asymmetry in ETH volatility.

In the empirical application, we consider two kernel configurations. The first uses an EWMA-based volatility state and is denoted by Kernel (EWMA vol). The second uses a GARCH-based volatility state and is denoted by Kernel (GARCH vol). More precisely, define the raw state vector for model $s \in \{E, G\}$ as:

$$Z_t^{(s)} = \left( Z_{1t}^{(s)}, Z_{2t} \right), \tag{30}$$

where $Z_{2t}$ is the smoothed signed-return predictor defined in subsection 3.2, while $Z_{1t}^{(s)}$ is the volatility-state component. For the EWMA kernel,

$$Z_{1t}^{(E)} = \sqrt{\widehat{f}_{\delta,t}^{EWMA}}, \tag{31}$$

where $\widehat{f}_{\delta,t}^{EWMA}$ is the EWMA volatility estimate. For the GARCH kernel,

$$Z_{1t}^{(G)} = \sqrt{\widehat{h}_t^{GARCH}}, \tag{32}$$

where $\widehat{h}_t^{GARCH}$ is the fitted conditional variance from the standard GARCH(1,1) model. In both cases, the volatility component is expressed in standard-deviation units rather than variance units. This makes the scale of the first state variable comparable across the EWMA and GARCH configurations.

Each raw state vector is then transformed marginally into normal scores:

$$X_{\nu t}^{(s)} = \Phi^{-1}\left[ \frac{\operatorname{rank}(Z_{\nu t}^{(s)}) - \frac{1}{2}}{N_t} \right], \qquad \nu = 1, 2, \qquad s \in \{E, G\}. \tag{33}$$

This transformation follows the rank-based construction in subsection 3.2. It preserves the ordering of the volatility and signed-return states while improving the homogeneity of the predictor space. The two kernel predictors are therefore

$$\widehat{f}_h^E(t,\eta) = \sum_{i=k}^{t-\eta} w_i^E(t) R_{i+\eta}^2, \qquad \widehat{f}_h^G(t,\eta) = \sum_{i=k}^{t-\eta} w_i^G(t) R_{i+\eta}^2, \tag{34}$$

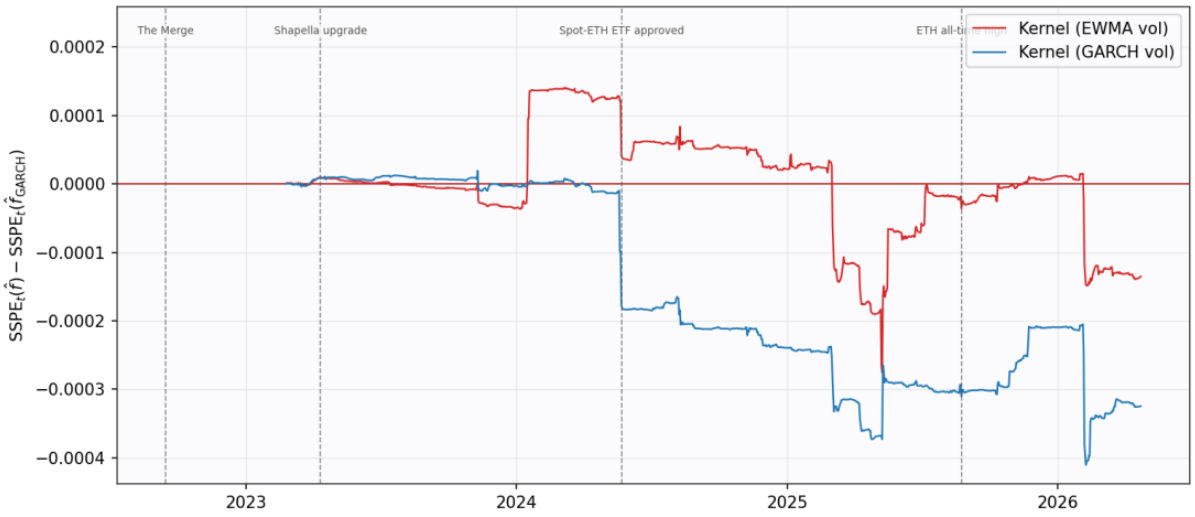
where the weights are computed using $X_t^{(E)}$ and $X_t^{(G)}$, respectively.

Figure 10 compares the two kernel-regression predictors against the standard GARCH(1,1) benchmark at the one-day horizon, $\eta = 1$. The figure plots the differenced cumulative squared prediction error,

$$D_{KR,t}^{(1)} = \mathrm{CSSPE}_{KR,t}^{(1)} - \mathrm{CSSPE}_{GARCH,t}^{(1)}\,. \tag{35}$$

Negative values indicate that the kernel predictor has accumulated a lower squared prediction error than GARCH. The figure confirms the ranking reported in Table 3: the Kernel (GARCH vol) predictor is the best final performer at $\eta = 1$, reducing the final CSSPE by 3.39% relative to the standard GARCH benchmark. The Kernel (EWMA vol) predictor also improves on GARCH at the final date, but its gain is smaller, at 1.40%.

Figure 10: Differenced CSSPE for the two kernel-regression predictors against GARCH(1,1), $\eta = 1$




Note: The figure plots $\mathrm{CSSPE}_{KR,t}^{(1)} - \mathrm{CSSPE}_{GARCH,t}^{(1)}$. Negative values indicate that the kernel predictor outperforms the standard GARCH(1,1) benchmark.

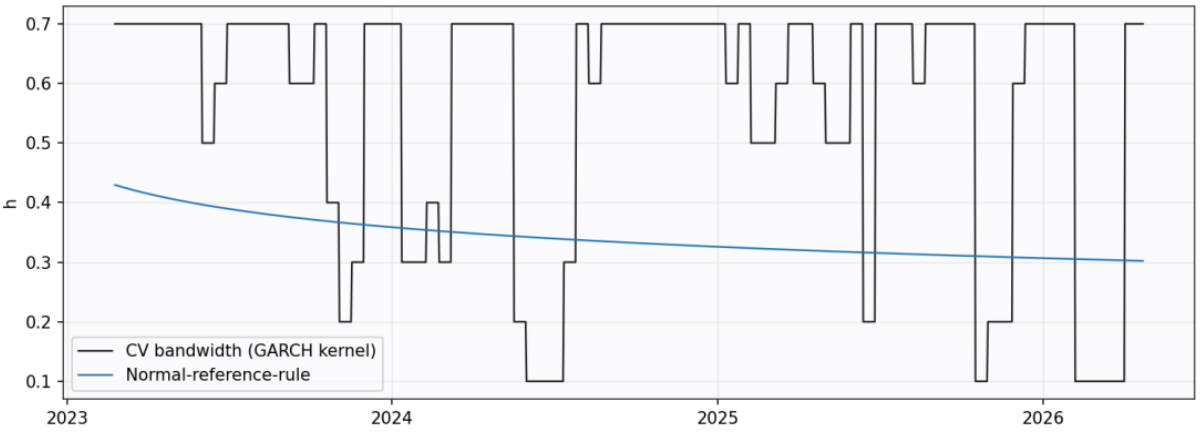
The difference between the two kernel configurations is economically informative. The GARCH-based kernel uses a volatility state that already contains a parametric summary of shock persistence and mean reversion. The kernel estimator then uses this state nonparametrically, allowing the relation between the current volatility state and future squared returns to be nonlinear. This combination performs well at the daily horizon because it retains the information content of the GARCH volatility recursion but relaxes the restriction that the forecast must be a linear affine function of the current variance.

The EWMA-based kernel is weaker because its volatility state is more directly tied to recent squared returns and less explicitly disciplined by mean reversion. Figure 10 shows that the EWMA-kernel predictor is competitive early in the sample, loses ground during the 2024 ETF-related repricing period, and then recovers later. This pattern suggests that the EWMA volatility state is useful when local volatility clustering is stable, but it can be less reliable when the market undergoes a sharp regime transition.

In contrast, the GARCH-kernel predictor remains more robust because the GARCH state filters recent volatility shocks through a persistent but mean-reverting structure.

Figure 11 reports the cross-validated bandwidth for the GARCH-kernel predictor together with the normal-reference-rule bandwidth. The normal-reference bandwidth declines smoothly as the effective sample size expands, reflecting the standard nonparametric rate. The cross-validated bandwidth is much more irregular and frequently moves between relatively low and high values. This is not a defect of the method. Rather, it indicates that the optimal degree of smoothing is state-dependent and regime-dependent. When the market is relatively stable, a larger bandwidth pools more historical observations and reduces noise. During turbulent episodes, a smaller bandwidth allows the estimator to focus on more local states and to respond more sharply to changing volatility conditions.

Figure 11: Cross-validation bandwidth (black) and normal-reference-rule bandwidth (blue) for the GARCH-kernel predictor, $\eta = 1$



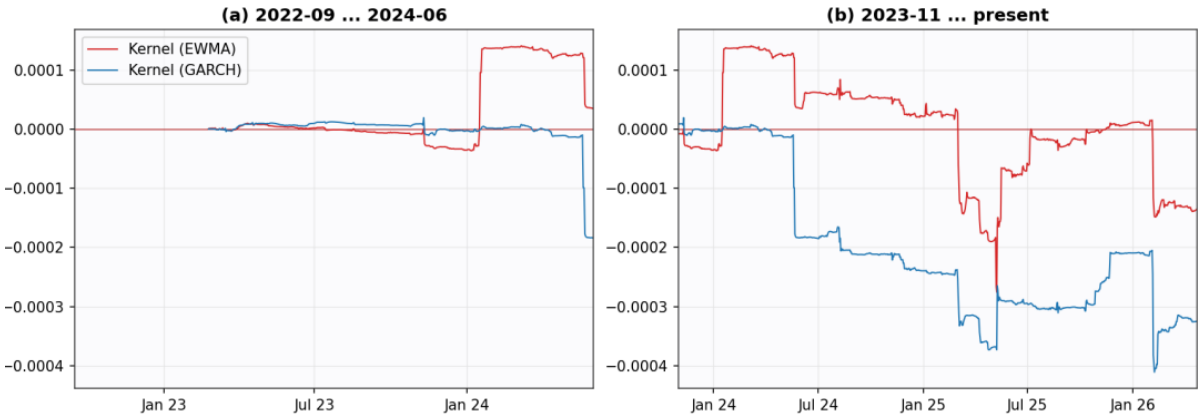

Note: The black line reports the bandwidth selected by expanding-window cross-validation. The blue line reports the normal-reference-rule bandwidth. Larger values of $h$ imply smoother kernel forecasts.

Figure 12 splits the one-day kernel comparison into two overlapping sub-samples. Panel (a) shows that the early part of the evaluation period contains relatively small differences between the kernel predictors and GARCH. Panel (b) shows that most of the advantage of the GARCH-kernel predictor is accumulated in the later part of the sample, especially around the ETF-related repricing period and the high-volatility 2025 episode. This is consistent with the main interpretation from subsections 3.4 and 3.5: nonparametric flexibility is most valuable when the market enters regimes in which the mapping from current volatility states to future squared returns changes over time.

At the ten-day prediction horizon, the results are substantially different. Figure 13 plots the two kernel predictors against GARCH for $\eta = 10$. In contrast to the one-day results, both kernel configurations accumulate positive differenced CSSPEs by the end of the sample. Therefore, the kernel approach does not dominate the standard GARCH benchmark at the ten-day-ahead horizon. This result is consistent with subsection 3.5, where the standard GARCH model performs well because its closed-form multi-step forecast shrinks volatility toward the unconditional variance. At longer prediction horizons, this built-in mean reversion becomes valuable.

The weaker performance of the kernel predictors at $\eta = 10$ should not be interpreted as evidence against nonlinear volatility modelling in general. Instead, it shows that the value of local nonparametric information is horizon-dependent. A kernel forecast is naturally local: it places high weight on historical observations whose current states resemble today's state. This is highly useful when forecasting tomorrow's squared return because the current state contains strong information about near-term volatility. However, when the target is $R^2_{t+10}$, the current state is less informative. Over ten days, the market

Figure 12: Differenced CSSPE for the kernel predictors at $\eta = 1$ across two overlapping sub-samples




Note: Panel (a) reports the early evaluation period from 2022-09 to 2024-06. Panel (b) reports the later evaluation period from 2023-11 to 2026-04. Negative values indicate that the kernel predictor outperforms the standard GARCH(1,1) benchmark.

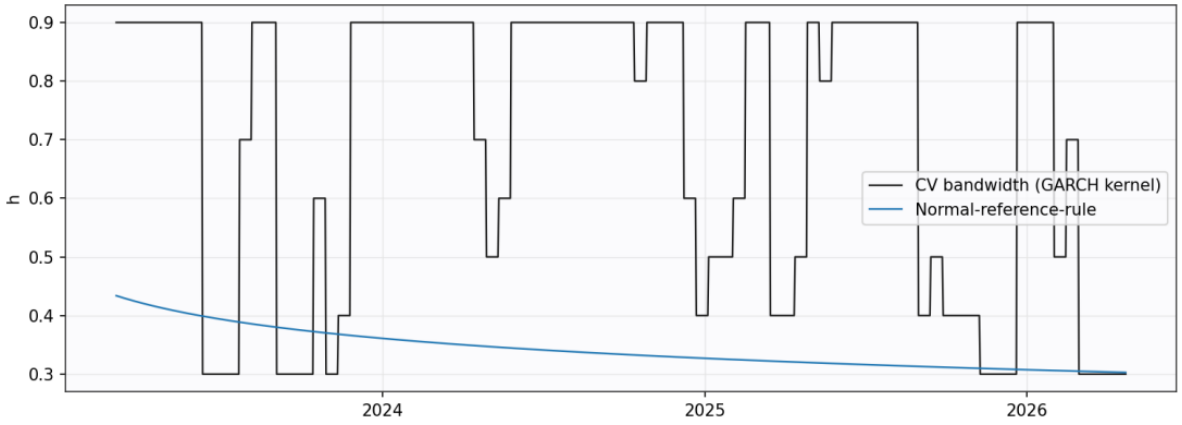
Figure 13: Differenced CSSPE for the two kernel-regression predictors against GARCH(1,1), $\eta = 10$


Note: The figure plots $\text{CSSPE}^{(10)}_{KR,t} - \text{CSSPE}^{(10)}_{GARCH,t}$. Positive values indicate that the standard GARCH(1,1) benchmark accumulates a smaller squared prediction error.

may move away from the current local state, and a model that explicitly imposes mean reversion can outperform a purely local nonparametric smoother.

Figure 14 reports the bandwidth behaviour for the GARCH-kernel predictor at $\eta = 10$. The cross-validated bandwidth is again irregular, indicating that the optimal amount of smoothing changes over time. However, relative to the one-day case, bandwidth variation at the ten-day horizon is less directly associated with sustained forecast gains. This reinforces the interpretation that choosing the right local neighbourhood is more difficult when the forecast target is further away from the current information set.

Figure 14: Cross-validation bandwidth (black) and normal-reference-rule bandwidth (blue) for the GARCH-kernel predictor, $\eta = 10$




The black line reports the cross-validated bandwidth and the blue line reports the normal-reference-rule bandwidth for the GARCH-state kernel predictor.

Figure 15 splits the ten-day kernel comparison into two overlapping sub-samples. The early sub-sample shows that the kernel predictors are not uniformly inferior throughout the entire evaluation period. There are episodes in which both kernel configurations are competitive with, or temporarily better than, GARCH. The later sub-sample, however, shows a clear deterioration after the market enters the more volatile 2025–2026 regime. This deterioration is consistent with the evidence in Figure 8: at longer prediction horizons, the GARCH benchmark benefits from its parametric mean-reversion structure, whereas local kernel forecasts can overreact to recent volatility states.

Finally, Figure 16 considers a mixed forecasting problem with return horizon $m = 10$ and prediction horizon $\eta = 1$. This case is distinct from the $\eta = 10$ exercise above. Instead of forecasting the single squared return $R^2_{t+10}$, the target is the squared cumulative return over the next ten days:

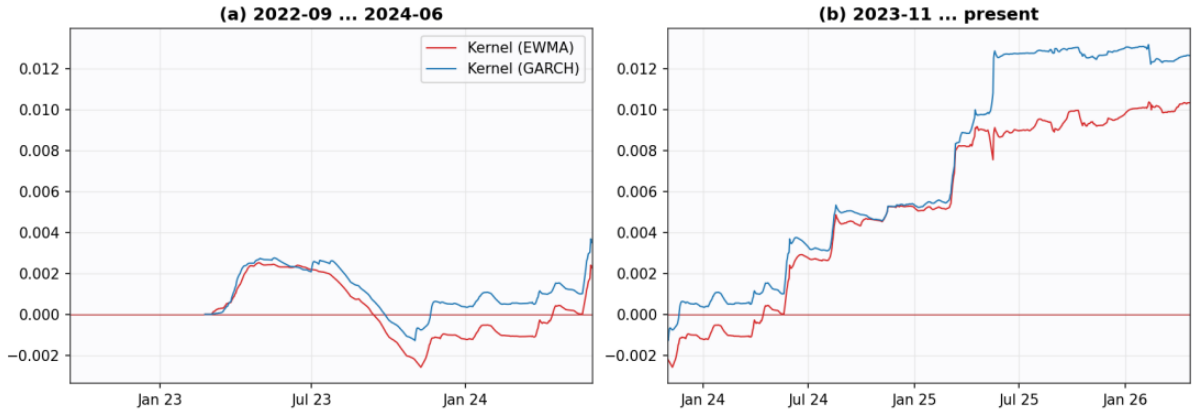
$$Y_{t+1,10} = \left( \sum_{\ell=1}^{10} R_{t+\ell} \right)^2 . \tag{36}$$

The forecast is formed at date $t$, so the target starts immediately after the information set. The corresponding kernel forecast is:

$$\widehat{f}_h(t; m = 10, \eta = 1) = \sum_{i=k}^{t-10} w_i(t) \left( \sum_{\ell=1}^{10} R_{i+\ell} \right)^2 . \tag{37}$$

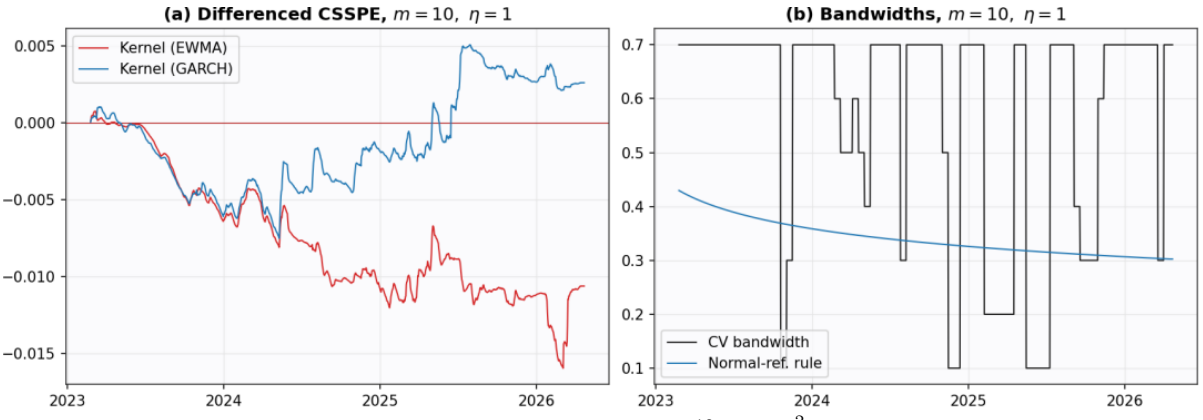
This mixed case is therefore a short prediction-horizon problem with a longer return-measurement window. It asks whether the current state predicts near-future ten-day cumulative volatility, rather than whether the current state predicts the single squared return ten days later.

Figure 15: Differenced CSSPE for the kernel predictors at $\eta = 10$ across two overlapping sub-samples




Note: Panel (a) reports the early evaluation period from 2022-09 to 2024-06. Panel (b) reports the later evaluation period from 2023-11 to 2026-04. Negative values indicate that the kernel predictor outperforms the standard GARCH(1,1) benchmark.

Figure 16: Kernel predictors versus GARCH(1,1) for return horizon $m = 10$ and prediction horizon $\eta = 1$




Note: Panel (a) reports differenced CSSPEs for the target $\left(\sum_{\ell=1}^{10} R_{t+\ell}\right)^2$. Panel (b) reports the bandwidths. Negative values in Panel (a) indicate that the kernel predictor outperforms the standard GARCH(1,1) benchmark.

The mixed-horizon results are important because both kernel configurations outperform GARCH substantially. This finding reconciles the strong one-day kernel results with the weak ten-day-ahead results. The key distinction is the distance between the information set and the target. When $\eta = 10$, the forecast target is far from the current information set, so the value of local state information decays. When $m = 10$ and $\eta = 1$, the target begins immediately after date $t$, so the current volatility state is highly informative even though the return measurement window spans ten days. In this case, the kernel method can exploit the near-term information in the state vector while allowing the relationship between the state and cumulative volatility to be nonlinear.

Overall, the kernel-regression evidence delivers four conclusions. First, at the one-day horizon, the GARCH-state kernel is the strongest predictor in the paper, outperforming the standard GARCH model and all EWMA-type benchmarks. Second, the choice of state variable matters: using the GARCH volatility state inside the kernel estimator produces stronger performance than using the EWMA volatility state. Third, the kernel advantage is horizon-dependent. Local nonparametric information is highly useful when the prediction horizon is short, but less useful when the forecast target is ten days away. Fourth, the mixed case $m = 10, \eta = 1$ shows that kernel regression is especially effective when the target begins close to the current information set, even if the return horizon itself is longer. These findings support the main methodological argument of the paper: Ethereum volatility is best forecast using economically meaningful state variables combined with a flexible, nonlinear forecasting map.

### 3.7 The estimated kernel surface

This subsection provides a visual and economic interpretation of the estimated two-dimensional kernel-regression surface. While the previous subsections evaluate forecasting performance using cumulative squared prediction errors, the purpose here is different. The kernel surface allows us to inspect the shape of the estimated conditional-volatility function directly. In particular, it shows how the predicted future squared return varies with two state variables: the rank-transformed volatility state and the rank-transformed signed-return state.

Let $x = (x_1, x_2)'$ denote a generic point in the transformed predictor space. The first coordinate, $x_1$, corresponds to the transformed moving average of past squared returns. It therefore measures the recent volatility state of the Ethereum market. A high value of $x_1$ indicates that recent ETH returns have been large in absolute value, while a low value of $x_1$ indicates a relatively calm recent market environment. The second coordinate, $x_2$, corresponds to the transformed moving average of signed returns. It captures the direction of recent market movements. Negative values of $x_2$ correspond to recent negative-return states, whereas positive values of $x_2$ correspond to recent positive-return states.

The estimated kernel-regression surface is defined as:

$$\widehat{f}_h(x_1, x_2) = \sum_{i=k}^{T-1} \omega_i(x_1, x_2) R_{i+1}^2, \tag{38}$$

where the kernel weights are:

$$\omega_i(x_1, x_2) = \frac{K_h((x_1, x_2)' - X_i)}{\sum_{j=k}^{T-1} K_h((x_1, x_2)' - X_j)}. \tag{39}$$

Here $X_i = (X_{1i}, X_{2i})'$ is the transformed predictor vector observed at date $i$, $R_{i+1}^2$ is the next-day realised squared return, and $K_h(\cdot)$ is a Gaussian product kernel with bandwidth $h$. In the surface visualisation, we use $h = 0.5$. Thus, $\widehat{f}_h(x_1, x_2)$ can be interpreted as the locally weighted average of next-day ETH

realised variance following historical market states similar to $(x_1, x_2)$.

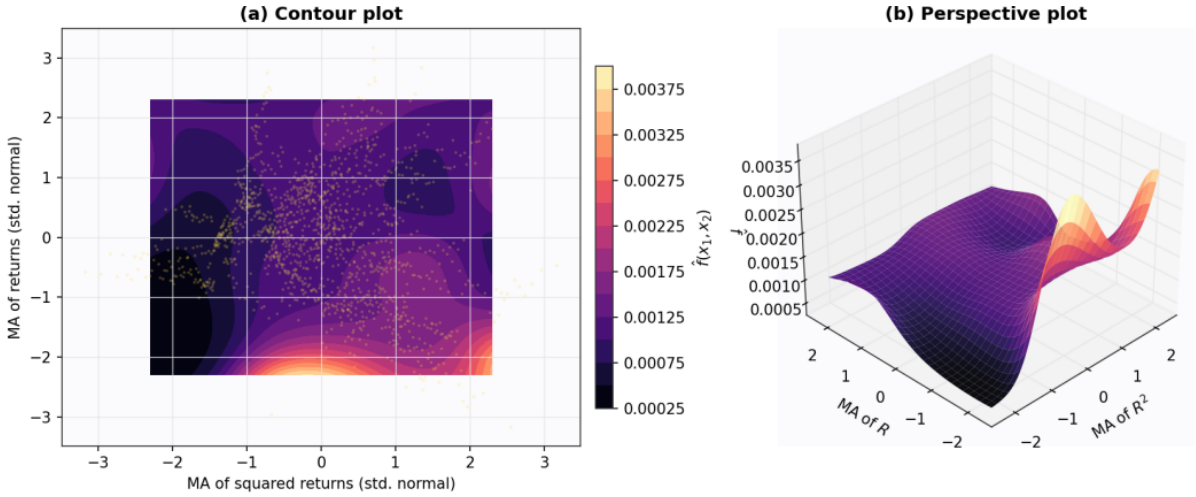
The Gaussian product kernel takes the form:

$$K_h((x_1, x_2)' - X_i) = \frac{1}{h^2} K\left(\frac{x_1 - X_{1i}}{h}\right) K\left(\frac{x_2 - X_{2i}}{h}\right), \tag{40}$$

where $K(\cdot)$ is the standard Gaussian density. The forecast at each point of the surface is therefore constructed by giving the largest weight to historical observations whose transformed volatility and signed-return states are closest to the point being evaluated. Observations far away from $(x_1, x_2)$ receive exponentially smaller weights. This local weighting mechanism is what gives the estimated surface its nonlinear shape.

Figure 17 reports the estimated surface using both contour and perspective plots. The contour plot is useful for reading the level sets of predicted future variance, while the perspective plot makes the curvature of the fitted surface visually clearer. The main feature of the surface is that predicted future ETH variance is low in the calm region of the state space and high in the stress region. When $x_1$ is low, the fitted values of $\widehat{f}_h(x_1, x_2)$ remain relatively small across most values of $x_2$. This indicates that when the recent volatility state is calm, the predicted next-day squared return is also low. By contrast, when $x_1$ is high, predicted future variance increases sharply.

Figure 17: Kernel regression surface $\hat{f}(x_1, x_2)$ of ETH realised daily variance on rank-transformed moving averages of squared returns and of signed returns




Note: The figure reports contour and perspective plots of the estimated kernel-regression function. The horizontal axis $x_1$ denotes the rank-transformed moving average of squared returns. The vertical axis $x_2$ denotes the rank-transformed moving average of signed returns. The fitted value $\widehat{f}_h(x_1, x_2)$ is the predicted next-day realised variance of ETH returns. The surface is estimated using a Gaussian product kernel with bandwidth $h = 0.5$.

The second important feature of the surface is asymmetry along the signed-return dimension. The largest values of $\widehat{f}_h(x_1, x_2)$ are concentrated in the region where $x_1$ is high and $x_2$ is strongly negative. Economically, this region corresponds to market states in which Ethereum has recently displayed both high volatility and negative return pressure. In such states, the model predicts a particularly large next-day squared return. This pattern is consistent with a leverage-type volatility response: adverse return states are followed by higher conditional volatility than neutral or positive return states with similar recent volatility.

The location of the peak is economically informative. A high value of $x_1$ alone means that recent absolute returns have been large. A negative value of $x_2$ means that recent signed returns have been

predominantly negative. The surface shows that the combination of these two conditions is especially important. Volatility does not rise only because the market has been turbulent; it rises most strongly when turbulence coincides with downside price pressure. This suggests that negative-return episodes in Ethereum are not merely large observations in the return distribution, but are associated with a distinct future-volatility state.

This interaction can be expressed by considering the local partial effects of the two state variables. The effect of the volatility state is:

$$\frac{\partial \widehat{f}_h(x_1, x_2)}{\partial x_1}, \tag{41}$$

while the effect of the signed-return state is

$$\frac{\partial \widehat{f}_h(x_1, x_2)}{\partial x_2}. \tag{42}$$

The estimated surface suggests that $\partial \widehat{f}_h / \partial x_1$ is positive over a large part of the state space: higher recent volatility predicts higher future squared returns. However, this effect is not constant. The slope with respect to $x_1$ is steeper when $x_2$ is negative, meaning that the effect of volatility clustering is amplified when recent returns have been adverse. Similarly, the effect of $x_2$ depends on $x_1$. When $x_1$ is low, changes in $x_2$ have limited predictive content. When $x_1$ is high, moving from positive to negative values of $x_2$ substantially raises the fitted variance.

The estimated kernel surface therefore identifies a nonlinear interaction between volatility clustering and signed-return asymmetry. A purely symmetric volatility model would allow $x_1$ to matter but would not allow the sign of past returns to change the predicted variance in such a flexible way. A purely linear asymmetric model would allow $x_2$ to matter, but would impose a single slope or a single asymmetry coefficient across all market states. By contrast, the kernel surface allows the strength of asymmetry to vary across the volatility distribution. This is precisely why the nonparametric representation is useful: it does not force the effect of signed returns to be the same in calm markets and turbulent markets.

Figure 18 gives a more transparent view of the same object by plotting slices of the estimated surface. These slices convert the three-dimensional surface into a set of one-dimensional curves. Panel (a) fixes $x_2$ at several representative values and plots

$$x_1 \mapsto \widehat{f}_h(x_1, x_2).$$

Panel (b) fixes $x_1$ at several representative values and plots

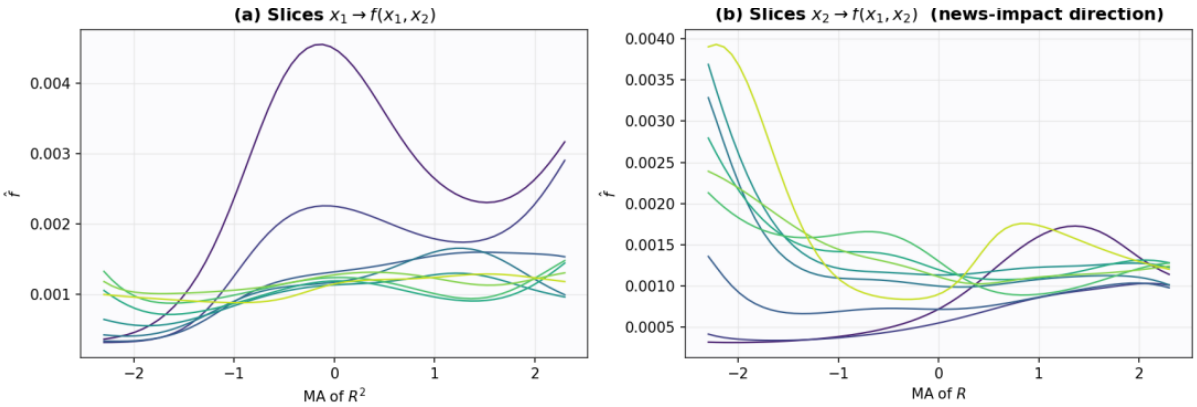
$$x_2 \mapsto \widehat{f}_h(x_1, x_2).$$

The slices are useful because they show how the marginal relation between one state variable and future variance changes when the other state variable is held fixed.

Panel (a) shows that predicted future variance generally increases with $x_1$. This confirms that the kernel estimator recovers the standard volatility-clustering mechanism: high recent volatility tends to be followed by high future volatility. However, the steepness of the curve depends on the value of $x_2$. For strongly negative values of $x_2$, the curve is steeper and reaches higher fitted values. For neutral or positive values of $x_2$, the increase in predicted variance is more moderate. This means that the effect of recent volatility is conditional on the direction of recent returns.

Panel (b) highlights the asymmetry more directly. When $x_1$ is high, the fitted curve is higher for negative values of $x_2$, implying that recent negative-return states predict larger future squared returns.

Figure 18: Slices of the estimated kernel regression surface




Note: Panel (a) plots $\widehat{f}_h(x_1, x_2)$ as a function of $x_1$ for several fixed values of $x_2$. Panel (b) plots $\widehat{f}_h(x_1, x_2)$ as a function of $x_2$ for several fixed values of $x_1$. The slices are computed from the same estimated surface shown in Figure 17.

When $x_1$ is low, the curve is much flatter, implying that signed-return information has only limited additional forecasting power in calm periods. Thus, the asymmetry is not uniform across the whole state space. It is concentrated in high-volatility regimes, where negative return pressure carries strong information about near-term volatility.

This surface interpretation is useful because it translates the kernel-regression estimator from a purely statistical forecasting device into an economically interpretable volatility map. The high-$x_1$, low-$x_2$ region can be interpreted as a stress region: the market has recently experienced large movements and those movements have been predominantly negative. The low-$x_1$, near-zero-$x_2$ region can be interpreted as a calm region: recent returns have been small and directionally neutral. The high-$x_1$, positive-$x_2$ region corresponds to an upside-turbulence regime, where large movements occur during positive price runs. The fitted surface suggests that all three regions have different volatility implications.

It is also useful to compare the estimated kernel surface with the intuition of standard parametric volatility models. In a standard GARCH model, high past squared returns increase future conditional variance, but the sign of the return does not directly matter. In an asymmetric volatility model, the sign of the return matters, but the asymmetry is usually imposed through a fixed parameter. The kernel surface generalises both ideas. It captures volatility clustering through the $x_1$ direction, captures signed-return effects through the $x_2$ direction, and allows the interaction between the two to be nonlinear. Thus, the surface provides a richer description of the ETH volatility process than either a symmetric GARCH recursion or a constant-parameter asymmetry model.

The surface also explains why the kernel-regression approach can be valuable for cryptocurrency volatility prediction. Ethereum markets are characterised by rapid changes in liquidity, leverage, market sentiment, and institutional participation. These features can generate volatility regimes that are difficult to summarise with a single persistence coefficient. The estimated surface shows that the relevant state space contains calm regions, upside-turbulence regions, and downside-stress regions. A flexible nonparametric method can adapt to these regions by using nearby historical observations rather than forcing the same volatility-response rule to hold everywhere.

Overall, Figures 17 and 18 show that the estimated kernel-regression function is economically interpretable. The predicted next-day realised variance is lowest in calm, neutral-return states and highest in high-volatility states associated with recent negative returns. The shape of the surface provides vi-

sual evidence of three empirical features of ETH volatility: persistence, asymmetry, and nonlinear state dependence. These features are central to the motivation for using kernel regression. The method does not merely improve forecast accuracy mechanically; it also reveals how future volatility changes across different regions of the Ethereum market-state space.

### 3.8 Conditional quantile (VaR) estimation

We now turn from conditional variance prediction to conditional tail-risk prediction. This step is important because a volatility forecast is not itself a risk-management object. A volatility forecast measures the expected scale of future return fluctuations, whereas a value-at-risk forecast measures a specific lower-tail quantile of the future return distribution. Since the ETH return distribution in Table 1 is strongly leptokurtic, a direct translation from volatility into VaR requires careful treatment of the innovation distribution. For this reason, we compare GARCH-based and kernel-regression-based VaR forecasts using two alternative residual inversions: a parametric Student-$t$ inversion and a fully empirical residual-quantile inversion.

Let $q_{j,t+\eta|t}(\tau)$ denote the $\tau$-level conditional return quantile produced by model $j$, where $j \in \{\text{GARCH}, \text{kernel}\}$ and $\tau \in \{0.01, 0.05\}$. In this paper, VaR is reported as a lower-tail return threshold. Thus, a more negative value of $q_{j,t+\eta|t}(\tau)$ corresponds to a larger downside loss. If one wishes to express VaR as a positive loss number, the corresponding loss-based VaR is simply

$$\text{VaR}_{j,t+\eta|t}(\tau) = -q_{j,t+\eta|t}(\tau).$$

For consistency with the return-based forecasting exercise above, we work directly with $q_{j,t+\eta|t}(\tau)$.

For each model $j$, the VaR forecast is constructed by combining a conditional mean estimate, a conditional volatility estimate, and a standardized innovation quantile:

$$\widehat{q}^{(a)}_{j,t+\eta|t}(\tau) = \widehat{\mu}_{t+\eta|t} + \widehat{\sigma}_{j,t+\eta|t}\, \widehat{z}^{(a)}_{j,t}(\tau), \qquad a \in \{\text{St}, \text{Emp}\},$$

where $a = \text{St}$ denotes the Student-$t$ residual inversion and $a = \text{Emp}$ denotes the empirical residual-quantile inversion. The conditional mean $\widehat{\mu}_{t+\eta|t}$ is estimated using the available past returns. In practice, the mean component is small relative to the volatility component for daily ETH returns, but it is retained in the notation to make clear that VaR is a forecast of the conditional return quantile rather than of volatility alone.

The two volatility forecasts are obtained as follows. For the GARCH benchmark, the conditional volatility is:

$$\widehat{\sigma}_{\text{GARCH},t+\eta|t} = \left[\widehat{\bar{\sigma}}^{\,2} + (\widehat{\alpha} + \widehat{\beta}_1)^{\eta-1} \left(\widehat{\sigma}^2_{t+1} - \widehat{\bar{\sigma}}^{\,2}\right)\right]^{1/2},$$

where

$$\widehat{\bar{\sigma}}^{\,2} = \frac{\widehat{\beta}_0}{1 - \widehat{\alpha} - \widehat{\beta}_1}.$$

For the kernel-regression predictor, the conditional volatility is obtained from the nonparametric forecast of future squared returns:

$$\widehat{\sigma}_{\text{kernel},t+\eta|t} = \left[\max\left\{\widehat{f}_h(t,\eta), \varepsilon\right\}\right]^{1/2},$$

where $\varepsilon > 0$ is a small numerical lower bound used only to avoid taking the square root of a non-positive fitted value. This definition makes the VaR exercise directly comparable to the volatility forecasting exercise in the previous subsections, because the same forecast object, $\widehat{f}_h(t,\eta)$, is used to construct the kernel-based tail-risk forecast.

The standardized residuals for model $j$ are computed recursively as:

$$\widehat{\varepsilon}_{j,s} = \frac{R_s - \widehat{\mu}_{s|s-\eta}}{\widehat{\sigma}_{j,s|s-\eta}}, \qquad s \le t.$$

The Student-$t$ inversion estimates the left-tail innovation quantile by fitting a Student-$t$ distribution to the past standardized residuals. Since the Student-$t$ distribution with degrees of freedom $\nu$ has variance $\nu/(\nu-2)$, the unit-variance standardized Student-$t$ quantile is:

$$\widehat{z}_{j,t}^{(\mathrm{St})}(\tau) = \sqrt{\frac{\widehat{\nu}_{j,t}-2}{\widehat{\nu}_{j,t}}}\, F^{-1}_{t_{\widehat{\nu}_{j,t}}}(\tau),$$

where $F^{-1}_{t_{\widehat{\nu}_{j,t}}}(\tau)$ is the $\tau$-quantile of a Student-$t$ distribution with $\widehat{\nu}_{j,t}$ degrees of freedom. This approach imposes a smooth parametric tail shape while allowing for excess kurtosis relative to the Gaussian benchmark.

The empirical inversion avoids imposing a parametric innovation distribution. It estimates the residual quantile directly from the historical standardized residuals:

$$\widehat{z}_{j,t}^{(\mathrm{Emp})}(\tau) = \inf\left\{ z : \frac{1}{N_t}\sum_{s\le t} \mathbf{1}\{\widehat{\varepsilon}_{j,s} \le z\} \ge \tau \right\},$$

where $N_t$ is the number of standardized residuals available at date $t$. This empirical residual method is particularly attractive for ETH returns because the lower tail may be shaped by discrete market events, liquidation cascades, exchange-specific stress, and ETF-related repricing episodes that are not well approximated by a fixed Student-$t$ law.

Forecast accuracy is evaluated using the quantile, or check, loss function

$$\rho_\tau(u) = u\left(\tau - \mathbf{1}\{u<0\}\right),$$

where the forecast error is:

$$u_{j,s+\eta} = R_{s+\eta} - \widehat{q}^{(a)}_{j,s+\eta|s}(\tau).$$

The cumulative quantile loss for model $j$ is therefore

$$SL^{(a)}_{j,t}(\tau) = \sum_{s=t_0}^{t-\eta} \rho_\tau\left(R_{s+\eta} - \widehat{q}^{(a)}_{j,s+\eta|s}(\tau)\right).$$

A smaller value of $SL^{(a)}_{j,t}(\tau)$ indicates a more accurate conditional quantile forecast. To compare the kernel-regression VaR forecast with the GARCH VaR benchmark, we plot the differenced cumulative quantile loss:

$$D^{(a)}_t(\tau) = SL^{(a)}_{\mathrm{kernel},t}(\tau) - SL^{(a)}_{\mathrm{GARCH},t}(\tau).$$

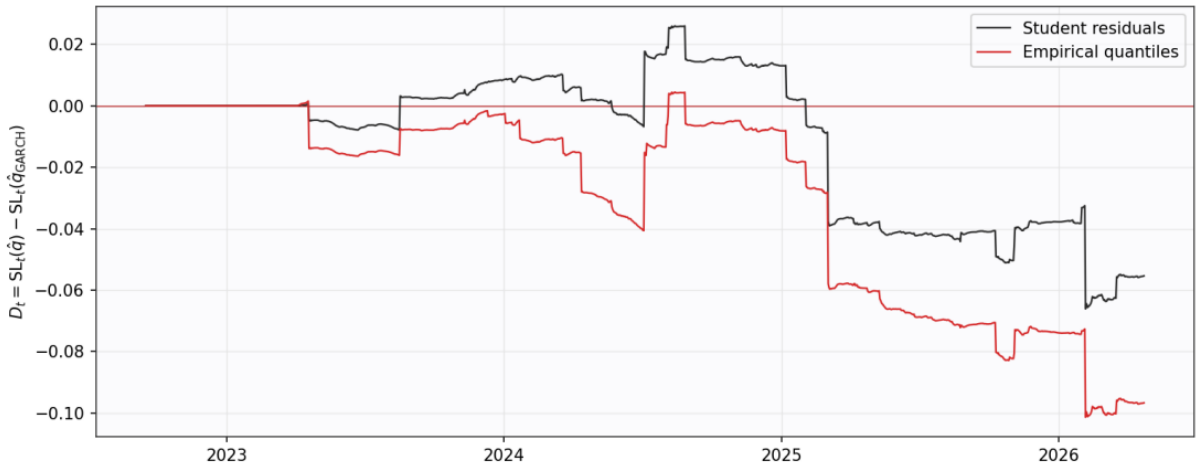
Hence,

$$D^{(a)}_t(\tau) < 0 \quad \Longleftrightarrow \quad SL^{(a)}_{\mathrm{kernel},t}(\tau) < SL^{(a)}_{\mathrm{GARCH},t}(\tau),$$

so a negative value means that the kernel-regression VaR forecast has accumulated a smaller quantile loss than the GARCH-based VaR forecast. Conversely, a positive value means that the GARCH-based VaR forecast is more accurate up to date $t$. This sign convention is important for interpreting Figures 19–22.

Figure 19 reports the one-day-ahead VaR comparison at the extreme lower tail, $\tau = 0.01$ and $\eta = 1$. Both residual-inversion methods eventually generate negative values of $D_t^{(a)}(\tau)$, implying that the kernel-regression VaR forecast outperforms the GARCH-based VaR forecast over most of the evaluation sample. The improvement is especially strong after the 2024 spot-ETH-ETF approval period, when ETH volatility dynamics became more nonlinear and more state-dependent. The empirical residual-quantile inversion performs particularly well: its cumulative loss difference lies below the Student-$t$ path for much of the sample and ends with the largest negative gap relative to GARCH. This indicates that, for the most extreme one-day tail events, the historical empirical tail of standardized residuals contains information that is not fully captured by a smooth Student-$t$ approximation.

Figure 19: Quantile estimation of 1-day returns, $\tau = 0.01$, $\eta = 1$

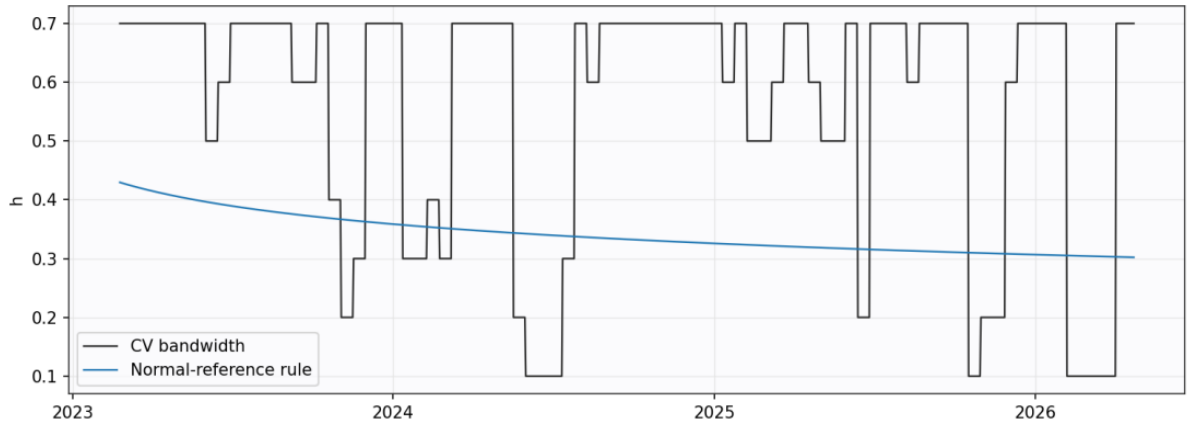



Note: The figure reports the differenced cumulative quantile loss for one-day-ahead conditional VaR forecasts at the 1% lower-tail quantile, $\tau = 0.01$. The plotted series are $D_t^{(a)}(0.01) = SL^{(a)}_{\mathrm{kernel},t}(0.01) - SL^{(a)}_{\mathrm{GARCH},t}(0.01)$, $a \in \{\mathrm{St}, \mathrm{Emp}\}$, where $a = \mathrm{St}$ denotes the Student-$t$ residual inversion and $a = \mathrm{Emp}$ denotes the empirical residual-quantile inversion. Negative values indicate that the kernel-regression VaR forecast accumulates a lower quantile loss than the GARCH-based VaR forecast.

Figure 20 plots the bandwidths used for the quantile estimation in Figure 19. The normal-reference-rule bandwidth declines smoothly over time as the effective sample size expands. In contrast, the cross-validated bandwidth is much more irregular and frequently switches between high and low values. This behaviour is expected in a tail-risk problem. At $\tau = 0.01$, the loss function is driven by relatively few exceedance observations, so the cross-validation criterion can be sensitive to the arrival or removal of a small number of extreme return days. Economically, this instability is informative: the optimal degree of smoothing for ETH tail risk is not constant over time, especially around periods of sharp repricing and heavy liquidation activity.

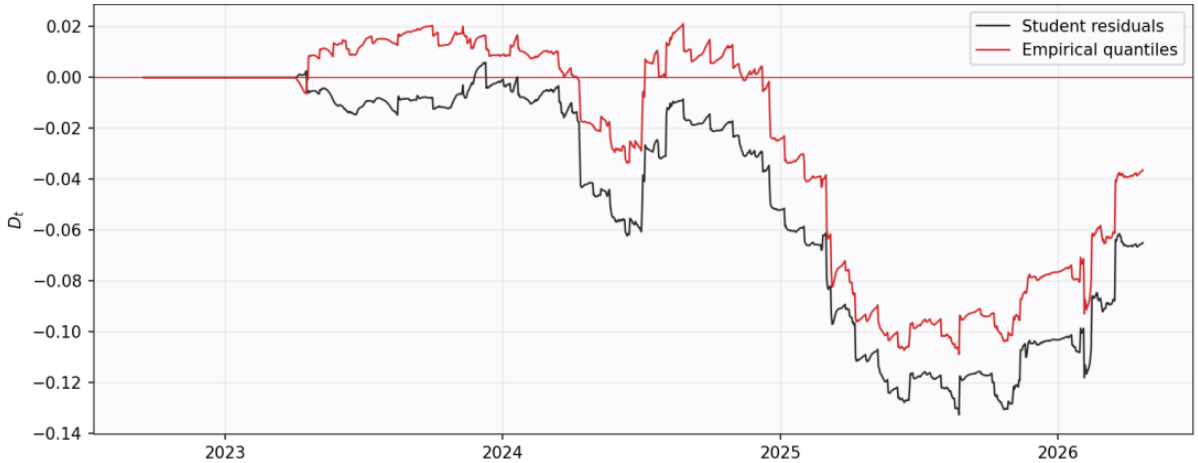
Figure 21 repeats the one-day-ahead VaR exercise at the less extreme quantile, $\tau = 0.05$. The kernel-regression VaR forecast still performs competitively against the GARCH benchmark, but the advantage is less uniform than at $\tau = 0.01$. This difference is intuitive. The 1% tail is dominated by rare stress events, where nonlinear state dependence and empirical tail information are especially valuable. The 5% tail contains more ordinary downside observations, so the parametric GARCH benchmark becomes harder to dominate. The figure shows that the kernel advantage is concentrated mainly around the 2024–2025 high-volatility window and narrows toward the end of the sample. Moreover, the ranking between the Student-$t$ and empirical residual inversions is less clear at $\tau = 0.05$ than at $\tau = 0.01$. This suggests that the empirical residual method is most valuable when the target is the far-left tail rather than the broader lower tail.

Figure 20: Cross-validation and normal-reference-rule bandwidths for the quantile estimation of Figure 19




Note: The figure reports the bandwidths used for the one-day-ahead kernel VaR estimation at $\tau = 0.01$. The cross-validated bandwidth is selected by minimising the expanding-window quantile loss, while the normal-reference-rule bandwidth follows the rule-of-thumb rate for a two-dimensional kernel predictor. Larger values of $h$ imply smoother VaR forecasts because more historical states receive non-negligible kernel weights. Smaller values of $h$ imply more local forecasts that respond more sharply to nearby volatility states.

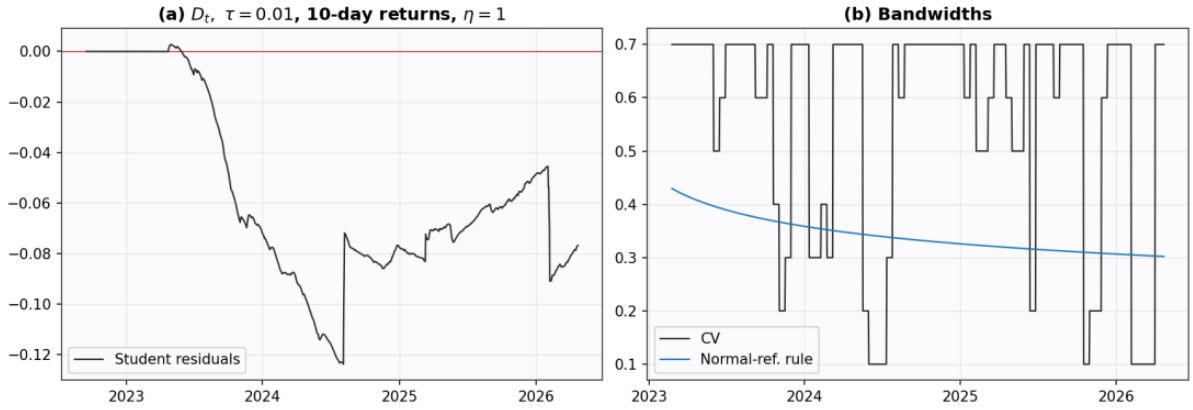
Figure 21: Quantile estimation of 1-day returns, $\tau = 0.05$, $\eta = 1$




Note: The figure reports the differenced cumulative quantile loss for one-day-ahead conditional VaR forecasts at the 5% lower-tail quantile, $\tau = 0.05$. The plotted object is $D_t^{(a)}(0.05) = SL_{\text{kernel},t}^{(a)}(0.05) - SL_{\text{GARCH},t}^{(a)}(0.05)$, $a \in \{\text{St}, \text{Emp}\}$. Negative values indicate that the kernel-regression VaR forecast outperforms the GARCH-based VaR forecast under the quantile loss. Compared with the 1% VaR exercise, the 5% quantile evaluates a broader lower-tail region and is therefore less dominated by rare extreme-loss observations.

Figure 22 considers the same quantile-loss comparison for ten-day cumulative returns at $\tau = 0.01$, using information available at date $t$. Thus, the object being forecast is a ten-day return quantile, while the information set remains one-step ahead in the sense that the forecast is formed before the ten-day return is realised. This distinction matters because the label "10-day returns" should not be confused with the one-day return horizon used in Figures 19–21. The cumulative loss difference falls sharply during the 2023–2024 period and remains negative for most of the evaluation sample, implying that the kernel-regression VaR remains competitive for multi-day downside risk. Although part of the kernel advantage reverses after the 2025 peak, the final value remains below zero. The ten-day evidence therefore supports the same broad conclusion as the one-day extreme-tail exercise: the nonparametric state-dependent volatility forecast can be translated into a useful tail-risk forecast, particularly when the target is severe downside risk.

Figure 22: Quantile estimation of 10-day returns, $\tau = 0.01$, $\eta = 1$: (a) differenced cumulative loss, (b) bandwidths

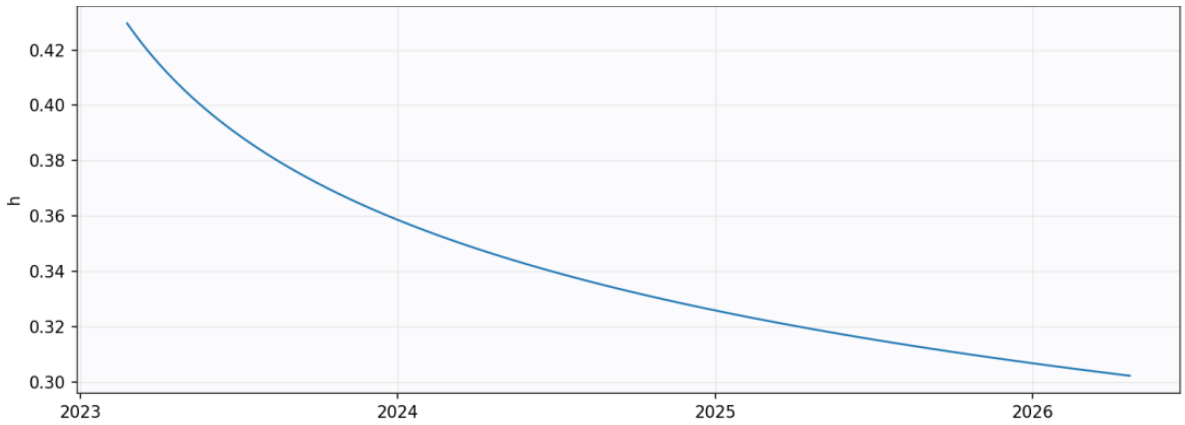



Note: The figure reports the differenced cumulative quantile loss for conditional VaR forecasts of ten-day cumulative ETH returns at the 1% lower-tail quantile. The target return is $R_{t,t+10} = \sum_{\ell=1}^{10} R_{t+\ell}$. The plotted series compare the kernel-regression VaR forecast with the GARCH-based VaR benchmark: $D_t^{(a)}(0.01) = SL_{\text{kernel},t}^{(a)}(0.01) - SL_{\text{GARCH},t}^{(a)}(0.01)$. Negative values indicate that the kernel VaR forecast accumulates a lower quantile loss. This figure therefore evaluates multi-day downside-risk forecasting rather than the single-day squared-return forecasting horizon used in the volatility-prediction exercises.

Figure 23 shows the normal-reference-rule bandwidth for the ten-day quantile exercise. As in the one-day case, the rule-of-thumb bandwidth declines gradually as the sample expands. This smooth path provides a useful benchmark against the more erratic cross-validated bandwidths. The contrast between the two bandwidth choices highlights a practical trade-off in VaR estimation. A larger bandwidth stabilises the estimated tail quantile by pooling more historical states, but may oversmooth regime changes. A smaller bandwidth adapts more quickly to local market conditions, but can become noisy when tail observations are sparse. The empirical results suggest that the kernel VaR gains are not driven solely by one bandwidth rule; rather, they reflect the broader usefulness of conditioning tail forecasts on the transformed ETH volatility state $X_t = (X_{1t}, X_{2t})$.

Overall, the conditional quantile results reinforce the volatility-forecasting findings. The kernel regression does not merely improve squared-return prediction in a mean-square-error sense; it also produces economically meaningful improvements in lower-tail risk forecasts. The gains are strongest for the 1% VaR, where extreme ETH losses are most sensitive to nonlinear volatility states and empirical tail behaviour. At the 5% quantile, the kernel advantage is weaker but remains visible during the 2024–2025 stress period. For ten-day cumulative returns, the kernel-based VaR also accumulates lower quantile loss than the GARCH benchmark over most of the sample. These findings suggest that state-dependent

Figure 23: Normal-reference-rule bandwidth for the quantile estimation of 10-day returns, $\tau = 0.01$, $\eta = 1$




Note: The figure reports the normal-reference-rule bandwidth used in the kernel-based VaR estimation for ten-day cumulative returns. The rule-of-thumb bandwidth declines gradually as the effective sample size increases, reflecting the standard nonparametric bias–variance trade-off. A larger bandwidth produces smoother tail-risk forecasts by pooling more historical observations, whereas a smaller bandwidth gives more weight to observations with state variables closer to the current ETH volatility state.

nonparametric volatility modelling is especially useful for measuring ETH downside risk in periods where conventional GARCH dynamics are too restrictive.

# 4 Conclusion

This paper studies volatility prediction for Ethereum in the post-Merge era. The empirical question is motivated by a simple observation: after the Merge, ETH continued to trade as a cryptoasset, but the economic environment surrounding Ethereum changed substantially. The network moved from proof-of-work to proof-of-stake, staking became central to the asset's economic narrative, and the market became increasingly exposed to institutional participation, exchange-traded-product expectations, decentralised-finance liquidity, and regulatory news. These features make post-Merge ETH an important testing ground for volatility models that can accommodate persistence, heavy tails, asymmetry, and nonlinear regime dependence.

Using daily ETH/USD returns from 15 September 2022 to 23 April 2026, the paper compares several volatility predictors: standard GARCH(1,1), Heston–Nandi GARCH(1,1), cross-validated EWMA predictors, aggregated EWMA predictors, and Nadaraya–Watson kernel-regression predictors. The kernel approach uses a two-dimensional state vector that summarises recent volatility and recent signed returns. These raw predictors are transformed into normal scores using empirical ranks, producing a more homogeneous predictor space for nonparametric forecasting. The empirical design therefore asks whether local, state-dependent information improves volatility forecasts beyond conventional parametric benchmarks.

The first main finding is that post-Merge ETH volatility is persistent, but not adequately described by a single fixed volatility-updating rule. The standard GARCH(1,1) model provides a useful benchmark, and its fitted volatility path captures the broad clustering pattern in ETH returns. However, the rolling estimates show that volatility persistence changes substantially across market regimes. The Heston–Nandi estimates further indicate that signed-return asymmetry is unstable over time. This instability is economically important. It means that ETH volatility cannot be fully summarised by one constant

persistence parameter or one constant leverage parameter. Rather, the volatility process appears to shift across calm states, upside-turbulence states, and downside-stress states.

The second main finding is that forecast performance is strongly horizon-dependent. At the one-day prediction horizon, $\eta = 1$, several alternatives improve on the standard GARCH benchmark. The best-performing model is the kernel predictor that uses the fitted GARCH volatility state as its first state variable. This result is central to the paper. It shows that the GARCH volatility state contains valuable information, but that the most effective use of this information is not necessarily a linear parametric recursion. Instead, forecast accuracy improves when the GARCH state is embedded inside a nonparametric forecasting map that allows the relation between current market conditions and future squared returns to vary across the state space.

The third main finding is that the kernel advantage weakens when the prediction target is further away from the information set. At the ten-day-ahead horizon, $\eta = 10$, the standard GARCH model becomes harder to beat. This reflects the value of the GARCH model's mean-reverting multi-step forecast. When recent volatility shocks are sharp but partly transitory, a model that mechanically shrinks forecasts back toward the unconditional variance can perform better than a highly local adaptive method. The ten-day results therefore qualify the main forecasting message: nonlinear local information is valuable, but its value decays as the distance between the current state and the forecast target increases.

The mixed-horizon exercise sharpens this interpretation. When the target is the squared cumulative return over the next ten days, but the prediction horizon begins immediately after date $t$, the kernel predictors again perform strongly. This distinction is important. The weak performance at $\eta = 10$ does not imply that kernel regression is ineffective for multi-day risk. Rather, it implies that the timing of the target matters. Kernel regression is most useful when the forecast target begins close to the current information set, because current volatility states contain strong information about near-term market risk.

The estimated kernel surface provides an economic interpretation of these forecasting results. The surface shows that predicted next-day realised variance is lowest in calm, neutral-return states and highest when high recent volatility coincides with negative signed-return pressure. This pattern reveals three key features of the post-Merge ETH volatility process: volatility persistence, signed-return asymmetry, and nonlinear state dependence. The surface therefore converts the kernel estimator from a black-box forecasting device into an interpretable volatility map. ETH volatility is not simply high after large past returns; it is especially high in particular regions of the market-state space.

The conditional quantile results extend the analysis from volatility prediction to tail-risk measurement. The kernel-based VaR forecasts perform especially well at the 1% lower-tail quantile. The gains are strongest when the empirical residual-quantile inversion is used, suggesting that extreme ETH losses are not fully captured by a smooth parametric Student-$t$ residual distribution. At the 5% quantile, the kernel advantage is weaker and more time-varying, which is consistent with the idea that nonlinear state dependence matters most for rare stress events. For ten-day cumulative returns, the kernel-based VaR also remains competitive, reinforcing the conclusion that state-dependent nonparametric volatility forecasting is useful for downside-risk management.

These findings have several policy and risk-management implications. First, exchanges, brokers, and clearing platforms should not rely exclusively on fixed-parameter volatility models when setting margins for ETH-related products. A standard GARCH model is useful, especially for longer-horizon forecasts, but the evidence suggests that near-term ETH risk is highly state-dependent. Margin systems that condition only on recent variance may miss the additional information contained in signed-return states. A more robust framework should combine conventional parametric volatility filters with nonlinear

state-dependent adjustments.

Second, regulators and market supervisors should monitor crypto volatility using models that distinguish between ordinary volatility clustering and stress-state amplification. The estimated kernel surface shows that high-volatility, negative-return states are associated with especially large future realised variance. This means that downside-stress regimes may require closer surveillance than calm or upside-turbulence regimes with similar average volatility. For supervisory purposes, the relevant question is not only whether volatility is high, but also where the market is located in the volatility–return state space.

Third, institutional investors, ETF issuers, and risk managers should evaluate ETH risk using horizon-specific forecasting models. The paper shows that a model that performs well at the one-day horizon may not perform well at the ten-day-ahead horizon. This has direct implications for portfolio risk limits, hedging intervals, and liquidity planning. Short-horizon risk systems should place greater weight on local state-dependent information, while longer-horizon systems should retain mean-reverting structure to avoid over-extrapolating temporary volatility bursts.

Fourth, VaR systems for ETH should treat the far-left tail differently from the broader lower tail. The stronger performance of kernel-based VaR at the 1% quantile indicates that extreme downside risk is particularly sensitive to nonlinear volatility states and empirical tail behaviour. Risk managers should therefore be cautious about relying only on parametric residual distributions when measuring extreme ETH losses. Empirical residual quantiles, stress scenarios, and state-dependent volatility forecasts should be used as complements to standard parametric VaR models.

Finally, the results have implications for the design of crypto-market disclosure and stress-testing frameworks. Major institutional and protocol events, such as ETF-related repricing, staking-related developments, and large market drawdowns, can alter the local relationship between current states and future volatility. A useful stress-testing framework should therefore incorporate event-sensitive volatility states rather than treating ETH as if it follows a stable unconditional volatility process. This is especially important as Ethereum becomes more integrated with traditional financial markets.

The paper has several limitations that point to future research. The analysis focuses on daily ETH/USD returns in the post-Merge period. Future work could extend the framework to intraday data, option-implied volatility, staking yields, decentralised-exchange liquidity, futures funding rates, and cross-asset spillovers from Bitcoin or equity markets. Another useful extension would be to compare kernel regression with modern machine-learning methods, such as random forests, gradient boosting, neural networks, or Gaussian processes, while preserving the interpretability of the volatility surface. Finally, future research could use formal backtesting tests for VaR exceedances and expected shortfall to examine whether the tail-risk gains documented here translate into stronger regulatory risk measures.

Overall, the evidence shows that post-Merge Ethereum volatility is persistent, asymmetric, heavy-tailed, and nonlinear. Standard GARCH remains a valuable benchmark, particularly at longer horizons, but it does not fully capture the local state dependence of near-term ETH risk. Kernel regression performs best when it combines an economically meaningful volatility state with a flexible nonparametric forecasting map. The broader message is therefore not that nonparametric methods always dominate parametric models. Rather, the main conclusion is that Ethereum volatility forecasting benefits from a hybrid approach: parametric structure provides discipline, while nonparametric local learning captures the regime-dependent features of crypto-market risk.

# References


Dirk G. Baur, KiHoon Hong, and Adrian D. Lee. Bitcoin: Medium of exchange or speculative assets? *Journal of International Financial Markets, Institutions and Money*, 54:177–189, 2018. doi: 10.1016/j.intfin.2017.12.004.

Rainer Böhme, Nicolas Christin, Benjamin Edelman, and Tyler Moore. Bitcoin: Economics, technology, and governance. *Journal of Economic Perspectives*, 29(2):213–238, 2015. doi: 10.1257/jep.29.2.213.

Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, 1986a. doi: 10.1016/0304-4076(86)90063-1.

Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, 1986b. doi: 10.1016/0304-4076(86)90063-1.

Nicola Borri. Conditional tail-risk in cryptocurrency markets. *Journal of Empirical Finance*, 50:1–19, 2019. doi: 10.1016/j.jempfin.2018.11.002.

Elie Bouri, Peter Molnár, Georges Azzi, David Roubaud, and Lars Ivar Hagfors. On the hedge and safe haven properties of bitcoin: Is it really more than a diversifier? *Finance Research Letters*, 20:192–198, 2017. doi: 10.1016/j.frl.2016.09.025.

Alexander Brauneis and Roland Mestel. Price discovery of cryptocurrencies: Bitcoin and beyond. *Economics Letters*, 165:58–61, 2018. doi: 10.1016/j.econlet.2018.02.001.

Guglielmo Maria Caporale, Luis Gil-Alana, and Alex Plastun. Persistence in the cryptocurrency market. *Research in International Business and Finance*, 46:141–148, 2018. doi: 10.1016/j.ribaf.2018.01.002.

Eng-Tuck Cheah and John Fry. Speculative bubbles in bitcoin markets? an empirical investigation into the fundamental value of bitcoin. *Economics Letters*, 130:32–36, 2015. doi: 10.1016/j.econlet.2015.02.029.

Peter F. Christoffersen. Evaluating interval forecasts. *International Economic Review*, 39(4):841–862, 1998. doi: 10.2307/2527341.

Gerald P. Dwyer. The economics of bitcoin and similar private digital currencies. *Journal of Financial Stability*, 17:81–91, 2015. doi: 10.1016/j.jfs.2014.11.006.

Anne Haubo Dyhrberg. Bitcoin, gold and the dollar: A garch volatility analysis. *Finance Research Letters*, 16: 85–92, 2016. doi: 10.1016/j.frl.2015.10.008.

Robert F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4):987–1007, 1982. doi: 10.2307/1912773.

Robert F. Engle and Simone Manganelli. Caviar: Conditional autoregressive value at risk by regression quantiles. *Journal of Business and Economic Statistics*, 22(4):367–381, 2004. doi: 10.1198/073500104000000370.

Robert F. Engle and Victor K. Ng. Measuring and testing the impact of news on volatility. *The Journal of Finance*, 48(5):1749–1778, 1993a. doi: 10.1111/j.1540-6261.1993.tb05127.x.

Robert F. Engle and Victor K. Ng. Measuring and testing the impact of news on volatility. *Journal of Finance*, 48(5):1749–1778, 1993b. doi: 10.1111/j.1540-6261.1993.tb05127.x.

Konstantinos Gkillas and Paraskevi Katsiampa. An application of extreme value theory to cryptocurrencies. *Economics Letters*, 164:109–111, 2018. doi: 10.1016/j.econlet.2018.01.020.

Lawrence R. Glosten, Ravi Jagannathan, and David E. Runkle. On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, 48(5):1779–1801, 1993. doi: 10.1111/j.1540-6261.1993.tb05128.x.

Wolfgang Härdle. *Applied Nonparametric Regression*. Cambridge University Press, Cambridge, 1990.

Steven L. Heston and Saikat Nandi. A closed-form GARCH option valuation model. *The Review of Financial Studies*, 13(3):585–625, 2000a. doi: 10.1093/rfs/13.3.585.

Steven L. Heston and Saikat Nandi. A closed-form garch option valuation model. *Review of Financial Studies*, 13(3):585–625, 2000b. doi: 10.1093/rfs/13.3.585.

J.P. Morgan/Reuters. Riskmetrics technical document. Technical report, J.P. Morgan, 1996.

Paraskevi Katsiampa. Volatility estimation for bitcoin: A comparison of garch models. *Economics Letters*, 158: 3–6, 2017. doi: 10.1016/j.econlet.2017.06.023.

Paraskevi Katsiampa, Shaen Corbet, and Brian Lucey. Volatility spillover effects in leading cryptocurrencies: A bekk-mgarch analysis. *Finance Research Letters*, 29:68–74, 2019. doi: 10.1016/j.frl.2019.03.009.

Tony Klein, Hien Pham Thu, and Thomas Walther. Bitcoin is not the new gold: A comparison of volatility, correlation, and portfolio performance. *International Review of Financial Analysis*, 59:105–116, 2018. doi: 10.1016/j.irfa.2018.07.010.

Jussi Klemelä. Volatility prediction using kernel regression. Working paper, 2017a. URL http://jklm.fi/art/volapred/volapred.pdf. Accessed 27 February 2020.

Jussi Klemelä. Volatility prediction using kernel regression, 2017b. URL https://jklm.fi/art/volapred/volapred.pdf. Working paper.

Roger Koenker and Gilbert Bassett. Regression quantiles. *Econometrica*, 46(1):33–50, 1978. doi: 10.2307/1913643.

Dimitrios Koutmos. Return and volatility spillovers among cryptocurrencies. *Economics Letters*, 173:122–127, 2018. doi: 10.1016/j.econlet.2018.10.004.

Paul H. Kupiec. Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives*, 3(2): 73–84, 1995. doi: 10.3905/jod.1995.407942.

Yukun Liu and Aleh Tsyvinski. Risks and returns of cryptocurrency. *Review of Financial Studies*, 34(6):2689–2727, 2021. doi: 10.1093/rfs/hhaa113.

Alexander J. McNeil and Rüdiger Frey. Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of Empirical Finance*, 7(3–4):271–300, 2000. doi: 10.1016/S0927-5398(00)00012-8.

Saralees Nadarajah and Jeffrey Chu. On the inefficiency of bitcoin. *Economics Letters*, 150:6–9, 2017. doi: 10.1016/j.econlet.2016.10.033.

Elizbar A. Nadaraya. On estimating regression. *Theory of Probability and Its Applications*, 9(1):141–142, 1964. doi: 10.1137/1109020.

Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system, 2008. URL https://bitcoin.org/bitcoin.pdf. White paper.

Daniel B. Nelson. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2):347–370, 1991. doi: 10.2307/2938260.

Andrew Phillip, Jennifer S. K. Chan, and Shelton Peiris. A new look at cryptocurrencies. *Economics Letters*, 163:6–9, 2018. doi: 10.1016/j.econlet.2017.11.020.

Sera Şanlı, Mehmet Balcılar, and Mehmet Özmen. Predicting the volatility of Bitcoin returns based on kernel regression. *Annals of Operations Research*, 352(1):505–542, 2025. doi: 10.1007/s10479-023-05490-4.

Sera Sanli, Mehmet Balcilar, and Mehmet Ozmen. Predicting the volatility of bitcoin returns based on kernel regression. *Annals of Operations Research*, 352(3):505–542, 2025. doi: 10.1007/s10479-023-05490-4.

Andrew Urquhart. The inefficiency of bitcoin. *Economics Letters*, 148:80–82, 2016. doi: 10.1016/j.econlet.2016.09.019.

Geoffrey S. Watson. Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, 26(4): 359–372, 1964.

Wang Chun Wei. Liquidity and market efficiency in cryptocurrencies. *Economics Letters*, 168:21–24, 2018. doi: 10.1016/j.econlet.2018.04.003.

David Yermack. Is bitcoin a real currency? an economic appraisal. In *Handbook of Digital Currency*, pages 31–43. Academic Press, 2015. doi: 10.1016/B978-0-12-802117-0.00002-3.