University of Hohenheim

Time Series Approaches to Cocoa Price Risk

Forecasting and Hedging Applications

Master Thesis

submitted in partial fulfillment of the requirements for the degree of

Master of Agricultural Sciences - Advisory and Innovation Services in agri-food systems

Author: Felix Köhler

First Supervisor: Prof. Dr. Thomas Dimpfl

Second Supervisor: Prof. Johannes Bleher

Contents

Chapter 1

Introduction

Volatility forecasting is fundamental to financial risk management, portfolio allocation, and derivative pricing (Poon and Granger 2003). The literature has developed multiple approaches, including historical rolling averages, implied volatility from options prices, stochastic volatility models, and various parametric specifications. Among these, two modeling paradigms have become dominant. The first treats volatility as a latent process that must be inferred from returns. The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model of Bollerslev (1986) represents the canonical approach in this paradigm, modeling conditional variance as a function of past shocks and past variances. The second paradigm, enabled by the availability of high-frequency data, treats volatility as directly observable through realized measures computed from intraday returns (Andersen et al. 2001). Within this paradigm, the Heterogeneous Autoregressive (HAR) model of Corsi (2009) has emerged as the dominant forecasting framework, offering a parsimonious structure that captures the long-memory properties of volatility while remaining simple to estimate.

The majority of volatility forecasting research focuses on conditional mean forecasts: the expected level of future volatility given current information. For many applications, however, the mean provides incomplete guidance. Risk managers, portfolio allocators, and hedgers require information about the tails of the volatility distribution—specifically, bounds that volatility is unlikely to exceed. Value-at-Risk calculations, stress testing, and margin requirements all depend on quantile estimates rather than mean forecasts (Christoffersen 1998). Quantile regression, introduced by Koenker and Bassett (1978), provides a natural framework for estimating conditional quantiles directly, without distributional assumptions.

Recent work has combined these approaches. Haugom et al. (2016) introduced HAR-QREG, applying quantile regression to the HAR framework for direct Value-at-Risk estimation. Subsequent studies have extended this approach to equity markets (Lyócsa and Molnár 2021; Huang et al. 2022) and precious metals (Song et al. 2025). However, agricultural commodities remain understudied. While HAR models have been applied to agricultural futures (Degiannakis et al. 2022; Alfeus and Nikitopoulos 2022), no study applies HAR-based quantile regression to this asset class.

Cocoa futures present a particularly relevant case. The cocoa market is characterized by high geographic concentration of production, vulnerability to weather and disease shocks, and pronounced price volatility (Rogna and Tillie 2025). Between 2022 and 2024, international cocoa prices approximately quadrupled - an increase unprecedented in the commodity's modern trading history. For chocolate manufacturers, who commit to purchases months in advance through futures contracts, accurate volatility forecasts at long horizons are essential for hedging decisions and procurement planning. Yet the volatility forecasting literature provides limited guidance at such horizons: existing HAR-QR studies extend only to approximately one month (Lyócsa and Molnár 2021), and Poon and Granger (2003) conclude that forecast accuracy degrades substantially beyond six months.

At long forecast horizons, the autoregressive signal from lagged volatility fades, motivating the search for additional predictors. Climate variables offer a potential source of information that operates on precisely these longer timescales. Bouri et al. (2021) demonstrate that the El Niño–Southern Oscillation (ENSO) index improves oil volatility forecasts at horizons of two to four years. Bonato et al. (2023) extend this finding to agricultural commodities, documenting that ENSO predictive value strengthens at longer horizons. For cocoa, whose production is concentrated in West Africa and sensitive to rainfall patterns affected by ENSO phases (Ubilava 2018), climate augmentation may prove particularly valuable.

This thesis addresses the following research question:

> Can HAR-based quantile regression models forecast cocoa futures volatility at horizons relevant for procurement and hedging (1 month to 12 months), and does the ENSO climate index improve tail-risk forecasts at long horizons?

We address this question by estimating HAR-QR models at multiple quantiles ($\tau \in \{0.05, 0.25, 0.50, 0.75, 0.95\}$) and horizons (1 day, 1 week, 1 month, 3 months, 6 months, 12 months). At each horizon, we compare out-of-sample forecast accuracy against standard benchmarks: historical rolling volatility, GARCH(1,1), and HAR estimated by ordinary least squares. For long horizons, we test whether adding ENSO improves forecasts, with particular focus on the upper quantile ($\tau = 0.95$) relevant for risk management applications.

The thesis makes three contributions to the literature. First, it applies HAR-QR to an agricultural commodity. While HAR models have been applied to cocoa within broad commodity surveys (Alfeus and Nikitopoulos 2022; Bonato et al. 2024), and HAR-QR has been applied to equities and precious metals, no study has combined these approaches for agricultural commodities.

Second, the thesis extends HAR-QR to horizons beyond existing applications. The literature currently provides HAR-QR evidence only for horizons up to approximately one month (Lyócsa and Molnár 2021). This research tests whether quantile forecasts remain useful at horizons of 6 to 12 months, where industrial hedgers require guidance but the literature offers none.

Third, the thesis incorporates climate variables into a quantile regression framework for commodity volatility. While Bouri et al. (2021) demonstrate that ENSO improves mean volatility forecasts for oil, no study has tested whether ENSO enhances quantile forecasts for agricultural commodities. This extension is particularly relevant for cocoa, given the documented sensitivity of tropical agricultural production to ENSO phases.

The remainder of the thesis is organized as follows. Chapter 2 reviews the literature on volatility measurement, HAR models, quantile regression in finance, and climate-commodity linkages, establishing the research gap this thesis addresses. Chapter **??** specifies the HAR-QR model, the multi-horizon projection approach, and the evaluation framework. Chapter **??** describes the ICE London cocoa futures data and validates the daily volatility proxy against realized volatility from intraday data. Chapter **??** presents empirical results organized by forecast horizon. Chapter **??** interprets findings in the context of risk management and procurement applications. Chapter **??** summarizes key results and directions for future research.

Chapter 2

Literature Review

This chapter reviews the theoretical foundations underlying volatility forecasting for commodity markets. We begin with background on the cocoa market and the risk management context that motivates our focus on tail-risk forecasting. We then trace three strands of literature: volatility measurement, the Heterogeneous Autoregressive (HAR) model, and quantile regression methods. The intersection of these strands—HAR-based quantile regression for commodities—represents the gap this thesis addresses.

2.1 The Cocoa Market and Procurement Risk

Cocoa is produced almost exclusively in tropical regions. West Africa dominates global supply: Côte d'Ivoire and Ghana together account for over half of world production (Rogna and Tillie 2025). Production is geographically concentrated, dependent on rainfall patterns, and vulnerable to plant diseases - most recently the Cocoa Swollen Shoot Virus Disease (CSSVD), which contributed to the 2022–2024 price surge. This concentration creates structural volatility: disruptions in a single region can move global prices substantially.

Chocolate manufacturers hedge their exposure by buying futures contracts on exchanges such as ICE London. A futures contract specifies delivery of a standardized quantity (10 metric tonnes for London contracts) at a future date for a price agreed today. The futures price reflects spot prices, storage costs, interest rates, and the *convenience yield*—the benefit from holding physical inventory (Hull 2012). When convenience yield is high (typically during supply uncertainty), futures trade below spot (backwardation); when storage costs dominate, futures trade above spot (contango).

For a manufacturer hedging with futures, three distinct risks arise. First, *spot price volatility*: the risk that the price of physical cocoa moves adversely before purchase. Second, *futures price volatility*: the risk that the value of the hedging instrument itself fluctuates, triggering margin calls and complicating cash-flow planning. Third, *basis risk*: the risk that spot and futures prices diverge, causing the hedge to under- or overcompensate for spot price movements.

This thesis focuses on the second - futures price volatility - for three reasons. Manufacturers who hedge months ahead hold futures positions throughout the procurement window; volatility in the futures price directly determines margin requirements, option premiums (if options are used for additional protection), and budgeting uncertainty. Spot price risk is largely transferred to the exchange through the futures position itself. Basis risk, while relevant at delivery, is a secondary concern during the hedging window, and its analysis requires simultaneous modeling of spot and futures prices, which lies outside the scope of a single-instrument volatility study.

The procurement question is therefore: "how volatile could the futures price become over the hedging horizon?" This is a tail-risk question- the 95th percentile of the volatility distribution, not the mean. A mean forecast that says "volatility will likely be moderate" provides little comfort if the upper tail includes a repeat of 2022–2024 conditions. This tail-risk framing motivates quantile regression as the appropriate estimation method.

2.2 Volatility Measurement

A foundational challenge in volatility research is that volatility is latent: unlike prices, which are directly observable, volatility must be estimated from price movements (Poon and Granger 2003). The literature has developed progressively sophisticated estimators, each offering different trade-offs between data requirements, statistical efficiency, and practical applicability.

2.2.1 OHLC-Based Volatility Estimators

Daily Open-High-Low-Close (OHLC) data provide a practical foundation for volatility estimation, offering broad historical coverage and standardized availability across markets. This thesis employs OHLC-based volatility measures, and this subsection reviews the relevant estimators.

The simplest approach uses closing prices only. The close-to-close variance estimator is

$$\hat{\sigma}^2_{\mathrm{CC}} = \frac{1}{n-1}\sum_{i=1}^{n}(r_i - \bar{r})^2, \tag{2.1}$$

where $r_i = \ln(C_i/C_{i-1})$ denotes the log return on day i, C_i is the closing price, $\bar{r}$ is the sample mean of returns, and n is the number of observations. While unbiased, this estimator discards valuable information contained in intraday price movements. Poon and Granger (2003) note that close-to-close estimators remain widely used despite their inefficiency, partly due to data availability constraints.

Parkinson (1980) demonstrated that the daily high-low range contains substantial information about volatility, proposing an estimator with theoretical efficiency gains of approximately five times over close-to-close methods. Rogers and Satchell (1991) extended this work to allow for non-zero drift, addressing a key limitation of Parkinson's original formulation. However, both estimators assume continuous trading- an assumption violated by commodity futures markets, which experience regular overnight closures and can gap significantly at the open.

Yang and Zhang (2000) synthesized these approaches into an estimator that is unbiased, drift-independent, and robust to opening jumps. Their key insight was to estimate overnight and trading-session variance separately, combining them optimally. The Yang-Zhang estimator is

$$\hat{\sigma}^2_{\mathrm{YZ}} = \hat{\sigma}^2_o + k\,\hat{\sigma}^2_c + (1-k)\,\hat{\sigma}^2_{\mathrm{RS}}, \tag{2.2}$$

where $\hat{\sigma}^2_o$ is the overnight variance (open-to-previous-close), $\hat{\sigma}^2_c$ is the close-to-open variance, and $\hat{\sigma}^2_{\mathrm{RS}}$ is the Rogers-Satchell intraday variance component:

$$\hat{\sigma}^2_{\mathrm{RS}} = \frac{1}{n}\sum_{i=1}^{n}\Big[u_i(u_i - c_i) + d_i(d_i - c_i)\Big], \tag{2.3}$$

with $u_i = \ln(H_i/O_i)$, $d_i = \ln(L_i/O_i)$, and $c_i = \ln(C_i/O_i)$ denoting the normalized high, low, and close relative to the opening price. The weighting factor $k = 0.34/(1.34+(n+1)/(n-1))$ minimizes variance under the assumption of zero drift. The Yang-Zhang estimator achieves efficiency gains of 7–8 times over close-to-close methods, making it particularly valuable for commodity markets where overnight information (weather reports, inventory data, policy announcements) frequently causes opening gaps.

While high-frequency intraday data enable "realized volatility" estimators with superior statistical properties (Andersen et al. 2001), such data are often unavailable for commodity markets over extended sample periods. Clements et al. (2024) provide evidence that HAR models applied to daily OHLC-based volatility achieve forecasting performance comparable to realized volatility at medium and long horizons, validating the use of daily data for volatility modeling.

2.2.2 Conditional Volatility: The GARCH Framework

The OHLC estimators reviewed above treat volatility as an ex-post measure computed from observed prices. An alternative paradigm models volatility as a latent process that evolves over time. The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model of Bollerslev (1986) is the canonical specification in this paradigm.

The GARCH(1,1) model defines the conditional variance as

$$\sigma_t^2 = \omega + \alpha\,\varepsilon_{t-1}^2 + \beta\,\sigma_{t-1}^2, \tag{2.4}$$

where σ_t^2 is the conditional variance at time t, $\omega > 0$ is a constant, $\varepsilon_{t-1} = r_{t-1} - \mu$ is the mean-adjusted return innovation, $\alpha \geq 0$ governs the response to new shocks, and $\beta \geq 0$ captures persistence from past variance. The condition $\alpha + \beta < 1$ ensures stationarity, with the unconditional variance given by $\sigma^2 = \omega/(1 - \alpha - \beta)$.

The model captures two key stylized facts of financial returns: volatility clustering (large shocks are followed by large shocks) and mean reversion (variance returns to its long-run level after a shock). The sum $\alpha + \beta$ measures persistence: values close to one imply slow decay of volatility shocks, a pattern commonly observed in commodity markets.

Despite its simplicity, the GARCH(1,1) has proven remarkably difficult to beat in forecasting comparisons. Hansen and Lunde (2005) evaluate 330 volatility models on exchange rate data and conclude that "nothing beats a GARCH(1,1)." While subsequent work has identified settings where more complex specifications add value, this finding establishes GARCH(1,1) as the natural benchmark against which more elaborate models must demonstrate improvement. This thesis uses GARCH(1,1) in precisely this role: as a well-understood, parsimonious benchmark for evaluating the forecast accuracy of HAR-QR models.

2.3 The Heterogeneous Autoregressive Model

The HAR model has become the dominant framework for volatility forecasting since its introduction by Corsi (2009). Its success stems from a parsimonious structure that captures the long-memory properties of volatility without the estimation difficulties of fractionally integrated models.

2.3.1 Theoretical Foundations

The HAR model builds on the *Heterogeneous Market Hypothesis* of Müller et al. (1997). Müller et al. documented a striking empirical asymmetry: coarse-grained volatility (e.g., weekly) predicts fine-grained volatility (e.g., daily) more strongly than the reverse. This pattern cannot be explained by standard GARCH models, which treat volatility symmetrically across horizons.

The heterogeneous market hypothesis attributes this asymmetry to the presence of traders with different investment horizons. Intraday speculators, swing traders, and long-term institutional investors each respond to volatility at their characteristic frequency. When long-horizon traders adjust positions, they generate volatility that cascades down to shorter horizons. Müller et al. proposed the HARCH model to capture this cascade:

$$\sigma_t^2 = c_0 + \sum_{j=1}^{n} c_j \left(\sum_{i=1}^{j} r_{t-i} \right)^2, \tag{2.5}$$

where σ_t^2 is the conditional variance at time t, c_0 is a constant, c_j are coefficients for each aggregation horizon j, and r_{t-i} denotes the return i periods ago. The model conditions current variance on aggregated returns over multiple horizons, with longer aggregation windows capturing the influence of longer-horizon traders.

For agricultural commodities, Alfeus and Nikitopoulos (2022) find that "in silver, palladium, rice, and cocoa markets, monthly volatility matters the most." This suggests that longer-term participants- including producers hedging crop cycles and manufacturers managing procurement- exert particular influence on commodity volatility dynamics. Haugom et al. (2016) provide further evidence using quantile regression, finding that "the effect from the monthly volatility component is strongest for all assets when predicting conditional tails"- precisely the quantiles relevant for risk management.

2.3.2 The HAR-RV Specification

Corsi (2009) simplified the HARCH framework into an additive model with three volatility components:

$$\sigma_{t+1} = \beta_0 + \beta_d \sigma_t^{(d)} + \beta_w \sigma_t^{(w)} + \beta_m \sigma_t^{(m)} + \varepsilon_{t+1}, \tag{2.6}$$

where σ_{t+1} is the volatility to be forecast, β_0 is a constant intercept, and ε_{t+1} is the error term. The three volatility components are: daily volatility $\sigma_t^{(d)} = \sigma_t$; weekly volatility $\sigma_t^{(w)} = \frac{1}{5}\sum_{i=0}^{4} \sigma_{t-i}$, the average over the past 5 trading days; and monthly volatility $\sigma_t^{(m)} = \frac{1}{22}\sum_{i=0}^{21} \sigma_{t-i}$, the average over the past 22 trading days. The coefficients β_d, β_w, and β_m measure each horizon's contribution to future volatility.

Despite its simplicity- the model is a restricted AR(22) estimable by OLS- the HAR reproduces the hyperbolic decay of volatility autocorrelations that characterizes long-memory processes. Corsi demonstrated both theoretically and via simulation that an additive cascade of AR(1) processes with different persistence generates autocorrelation patterns indistinguishable from fractional integration over typical sample lengths.

The HAR model offers several practical advantages documented in the literature. First, estimation requires only OLS regression, avoiding the numerical optimization of GARCH or the specialized methods of fractional integration. Second, coefficients are interpretable: each measures the contribution of a specific trading horizon to future volatility. Third, the framework is extensible- additional predictors can be incorporated as regressors.

2.3.3 Extensions and Empirical Performance

Since Corsi's seminal contribution, the HAR framework has been extended in multiple directions. Andersen et al. (2007) decomposed realized volatility into continuous and jump components, finding that jumps add little to out-of-sample forecast accuracy. Patton and Sheppard (2015) separated positive and negative realized semivariance to capture asymmetric volatility responses to gains and losses.

For agricultural commodities specifically, Degiannakis et al. (2022) conducted a comprehensive comparison of HAR variants on five commodities (corn, rice, soybeans, sugar, and wheat). Their key finding is cautionary: "sophisticated HAR-type models are not capable of outperforming the simple HAR in an out-of-sample exercise." This suggests that parsimony is valuable and that additional complexity must be justified by clear forecast improvements. Tian et al. (2017) develop a time-varying HAR model for Chinese agricultural futures, finding that "the jump component is important for forecasting the RV in Chinese agricultural commodity futures markets." Their model allows coefficients to change over time, addressing potential structural instability in commodity markets.

HAR models have been applied to cocoa as part of broader commodity surveys. Alfeus and Nikitopoulos (2022) include cocoa among 22 commodities, finding that HAR outperforms GARCH at short horizons. However, no study focuses specifically on cocoa volatility dynamics or applies the HAR framework to the long horizons relevant for procurement risk management.

2.4 Quantile Regression for Volatility

Standard volatility models- whether GARCH or HAR- estimate the conditional mean of volatility. For risk management applications, however, the mean is often insufficient. Procurement managers, portfolio risk officers, and regulatory capital calculations require statements about extreme outcomes: how bad could volatility get?

2.4.1 Foundations of Quantile Regression

Quantile regression, introduced by Koenker and Bassett (1978), estimates conditional quantiles rather than conditional means. Koenker and Bassett developed quantile regression specifically to address the sensitivity of least squares to non-Gaussian errors, showing that their estimators "have comparable efficiency to least squares for Gaussian linear models while substantially out-performing the least-squares estimator over a wide class of non-Gaussian error distributions." For any probability level $\tau \in (0, 1)$, quantile regression estimates the τ-th conditional quantile $Q_\tau(Y|X) = X'\beta(\tau)$, where Y is the dependent variable (here, volatility), X is a vector of predictors, and $\beta(\tau)$ is a coefficient vector that varies with the quantile of interest. The parameters are obtained by minimizing the asymmetric "check" loss function:

$$\hat{\beta}(\tau) = \arg\min_{\beta} \sum_{t=1}^{T} \rho_\tau(Y_t - X_t'\beta), \tag{2.7}$$

where T is the sample size and $\rho_\tau(u) = u(\tau - \mathbf{1}_{u<0})$ is the check function. Here $\mathbf{1}_{u<0}$ is an indicator that equals 1 when $u < 0$ and 0 otherwise. This function penalizes positive residuals (underpredictions) with weight τ and negative residuals (overpredictions) with weight $1 - \tau$. For $\tau = 0.95$, underpredictions are penalized 19 times more heavily than overpredictions, pushing the fitted quantile toward the upper tail.

The key innovation is that each quantile has its own parameter vector, allowing the relationship between Y and X to differ across the distribution. At $\tau = 0.95$, quantile regression estimates the conditional 95th percentile- the level exceeded only 5% of the time. This maps directly to Value-at-Risk (VaR), making quantile regression natural for risk management applications.

2.4.2 Quantile Regression in Finance

The application of quantile regression to volatility forecasting has grown substantially. Taylor (2008) pioneered its use for VaR estimation, demonstrating that quantile regression provides robust forecasts even when the volatility distribution deviates from normality. Haugom et al. (2016) introduced HAR-QREG, combining the multi-horizon HAR structure with quantile regression for direct VaR estimation.

For equity markets, Lyócsa and Molnár (2021) developed the HAR-CSQR (complete subset quantile regression) model, finding significant improvements over benchmark HAR models at horizons up to 22 trading days. Huang et al. (2022) applied HAR-QREG to Chinese stock indices, confirming that "the model has better results for out-of-sample VaR forecasting" compared to GARCH alternatives. More recently, Song et al. (2025) combined HAR-QR with machine learning models for gold futures, finding that HAR-QR parameters significantly improve volatility forecasts across multiple evaluation criteria.

2.4.3 Literature Gap: HAR-QR for Commodities

Despite the success of HAR-QR in equity markets, its application to commodities remains limited. Degiannakis et al. (2022) test HAR models for agricultural commodities but focus exclusively on mean forecasts. No study applies HAR-QR to any agricultural commodity.

Furthermore, existing HAR-QR studies are confined to short horizons. Lyócsa and Molnár (2021) extend to 22 trading days (approximately one month)- the longest horizon in the literature. Poon and Granger (2003) review 93 forecasting studies and conclude that "for forecast horizons that are longer than 6 months, a simple historical method using low frequency data over a period at least as long as the forecast horizon works best." Whether quantile forecasts retain predictive value at procurement-relevant horizons (6–12 months) remains untested.

2.5 Climate Variables and Commodity Volatility

At long forecast horizons, the autoregressive signal from lagged volatility fades. Climate variables offer a potential source of predictive information that operates on precisely these longer timescales.

2.5.1 ENSO and Agricultural Commodities

The El Niño–Southern Oscillation (ENSO) is a periodic fluctuation in Pacific sea surface temperatures that affects weather patterns globally. Ubilava (2018) examines the relationship between ENSO and 43 commodity prices, finding that "the prices of tropically-grown beverages, including coffee varieties and cocoa, are also affected by SST anomalies." Importantly, Ubilava documents "more amplified price responses during El Niño events, and at the onset of the ENSO cycle"- suggesting that ENSO may be particularly relevant for volatility rather than just price levels.

For cocoa specifically, the transmission mechanism operates through West African rainfall patterns. ENSO phases affect precipitation during the growing season, influencing yields in subsequent harvests. With main crop (October–March) and mid-crop (May–August) seasons, ENSO effects materialize 6–12 months after the climate signal-aligning with procurement planning horizons.

2.5.2 ENSO and Volatility Forecasting

Recent research has extended ENSO analysis from price levels to volatility. Bouri et al. (2021) find that ENSO improves oil volatility forecasts at horizons of 2–4 years-far beyond where lagged volatility alone provides value. Their methodology uses current ENSO values to predict future volatility, with the forecast horizon providing the transmission lag.

Bonato et al. (2023) extend this approach to 16 agricultural commodities, documenting that "there is a general tendency that the evidence of predictive value of El Niño and La Niña events strengthens at the longer term forecast horizons." This finding is crucial: ENSO adds value precisely where standard autoregressive models lose power. Beyond ENSO specifically, Guo et al. (2025) demonstrate that a broader climate change concern index significantly improves commodity futures volatility forecasts across energy, metal, and agricultural markets, with models incorporating climate risk "outperforming traditional ones in economic value."

References

Alfeus, Mesias and Christina Sklibosios Nikitopoulos (2022). "Forecasting Volatility in Commodity Markets with Long-Memory Models". In: *Journal of Commodity Markets* 28, p. 100248. DOI: 10.1016/j.jcomm.2022.100248.

Andersen, Torben G., Tim Bollerslev, and Francis X. Diebold (2007). "Roughing It Up: Including Jump Components in the Measurement, Modeling, and Forecasting of Return Volatility". In: *Review of Economics and Statistics* 89.4, pp. 701–720. DOI: 10.1162/rest.89.4.701.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Paul Labys (2001). "The Distribution of Realized Exchange Rate Volatility". In: *Journal of the American Statistical Association* 96.453, pp. 42–55. DOI: 10.1198/016214501750332965.

Bollerslev, Tim (1986). "Generalized Autoregressive Conditional Heteroskedasticity". In: *Journal of Econometrics* 31.3, pp. 307–327. DOI: 10.1016/0304-4076(86)90063-1.

Bonato, Matteo, Oğuzhan Çepni, Rangan Gupta, and Christian Pierdzioch (2023). "El Niño, La Niña, and Forecastability of the Realized Variance of Agricultural Commodity Prices: Evidence from a Machine Learning Approach". In: *Journal of Forecasting* 42.4, pp. 785–801. DOI: 10.1002/for.2914.

— (2024). "Forecasting the Realized Volatility of Agricultural Commodity Prices: Does Sentiment Matter?" In: *Journal of Forecasting* 43.6, pp. 2088–2125. DOI: 10.1002/for.3106.

Bouri, Elie, Rangan Gupta, Christian Pierdzioch, and Afees A. Salisu (2021). "El Niño and Forecastability of Oil-Price Realized Volatility". In: *Theoretical and Applied Climatology* 144, pp. 1173–1180. DOI: 10.1007/s00704-021-03569-1.

Christoffersen, Peter F. (1998). "Evaluating Interval Forecasts". In: *International Economic Review* 39.4, pp. 841–862. DOI: 10.2307/2527341.

Clements, Adam, Stan Hurn, and Vladimir Volkov (2024). "Harvesting HAR-X Volatility Forecasts Using OHLC Data". In: *Journal of Empirical Finance* 79, p. 101555. DOI: 10.1016/j.jempfin.2024.101555.

Corsi, Fulvio (2009). "A Simple Approximate Long-Memory Model of Realized Volatility". In: *Journal of Financial Econometrics* 7.2, pp. 174–196. DOI: 10.1093/jjfinec/nbp001.

Degiannakis, Stavros, George Filis, Tony Klein, and Thomas Walther (2022). "Forecasting Realized Volatility of Agricultural Commodities". In: *International Journal of Forecasting* 38.1, pp. 74–96. DOI: 10.1016/j.ijforecast.2019.08.011.

Guo, Yangli, Pei Peng, Ling Zhou, and Yusui Tang (2025). "Forecasting Volatility in Commodity Markets with Climate Risk". In: *Finance Research Letters* 71, p. 106357. DOI: 10.1016/j.frl.2024.106357.

Hansen, Peter R. and Asger Lunde (2005). "A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1)?" In: *Journal of Applied Econometrics* 20.7, pp. 873–889. DOI: 10.1002/jae.800.

Haugom, Erik, Rina Ray, Carl J. Ullrich, Steinar Veka, and Sjur Westgaard (2016). "A Parsimonious Quantile Regression Model to Forecast Day-Ahead Value-at-Risk". In: *Finance Research Letters* 16, pp. 196–207. DOI: 10.1016/j.frl.2015.12.006.

Huang, Jiefei, Yang Xu, and Yuping Song (2022). "A High-Frequency Approach to VaR Measures and Forecasts Based on the HAR-QREG Model with Jumps". In: *Physica A: Statistical Mechanics and its Applications* 608, p. 128253. DOI: 10.1016/j.physa.2022.128253.

Hull, John C. (2012). *Options, Futures, and Other Derivatives.* 8th. Pearson Education.

Koenker, Roger and Gilbert Bassett (1978). "Regression Quantiles". In: *Econometrica* 46.1, pp. 33–50. DOI: 10.2307/1913643.

Lyócsa, Štefan and Peter Molnár (2021). "Stock Market Volatility Forecasting: Do We Need High-Frequency Data?" In: *International Journal of Forecasting* 37.3, pp. 1092–1110. DOI: 10.1016/j.ijforecast.2020.12.001.

Müller, Ulrich A., Michel M. Dacorogna, Rakhal D. Davé, Richard B. Olsen, Olivier V. Pictet, and Jakob E. von Weizsäcker (1997). "Volatilities of Different Time Resolutions – Analyzing the Dynamics of Market Components". In: *Journal of Empirical Finance* 4.2-3, pp. 213–239. DOI: 10.1016/S0927-5398(97)00007-8.

Parkinson, Michael (1980). "The Extreme Value Method for Estimating the Variance of the Rate of Return". In: *Journal of Business* 53.1, pp. 61–65. DOI: 10.1086/296071.

Patton, Andrew J. and Kevin Sheppard (2015). "Good Volatility, Bad Volatility: Signed Jumps and the Persistence of Volatility". In: *Review of Economics and Statistics* 97.3, pp. 683–697. DOI: 10.1162/REST_a_00503.

Poon, Ser-Huang and Clive W. J. Granger (2003). "Forecasting Volatility in Financial Markets: A Review". In: *Journal of Economic Literature* 41.2, pp. 478–539. DOI: 10.1257/002205103765762743.

Rogers, L. C. G. and S. E. Satchell (1991). "Estimating Variance from High, Low, and Closing Prices". In: *Annals of Applied Probability* 1.4, pp. 504–512. DOI: 10.1214/aoap/1177005835.

Rogna, Marco and Pierre Tillie (2025). "An Analysis of Cocoa Market Fundamentals and Price Transmission". In: *Journal of Commodity Markets* 37, p. 100421. DOI: 10.1016/j.jcomm.2024.100421.

Song, Yan, Jiahui Pei, Yufeng Li, and Yuping Song (2025). "Gold Futures Volatility Forecasting Based on HAR-QR and Machine Learning Model". In: *Artificial Intelligence and Human-Computer Interaction*. IOS Press. DOI: 10.3233/FAIA250131.

Taylor, James W. (2008). "Using Exponentially Weighted Quantile Regression to Estimate Value at Risk and Expected Shortfall". In: *Journal of Financial Econometrics* 6.3, pp. 382–406. DOI: 10.1093/jjfinec/nbn007.

Tian, Fengping, Ke Yang, and Langnan Chen (2017). "Realized Volatility Forecasting of Agricultural Commodity Futures Using the HAR Model with Time-Varying Sparsity". In: *International Journal of Forecasting* 33.1, pp. 132–152. DOI: 10.1016/j.ijforecast.2016.08.002.

Ubilava, David (2018). "The Role of El Niño Southern Oscillation in Commodity Price Movement and Predictability". In: *American Journal of Agricultural Economics* 100.1, pp. 239–263. DOI: 10.1093/ajae/aax060.

Yang, Dennis and Qiang Zhang (2000). "Drift-Independent Volatility Estimation Based on High, Low, Open, and Close Prices". In: *Journal of Business* 73.3, pp. 477–491. DOI: 10.1086/209650.