Article

Not peer-reviewed version

Extreme Tail Risk in NOK Exchange Rates: A GARCH–Block-Maxima EVT Assessment

Kristian Mjøen , Haakon Sandvik , Petter Eilif de Lange * , Sjur Westgaard

Posted Date: 4 February 2026

doi: 10.20944/preprints202602.0321.v1

Keywords: GARCH; extreme value theory; block maxima; GEV; value-at-risk; EUR/NOK; USD/NOK; EUR/USD; tail risk; procyclicality

Article

Extreme Tail Risk in NOK Exchange Rates: A GARCH–Block-Maxima EVT Assessment

Kristian Mjøen [1], Haakon Sandvik [1], Petter Eilif de Lange [2,*] and Sjur Westgaard [3]

[1] Faculty of Economics and Management, Department of Economics, NTNU
[2] Faculty of Economics and Management, NTNU School of International Business
[3] Faculty of Economics and Management, Department of Industrial Economics and Technology Management, NTNU
* Correspondence: Petter.e.delange@ntnu.no

Abstract

We study extreme tail risk in EUR/NOK, USD/NOK, and EUR/USD using an integrated GARCH–Block-Maxima EVT (GEV) framework. Layer A filters time-varying volatility via GARCH(1,1) to obtain standardized residuals; Layer B models monthly block maxima with a GEV distribution. To produce one-day VaR, we map daily confidence levels α to block-maxima probabilities $\alpha_B = \alpha^B$ (with $B = 21$ trading days) and rescale by next-day conditional volatility. Using daily data from 2015–2025, we find heavy tails across all pairs, with EUR/NOK exhibiting the heaviest tail (larger $\hat{\xi}$) and USD/NOK the largest extreme-scale. Back tests show reliable 99% VaR across pairs (coverage not rejected; no systematic clustering), while 95% reveals under-coverage for EUR/USD and mild clustering for USD/NOK—consistent with Block-Maxima's emphasis on far tails. The framework is transparent and auditable; for moderate-tail control (95%), a light PoT overlay or t/skew-t innovations can improve calibration. Our results document economically meaningful cross-pair differences in NOK risk and provide a practical pipeline from filtered returns to daily far-tail capital metrics.

Keywords: GARCH; extreme value theory; block maxima; GEV; value-at-risk; EUR/NOK; USD/NOK; EUR/USD; tail risk; procyclicality

1. Introduction

As global financial systems have become increasingly interconnected, risk management has become a crucial component of informed financial decision-making. Among key risk factors, exchange rate fluctuations play a significant role due to their broad market implications. These movements affect international competitiveness, profit margins, capital flow, inflation, and importantly interest rate decisions. Understanding and managing exchange rate risk is therefore essential not only for investors, but also for financial institutions as well as public and corporate policymakers.

The EUR/USD currency pair recorded the highest global turnover as of April 2022 (Bank for International Settlements, 2022). Due to the high liquidity and the scale of the underlying economies, the pair is generally considered more stable than EUR/NOK and USD/NOK. Including EUR/USD as a benchmark enables a clearer assessment of the relative risk associated with the Norwegian krone over the analyzed period.

Exchange rates transmit global shocks to a small, open economy like Norway, shaping inflation, margins, and policy. Conventional models that assume conditional normality understate rare losses, due to volatility clustering and heavy tails. The study investigates differences in tail behavior and volatility patterns across the selected currency pairs. We fuse GARCH filtering with EVT's Block Maxima Method (BMM) to produce daily VaR forecasts for EUR/NOK, USD/NOK, and EUR/USD. We ask the following research questions:

 doi:10.20944/preprints202602.0321.v1

- RQ1: How heavy are NOK tails relative to EUR/USD?
- RQ2: Do GARCH–BMM VaR forecasts pass strict 95%/99% backtests?
- RQ3: What do rolling estimates reveal during stress (e.g., COVID-19)?

Preview of findings: The tail is heaviest for EUR/NOK; extreme scale is largest for USD/NOK; and 99% performance is strong across pairs, with a moderate 95% shortfall for EUR/USD.

2. Literature Review

Financial risk management has been shaped by international standards, notably the Basel framework. Basel II modernized capital rules by linking requirements to banks' internal risk measures (Bank for International Settlements; Cayman Islands Monetary Authority). After the 2007–2009 crisis revealed excess leverage and thin liquidity, Basel III raised the quality/quantity of capital and added buffers and liquidity metrics, with a stronger emphasis on forward-looking risk control (BIS).

Industry practice long leaned on RiskMetrics (J.P. Morgan/Reuters, 1996), which assumes conditional normality. Empirically, however, returns feature fat tails and volatility clustering (Cont, 2001), motivating models that handle both dynamics and extremes. Two complementary strands address this: GARCH for time-varying volatility (Bollerslev, 1986) and Extreme Value Theory (EVT) for tail behavior (Coles, 2001; Embrechts, Klüppelberg & Mikosch, 1997; McNeil, Frey & Embrechts, 2015).

EVT can be implemented via Block Maxima (BMM), modeling block extremes with the GEV distribution, whose shape parameter ξ governs tail heaviness (Coles; Embrechts et al.; McNeil et al.). Alternatively, Peak-over-Threshold (PoT) fits exceedances to the GPD and is often more data-efficient (McNeil, Frey & Embrechts, 2015). A widely used integration is the two-step GARCH–EVT approach: filter returns with GARCH to obtain standardized residuals, then apply EVT to those residuals (McNeil & Frey, 2000). We follow this structure but evaluate BMM within that pipeline, offering a transparent alternative to PoT.

Recent evidence underscores EVT's relevance: Chikobvu & Ndlovu (2023), using monthly BMM and MLE, finds heavy tails and strong VaR performance for ZAR/USD and BTC/USD. For evaluation, we adopt Christoffersen's (1998) backtests of VaR coverage and independence. Overall, the academic literature motivates our choice to combine GARCH for dynamics with EVT for extremes, while explicitly assessing the practicality of BMM in a two-step framework for FX risk.

Our contribution is a clear, auditable GARCH–BMM pipeline for major NOK pairs with out-of-sample VaR validation and a narrative that connects volatility regimes, far-tail dynamics, and backtesting outcomes to real episodes.

3. Data and Data Analysis

Sample and usage. We use daily spot exchange rates for EUR/NOK, USD/NOK, and EUR/USD from 2 Jan 2015 to 21 Feb 2025, see Figures 1, 2, 3. The plots below (both levels and returns) reflect major events affecting markets such as COVID-19 and the outbreak of the war in Ukraine, supporting data plausibility. The full sample underpins backtesting (first 1,000 observations for initial calibration), while 4 Jan 2021–21 Feb 2025 is used for a focused characteristics analysis.

Sources and data quality. NOK-denominated pairs are taken from Norges Bank and EUR/USD from FRED, after rejecting Yahoo Finance due to anomalies (e.g., 20 Mar 2020: USD/NOK 7.73 on Yahoo vs 11.32 at Norges Bank). Central-bank data offered greater consistency and reliability.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 February 2026 doi:10.20944/preprints202602.0321.v1

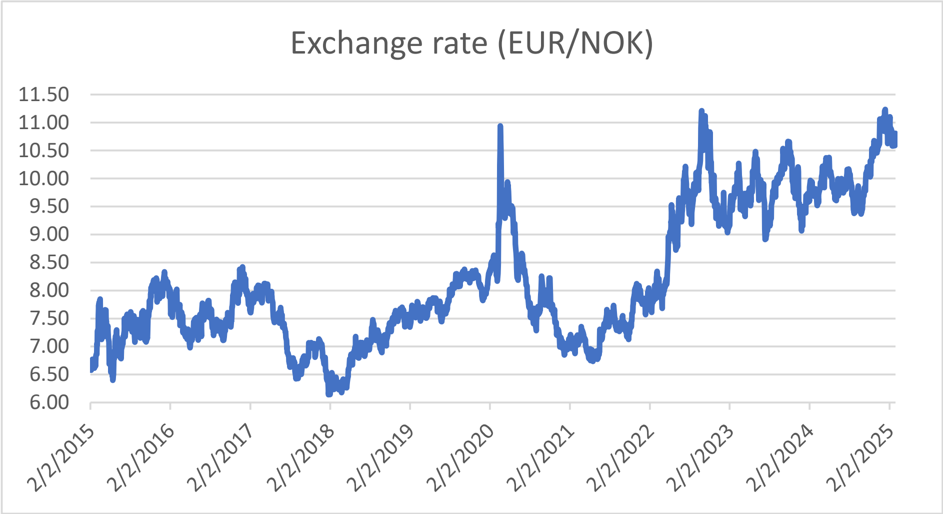


Figure 1. Plot of daily spot exchange rates. Source Federal Reserve Economic Data (FRED).

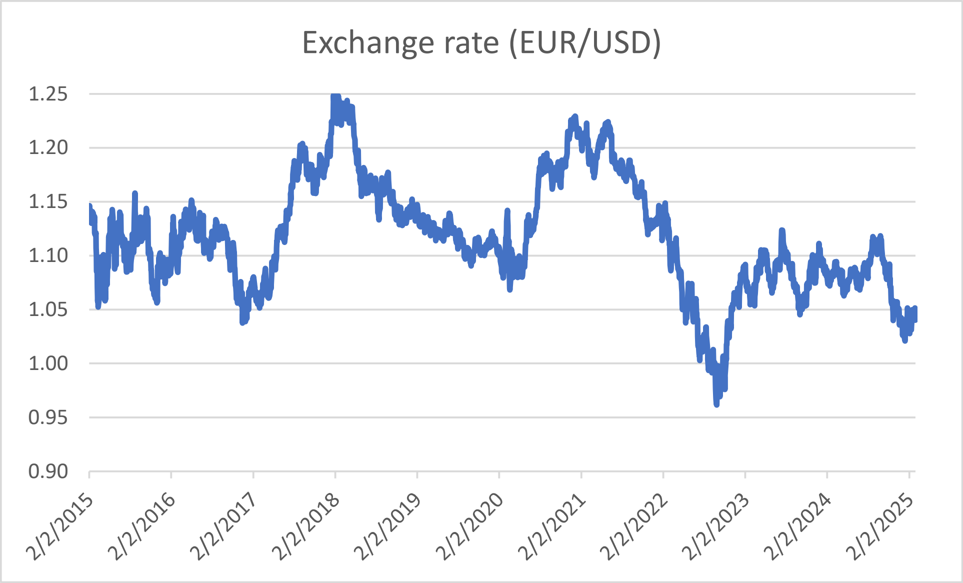


Figure 2. Plot of daily spot exchange rates. Federal Reserve Economic Data (FRED).

 doi:10.20944/preprints202602.0321.v1

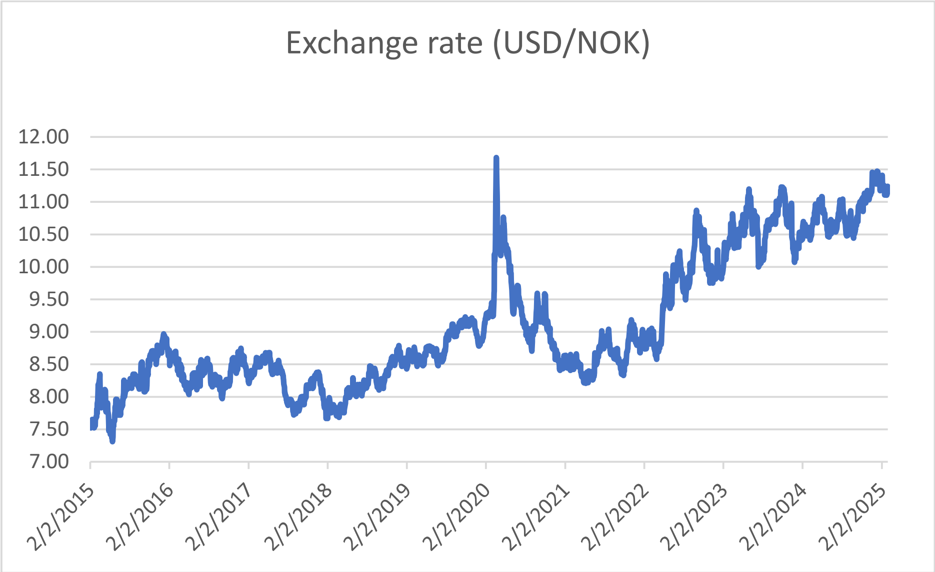


Figure 3. Plot of daily spot exchange rates. Federal Reserve Economic Data (FRED).

Preprocessing. We validate, align business days, handle missing values, and compute log returns for stationarity and compatibility with the modeling framework:

$$r_t = ln\,(\frac{P_t}{P_{t-1}})$$

Here r_t is the log return at time t, and P_t and P_{t-1} are the exchange rates at times t and t-1 respectively.

Exploratory patterns. Returns display volatility clustering (figure 4 and 5): tranquil periods alternate with bursts of high variability. Autocorrelation is negligible in raw returns but significant in squared returns, indicating time-varying conditional variance—classic motivation for GARCH filtering ahead of EVT.

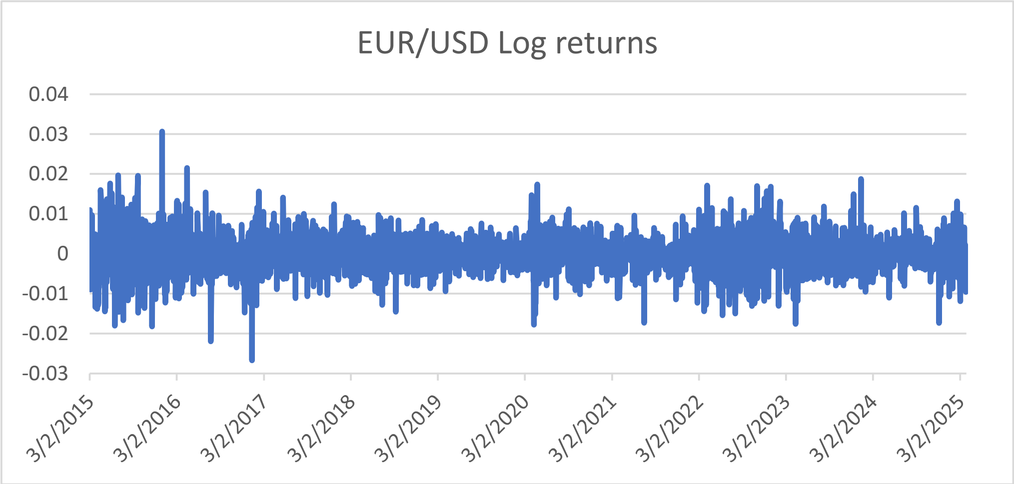


Figure 4. Log Returns of daily spot exchange rates. Federal Reserve Economic Data (FRED).

 doi:10.20944/preprints202602.0321.v1

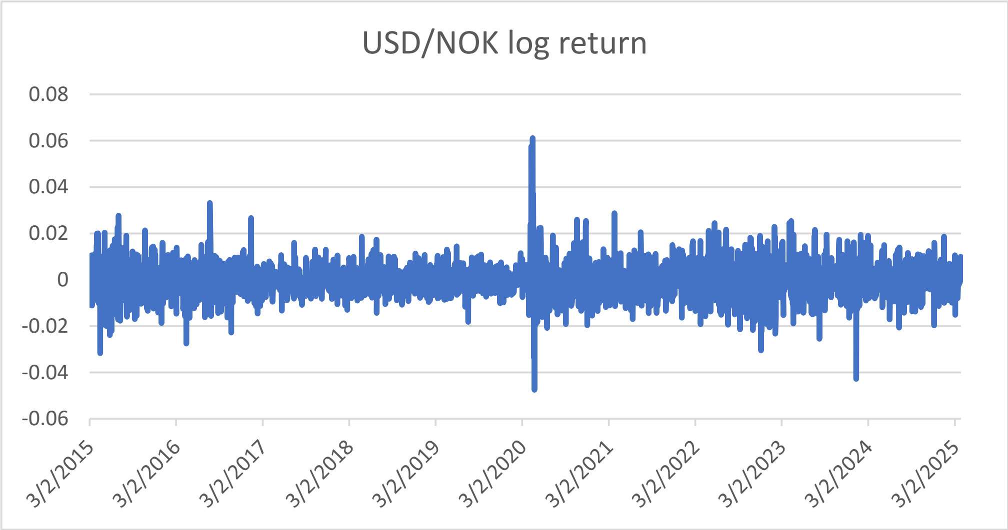


Figure 5. Log Returns of daily spot exchange rates. Federal Reserve Economic Data (FRED).

Descriptive statistics (highlights). Mean returns are near zero for all pairs; USD/NOK is the most volatile and EUR/USD, which is consistently the most frequently traded currency pair in the world by turnover, is the least volatile. EUR/NOK exhibits substantial excess kurtosis, and all series show positive skews, reinforcing the need for fat-tailed models (see table 1 for full metrics).

Table 1. Descriptive statistics for daily log-returns (2015–2025). All three series have near-zero mean returns. USD/NOK is most volatile; EUR/USD is least. EUR/NOK exhibits substantial excess kurtosis; all three show positive skewness and volatility clustering.

Pair	Count	Mean	Std	Min	Max	Skewness	Kurtosis
EUR/NOK	2553	0.0000	0.0056	-0.3016	0.0513	0.5986	10.1221
USD/NOK	2553	0.0001	0.0077	-0.04445	0.0635	0.2099	7.7873
EUR/USD	2532	-5.49e-05	0.0049	-0.0267	0.0306	0.0469	5.1546

Diagnostics. On standardized residuals from GARCH(1,1), Ljung-Box (levels/squared) and ARCH–LM indicate approximate i.i.d. behavior in the initial calibration sample across all pairs. In the 4-year subsample, tests remain non-significant for squared residuals; EUR/USD shows some residual autocorrelation in levels, but block-maxima of those residuals pass Ljung-Box ($p = 0.45$), supporting continued analysis. More details are provided in appendix B.

Implication. The combination of clustering, heavy tails, and diagnostics justifies our two-step pipeline: GARCH to filter volatility and EVT (Block Maxima/GEV) to model extremes for VaR.

4. Methodology

4.1. Motivation and Framework.

High-frequency financial returns display volatility clustering, heavy tails, and moderate asymmetries that violate conditional normality (e.g., Cont, 2001). To address these features, we adopt a two-layer GARCH–EVT pipeline: (i) a GARCH(1,1) filter extracts time-varying conditional volatility and standardized residuals; (ii) Block Maxima EVT (BMM) fits a GEV distribution to monthly maxima of those residuals to infer far-tail quantiles. Daily one-step-ahead VaR is then produced by combining the GEV tail quantile with the next-day GARCH forecast. This structure follows McNeil and Frey's two-step idea while implementing BMM rather than PoT for transparency and reproducibility.

 doi:10.20944/preprints202602.0321.v1

4.1.1. Why choose BMM rather than PoT

The choice involves the following trade-offs:

BMM-strengths

- Transparency and governance: No subjective threshold to pick; one extreme per block → easy to audit and replicate.
- Stability in far tails: Focuses on the most extreme realizations; often yields robust 99% tail behavior (aligns with regulatory emphasis).
- Tolerance to mild dependence: Maxima are less sensitive to short-run serial correlation that may remain after filtering.
- Clean rolling implementation: Block definition is fixed (e.g., 21 trading days); avoids re-optimizing thresholds day by day.

BMM-limitations

- Data inefficiency: Discards sub-maximal exceedances; estimates can have higher variance, especially with short samples.
- Block-size choice matters: Too long → few blocks; too short → noisier extremes. It is a bias-variance trade-off.
- Horizon mismatch needs care: Maxima are over a block horizon; mapping to daily VaR requires an explicit adjustment.

PoT-strengths

- Data efficiency: Uses all exceedances above a threshold → lower variance, especially for 95% tails.
- Flexibility: Can target specific tail probabilities and adapt using threshold diagnostics (mean residual life plots, stability plots).
- Asymptotic justification: GPD approximation above a high threshold is well-understood.

PoT-limitations

- Threshold selection risk: Results can be sensitive to the chosen cutoff; diagnostics are judgment-heavy.
- Declustering and dependence: Requires careful treatment when exceedances cluster; otherwise inference is biased.
- Operational fragility in rolling setups: Re-selecting thresholds over time can induce instability, look-ahead bias, or parameter jumps across windows.

4.1.2. Why BMM suits this paper

Weighing these trade-offs, we select BMM over PoT for this study:

- Auditability and reproducibility were priorities in this study: BMM's "no-threshold" design avoids subjective tuning and is easy to explain to regulators and practitioners.
- The study emphasizes far-tail (99%) performance and cross-pair comparability; BMM's focus on extremes and fixed block design serve both aims.
- With a GARCH filter first, residual dependence is already reduced; BMM's robustness to mild residual dependence supports reliable tail estimation.
- Empirically, our results show strong 99% coverage with BMM, while the main shortfall appears at 95% (where PoT is often superior). That trade-off is acceptable given our far-tail objective and governance constraints.

Below we define two layers in our methodology pipeline. Layer A standardizes the data (making fluctuations comparable over time), while Layer B extracts tail information; together they deliver a daily VaR that combines the current risk level $\hat{\sigma}_{t+1}$ with tail thickness ξ.

4.2. Layer A: Conditional Volatility via GARCH(1,1)

Layer A - Volatility filter (GARCH):

Estimate a GARCH(1,1) on daily log returns to obtain the next-day conditional volatility σ_{t+1} and standardized residuals $\hat{z}_t$. Purpose: remove heteroskedasticity so the residuals are closer to i.i.d. and suitable for extreme-value analysis. Output: $\hat{\mu}_{t+1}$, $\hat{\sigma}_{t+1}$, and $\hat{z}_t = \hat{\epsilon}_t/\hat{\sigma}_t$.

Let r_t denote daily log-returns. We estimate

$$r_t = \mu + \epsilon_t,\ \epsilon_t = \sigma_t z_t$$

and

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

Here, r_t is the daily log return at time t, μ is the conditional mean of the return series, ϵ_t is the conditional residual, z_t is the standardized residual, assumed to be i.i.d. (mean 0, variance 1). The parameters $\omega > 0$, $\alpha \geq 0$ and $\beta \geq 0$ represent the persistence of volatility.

The GARCH filter delivers $\hat{\sigma}_{t+1}$ and standardized residuals $\hat{z}_t = \hat{\epsilon}_t/\hat{\sigma}_t$ reducing heteroskedasticity so that EVT assumptions are closer to i.i.d. behavior. We employ a rolling window of 1,000 observations to accommodate regime shifts and produce daily one-step-ahead forecasts $(\hat{\mu}_{t+1}, \hat{\sigma}_{t+1})$.

4.3. Layer B: Extremes via Block Maxima and the GEV

Layer B contains the following steps:

Partition $\hat{z}_t$ into non-overlapping 21-trading-day blocks and take the maximum per block. Fit a GEV distribution (μ, σ, ξ) to the block maxima to obtain tail quantiles q_p. Map these back to one-day VaR using $\hat{\sigma}_{t+1}$ (with the month→day adjustment). Output: one-day VaR thresholds via §4.4.1 mapping.

We partition standardized residuals into non-overlapping 21-trading-day blocks (≈ monthly) and collect the maximum from each block. Under the extremal types theorem, block maxima converge (after normalization) to the GEV family with cdf:

$$G(z) = exp\left\{-\left[1 + \xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\}, for\ 1 + \xi\left(\frac{z-\mu}{\sigma}\right) > 0$$

Here, μ is the location parameter, $\sigma > 0$ is the scale parameter and ξ (shape) govern tail behavior. $\xi > 0$ indicates Fréchet heavy tails, typical in asset returns. Parameters (μ, σ, ξ) are estimated by MLE.

4.4. Integrating the Layers: From GEV Quantiles to Daily VaR

We model the lower tail by applying BMM to the negated standardized residuals $-\hat{z}_t$; block maxima of $-\hat{z}_t$ therefore correspond to lower-tail quantiles of returns, and the VaR mapping below is written in terms of these block-maxima quantiles.

4.4.1. Mapping Block (Monthly) Probabilities to Daily VaR

Let B denote the block size in trading days (here $B = 21$). If $M = max\{Z_1, \dots, Z_B\}$ are block maxima of standardized daily residuals with daily cdf F, then

$$\Pr(M \leq x) = \Pr(Z_1 \leq x, \dots, Z_B \leq x) = \prod_{i=1}^{B} \Pr(Z_i \leq x) = F(x)^B$$

where the product uses (approximate) independence after GARCH filtering.

If x is the daily α-quantile, i.e., $F(x) = \alpha$, then the block-maxima cdf at the same x is:

$$\Pr(M \leq x) = F(x)^B = \alpha^B$$

So, the probability level we must use in the GEV (block-maxima) quantile to obtain the daily α-quantile is

 doi:10.20944/preprints202602.0321.v1

$$\alpha_B = \alpha^B$$

Operationally:

1. Pick the daily confidence level $\alpha \in \{0.95, 0.99\}$
2. Compute $\alpha_B = \alpha^B$
3. Take the GEV quantile at α_B: $q_{\alpha_B} \coloneqq Q_{GEV}(\alpha_B)$.
4. Map to one-day VaR:

$$VaR_{t+1}^{(\alpha)} = \hat{\mu}_{t+1} + \hat{\sigma}_{t+1} q_{\alpha_B}$$

Example (using $\boldsymbol{B = 21}$)

$$\alpha = 0.99 \Rightarrow \alpha_B = 0.99^{21} \approx exp(21 ln 0.99) \approx exp(21 \times (-0.0100503)) = exp(-0.2111) \approx 0.8097 \approx 0.81$$

$$\alpha = 0.95 \Rightarrow \alpha_B = 0.95^{21} \approx 0.34$$

Interpretation: for daily 99% VaR we evaluate the GEV quantile at ~ 0.81, not at 0.99; for daily 95% VaR we evaluate it at ~ 0.34.

If residual dependence persists, a standard refinement replaces B with an effective block size $\tilde{B} = \theta B$ where $\theta \in (0,1)$ is the extremal index. Our block-maxima diagnostics (Appendix B) are non-significant (p=0.45), so we proceed with $\theta \approx 1$.

4.4.2. GEV Quantiles and VaR Forecasts

The fitted GEV yields a block-maxima quantile q_p at probability level $p \in (0,1)$. The GEV quantile for a given non-exceedance probability p is given by:

$$q_p = \begin{Bmatrix} \mu + \frac{\sigma}{\xi}\{[-lnp]^{-\xi} - 1\}, for \ \xi \neq 0 \\ \mu + \sigma ln[-lnp], \ \ for \qquad \xi = 0 \end{Bmatrix}$$

These quantiles define the VaR threshold used to assess extreme losses. To obtain one-day-ahead VaR, we map that tail quantile back to returns using the next-day GARCH forecast

$$VaR_{t+1}^{(p)} = \hat{\mu}_{t+1} + \hat{\sigma}_{t+1} q_{p_B}, \qquad p_B = p^B$$

with a monthly→daily probability adjustment to align the block-maxima scale with the daily horizon. This coupling lets current volatility (GARCH) and tail thickness (GEV) jointly determine the daily loss threshold.

4.5. Practical Choices

- Block size: 21 trading days (~monthly) balances tail-signal extraction and sample size of maxima within our rolling design. Larger blocks would unduly reduce the number of extremes for backtesting
- Rolling estimation: Each day, we re-estimate GARCH on the latest 1,000 observations, update residuals, refresh block maxima, re-fit GEV, and produce $VaR_{t+1}^{(p)}$. This design lets the model react to evolving regimes (e.g., crisis periods).

4.6. Risk Metric and Regulatory Context

We focus on *Value-at-Risk (VaR)* because it remains widely used in practice and aligns with standard backtesting tools. The GEV-based tail modeling is consistent with supervisory emphasis on better capturing extreme losses.

4.7. Backtesting and Model Validation

We evaluate the forecasts using Christoffersen's framework: (i) unconditional coverage (does the total number of breaches match the nominal rate?) and (ii) independence (are breaches randomly scattered or clustered?), with the Kupiec-type LR for coverage. Tests are conducted at 95% and 99% confidence levels over the out-of-sample period. Further implementation details and test statistics are provided in Appendix B.

5. Results and Contribution

5.1. Visual back tests

Figures 6-11 plot one-day-ahead VaR at 95% and 99% (colored lines) against realized daily log returns (black). The forecasts track the broad risk environment: a breach cluster in Feb–Apr 2020 signals model strain during the COVID shock—especially at 95%—while widened VaR bands post-2020 indicate the framework's adaptive response to a higher-volatility regime and the subsequent decline in breaches.

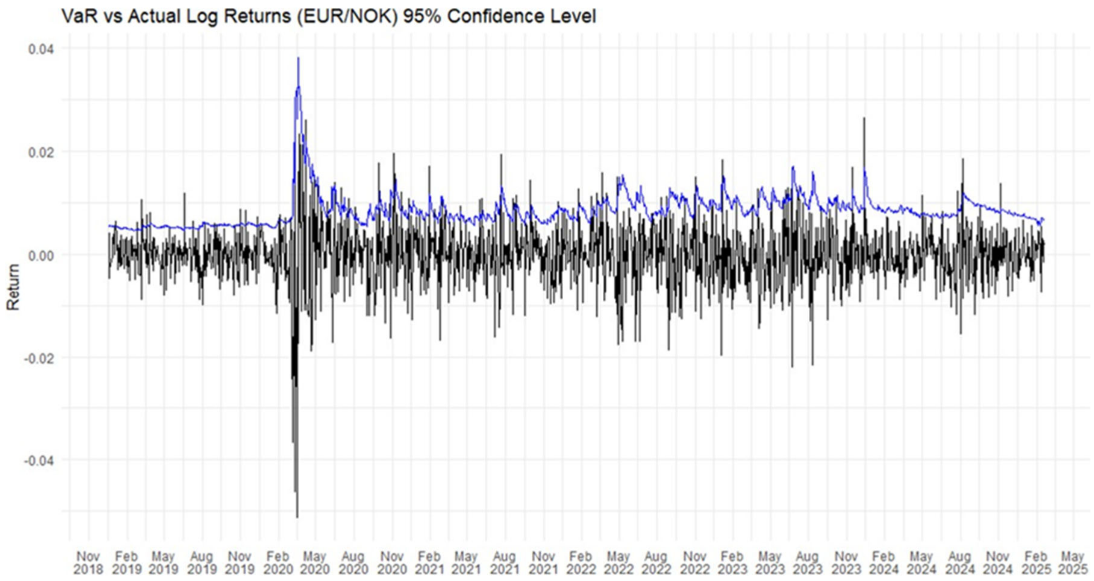
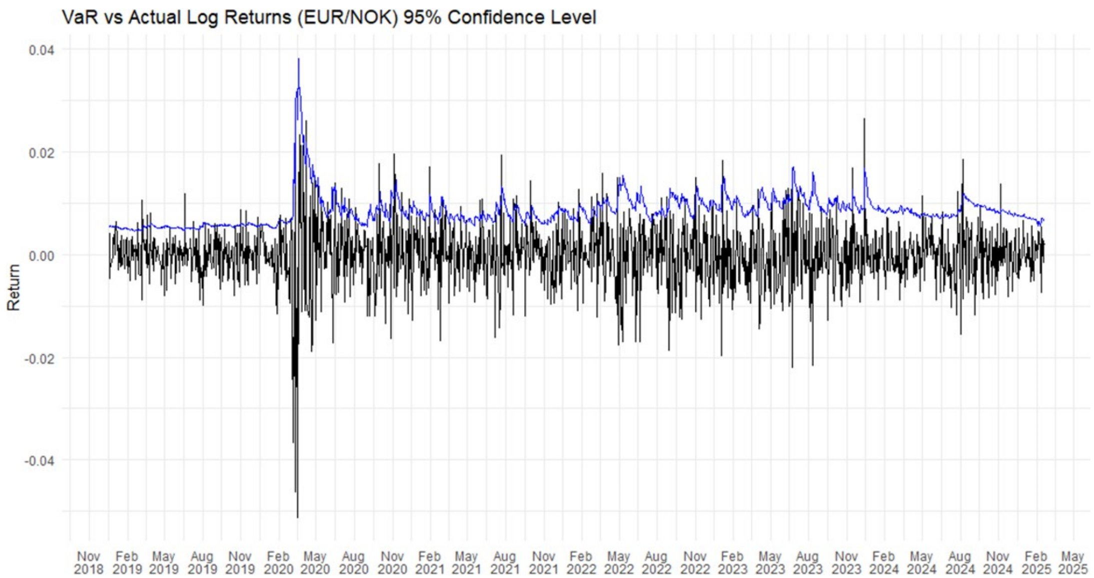


Figure 6. Visual back test.



 doi:10.20944/preprints202602.0321.v1

Figure 7. Visual back test.

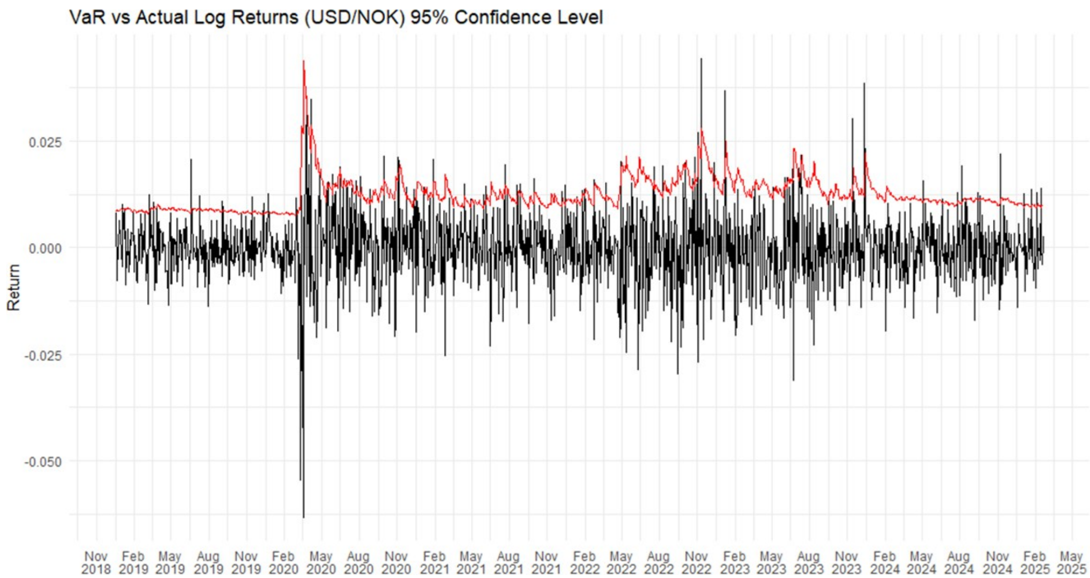


Figure 8. Visual back test.

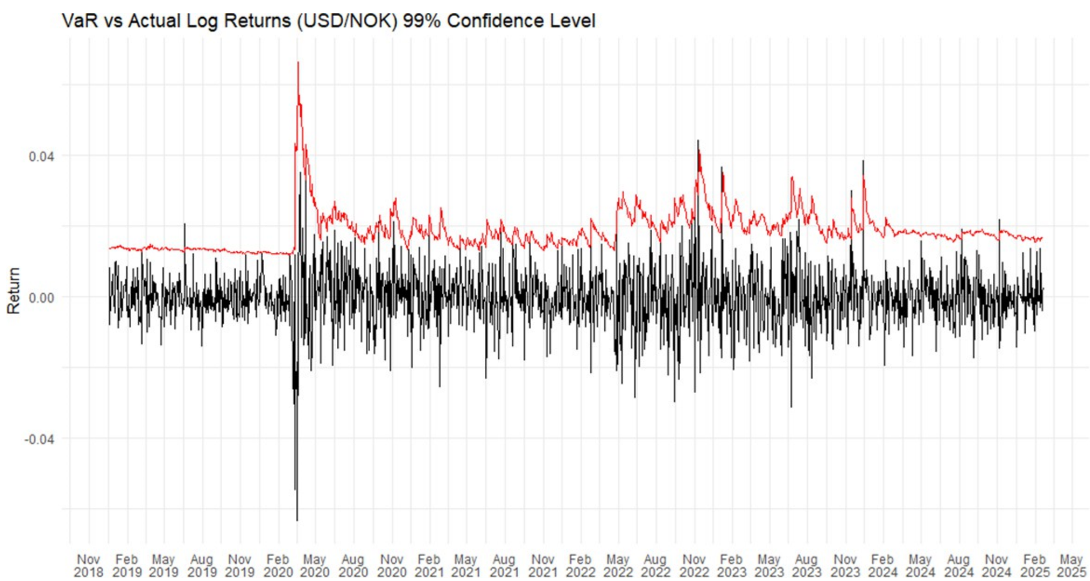


Figure 9. Visual back test.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 February 2026 doi:10.20944/preprints202602.0321.v1

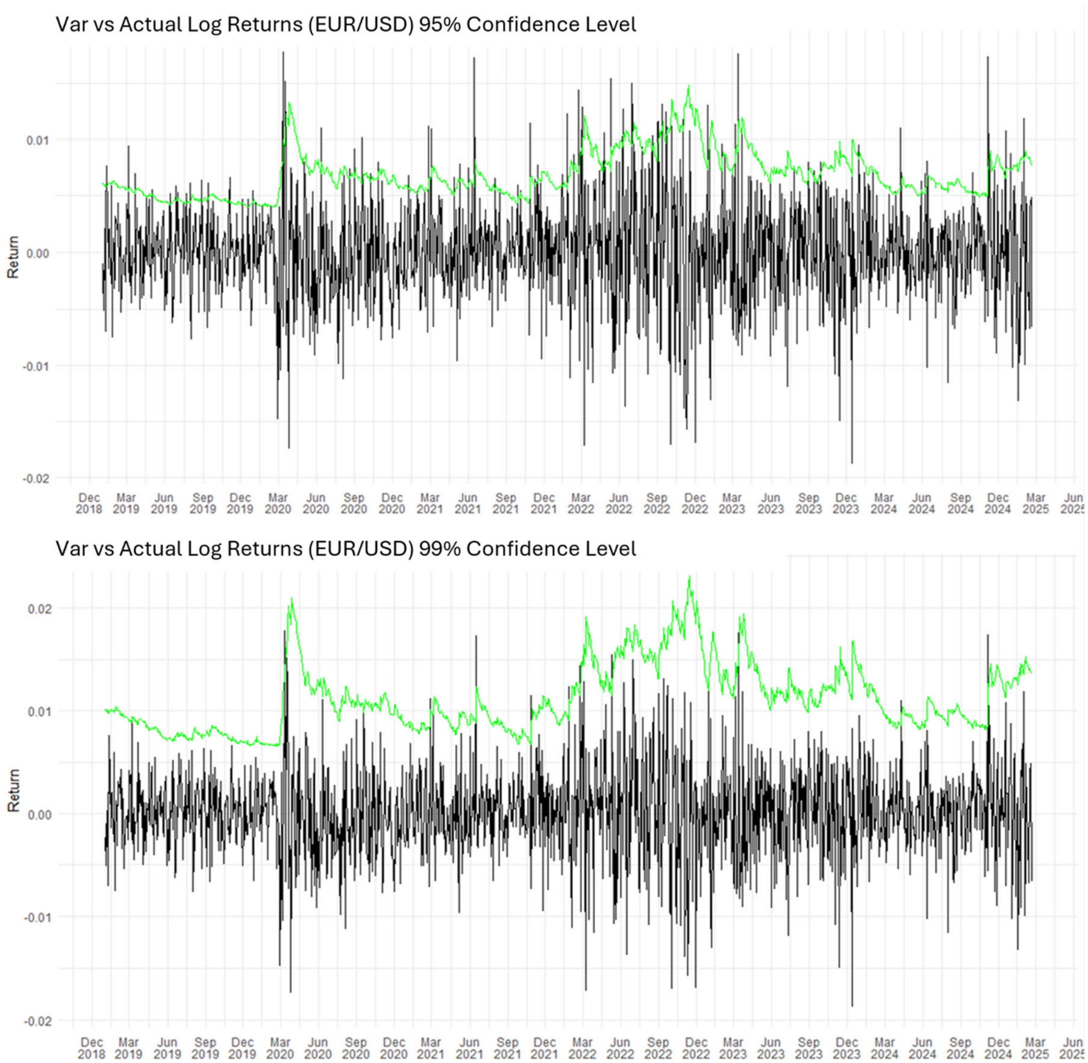


Figure 10. and 11 Visual back test.

Figures 6-11 show the 95% and 99% VaR lines (blue, red and green) against the actual negative log returns (black bars) for the EUR/USD pair using the block size of 21 days.

5.2. Unconditional Coverage (Kupiec)

The Kupiec unconditional coverage test reported in Table 2 shows that EUR/USD at 95% fails the coverage test (Expected 78, Actual 94, $p = 0.0478$), implying underestimation in the moderate tail for that pair. All other pair–level combinations at 95% and 99% do not reject correct coverage ($p > 0.05$), though realized exceedances are slightly above target in several cases. At 99%, EUR/USD even shows one fewer breach than expected (14 vs 15). Overall, the 99% VaR performs robustly across pairs; the 95% EUR/USD shortfall is the notable exception.

Table 2. Kupiec unconditional coverage test.

Time series	ıfidence level	Predictions	Expected Exceedances	Actual Exceedances	P-value
USD/NOK	95%	1552	78	82	0.6115

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 February 2026 doi:10.20944/preprints202602.0321.v1

USD/NOK	99%	1552	16	17	0.71
EUR/NOK	95%	1552	78	87	0.2824
EUR/NOK	99%	1552	16	21	0.1846
EUR/USD	95%	1531	78	94	0.0478
EUR/USD	99%	1531	15	14	0.7328

5.3. Independence of exceedances (Christoffersen)

Christoffersen's independence of exceedances test reported in Table 3 indicates that USD/NOK at 95% exhibits clustering of breaches (LR = 4.0275, p = 0.0448). All other cases fail to reject independence, suggesting exceedances are approximately serially independent outside this single pairing/confidence level.

Table 3. Christoffersen's independence of exceedances test.

Time series	ıfidence level	Predictions	LR-stat	P- value
USD/NOK	95%	1552	4.0275	0.0448
USD/NOK	99%	1552	0.3768	0.5393
EUR/NOK	95%	1552	0.9289	0.3351
EUR/NOK	99%	1552	0.5765	0.4477
EUR/USD	95%	1531	0.2785	0.5979
EUR/USD	99%	1531	0.2586	0.6111

5.4. Conditional Volatility Dynamics (GARCH)

Rolling GARCH(1,1) estimates (2021–2025) confirm high persistence ($\alpha_1+\beta_1 \geq 0.985$), with USD/NOK displaying the highest persistence and the only statistically significant ARCH response among the three pairs—consistent with a sharper near-term reaction to shocks (table 4). Mean (μ) and variance intercept (ω) are small and not statistically different from zero, supporting the use of GARCH filtering prior to EVT.

Table 4. Rolling GARCH(1,1) parameters (2021–2025). All pairs exhibit high persistence ($\alpha_1+\beta_1 \geq 0.985$); USD/NOK shows the highest persistence and significant ARCH, indicating stronger short-term sensitivity.

Pair	μ	ω	α1	β1	α1+β1
EUR/NOK	-6.280e-05	4.591e-07	0.034014	0.951233	0.98527
USD/NOK	-0.000203	2.536e-07	0.017716	0.978044	0.99576
EUR/USD	0.000150	2.622e-07	0.038504	0.950560	0.98906

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 February 2026 doi:10.20944/preprints202602.0321.v1

5.5. Tail Behavior (GEV on Block Maxima)

GEV fits to monthly block maxima of standardized residuals yield $\xi > 0$ (Fréchet class) for all pairs, confirming heavy tails (table 5). EUR/NOK posts the largest shape ($\xi \approx 0.128$), implying the heaviest tail; USD/NOK has the largest scale, indicating the widest dispersion of extremes; EUR/USD lies between on shape and scale. These differences anticipate cross-pair variation in VaR and breach patterns.

Table 5. GEV parameters for block maxima of standardized residuals (2021–2025). All $\xi>0$ (Fréchet heavy tails). EUR/NOK has the heaviest tail (largest ξ); USD/NOK the largest scale (widest spread of extremes).

Pair	Location (μ)	Scale (σ)	Shape (ξ)
EUR/NOK	1.640106	0.520822	0.127542
USD/NOK	1.632041	0.616371	0.074646
EUR/USD	1.634206	0.602279	0.109836

5.6. One-day VaR levels

Average VaR (2021–2025) is 0.66%/1.04% (EUR/NOK), 0.97%/1.64% (USD/NOK), and 0.84%/1.39% (EUR/USD) at 95%/99%, respectively (table 6). With returns defined so that negative values represent losses, the reported VaR magnitudes represent daily loss thresholds at the stated confidence levels.

Table 6. One-day VaR from GARCH–BMM (2021–2025).

Pair	VaR 95%	VaR 99%
EUR/NOK	0.66%	1.04%
USD/NOK	0.97%	1.64%
EUR/USD	0.84%	1.39%

Contribution: The pipeline delivers robust far-tail VaR estimates—especially at 99%—while revealing moderate-tail underestimation for EUR/USD at 95%. It balances transparency (no threshold selection) with practical performance, offering an auditable alternative to PoT for FX risk.

6. Discussion

6.1. Mechanisms Behind the Tail Patterns

Our results point to distinct mechanisms across pairs. The heavier tail for EUR/NOK (larger ξ) is consistent with jump-like repricing when euro-area news interacts with Norway's commodity exposure and relatively shallow NOK market depth. By contrast, USD/NOK's larger extreme-scale aligns with high volatility persistence ($\alpha + \beta \approx 1$): once volatility rises, extreme moves remain elevated for longer. EUR/USD's moderate-tail under-coverage at 95% suggests that, outside the far tail, additional information is needed beyond block maxima—either through a denser sampling of exceedances (e.g., PoT) or heavier-tailed innovations in the volatility filter.

6.2. Regime Sensitivity and Model Dynamics

Layer A (GARCH) and Layer B (BMM/GEV) respond at different speeds in stress. After a shock, Layer A widens next-day conditional volatility $\hat{\sigma}_{t+1}$ immediately, while Layer B updates only when new block maxima enter the estimation window. This sequencing—fast scale update, slower tail-shape update—explains the temporary rise in 95% breaches around regime shifts (e.g., early COVID-19), alongside the preservation of 99% coverage once blocks incorporate post-shock extremes.

6.3. Implications for risk management and policy

- Risk governance. Use BMM-based VaR for far-tail capital (e.g., 99%), where auditability and threshold-free estimation reduce model risk.
- Monitoring. Track a tail-thickness dashboard—a small set of live indicators showing whether the return distribution is growing fatter-tailed (more extreme-event risk) and more persistent (volatility lingers). This includes rolling $\hat{\xi}$, extreme quantiles (e.g., 99% VaR), and volatility persistence $(\alpha + \beta)$ to inform procyclicality safeguards.
- Hedging design. For EUR/NOK (heavier tail), prioritize gap-risk protection (options that load on jumps); for USD/NOK (persistent volatility), emphasize longer-dated volatility hedges.

6.4. Why 95% Can Underperform, and How to Shore It Up

BMM samples one extreme per block; at 95% the tail is "moderate" rather than extreme, so information is thin. Two practical refinements improve 95% calibration without sacrificing far-tail clarity: (i) overlay a lightweight PoT check (select a high threshold using stability/mean-residual-life diagnostics and model exceedances with GPD), or (ii) allow heavier-tailed/skewed innovations in Layer A (e.g., Student-t or skew-t), which can thicken the moderate tail while preserving the GEV-based far-tail mapping.

6.5. Dependence and the Effective Block Size

If short-run clustering persists in standardized residuals, the effective block size becomes $\tilde{B} = \theta B$ with extremal index $\theta \in (0,1]$. A lower θ raises the implied daily tail probability for a given block quantile, tightening 95% coverage. Our block-maxima diagnostics do not reject independence for the series we use, supporting $\theta \approx 1$, but periodic checks are advisable during stress episodes.

6.6. External Validity and Scope Conditions

The GARCH→BMM pipeline should transfer well to other liquid FX pairs and commodities, where volatility clustering and heavy tails are pervasive. In illiquid markets with sparse trading and microstructure frictions, block maxima may be dominated by noise; in such settings, a PoT approach with declustering (and possibly realized-measure volatility filters) may be preferable.

6.7. Practical takeaway

- Use BMM-based 99% VaR for far-tail capital and governance—keep it as the primary tail metric.
- Supplement 95% limits with a PoT overlay or t/skew-t innovations to better capture the moderate tail.
- Maintain a regime-switch alert: widen monitoring when rolling $\alpha + \beta$ or $\hat{\xi}$ crosses pre-set thresholds.
- Document the month→day mapping ($\alpha_B = \alpha^B$) and, if needed, adjust for $\theta < 1$ during clustered extremes.

7. Conclusion

In this section we answer briefly our main research questions:

RQ1 (Tail heaviness). NOK crosses exhibit materially heavy tails. EUR/NOK shows the heaviest tail (larger $\hat{\xi}$), while USD/NOK displays the largest extreme-scale (wider dispersion of block maxima); EUR/USD is comparatively thinner-tailed in the far tail but not immune to moderate-tail risk.

RQ2 (Backtests). The integrated GARCH–BMM (GEV) framework delivers reliable 99% VaR across pairs (coverage not rejected; no systematic clustering), while 95% shows under-coverage for EUR/USD in some windows and mild clustering for USD/NOK—consistent with BMM's emphasis on far tails over moderate tails.

RQ3 (Rolling estimates in stress). During stress (e.g., COVID onset), Layer A (GARCH) adjusts scale quickly ($\hat{\sigma}_{t+1}$ widens), whereas Layer B (BMM/GEV) updates tail shape more slowly as new block maxima arrive. This sequencing explains transient 95% breaches during regime shifts, with 99% performance stabilizing once blocks incorporate post-shock extremes.

Implications: For governance and auditability, BMM-based 99% VaR is a practical far-tail metric. To shore up the moderate tail (95%), consider a PoT overlay or t/skew-t innovations in Layer A. Tracking a tail-thickness dashboard (rolling ν, far-tail quantiles, and $\alpha + \beta$) can trigger procyclicality safeguards when tail risk and persistence rise.

Summary of contribution: We provide a transparent, self-contained pipeline: GARCH filter → BMM/GEV tail → month → day mapping ($\alpha_B = \alpha^B$) → VaR → backtests, and document economically meaningful cross-pair differences in NOK risk.

Author Contributions: Conceptualization, K.M., H.S., PdL, and S.W.; methodology, K.M. and H.S., with input and model suggestions from S.W and PdL.; software, K.M. and H.S.; validation, K.M., H.S., PdL, and S.W.; formal analysis, K.M. and H.S.; investigation, K.M. and H.S.; data curation, K.M. and H.S.; writing—original draft preparation, K.M. and H.S.; writing—review and editing, PdL and S.W., including detailed checks of methods, models, and results; visualization, K.M. and H.S.; supervision, PdL and S.W.; project administration and quality control, PdL and S.W. All authors have read and agreed to the published version of the manuscript.

Funding: No funding was received.

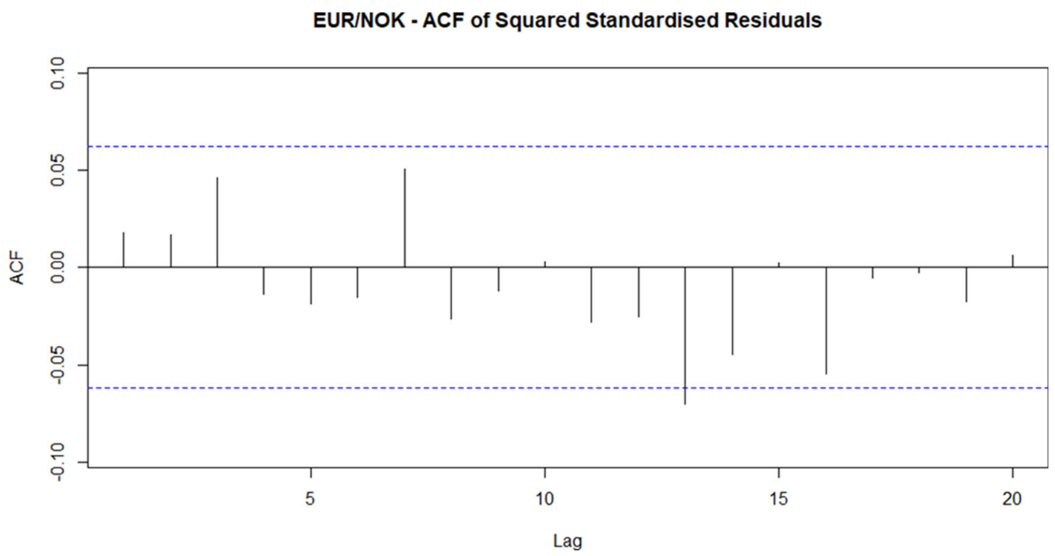
Data Availability Statement: All data is publicly available. See section 3.

Acknowledgments: No additional support was received.

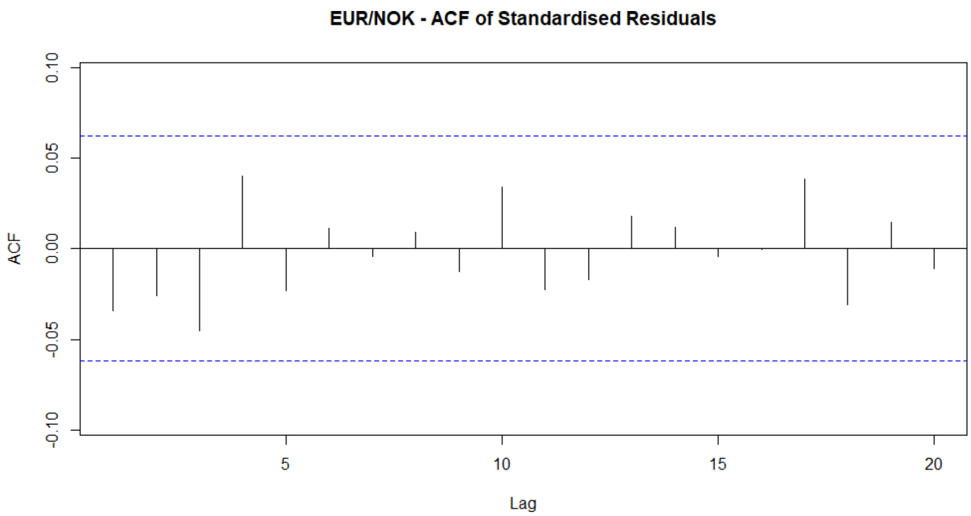
Conflicts of Interest: No conflicts of interest.

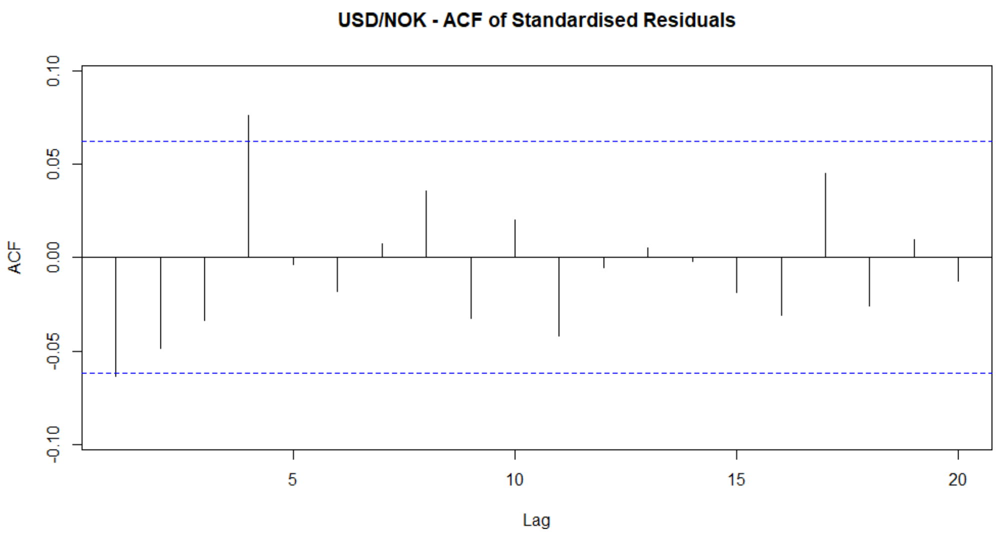
Appendix A. Stationarity and Autocorrelation Diagnostics

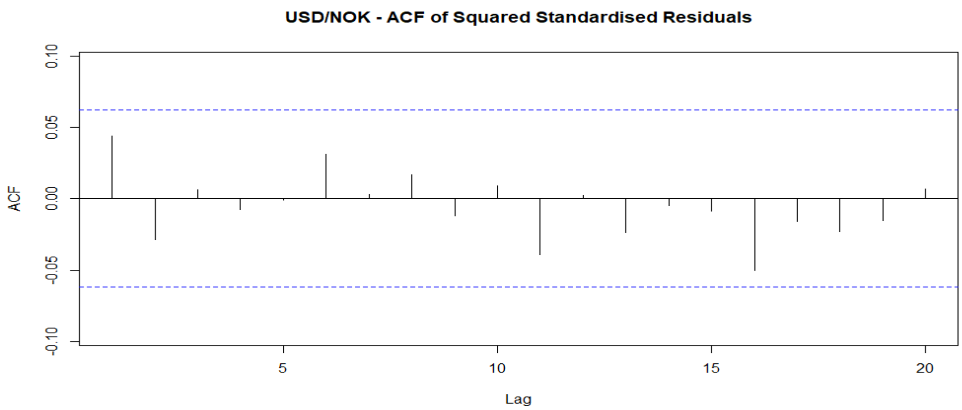
ACF plots (4-year risk-characteristics window). Visual inspection of the ACFs reveals no systematic autocorrelation for EUR/NOK and USD/NOK: autocorrelation coefficients lie within the 95% significance bands at most lags and display no persistent decay. By contrast, EUR/USD shows a few isolated spikes, consistent with the Ljung–Box test on the standardized residuals $\hat{z}_t$reported in Appendix B. Overall, the evidence supports approximate serial independence for EUR/NOK and USD/NOK, with limited, localized dependence in EUR/USD.

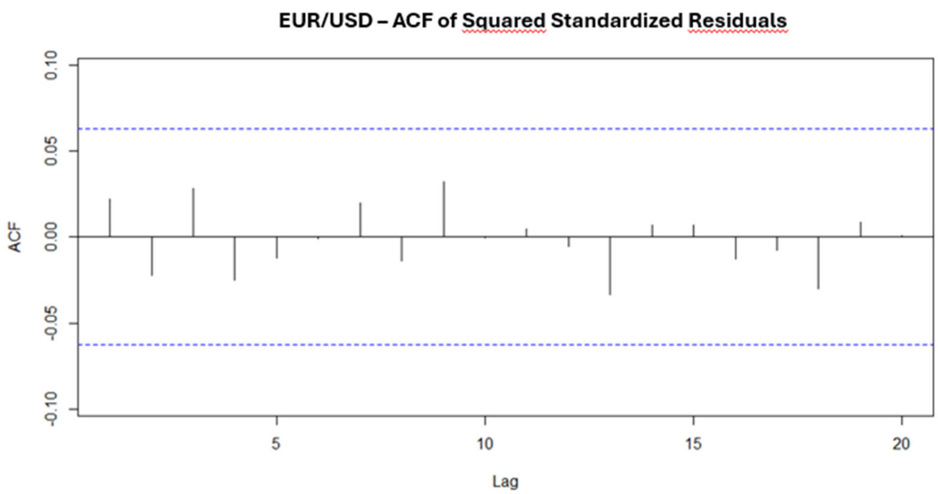


Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 February 2026 doi:10.20944/preprints202602.0321.v1

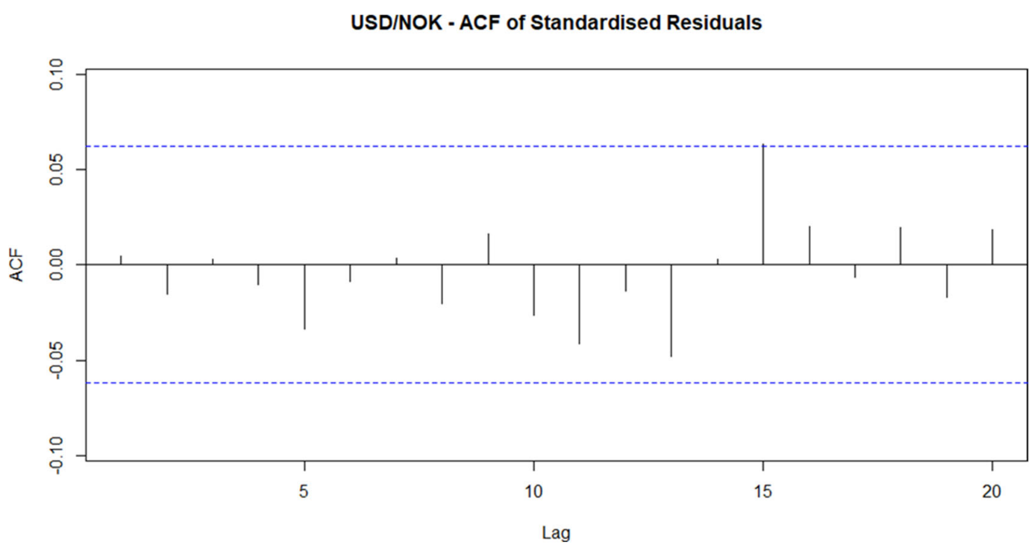


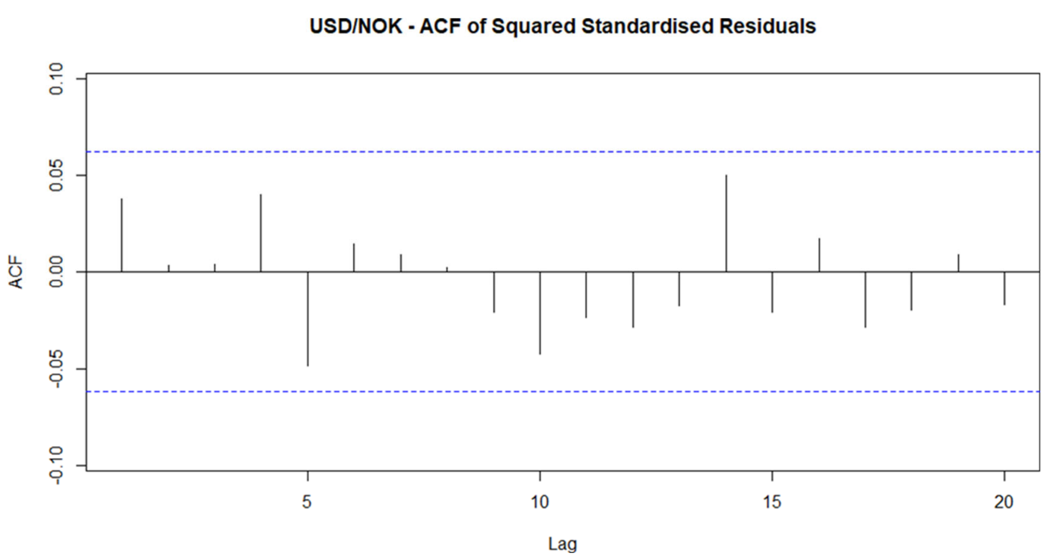






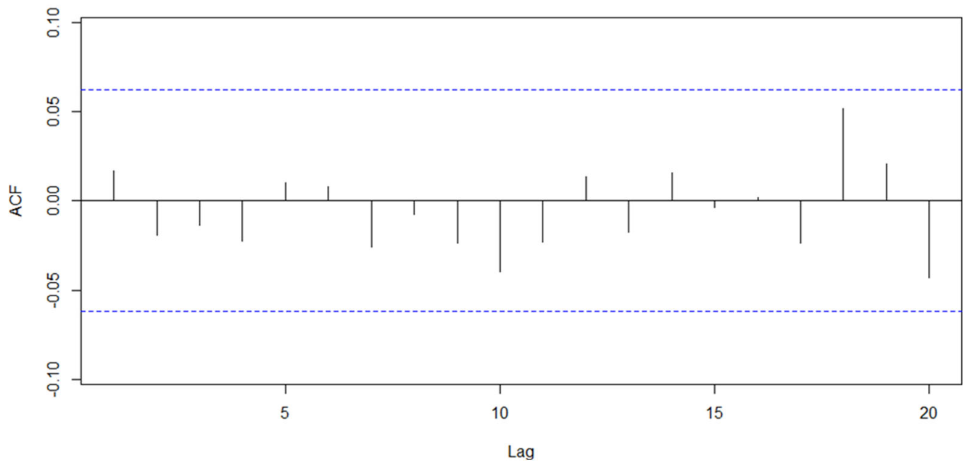
First calibration period for backtesting, no serial correlation visual in these plots either:



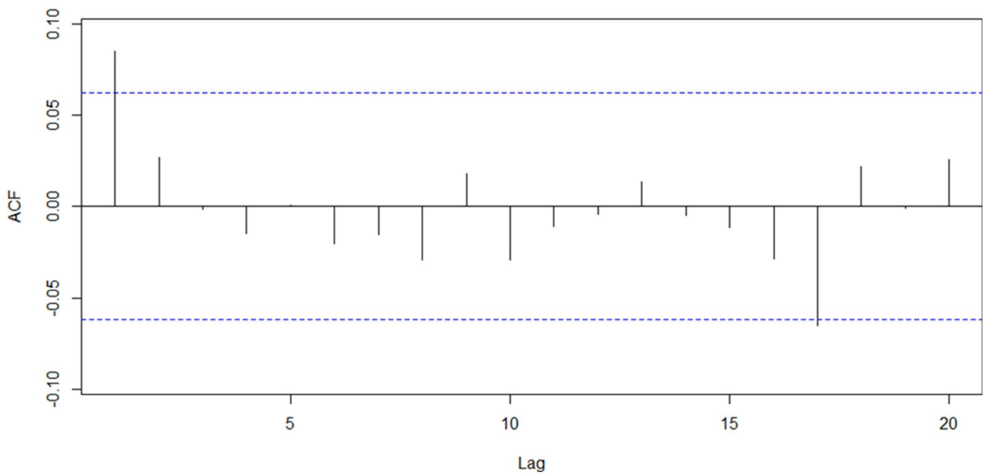


 doi:10.20944/preprints202602.0321.v1

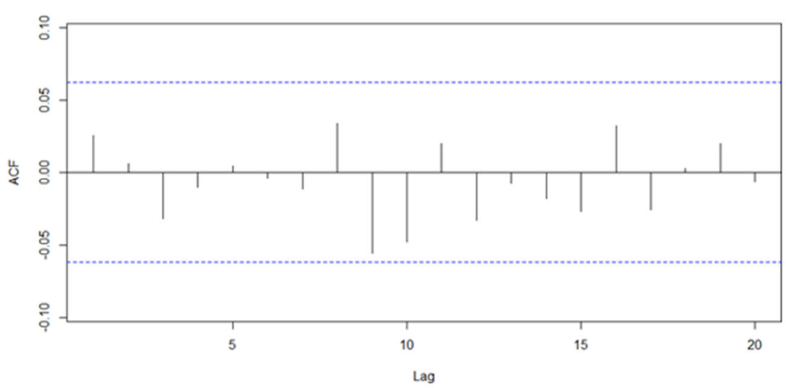
EUR/NOK - ACF of Standardised Residuals



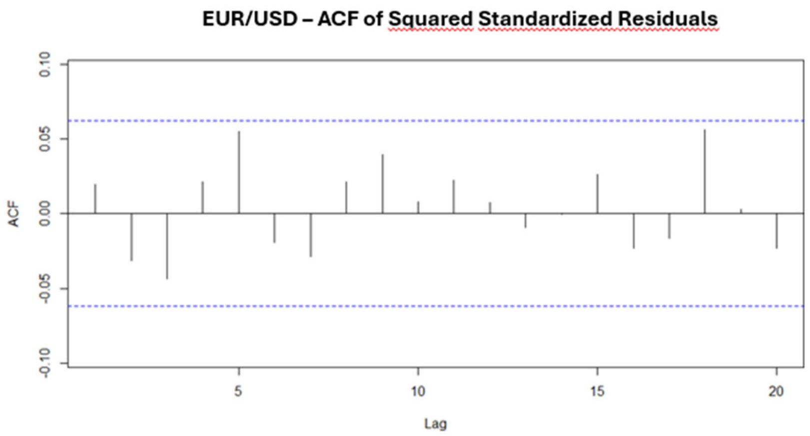
EUR/NOK - ACF of Squared Standardised Residuals



EUR/USD – ACF plots of Standardized Residuals



 doi:10.20944/preprints202602.0321.v1



Appendix B

Autocorrelation and Ljung Box test

We assess residual diagnostics on the standardized residuals $\hat{z}_t$ from the GARCH(1,1) fit. Specifically, we test for serial correlation using the Ljung–Box statistic on $\hat{z}_t$ and for remaining ARCH effects using the Ljung–Box on $\hat{z}_t^2$ and the ARCH–LM test. In the initial calibration window, all three currency pairs fail to reject the null of no autocorrelation and no remaining conditional heteroskedasticity, indicating residuals that are approximately i.d. Over the four-year evaluation window, the Ljung–Box on $\hat{z}_t^2$ and the ARCH–LM remain non-significant for all pairs, while EUR/USD shows a significant Ljung–Box on $\hat{z}_t$ implying some residual linear dependence in that series.

Because our tail modeling uses block maxima of standardized residuals, we further tested serial dependence on the extracted maxima for EUR/USD. The Ljung–Box on maxima is non-significant ($p = 0.45$), indicating no evidence of dependence at the frequency relevant for the EVT step. We therefore proceeded with the analysis, noting that any mild short-run autocorrelation in EUR/USD residual levels does not appear to contaminate the block-maxima series used for the GEV fit.

References

Bank for International Settlements (2022) *Triennial central bank survey: OTC foreign exchange turnover in April 2022*. Basel: BIS. Available at: https://www.bis.org/statistics/rpfx22_fx.htm

Bank for International Settlements (no date) *History of the Basel Committee*. Available at: https://www.bis.org/bcbs/history.htm

Bollerslev, T. (1986) 'Generalized autoregressive conditional heteroskedasticity', *Journal of Econometrics*, 31(3), pp. 307–327. doi: https://doi.org/10.1016/0304-4076(86)90063-1

Cayman Islands Monetary Authority (no date) *Basel I & II*. Available at: https://www.cima.ky/basel-ii-iii

Chikobvu, D. and Ndlovu, T. (2023) 'The generalized extreme value distribution approach to comparing the riskiness of Bitcoin/US dollar and South African rand/US dollar returns', *Journal of Risk and Financial Management*, 16(4), p. 253. doi: https://doi.org/10.3390/jrfm16040253

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

Coles, S. (2001) *An introduction to statistical modeling of extreme values*. London: Springer.

Cont, R. (2001) 'Empirical properties of asset returns: stylized facts and statistical issues', *Quantitative Finance*, 1(2), pp. 223–236. doi: https://doi.org/10.1080/713665670

Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling extremal events: For insurance and finance*. Berlin: Springer.

 doi:10.20944/preprints202602.0321.v1

Engle, R.F. (1982) 'Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation', *Econometrica*, 50(4), pp. 987–1007. doi: https://doi.org/10.2307/1912773

J.P. Morgan and Reuters (1996) *RiskMetrics—Technical Document*. 4th edn. New York: Morgan Guaranty Trust Company of New York.

Ljung, G.M. and Box, G.E.P. (1978) 'On a measure of lack of fit in time series models', *Biometrika*, 65(2), pp. 297–303.

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

McNeil, A.J., Frey, R. and Embrechts, P. (2015) *Quantitative risk management: Concepts, techniques and tools*. 2nd edn. Princeton, NJ: Princeton University Press.