Cryptocurrency Volatility and Tail Risk: An Empirical Investigation Using ARMA-GARCH-X Models

Volatilidade das criptomoedas e risco nas caudas: uma investigação empírica utilizando modelos ARMA-GARCH-X

Volatilidad de las criptomonedas y riesgo en las colas: una investigación empírica utilizando modelos ARMA-GARCH-X

Como citar:

Abreu, Daniel P. A.; Campos, Octávio V. & Bressan, Aureliano A. (2026). Cryptocurrency Volatility and Tail Risk: An Empirical Investigation Using ARMA-GARCH-X Models. Revista Gestão & Tecnologia, vol. 26, nº 1, p.07-40

Daniel Pereira Alves de Abreu, Professor especialização da Universidade Federal de Minas Gerais e bolsista doutorado da Universidade Federal de Minas Gerais
danielpabreu22@gmail.com
http://orcid.org/0000-0002-9820-8453

Octávio Valente Campos, Doutor em Controladoria e Contabilidade pela Universidade Federal de Minas Gerais (UFMG). Professor do Centro de Pós-graduação e Pesquisas em Controladoria e Contabilidade (CEPCON)
octaviovc@yahoo.com.br
https://orcid.org/0000-0002-0354-0796

Aureliano Angel Bressan, Professor no Departamento de Administração da Universidade Federal de Minas Gerais, Doutor em Economia pela Universidade Federal de Viçosa.
aureliano.bressan@gmail.com
https://orcid.org/0000-0002-9333-3394

"Os autores declaram não haver qualquer conflito de interesse de natureza pessoal ou corporativa, em relação ao tema, processo e resultado da pesquisa".

Scientific Editor: José Edson LaraOrganization Scientific CommitteeDouble Blind Review by SEER/OJS
Received on 11/02/2026
Approved on 10/06/2026

Abstract

Objective: This study aims to evaluate the performance of different ARMA-GARCH model specifications in the risk management of major cryptocurrencies, investigating whether the inclusion of exogenous variables improves the calibration of risk measures such as Value-at-Risk (VaR) and Expected Shortfall (ES).

Methodology: To achieve this objective, 4,032 specifications of the ARMA-GARCH model applied to the ten main cryptocurrencies in trading were tested. The study incorporated the Fear and Greed Index and Bitcoin Trading Volume as exogenous variables in an ARMA-GARCH-X framework, comparing the performance of the different specifications against an ARMA(1,1)-GARCH(1,1) benchmark.

Originality: Despite growing interest in crypto asset risk management, there are still gaps in the literature regarding the effectiveness of incorporating exogenous variables into forecasting models, as well as the increase in the quality of forecasts when using more complex models.
Main results: The results indicate that the inclusion of external variables improves risk calibration in some assets, although the gains are marginal and heterogeneous. There is also no single optimal parameterization, requiring ARMA orders, GARCH specifications, and error distributions to be adjusted for each cryptocurrency.

Theoretical/methodological contributions: From a methodological point of view, the study contributes by demonstrating the importance of specific calibration of ARMA-GARCH models for different cryptocurrencies in risk estimation. Furthermore, the results suggest that, although more complex models can improve tail risk estimation, the gains in predictive power over simpler models are limited.

Keywords: Cryptocurrencies; Risk Management; ARMA-GARCH; Value-at-Risk; Expected Shortfall.

Resumo

Objetivo: Este estudo tem como objetivo avaliar o desempenho de diferentes especificações do modelo ARMA-GARCH na gestão de risco das principais criptomoedas, investigando se a inclusão de variáveis exógenas melhora a calibração de medidas de risco como Value-at-Risk (VaR) e Expected Shortfall (ES).

Metodologia: Para atingir esse objetivo, foram testadas 4.032 especificações do modelo ARMA-GARCH aplicadas às dez principais criptomoedas em negociação. O estudo incorporou o Índice de Medo e Ganância e o Volume de Negócios do Bitcoin como variáveis exógenas em uma estrutura ARMA-GARCH-X, comparando o desempenho das diferentes especificações em relação a um benchmark ARMA(1,1) -GARCH(1,1).

Originalidade: Apesar do crescente interesse na gestão de risco de criptoativos, ainda há lacunas na literatura quanto à eficácia da incorporação de variáveis exógenas em modelos de previsão, bem o incremento na qualidade das previsões à medida em que se utiliza modelagens mais complexas.

Principais resultados: Os resultados indicam que a inclusão de variáveis externas melhora a calibração de risco em alguns ativos, embora os ganhos sejam marginais e heterogêneos. Também se observa a ausência de uma parametrização ótima única, sendo necessário ajustar ordens ARMA, especificações GARCH e distribuições de erro para cada criptomoeda.

Contribuições teóricas / metodológicas: Do ponto de vista metodológico, o estudo contribui ao demonstrar a importância da calibração específica de modelos ARMA-GARCH para diferentes criptomoedas na estimação de risco. Ademais, os resultados sugerem que, embora modelos mais complexos possam melhorar a estimação do risco de cauda, os ganhos em capacidade preditiva em relação a modelos mais simples são limitados.

Palavras-Chave: Criptoativos; Gestão de Risco; ARMA-GARCH, Value-at-Risk; Expected Shortfall.

Resumen

Objetivo: El objetivo de este estudio es evaluar el rendimiento de diferentes especificaciones del modelo ARMA-GARCH en la gestión del riesgo de las principales criptomonedas, investigando si la inclusión de variables exógenas mejora la calibración de medidas de riesgo como el valor en riesgo (VaR) y el déficit esperado (ES).

Metodología: Para alcanzar este objetivo, se probaron 4032 especificaciones del modelo ARMA-GARCH aplicadas a las diez principales criptomonedas en negociación. El estudio incorporó el Índice de Miedo y Ganancia y el Volumen de Negociación de Bitcoin como variables exógenas en una estructura ARMA-GARCH-X, comparando el rendimiento de las diferentes especificaciones en relación con un benchmark ARMA(11)GARCH(1,1).

Originalidad: A pesar del creciente interés en la gestión del riesgo de los criptoactivos, todavía existen lagunas en la literatura sobre la eficacia de la incorporación de variables exógenas en los modelos de predicción, así como sobre el aumento de la calidad de las predicciones a medida que se utilizan modelos más complejos.

Principales resultados: Los resultados indican que la inclusión de variables externas mejora la calibración del riesgo en algunos activos, aunque las ganancias son marginales y heterogéneas. También se observa la ausencia de una parametrización óptima única, siendo necesario ajustar las órdenes ARMA, las especificaciones GARCH y las distribuciones de error para cada criptomoneda.

Contribuciones teóricas/metodológicas: Desde el punto de vista metodológico, el estudio contribuye al demostrar la importancia de la calibración específica de los modelos ARMA-GARCH para diferentes criptomonedas en la estimación del riesgo. Además, los resultados sugieren que, aunque los modelos más complejos pueden mejorar la estimación del riesgo de cola, las ganancias en capacidad predictiva en relación con los modelos más simples son limitadas.

Palabras clave: Criptoactivos; Gestión de riesgos; ARMA-GARCH, Value-at-Risk; Expected Shortfall.

Jel Code: **G170 Financial Forecasting and Simulation**

1 Introduction

The financial market has undergone significant transformations driven by the internet, which has facilitated interactions and the sharing of information at reduced costs (Suryono et al., 2020). One of the most notable innovations in the last decade has been the rise of cryptocurrencies, a new decentralized financial product that can be identified as a hybrid between a currency, a commodity and a stock (Charfeddine et al., 2020). Bitcoin, developed by Nakamoto (2008) and launched in 2009, was responsible for inaugurating this new branch of the financial market and, due to its gain in popularity, several other cryptocurrency systems began to be developed and commercialized, known as altcoins.

The rapid growth and global dissemination of cryptocurrencies has gained academic interest, in order to explain and understand the evolution of this new financial segment as well as finding ways to calibrate the usage of these new assets in investment portfolios. Given the high volatility and speculative dynamics commonly observed in digital assets, robust risk measurement tools have become essential, particularly those capable of capturing tail risks.

In this context, this study aims to measure the risk of these assets using the Value-at-Risk (VaR) and Expected Shortfall (ES) of the 10 largest cryptocurrencies. While VaR remains the most widely used market risk metric, it presents well-documented limitations, especially regarding the lack of subadditivity and its inability to fully capture extreme tail losses. In contrast, Expected Shortfall overcomes these limitations by incorporating the magnitude of losses beyond the VaR threshold (Acerbi & Tasche, 2002), providing a more coherent and comprehensive measure of downside risk. To do this, volatility forecasting techniques using GARCH models were applied.

Although the literature on cryptocurrency volatility modelling has advanced considerably, most empirical applications of ARMA-GARCH frameworks rely predominantly on past return information, treating volatility as an endogenous process driven solely by its own historical dynamics. However, cryptocurrency markets are particularly sensitive to behavioral sentiment shocks (López-Cabarcos et al., 2021) and structural valuation imbalances (Liu, Tsyvinski, & Wu, 2020), suggesting that purely autoregressive volatility specifications may be incomplete (Wei, Sermpinis & Stasinakis, 2023).

as a proxy for investor sentiment rooted in Behavioral Finance theory, and the Bitcoin's Turnover (TO), as a proxy for market liquidity and speculative intensity reflecting the velocity of information flow in cryptoassets.

The inclusion of the FGI is theoretically grounded in Behavioral Finance, particularly in the literature on investor sentiment, which argues that emotional and psychological factors systematically affect price dynamics and volatility in financial markets. Recent evidence suggests that sentiment indicators are key predictors of cryptocurrency volatility (López-Cabarcos et al., 2021), reinforcing the relevance of incorporating such measures into volatility models to capture "noise trader" effects.

Similarly, the TO ratio is supported by the theoretical framework of market microstructure and the Mixture of Distributions Hypothesis. As an on-chain metric that scales trading volume relative to circulating supply, TO serves as a standardized indicator of market "heat" and participation levels. Unlike absolute volume, turnover captures the relative intensity of asset turnover, signaling periods of information arrival or speculative exhaustion that typically precede volatility clusters. Thus, its inclusion allows the volatility process to incorporate both behavioral shocks and liquidity-driven dynamics, bridging traditional financial econometrics with blockchain-specific activity indicators.

Finally, the number of cryptocurrencies analyzed in the study was increased to 10, whereas related empirical studies usually focus on 3 to 5 currencies, thereby expanding the empirical scope and enhancing the generalizability of the findings.

2 Theoretical background

2.1 Cryptocurrencies

According to Lánský (2017), since the creation of Bitcoin by Nakamoto (2008), several investors have become interested in this new type of asset, mainly due to its unique characteristics such as:

1. Being decentralized, i.e., not regulated by governments, banks or companies;
2. The realization of peer-to-peer transactions, eliminating the need for intermediaries and consequently part of the transaction costs;
3. Use cryptographic techniques to secure transactions and protect the integrity of the network, as well as making transactions virtually anonymous between the parties;
4. Blockchain system, which records all transactions made in a public ledger in order to avoid possible fraudulent transactions.

Due to Bitcoin's rise in popularity and the spread of the blockchain system, new cryptocurrencies have been created and spread around the world, known as altcoins. These are generally designed to change some aspectaspects of the system offered by Bitcoin. The main changes they propose concern the speed of transactions, changes to the mining system and the volume of availability (Gandal & Halaburda, 2014).

However, as Wang et al. (2020) point out, the high degree of volatility in these assets ended up creating a barrier to investors making stable gains. With this in mind, the so-called stablecoins were created, cryptocurrencies that aim to peg their prices to currencies (Theter, BitUSD and Nubits, for example) or even commodities (HelloGold, DigixDAO and Xaurum, for example).

Although this market is relatively new, it is still booming. In this sense, several new cryptocurrencies are created all the time, and in May 2024 there were a total of 10,043 cryptocurrencies, totaling a capitalization of around USD 2.5 trillion. Of this amount, around

54.28 per cent corresponds to Bitcoins, 15.44 per cent to Ethereum, 4.57 per cent to Tether, 3.53 per cent to Binance and 3.19 per cent to Solana. Finally, Figure 1 shows the timeline of the main events related to this market.

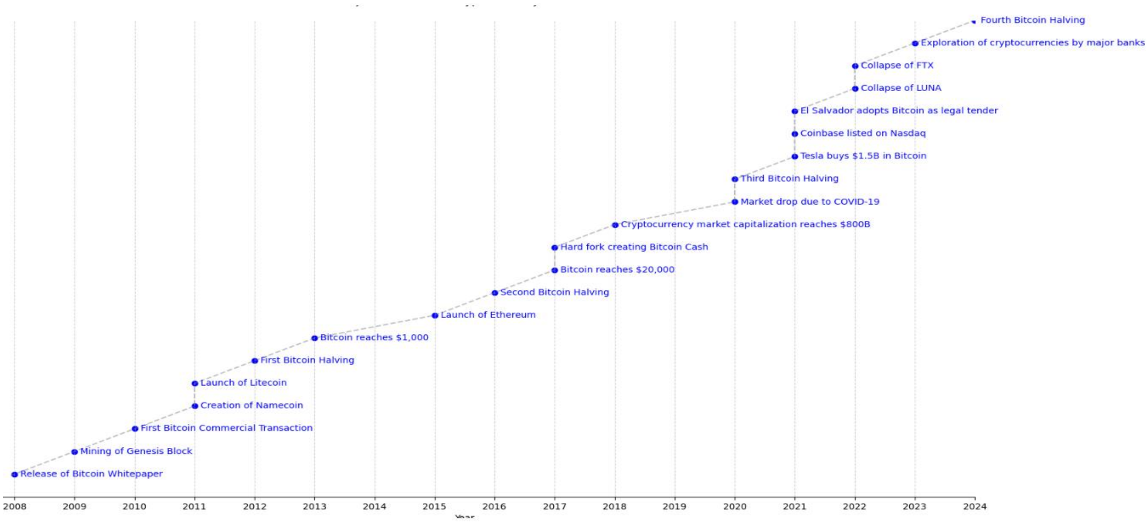

Figure 1 - Main Events in the Cryptocurrency Market

In general, one of the most important events for this market is Bitcoin's halving periods. This process occurs every time 210,000 blocks are added to the blockchain and affects the amount paid for the mining process, thus limiting the creation of more Bitcoins. Initially, the mining process was paid 50 BTC and with each halving this amount is divided by two. A total of four such processes have already taken place, the first in 2012, one in 2016 and another in 2020, with the first half of 2024 being the last recorded halving.

Other important recent events for these markets include: the ban on cryptocurrency mining and transactions in China, causing a significant drop in prices; the collapse of LUNA, UST, crypto hedge fund Three Arrows Capital and FTX, which affected the degree of confidence in the cryptocurrency market; and the creation of CBDCs (Central Bank Digital Currencies), digital currencies created by central banks, which in principle goes against the decentralization aspect of cryptocurrencies.

2.2 Risk Measures

Risk is essentially the possibility that the actual return on an investment will differ from the expected return. Markowitz (1952), a pioneer in modern portfolio theory, defines risk in terms of the volatility of returns, measured by metrics such as standard deviation and variance, which quantify the dispersion of results around the expected average. In this sense, risk management becomes an essential tool for making investment decisions, aimed at helping investors maximize their returns adjusted to a certain degree of risk that is compatible with their preferences.

According to Acereda et al. (2020), Value-at-Risk (VaR) is one of the traditional measures most commonly used to measure the risk of an asset. It can be calculated using Equation 1, in which μ and σ are the mean and standard deviation of the asset in question and Z_{α} is the critical value of the normal distribution corresponding to the confidence level α, generating the well known relative VaR:

$$VaR_{\alpha} = \mu + \sigma * Z_{\alpha} \quad (1)$$

In this way, this metric helps to identify the maximum loss value of an asset given an error tolerance. It is worth noting that VaR can also be calculated based on historical data or via simulation models.

However, as criticized by Acerbi and Tasche (2002), this metric may not be the most suitable for analyzing risk as it is limited to identifying the maximum potential loss. The authors therefore present Expected Shortfall (ES) as a more assertive alternative for risk management.

Also known as Conditional Value-at-Risk (CVaR), it differs from traditional VaR in that ES estimates the average loss in worst-case scenarios in addition to VaR. In other words, ES calculates the average loss that exceeds VaR, providing a more complete view of extreme risk (Acerbi & Tasche, 2002; Artzner et al., 1999). In general terms, the ES is given by.

$$ES_{\alpha} = \mathbb{E}[X|X > VaR_{\alpha}] = \frac{1}{1-\alpha}\int_{0}^{1-\alpha} VaR_{\alpha} \quad (2)$$

It is worth noting that these methodologies originally defined asset returns as belonging to a normal distribution. However, empirical studies such as those by Malek et al.(2023) , Fung et al. (2022) Acereda et al. (2020) and Conlon and McGee (2020), identifying leptokurtic properties and asymmetries incompatible with Gaussian models, opted to use alternative distributions to describe cryptocurrencies' returns, such as the stable alpha distribution, Student's t, Azzalini-Skew-t and inverse normal. In addition, these studies identified clusters of return volatility, so ARMA-GARCH models were used to model the heteroscedasticity of cryptocurrency returns.

2.3 Previous studies

In general, studies on the cryptocurrency market can be divided into four groups:

1. Analyzing the characteristics of this market, comparing cryptocurrencies with financial assets such as commodities, shares and gold (Baur et al., 2018; Cai et al., 2023; Charfeddine et al., 2020);
2. Analyzing the efficiency of this market (Abreu et al., 2022; Kristoufek & Vosvrda, 2019; Urquhart, 2016);
3. Analyzing the volatility of these assets and their hedging capacity (Baur & Dimpfl, 2018; Hasan et al., 2024; Trimborn & Härdle, 2018).
4. Examining behavioral and systematic determinants of cryptocurrency returns and volatility, including investor sentiment indicators and common risk factors (Liu et al., 2022; López-Cabarcos et al., 2021; Wei et al., 2023).

Of these four typologies, the present study is primarily situated within the third strand, as its core objective is the modelling of conditional volatility and tail risk. However, it extends this framework by incorporating behavioral and valuation-based exogenous variables, thereby integrating elements from the fourth strand into a volatility modelling context.

In order to study the hedging properties of stablecoins, Wang et al. (2020) analyzed the relationship between two groups of these cryptocurrencies, one linked to the dollar and the other to gold, and Bitcoin, Litecoin and Ripple using a GARCH model with dynamic conditional correlations (DCC-GARCH) proposed by Engle (2002). As a result, the authors identified that stablecoins pegged to the dollar had a greater hedging capacity than the three traditional cryptocurrencies. Subsequently, in order to empirically verify this fact, portfolios were generated combining Bitcoin, Litecoin or Ripple with one of the stablecoins. Finally, the VaRs and ES of the portfolios were analyzed, confirming the hypothesis discussed above.

In order to identify more efficient alternatives for modelling the volatility of Bitcoin, Litecoin, Ripple and Ethereum, Acereda et al. (2020) estimated the ES using traditional GARCH modelling and three of its variations: component GARCH (csGARCH), non-linear GARCH (nGARCH) and threshold GARCH (tGARCH). The conclusions identified a better fit for the models when considering that cryptocurrency returns would follow an Azzalini-Skew-t distribution. Furthermore, it was concluded that csGARCH and nGARCH models provided more accurate predictions for cryptocurrencies, especially Bitcoin.

The study by Fung et al. (2022) draws an important conclusion about forecasting and risk analysis models. Analyzing eight different models from the GARCH family, the authors found that of the 254 cryptocurrencies analyzed, around one third had their returns best explained by a tGARCH model. However, when analyzing the VaR of the returns, the authors report that the choice of an appropriate distribution for modelling the errors is more important for the accuracy of the VaR than the GARCH specification itself. This reinforces similar conclusions, such as those of Troster et al. (2019) and Ngunyi et al. (2019), regarding the importance of distributional assumptions when modelling cryptocurrencies.

Using data from Bitcoin, Ethereum, Ripple, Litecoin and Bitcoin Cash, Malek et al. (2023) proposed analyzing changes in the VaR and ES of these cryptocurrencies before, during and after the COVID-19 pandemic. Their results indicated that the stable alpha distribution appears to be a viable alternative for estimating VaR and ES, as well as confirming the increase in downside risk of these assets after the start of the pandemic. This conclusion corroborates the similar study carried out previously by Conlon and McGee (2020).

Finally, Huang et al. (2024) analyzed the volatilities of Bitcoin, Ethereum and Binance Coin returns using GARCH modelling and concluded that in only 6, 3 and 2 days, respectively for each asset, the VaR identified was underestimated. Thus, the authors confirm the validity of applying GARCH modelling to cryptocurrency volatility analysis, but at the same time recognize that the use of VaR alone may not be sufficient to capture extreme volatility. This conclusion supports the joint use of ES as a complementary and more coherent measure of tail risk.

Beyond traditional volatility modelling, recent studies have emphasized the role of systematic risk components and behavioral dynamics in cryptocurrency markets. Liu et al. (2022) identify common risk factors that explain cross-sectional cryptocurrency returns, suggesting that digital assets exhibit systematic structures comparable to those observed in traditional financial markets. From a behavioral perspective, López-Cabarcos et al. (2021) provide evidence that investor sentiment significantly influences Bitcoin volatility. Similarly, Wei et al. (2023) combine volatility modelling with sentiment indicators and machine learning techniques, demonstrating that hybrid approaches can improve forecasting performance.

Although prior research validates the application of GARCH-type models for cryptocurrency risk estimation and highlights the importance of distributional assumptions and crisis periods, most studies rely primarily on univariate return dynamics or focus on forecasting accuracy rather than tail-risk measurement. Moreover, there remains limited empirical evidence on whether integrating behavioral sentiment indicators and blockchain-based valuation metrics directly into ARMA-GARCH specifications enhances the estimation of extreme downside risk measured through VaR and Expected Shortfall across a broad set of leading cryptocurrencies. This gap motivates the present study, which evaluates an extensive set of GARCH-family models while incorporating both proxies for market liquidity and for investor sentiment."

3 Methodology

3.1 Data

For this study, data was collected on the dollar quotations of 10 cryptocurrencies from 05/14/2019 to 02/14/2026, thus almost 7 years for the analyses. The assets chosen for the sample were based on the market capitalization of the cryptocurrencies, excluding stablecoins and cryptocurrencies that did not exist throughout the period.

The following cryptocurrencies were therefore selected for this study: Bitcoin (BTC), Ethereum (ETH), Ripple (XRP), Dogecoin (), Binance (BNB), Cardano (ADA), Tronix (TRX), Bitcoin Cash (BCH), Chainlink (LINK) and Litecoin (LTC). The price data for these assets was collected via Yahoo Finance and the logarithmic returns for each series were calculated from this data.

In addition, data was collected on two series of exogenous variables: the Fear and Greed Index (FGI) and the Bitcoin's Turnover (TO). The TO was employed as a proxy for market liquidity and speculative intensity, as it reflects the velocity of trading activity relative to the circulating supply, providing a standardized measure of market 'heat' across different price cycles. Complementarily, the FGI was utilized as a proxy for investor sentiment and behavioral bias, which captures the psychological state of the market. Its inclusion in the variance equation is justified by the 'noise trader' theory, suggesting that periods of extreme fear or greed often lead to asymmetric volatility clusters that traditional GARCH models might not fully capture through returns alone.

3.2 Time Series Models

In order to calculate VaR and ES, GARCH-type models were employed, as they allow not only the estimation of asset return forecasts but also the prediction of the conditional variance σ_t^2, which is fundamental for tail-risk measurement. The traditional ARMA(p,q)-GARCH(r,m) specification is represented by Equation 3, where the error $\varepsilon_t \sim D(0, \sigma_t^2)$ and D represents a predefined distribution.

$$R_t = \phi_0 + \sum_{i=1}^{p} \phi_i R_{t-i} + \varepsilon_t + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j}$$

$$Var(\varepsilon_t) = \sigma_t^2 = \alpha_0 + \textstyle\sum_{k=1}^{r} \alpha_k \varepsilon_{t-k}^2 + \epsilon_t + \sum_{l=1}^{m} \beta_l \sigma_{t-l}^2 \qquad (3)$$

As studies of the limitations of the GARCH model have progressed, alternatives have been developed, aimed above all at increasing the robustness of the forecasts. To illustrate the modelling used in this study, Table 1 summarizes the modifications made to each of the 6 variations selected.

Table 1
GARCH Modelling Variations Used in the Study

Variação GARCH	Equação
Exponential Generalized Autoregressive Conditional Heteroscedastic - eGARCH	$\ln \sigma_t^2 = \alpha_0 + \sum_{k=1}^{r} \alpha_k \frac{\lvert\varepsilon_{t-k}\rvert}{\sigma_{t-k}} + \sum_{k=1}^{r} \gamma_k \frac{\lvert\varepsilon_{t-k}\rvert}{\sigma_{t-k}} \epsilon_t + \sum_{l=1}^{m} \beta_l \ln \sigma_{t-l}^2$
Threshold Generalized Autoregressive Conditional Heteroscedastic - tGARCH	$\sigma_t^2 = \alpha_0 + \sum_{k=1}^{r} \alpha_k \varepsilon_{t-k}^2 + \sum_{o=1}^{g} I_{t-o} \gamma_k \varepsilon_{t-o}^2 + \sum_{n=1}^{h} J_{t-n} \delta_n \sigma_{t-n}^2 \epsilon_t + \sum_{l=1}^{m} \beta_l \sigma_{t-l}^2$
Asymmetric Power Autoregressive Conditional Heteroscedastic - apARCH	$\sigma_t^\delta = \alpha_0 + \sum_{k=1}^{r} \alpha_k (\lvert\varepsilon_{t-k}\rvert - \gamma_k \varepsilon_{t-k})^\delta + \epsilon_t + \sum_{l=1}^{m} \beta_l \sigma_{t-l}^\delta$
Component Generalized Autoregressive Conditional Heteroscedastic - csGARCH	$\sigma_t^t = q_t + h_t$ Long-term component: $q_t = \omega_0 + p_0 q_{t-1} + \sum_{k=1}^{r} \phi_k (\varepsilon_{t-k}^2 - \sigma_{t-k}^2)$ Short-term component: $h_t = \beta_0 h_{t-1} + \sum_{l=1}^{m} \alpha_k (\varepsilon_{t-k}^2 - q_{t-k})$
Glosten, Jagannathan and Runkle Generalized Autoregressive Conditional Heteroscedastic - gjrGARCH	$\sigma_t^2 = \alpha_0 + \sum_{k=1}^{r} \alpha_k \varepsilon_{t-k}^2 + \sum_{o=1}^{g} I_{t-o} \gamma_o \varepsilon_{t-o}^2 + \epsilon_t + \sum_{l=1}^{m} \beta_l \sigma_{t-l}^2$
Non Linear Generalized Autoregressive Conditional Heteroscedastic - nGARCH	$\sigma_t^2 = \alpha_0 + \sum_{k=1}^{r} \alpha_k f(\varepsilon_{t-k})^2 + \epsilon_t + \sum_{l=1}^{m} \beta_l \sigma_{t-l}^2$

In general, both the gjrGARCh and tGARCH models incorporate a threshold component when modelling conditional volatility, allowing it to respond differently to positive and negative shocks. The eGARCH, nGARCH and apARCH models work with non-linear variations for volatility modelling. Finally, the csGARCH model divides conditional volatility into distinct components, each modelled separately. This provides greater flexibility in modelling different aspects of volatility, such as short-term and long-term volatility (Charles & Darné, 2019).

Furthermore, it is important to emphasize that the error ε_t must be assumed to follow a predefined distribution, originally defined as the normal distribution. However, previous studies, as described in section 2.3, identified a divergence between the empirical and theoretical distribution of errors, so other distributions that take into account asymmetry and kurtosis began to be used to estimate GARCH models. Table 2 summarizes some of the distributions commonly and their characteristics.

Tabela 2

Characteristics of Some Distributions Used in GARCH Models

Distribution	Symmetry	Tails	Kurtosis	Skewness	Parameters
Normal (N)	Symmetric	Light	Mesokurtic	No	μ - mean σ^2 - variance
Skewed Normal (SN)	Asymmetric	Light	Close to mesokurtic	Yes	μ - location σ^2 - scale λ - skewness
Student's t (St)	Symmetric	Heavy	Leptokurtic	No	μ - mean σ^2 - variance ν - degrees of freedom
Skewed Student's t (SSt)	Asymmetric	Heavy	Leptokurtic	Yes	μ - location σ^2 - scale ν - degrees of freedom λ - skewness
Generalized Error Distribution (GED)	Symmetric	Flexible	Depends on shape parameter κ	No	μ - mean σ^2 - variance κ - shape
Skewed GED (SGED)	Asymmetric	Flexible	Depends on shape parameter κ	Yes	μ - location σ^2 - scale κ - shape λ - skewness
Inverse Normal (IN)	Asymmetric (Positive Skew)	Right-heavy	Leptokurtic	Yes (positive skew)	μ - mean λ - shape

Generalized Hyperbolic (GH)	Symmetric or asymmetric	Very heavy	Typically Leptokurtic	Yes	μ - location δ - scale γ - shape δ - shape λ - tail index
Johnson's SU (JSU)	Symmetric or asymmetric	Very flexible	Leptokurtic or Platykurtic	Yes	ξ - location λ -scale γ - skewness δ - kurtosis

Using the free R software, ARMA(p,q)-GARCH(r,m) models were run for each of the 10 cryptocurrencies selected. To analyze the best modelling, three variations were considered:

1. Parameters *p,q,r,m*: were allowed to vary within the interval [0,2]. Configurations in which $(p, q) = (0,0)$ or $(r, m) = (0,0)$ were excluded, as they imply the absence of conditional mean and/or variance dynamics. After applying this restriction, 64 admissible ARMA-GARCH order combinations remained.

2. Error Distribution: the following distributions were considered: normal, skewed normal, Student's t, skewed Student's t, generalized error distribution, skewed generalized error distribution, inverse normal, Generalized Hyperbolic and Johnson's SU distribution.

3. GARCH variations: finally, in addition to traditional GARCH modelling, eGARCH, apARCH, csGARCH, tGARCH, gjrGARCH and nGARCH models were also run.

All additional model-specific parameters (e.g., asymmetry, power and component terms) and distribution-specific parameters (e.g., degrees of freedom, skewness and shape parameters) were jointly estimated with the mean and volatility parameters via maximum likelihood within the rugarch framework. Combining 64 order configurations, 9 alternative distributions and 7 volatility specifications results in 4,032 potential model structures for each cryptocurrency. It is worth noting that the limitations of the parameters p, q, r, and m between [0,2] align with the findings of Silva and Maciel (2023), which concludes that investors should prioritize parsimonious GARCH structures, as excessive complexity often fails to capture the idiosyncratic volatility behavior of digital assets

Furthermore, for each model, 80% of the sample was used for estimation and the remaining 20% for out-of-sample evaluation. Model selection was primarily based on forecasting performance, with the Mean Absolute Error (MAE) serving as the main ranking criterion. To ensure robustness, additional performance measures, including the Mean Squared Error (MSE) and the Root Mean Square Error (RMSE), were also analyzed. In-sample model adequacy was assessed using the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), which enable comparison across competing specifications while penalizing excessive model complexity.

After identifying the best-performing specification for each cryptocurrency, the selected model was employed to estimate Value at Risk (VaR) and Expected Shortfall (ES). To allow risk measures to adapt to evolving market conditions, a rolling window of 250 observations was implemented. Risk management performance was evaluated through backtesting procedures, computing violation frequencies relative to the total number of out-of-sample observations.

The statistical adequacy of the VaR forecasts was rigorously assessed using a battery of hypothesis tests. We employed the Kupiec's Unconditional Coverage test to verify if the empirical proportion of violations was consistent with the theoretical significance level. Furthermore, to ensure that violations were not serially dependent, a critical requirement for risk management during market crashes, the Christoffersen's Conditional Coverage test was also applied. This test simultaneously evaluates both the correct frequency of exceptions and their independence over time.

Complementing these, the Dynamic Quantile DQ test was utilized to detect higher-order serial correlation in the hit sequence. Finally, the accuracy of the ES estimates was validated through Exceedance Residuals backtesting, ensuring that the model captures the magnitude of losses in the distribution's tail beyond the VaR threshold

3.3 Inclusion of Exogenous variables

The final stage of the analysis consists of assessing whether GARCH modelling performance improves after incorporating the FGI and TO as exogenous variables. The objective is to examine whether these indicators contribute to enhancing volatility forecasting and, consequently, tail-risk estimation. To this end, the procedure described in Section 3.2 was repeated under the augmented GARCH-X specification. After estimating the models and computing VaR and ES, the resulting risk measures were compared with those obtained from the baseline univariate models to evaluate the marginal predictive power of market sentiment and liquidity.

Beyond the traditional ARMA-GARCH framework, the models were extended by including exogenous variables directly into the conditional variance equation, yielding a GARCH-X specification. Specifically, the FGI and the TO were incorporated as predictors of conditional volatility rather than expected returns. The extended variance equation is given by Equation 4, where the coefficients for FGI and TO measure the marginal impact of investor sentiment and market liquidity on the asset's risk dynamics.

$$R_t = \phi_0 + \sum_{i=1}^{p} \phi_i R_{t-i} + \varepsilon_t + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j}$$

$$Var(\varepsilon_t) = \sigma_t^2 = \alpha_0 + \sum_{k=1}^{r} \alpha_k \varepsilon_{t-k}^2 + \epsilon_t + \sum_{l=1}^{m} \beta_l \sigma_{t-l}^2 + \gamma_1 FGI_{t-1} + \gamma_2 TO_{t-1} \quad (4)$$

Both exogenous variables were included with a one-period lag to mitigate potential simultaneity concerns and preserve the predictive structure of the model. By incorporating these variables into the variance equation, the model explicitly accounts for how external behavioral shocks and trading intensity drive the clustering of volatility, providing a more robust framework for risk management in highly speculative environments like the cryptocurrency market.

4 Results

4.1 Descriptive Statistics

From the data on the logarithmic returns of the cryptoassets, the average returns and annualized standard deviation were calculated, as well as the Sharpe Ratio (SR) and the historical VaR and ES. In addition, the Kwiatkowski-Phillips-Schmidt-Shin (KPSS), Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests were estimated to analyze the stationarity of the series. The results are shown in Table 3.

Table 3

Basic Statistics of the Cryptocurrency Returns Series

	Average Annualized Returns	Annualized Standard Deviation	SR	VaR 5%	ES 5%	KPSS	ADF	PP
BTC	0.3204	0.6314	0.5074	-0.0488	-0.0777	0.1184	-12.5183**	-2717.063**
ETH	0.3346	0.8248	0.4057	-0.0663	-0.1029	0.2155	-12.7696**	-2711.246**
XRP	0.1931	0.9959	0.1939	-0.0697	-0.1158	0.0565	-12.6431**	-2691.213**
DOGE	0.5309	1.2841	0.4135	-0.0763	-0.1258	0.1995	-12.7267**	-2510.619**
BNB	0.4864	0.8306	0.5856	-0.0613	-0.0991	0.1699	-12.1315**	-2844.421**
ADA	0.185	0.981	0.1886	-0.0761	-0.1128	0.2509	-11.8804**	-2713.377**
TRX	0.3444	0.8434	0.4084	-0.0609	-0.1055	0.0527	-14.1579**	-2684.058**
BCH	0.0547	0.9593	0.057	-0.0704	-0.116	0.0812	-14.135**	-2803.242**
LINK	0.3552	1.0882	0.3264	-0.082	-0.1283	0.3003	-13.702**	-2642.496**
LTC	-0.0728	0.8866	-0.0821	-0.0729	-0.1148	0.0521	-14.0464**	-2523.822**

Note: ***; **; * indicate significance at 1%, 5% and 10% respectively.

Analyzing the table provided reveals several important characteristics about the different cryptocurrencies in terms of returns, volatility, adjusted risk and the statistical properties of the time series. Firstly, when looking at the average annualized returns, DOGE stands out with the highest return (0.5309), followed by BNB (0.4864) and LINK (0.3552). In contrast, LTC shows a negative return (-0.0728), indicating an average annual loss.

In terms of volatility, DOGE again stands out, but this time negatively, with the highest annualized standard deviation (1.2841), signaling high volatility. BTC and ETH, on the other hand, show lower volatility relative to the group, with standard deviations of 0.6314 and 0.8248 respectively.

The SR, which measures risk-adjusted return, indicates that BNB has the best performance (0.5856), suggesting a good balance between return and risk. On the other hand, LTC has a negative Sharpe Ratio (-0.0821), suggesting that the associated risk does not compensate for the return.

Analyzing the VaR at 5%, which estimates the maximum loss expected under normal market conditions for 5% of the worst-case scenarios, we see that LINK has the highest negative VaR (-0.0820), indicating greater potential losses. BTC has the lowest negative VaR (-0.0488), suggesting lower potential losses. The ES at 5%, which calculates the average expected loss in the 5% worst cases, shows that LINK also has the highest expected loss (-0.1283), closely followed by DOGE (-0.1258), indicating greater extreme risk. BTC again stands out positively with the lowest expected loss (-0.0777).

The KPSS, ADF and PP stationarity tests provide insight into the nature of the time series of cryptocurrency returns. It is important to emphasize that the KPSS establishes the null hypothesis that the series is stationary and therefore has no unit root, while the ADF and PP tests have the null hypothesis that the series has a unit root. Since the ADF and PP statistics are highly significant (as indicated by the asterisks) and the KPSS values fail to reject the null hypothesis of stationarity at conventional levels, it is possible to corroborate that cryptocurrency returns are stationary series, which supports the use of ARMA-GARCH models.

4.2 ARIMA models and ARCH tests

Continuing the analyses of the time series, the validity of applying models for heteroscedasticity of the variances of the time series was checked. To do this, the Lagrange Multiplier (LM) test was applied to ARCH modelling.

However, to carry out this procedure, it is first necessary to calculate the best ARMA model for each time series so that the LM test can then be applied, testing the null hypothesis of no correlation between the model's residuals. The best ARMA models were estimated for the 10 series of cryptocurrency returns using the auto.arima method in the forecast package, which identifies the values of p, d and q that minimise the AIC and SBIC criteria. Based on the best modelling, LM tests were carried out, the results of which are shown in Table 4.

Table 4

LM ARCH Test Results

	Model	Order of the LM ARCH tests					
		4	**8**	**12**	**16**	**20**	**24**
BTC	ARIMA(0,0,4)	14.01***	45.19***	47.02***	48.81***	50.89***	51.69***
ETH	ARIMA(1,0,1)	42.43***	59.05***	60.16***	62.25***	64.82***	65.61***
XRP	ARIMA(2,0,2)	65.88***	79.96***	80.2***	89.03***	91.41***	102.37***
	ARIMA(2,0,2)	30.19***	35.05***	37.16***	37.16***	37.3***	37.47***
BNB	ARIMA(2,0,2)	120.86***	136.35***	187.68***	191.07***	199.94***	200.42***
ADA	ARIMA(2,0,2)	62.77***	79.54***	80.22***	82.87***	82.74***	82.79***
TRX	ARIMA(0,0,5)	23.45***	32.93***	33.31***	37.55***	38.01***	38.84***
BCH	ARIMA(1,0,4)	24.64***	58.53***	59.14***	82.14***	94.06***	96.67***
LINK	ARIMA(0,0,5)	40.86***	64.48***	70.16***	80.72***	82.25***	83.65***
LTC	ARIMA(2,0,3)	52.85***	87.4***	90.7***	96.21***	96.76***	98.04***

Note: ***; **; * indicate significance at 1%, 5% and 10% respectively.

The results show that there is a correlation between the residuals of the ARIMA models that minimize the AIC and SBIC criteria. This corroborates the application of modelling for the volatility of the residuals, which is discussed in the next section.

4.3 GARCH models

Based on the previous analyses, the validity of applying models to estimate the heteroscedasticity of the selected cryptocurrency series was verified. Thus, using the rugarch package, 4,032 variations of ARMA-GARCH models were estimated, considering 80% of the total sample of return series. Forecasts were then made for the remaining 20% of the sample, selecting the configuration that minimised the MAE for each asset. The results are shown in Table 5**.**

Table 5

Results of the Best ARMA-GARCH Models for Cryptocurrencies

Cripto	Error Distribution	GARCH	Model	AIC	BIC	MAE	MSE	RMSE
BTC	GED	eGARCH	ARMA(1,2) GARCH(2,1)	4.978	5.01	0.016966	0.000594	0.024373
ETH	NI	apARCH	ARMA(2,2)	5.522	5.553	0.027532	0.001543	0.03928

			GARCH(1,0)					
XRP	St	gjrGARCH	ARMA(1,0) GARCH(2,0)	5.52	5.543	0.031667	0.002374	0.048727
	St	csGARCH	ARMA(1,1) GARCH(2,0)	5.643	5.669	0.037657	0.002774	0.052668
BNB	GH	apARCH	ARMA(2,2) GARCH(2,0)	5.386	5.425	0.019524	0.000802	0.028316
ADA	JSU	gjrGARCH	ARMA(2,2) GARCH(1,0)	5.837	5.865	0.037213	0.003028	0.055026
TRX	NI	gjrGARCH	ARMA(2,2) GARCH(1,0)	5.286	5.314	0.01774	0.001544	0.039299
BCH	JSU	sGARCH	ARMA(2,2) GARCH(2,0)	5.708	5.736	0.029038	0.001606	0.040073
LINK	St	sGARCH	ARMA(2,2) GARCH(0,2)	6.158	6.183	0.036894	0.002563	0.050621
LTC	GH	sGARCH	ARMA(2,2) GARCH(2,0)	5.654	5.686	0.030338	0.001861	0.043141

Firstly, concerning predictive accuracy and parsimony, the results confirm a high level of precision across all assets. The low values of MAE, ranging from 0.0169 for BTC to 0.0376 for DOGE, demonstrate that the models, even under parsimonious constraints where the parameters p, q, r, and m are restricted to between 0 and 2, provide robust predictive power for return dynamics. Secondly, the selection of error distributions underscores that the standard Normal distribution is insufficient for the cryptocurrency market. The predominant selection of complex distributions, such as GH and JSU, indicates that the models effectively account for the fat-tailed and asymmetric nature of crypto returns, which is essential for accurate tail-risk measurement.

In addition, the variation in model architecture highlights significant model specialization. Assets exhibiting higher volatility, such as DOGE and ADA, frequently utilized asymmetric GARCH variants like csGARCH and gjrGARCH. This suggests that leverage effects are a fundamental feature of these assets, requiring specialized modelling to capture extreme conditional variance correctly. Finally, assets with lower standard deviations and higher market maturity, such as BTC, utilize more compact models like eGARCH, which are highly effective at capturing volatility persistence without necessitating over-parameterized structures. In contrast, emerging or higher-risk assets like BNB or ETH rely on apARCH models, indicating that both volatility and returns require more flexible modelling frameworks to mitigate forecast errors effectively.

These conclusions corroborate analyses by Fung et al. (2022), Acereda et al. (2020), and Ngunyi et al. (2019) regarding the importance of tailoring the GARCH specification and the assumed error distribution individually for each asset. Also, the consistency across these different error distributions reinforces the conclusion that a "one-size-fits-all" approach to cryptocurrency risk estimation is inadequate, necessitating the heterogeneous modelling strategy adopted in this study (Sözen, 2025).

4.4 Inclusion of External Regressors

In order to check whether the inclusion of two exogenous variables, the FGI and the TO, would increase the predictive capacity of the volatility of cryptocurrency returns, ARMA-GARCH-X models were estimated, and Table 6 shows the results of the models that minimised the MAE;

Table 6
Results of Best ARMA-GARCH-X Models with Exogenous variables for Cryptocurrencies

Cripto	Error Distribution	GARCH	Model	AIC	BIC	MAE	MSE	RMSE
BTC	GED	eGARCH	ARMA(1,2) GARCH(2,0)	5.02	5.054	0.016966	0.00059428	0.024378
ETH	GH	gjrGARCH	ARMA(1,1) GARCH(2,0)	5.516	5.553	0.027387	0.00153161	0.039136
XRP	SSt	apARCH	ARMA(2,2) GARCH(1,0)	5.55	5.587	0.031798	0.00242896	0.049284
	St	csGARCH	ARMA(2,2) GARCH(2,0)	5.646	5.683	0.037659	0.00277028	0.052633
BNB	St	gjrGARCH	ARMA(2,2) GARCH(2,0)	5.371	5.408	0.019533	0.00080562	0.028383
ADA	GH	sGARCH	ARMA(2,2) GARCH(0,1)	5.816	5.85	0.036289	0.00296826	0.054482
TRX	JSU	eGARCH	ARMA(2,2) GARCH(1,0)	5.243	5.277	0.017419	0.00150107	0.038744
BCH	St	csGARCH	ARMA(1,2) GARCH(2,1)	5.674	5.711	0.02906	0.00160396	0.040049
LINK	GH	sGARCH	ARMA(2,2) GARCH(2,0)	6.124	6.16	0.036966	0.00257974	0.050791
LTC	GED	sGARCH	ARMA(2,0) GARCH(2,2)	5.63	5.662	0.029629	0.00184531	0.042957

From the data in Table 6, it can be seen that the inclusion of FGI and TO did not generate a uniform effect on the indicators of the ARMA-GARCH models used. Regarding the predictive performance, assets such as ETH, ADA, TRX, and LTC presented a reduction in their error metrics (MAE, MSE, and RMSE) when compared to the baseline models in Table 5, reflecting improvements in forecasts. Conversely, other cryptocurrencies showed a slight increase in these values.

However, it is worth noting that all variations obtained were marginal, signaling, a priori, a low forecasting gain when adding the external regressors. Such results are consistent with Liu et al. (2022) and Orujov et al. (2025), which emphasize that including more covariates than necessary may lead to overfitting.

Furthermore, it can be seen that the inclusion of external regressors in the modelling altered the parameters of the time series models. Thus, it can be concluded that the inclusion of exogenous variables can alter the dynamics of cryptocurrency returns, reflecting new patterns of error variance, which in turn require more flexible distributions to be properly captured.

A fundamental observation lies in the sensitivity of the model structure to the inclusion of these variables. When incorporating FGI and TO, a reconfiguration occurs not only in the parameters but in the taxonomy of the GARCH specification and the error distribution. This transition indicates that exogenous variables act as "fine-tuners": by capturing part of the variability that previously remained in the error term, they allow the model to isolate volatility shocks more effectively. As the model becomes more efficient in absorbing the impact of external information, the selection of the GARCH structure and the error distribution adjust dynamically to this new information.

For instance, in several cases, such as ETH, which shifted from NI to GH, and BCH, which migrated from JSU to St, this reconfiguration signals that exogenous information permitted the model to refine its risk-capture mechanism. This adjustment process suggests that the choice between a simpler or more complex GARCH structure is not static. Rather, it is a response to the information provided by the regressors. In some instances, the inclusion of exogenous variables reduced the need for over-parameterization, allowing more parsimonious models to achieve predictive performance comparable to more complex structures that operated without these variables.

In summary, the inclusion of exogenous variables in the ARMA-GARCH models resulted in adjustments to the model structures and distributions, reflecting occasional improvements in the prediction of some cryptocurrencies and minimal variation in overall predictive performance. The models indicate that while the FGI and TO are valuable diagnostic tools, the idiosyncratic volatility of each cryptocurrency remains the primary driver of forecast errors. Ultimately, these results demonstrate that external regressors are most effective when paired with GARCH specifications uniquely tailored to the asset's specific volatility profile, rather than serving as a universal booster for predictive accuracy.

4.5 VaR and ES estimates

As discussed above, the inclusion of external regressors affects the choice of time series model parameters that minimize the MAE. In this sense, it is essential to comparatively analyze the volatility estimation performance of models while controlling the inclusion of external regressors. To assess the effective gains in risk management, we first estimated a standard ARMA(1,1)-GARCH(1,1) model assuming that errors are distributed according to a Gaussian distribution (Panel A) as a benchmark. This was compared against the optimized models selected in the previous sections: the "Best ARMA-GARCH" (Panel B) and the "Best ARMA-GARCH-X" (Panel C), which include the exogenous variables. All models were trained on 80% of the sample, with forecasts generated for the remaining 20% out-of-sample period. Table 7 summarizes the Value at Risk (VaR) and Expected Shortfall (ES) results for these configurations

Table 7

Summary of the VaR and ES Estimated

Panel A: ARMA(1.1)-GARCH(1.1)					
	VaR 5% Failure Ration and Kupiec's Unconditional Coverage test	**Christoffersen's Conditional Coverage test**	**ES 5%Failure Ratio and Exceedance Residuals Backtesting**	**VaR 5%**	**ES 5%**
BTC	4.05%	1.05	1.42%***	-0.0448	-0.0566
ETH	4.66%	0.84	3.44%***	-0.0646	-0.0814
XRP	3.24%*	4.74*	1.82%***	-0.0814	-0.1021
DOGE	4.86%	0.05	2.43%***	-0.089	-0.1115
BNB	4.25%	0.63	2.63%***	-0.0499	-0.0628
ADA	4.66%	0.13	2.02%***	-0.0827	-0.1037
TRX	2.43%***	9.01**	1.62%***	-0.0472	-0.0596
BCH	2.23%***	10.5***	1.62%***	-0.0761	-0.0952
LINK	5.06%	0.07	1.62%***	-0.0842	-0.1059
LTC	3.44%*	4.04	2.43%***	-0.072	-0.0903
Panel B: Best ARMA-GARCH					
	VaR 5% Failure Ration and Kupiec's Unconditional Coverage test	**Christoffersen's Conditional Coverage test**	**ES 5%Failure Ratio and Exceedance Residuals Backtesting**	**VaR 5%**	**ES 5%**
BTC	4.45%	0.32	1.42%***	-0.0415	-0.0613
ETH	4.05%	1.05	1.01%***	-0.0659	-0.1031
XRP	5.06%	2.68	1.82%***	-0.0681	-0.1189
DOGE	6.88%*	3.38	3.24%***	-0.0726	-0.1181
BNB	2.63%***	7.9**	0.4%***	-0.0588	-0.0969
ADA	5.47%	3.35	1.21%***	-0.0774	-0.1259
TRX	1.62%***	16.21***	0.2%***	-0.0584	-0.1008
BCH	3.24%*	4.05	1.01%***	-0.0721	-0.1182
LINK	6.48%	2.51	2.23%***	-0.0671	-0.0974
LTC	3.04%**	5.58*	1.42%***	-0.0788	-0.12
Panel C: Best ARMA-GARCH-X					
	VaR 5% Failure Ration and Kupiec's Unconditional Coverage test	**Christoffersen's Conditional Coverage test**	**ES 5%Failure Ratio and Exceedance Residuals Backtesting**	**VaR 5%**	**ES 5%**
BTC	3.04%**	5.58*	1.01%***	-0.0461	-0.0684
ETH	4.25%	2.48	1.42%***	-0.0657	-0.101
XRP	5.67%	3.82	1.42%***	-0.0672	-0.124

DOGE	7.09%**	4.14	3.04%***	-0.0731	-0.1202
BNB	4.25%	0.63	1.82%***	-0.0465	-0.0771
ADA	9.31%***	17.4***	3.04%***	-0.0589	-0.0878
TRX	4.05%	1.05	1.01%**	-0.0446	-0.0786
BCH	3.85%	1.59	1.42%***	-0.0621	-0.1038
LINK	3.44%*	3.09	1.21%***	-0.0885	-0.1471
LTC	3.64%	3.47	2.43%***	-0.0692	-0.0985

Note: ***; **; * indicate significance at 1%, 5% and 10% respectively.

Overall, the calculated VaRs achieved a global failure rate of 4.35% in risk estimation, implying a global success margin of 95.65%. For the ES, these values are 1.76% and 98.24%, respectively. This confirms that the ES provides a lower margin of error for extreme outcomes compared to VaR, consistent with the literature (Acerbi & Tasche, 2002; Acereda et al., 2020; Malek et al., 2023). Thus, it is corroborated that the ES is a more adequate tool for risk management for cryptocurrency investors, aligning with the conclusions of Huang et al. (2024).

Regarding the coverage tests, the Kupiec (1995) unconditional coverage test reveals that the transition from standard benchmarks (Panel A) to optimized specifications (Panels B and C) significantly reshapes risk management profiles. In Panels A and B, we observe a systemic issue for certain assets: TRX, BCH, and LTC (in Panel A), as well as BNB, TRX, BCH, and LTC (in Panel B), exhibited failure rates consistently exceeding the 5% threshold, indicating that baseline ARMA-GARCH models were failing to capture the full extent of tail risks for these cryptocurrencies.

The inclusion of external regressors in Panel C (the Best ARMA-GARCH-X models) successfully addressed these specific coverage failures. By integrating FGI and TO, the models for TRX, BCH, and LTC achieved a more robust risk calibration, effectively bringing their failure rates within acceptable statistical limits. However, this refinement introduces a distinct trade-off in model behavior. While coverage improved for several assets, the model calibration for others became notably polarized. Specifically, DOGE and ADA shifted toward highly conservative estimates, resulting in a significant reduction in observed exceedances that, paradoxically, leads to the rejection of correct coverage due to over-estimation of risk.

Furthermore, the BTC case highlights the sensitivity of these models: while it remained stable in Panels A and B, the inclusion of exogenous variables in Panel C introduced a failure rate that, while marginal, deviates from the perfect coverage target. Ultimately, these results demonstrate that exogenous variables act as "risk-calibration dampers." They are highly effective at correcting model under-coverage for volatile assets, but this adaptive mechanism can induce a conservative bias in assets driven by idiosyncratic liquidity spikes.

The analysis of the ES further validates the efficacy of specialized models in controlling tail risk. Comparing the models, 80% of the assets showed improved ES performance when moving from Panel A to Panel B, with notable gains in BNB (2.63% to 0.40%) and TRX (1.62% to 0.20%). The inclusion of exogenous variables in Panel C provided incremental gains for assets like BTC and XRP, yet for others, such as ADA and BCH, the internal GARCH structure (Panel B) proved more effective than the GARCH-X specification.

This suggests that while GARCH-X models refine the capture of volatility dynamics, the primary driver of ES accuracy remains the asset's specific distribution and GARCH structure. Ultimately, these results confirm that even within an optimized framework, the idiosyncratic nature of each cryptocurrency requires a tailored approach to risk management, as no single ARMA-GARCH-X specification can simultaneously optimize coverage for the entire spectrum of digital assets.

4.6 Robustness Tests

In order to evaluate the significance of the actual gain of more complex models compared to simpler models, the metrics available in Table 8 were computed. The Panel A presents the predictive robustness analysis through the Diebold-Mariano (DM) test, comparing the benchmark ARMA(1,1)-GARCH(1,1) against the optimized "Best ARMA-GARCH" and "Best ARMA-GARCH-X" specifications. The results indicate that for the majority of the cryptocurrencies analyzed, such as BTC, ETH, and LTC, the differences in predictive accuracy are not statistically significant, as evidenced by DM statistics that fail to reach conventional critical thresholds for most of the cases.

This leads that while the optimized models often yield lower Loss Function values, reaching, they do not necessarily provide a statistically superior point-forecast over the simpler benchmark. However, specific idiosyncratic behaviors emerge in the cases of ADA and XRP, which suggests that for this specific asset, the inclusion of FGI and TO may lead to a statistically different and potentially less efficient predictive performance than the optimized pure GARCH model.

Table 8

Summary of Models Metrics

Panel A: Comparations Metrics				*Panel B: Models Metrics*				
Crypto	**Comparation Models**	**Diebold-Mariano**	**Loss Function**	**Crypto**	**Models**	**Mincer-Zarnowitz R2**	**LogLik**	**Time Estimation**
BTC	Best ARMA-GARCH vs ARMA(1,1-)GARCH(1,1)	0.48	-113775	BTC	ARMA(1,1)-GARCH(1,1)	0.52%	-6456.36	1.39
	Best ARMA-GARCH-X vs ARMA(1,1-)GARCH(1,1)	0.62	-118485		Best ARMA-GARCH	0.46%	-6138.17	16.59
	Best ARMA-GARCH-X vs Best ARMA-GARCH	0.3	-125707		Best ARMA-GARCH-X	0.50%	-6189.81	81.78
ETH	Best ARMA-GARCH vs ARMA(1,1-)GARCH(1,1)	-0.7	-174276	ETH	ARMA(1,1)-GARCH(1,1)	0.02%	-6941.87	0.38
	Best ARMA-GARCH-X vs ARMA(1,1-)GARCH(1,1)	0.61	-191949		Best ARMA-GARCH	0.01%	-6808.4	14.59
	Best ARMA-GARCH-X vs Best ARMA-GARCH	0.72	-191807		Best ARMA-GARCH-X	0.23%	-6801.21	21.46
XRP	Best ARMA-GARCH vs ARMA(1,1-)GARCH(1,1)	-0.88	-20822	XRP	ARMA(1,1)-GARCH(1,1)	0.15%	-7296.53	1.09
	Best ARMA-GARCH-X vs ARMA(1,1-)GARCH(1,1)	-1.41	-223603		Best ARMA-GARCH	0.01%	-6806.11	4.05
	Best ARMA-GARCH-X vs Best ARMA-GARCH	-2.1**	-226733		Best ARMA-GARCH-X	0.14%	-6843.42	10.64
DOGE	Best ARMA-GARCH vs ARMA(1,1-)GARCH(1,1)	-1.53	-23444	DOGE	ARMA(1,1)-GARCH(1,1)	0.97%	-7593.59	0.72
	Best ARMA-GARCH-X vs ARMA(1,1-)GARCH(1,1)	-1.51	-24.48		Best ARMA-GARCH	0.70%	-6958.3	6.41
	Best ARMA-GARCH-X vs Best ARMA-GARCH	0.6	-246805		Best ARMA-GARCH-X	0.66%	-6961.38	2.53
BNB	Best ARMA-GARCH vs ARMA(1,1-)GARCH(1,1)	-0.73	-129439	BNB	ARMA(1,1)-GARCH(1,1)	0.58%	-6800.06	0.4
	Best ARMA-GARCH-X vs ARMA(1,1-)GARCH(1,1)	-1.3	-16941		Best ARMA-GARCH	0.03%	-6640.67	90.42
	Best ARMA-GARCH-X vs Best ARMA-GARCH	-1.15	-146687		Best ARMA-GARCH-X	0.16%	-6622.74	6.69
ADA	Best ARMA-GARCH vs ARMA(1,1-)GARCH(1,1)	-1.37	-21759	ADA	ARMA(1,1)-GARCH(1,1)	0.05%	-7292.72	0.56
	Best ARMA-GARCH-X vs ARMA(1,1-)GARCH(1,1)	-0.5	-240676		Best ARMA-GARCH	0.02%	-7196.95	3.35
	Best ARMA-GARCH-X vs Best ARMA-GARCH	1.79*	-205978		Best ARMA-GARCH-X	0.07%	-7171.14	18.06
TRX	Best ARMA-GARCH vs ARMA(1,1-)GARCH(1,1)	-0.48	-114543	TRX	ARMA(1,1)-GARCH(1,1)	0.45%	-6657.17	0.43

	Best ARMA-GARCH-X vs ARMA(1,1-)GARCH(1,1)	0.76	-166875		Best ARMA-GARCH	0.21%	-6517.1	12.33
	Best ARMA-GARCH-X vs Best ARMA-GARCH	0.91	-141422		Best ARMA-GARCH-X	3.41%	-6465.04	4.23
BCH	Best ARMA-GARCH vs ARMA(1,1-)GARCH(1,1)	1.2	-189229	BCH	ARMA(1,1)-GARCH(1,1)	0.51%	-7365.52	0.46
	Best ARMA-GARCH-X vs ARMA(1,1-)GARCH(1,1)	1.19	-209401		Best ARMA-GARCH	1.39%	-7037.53	2.87
	Best ARMA-GARCH-X vs Best ARMA-GARCH	0.29	-191767		Best ARMA-GARCH-X	1.44%	-6995.86	3.24
LINK	Best ARMA-GARCH vs ARMA(1,1-)GARCH(1,1)	-1.14	-219197	LINK	ARMA(1,1)-GARCH(1,1)	0.02%	-7633.91	0.69
	Best ARMA-GARCH-X vs ARMA(1,1-)GARCH(1,1)	-0.85	-209378		Best ARMA-GARCH	0.02%	-7592.4	1.5
	Best ARMA-GARCH-X vs Best ARMA-GARCH	-0.38	-263862		Best ARMA-GARCH-X	0.74%	-7550.34	30.72
LTC	Best ARMA-GARCH vs ARMA(1,1-)GARCH(1,1)	-0.67	-198841	LTC	ARMA(1,1)-GARCH(1,1)	0.22%	-7131.78	0.41
	Best ARMA-GARCH-X vs ARMA(1,1-)GARCH(1,1)	-0.28	-231421		Best ARMA-GARCH	1.36%	-6971.94	28.84
	Best ARMA-GARCH-X vs Best ARMA-GARCH	0.52	-206088		Best ARMA-GARCH-X	0.15%	-6942.22	2.75

As for the Panel B, provides the Mincer-Zarnowitz R2, Log-Likelihood and the models estimating time, which evaluate the calibration quality of the estimated models and its cost-benefits analysis. The results indicate a consistent improvement in Log-Likelihood values when moving from the benchmark ARMA(1,1)-GARCH(1,1) to the optimized specifications. Regarding the Loss Function, the values show a systematic reduction as the model complexity increases, which numerically supports the transition toward more specialized architectures.

However, the Mincer-Zarnowitz R2 values remain generally low across all specifications, characteristic of the inherent noise and stochastic volatility in cryptocurrency markets. Nevertheless, specific assets show notable improvements in explanatory power: TRX exhibits a significant increase from 0.45% in the benchmark to 3.41% in the ARMA-GARCH-X specification, while BCH shows a steady progression from 0.51% to 1.44%.

These improvements suggests that for assets with complex volatility regimes, the integration of exogenous liquidity and sentiment indicators (FGI and TO) enhances the model's ability to map the return distribution. This supports the argument of López-Cabarcos et al. (2021) and Wei et al. (2023), confirming that cryptocurrency volatility is not merely an endogenous process but is significantly influenced by 'noise trader' effects and behavioral sentiment shocks.

Finally, regarding estimation time, there was a general increase in computational requirements as model complexity rose. While assets such as BTC and BNB demonstrate improved Log-Likelihood values and enhanced stability in risk metrics, it is important to note that these gains are often marginal.

Consequently, the significant computational overhead associated with ARMA-GARCH-X specifications may only be justified in high-stakes professional risk management frameworks, rather than as a universal requirement. Ultimately, while these calibration results align with backtesting performance, which suggest that the optimized models achieve a more precise statistical representation of the market's risk environment, the incremental improvements in predictive accuracy remain limited, confirming that the added complexity does not represent a systemic breakthrough in point-forecasting as in Silva e Maciel (2023).

5 Final considerations

Investor interest in the cryptocurrency market has grown exponentially in recent years, driven by the potential for high returns and the consolidation of digital assets as a new financial frontier. However, the well-documented high volatility of these assets demands robust and innovative risk management approaches for their inclusion in investment portfolios. In line with this concern, the present study focused on analyzing the capability of ARMA-GARCH family models to estimate market risk for the top 10 cryptocurrencies by market capitalization over a period of nearly seven years.

To achieve this, an extensive computational experiment was conducted, estimating over 4,032 model specifications per cryptocurrency. This approach combined 7 GARCH-family variants, 9 error distributions, and 64 ARMA-GARCH order combinations, aiming to identify, for each asset, the parameterization that minimized the Mean Absolute Error in out-of-sample volatility forecasts. Using these optimized specifications, dynamic estimates of Value-at-Risk and Expected Shortfall were generated and subjected to rigorous backtesting procedures, including unconditional and conditional coverage tests, which validated the statistical adequacy of the risk measures produced.

A key contribution of this study was the evaluation of whether incorporating exogenous variables could enhance predictive accuracy and risk estimation. The analysis was revisited by including the Fear and Greed Index, as a proxy for investor sentiment, and Bitcoin's Turnover, as a proxy for market liquidity and speculative intensity, within a GARCH-X framework. The comparative analysis revealed that the inclusion of these variables did not produce a uniform effect across all cryptocurrencies. While some assets exhibited marginal improvements in forecast errors and better calibration of downside risk, others showed slight deteriorations in predictive performance.

Overall, the gains obtained with the inclusion of exogenous variables were modest and heterogeneous, suggesting that these variables act more as diagnostic tools or fine-tuning mechanisms than as universal performance boosters. This finding indicates that the predictive gains from incorporating external information are conditional on the specific volatility profile of each asset and do not represent a systemic breakthrough in forecasting accuracy.

An equally important finding concerns the marginal gain of moving from a simple benchmark specification to more complex optimized models. When comparing the predictive performance of the best ARMA-GARCH and best ARMA-GARCH-X models against a standard ARMA(1,1)-GARCH(1,1) benchmark, the improvements in forecast accuracy, while present, were generally marginal.

In the other hands, the optimized models, particularly those selected through the exhaustive search procedure, demonstrated superior calibration of VaR and ES, with failure rates more closely aligned with theoretical expectations and better performance in conditional coverage tests. This distinction is crucial: while the gains in forecasting volatility may be incremental, the benefits for tail risk estimation are far more pronounced, justifying the additional computational effort for risk management applications.

Limitations of this study are related to three methodological choices. First, cryptocurrency returns were calculated based on USD quotes. Additionally, only two exogenous variables were considered, while other relevant effects could be included in ARMA-GARCH modeling adjustments. Finally, this study focused on cryptocurrencies with the highest market capitalization, selecting 10 out of over 9,000 coins.

Therefore, for future research on this topic, it is suggested to first replicate this study considering the exchange rate risk effect of the dollar to expand the study's conclusions. Furthermore, it is also suggested to consider quotations in terms of BTC rather than USD, as this change in scale may alter the volatility of time series data.

Secondly, expanding the scope of exogenous variables used aims to further enhance the predictive capability of models and the efficiency of estimated VaR and ES. Finally, it is suggested to study risk management comparatively between cryptocurrencies with higher and lower market capitalization, aiming to identify similarities and differences in the conclusions obtained in this study when analyzing smaller cryptocurrencies.

References

Abreu, D. P. A. de, Coaguila, R. A. I., & Camargos, M. A. de. (2022). Evolution of the degree of efficiency of the cryptocurrency market from 2014 to 2020: An analysis based on its fractal components. Revista de Administração Da UFSM, 15(2), Artigo 2.

Acerbi, C., & Tasche, D. (2002). On the coherence of expected shortfall. Journal of Banking & Finance, 26(7), 1487–1503. https://doi.org/10.1016/S0378-4266(02)00283-2

Acereda, B., Leon, A., & Mora, J. (2020). Estimating the expected shortfall of cryptocurrencies: An evaluation based on backtesting. Finance Research Letters, 33, 101181. https://doi.org/10.1016/j.frl.2019.04.037

Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999). Coherent Measures of Risk. Mathematical Finance, 9(3), 203–228. https://doi.org/10.1111/1467-9965.00068

Baur, D. G., & Dimpfl, T. (2018). Asymmetric volatility in cryptocurrencies. Economics Letters, 173, 148–151. https://doi.org/10.1016/j.econlet.2018.10.008

Baur, D. G., Hong, K., & Lee, A. D. (2018). Bitcoin: Medium of exchange or speculative assets? Journal of International Financial Markets, Institutions and Money, 54, 177–189. https://doi.org/10.1016/j.intfin.2017.12.004

Cai, C. W., Xue, R., & Zhou, B. (2023). Cryptocurrency puzzles: A comprehensive review and re-introduction. Journal of Accounting Literature, 46(1), 26–50. https://doi.org/10.1108/JAL-02-2023-0023

Charfeddine, L., Benlagha, N., & Maouchi, Y. (2020). Investigating the dynamic relationship between cryptocurrencies and conventional assets: Implications for financial investors. Economic Modelling, 85, 198–217. https://doi.org/10.1016/j.econmod.2019.05.016

Charles, A., & Darné, O. (2019). The accuracy of asymmetric GARCH model estimation. International Economics, 157, 179–202. https://doi.org/10.1016/j.inteco.2018.11.001

Conlon, T., & McGee, R. (2020). Safe haven or risky hazard? Bitcoin during the Covid-19 bear market. Finance Research Letters, 35, 101607. https://doi.org/10.1016/j.frl.2020.101607
Engle, R. (2002). Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. Journal of Business & Economic Statistics, 20(3), 339–350. https://doi.org/10.1198/073500102288618487
Fung, K., Jeong, J., & Pereira, J. (2022). More to cryptos than bitcoin: A GARCH modelling of heterogeneous cryptocurrencies. Finance Research Letters, 47, 102544. https://doi.org/10.1016/j.frl.2021.102544
Gandal, N., & Halaburda, H. (2014). Competition in the Cryptocurrency Market (SSRN Scholarly Paper 2501640). https://papers.ssrn.com/abstract=2501640
Hasan, F., Al-Okaily, M., Choudhury, T., & Kayani, U. (2024). A comparative analysis between FinTech and traditional stock markets: Using Russia and Ukraine war data. Electronic Commerce Research, 24(1), 629–654. https://doi.org/10.1007/s10660-023-09734-0
Huang, Y., Wang, H., Chen, Z., Feng, C., Zhu, K., Yang, X., & Yang, W. (2024). Evaluating Cryptocurrency Market Risk on the Blockchain: An Empirical Study Using the ARMA-GARCH-VaR Model. IEEE Open Journal of the Computer Society, 5, 83–94. https://doi.org/10.1109/OJCS.2024.3370603
Kristoufek, L., & Vosvrda, M. (2019). Cryptocurrencies market efficiency ranking: Not so straightforward. Physica A: Statistical Mechanics and Its Applications, 531, 120853. https://doi.org/10.1016/j.physa.2019.04.089
Lánský, J. (2017). Bitcoin System. Acta Informatica Pragensia, 6(1), 20–31. https://doi.org/10.18267/j.aip.97
Liu, Y., Tsyvinski, A., & Wu, X. (2022). Common risk factors in cryptocurrency. The Journal of Finance, 77(2), 1133-1177. https://doi.org/10.1111/jofi.13119
López-Cabarcos, M. Á., Pérez-Pico, A. M., Piñeiro-Chousa, J., & Šević, A. (2021). Bitcoin volatility, stock market and investor sentiment. Are they connected?. Finance Research Letters, 38, 101399. https://doi.org/10.1016/j.frl.2019.101399
Malek, J., Nguyen, D. K., Sensoy, A., & Tran, Q. V. (2023). Modeling dynamic VaR and CVaR of cryptocurrency returns with alpha-stable innovations. Finance Research Letters, 55, 103817. https://doi.org/10.1016/j.frl.2023.103817
Markowitz, H. (1952). Portfolio Selection. The Journal of Finance, 7(1), 77–91. https://doi.org/10.1111/j.1540-6261.1952.tb01525.x
Nakamoto, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System. https://bitcoin.org/bitcoin.pdf
Ngunyi, A., Mundia, S., & Omari, C. (2019). Modelling Volatility Dynamics of Cryptocurrencies Using GARCH Models. http://repository.dkut.ac.ke:8080/xmlui/handle/123456789/978
Orujov, S., Elvira, V., Poterie, A., Rajabov, F., & Septier, F. (2025). VS-LTGARCHX: A Flexible Variable Selection in Log-TGARCHX Models. Journal of Time Series Econometrics, 17(1), 1-34. https://doi.org/10.1515/jtse-2023-0035

Silva, L., & Maciel, L. (2025). Cryptocurrency price returns volatility modeling and forecasting with GARCH models. RAUSP Management Journal, 60(1), 220-234. https://doi.org/10.1108/RAUSP-04-2023-0056

Sözen, Ç. (2025). Volatility dynamics of cryptocurrencies: a comparative analysis using GARCH-family models. Future Business Journal, 11(1), 166. https://doi.org/10.1186/s43093-025-00568-w

Suryono, R. R., Budi, I., & Purwandari, B. (2020). Challenges and Trends of Financial Technology (Fintech): A Systematic Literature Review. Information, 11(12), Artigo 12. https://doi.org/10.3390/info11120590

Trimborn, S., & Härdle, W. K. (2018). CRIX an Index for cryptocurrencies. Journal of Empirical Finance, 49, 107–122. https://doi.org/10.1016/j.jempfin.2018.08.004

Troster, V., Tiwari, A. K., Shahbaz, M., & Macedo, D. N. (2019). Bitcoin returns and risk: A general GARCH and GAS analysis. Finance Research Letters, 30, 187–193. https://doi.org/10.1016/j.frl.2018.09.014

Urquhart, A. (2016). The inefficiency of Bitcoin. Economics Letters, 148, 80–82. https://doi.org/10.1016/j.econlet.2016.09.019

Wang, G.-J., Ma, X., & Wu, H. (2020). Are stablecoins truly diversifiers, hedges, or safe havens against traditional cryptocurrencies as their name suggests? Research in International Business and Finance, 54, 101225. https://doi.org/10.1016/j.ribaf.2020.101225

Wei, M., Sermpinis, G., & Stasinakis, C. (2023). Forecasting and trading Bitcoin with machine learning techniques and a hybrid volatility/sentiment leverage. Journal of Forecasting, 42(4), 852-871. https://doi.org/10.1002/for.2922