

UNIVERSIDADE ABERTA



ANÁLISE DE RELAÇÕES DINÂMICAS ENTRE *TOKENS* EM REDES BLOCKCHAIN COM FERRAMENTAS DE SÉRIES TEMPORAIS

Ricardo João Lourenço de Abreu

Dissertação de Mestrado em Estatística, Matemática e Computação

Dissertação orientada pela
Professora Doutora Maria do Rosário Olaia Duarte Ramos

Dezembro de 2025

UNIVERSIDADE ABERTA



UNIVERSIDADE
AbERTA
www.uab.pt

ANÁLISE DE RELAÇÕES DINÂMICAS ENTRE *TOKENS* EM REDES BLOCKCHAIN COM FERRAMENTAS DE SÉRIES TEMPORAIS

Ricardo João Lourenço de Abreu

Dissertação de Mestrado em Estatística, Matemática e Computação

Dissertação orientada pela
Professora Doutora Maria do Rosário Olaia Duarte Ramos

Dezembro de 2025

Declaração de Direitos de Autor e Licença de Utilização

O presente trabalho está licenciado nos termos da licença **Creative Commons Atribuição–Não Comercial–Sem Derivações 4.0 Internacional (CC BY-NC-ND 4.0)**.

Esta licença permite a terceiros copiar, distribuir e partilhar o trabalho, desde que seja feita a devida atribuição ao autor, não seja utilizado para fins comerciais e não sejam realizadas alterações ou obras derivadas, sem autorização expressa do autor.

O texto integral da licença encontra-se disponível em:

<https://creativecommons.org/licenses/by-nc-nd/4.0/>



Agradecimentos

Este trabalho é fruto de uma longa viagem académica de aprendizagem, pautada pelo apoio e rigor da Professora Doutora Maria do Rosário Olaia Duarte Ramos, que orientou os momentos mais difíceis da elaboração desta dissertação.

Agradece-se também a todos os docentes do Mestrado em Estatística, Matemática e Computação, em particular ao Professor Doutor Mário Júlio Pereira Bessa da Costa, que na sua Unidade Curricular encontrei a inspiração e vontade para realizar esta dissertação.

Manifesta-se uma profunda gratidão à família e entes queridos, em particular à mãe do autor e à sua mulher Raquel.



DECLARAÇÃO DE INTEGRIDADE

STATEMENT OF INTEGRITY

Declaro ter atuado com integridade na elaboração da presente dissertação/tese. Confirmo que em todo o trabalho conducente à sua elaboração não recorri à prática de plágio ou a qualquer outra forma de falsificação de resultados.

Mais declaro que tomei conhecimento integral do Regulamento Disciplinar da Universidade Aberta, publicado no Diário da República, 2.^a série, n.º 215, de 6 de novembro de 2013.

I hereby declare having conducted my thesis with integrity. I confirm that I have not used plagiarism or any form of falsification of results in the process of the thesis elaboration.

I further declare that I have fully acknowledged Disciplinary Regulations of the Universidade Aberta (regulation published in the official journal Diário da República, 2.^a série, N.º 215, de 6 de novembro de 2013).

Universidade Aberta, 28 de dezembro de 2025

Nome completo/Full name: Assinatura/Signature:


manuscrita ou digital / handwritten or digital

ANÁLISE DE RELAÇÕES DINÂMICAS ENTRE *TOKENS* EM REDES BLOCKCHAIN COM FERRAMENTAS DE SÉRIES TEMPORAIS

Esta dissertação explora, sob uma perspectiva quantitativa, a estrutura das dependências dinâmicas entre *tokens* de Camada 1 (L1) e Camada 2 (L2) no ecossistema *blockchain* do Ethereum, colocando a questão central de saber se a hierarquia tecnológica entre camadas se reflete em relações econométricas observáveis. Em particular, analisa-se a interação entre os preços de Camada 1, ETH, e de Camada 2, MATIC e OP através de um enquadramento unificado de séries temporais multivariadas, combinando métodos clássicos de econometria financeira com abordagens modernas de aprendizagem profunda.

Do ponto de vista metodológico, o estudo articula testes de raiz unitária e cointegração (Johansen), modelos VAR/VECM para a dinâmica de curto e longo prazo, testes de causalidade de Granger, e modelos de volatilidade condicional GARCH e DCC-GARCH para caracterização do risco sistêmico e das correlações dinâmicas. Em paralelo, propõe-se um modelo híbrido ARIMA–LSTM multivariado com mecanismo de *self-attention*, concebido para decompor a dinâmica observada em componentes lineares e não lineares, maximizando a capacidade preditiva num contexto de elevada não estacionariedade e heterocedasticidade condicional.

Os resultados empíricos indicam ausência de cointegração robusta entre *tokens* L1 e L2, sugerindo que a dependência tecnológica não implica, por si só, um equilíbrio económico de longo prazo. Contudo, identificam-se relações causais significativas e efeitos de transmissão assimétricos no curto prazo, bem como elevados níveis de correlação condicional variável no tempo, evidenciando contágio de volatilidade entre camadas. A análise comparativa de desempenho confirma que o modelo híbrido ARIMA–LSTM supera os modelos lineares tradicionais, demonstrando a relevância de arquiteturas não lineares para capturar dependências complexas.

Este trabalho contribui para uma formalização matemática mais profunda da economia de *tokens* em arquiteturas *blockchain* modulares, oferecendo implicações relevantes para previsão, gestão de risco e modelação económica de sistemas descentralizados.

Palavras-chave: Séries temporais multivariadas; Blockchain; *Tokens* L1–L2; Volatilidade condicional; Modelos híbridos ARIMA–LSTM

ANALYSIS OF DYNAMIC RELATIONSHIPS BETWEEN TOKENS IN BLOCKCHAIN NETWORKS USING TIME SERIES TOOLS

This dissertation explores, from a rigorous quantitative perspective, the dynamic dependence structure between Layer 1 (L1) and Layer 2 (L2) tokens within the Ethereum ecosystem, addressing the central question of whether technological hierarchy translates into measurable econometric relationships. Specifically, it examines the interactions among Layer 1, ETH, and layer 2, MATIC and OP prices, within a unified multivariate time series framework that integrates classical financial econometrics with modern deep learning techniques.

Methodologically, the study combines unit root and cointegration testing (Johansen), VAR/VECM models for short and long-run dynamics, Granger causality analysis, and conditional volatility modeling using GARCH and DCC-GARCH specifications to capture time-varying risk transmission and dynamic correlations. In parallel, it proposes a multivariate hybrid ARIMA–LSTM model augmented with a self-attention mechanism, designed to decompose observed price dynamics into linear and nonlinear components, thereby enhancing predictive performance under conditions of non-stationarity and conditional heteroskedasticity.

Empirical results reveal no robust cointegration between L1 and L2 tokens, indicating that technological dependence does not imply a stable long-run economic equilibrium. Nevertheless, significant short-run causal relationships and asymmetric shock transmission patterns are identified, predominantly from ETH to Layer 2 tokens. Moreover, the DCC-GARCH analysis uncovers high and persistent time-varying conditional correlations, providing strong evidence of volatility spillovers across layers. Forecasting comparisons demonstrate that the proposed hybrid ARIMA–LSTM model consistently outperforms traditional linear benchmarks, highlighting the importance of nonlinear architectures for capturing complex dependencies in crypto-asset markets.

Overall, this research contributes to a deeper mathematical formalization of token economics in modular blockchain architectures, offering relevant implications for forecasting, risk management, and the economic modeling of decentralized systems

Keywords: Multivariate time series; Blockchain; Layer 1 and Layer 2 tokens; Conditional volatility; Hybrid ARIMA–LSTM models

ÍNDICE

1. INTRODUÇÃO.....	1
2. REVISÃO DA LITERATURA.....	4
2.1. Tecnologia de <i>Blockchain</i> e economia de <i>tokens</i>.....	4
2.1.1. O surgimento da economia de <i>tokens</i>	4
2.1.2. Dinâmica de Token de Camada 1 e Camada 2.....	4
2.1.3. Modelos económicos e mecanismos de incentivo.....	5
2.1.4. Desafios na economia de Tokens.....	5
2.2. Fundamentos matemáticos e estatísticos dos modelos de séries temporais.....	6
2.2.1. Modelos ARIMA (Autoregressive Integrated Moving Average).....	7
2.2.2. Modelos multivariados, cointegração e VECM.....	7
2.2.3. Modelos de volatilidade condicional: GARCH e DCC-GARCH.....	9
2.2.4. Modelos de aprendizagem profunda: Redes LSTM.....	10
2.2.5. Modelos híbridos ARIMA–LSTM.....	12
2.2.6. Comparação entre modelos lineares e não lineares na análise de séries temporais.....	12
2.3. Análise de séries temporais na investigação sobre <i>blockchain</i> e criptomoedas	14
2.3.1. Modelos tradicionais de séries temporais.....	14
2.3.2. Abordagens de aprendizagem automática e aprendizagem profunda.....	15
2.3.3. Análise de Cointegração e Causalidade.....	15
2.4. Enquadramento teórico e questões de investigação.....	16
3. METODOLOGIA.....	17
3.1. Dados, seleção dos tokens e período de análise.....	17
3.2. Transformação das séries e análise preliminar.....	18
3.3. Análise de cointegração e causalidade.....	23
3.4. Modelos para as relações dinâmicas de preços (VAR/VECM e ARIMA–LSTM).....	24
3.4.1. Modelos VAR e VECM.....	24
3.4.2. Modelo Híbrido ARIMA–LSTM com Mecanismo de Atenção.....	25
3.5. Modelos de volatilidade condicional e correlações dinâmicas (GARCH e DCC-GARCH).....	27
3.6. Estratégia de Avaliação e Métricas de Desempenho.....	27
3.7. Implementação Computacional.....	29
4. ANÁLISE DOS RESULTADOS.....	31
4.1. Estrutura de Longo prazo.....	31
4.1.1. Estacionariedade e Teste de Cointegração.....	31
4.2. Dinâmica de Curto prazo.....	32
4.2.1. Seleção e estimação dos modelos VAR.....	32
4.2.2. Causalidade de Granger.....	35
4.3. Dinâmica Estrutural (IRF e FEVD).....	35
4.4. Análise Preditiva.....	40
4.4.1. Componente Linear (ARIMA).....	40
4.4.2. Componente não-linear: Modelos híbridos ARIMA-LSTM.....	41
4.4.3. Análise comparativa do desempenho preditivo.....	42

4.4.4.	Resultados Empíricos	42
4.5.	<i>Volatilidade e Contágio</i>	45
4.5.1.	Testes ARCH-LM e Fundamentação da Modelação GARCH	45
4.5.2.	Dinâmica de Volatilidade ETH-MATIC	46
4.5.3.	Dinâmica de Volatilidade ETH-OP	48
5.	<i>DISCUSSÃO DOS RESULTADOS</i>	50
6.	<i>CONCLUSÃO</i>	52
7.	<i>REFERÊNCIAS</i>	55
8.	<i>ANEXOS</i>	60
	Anexo A - Procedimento do Modelo Híbrido ARIMA–LSTM com <i>Self-Attention</i>	60
	Anexo B – Prompts utilizadas no Google® Gravity	62
	Anexo C – Resultados dos Modelos ARIMA	64
	Anexo D – Resultados dos Testes Augmented Dickey–Fuller	66

Tabela 2-1 Comparação entre modelos lineares e modelos não lineares de previsão de séries temporais.....	13
Tabela 3-1 Estatísticas descritivas dos retornos logarítmicos diários de ETH, MATIC e OP	20
Tabela 3-2 Matriz de correlação dos retornos logarítmicos diários de ETH, MATIC e OP ...	21
Tabela 4-1 Critérios de informação para seleção da ordem do VAR	32
Tabela 4-2 Testes ARCH-LM aplicados aos resíduos dos modelos VAR(1).....	45
Tabela 4-3 Parâmetros do Modelo DCC-GARCH (ETH-MATIC).....	46
Tabela 4-4 Parâmetros DCC	46
Tabela 4-5 Parâmetros do Modelo DCC-GARCH (ETH-OP)	48
Tabela 4-6 Parâmetros DCC	49
Tabela C-1 Comparação de modelos ARIMA para ETH	64
Tabela C-2 Comparação de modelos ARIMA para MATIC	64
Tabela C-3 Comparação de modelos ARIMA para OP.....	65

ÍNDICE FIGURAS

Figura 2-1 Exemplo de rede neural.....	11
Figura 3-1 Log-preços e retornos de ETH, MATIC e OP diários	19
Figura 3-2 Box-plot dos retornos logarítmicos diários dos tokens	21
Figura 3-3 Dispersão de retornos ETH vs MATIC e ETH vs OP	22
Figura 3-4 ACF e PACF dos retornos de ETH, MATIC e OP. Funções de autocorrelação dos retornos logarítmicos	23
Figura 3-5 Procedimento do Modelo Híbrido ARIMA-LSTM	27
Figura 3-6 Fluxo da Metodologia	29
Figura 4-1 - Funções de Impulso-Resposta (IRF) para o sistema ETH-MATIC.....	36
Figura 4-2 Decomposição da Variância (FEVD) para o sistema ETH-MATIC.....	37
Figura 4-3 Análise do Par ETH-MATIC	38
Figura 4-4 Análise do Par ETH-OP	39
Figura 4-5 Desempenho preditivo out-of-sample dos modelos para o sistema ETH-MATIC	43
Figura 4-6 Desempenho preditivo out-of-sample dos modelos para o sistema ETH-OP.....	43
Figura 4-7 Volatilidades Condicionais Individuais do ETH e MATIC	47
Figura 4-8 Volatilidades Condicionais Individuais do ETH e OP.....	47
Figura 4-9 Correlação Condicional Dinâmica (DCC) entre ETH e MATIC.....	48
Figura 4-10 Correlação Condicional Dinâmica (DCC) entre ETH e OP.....	49

GLOSSÁRIO

API – Application Programming Interface

ACF – Função de Autocorrelação, que mede a correlação de uma série temporal com os seus próprios desfasamentos.

ADF – Augmented Dickey–Fuller, teste de raiz unitária usado para avaliar estacionariedade das séries.

ADAM – **Adaptive Moment Estimation**, algoritmo de otimização em deep learning

AIC – Akaike Information Criterion, critério de informação para seleção de modelos.

ANN – Artificial Neural Network, classe geral de modelos de redes neuronais artificiais.

ARCH – Autoregressive Conditional Heteroskedasticity, modelo de volatilidade com variância condicional dependente de erros passados.

ARCH-LM – Lagrange Multiplier Test para heterocedasticidade ARCH

AR – Autoregressive, componente autoregressiva de um modelo de séries temporais em que o valor atual depende de valores passados.

ARIMA – Autoregressive Integrated Moving Average, modelo linear de séries temporais que combina componentes AR, diferenciação (I) e MA.

ARMA – Autoregressive Moving Average, modelo de séries temporais com componentes AR e MA, sem diferenciação.

BIC – Bayesian Information Criterion, critério de informação bayesiano para seleção de modelos.

BTC – Bitcoin, primeira criptomoeda baseada em blockchain.

DApps – Decentralized Applications, aplicações que correm sobre uma rede blockchain, tipicamente usando contratos inteligentes.

DCC – Dynamic Conditional Correlation, especificação para modelar correlações condicionais que variam no tempo.

DCC-GARCH – Dynamic Conditional Correlation Generalized Autoregressive Conditional Heteroskedasticity, modelo multivariado que combina GARCH univariado com correlações dinâmicas.

DeFi – Decentralized Finance, ecossistema de serviços financeiros construídos sobre blockchains públicos.

DLT – Distributed Ledger Technology, tecnologia de registo distribuído que suporta blockchains.

ECM – Error Correction Model, modelo de correção de erros usado quando há cointegração, captando ajuste de curto prazo ao equilíbrio de longo prazo.

ERC-20 – Padrão de token fungível na rede Ethereum para ativos compatíveis com contratos inteligentes.

ETH – Ether, criptomoeda nativa da rede Ethereum.

FEVD – Forecast Error Variance Decomposition, decomposição da variância do erro de previsão

GARCH – Generalized Autoregressive Conditional Heteroskedasticity, modelo de volatilidade condicional que generaliza o ARCH.

H – Teste de heterocedasticidade, normalmente associado ao teste de Breusch–Pagan ou White dependendo do contexto

ICO – Initial Coin Offering, mecanismo de financiamento em que projetos emitem tokens a investidores.

IRF – Impulse Response Function, função de resposta ao impulso num modelo VAR

JB – Jarque–Bera – teste de normalidade baseado em assimetria e curtose.

KPSS – Kwiatkowski–Phillips–Schmidt–Shin, teste de estacionariedade alternativo ao ADF, usado para avaliar a hipótese nula de estacionariedade.

L – Operador de defasagem (lag operator), usado para expressar modelos ARIMA

L1 – Layer 1, camada base de blockchain (por exemplo, Ethereum) responsável por consenso e segurança.

L2 – Layer 2, soluções de escalabilidade construídas sobre uma L1 que processam transações fora da cadeia principal.

LSTM – Long Short-Term Memory, tipo de rede neuronal recorrente capaz de capturar dependências de longo prazo em séries temporais.

MA – Moving Average, componente de média móvel em modelos de séries temporais baseada em erros passados.

MAE – Mean Absolute Error, métrica de desempenho preditivo que calcula o erro absoluto médio.

MATIC – Token nativo da rede Polygon, solução de Camada 2 construída sobre Ethereum.

MQO – Mínimos Quadrados Ordinários, método de estimação por regressão linear (Ordinary Least Squares).

MSE – Mean Squared Error, erro quadrático médio, usado como função de perda em modelos de previsão.

NFT – Non-Fungible Token, token não fungível que representa ativos digitais únicos.

OP – Optimism, token associado à solução de escalabilidade Optimism na Camada 2 do Ethereum.

PACF – Função de Autocorrelação Parcial, mede a correlação entre a série e os seus desfasamentos eliminando efeitos intermediários.

PoS – Proof of Stake, mecanismo de consenso em que validadores apostam (stake) tokens para participar na validação de blocos.

R² – Coeficiente de determinação, mede a proporção da variância explicada pelo modelo.

RMSE – Root Mean Squared Error, raiz do erro quadrático médio, métrica de avaliação de previsões.

RNN – Recurrent Neural Network, arquitetura de rede neuronal que processa sequências, com ligações recorrentes no tempo.

STO – Security Token Offering, emissão de tokens classificados como valores mobiliários sob enquadramento regulatório específico.

SVR – Support Vector Regression, técnica de regressão baseada em máquinas de vetores de suporte, mencionada em contexto de modelos híbridos ARIMA-SVR na revisão da literatura.

VAR – Vector Autoregression, modelo multivariado em que cada variável depende dos seus próprios desfasamentos e dos desfasamentos das restantes.

VECM – Vector Error Correction Model, versão do VAR que inclui termos de correção de erro quando as séries são cointegradas.

σ^2 – Variância condicional estimada pelo modelo

Q, K, V – Vetores de consulta (query), chave (key) e valor (value) usados nos mecanismos de atenção em redes neurais

1. INTRODUÇÃO

A tecnologia *blockchain*, desde o seu início com a introdução da Bitcoin por Nakamoto (2008), emergiu como uma força transformadora na economia digital. Ao permitir métodos descentralizados, transparentes e seguros para gerir transações e ativos digitais, a cadeia de blocos abriu caminho a aplicações inovadoras em vários setores, incluindo finanças, cadeia de abastecimento e governança (Zheng et al., 2018). No centro desta revolução tecnológica está a *economia de tokens*, em que os *tokens* digitais representam valor e facilitam uma vasta gama de funções de rede, desde taxas de transação a governança e incentivo (Catalini & Gans, 2018; Y. Chen, 2018). Os *tokens*, quer sejam nativos de uma cadeia de blocos (Camada 1 ou L1) ou construídos sobre ela (Camada 2 ou L2), desempenham um papel fundamental no funcionamento e crescimento dos ecossistemas de cadeias de blocos.

Apesar do seu potencial, a tecnologia *blockchain* enfrenta desafios significativos de escalabilidade devido à sua natureza descentralizada. A crescente procura de transações mais rápidas e mais baratas levou ao desenvolvimento de estruturas hierárquicas, distinguindo entre soluções da Camada 1 e da Camada 2 (Sguanci et al., 2021). A Camada 1 refere-se ao protocolo de base da cadeia de blocos, como o Bitcoin ou o Ethereum, que fornece os mecanismos fundamentais de segurança e consenso (Buterin, 2014). As soluções da Camada 1 visam melhorar o rendimento das transações diretamente na cadeia de blocos através de métodos como o aumento da dimensão dos blocos ou a implementação de técnicas de fragmentação (Song et al., 2024). Por outro lado, as soluções da Camada 2, construídas sobre as redes da Camada 1, aumentam a velocidade das transações e reduzem os custos, tirando partido da segurança da cadeia de blocos subjacente. Estas incluem tecnologias como *rollups*, canais de pagamento e *sidechains*, que processam transações fora da cadeia e as liquidam periodicamente na cadeia de blocos principal (Gudgeon et al., 2020; Poon & Dryja, 2015).

A interação entre os *tokens* da Camada 1 e da Camada 2 introduz uma dinâmica complexa que ainda não é totalmente compreendida. Os modelos financeiros tradicionais e as técnicas de análise de séries temporais são frequentemente insuficientes quando aplicados a dados de cadeias de blocos devido às suas características únicas, como a elevada volatilidade, a não estacionariedade e às interdependências complexas (Dyhrberg, 2016; Malladi & Dheeriyaa, 2021). Por exemplo, foi demonstrado que a volatilidade dos preços das criptomoedas exhibe padrões de agrupamento e variação no tempo, que podem ser modelados de forma eficaz usando modelos autorregressivos generalizados condicionalmente heteroscedásticos (GARCH) (Malladi & Dheeriyaa, 2021), contudo as relações hierárquicas entre *tokens* não foram abordadas pelos autores. Além disso, as relações entre os *tokens* da Camada 1 e da Camada 2 podem envolver equilíbrios de longo prazo, que podem ser explorados por meio da análise de cointegração (Nair, 2021). Os testes de causalidade, como o teste de causalidade de Granger, podem fornecer informações sobre as relações direcionais entre *tokens* em diferentes camadas (Malladi & Dheeriyaa, 2021). Por exemplo, compreender os efeitos volatilidade e a sua contribuição para a gestão de riscos das Finanças Descentralizadas (DeFi) e para a concepção de incentivos da Camada 2.

Apesar do crescimento acelerado dos ecossistemas de Camada 2 e da sua dependência operacional das redes de Camada 1, a literatura ainda trata estas relações de forma fragmentada. A maioria dos estudos analisa criptomoedas como ativos independentes ou compara grupos

amplios de *tokens*, mas raramente considera a dimensão tecnológica da hierarquia L1–L2 como um fator estruturante dos seus comportamentos de mercado. Esta ausência é particularmente evidente no ecossistema Ethereum, onde soluções como Polygon e Optimism desempenham hoje um papel económico significativo, mas continuam pouco estudadas no contexto da transmissão de preços, da volatilidade e da previsibilidade.

Apesar da crescente relevância económica das soluções de Camada 2 no ecossistema Ethereum, a literatura continua a tratar as relações entre *tokens* L1–L2 de forma fragmentada. Os estudos existentes analisam, em geral, volatilidade, causalidade ou previsibilidade de criptomoedas de forma isolada, mas raramente integram estas dimensões num enquadramento único que considere a hierarquia tecnológica subjacente. Assim, permanece por esclarecer se a ligação técnica entre Camada 1 e Camada 2, baseada em segurança partilhada, dependência operacional e fluxo de transações, se traduz também numa ligação económica mensurável. Este vazio conceptual define o gap de investigação central desta dissertação: determinar se os mecanismos tecnológicos que unem os *tokens* L1 e L2 se refletem em equilíbrios de longo prazo, em padrões direcionais de causalidade ou em efeitos de contágio na volatilidade, e de que forma esta estrutura influencia a dinâmica conjunta dos seus preços.

A motivação técnica para esta investigação assenta no facto de que a relação entre Camada 1 e Camada 2, embora bem definida do ponto de vista computacional, não implica necessariamente uma relação econométrica equivalente. *Tokens* de Camada 2 dependem estruturalmente do Ethereum para segurança, liquidação e disponibilidade de blockspace, mas essa dependência não garante, por si só, que os seus preços sigam padrões comuns ou exibam equilíbrios estáveis. Da mesma forma, o rápido desenvolvimento de soluções de escalabilidade, como por exemplo, os *rollups* e *sidechains*¹, altera o fluxo económico dentro do ecossistema, redistribuindo taxas, liquidez e procura de transações entre camadas. Assim, compreender se estas interdependências técnicas se refletem em interdependências estatísticas observáveis é essencial para avaliar o papel económico dos *tokens* da Camada 2, medir o grau em que influenciam o *token* base da rede e caracterizar a estrutura de risco e transmissão de choques num sistema cada vez mais modularizado.

Do ponto de vista aplicado, a relevância prática desta investigação é clara. Para investidores, desenvolvedores e gestores de risco, conhecer a direção dos fluxos de informação entre ETH, MATIC e OP permite antecipar movimentos de mercado, melhorar modelos de alocação de ativos e calibrar estratégias de mitigação de risco. De forma semelhante, reguladores e equipas de projeto podem beneficiar de uma visão mais rigorosa sobre como choques em diferentes níveis da arquitetura blockchain se propagam economicamente. Ao integrar métodos econométricos tradicionais com abordagens de aprendizagem profunda, este estudo fornece ainda um contributo prático para a melhoria da capacidade preditiva num mercado conhecido pela sua elevada volatilidade e complexidade estrutural.

¹ Rollups são soluções de escalabilidade de Camada 2 que processam transações fora da cadeia principal e enviam apenas dados consolidados para a blockchain de Camada 1, reduzindo custos e aumentando o rendimento. Sidechains são blockchains paralelas que operam independentemente da Camada 1, mas ligadas a ela através de pontes, permitindo maior flexibilidade e capacidade de processamento

Objetivos da investigação

Esta investigação tem como objetivo colmatar estas lacunas através da utilização de técnicas avançadas de análise de séries temporais para explorar as relações dinâmicas entre os *tokens* da Camada 1 e da Camada 2. Especificamente, a investigação terá os seguintes objetivos:

1. Caracterizar a estrutura das relações dinâmicas de longo prazo ao analisar a existência de equilíbrio de longo prazo (cointegração) e determinar a direção dos fluxos de informação (causalidade de Granger) entre o token de Camada 1 e os tokens de Camada 2.
2. Avaliar o desempenho preditivo comparativo ao determinar em que medida a utilização de modelos híbridos de Aprendizagem Profunda (*Deep Learning*), enriquecidos com mecanismos de atenção e variáveis exógenas, supera a capacidade de previsão dos modelos lineares tradicionais.
3. Analisar a dinâmica de risco e contágio ao modelar a estrutura de volatilidade condicional e as correlações dinâmicas ao longo do tempo, para compreender os efeitos de spillover entre ativos.

Em conjunto, estes objetivos permitem que esta dissertação ofereça três contributos específicos para a literatura e para a prática. Em primeiro lugar, fornece evidência empírica sistemática sobre a dinâmica económica entre tokens de Camada 1 e Camada 2, clarificando se a dependência tecnológica entre Ethereum, Polygon e Optimism se traduz em relações econométricas observáveis. Em segundo lugar, introduz uma arquitetura híbrida ARIMA–LSTM multivariada com mecanismo de atenção (*Deep Learning*), constituindo um avanço metodológico face aos modelos tradicionalmente aplicados ao estudo de criptoativos e permitindo captar dependências não lineares entre camadas. Por fim, ao integrar análise de causalidade, previsão e volatilidade num enquadramento unificado, o estudo contribui para uma compreensão mais profunda da estrutura de transmissão de preços e de risco no ecossistema Ethereum, oferecendo insights relevantes para investidores, programadores de protocolos, instituições financeiras e investigadores que analisam a evolução da economia de *tokens* (DeFi).

O restante desta dissertação está organizado de forma a acompanhar progressivamente os objetivos definidos. O Capítulo 2 apresenta a revisão da literatura relevante, abordando os fundamentos da tecnologia *blockchain*, a evolução das economias de tokens e os principais modelos estatísticos aplicados ao estudo de criptoativos. O Capítulo 3 descreve a metodologia adotada, detalhando o processo de recolha e preparação dos dados, a análise preliminar e a especificação dos modelos econométricos e de aprendizagem profunda utilizados. O Capítulo 4 expõe os resultados empíricos, incluindo a análise das relações dinâmicas entre os tokens, a avaliação preditiva e a caracterização da volatilidade. O Capítulo 5 discute as implicações teóricas e práticas dos resultados, bem como as limitações do estudo e possíveis direções para investigação futura. Finalmente, o Capítulo 6 apresenta a conclusão geral, sintetizando as principais contribuições desta investigação para a compreensão da dinâmica económica entre Camada 1 e Camada 2 no ecossistema Ethereum.

2. REVISÃO DA LITERATURA

2.1. Tecnologia de *Blockchain* e economia de *tokens*

A tecnologia *Blockchain*, introduzida por Nakamoto (2008) através do livro branco Bitcoin, revolucionou a economia digital ao permitir métodos descentralizados, transparentes e seguros para gerir transações e ativos digitais. Na sua essência, a *blockchain* funciona como uma tecnologia de livro-razão distribuído (DLT), em que as transações são registradas numa cadeia de blocos, protegidas através de técnicas criptográficas e validadas por uma rede de nós sem a necessidade de uma autoridade central (Zheng et al., 2018). Esta descentralização abriu caminho a aplicações inovadoras em vários setores, incluindo o financeiro, a cadeia de abastecimento e a governança, ao mesmo tempo que introduziu novos modelos económicos centrados em *tokens* digitais (Sguanci et al., 2021).

2.1.1. O surgimento da economia de *tokens*

A economia de *tokens*, um componente fundamental dos ecossistemas de *blockchain*, refere-se ao sistema de *tokens* digitais que representam valor e facilitam várias funções de rede. Os *tokens* podem servir a múltiplos propósitos, inclusive como meio de troca, reserva de valor ou meio de governança em redes descentralizadas (Malladi & Dheeriyaa, 2021). Catalini & Gans (2018) destacam que os *tokens* não são meramente ativos digitais, mas também desempenham um papel crítico no incentivo à participação na rede, alinhando os interesses das partes interessadas e permitindo novas formas de atividade económica. Por exemplo, na rede Ethereum, *tokens* como o Ether (ETH) são utilizados para pagar taxas de transação e executar contratos inteligentes, enquanto outros *tokens* construídos no Ethereum (por exemplo, *tokens* ERC-20) permitem o funcionamento de aplicações descentralizadas (DApps) (Freni et al., 2022).

A economia dos *tokens* evoluiu significativamente desde a criação da Bitcoin, com o surgimento das *Initial Coin Offerings* (ICOs), das *Security Token Offerings* (STOs) e das *Non-Fungible Tokens* (NFTs) (Au & Power, 2018; Mohamed, 2022). Estes mecanismos de tokenização permitiram às empresas em fase de arranque e aos projetos angariar capital, criar ativos digitais e estabelecer novos mercados. No entanto, a economia dos *tokens* também enfrenta desafios, como a incerteza regulamentar, a volatilidade do mercado e a necessidade de modelos económicos sólidos para garantir a sustentabilidade a longo prazo (Malladi & Dheeriyaa, 2021; Z. Zhang, 2019).

2.1.2. Dinâmica de Token de Camada 1 e Camada 2

As redes de cadeias de blocos são frequentemente categorizadas em protocolos da Camada 1 (camada de base ou *Layer 1*) e da Camada 2 (soluções de escalonamento ou *Layer 2*) (Hafid et al., 2020; Song et al., 2024). A Camada 1 refere-se aos protocolos de cadeia de blocos fundamentais, como o Bitcoin e o Ethereum, que fornecem a segurança subjacente e os mecanismos de consenso (Song et al., 2024). As soluções da Camada 2, construídas sobre as redes da Camada 1, visam melhorar a escalabilidade, reduzir os custos de transação e aumentar o rendimento através do processamento de transações fora da cadeia e da sua liquidação periódica na cadeia de blocos principal (Sguanci et al., 2021). Exemplos de soluções da

Camada 2 incluem rollups, canais de pagamento e *sidechains*, que ganharam força nos últimos anos devido à sua capacidade de lidar com as limitações de escalabilidade das redes da Camada 1 (Katsika et al., 2024; Sguanci et al., 2021).

A interação entre os *tokens* da Camada 1 e da Camada 2 introduz dinâmicas complexas que ainda não são totalmente compreendidas. Por exemplo, os *tokens* da Camada 1 (e.g. ETH) servem frequentemente como a principal reserva de valor e mecanismo de segurança, enquanto os *tokens* da Camada 2 (e.g. *token* OP do Optimism) podem ser utilizados para governança ou como incentivos à participação na rede (Han et al., 2025; Sguanci et al., 2021). A relação entre estes *tokens* pode influenciar o comportamento do mercado, a volatilidade e o crescimento da rede, tornando-a uma área de estudo fundamental para a economia da cadeia de blocos (Z. Zhang, 2019).

2.1.3. Modelos económicos e mecanismos de incentivo

A concepção de economias de *tokens* exige uma análise cuidadosa dos modelos económicos e dos mecanismos de incentivo para garantir a estabilidade e o crescimento da rede. Zhang (2019) propõe uma estrutura para a engenharia de economias de *tokens* utilizando modelação de sistemas, que permite a simulação de dinâmicas complexas e a avaliação de diferentes estruturas de incentivo. Por exemplo, os modelos de *token* inflacionários, em que novos *tokens* são cunhados ao longo do tempo, podem incentivar a participação na rede, mas também podem levar à desvalorização. Inversamente, os modelos deflacionários, que limitam a oferta total de *tokens*, podem criar escassez e impulsionar a valorização, mas podem limitar o crescimento da rede (Caton, 2020; Z. Zhang, 2019).

Os mecanismos de incentivo são cruciais para alinhar os interesses dos participantes na rede, tais como mineradores, validadores e utilizadores. Por exemplo, nas redes Proof-of-Stake (PoS), os validadores são incentivados a atuar honestamente por meio de mecanismos de staking, em que bloqueiam *tokens* como garantia para participar no consenso (Freni et al., 2022; Saleh, 2021). Do mesmo modo, nas plataformas de finanças descentralizadas (DeFi), os fornecedores de liquidez são recompensados com *tokens* por contribuírem para os pools de liquidez, o que, por sua vez, facilita a negociação e reduz a derrapagem (Qin et al., 2021; Schueffel, 2021).

2.1.4. Desafios na economia de Tokens

Apesar do potencial da economia dos *tokens*, subsistem vários desafios. A incerteza regulamentar, em especial no que respeita à classificação dos *tokens* como valores mobiliários ou mercadorias, constitui um obstáculo significativo à sua adoção (Gans, 2024; Lesche et al., 2022). Além disso, a elevada volatilidade dos preços dos *tokens* pode comprometer a sua utilidade como meio de troca ou reserva de valor (Chen et al., 2023). Além disso, a falta de modelos económicos sólidos e de investigação empírica sobre a dinâmica dos *tokens* limita a nossa compreensão da forma como estes sistemas evoluem ao longo do tempo (Cong et al., 2021; Fernandez et al., 2024).

A investigação futura deve centrar-se no desenvolvimento de modelos económicos mais sofisticados para analisar a dinâmica dos *tokens*, em especial no contexto das interações das camadas 1 e 2. Além disso, abordagens interdisciplinares que combinem conhecimentos de economia, ciências da computação e séries temporais poderão proporcionar uma compreensão

mais abrangente da economia dos *tokens* e das suas implicações para a economia digital em geral.

2.2. Fundamentos matemáticos e estatísticos dos modelos de séries temporais

A tecnologia *Blockchain*, com a sua natureza descentralizada e imutável, gera grandes quantidades de dados de séries temporais, tais como volumes de transação, preços de *tokens* e métricas de rede. A análise desses dados requer modelos matemáticos e estatísticos robustos que possam capturar as características únicas dos sistemas de *blockchain*, incluindo alta volatilidade, não estacionariedade e interdependências complexas.

Uma série temporal pode ser formalmente definida como uma realização de um processo estocástico $\{X_t\}_{t \in T}$, em que as observações são indexadas pelo tempo e apresentam dependência estatística entre instantes distintos (Lütkepohl, 2005). Do ponto de vista analítico, uma série temporal observada pode ser representada como combinação de diferentes componentes estruturais. Numa formulação clássica de decomposição aditiva, esta representação assume a forma:

$$X_t = T_t + S_t + \varepsilon_t$$

Onde T_t denota a componente de tendência, responsável pela evolução média de longo prazo, S_t corresponde à componente sazonal ou periódica, associada a flutuações regulares ao longo do tempo, e ε_t representa a componente estocástica ou irregular, que incorpora choques aleatórios não explicados pelas restantes componentes.

Em determinadas séries temporais, a presença de tendência e de sazonalidade pode originar que o processo seja não estacionário. Um processo estocástico é considerado fracamente estacionário quando a sua média e variância são constantes ao longo do tempo, e a covariância entre duas observações depende apenas da distância temporal entre elas, e não dos instantes específicos em que são observadas (Box et al., 2015; Lütkepohl, 2005). A não estacionariedade pode comprometer a aplicação de modelos autorregressivos lineares que pressupõem a estabilidade dessas propriedades estatísticas. Por esse motivo, recorre-se frequentemente a transformações como a diferenciação regular, destinada a atenuar ou eliminar tendências, e a diferenciação sazonal, ou a outros procedimentos de ajustamento sazonal, destinados a remover padrões periódicos. Essas transformações procuram obter uma série estacionária, ou suficientemente próxima da estacionariedade para efeitos de modelação. É neste enquadramento que se inserem os modelos Autorregressivos Integrados de Médias Móveis (ARIMA), nos quais a componente integrada corresponde ao número de diferenciações necessárias para tornar a série estacionária, permitindo que a dinâmica remanescente seja representada por componentes autorregressivos e de médias móveis (Hyndman & Athanasopoulos, 2018a; Tunnicliffe Wilson, 2016).

Do ponto de vista estatístico, a análise de séries temporais centra-se na caracterização das propriedades probabilísticas do processo gerador de dados, em particular conceitos como estacionariedade, autocorrelação, dependência temporal e heterocedasticidade condicional. No contexto financeiro e, em particular, no estudo de criptoativos, estas séries são frequentemente não estacionárias, exibem distribuições com caudas pesadas e apresentam variância condicional dependente do tempo, bem como interdependências dinâmicas entre múltiplos

processos, exigindo modelos capazes de capturar simultaneamente dependências lineares, não lineares e estruturas de variância condicionais (Box et al., 2015; Kulik & Soulier, 2015; Lütkepohl, 2005).

Esta secção apresenta, de forma sintética, os principais modelos de séries temporais utilizados na análise de preços e volatilidade de criptomoedas: modelos lineares (ARIMA, VAR/VECM), modelos de volatilidade condicional (GARCH e DCC-GARCH) e modelos não lineares baseados em redes neuronais recorrentes (LSTM), incluindo abordagens híbridas ARIMA–LSTM. O objetivo desta secção é fornecer a base matemática estritamente necessária para enquadrar a metodologia aplicada na investigação, articulando-a com a literatura recente sobre séries temporais e mercados financeiros (Arumugam & Natarajan, 2023a; Y. Chen, 2024a; Lara-Benítez et al., 2021; Lütkepohl, 2022).

2.2.1. Modelos ARIMA (Autoregressive Integrated Moving Average)

Os modelos ARIMA (AutoRegressive Integrated Moving Average) representam uma classe flexível para séries temporais estacionárias ou tornadas estacionárias por diferenciação. Introduzidos por Box e Jenkins, estes modelos continuam a ser uma referência para modelação e previsão de séries temporais em vários domínios (Arumugam & Natarajan, 2023a; Box & Jenkins, 1976; Gunawan & Febrianti, 2023; Rizvi, 2024; Upadhyay & Pradhan, 2023). Seja y_t o logaritmo do preço ou o retorno de um ativo. Um processo ARIMA(p,d,q) é caracterizado por (p), a ordem da componente autorregressiva; (d), o número de diferenciações necessárias para tornar a série estacionária; e (q), a ordem da componente de médias móveis, correspondente ao número de impulsos aleatórios, ou inovações passadas, considerados na representação do valor atual da série, a qual pode ser expressa, em notação compacta, como:

$$\varphi(L)(1 - L)^d y_t = \theta(L)\varepsilon_t$$

em que L é o operador de defasagem, $\varphi(L) = 1 - \varphi_1 L - \dots - \varphi_p L^p$ é o polinómio autorregressivo, $\theta(L) = 1 + \theta_1 L + \dots + \theta_q L^q$ é o polinómio de média móvel, e ε_t representa um processo de ruído, entendido como um processo estocástico com média nula, variância constante σ^2 e ausência de autocorrelação, geralmente assumido como ruído branco. Assim, a ordem p controla o número de defasagens autorregressivas, d o número de diferenças necessárias para obter estacionariedade e q o número de termos de média móvel incluídos no modelo (Arumugam & Natarajan, 2023b).

Nos mercados de criptomoedas, os modelos ARIMA são frequentemente utilizados como benchmark linear para previsão de preços ou retornos, servindo de base de comparação para modelos mais sofisticados de volatilidade ou de aprendizagem profunda (Cao & Wang, 2024; Santoso & Widodo, 2024).

2.2.2. Modelos multivariados, cointegração e VECM

Quando se analisam conjuntamente várias séries temporais, como os preços ou os retornos de *tokens* de Camada 1 e de Camada 2, podem utilizar-se modelos vetoriais autorregressivos, designados por VAR (Vector Autoregressive). Seja $(y_{i,t})$ o valor observado da (i)-ésima série temporal no instante (t), com $(i = 1, \dots, k)$. O conjunto das (k) séries é representado pelo vetor coluna:

$$\mathbf{y}_t = \begin{pmatrix} y_{1,t} \\ y_{2,t} \\ \vdots \\ y_{k,t} \end{pmatrix} \in \mathbb{R}^k$$

que, no contexto desta investigação, pode incluir os retornos de ETH, MATIC e OP. Um modelo VAR de ordem (p), denotado por VAR(p), é definido por

$$y_t = \mathbf{c} + \mathbf{A}_1 \mathbf{y}_{t-1} + \mathbf{A}_2 \mathbf{y}_{t-2} + \cdots + \mathbf{A}_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t,$$

em que $(\mathbf{y}_t)$ é o vetor $(k \times 1)$ das variáveis observadas no instante (t) , $(\mathbf{c})$ é um vetor $(k \times 1)$ de constantes, $(\mathbf{A}_j)$, para $(j = 1, \dots, p)$, são matrizes $(k \times k)$ de coeficientes autorregressivos e $(\boldsymbol{\varepsilon}_t)$ é um vetor $(k \times 1)$ de inovações, geralmente assumido com média nula e matriz de covariâncias constante. Deste modo, cada variável do sistema é explicada pelos seus próprios valores desfasados e pelos valores desfasados das restantes variáveis. Os modelos VAR permitem, assim, representar dependências dinâmicas de curto prazo entre várias séries temporais e constituem uma ferramenta central da econometria multivariada (Diebold & Yilmaz, 2023; Lütkepohl, 2022).

Se algumas combinações lineares das séries forem estacionárias, apesar de cada série isolada ser integrada, diz-se que existe cointegração. Uma série temporal (y_t) diz-se integrada de ordem (d), denotada por $(y_t \sim I(d))$, quando se torna estacionária após ser diferenciada (d) vezes. Em particular, uma série integrada de ordem um, $(I(1))$, é não estacionária em níveis, mas a sua primeira diferença,

$$[\Delta y_t = y_t - y_{t-1},]$$

é estacionária. Quando duas ou mais séries são individualmente integradas da mesma ordem, geralmente $(I(1))$, mas há uma combinação linear entre elas que é estacionária, diz-se que as séries são cointegradas. O conceito de cointegração, introduzido por Granger (1981) e formalizado por Engle e Granger (1987), traduz a existência de uma relação de equilíbrio de longo prazo entre variáveis não estacionárias, cujos desvios em relação a esse equilíbrio permanecem estacionários. Um vetor de cointegração $\beta \in \mathbb{R}^k$ tal que:

$$\beta' y_t \text{ é estacionário } (I(0))$$

no qual y_t é o vetor de series integradas de ordem 1, $I(1)$. β o vetor (ou matriz) de cointegração e $\beta' y_t$ a combinação linear que é estacionária, ou seja, define uma relação de equilíbrio de longo prazo entre as variáveis (Engle & Granger, 1987; Pesaran et al., 2022). Neste caso, a dinâmica conjunta pode ser representada por um modelo VECM (Vector Error Correction Model):

$$\Delta y_t = \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + \cdots + \Gamma_{p-1} \Delta y_{t-p+1} + \varepsilon_t$$

Se o posto da matriz $(\Pi \in \mathbb{R}^{k \times k})$ for igual a (r) , com $(0 < r < k)$, então (Π) apresenta posto reduzido e pode ser decomposta como

$$[\Pi = \alpha\beta',]$$

em que $(\alpha \in \mathbb{R}^{k \times r})$ é a matriz dos coeficientes de ajustamento e $(\beta \in \mathbb{R}^{k \times r})$ contém, nas suas colunas, os (r) vetores de cointegração linearmente independentes. Cada vetor de cointegração define uma combinação linear estacionária das variáveis, interpretada como uma relação de equilíbrio de longo prazo. A matriz (α) descreve a velocidade e a direção com que cada variável se ajusta aos desvios em relação a esses equilíbrios, enquanto os termos $(\Gamma_i \Delta \mathbf{y}_{t-i})$ representam a dinâmica de curto prazo. Os testes de máxima verosimilhança de Johansen permitem determinar o posto (r) de (Π) e, conseqüentemente, o número de relações de cointegração existentes no sistema (Johansen, 1988; Johansen & Juselius, 2023; Johansen & Nielsen, 2022; Lütkepohl, 2005).

No contexto desta investigação, estes modelos permitem avaliar se os preços de tokens de Camada 2 partilham um equilíbrio de longo prazo com o token de Camada 1 (ETH) e estudar a causalidade de Granger entre as séries, seguindo a formulação clássica de Granger (1969) e desenvolvimentos recentes em testes de causalidade para séries potencialmente não lineares (Diks & Wolski, 2023; Ghysels & Marcellino, 2023; Narayan & Smyth, 2023).

2.2.3. Modelos de volatilidade condicional: GARCH e DCC-GARCH

Séries financeiras exibem frequentemente heterocedasticidade condicional, com períodos de maior e menor turbulência. Os modelos GARCH (Generalized AutoRegressive Conditional Heteroskedasticity) capturam esta dinâmica e foram desenvolvidos a partir do modelo ARCH de Engle (1982), posteriormente generalizado por Bollerslev (1986, 1987). Estes modelos tornaram-se uma pedra angular na análise da volatilidade em finanças, incluindo em mercados de criptomoedas (Chen, 2024; Nugroho et al., 2021; Osho & Oloyede, 2024). Para os retornos r_t , considera-se:

$$r_t = \mu_t + \varepsilon_t, \quad \varepsilon_t = z_t \sigma_t$$

em que μ_t é a média condicional (e.g. constante ou ARMA), ε_t o erro (i.e. inovação) ou ruído branco condicional, z_t é uma sequência i.i.d. com média zero e variância 1 (e.g. normal ou t-Student), e σ_t^2 é a variância condicional. No caso GARCH(1, 1):

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

com $\omega > 0$, $\alpha \geq 0$, $\beta \geq 0$. O parâmetro α mede a reação da volatilidade a choques recentes (i.e. efeito ARCH), enquanto β mede a persistência de volatilidade (i.e. efeito GARCH) e a soma $\alpha + \beta$ indica o grau de memória da volatilidade. Valores próximos de 1 sugerem uma persistência elevada, frequentemente observada em dados financeiros (Chen, 2024; Lindner, 2009; Wang, 2021).

Para estudar a volatilidade conjunta de vários ativos e o *spillover* de risco entre eles, utilizam-se extensões multivariadas dos modelos GARCH. O modelo DCC-GARCH (Dynamic Conditional Correlation GARCH) é particularmente popular por ser parcimonioso e permitir correlações condicionais dinâmicas entre ativos (Chou & Kroner, 1992). Seja $r_{i,t}$ o retorno do

ativo (i), no instante (t), com ($i = 1, \dots, k$). O vetor dos retornos dos (k) ativos é representado por

$$\mathbf{r}_t = (r_{1,t}, r_{2,t} : r_{k,t}) \in \mathbb{R}^k.$$

No modelo DCC-GARCH, assume-se que

$$\mathbf{r}_t = \boldsymbol{\mu}_t + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t | \mathcal{F}_{t-1} \sim (\mathbf{0}, \mathbf{H}_t),$$

em que ($\boldsymbol{\mu}_t$) é o vetor das médias condicionais, ($\boldsymbol{\varepsilon}_t$) é o vetor das inovações e ($\mathbf{H}_t \in \mathbb{R}^{k \times k}$) é a matriz de covariâncias condicionais, a qual é decomposta num produto de matrizes:

$$\mathbf{H}_t = \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t$$

com $\mathbf{D}_t = \text{diag}(\sigma_{1t}, \dots, \sigma_{kt})$ contendo os desvios-padrão condicionais univariados e $\mathbf{R}_t$ a matriz de correlação condicional dinâmica. No DCC(1, 1), a normalização do $\mathbf{R}_t$ é obtida a partir de uma matriz intermédia $\mathbf{Q}_t$:

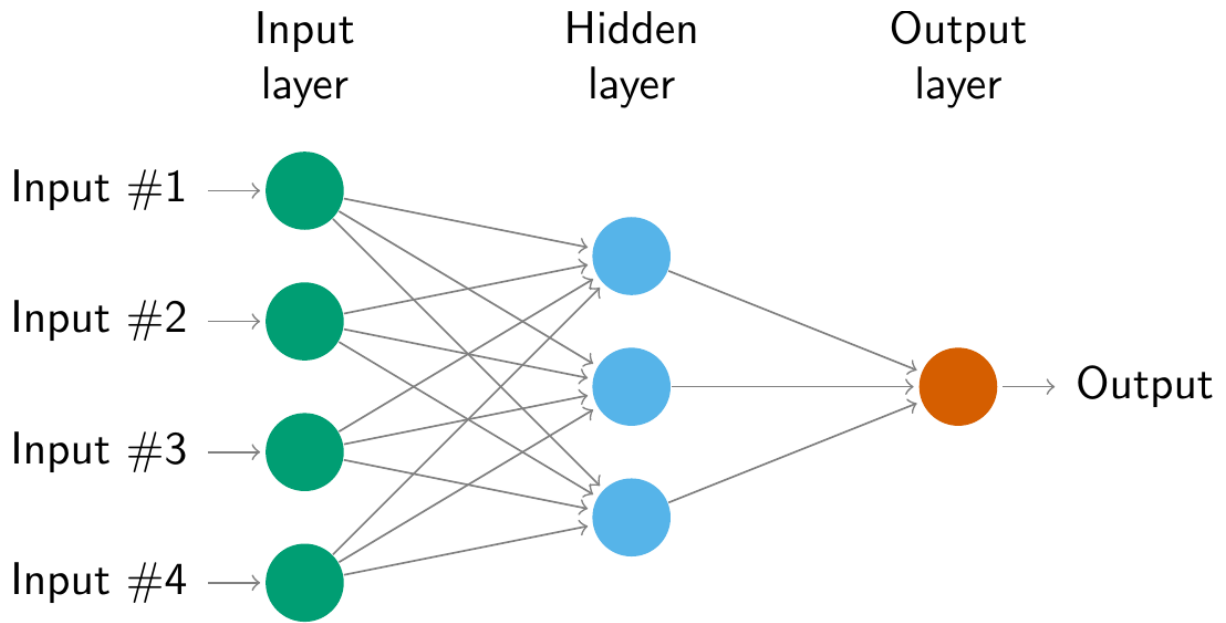
$$\mathbf{Q}_t = (1 - a - b)\bar{\mathbf{Q}} + a(z_{t-1}z'_{t-1}) + b\mathbf{Q}_{t-1}$$

em que $\mathbf{Q}_t \in \mathbb{R}^{k \times k}$ é a matriz intermédia de correlações não normalizadas, a e b os parâmetros DCC (i.e. reação a choques e persistência da correlação), os $z_t = \mathbf{D}_t^{-1}\boldsymbol{\varepsilon}_t$ são resíduos padronizados, $\bar{\mathbf{Q}}$ é a matriz de correlação incondicional de z_t , e $a, b \geq 0$ são parâmetros que controlam a reação da correlação a choques e a sua persistência. A matriz de correlação é depois normalizada para obter $\mathbf{R}_t$. Este modelo permite analisar como a correlação entre ETH e tokens de Camada 2 evolui ao longo do tempo e em que medida choques num ativo se propagam aos restantes (Diebold & Yilmaz, 2023; Engle, 2002).

2.2.4. Modelos de aprendizagem profunda: Redes LSTM

As redes neuronais artificiais (ANN) são modelos computacionais inspirados na estrutura dos neurónios biológicos, compostos por camadas de unidades interligadas que aplicam transformações não lineares aos dados de entrada. Em tarefas supervisionadas, estes modelos aprendem um mapeamento entre entradas e saídas ajustando os pesos das ligações através de algoritmos de otimização, como o gradiente descendente, com base numa função de perda (Goodfellow et al., 2014).

Figura 2-1 Exemplo de rede neural



Fonte: Hyndman & Athanasopoulos (2018b). A figura ilustra a arquitetura genérica de uma rede neuronal artificial do tipo feedforward, composta por uma camada de entrada (inputs), uma camada oculta (hidden layer) e uma camada de saída (output layer)².

As redes neurais recorrentes (RNN) estendem esta ideia introduzindo ligações temporais no estado oculto, permitindo que a rede mantenha uma “memória” de estados anteriores e, assim, seja particularmente adequada para dados sequenciais, como séries temporais ou texto (Lara-Benítez et al., 2021). No entanto, as RNN clássicas sofrem frequentemente de problemas de vanishing e exploding gradients, o que dificulta a aprendizagem de dependências de longo prazo em séries extensas.

Por outro lado, as redes LSTM (Long Short-Term Memory) foram concebidas para ultrapassar essa limitação, incorporando mecanismos de gating que controlam o fluxo de informação e permitem capturar dependências de longo prazo (Dung et al., 2020; Hochreiter & Schmidhuber, 1997; Lara-Benítez et al., 2021). Para uma unidade LSTM, no instante t , com entrada x_t , estado oculto h_t e estado da célula c_t , as equações principais podem ser resumidas como:

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f)$$

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i)$$

$$\tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t$$

²Uma rede neuronal artificial, evidenciando o fluxo de informação entre a camada de entrada, a camada oculta e a camada de saída. As ligações ponderadas entre neurónios permitem a aprendizagem de relações não lineares entre variáveis, sendo esta estrutura a base conceptual das arquiteturas mais avançadas, como as redes LSTM, utilizadas nesta investigação.

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o)$$

$$h_t = o_t \odot \tanh(c_t)$$

Aqui, $\sigma(\cdot)$ é a função sigmóide, $\odot$ representa o produto elemento a elemento e W , U , b são os pesos e *biases*. O x_t a entrada no tempo t (e.g. retorno), h_t o estado oculto (i.e. *output*), c_t o “estado da célula” (i.e. memória). Esta estrutura permite às LSTM modelar não linearidades e relações complexas nos retornos de criptomoedas, capturando padrões que modelos lineares como ARIMA ou VAR não conseguem representar. Estudos recentes mostram a eficácia das LSTM em previsão financeira, detecção de anomalias e previsão de procura de energia (Acuna-García, 2022; Hochreiter & Schmidhuber, 1997; Li & Dai, 2024; Ormaniec et al., 2022; Raut, 2024).

2.2.5. Modelos híbridos ARIMA–LSTM

Modelos híbridos combinam a capacidade de modelação linear dos ARIMA com a capacidade de captura de componentes não lineares das LSTM. Uma abordagem comum consiste em utilizar um modelo ARIMA para capturar a estrutura linear de curto prazo e, em seguida, aplicar uma LSTM aos resíduos, de modo a modelar a componente não linear remanescente (Jayawardhana & Colombage, 2024)

De forma esquemática, procede-se da seguinte forma: (i) ajusta-se um modelo ARIMA à série y_t e obtêm-se os resíduos $\hat{\varepsilon}_t$; (ii) treina-se uma rede LSTM para modelar a componente não linear dos resíduos; e (iii) a previsão híbrida resulta da soma entre a previsão ARIMA e a previsão da LSTM aplicada aos resíduos. Em termos genéricos:

$$\hat{\varepsilon}_t \approx f_{LSTM}(\hat{\varepsilon}_{t-1}, \dots, \hat{\varepsilon}_{t-k}),$$

$$\hat{y}_{t+h}^{(H)} = \hat{y}_{t+h}^{(ARIMA)} + \hat{\varepsilon}_{t+h}^{(LSTM)}$$

na qual, $\hat{\varepsilon}_t$ são os resíduos do ARIMA (i.e. parte não explicada), $\hat{y}_{t+h}^{(ARIMA)}$ representa a previsão linear ARIMA, f_{LSTM} a função aprendida pela LSTM para modelar a parte não linear dos resíduos e $\hat{y}_{t+h}^{(H)}$ a previsão híbrida (i.e. componente linear + componente não linear).

Esta estratégia permite utilizar o modelo ARIMA para capturar a estrutura linear de curto prazo, enquanto a LSTM aprende padrões residuais não lineares. Em mercados jovens e voláteis como os de criptoativos, esta combinação tende a apresentar melhor desempenho preditivo do que modelos puramente lineares ou puramente neuronais, em linha com evidência recente em previsão de preços de criptomoedas e outros ativos financeiros (Fleischer et al., 2022; Jayawardhana & Colombage, 2024).

2.2.6. Comparação entre modelos lineares e não lineares na análise de séries temporais

A literatura recente evidencia uma distinção fundamental entre modelos lineares tradicionais e modelos não lineares baseados em aprendizagem profunda, ambos amplamente aplicados na análise de séries temporais financeiras e, mais recentemente, de criptoativos. Os primeiros, como ARIMA, VAR e GARCH, oferecem estrutura econométrica clara, elevada

interpretabilidade e propriedades estatísticas bem estabelecidas (Lütkepohl, 2022; Engle, 2002). Já os segundos, como LSTM e variantes recorrentes, destacam-se pela capacidade de modelar dependências complexas, não linearidades e relações de longo prazo, embora com maior custo computacional e menor interpretabilidade (Giantsidi & Tarantola, 2025; Lara-Benítez et al., 2021). Para efeitos de sistematização, a Tabela 2-1 sintetiza as principais diferenças entre estas abordagens, de acordo com a literatura recente sobre econometria financeira, aprendizagem automática e previsão de séries temporais.

Tabela 2-1 Comparação entre modelos lineares e modelos não lineares de previsão de séries temporais

Característica	Modelos Lineares (ARIMA, VAR, GARCH)	Modelos Não Lineares (LSTM, RNN, Modelos Híbridos)
Natureza do modelo	Estruturas paramétricas lineares que assumem relações proporcionais entre variáveis (Lütkepohl, 2022).	Arquiteturas flexíveis capazes de representar funções altamente não lineares (Lara-Benítez et al., 2021).
Pressupostos principais	Estacionariedade, linearidade e relações temporais de curta memória.	Poucas assunções; capacidade de capturar dependências não lineares e de longo prazo (Giantsidi & Tarantola, 2025).
Interpretação dos parâmetros	Elevada interpretabilidade; coeficientes têm significado económico claro.	Baixa interpretabilidade; modelos são tratados como “caixa-preta”.
Sensibilidade a mudanças de regime	Moderada; podem ser ajustados com reestimação frequente.	Elevada vulnerabilidade a regime changes e concept drift (Giantsidi & Tarantola, 2025).
Risco de <i>overfitting</i> ³	Baixo a moderado devido ao número limitado de parâmetros.	Elevado, sobretudo em séries curtas e ruidosas (Vuković et al., 2024)
Custo computacional	Baixo; estimação rápida com poucos recursos.	Alto; requer treino intensivo e, por vezes, GPU (Sezgin et al., 2025).
Capacidade de capturar não linearidades	Limitada, exceto com extensões (e.g., modelos não lineares ou Threshold VAR).	Alta; conseguem modelar estruturas complexas e padrões irregulares.
Aplicabilidade em criptoativos	Muito adequados para análise estrutural: cointegração, causalidade, volatilidade (Dyhrberg, 2016; Engle et al., 2013; Malladi & Dheeriy, 2021).	Adequados para previsão em ambientes voláteis, mas exigem regularização e cuidado metodológico (Giantsidi & Tarantola, 2025).

A comparação mostra que modelos lineares continuam a desempenhar um papel essencial na análise estrutural de mercados financeiros, sobretudo quando o objetivo passa por interpretar relações económicas, testar hipóteses ou modelar volatilidade de forma parcimoniosa. No entanto, a crescente complexidade dos mercados de criptoativos, caracterizados por elevada volatilidade, mudanças estruturais rápidas e comportamentos não lineares, tem motivado o recurso a abordagens de aprendizagem profunda (*Deep Learning*), que demonstram vantagens claras em tarefas preditivas. Ainda assim, a literatura alerta para riscos importantes, como a sensibilidade a hiperparâmetros, vulnerabilidade a mudanças de regime e propensão para *overfitting*, que justificam o uso complementar de métodos lineares e não lineares em investigações contemporâneas (Giantsidi & Tarantola, 2025; Vuković et al., 2024).

Embora os modelos de *deep learning* tenham demonstrado ganhos relevantes na previsão de séries temporais financeiras, a literatura destaca várias limitações que devem ser consideradas.

³ Risco de *overfitting* corresponde à situação em que um modelo se ajusta excessivamente aos dados de treino, capturando ruído em vez da estrutura subjacente, o que compromete a sua capacidade de generalização fora da amostra (Hastie et al., 2009)

Em particular, estes modelos revelam fragilidade quando confrontados com mudanças de regime, uma vez que padrões aprendidos num determinado contexto de mercado nem sempre se mantêm válidos sob novas condições estruturais (Giantsidi & Tarantola, 2025). Além disso, o seu desempenho é altamente sensível à escolha de hiperparâmetros, como a dimensão das camadas ocultas, o número de épocas de treino ou a janela temporal de entrada, o que pode resultar em variações significativas nos resultados e exigir processos complexos de otimização (Sezgin et al., 2025). Por fim, a elevada flexibilidade destas arquiteturas aumenta o risco de *overfitting*, sobretudo em séries curtas e ruidosas, levando os modelos a capturar padrões não casualísticos que não se generalizam fora da amostra (Vuković et al., 2024). Estas limitações sublinham a importância de uma abordagem cautelosa e devidamente regularizada quando se aplicam métodos de deep learning ao estudo de criptoativos.

2.3. Análise de séries temporais na investigação sobre *blockchain* e criptomoedas

Enquanto a secção anterior apresentou os fundamentos teóricos dos modelos de séries temporais, a presente secção analisa como estes modelos têm sido aplicados especificamente ao estudo de mercados de criptomoedas. O objetivo não é descrever novamente os modelos, mas contextualizar a sua utilização empírica, identificando padrões, limitações e oportunidades que informam a presente investigação.

2.3.1. Modelos tradicionais de séries temporais

Modelos tradicionais de séries temporais, como a Autoregressive Integrated Moving Average (ARIMA) e a Generalized Autoregressive Conditional Heteroskedasticity (GARCH), têm sido amplamente utilizados para analisar os movimentos dos preços das criptomoedas. Os modelos ARIMA são eficazes na captura de dependências lineares em dados de séries temporais, tornando-os adequados para previsões de curto prazo dos preços das criptomoedas. Por exemplo, Azari (2019) demonstrou que os modelos ARIMA podem prever os preços do Bitcoin com baixo erro quadrático médio, estabelecendo sua viabilidade para aplicações de previsão de curto prazo (Freni et al., 2022). No entanto, os modelos ARIMA assumem linearidade e estacionariedade, o que pode limitar a sua capacidade de capturar totalmente o comportamento errático dos mercados de criptomoedas.

Em contrapartida, os modelos GARCH são particularmente adequados para modelar o agrupamento de volatilidade, uma característica prevalente nos mercados de criptomoedas, caracterizada por períodos de alta e baixa volatilidade. Dyhrberg (2016) aplicou modelos GARCH para comparar a volatilidade do Bitcoin com ativos tradicionais, como o ouro e o dólar americano, destacando os padrões de volatilidade únicos do Bitcoin. Da mesma forma, Zhang (2019) utilizou modelos GARCH para analisar a dinâmica da volatilidade de várias criptomoedas, ressaltando sua capacidade de modelar a heterocedasticidade em séries temporais financeiras. Esses estudos indicam que, embora modelos tradicionais como ARIMA e GARCH forneçam insights valiosos, suas limitações no tratamento de padrões não lineares e dependências de longo prazo levaram à exploração de técnicas mais avançadas.

2.3.2. Abordagens de aprendizagem automática e aprendizagem profunda

O advento da aprendizagem automática e da aprendizagem profunda levou ao desenvolvimento de modelos mais sofisticados, capazes de capturar padrões não lineares e complexos em dados de criptomoedas. As redes de memória longa de curto prazo (LSTM), um tipo de rede neural recorrente (RNN), ganharam destaque devido à sua capacidade de modelar dependências de longo prazo em dados sequenciais. Fleischer et al. (2022) aplicaram modelos LSTM para prever os preços de fechamento de várias criptomoedas, demonstrando seu desempenho superior em comparação com modelos tradicionais como ARIMA. Da mesma forma, Malladi e Dheeriyaa (2021) utilizaram modelos LSTM para analisar a volatilidade das criptomoedas, mostrando maior precisão de previsão em relação às abordagens tradicionais.

Além disso, modelos híbridos que combinam métodos tradicionais de séries temporais com técnicas de aprendizado de máquina têm se mostrado promissores no aprimoramento do desempenho preditivo. Por exemplo, Jayawardhana e Colombage (2024) propuseram uma abordagem híbrida que integra modelos ARIMA e LSTM para prever os preços das criptomoedas, alcançando uma precisão melhor do que qualquer um dos modelos isoladamente. Essas descobertas destacam o potencial do aprendizado de máquina e do aprendizado profundo para lidar com as limitações dos modelos tradicionais, particularmente na captura da natureza não linear e dinâmica dos mercados de criptomoedas.

2.3.3. Análise de Cointegração e Causalidade

As análises de cointegração e causalidade são fundamentais para compreender as relações de longo prazo e as dependências direcionais entre criptomoedas ou entre *tokens* de Camada 1 e Camada 2. A análise de cointegração tem sido usada para identificar relações de equilíbrio de longo prazo entre criptomoedas. Nair (2021) aplicou a análise de cointegração para explorar estratégias de negociação em pares nos mercados de criptomoedas, identificando relações estáveis de longo prazo entre diferentes *tokens*. Essa abordagem fornece insights sobre potenciais oportunidades de arbitragem e ineficiências do mercado.

Testes de causalidade de Granger também foram utilizados para investigar relações causais entre os preços das criptomoedas. Malladi e Dheeriyaa (2021) utilizaram testes de causalidade de Granger para examinar as relações entre retornos e volatilidades de diferentes criptomoedas, encontrando evidências de causalidade bidirecional em alguns casos. Essas análises são essenciais para compreender a interconectividade dos mercados de criptomoedas e o fluxo direcional de informações, o que pode informar estratégias de negociação e práticas de gestão de risco.

Esta secção destaca a evolução da análise de séries temporais na investigação sobre criptomoedas, desde modelos tradicionais como ARIMA e GARCH até abordagens avançadas de aprendizagem automática e análises de cointegração. Embora os modelos tradicionais continuem a ser valiosos pela sua simplicidade e interpretabilidade, as técnicas de aprendizagem automática e aprendizagem profunda oferecem um desempenho superior na captura de dinâmicas de mercado complexas. As análises de cointegração e causalidade melhoram ainda mais a nossa compreensão das relações entre mercados, fornecendo uma estrutura robusta para analisar o panorama em rápida evolução das criptomoedas.

2.4. Enquadramento teórico e questões de investigação

Apesar do avanço significativo da investigação sobre mercados de criptoativos, persistem limitações importantes que restringem a compreensão aprofundada das suas dinâmicas económicas. Em primeiro lugar, a maioria dos estudos empíricos continua excessivamente centrada no Bitcoin, dada a sua antiguidade e dominância de mercado. Embora esta escolha seja metodologicamente conveniente, conduz a conclusões dificilmente generalizáveis a ecossistemas tecnologicamente distintos, como o da Ethereum, nos quais os mecanismos de segurança, incentivo e escalabilidade assumem formas substancialmente diferentes (Dyhrberg, 2016; Zhang, 2019). Como resultado, a literatura permanece desproporcionalmente focada num único ativo, deixando por explorar relações estruturais mais complexas entre criptomoedas interdependentes.

Em segundo lugar, verifica-se uma escassez de estudos que analisem explicitamente as interações entre tokens de Camada 1 e Camada 2. Investigação recente aborda volatilidade, causalidade ou previsibilidade de forma isolada (Malladi & Dheeriy, 2021; Nair, 2021) mas raramente considera a arquitetura hierárquica que liga as soluções de escalabilidade às redes base. Assim, permanece por esclarecer se a dependência técnica das L2 relativamente à L1, incluindo a partilha de segurança, liquidação e blockspace, se traduz também em relações económicas mensuráveis, como equilíbrios de longo prazo, causalidade bidirecional ou efeitos de contágio na volatilidade.

Por fim, apesar da crescente utilização de aprendizagem profunda na previsão de preços e volatilidade de criptoativos (Fleischer et al., 2022; Lara-Benítez et al., 2021) a integração destas abordagens com modelos económicos clássicos permanece limitada. A literatura tende a separar modelos estruturais, como VAR/VECM ou GARCH, de modelos não lineares, como LSTM ou arquiteturas híbridas, o que impede a exploração do seu potencial complementar. Esta fragmentação metodológica reduz a capacidade de capturar simultaneamente relações de equilíbrio, dependências dinâmicas e padrões não lineares característicos dos mercados de criptomoedas.

Com base neste enquadramento, formulam-se as seguintes questões de investigação, alinhadas com os objetivos definidos na Introdução e retomadas na análise empírica:

Q1. Existem relações de equilíbrio de longo prazo (cointegração) e padrões de causalidade direcionais entre o token de Camada 1 (ETH) e os tokens de Camada 2 (MATIC e OP), e como se caracterizam estas dinâmicas estruturais?"

Q2. Em que medida modelos híbridos de Aprendizagem Profunda (ARIMA–LSTM), integrando mecanismos de atenção e variáveis exógenas, superam o desempenho preditivo de modelos lineares tradicionais (VAR/ARIMA) na previsão de preços destes tokens?

Q3. Como se caracteriza a estrutura de volatilidade condicional e os efeitos de spillover entre as diferentes camadas, e qual a dinâmica das correlações em períodos de instabilidade de mercado?

3. METODOLOGIA

Este capítulo apresenta o enquadramento metodológico delineado para responder aos objetivos da investigação, articulando a robustez da econometria financeira clássica com a flexibilidade das técnicas de aprendizagem profunda. A estratégia empírica foi concebida para analisar as interações entre a Camada 1 e a Camada 2 em três dimensões complementares, organizadas de forma sequencial e alinhadas com as questões de investigação previamente formuladas.

Numa primeira vertente, procede-se à caracterização estrutural dos ativos e à avaliação de relações causais. Esta etapa centra-se na identificação de equilíbrios de longo prazo e na determinação da direção dos fluxos de informação, recorrendo aos testes de cointegração de Johansen e à análise de causalidade de Granger. O objetivo é esclarecer a hierarquia económica subjacente à infraestrutura tecnológica e os mecanismos estruturais que a sustentam.

A segunda vertente incide sobre a modelação dinâmica e a previsão. Aqui avalia-se a capacidade preditiva e as relações de curto prazo, comparando modelos lineares tradicionais, como VAR e Auto-ARIMA, com uma abordagem híbrida multivariada (ARIMA-LSTM) enriquecida com mecanismos de atenção e variáveis exógenas. Esta arquitetura visa captar padrões não lineares e dependências complexas que caracterizam a interação entre os ativos.

Por fim, a terceira vertente dedica-se à análise da volatilidade e do contágio. A estrutura de risco e a evolução da correlação dinâmica entre os tokens são modeladas através de estratégias multivariadas de heterocedasticidade condicional, em particular o modelo DCC-GARCH, permitindo identificar efeitos de spillover e variações na interdependência durante períodos de instabilidade.

As secções seguintes detalham os dados utilizados, os procedimentos de pré-processamento e a especificação técnica de cada um dos modelos que compõem esta abordagem metodológica.

3.1. Dados, seleção dos tokens e período de análise

A investigação focou-se em três tokens específicos: Ethereum (ETH) como protocolo de Camada 1 e Polygon (MATIC) e Optimism (OP) como soluções de Camada 2. A seleção destes tokens baseou-se em critérios de relevância de mercado, volume de transações e maturidade tecnológica. O Ethereum foi escolhido por ser a principal plataforma de contratos inteligentes com o ecossistema de Camada 2 mais desenvolvido. Polygon e Optimism representam duas abordagens técnicas distintas para escalabilidade: Polygon como sidechain e Optimism como rollup otimista.

Os dados históricos de preços foram obtidos através da plataforma CoinGecko®, utilizada como fonte única para as três séries analisadas, de forma a assegurar a uniformidade do procedimento de recolha. Para cada token, foram recolhidos os preços de fecho diários, expressos em dólares norte-americanos, referentes ao período de 30 de junho de 2022 a 30 de junho de 2025. A utilização da mesma fonte, da mesma frequência temporal e da mesma moeda de referência para todos os ativos permitiu preservar a comparabilidade entre as séries. Após a recolha, foram realizados procedimentos de verificação da continuidade cronológica, identificação de valores ausentes e controlo de observações potencialmente anómalas.

3.2. Transformação das séries e análise preliminar

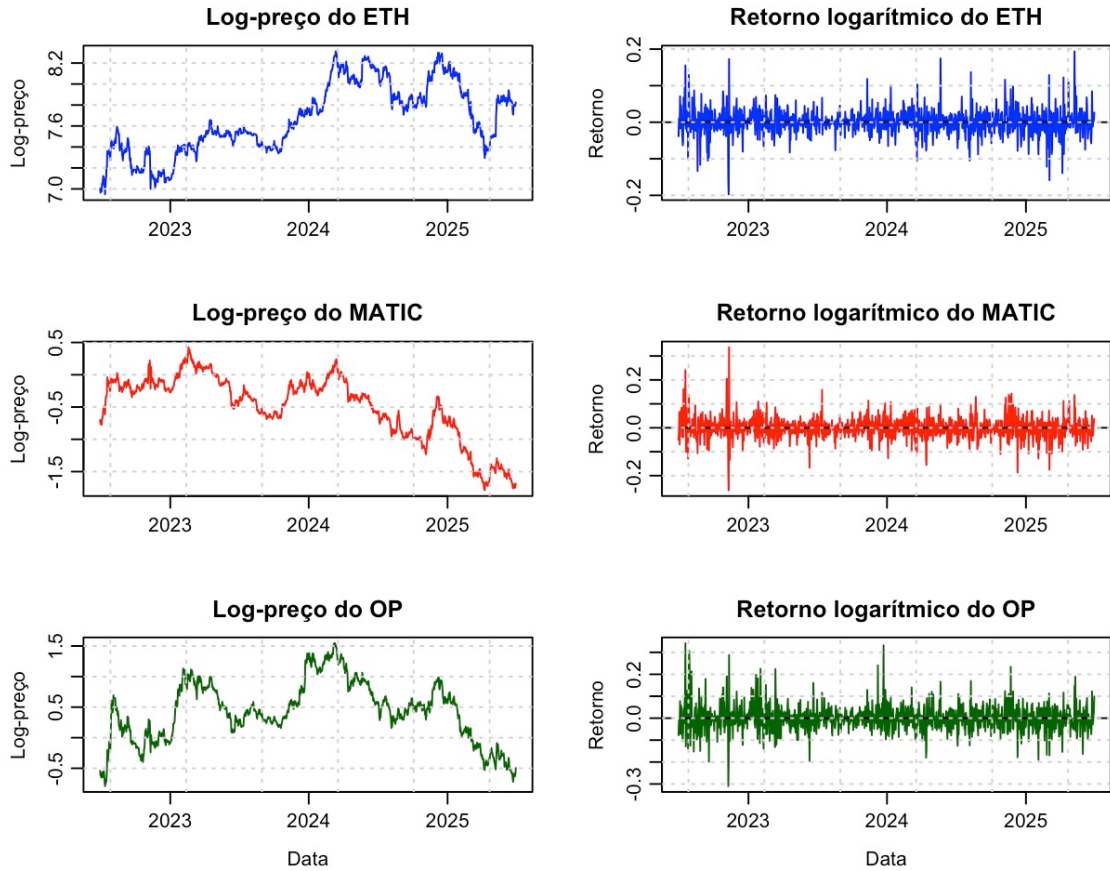
As séries de preços foram transformadas em retornos logarítmicos para estabilizar a variância e facilitar a interpretação económica. Para um dado token, o retorno logarítmico diário é definido por:

$$R_t = \ln(P_t) - \ln(P_{t-1})$$

Onde P_t representa o preço no momento t . Esta transformação é padrão na literatura financeira e permite a comparabilidade entre diferentes escalas de preços dos tokens, além de aproximar os retornos de uma distribuição mais simétrica.

A Figura 3-1 apresenta a evolução dos log-preços e dos retornos logarítmicos diários de ETH, MATIC e OP entre 30/06/2022 e 30/06/2025. Os log-preços revelam trajetórias distintas ao longo do período analisado: o ETH apresenta uma evolução globalmente ascendente até meados de 2024, seguida de uma correção acentuada e de uma recuperação parcial; o MATIC evidencia uma tendência descendente mais marcada, sobretudo na parte final da amostra; e o OP regista uma valorização inicial expressiva, seguida de reversão. Em contraste, os retornos logarítmicos oscilam em torno de zero e não exibem uma tendência persistente, embora apresentem episódios de elevada variabilidade e observações extremas, particularmente em determinados momentos da amostra. Este comportamento sugere maior proximidade à estacionariedade nos retornos do que nos log-preços e evidencia a presença de agrupamentos de volatilidade, justificando a utilização dos retornos nas análises de curto prazo e nos modelos de volatilidade condicional, enquanto os log-preços são mantidos para os testes de raiz unitária e de cointegração destinados à avaliação das relações de longo prazo.

Figura 3-1 Log-preços e retornos de ETH, MATIC e OP diários



Fonte: Elaboração própria com base nos dados do CoinGecko®

A estacionariedade das séries foi avaliada por meio do teste Augmented Dickey–Fuller (ADF), aplicado aos log-preços e aos retornos logarítmicos de ETH, MATIC e OP. O teste ADF permite avaliar a presença de uma raiz unitária, tendo como hipótese nula que a série é não estacionária (Dickey & Fuller, 1979; Lütkepohl, 2005):

$$H_0: \gamma = 0$$

contra a hipótese alternativa de estacionariedade,

$$H_1: \gamma < 0$$

A regressão auxiliar do teste pode ser representada, na sua forma geral, por

$$\alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^p \delta_i \Delta y_{t-i} + \varepsilon_t$$

em que (y_t) representa a série analisada, (α) é uma constante, (βt) corresponde a uma eventual tendência determinística, (p) é o número de defasamentos das primeiras diferenças e (ε_t) é o termo de inovação.

Para os log-preços, o teste foi especificado com constante e tendência determinística, atendendo à evolução persistente observada nas séries em níveis. Para os retornos logarítmicos,

utilizou-se uma especificação com constante, uma vez que essas séries oscilam em torno de um nível aproximadamente constante e não apresentam uma tendência determinística evidente. O número de defasamentos foi selecionado pelo critério de informação de Akaike, considerando um número máximo de 30 defasamentos..

A decisão foi tomada por comparação entre a estatística do teste e os valores críticos correspondentes. A hipótese nula de raiz unitária foi rejeitada sempre que a estatística ADF se revelou inferior, isto é, mais negativa, do que o valor crítico correspondente. Esta análise permitiu determinar a ordem de integração das séries e fundamentar a utilização dos log-preços nos testes de cointegração e dos retornos logarítmicos nos modelos VAR e de volatilidade condicional.

As estatísticas descritivas dos retornos logarítmicos são apresentadas na Tabela 3.1. Os valores médios são muito próximos de zero para os três tokens, com uma ligeira média positiva para o ETH (0.000751) e o OP (0.000042) e uma média ligeiramente negativa para o MATIC (-0.000898), refletindo uma rentabilidade diária média reduzida no período analisado. Os quartis mostram que a maior parte das observações se concentra em torno de valores pequenos, mas os valores mínimos e máximos evidenciam retornos extremos: por exemplo, o ETH varia entre aproximadamente -19.7% e +19.4% num único dia, o MATIC entre cerca de -26.1% e +33.6%, e o OP entre -31.0% e +34.2%. Estes resultados são consistentes com a presença de caudas pesadas e episódios de elevada volatilidade típicos de mercados de criptoativos.

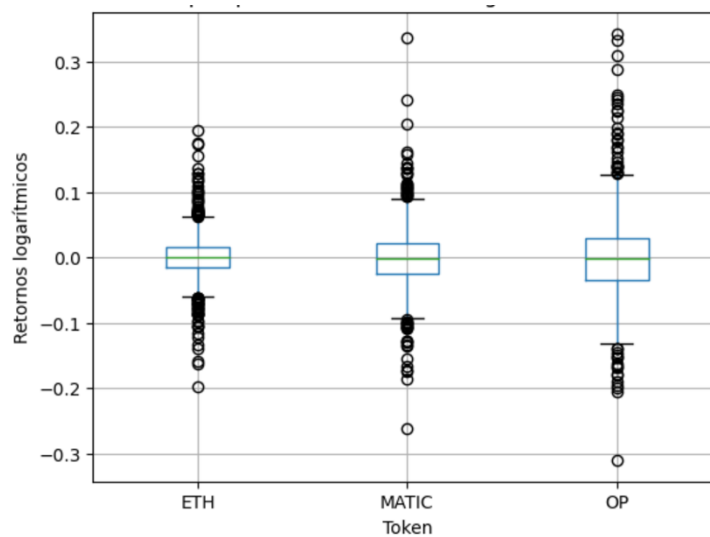
Tabela 3-1 Estatísticas descritivas dos retornos logarítmicos diários de ETH, MATIC e OP

Token	Min	1.º Quartil	Mediana	Média	3.º Quartil	Máx
ETH	-0.1975	-0.0145	0.0002	0.0008	0.0162	0.1938
MATIC	-0.2612	-0.0248	-0.0009	-0.0009	0.0217	0.3364
OP	-0.3102	-0.0346	-0.0020	0.0000 ¹	0.0297	0.3419

Fonte: Elaboração própria com base nos dados do CoinGecko®

A Figura 3-2 apresenta *box-plots* paralelos dos retornos logarítmicos diários de ETH, MATIC e OP. Esta representação gráfica permite comparar a dispersão, a posição central e a presença de valores extremos entre os três tokens, evidenciando a existência de caudas pesadas e uma volatilidade mais elevada nos tokens de Camada 2, particularmente no OP, em comparação com o ETH.

Figura 3-2 Box-plot dos retornos logarítmicos diários dos tokens



Fonte: Elaboração própria com base nos dados do CoinGecko®

A matriz de correlação entre retornos, apresentada na Tabela 3-2, revela correlações positivas de magnitude moderada a forte entre os três tokens. A correlação entre ETH e MATIC é aproximadamente 0.76, entre ETH e OP cerca de 0.72, e entre MATIC e OP cerca de 0.68, indicando uma co-movimentação relevante, embora não perfeita. Estes valores sugerem a presença de fatores sistemáticos comuns, sem, contudo, implicar uma correlação extremamente elevada (superior a 0.80), o que aponta para alguma diversificação residual entre os ativos.

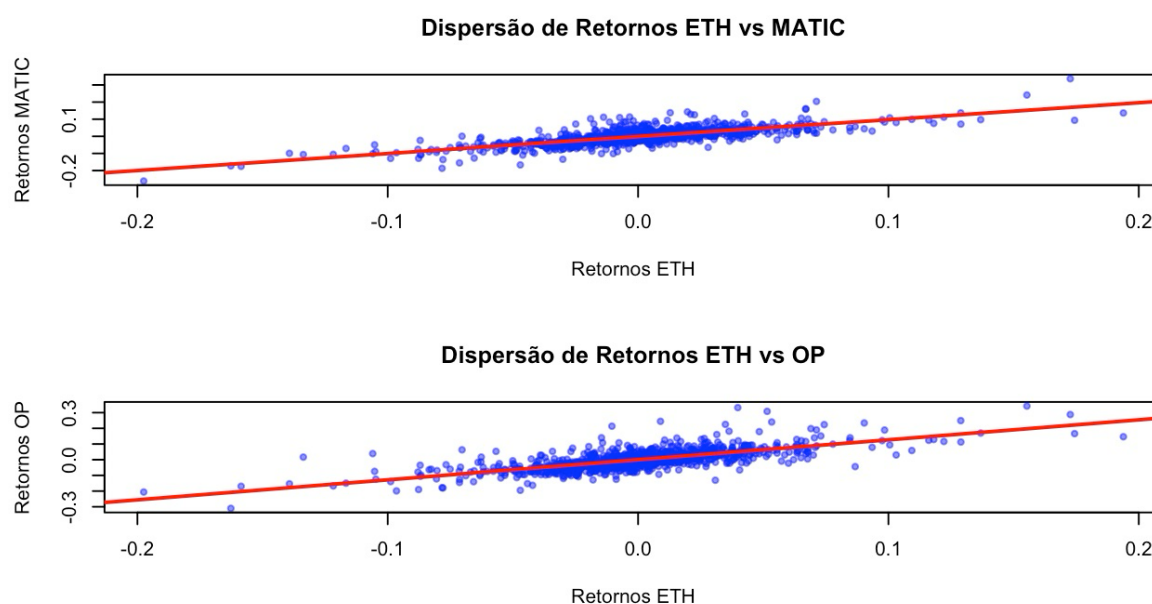
Tabela 3-2 Matriz de correlação dos retornos logarítmicos diários de ETH, MATIC e OP

	ETH	MATIC	OP
ETH	1.000	0.762	0.722
MATIC	0.762	1.000	0.685
OP	0.722	0.685	1.000

Fonte: Elaboração própria com base nos dados do CoinGecko®

Esta interdependência é também visível nos gráficos de dispersão dos retornos, apresentados na figura 3-3. Os pares ETH–MATIC e ETH–OP exibem nuvens de pontos com inclinação positiva bem definida, reforçada pelas retas de regressão ajustadas. Apesar de alguma dispersão, é evidente uma relação linear positiva entre os retornos do token de Camada 1 e os tokens de Camada 2, o que antecipa a existência de dependência contemporânea e justifica a utilização de modelos multivariados para estudar as relações dinâmicas entre estas séries.

Figura 3-3 Dispersão de retornos ETH vs MATIC e ETH vs OP

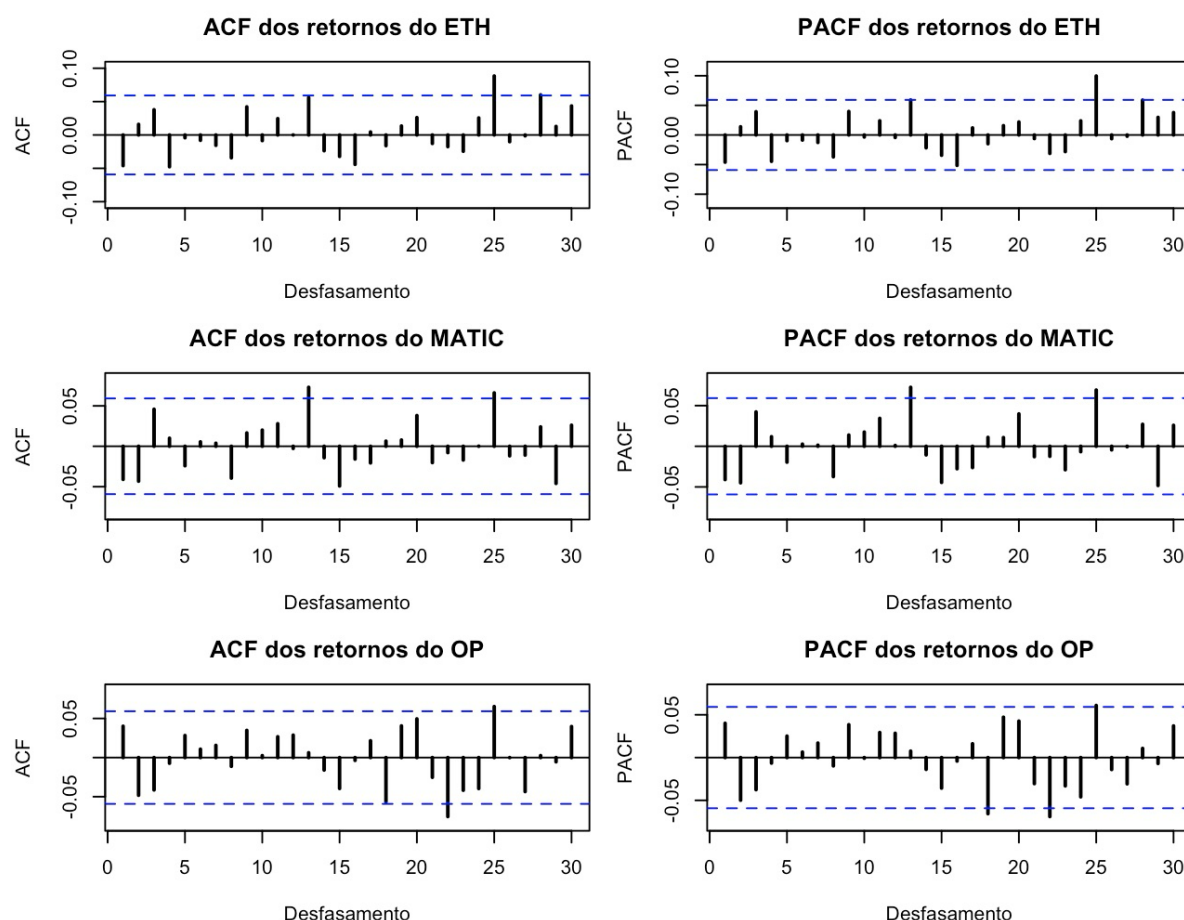


Fonte: Elaboração própria com base nos dados do CoinGecko®

Por fim, a Figura 3-4 apresenta as funções de autocorrelação (ACF) e de autocorrelação parcial (PACF) dos retornos logarítmicos de ETH, MATIC e OP para os primeiros 30 desfasamentos⁴. A exclusão do desfasamento zero da representação da ACF permite ampliar a escala vertical e tornar mais visíveis os coeficientes de autocorrelação dos restantes desfasamentos. Em termos gerais, a maioria dos coeficientes encontra-se dentro dos limites de confiança aproximados de 95% e assume valores próximos de zero, não se observando um padrão sistemático de decaimento nem uma estrutura persistente de dependência linear. Contudo, identificam-se alguns coeficientes isolados que ultrapassam os limites de confiança, tanto na ACF como na PACF, indicando a existência de dependência temporal pontual em determinados desfasamentos. Este comportamento sugere que os retornos apresentam uma dependência linear globalmente reduzida, embora não possam ser considerados estritamente não autocorrelacionados. A ausência de uma estrutura autorregressiva clara sustenta a utilização de modelos parcimoniosos para a equação da média, sem excluir a necessidade de testes formais de autocorrelação dos resíduos e de modelação da variância condicional

⁴ As funções de autocorrelação (ACF) e autocorrelação parcial (PACF) são ferramentas fundamentais na análise de séries temporais para avaliar a dependência linear entre observações separadas por diferentes desfasagens; a ACF mede a correlação total em cada desfasagem, enquanto a PACF isola a correlação direta, eliminando o efeito das desfasagens intermédias (Box et al., 2015).

Figura 3-4 ACF e PACF dos retornos de ETH, MATIC e OP. Funções de autocorrelação dos retornos logarítmicos



Fonte: Elaboração própria com base nos dados do CoinGecko®

Esta análise preliminar fornece, assim, uma primeira caracterização empírica dos dados, evidenciando: (i) não estacionariedade dos preços em níveis, (ii) presença de retornos extremos e correlações elevadas entre tokens do mesmo ecossistema, e (iii) fraca autocorrelação dos retornos, mas potencial heterocedasticidade condicional. A referência a potencial heterocedasticidade condicional baseia-se na elevada variabilidade dos retornos ao longo do tempo e na presença de episódios de volatilidade persistente observados na análise exploratória, características amplamente documentadas na literatura sobre mercados de criptoativos e que justificam a utilização de modelos GARCH para a modelação da variância condicional (Dyhrberg, 2016; Engle, 2002).

3.3. Análise de cointegração e causalidade

A **Q1** investiga se as relações entre ETH e os tokens de Camada 2 se traduzem em dependências de longo prazo ou apenas em interdependência de curto prazo. Para tal, são combinados testes de cointegração e de causalidade de Granger.

A cointegração entre os preços em níveis de ETH–MATIC e ETH–OP é avaliada com o procedimento de Johansen, através das estatísticas trace e máximo valor próprio, permitindo

identificar o número de vetores de cointegração. A ausência de cointegração indica que, apesar da ligação tecnológica L1/L2, não existe um equilíbrio de longo prazo estável entre os preços.

Em complemento, a causalidade de Granger é testada em modelos VAR estimados nos retornos. Diz-se que uma série “causa no sentido de Granger” outra série se as suas defasagens melhoram significativamente a previsão desta última. Os testes são realizados para ambas as direções (ETH $\rightarrow$ L2 e L2 $\rightarrow$ ETH), permitindo identificar padrões de influência unidirecional ou bidirecional entre o token de Camada 1 e os tokens de Camada 2.

3.4. Modelos para as relações dinâmicas de preços (VAR/VECM e ARIMA–LSTM)

Para responder à **Q2**, as relações dinâmicas de preços entre o token de Camada 1 (ETH) e os tokens de Camada 2 (MATIC e OP) são analisadas em dois planos complementares: (i) modelos lineares VAR/VECM, que captam interdependências multivariadas de curto e longo prazo, e (ii) um modelo híbrido ARIMA–LSTM, que explora componentes não lineares e a capacidade preditiva da aprendizagem profunda.

3.4.1. Modelos VAR e VECM

Os vetores de retornos $\mathbf{r}_t = (r_t^{ETH}, r_t^{L2})$ são modelados através de modelos VAR(p), em que a ordem (p) é escolhida com base em critérios de informação (AIC e BIC) e em testes de diagnóstico dos resíduos. Para as séries de preços em níveis, a existência de equilíbrios de longo prazo é avaliada com o teste de cointegração de Johansen; caso sejam detetados vetores de cointegração, a dinâmica é especificada como VECM, integrando termos de correção de erro. O VAR, nesta fase, foi estimado usando a amostra completa (ou a grande maioria dela) para entender a estrutura do mercado. Considerando uma análise estrutural, o objetivo foi calcular os coeficientes (β) para escrever as equações e fazer as Funções Impulso-Resposta. A partir destes modelos estimados foram obtidas funções impulso-resposta (IRF) e decomposições da variância dos erros de previsão (FEVD), que permitiram quantificar o impacto de choques no ETH sobre os tokens de Camada 2 em diferentes horizontes temporais.

Após a estimação dos modelos VAR, a dinâmica estrutural entre os retornos dos tokens foi analisada por meio das funções impulso–resposta e da decomposição da variância do erro de previsão. Estas ferramentas permitem complementar a interpretação dos coeficientes estimados, avaliando a propagação temporal dos choques e a importância relativa de cada variável na explicação da incerteza de previsão das restantes variáveis (Lütkepohl, 2005, 2022).

As funções impulso–resposta, ou IRF (Impulse Response Functions), medem a reação presente e futura das variáveis do sistema a um impulso aplicado a uma das inovações do modelo. Considerando a representação de médias móveis de um VAR estável,

$$\mu + \sum_{h=0}^{\infty} \Psi^h \varepsilon_{t-h}$$

a matriz (Ψh) contém as respostas das variáveis, no horizonte (h) , a choques ocorridos nas diferentes componentes do vetor de inovações. Assim, o elemento (ψ_{ij}, h) representa o efeito, após (h) períodos, de um choque associado à variável (j) sobre a variável (i) .

Como as inovações das diferentes equações podem estar correlacionadas simultaneamente, procedeu-se à sua ortogonalização por meio da decomposição de Cholesky da matriz de covariâncias dos resíduos. Esta abordagem permite interpretar separadamente os efeitos dos choques, embora os resultados possam depender da ordenação das variáveis no sistema. Atendendo à estrutura tecnológica analisada, o ETH foi colocado em primeiro lugar nos sistemas ETH–MATIC e ETH–OP, refletindo a hipótese de que choques no ativo de Camada 1 podem transmitir-se contemporaneamente aos tokens de Camada 2.

As respostas foram analisadas ao longo de um horizonte de 10 períodos. A incerteza associada às funções impulso–resposta foi avaliada através de intervalos de confiança obtidos por procedimentos de reamostragem, permitindo verificar se as respostas estimadas eram estatisticamente distintas de zero.

Complementarmente, foi utilizada a decomposição da variância do erro de previsão, ou FEVD (Forecast Error Variance Decomposition), para quantificar a proporção da variância do erro de previsão de cada variável que pode ser atribuída aos seus próprios choques e aos choques das restantes variáveis do sistema. Para a variável (i) , no horizonte (H) , esta proporção pode ser representada por

$$\frac{\sum_{h=0}^{H-1} (\mathbf{e}_i' \Psi_h \mathbf{P} \mathbf{e}_j)^2}{\sum_{h=0}^{H-1} \mathbf{e}_i' \Psi_h \Sigma_\varepsilon \Psi_h' \mathbf{e}_i}$$

em que $(\mathbf{P})$ resulta da decomposição de Cholesky de (Σ_ε) , e $(\mathbf{e}_i)$ e $(\mathbf{e}_j)$ são vetores de seleção. Os valores de $(\theta_{ij}(H))$ permitem avaliar o contributo relativo dos choques da variável (j) para o erro de previsão da variável (i) .

A análise conjunta das IRF e da FEVD foi utilizada para identificar a direção, a intensidade e a persistência da transmissão de choques entre ETH e os tokens de Camada 2. Enquanto as IRF permitem observar a trajetória temporal da resposta a um choque, a FEVD fornece uma medida quantitativa da relevância de cada choque na variabilidade futura das séries (Lütkepohl, 2005, 2022; Diebold & Yilmaz, 2023).

3.4.2. Modelo Híbrido ARIMA–LSTM com Mecanismo de Atenção

Para capturar a complexidade não-linear e a interdependência entre camadas, esta investigação propõe uma arquitetura híbrida avançada que expande a abordagem clássica de Jayawardhana & Colombage (2024). O modelo proposto combina uma componente linear (ARIMA) com uma rede neuronal de memória de longo prazo (LSTM) enriquecida com mecanismos de atenção e variáveis exógenas.

A especificação do modelo define-se pela soma de duas componentes:

$$\hat{y}_t = \hat{L}_t + \hat{N}_t$$

E representa uma decomposição conceptual da previsão no modelo híbrido, em que a componente linear $\hat{L}_t$ é estimada através de um modelo ARIMA e a componente não linear $\hat{N}_t$ corresponde à previsão da rede LSTM aplicada aos resíduos, não constituindo uma equação estrutural estimada num único passo.

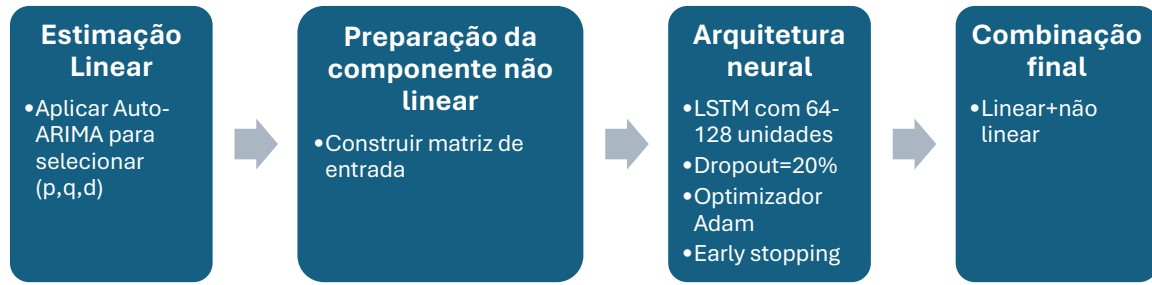
A componente Linear é realizada através de um algoritmo Auto-ARIMA que seleciona automaticamente a ordem (p, d, q) que minimiza o Critério de Informação de Akaike (AIC)⁵. Para a série do Ethereum, por exemplo, o modelo base identificado foi um ARIMA(1,0,0), capturando a reversão à média. Os resíduos resultantes, $\varepsilon_t = y_t - \hat{L}_t$, contêm a informação não explicada linearmente e servem de input para a rede neuronal.

A componente Não-Linear neuronal (LSTM Multivariada e Atenção) foi desenvolvida para superar as limitações das arquiteturas univariadas. Para os tokens de Camada 2 (MATIC e OP), o vetor de entrada X_t é multivariado, incorporando: os resíduos da própria série (ε_{t-1}); a volatilidade rolante (desvio-padrão de 7 dias); e os retornos defasados do Ethereum ($R_{ETH,t-1}$), operacionalizando a causalidade de Granger detetada.

A arquitetura inclui uma camada de *Self-Attention* após as células LSTM. Modelos AT-LSTM (attention-based LSTM) têm sido propostos para previsão de séries temporais financeiras (Kardakis et al., 2021; Vaswani et al., 2017; Zhang et al., 2019), este mecanismo permite à rede atribuir pesos dinâmicos (α_t) aos estados ocultos passados, focando a capacidade preditiva nos choques históricos mais relevantes (identificados empiricamente nas janelas de 3 a 5 dias). Para os procedimentos de treino, a rede foi configurada com uma camada oculta de 64 a 128 neurónios (*grid search*), utilizando o otimizador Adam (Adaptive Moment Estimation) com taxa de aprendizagem adaptativa (Dhake et al., 2023; Hidayat et al., 2023). Para mitigar o *overfitting* em dados financeiros ruidosos, aplicaram-se técnicas de regularização, incluindo *Dropout* (20%) e *Early Stopping* baseado na perda de validação e garantir o melhor desempenho da amostra (*out-of-sample*) (Srivastava et al., 2014). A figura 3-5 representa o procedimento e sua arquitetura.

⁵ A implementação empírica foi realizada em Python. A componente ARIMA foi estimada recorrendo à biblioteca statsmodels e à seleção automática de modelos através do algoritmo Auto-ARIMA, conforme implementado na biblioteca pmdarima, com base no critério de informação de Akaike (AIC). A componente não linear foi modelada através de redes neurais LSTM, incluindo uma extensão com mecanismo de atenção, implementadas em PyTorch. O pré-processamento dos dados e a avaliação do desempenho preditivo recorreram às bibliotecas pandas, numpy e scikit-learn.

Figura 3-5 Procedimento do Modelo Híbrido ARIMA-LSTM



Fonte: elaboração própria (ver Anexo A os passos do procedimento)

3.5. Modelos de volatilidade condicional e correlações dinâmicas (GARCH e DCC-GARCH)

A **Q3** centra-se na caracterização da volatilidade condicional e dos spillovers entre ETH, MATIC e OP. Para isso são utilizados, em primeiro lugar, modelos GARCH(1,1) univariados para cada série de retornos:

$$r_t = \mu_t + \varepsilon_t, \quad \varepsilon_t = z_t \sigma_t, \\ \sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2,$$

onde σ_t^2 é a variância condicional e $\alpha + \beta$ mede a persistência da volatilidade. A presença de efeitos ARCH é verificada por testes ARCH-LM aplicados aos resíduos dos modelos lineares.

Em seguida, para analisar a volatilidade conjunta e as correlações dinâmicas, estima-se um modelo DCC-GARCH multivariado para os pares ETH–MATIC e ETH–OP. A matriz de covariâncias condicionais H_t é decomposta como $H_t = D_t R_t D_t$, em que D_t contém as volatilidades condicionais GARCH individuais e R_t é a matriz de correlação condicional dinâmica, que evolui segundo a estrutura DCC(1,1) de Engle (2002). Esta abordagem permite estudar como a correlação e os spillovers de volatilidade entre tokens de Camada 1 e Camada 2 variam ao longo do tempo.

3.6. Estratégia de Avaliação e Métricas de Desempenho

A avaliação empírica desenvolvida assenta numa estratégia dual concebida para dar resposta aos distintos objetivos de investigação. No que respeita à análise estrutural, que inclui a caracterização das relações dinâmicas entre os ativos, bem como os testes de cointegração e de causalidade de Granger, recorreu-se à amostra completa disponível, abrangendo o período

compreendido entre 30/06/2022 e 30/06/2025. Esta opção metodológica visa maximizar a informação utilizada na estimação, permitindo identificar de forma mais robusta os coeficientes de longo prazo e a configuração estrutural subjacente ao funcionamento do mercado.

Para a avaliação preditiva *out-of-sample*, procedeu-se à definição de uma janela de teste independente correspondente aos últimos 30 dias da amostra. Todos os modelos foram reestimados exclusivamente com base nos dados anteriores a este intervalo, assegurando, assim, uma comparação da sua capacidade de generalização. Esta abordagem evita qualquer forma de enviesamento decorrente de “olhar para o futuro” (*data leakage*) e garante que o desempenho preditivo reflete condições de mercado efetivamente desconhecidas no momento da estimação.

Embora o Erro Percentual Absoluto Médio (MAPE) seja frequentemente utilizado em problemas de previsão, a sua aplicação a séries de retornos financeiros é limitada devido à presença de valores próximos de zero e mudanças de sinal, que podem tornar a métrica instável ou pouco informativa (Hyndman & Athanasopoulos, 2018c). Neste sentido, a avaliação das previsões foi quantificada utilizando duas métricas amplamente consolidadas na literatura econométrica. A Raiz do Erro Quadrático Médio (RMSE) que penaliza de forma mais acentuada erros de grande magnitude, mostrando-se particularmente sensível a desvios extremos ou episódios de elevada volatilidade:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}$$

E complementarmente, o Erro Absoluto Médio (MAE) que fornece uma medição linear da magnitude média dos erros de previsão, revelando menor sensibilidade à presença de outliers isolados:

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|$$

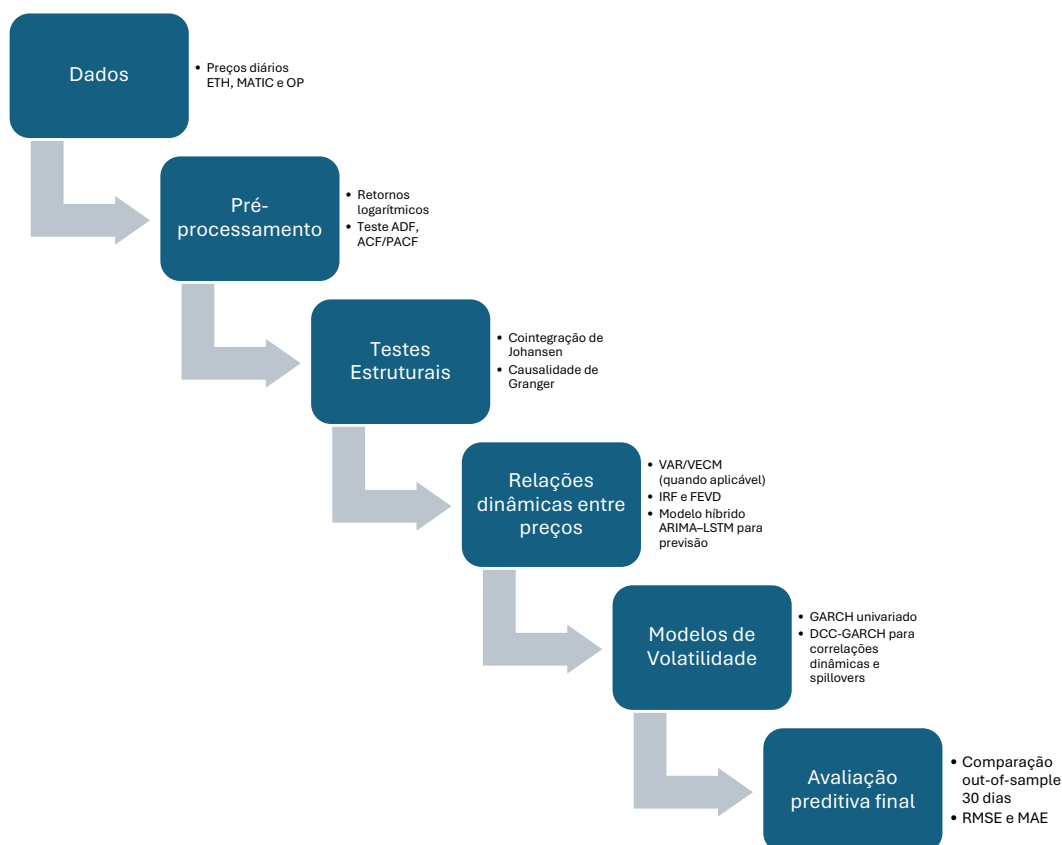
Esta distinção metodológica assegura que as conclusões relativas à estrutura de mercado, nomeadamente no domínio da cointegração, assentam em evidência empírica abrangente e robusta, enquanto as conclusões referentes à previsibilidade e ao potencial uso em contextos de trading ou investimento refletem condições realistas e são validadas com recurso a informação totalmente desconhecida pelos modelos no momento da previsão.

Os resultados destas métricas, apresentadas no Capítulo 4, permitem quantificar a melhoria relativa proporcionada pela abordagem híbrida face aos modelos puramente lineares ou puramente neuronais.

A estratégia metodológica adotada nesta investigação segue um processo estruturado e sequencial, concebido para analisar de forma abrangente as relações dinâmicas entre tokens de Camada 1 e Camada 2. A Figura 3-5 sintetiza este fluxo, desde a recolha e preparação dos dados até à modelação econométrica, análise de volatilidade e desenvolvimento do modelo híbrido de previsão. Este diagrama permite visualizar, de modo conciso, como cada etapa contribui para responder às questões de investigação: primeiro caracterizando as propriedades

estatísticas das séries, depois identificando relações estruturais e dinâmicas, e finalmente avaliando o desempenho preditivo dos modelos. Assim, o fluxo metodológico (ver figura 3-6) constitui o núcleo operacional da dissertação, articulando de forma coerente os diferentes blocos analíticos utilizados ao longo do estudo.

Figura 3-6 Fluxo da Metodologia



3.7. Implementação Computacional

A implementação computacional desta investigação recorreu a diferentes linguagens e ambientes de programação, selecionados de acordo com a adequação técnica e eficiência na execução dos procedimentos analíticos. A natureza híbrida da metodologia, combinando testes econométricos, modelos de séries temporais e redes neuronais profundas, justificou a utilização complementar de Python e R, bem como o recurso a ferramentas de desenvolvimento assistido.

As análises de cointegração, os testes de causalidade de Granger e os modelos VAR/VECM foram implementados em Python, usando o ambiente Google® Colab e Gravity, devido à maior flexibilidade desta linguagem para integração entre técnicas econométricas clássicas e métodos de aprendizagem profunda. Bibliotecas como *statsmodels*, *pandas*, *numpy* e *tensorflow/keras* permitiram estruturar um fluxo de análise coerente, especialmente no desenvolvimento do modelo híbrido ARIMA-LSTM, cuja implementação não encontra suporte igualmente robusto na linguagem R.

Para acelerar e padronizar o processo de programação, foram utilizados *prompts* iterativos no Google® Gravity, que serviram para gerar, adaptar e validar automaticamente scripts Python adequados às necessidades da análise (ver Anexo B). Esta abordagem facilitou o desenvolvimento modular do código e permitiu assegurar consistência entre as várias etapas empíricas.

Os modelos de volatilidade foram implementados exclusivamente em R, recorrendo às bibliotecas *rugarch* e *rmgarch*. A escolha desta linguagem deveu-se à maturidade, estabilidade e eficiência das suas ferramentas de estimação GARCH, amplamente validadas na literatura financeira. A aplicação de modelos DCC-GARCH, que requer procedimentos iterativos sensíveis à inicialização e à convergência, beneficiou da robustez das funções disponíveis em R.

Todos os scripts desenvolvidos, incluindo pré-processamento, testes, estimação dos modelos, validação e geração de gráficos, encontram-se disponibilizados integralmente num repositório GitHub associado a esta dissertação (<https://github.com/ricabreu/Uab>). A disponibilização integral do código segue recomendações internacionais sobre reprodutibilidade científica, transparência computacional e boas práticas em ciência de dados, assegurando que qualquer investigador possa replicar ou estender os resultados aqui apresentados.

Iniciativas como a de Peng (2011) e Nosek et al. (2015) defendem explicitamente que a reprodutibilidade é um elemento central da credibilidade científica, especialmente em investigações que dependem de computação intensiva. Do mesmo modo, Sandve et al. (2013) estabelecem princípios fundamentais, como as versões, registo de dependências, documentação clara e partilha aberta do código, que foram seguidos na estruturação do repositório GitHub desta dissertação. Assim, a transparência metodológica não só reforça a robustez dos resultados, como contribui para o avanço cumulativo da investigação no domínio dos ativos digitais.

4. ANÁLISE DOS RESULTADOS

Este capítulo apresenta os resultados empíricos decorrentes da aplicação da metodologia anteriormente definida, com o objetivo de caracterizar as relações dinâmicas, a previsibilidade e os padrões de volatilidade entre o token de Camada 1, Ethereum, e os tokens de Camada 2 selecionados, Polygon e Optimism. A análise encontra-se organizada em cinco seções complementares. Na Seção 4.1, avaliam-se a estacionariedade das séries e a existência de relações de cointegração, com vista à caracterização da estrutura de longo prazo. Na Seção 4.2, examina-se a dinâmica de curto prazo através da estimação dos modelos VAR e da análise da causalidade de Granger. A Seção 4.3 aprofunda a dinâmica estrutural dos sistemas ETH–MATIC e ETH–OP por meio das funções impulso–resposta e da decomposição da variância do erro de previsão, permitindo avaliar a direção, a intensidade e a persistência da transmissão de choques entre os ativos. Na Seção 4.4, procede-se à análise preditiva, comparando o desempenho dos modelos lineares tradicionais com o das abordagens híbridas ARIMA–LSTM numa amostra de teste independente. Por fim, a Seção 4.5 analisa a volatilidade condicional e os efeitos de contágio (*spillover*), recorrendo aos modelos GARCH e DCC-GARCH para caracterizar a evolução das volatilidades individuais e das correlações dinâmicas entre os tokens.

4.1. Estrutura de Longo prazo

Esta seção apresenta a análise estrutural das relações dinâmicas entre o Ethereum (ETH) e as duas soluções de escalabilidade de Camada 2 selecionadas: Polygon (MATIC) e Optimism (OP). A análise abrange a amostra completa de preços de fecho diários entre 30 de junho de 2022 e 30 de junho de 2025.

4.1.1. Estacionariedade e Teste de Cointegração

Preliminarmente, a estacionariedade das séries foi avaliada através do teste Augmented Dickey-Fuller (ADF). Os resultados confirmaram que, embora os preços em nível não sejam estacionários, os retornos logarítmicos de todos os tokens (ETH, MATIC e OP) são estacionários ao nível de significância de 1% ($p < 0.01$), cumprindo os pressupostos necessários para a modelação VAR (ver Anexo D).

A existência de relações de equilíbrio de longo prazo foi testada através do método de Johansen. Conforme apresentado na Tabela 4-1, as estatísticas do teste Trace para ambos os pares (ETH–MATIC e ETH–OP) são inferiores aos valores críticos a 5%, levando à não rejeição da hipótese nula de $r = 0$.

Table 4.1-1 Teste de Cointegração de Johansen

Par das Variáveis	Hipótese Nula	Estatística Trace	Valor (5%)	Crítico	Conclusão
ETH – MATIC	Nenhuma ($r = 0$)	11.8064	15.4943		Sem Cointegração

Par das Variáveis	Hipótese Nula	Estatística Trace	Valor (5%)	Crítico	Conclusão
ETH – OP	Nenhuma ($r = 0$)	10.2078	15.4943		Sem Cointegração

Fonte: Elaboração própria com base nos dados do CoinGecko

Estes resultados indicam a ausência de cointegração entre o protocolo de Camada 1 e as suas soluções de Camada 2. Este resultado sugere que, apesar da ligação tecnológica intrínseca descrita por Sguanci et al. (2021), os preços destes ativos não mantêm uma relação de equilíbrio estável no longo prazo, divergindo ao longo do tempo devido a fatores idiossincráticos e dinâmicas de mercado distintas. Neste sentido, a modelação prossegue utilizando modelos VAR nos retornos diferenciados, em detrimento de modelos VECM.

Os resultados desta secção mostram que, apesar de os retornos dos três *tokens* serem estacionários, não existe cointegração entre ETH, OP e MATIC. Assim, não se verifica qualquer forma de equilíbrio de longo prazo entre tokens de Camada 1 e Camada 2. Esta evidência sugere que a relação económica entre estes ativos não é estrutural nem persistente no tempo, respondendo negativamente à primeira questão de investigação. Tendo confirmado que não existe um equilíbrio de longo prazo entre os três *tokens*, a secção seguinte avança para a análise das relações dinâmicas de curto prazo, onde poderão emergir padrões de interação económica não detetáveis pelos testes estruturais.

4.2. Dinâmica de Curto prazo

4.2.1. Seleção e estimação dos modelos VAR

Foram estimados dois modelos VAR bivariados (ETH-MATIC e ETH-OP). Com base no Critério de Informação de Akaike (AIC), determinou-se que a estrutura de defasamento ótima para ambos os modelos é de 1 lag ($p = 1$), sugerindo que a informação de preços neste mercado é absorvida rapidamente, com os retornos do dia anterior a influenciarem significativamente a dinâmica atual (ver tabela 4-1).

Tabela 4.2-1 Critérios de informação para seleção da ordem do VAR

Ordem do VAR	AIC	BIC
VAR(0)	-20.19	-20.19
VAR(1)	-20.22	-20.15

Fonte: Elaboração própria com base nos dados do CoinGecko. Embora o BIC favoreça uma especificação sem defasagens (VAR(0)), foi adotado o modelo VAR(1) por permitir capturar interações dinâmicas entre os retornos dos tokens, em linha com o objetivo de analisar relações temporais e causalidade

Embora o critério BIC favoreça marginalmente uma especificação sem defasamentos, VAR(0), o critério AIC seleciona um modelo com um defasamento, VAR(1). Esta diferença decorre do facto de o AIC aplicar uma penalização menos severa à inclusão de parâmetros

adicionais, podendo favorecer modelos ligeiramente mais complexos quando estes melhoram o ajustamento. Atendendo a que o objetivo da análise consiste em captar as dependências dinâmicas de curto prazo entre os retornos, adotou-se a especificação VAR(1). A reduzida diferença entre os valores dos critérios recomenda, contudo, que esta escolha seja complementada pela análise da estabilidade do modelo e por testes de diagnóstico dos resíduos. Considerando os vetores de retornos y_t o vetor dos retornos considerados em cada sistema bivariado. Para o par ETH–MATIC, define-se

$$y_t = (R_{ETH,t} \ R_{MATIC,t})$$

enquanto, para o par ETH–OP,

$$y_t = (R_{ETH,t} \ R_{OP,t})$$

onde $R_{OP,t}$, $R_{ETH,t}$ e $R_{MATIC,t}$ representam, respetivamente, os retornos diários do Optimism, Ethereum e Polygon no instante t .

Em ambos os casos, a forma teórica do modelo VAR(1) pode ser representado por

$$y_t = c + A_1 y_{t-1} + \varepsilon_t$$

em que (c) é o vetor de constantes, (A_1) é a matriz dos coeficientes associados ao primeiro desfaseamento e (ε_t) é o vetor de inovações, tal que

$$\varepsilon_t \sim WN! (0, \Sigma_\varepsilon)$$

Para o sistema ETH–MATIC, define-se

$$\begin{pmatrix} R_{ETH,t} \\ R_{MATIC,t} \end{pmatrix}$$

sendo o modelo estimado dado por

$$c^{(ETH,MATIC)} + A_1^{(ETH,MATIC)} y_{t-1}^{(ETH,MATIC)} + \varepsilon_t^{(ETH,MATIC)}$$

com

$$c^{(ETH,OP)} = \begin{pmatrix} 0.0009 \\ -0.0007 \end{pmatrix}, A_1 = \begin{pmatrix} -0.101 & 0.055 \\ -0.150 & 0.047 \end{pmatrix}$$

e

$$\varepsilon_t^{(ETH,MATIC)} \sim WN! (0, \Sigma_\varepsilon^{(ETH,MATIC)})$$

Para o sistema ETH–OP, define-se

$$(R_{ETH,t} \ R_{OP,t})$$

sendo o modelo estimado dado por

$$\mathbf{c}^{(ETH,OP)} + A_1^{(ETH,OP)} \mathbf{y}_{t-1}^{(ETH,OP)} + \boldsymbol{\varepsilon}_t^{(ETH,OP)}$$

com

$$\mathbf{c}^{(ETH,OP)} = \begin{pmatrix} 0.0009 \\ 0.0003 \end{pmatrix}, A_1 = \begin{pmatrix} -0.148 & 0.080 \\ -0.220 & 0.130 \end{pmatrix}$$

e

$$\boldsymbol{\varepsilon}_t^{(ETH,OP)} \sim WN(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{(ETH,OP)})$$

No sistema bivariado $\mathbf{y}_t = (R_{ETH,t}, R_{MATIC,t})'$, os coeficientes estimados sugerem uma dinâmica assimétrica de curto prazo entre o ativo de Camada 1 (Ethereum) e o token de Camada 2 (Polygon). O coeficiente autorregressivo do ETH apresenta sinal negativo ($a_{11} = -0.101$), indicando um padrão de correção parcial após choques de curto prazo, compatível com um elevado grau de eficiência informacional do ativo base. Em contraste, o termo autorregressivo do MATIC é reduzido ($a_{22} = 0.047$), apontando para uma persistência linear limitada.

Os efeitos cruzados fornecem indícios de transmissão de informação entre camadas. Em particular, o coeficiente $a_{21} = -0.150$ sugere que variações passadas no retorno do ETH estão associadas a ajustamentos subsequentes no retorno do MATIC, refletindo um possível mecanismo de absorção de informação proveniente da Camada 1 por parte do token de Camada 2. Por outro lado, o coeficiente $a_{12} = 0.055$ indica que a informação passada do MATIC tem um impacto linear modesto sobre o retorno contemporâneo do ETH, coerente com uma hierarquia informacional em que o ativo base desempenha um papel dominante.

Resultados qualitativamente semelhantes emergem no modelo bivariado $\mathbf{y}_t = (R_{ETH,t}, R_{OP,t})'$, embora com magnitudes mais acentuadas. O coeficiente autorregressivo do ETH ($a_{11} = -0.148$) reforça a evidência de reversão à média de curto prazo, enquanto o coeficiente próprio do OP ($a_{22} = 0.130$) sugere maior persistência linear no token de Camada 2, quando comparado com o MATIC.

Os termos cruzados apontam para uma dependência direcional assimétrica. O coeficiente $a_{21} = -0.220$ indica que choques no retorno do ETH tendem a ser seguidos por ajustamentos negativos no retorno do OP, sugerindo um mecanismo de transmissão de choques da Camada 1 para a Camada 2. Em sentido inverso, o coeficiente $a_{12} = 0.080$ revela um efeito mais limitado do OP sobre o ETH, o que é consistente com uma estrutura hierárquica em que a informação relevante se propaga predominantemente do ativo base para as soluções de escalabilidade.

Importa salientar que estas interpretações se baseiam exclusivamente em dependências lineares de curto prazo captadas pelo modelo VAR e devem ser entendidas como indícios empíricos condicionais à especificação adotada, sendo aprofundadas nas análises subsequentes de causalidade de Granger e decomposição da variância do erro de previsão.

4.2.2. Causalidade de Granger

Os testes de causalidade de Granger (Tabela 4-2) revelaram diferenças estruturais profundas na forma como cada token de Camada 2 interage com o Ethereum.

Table 4.2-1 Testes de Causalidade de Granger (estatísticas Wald)

Hipótese Nula (H_0)	Estatística-F	p-valor	Conclusão
Relação ETH – MATIC			
MATIC não causa Granger ETH	2.370	0.124	Não Rejeitar H_0
ETH não causa Granger MATIC	6.132	0.013	Rejeitar H_0
Relação ETH – OP			
OP não causa Granger ETH	10.560	0.001	Rejeitar H_0
ETH não causa Granger OP	8.167	0.004	Rejeitar H_0

Fonte: Elaboração própria com base nos dados do CoinGecko.

Observa-se uma relação hierárquica unidirecional clara onde o ETH causa Granger ao MATIC, mas o inverso não se verifica ($p = 0.124$). Este resultado caracteriza o Polygon (MATIC) como um "ativo satélite", fortemente dependente da saúde do ecossistema Ethereum, mas incapaz de exercer influência significativa sobre o preço do protocolo base5. Esta dinâmica é consistente com a literatura que descreve sidechains como complementos que importam a descoberta de preço da rede principal (Sguanci et al., 2021).

Por outro lado, a relação entre Ethereum e Optimism apresenta uma causalidade bidirecional robusta (ambos os p-valores < 0.01). O Optimism não só segue o Ethereum, como também o influencia. Este padrão de *feedback* sugere uma relação simbiótica, possivelmente explicada pela natureza técnica dos *Optimistic Rollups*, que partilham a segurança da Camada 1 de forma mais direta e contribuem com receitas de sequencing para a rede base, integrando-se mais profundamente no fluxo de valor do Ethereum (Gudgeon et al., 2020; Han et al., 2025).

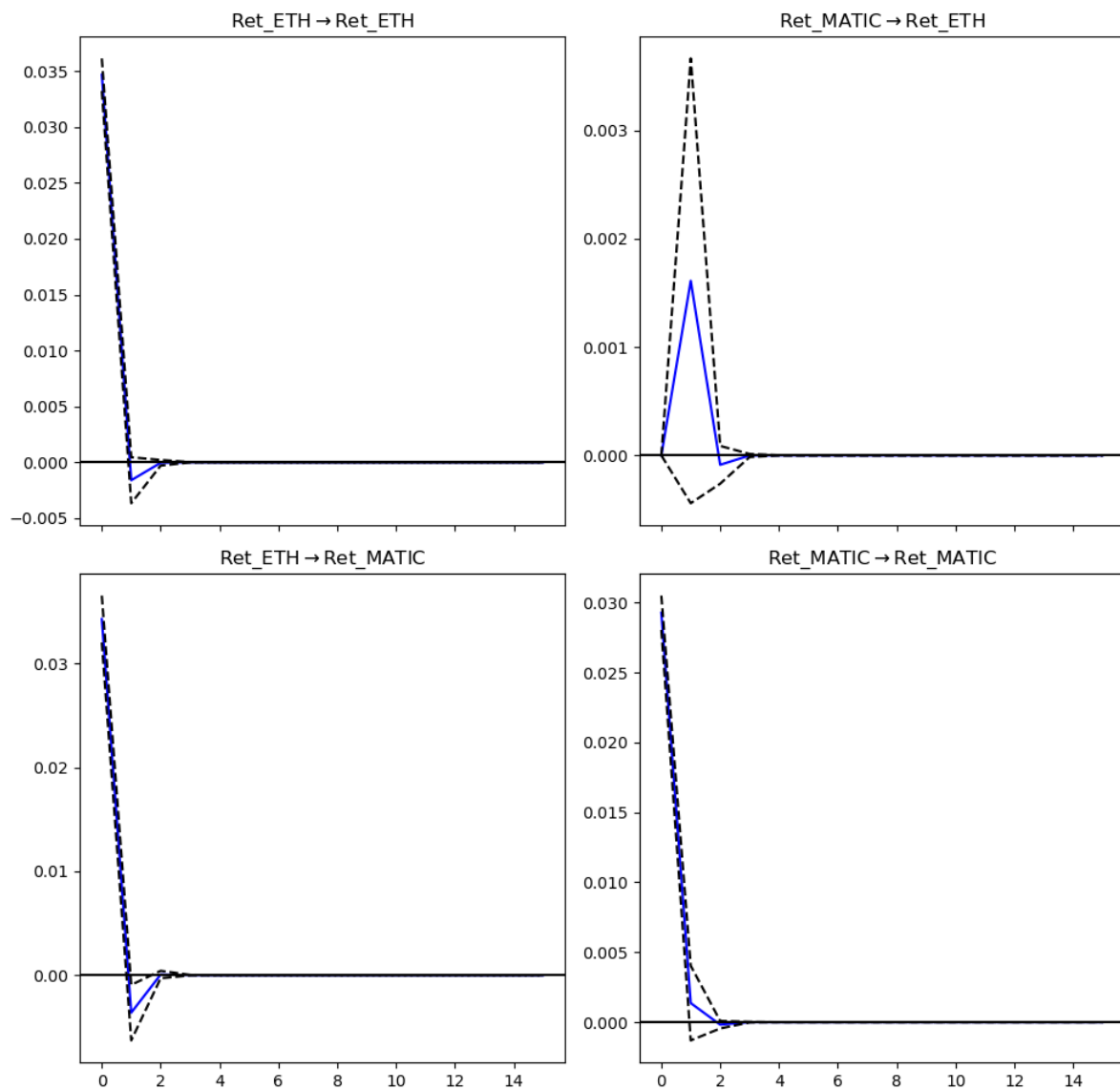
4.3. Dinâmica Estrutural (IRF e FEVD)

A análise das Funções de Impulso-Resposta (IRF) e da Decomposição da Variância (FEVD) corrobora as conclusões dos testes de causalidade, permitindo visualizar a magnitude e a persistência dos choques entre os ativos.

Análise do Par ETH-MATIC

No par ETH-MATIC, as Funções de Impulso-Resposta revelam uma reação assimétrica. Um choque nos preços do ETH gera uma resposta positiva e significativa no MATIC, que tende a persistir por 1 a 2 dias.

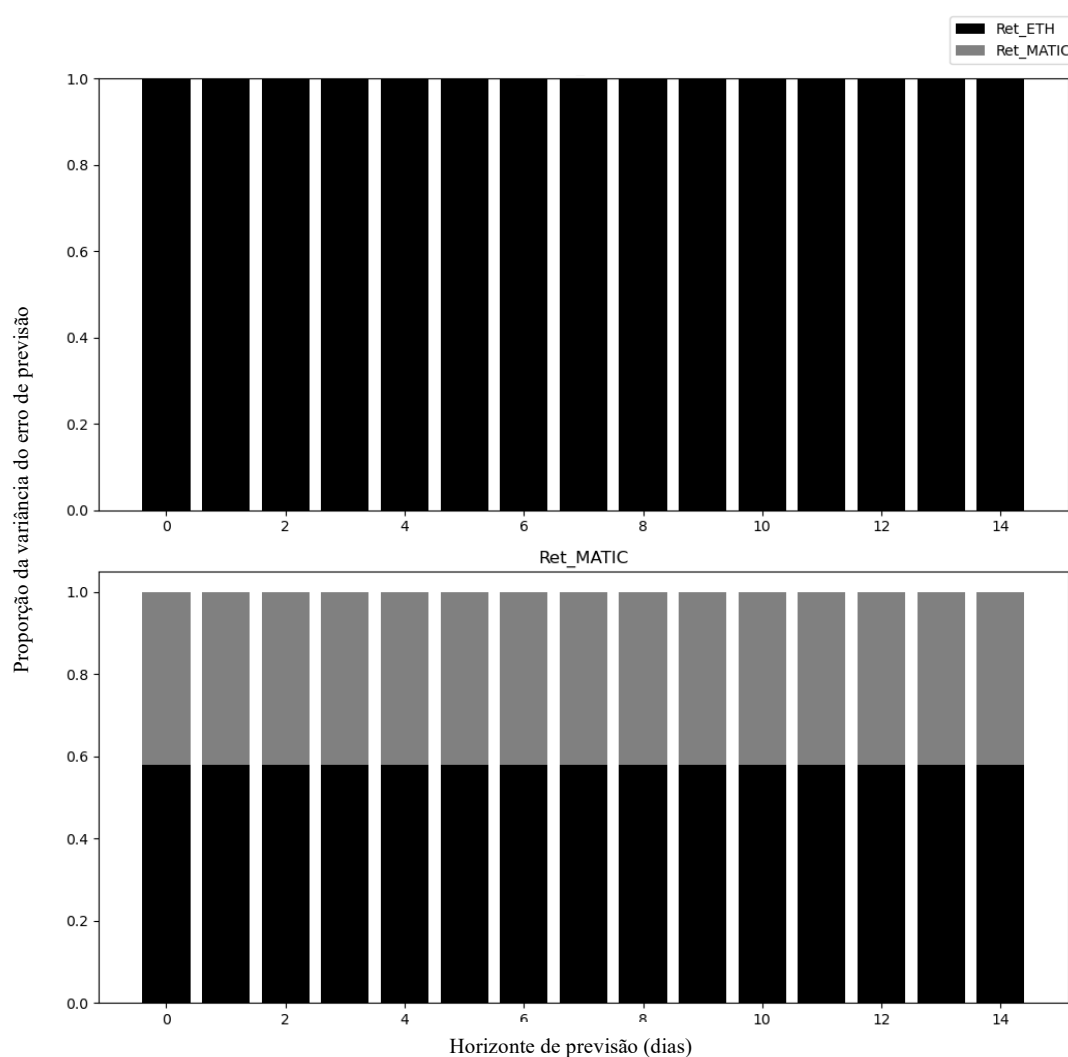
Figura 4.3-1 - Funções de Impulso-Resposta (IRF) para o sistema ETH-MATIC



Fonte: Elaboração própria com base nos dados do CoinGecko. As linhas tracejadas representam os intervalos de confiança de 95% bootstrap

A Decomposição da Variância confirma esta dependência: os choques no ETH explicam aproximadamente 58% da variância do erro de previsão do MATIC no longo prazo. Inversamente, choques no MATIC têm um impacto negligenciável no ETH (explicando perto de 0% da sua variância), o que confirma a dominância do *token* líder de mercado e a natureza de "ativo satélite" do Polygon.

Figura 4.3-2 Decomposição da Variância (FEVD) para o sistema ETH-MATIC

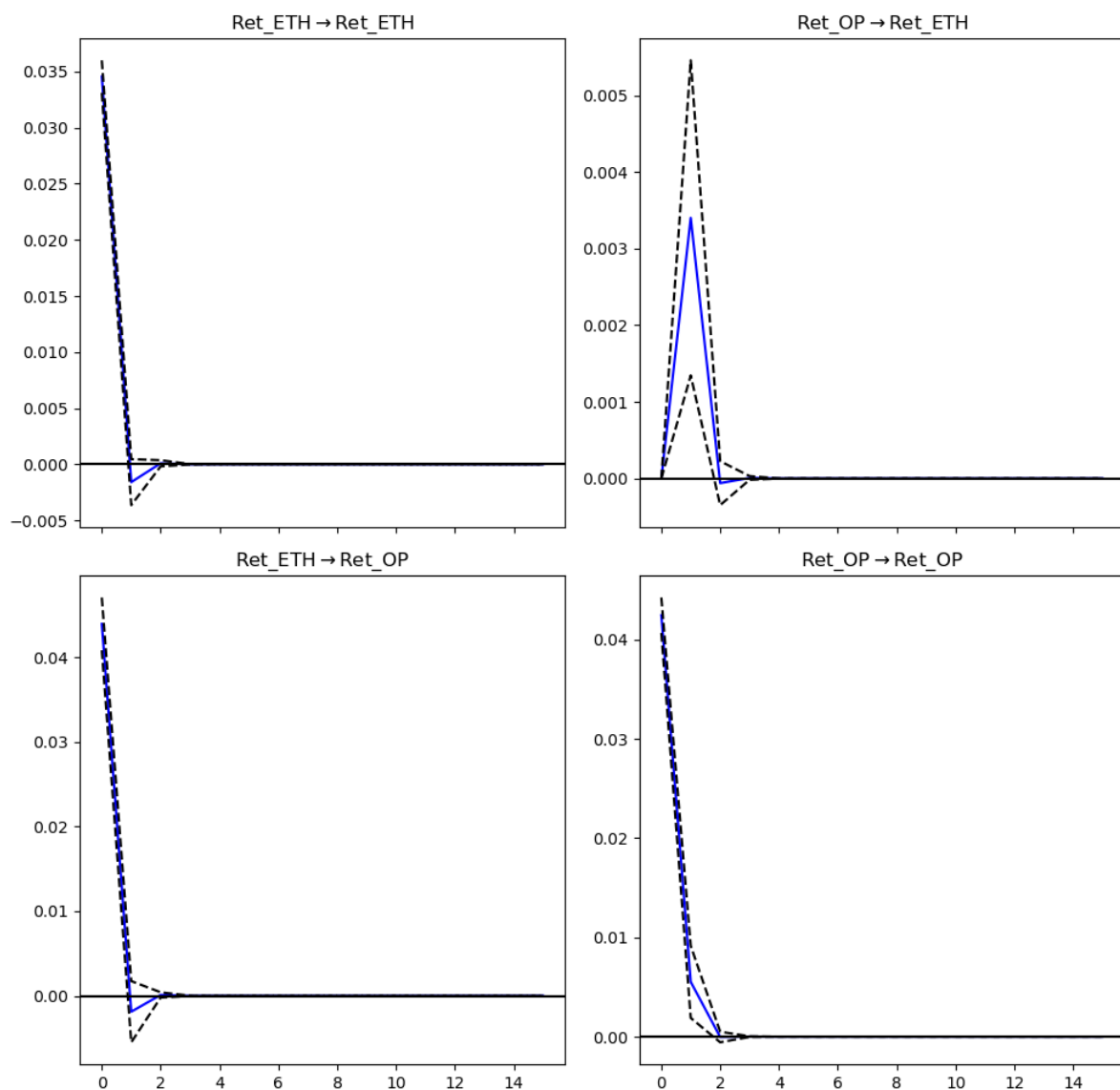


Fonte: Elaboração própria com base nos dados do CoinGecko. o eixo horizontal representa o horizonte de previsão, em dias à frente, enquanto o eixo vertical indica a proporção da variância do erro de previsão explicada por choques em cada variável. Cada barra empilhada soma a unidade, permitindo avaliar a importância relativa dos choques próprios e cruzados ao longo do tempo

Análise do Par ETH-OP

No par ETH-OP, a interdependência é visível nas IRF bidirecionais. Ao contrário do MATIC, o Ethereum reage de forma estatisticamente significativa a choques originados no Optimism.

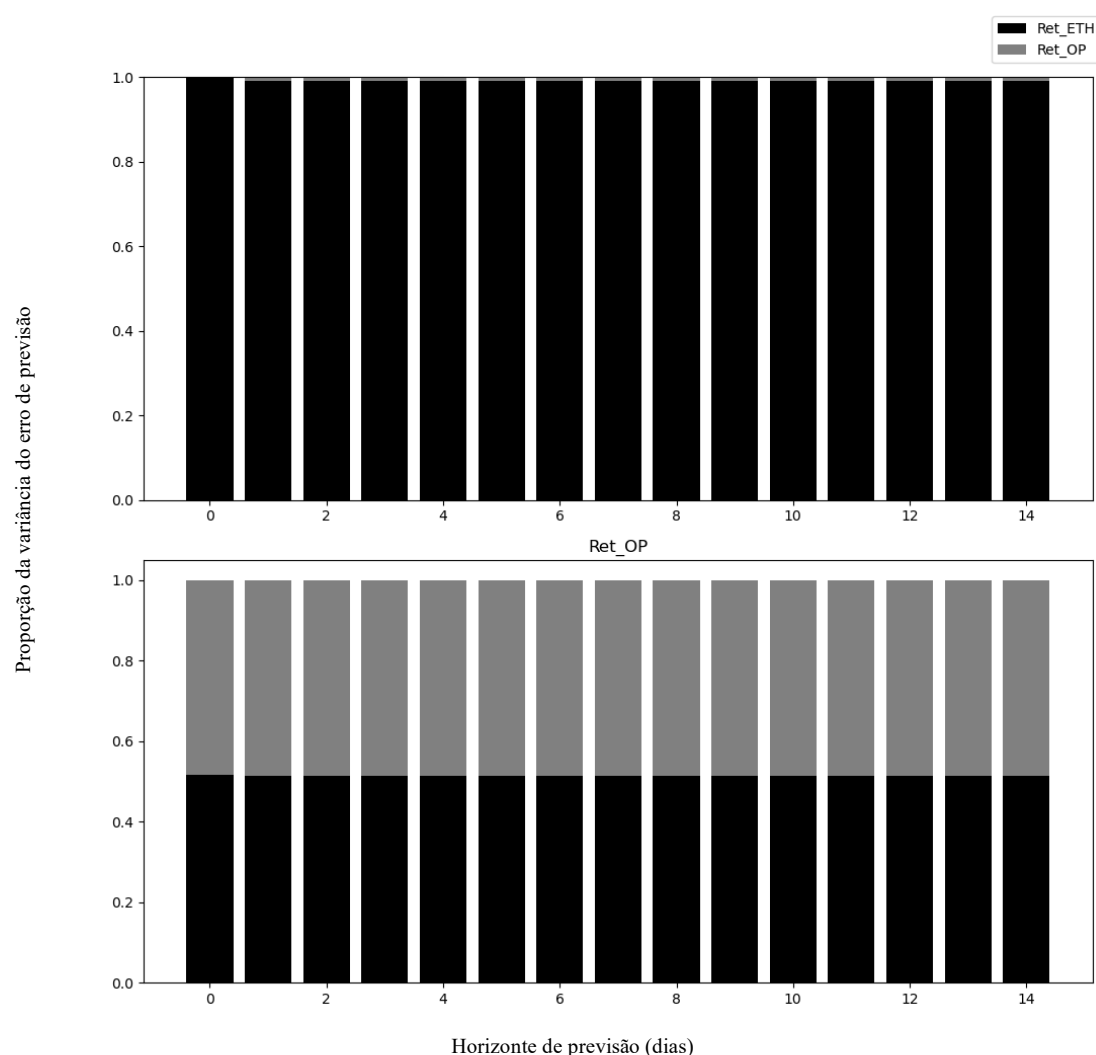
Figura 4.3-3 Análise do Par ETH-MATIC



Fonte: Elaboração própria com base nos dados do CoinGecko. As linhas tracejadas representam os intervalos de confiança de 95% bootstrap

Embora o ETH continue a ser o ativo dominante (explicando cerca de 51% da variância do OP), a Decomposição da Variância mostra que o Optimism demonstra capacidade de influenciar a formação de preço do Ethereum. Este resultado reflete a crescente importância da atividade económica transacionada nas Camadas 2 (especialmente *Rollups*) para a valorização do blockspace da Camada 1.

Figura 4.3-4 Análise do Par ETH-OP



Fonte: Elaboração própria com base nos dados do CoinGecko. o eixo horizontal representa o horizonte de previsão, em dias à frente, enquanto o eixo vertical indica a proporção da variância do erro de previsão explicada por choques em cada variável. Cada barra empilhada soma a unidade, permitindo avaliar a importância relativa dos choques próprios e cruzados ao longo do tempo.

Em suma, a análise estrutural evidenciou que, apesar da ausência de um equilíbrio de longo prazo (cointegração), existem mecanismos de transmissão de preços robustos entre as camadas. Por exemplo, o Ethereum atua como líder de mercado face ao Polygon e mantém uma relação de interdependência simbiótica com o Optimism. Estes resultados validam a existência de fluxos de informação cruciais que modelos univariados ignoram. Contudo, a natureza estritamente linear dos modelos VAR pode ser insuficiente para capturar a totalidade da dinâmica de volatilidade e os padrões complexos destes ativos. É para colmatar esta lacuna e explorar o potencial preditivo das dependências identificadas que se propõe na secção seguinte, uma abordagem Híbrida Multivariada ARIMA-LSTM.

A análise dinâmica através do VAR(1) revelou relações assimétricas entre os *tokens*, dominadas pela influência do ETH sobre OP e MATIC. As IRF e FEVD confirmam que os choques originados no Ethereum se propagam rapidamente para os ativos de Camada 2, enquanto os efeitos inversos são residuais. Estes resultados respondem à segunda questão de investigação, demonstrando que a interação económica ocorre sobretudo no curto prazo. Os

resultados dinâmicos do VAR sugerem assimetrias claras na forma como os *tokens* reagem aos retornos passados e aos choques contemporâneos. Para avaliar se estes padrões se refletem também na capacidade preditiva dos modelos, a secção seguinte prossegue com a análise de desempenho das abordagens lineares e híbridas.

4.4. Análise Preditiva

Para complementar a análise VAR e capturar potenciais padrões não-lineares nas relações dinâmicas, foi aplicado um modelo híbrido ARIMA-LSTM conforme proposto por Jayawardhana & Colombage (2024). Esta abordagem combina a capacidade dos modelos ARIMA de capturar dependências lineares com a habilidade dos modelos LSTM de identificar padrões temporais complexos de longo prazo.

A escolha de uma arquitetura híbrida multivariada é corroborada por Mudassir et al. (2024), que demonstraram que a integração de fatores exógenos (multivariado) com modelos híbridos lineares/não-lineares supera as abordagens univariadas na previsão de Bitcoin. Os autores validam especificamente o uso de variantes de LSTM com mecanismos de atenção para capturar dependências não lineares complexas que escapam aos modelos econométricos tradicionais. A introdução do mecanismo de atenção justifica-se pela necessidade de filtrar ruído e redundância inerentes a dados financeiros multivariados. Conforme detalhado por Liu et al. (2021), a arquitetura Attention-LSTM (ATT-LSTM) permite ao modelo atribuir pesos dinâmicos a diferentes instantes temporais, melhorando significativamente a capacidade de generalização face a LSTMs padrão, ao focar-se nas janelas temporais com maior poder preditivo.

Esta abordagem permite integrar duas componentes validadas pelos testes de causalidade de Granger: por um lado, a natureza multivariada e por outro a atenção dinâmica. Para os *tokens* de Camada 2 (MATIC e OP), o modelo deixa de olhar apenas para si próprio e passa a incorporar variáveis exógenas, especificamente os retornos passados do Ethereum e a volatilidade, operacionalizando assim a dependência económica detetada entre camadas. Simultaneamente, a arquitetura é reforçada com um mecanismo de *Self-Attention*, que permite à rede neuronal ponderar automaticamente a importância de diferentes choques históricos. Isto valida empiricamente as janelas temporais mais relevantes para a previsão, atribuindo pesos dinâmicos aos dados passados em vez de assumir uma influência uniforme ao longo do tempo.

4.4.1. Componente Linear (ARIMA)

Para capturar a estrutura linear e remover a autocorrelação de curto prazo da série temporal, foram ajustados modelos ARIMA. Para o Ethereum (ETH), o procedimento automático de seleção de ordem, implementado através do pacote `pmdarima`⁶, identificou o modelo ARIMA(1,0,0) como a especificação ótima segundo o critério de informação de Akaike (AIC) (ver anexo C).

Os parâmetros estimados indicam uma constante $c=0.000751$ e um coeficiente autoregressivo $\phi_1 = -0.045977$, com uma variância estimada das inovações $\sigma^2=0.001202$. Assim, a componente linear dos retornos logarítmicos diários do ETH pode ser representada por:

⁶ A estimação foi realizada utilizando o pacote `pmdarima`, que implementa o algoritmo Auto-ARIMA com pesquisa stepwise, permitindo a seleção conjunta da ordem (p,d,q) e a estimação dos parâmetros por máxima verosimilhança.

$$R_{ETH,t} = 0.000751 - 0.045977R_{ETH,t-1} + \varepsilon_t, \varepsilon_t \sim WN(0, 0.001202)$$

A reduzida magnitude do coeficiente autorregressivo sugere que a dependência linear entre o retorno atual e o retorno do dia anterior é fraca. O seu sinal negativo é compatível com uma ligeira correção de movimentos anteriores, embora a relevância estatística e a adequação do modelo devam ser avaliadas através dos resultados da estimação e dos testes de diagnóstico dos resíduos. A quantidade (0, 0.01202) corresponde à variância estimada das inovações do modelo, não sendo suficiente, por si só, para concluir que os resíduos constituem um processo de ruído branco.

Os resíduos obtidos após a estimação da componente ARIMA, que representam a parcela dos retornos não explicada pela estrutura linear selecionada, foram posteriormente utilizados como entrada da componente não linear baseada em redes LSTM.

4.4.2. Componente não-linear: Modelos híbridos ARIMA-LSTM

Diferentemente das abordagens univariadas tradicionais, a componente não linear do modelo recorre a uma rede neuronal recorrente do tipo Long Short-Term Memory (LSTM) configurada para processar um vetor de entrada multivariado, permitindo capturar dependências económicas dinâmicas entre a Camada 1 (Ethereum) e as soluções de Camada 2.

Os resíduos extraídos da componente linear, juntamente com variáveis exógenas relevantes nos modelos de Camada 2, constituem o vetor de entrada do modelo. Em cada instante temporal t , a rede recebe um vetor tridimensional dado por:

$$X_t = [\text{Residual}_t, \text{Volatility}_t, \text{ETH_Return}_t]^T, \quad X_t \in R^3$$

Onde Residual_t , corresponde ao resíduo do modelo ARIMA, Volatility_t , à volatilidade condicional estimada e ETH_Return_t , ao retorno do Ethereum enquanto variável exógena.

A arquitetura LSTM utilizada possui uma camada oculta com 64 neurónios (*Hidden Size*). A dinâmica interna da rede é governada pelas portas lógicas (*gates*), entrada e saída, bem como pelo estado de célula, formalmente descritas por:

$$\begin{aligned} f_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\ i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ \tilde{C}_t &= \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \\ C_t &= f_t \circ C_{t-1} + i_t \circ \tilde{C}_t \\ o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\ h_t &= o_t \circ \tanh(C_t) \end{aligned}$$

onde $\sigma(\cdot)$ denota a função logística e $\circ$ o produto elemento a elemento (Gers et al., 1999; Sherstinsky, 2020).

Para reforçar a capacidade do modelo em discriminar informação relevante ao longo do tempo, foi incorporado um mecanismo de atenção temporal. Este mecanismo atribui pesos $\alpha_{t,i}$ aos estados ocultos h_i , permitindo construir um vetor de contexto:

$$c_t = \sum_i \alpha_{t,i} h_i$$

onde os pesos de atenção são normalizados via função *softmax*, conforme proposto na literatura recente (Kardakis et al., 2021).

A previsão final $\hat{y}_t$ resulta de uma transformação linear do vetor de contexto (c_t), ponderado pelo mecanismo de atenção. Os parâmetros estimados para a camada de saída do modelo MATIC revelam a seguinte estrutura de pesos (W_{out}) e viés (b_{out}):

$$\hat{y}_t = W_{out} \cdot c_t + b_{out}$$

Os parâmetros da camada de saída refletem a combinação ótima aprendida pelo modelo durante o treino, não sendo interpretados individualmente, mas enquanto parte de uma representação não linear que ajusta dinamicamente a previsão dos resíduos. Esta abordagem permite capturar padrões temporais complexos e dependências não lineares que não são adequadamente modeladas por estruturas econométricas lineares tradicionais. Ao substituir pelos valores treinados (apresentando os primeiros 5 pesos significativos):

$$\hat{y}_t = [0.075, 0.105, -0.099, 0.057, -0.047 \dots] \cdot c_t + [-0.073]$$

Estes valores específicos refletem a calibração fina do modelo, onde o viés negativo (-0.073) sugere um ligeiro ajuste base na previsão dos resíduos, modulado dinamicamente pelos pesos da rede neuronal.

4.4.3. Análise comparativa do desempenho preditivo

Nesta secção procede-se à avaliação da capacidade preditiva dos diferentes modelos considerados na estimação dos retornos logarítmicos dos *tokens* analisados. O objetivo central consiste em comparar, de forma direta e rigorosa, o desempenho das abordagens lineares tradicionais (VAR e ARIMA) com o das arquiteturas híbridas propostas (modelos Univariado e Multivariado com Mecanismos de Atenção). Para tal, recorrem-se às métricas de erro mais comuns na literatura, a Raiz do Erro Quadrático Médio (RMSE) e o Erro Absoluto Médio (MAE), permitindo uma análise quantitativa consistente entre metodologias.

Metodologicamente, importa salientar que os resultados apresentados nesta secção se distinguem daqueles incluídos na análise estrutural da Secção 4.1. Enquanto os modelos VAR previamente discutidos foram estimados com a amostra completa, visando caracterizar as relações dinâmicas de longo prazo entre as variáveis, as métricas aqui reportadas referem-se exclusivamente ao desempenho preditivo *out-of-sample*, obtido numa janela de teste independente de 30 dias.

A reestimação dos modelos lineares (VAR e Auto-ARIMA) especificamente para esta janela de teste constitui um passo metodologicamente necessário para assegurar uma comparação equitativa com os modelos híbridos baseados em *Deep Learning*. Desta forma, garante-se que todas as abordagens enfrentam exatamente as mesmas condições de mercado e regimes de volatilidade, permitindo isolar e quantificar o contributo efetivo das componentes não lineares e dos mecanismos de atenção para a melhoria da precisão preditiva.

4.4.4. Resultados Empíricos

As Tabelas 4-5 e 4-6 apresentam a comparação do desempenho preditivo *out-of-sample* de quatro especificações: o modelo VAR, o modelo Auto-ARIMA, o modelo híbrido ARIMA–

LSTM univariado e o modelo híbrido ARIMA–LSTM multivariado com mecanismo de atenção. A avaliação foi realizada numa janela de teste independente de 30 dias, utilizando como métricas a raiz do erro quadrático médio (RMSE) e o erro absoluto médio (MAE). No sistema ETH–MATIC, reporta-se o RMSE para os retornos de ETH e MATIC e o MAE para os retornos de MATIC. No sistema ETH–OP, apresentam-se o RMSE para os retornos de ETH e OP e o MAE para os retornos de OP

Figura 4.4-1 Desempenho preditivo out-of-sample dos modelos para o sistema ETH-MATIC

Modelo	RMSE dos retornos ETH	RMSE dos retornos MATIC	MAE dos retornos MATIC
VAR	0.03265	0.03509	0.02819
Auto-ARIMA	0.03259	0.03472	0.02789
Híbrido Univariado	0.03260	0.03443	0.02759
Híbrido Multivariado (Atenção)	0.03260	0.03436	0.02751

Fonte: Elaboração própria com base nos dados do CoinGecko®. As métricas foram calculadas sobre os retornos logarítmicos previstos numa amostra de teste independente de 30 dias. Valores inferiores de RMSE e MAE indicam melhor desempenho preditivo.

Figura 4.4-2 Desempenho preditivo out-of-sample dos modelos para o sistema ETH-OP

Modelo	RMSE dos retornos ETH	RMSE dos retornos OP	MAE dos retornos OP
VAR	0.03265	0.05253	0.04143
Auto-ARIMA	0.03259	0.05210	0.04101
Híbrido Univariado	0.03260	0.05234	0.04068
Híbrido Multivariado (Atenção)	0.03260	0.05201	0.04055

Fonte: Elaboração própria com base nos dados do CoinGecko®. As métricas foram calculadas sobre os retornos logarítmicos previstos numa amostra de teste independente de 30 dias. Valores inferiores de RMSE e MAE indicam melhor desempenho preditivo

A análise dos resultados evidencia uma dicotomia estrutural consistente entre o ativo de Camada 1 e os ativos de Camada 2, sugerindo diferentes graus de eficiência de mercado. No caso do Ethereum (ETH), observa-se uma prevalência clara da eficiência na sua forma fraca. O modelo linear selecionado automaticamente pelo algoritmo Auto-ARIMA, correspondente a um ARIMA(1,0,0), registou o menor erro (RMSE = 0.03259), sendo que o modelo híbrido apresentou um desempenho estatisticamente equivalente. A ausência de ganhos preditivos relevantes por parte da LSTM indica que os retornos do ETH exibem fraca dependência temporal e reduzido conteúdo não linear explorável, característica consistente com um

processo próximo de passeio aleatório (*Random Walk*). Neste contexto, a previsibilidade incremental fornecida por modelos não lineares tende a ser limitada, uma vez que a informação passada não contribui significativamente para reduzir a incerteza futura. Este resultado não constitui uma limitação metodológica, mas antes uma evidência empírica de elevada eficiência informacional do mercado do ETH, convergindo com a literatura recente, que indica que o aumento da capitalização de mercado tende a reduzir a previsibilidade incremental proporcionada por modelos de *Deep Learning*, aproximando o seu desempenho ao de modelos econométricos clássicos (Alvi et al., 2025; B. B. Gupta et al., 2025).

Por contraste, a análise do Polygon (MATIC) revela uma dinâmica distinta e mais compatível com a presença de não-linearidades exploráveis. O Modelo Híbrido Multivariado evidenciou uma vantagem expressiva ao reduzir o RMSE para 0.03436, comparativamente aos 0.03509 obtidos pelo modelo VAR. Este comportamento confirma a capacidade das arquiteturas baseadas em LSTM para capturar padrões idiossincráticos e não-lineares característicos de *tokens* de menor capitalização e de *sidechains* relativamente independentes. A introdução de um mecanismo de atenção mostrou-se particularmente relevante, uma vez que permitiu ao modelo discriminar entre ruído e informação economicamente significativa, concentrando-se nos choques históricos mais influentes (Liu et al., 2021).

O caso do Optimism (OP) constitui o resultado mais revelador no contexto desta investigação. Embora o modelo híbrido univariado não tenha conseguido superar o desempenho do ARIMA (RMSE de 0.05234 face a 0.05210), a versão multivariada alcançou o melhor desempenho global (RMSE = 0.05201). Este resultado fornece evidência empírica da existência de uma dependência económica estrutural entre os ativos estudados, demonstrando que a previsão robusta do preço do OP requer a integração dos retornos do ETH como variável exógena. Tal constatação está em consonância com as conclusões de Mudassir et al. (2024), que defendem que arquiteturas híbridas beneficiam significativamente da incorporação de fatores externos em ecossistemas fortemente interligados.

Em síntese, os resultados sugerem que, embora os modelos lineares permaneçam adequados e robustos para a previsão de ativos pertencentes ao protocolo base (Camada 1), como o Ethereum, a obtenção de previsões mais precisas para soluções de escalabilidade (Camada 2) exige abordagens híbridas capazes de integrar mecanismos de atenção e captar interdependências hierárquicas entre preços. Esta diferenciação metodológica evidencia a natureza heterogénea dos mercados criptoativos e sublinha a importância de modelos flexíveis para captar as suas distintas estruturas informacionais.

Os modelos híbridos ARIMA–LSTM revelaram ganhos preditivos relevantes sobretudo para ativos de Camada 2, como o MATIC e o OP, evidenciando que a componente não linear capta padrões adicionais associados a estruturas de mercado menos eficientes e mais dependentes de interações interativas. Em contraste, o ETH mantém um comportamento mais próximo de um passeio aleatório, limitando o valor acrescentado de arquiteturas preditivas complexas. Estas conclusões reforçam a interpretação da segunda questão de investigação, ao mostrar que a dinâmica de curto prazo é complexa e heterogénea entre *tokens*. Embora os modelos preditivos revelem diferenças relevantes entre os *tokens* na sua dinâmica de curto prazo, a previsibilidade dos retornos não esclarece a intensidade da transmissão de risco entre eles. Neste sentido a secção seguinte, contempla a análise de volatilidade, onde são examinados os mecanismos de *spillover* e correlação dinâmica.

4.5. Volatilidade e Contágio

A análise da volatilidade constitui um elemento central nos mercados de criptoativos (Gupta & Chaudhary, 2022), caracterizados por fortes episódios de heterocedasticidade condicional, clustering de volatilidade e correlações que se intensificam em momentos de stress de mercado. Após avaliar as relações dinâmicas e preditivas entre os *tokens* do ecossistema Ethereum, este segmento aprofunda a forma como o risco é distribuído e transmitido entre o *token* de Camada 1 (ETH) e os *tokens* de Camada 2 (MATIC e OP). Para tal, recorre-se a modelos GARCH univariados e ao modelo DCC-GARCH, amplamente reconhecido pela sua capacidade em capturar dependências temporais na volatilidade e na correlação condicional.

4.5.1. Testes ARCH-LM e Fundamentação da Modelação GARCH

A presença de heterocedasticidade condicional foi avaliada através do teste ARCH-LM, aplicado separadamente aos resíduos de cada equação dos modelos VAR(1) ETH–MATIC e ETH–OP. Em todos os testes foram considerados 10 desfasamentos dos resíduos quadráticos, pelo que a estatística LM segue, sob a hipótese nula, uma distribuição qui-quadrado com 10 graus de liberdade. Os resultados apresentados na Tabela 4-2 conduzem à rejeição da hipótese nula de ausência de efeitos ARCH ao nível de significância de 1% em todas as equações. Verifica-se, portanto, dependência temporal nos quadrados dos resíduos, evidenciando que a variância condicional não permanece constante ao longo do tempo. Estes resultados fundamentam a utilização posterior de modelos GARCH para representar a dinâmica da volatilidade de ETH, MATIC e OP.

Tabela 4.5-1 Testes ARCH-LM aplicados aos resíduos dos modelos VAR(1)

Resíduos testados dos <i>tokens</i>	Estatística χ^2	Graus de Liberdade	P-Valor	Resultado
Equação do ETH no VAR ETH–MATIC	87.371	10	< 0.001 ***	Efeitos ARCH Presentes
Equação do MATIC no VAR ETH–MATIC	146.807	10	< 0.001 ***	Efeitos ARCH Presentes
Equação do ETH no VAR ETH–OP	81.748	10	< 0.001 ***	Efeitos ARCH Presentes
Equação do OP no VAR ETH–OP	89.215	10	< 0.001 ***	Efeitos ARCH Presentes

Fonte: Elaboração própria com base nos dados do CoinGecko®. A hipótese nula do teste ARCH-LM estabelece a ausência de autocorrelação nos quadrados dos resíduos até ao décimo desfasamento. O teste foi aplicado separadamente aos resíduos de cada equação dos modelos VAR(1). Sob a hipótese nula, a estatística LM segue assintoticamente uma distribuição qui-quadrado com 10 graus de liberdade *, ** e *** indicam significância estatística aos níveis de 10%, 5% e 1%, respetivamente, com base em testes t assintóticos.

4.5.2. Dinâmica de Volatilidade ETH-MATIC

Os parâmetros estimados do modelo DCC-GARCH para o par ETH–MATIC encontram-se detalhados nas Tabelas 4-3 e 4-4, evidenciando padrões de persistência elevados no Ethereum e uma dinâmica mais reativa no MATIC

Tabela 4.5-2 Parâmetros do Modelo DCC-GARCH (ETH-MATIC)

Parâmetro	ETH	MATIC	Significância
ω (omega)	0.000011	0.000288	**
α_1 (alpha1)	0.0306	0.1690	***
β_1 (beta1)	0.9603	0.6894	***
Persistência ($\alpha_1 + \beta_1$)	0.9909	0.8584	-

Fonte: Elaboração própria com base nos dados do CoinGecko. *, ** e *** indicam significância estatística aos níveis de 10%, 5% e 1%, respetivamente, com base em testes t assintóticos

Tabela 4.5-3 Parâmetros DCC

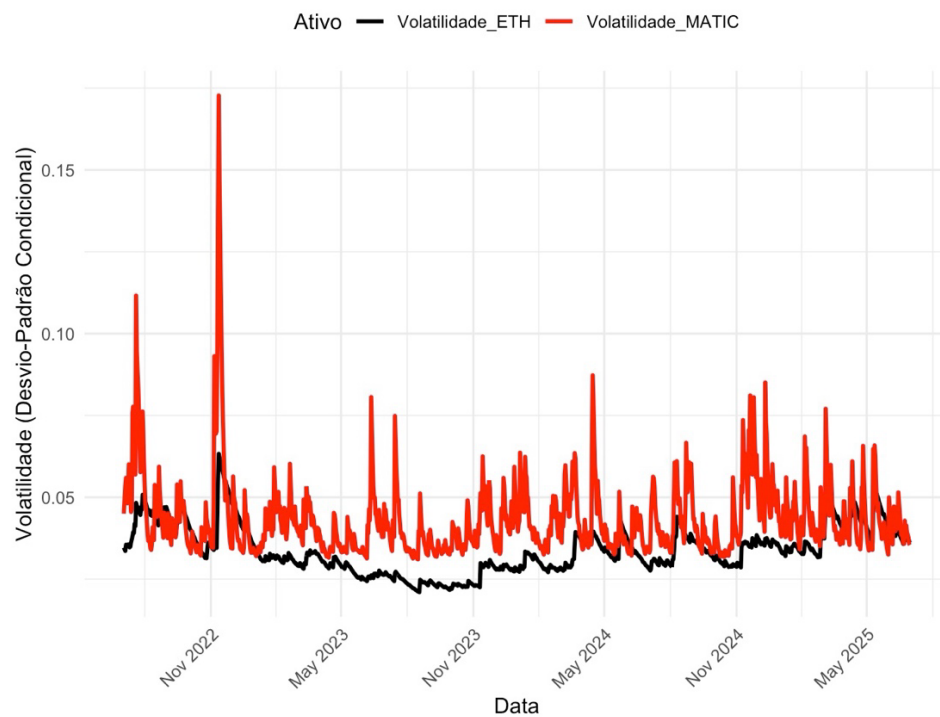
Parâmetro	Valor	Erro Padrão	P-Valor
dcca1	0.0554	0.0152	0.0003 ***
dccb1	0.8717	0.0326	0.0000 ***

Fonte: Elaboração própria com base nos dados do CoinGecko. *, ** e *** indicam significância estatística aos níveis de 10%, 5% e 1%, respetivamente, com base em testes t assintóticos

A elevada soma dos parâmetros $\alpha_1 + \beta_1$ para o ETH (0.9909) sugere um comportamento praticamente integrado, típico de séries financeiras de elevada frequência, em que os efeitos de choques na volatilidade se dissipam lentamente. Em contraste, o MATIC apresenta maior sensibilidade a choques recentes, refletindo a sua maior exposição a fluxos especulativos e menor capitalização de mercado, fenómeno igualmente reportado por Zhang (2019).

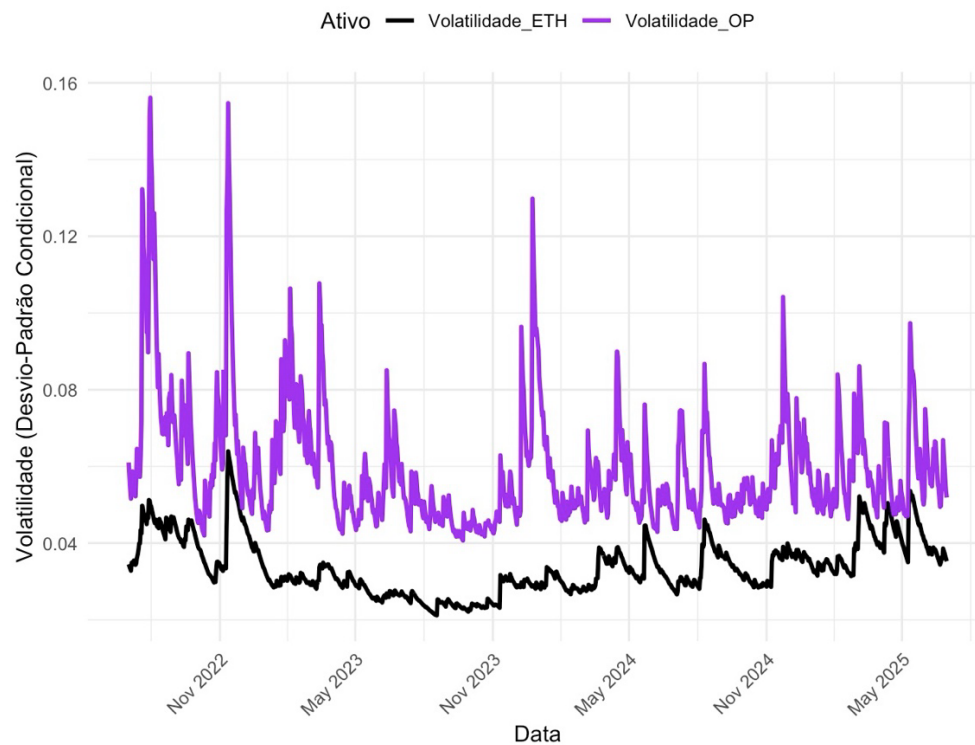
As volatilidades condicionais individuais estão representadas na Figura 4-7 (ETH–MATIC) e Figura 4-8 (ETH–OP), revelando episódios de volatilidade concentrada que coincidem temporalmente com eventos de mercado amplos.

Figura 4.5-1 Volatilidades Condicionais Individuais do ETH e MATIC



Fonte: Elaboração própria com base nos dados do CoinGecko

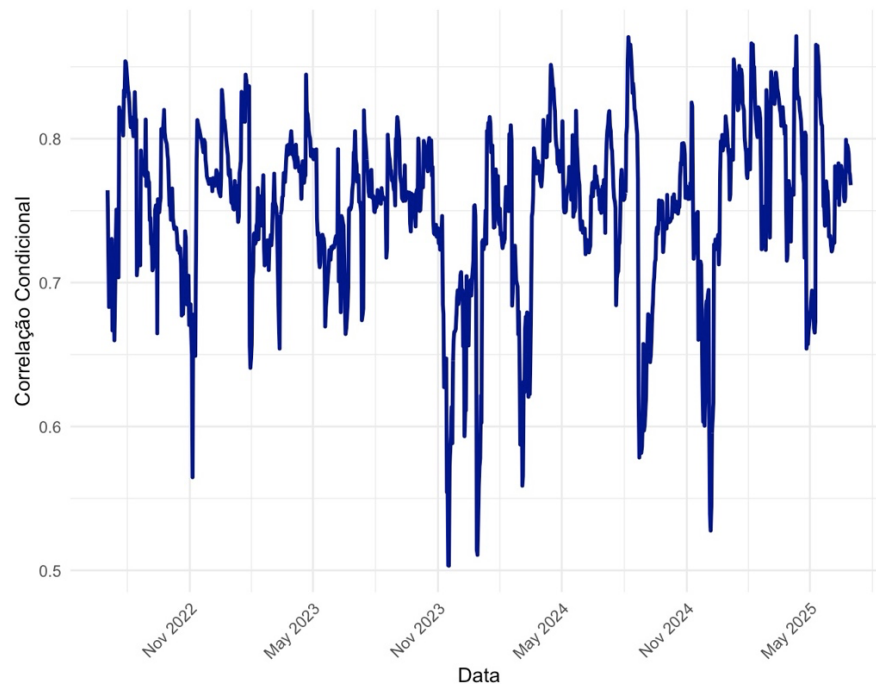
Figura 4.5-2 Volatilidades Condicionais Individuais do ETH e OP



Fonte: Elaboração própria com base nos dados do CoinGecko

Já a evolução da correlação condicional dinâmica encontra-se ilustrada na Figura 4-9, onde se verifica que a correlação entre ETH e MATIC se intensifica em períodos de maior incerteza. Este comportamento sugere a existência de efeitos de *spillover*, consistentes com a hipótese de que *tokens* de Camada 2 absorvem parte da pressão informacional originada no protocolo base (ETH).

Figura 4.5-3 Correlação Condicional Dinâmica (DCC) entre ETH e MATIC



Fonte: Elaboração própria com base nos dados do CoinGecko

4.5.3. Dinâmica de Volatilidade ETH-OP

De forma análoga, os resultados para o par ETH-OP, apresentados na Tabela 4-5 e 4-6, mostram padrões de volatilidade semelhantes, embora com algumas diferenças estruturais.

Tabela 4.5-4 Parâmetros do Modelo DCC-GARCH (ETH-OP)

Parâmetro	ETH	OP	Significância
ω (omega)	0.000015	0.000299	- / -
α_1 (alpha1)	0.0353	0.1231	* / **
β_1 (beta1)	0.9521	0.8003	*** / ***
Persistência ($\alpha_1 + \beta_1$)	0.9874	0.9234	-

Fonte: Elaboração própria com base nos dados do CoinGecko. *, ** e *** indicam significância estatística aos níveis de 10%, 5% e 1%, respetivamente, com base em testes t assintóticos

Tabela 4.5-5 Parâmetros DCC

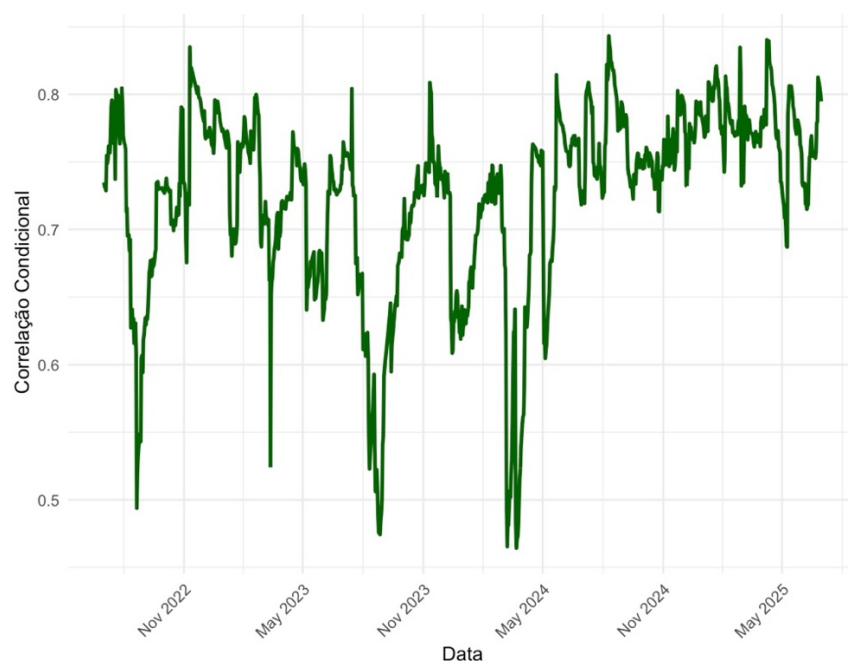
Parâmetro	Valor	Erro Padrão	P-Valor
dcca1	0.0326	0.0141	0.0213 **
dccb1	0.9297	0.0235	0.0000 ***

Fonte: Elaboração própria com base nos dados do CoinGecko. *, ** e *** indicam significância estatística aos níveis de 10%, 5% e 1%, respectivamente, com base em testes t assintóticos

O OP exibe uma persistência elevada ($\alpha_1 + \beta_1 = 0.9234$), mas inferior à observada para o ETH, sugerindo que choques na sua volatilidade têm menor duração. A comparação dos parâmetros DCC revela ainda uma estrutura de correlação condicional mais estável ao longo do tempo no par ETH–OP (soma dos parâmetros DCC = 0.9623), quando comparado com ETH–MATIC (0.9271). Este diferencial pode refletir diferenças na arquitetura das soluções de escalabilidade: enquanto o Optimism funciona como um rollup otimista, estreitamente acoplado à segurança do Ethereum, o Polygon opera como uma *sidechain* com maior autonomia (Gudgeon et al., 2020), resultando em respostas distintas a choques sistêmicos.

A Figura 4-10 apresenta a correlação condicional dinâmica entre ETH e OP, evidenciando períodos de maior sincronização durante fases de volatilidade extrema.

Figura 4.5-4 Correlação Condicional Dinâmica (DCC) entre ETH e OP



Fonte: Elaboração própria com base nos dados do CoinGecko

Os resultados dos modelos DCC-GARCH indicam que o Ethereum atua como o principal emissor de volatilidade, condicionando o comportamento dos *tokens* de Camada 2. O MATIC revela maior sensibilidade a choques, enquanto o OP apresenta correlações mais estáveis com o ETH, refletindo uma dependência estrutural mais estreita. Verifica-se ainda que períodos de stress aumentam significativamente as correlações condicionais, evidenciando contágio financeiro dentro do ecossistema. No conjunto, estes padrões confirmam a terceira questão de investigação, mostrando que a relação entre Camada 1 e Camada 2 se expressa sobretudo na transmissão de risco e não numa convergência de preços.

5. DISCUSSÃO DOS RESULTADOS

Este capítulo reúne e interpreta os principais resultados do estudo, articulando as evidências estruturais, preditivas e de volatilidade com o objetivo de responder às questões de investigação. A análise conjunta destas três dimensões permite caracterizar o ecossistema Ethereum como heterogêneo, evidenciando comportamentos significativamente distintos entre o protocolo base (Ethereum) e as suas soluções de escalabilidade de Camada 2 (Polygon e Optimism). Estes resultados indicam que a dinâmica económica e informacional entre as Camadas 1 e 2 não é homogênea, refletindo diferentes graus de eficiência de mercado, interdependência e previsibilidade.

Uma das contribuições mais significativas deste estudo reside na identificação de diferentes regimes de eficiência entre camadas, ou seja, neste contexto, a não existência de regimes estatísticos formais, mas sim diferentes graus de eficiência informacional dos mercados, inferidos a partir da capacidade (ou incapacidade) dos modelos em extrair previsibilidade a partir da informação passada. Os resultados da análise preditiva (Secção 4.2) demonstraram que, para o Ethereum (ETH), os modelos lineares simples (ARIMA)⁷ são tão eficazes quanto as arquiteturas complexas de *Deep Learning*. O coeficiente autoregressivo próximo de zero e a incapacidade do modelo híbrido em gerar ganhos significativos sugerem que o mercado de ETH opera num regime de Eficiência Fraca⁸. Como ativo dominante e de elevada liquidez, o Ethereum incorpora a informação passada de forma quase instantânea, comportando-se próximo de um Passeio Aleatório (*Random Walk*), limitando as oportunidades de arbitragem baseadas em padrões históricos.

Em contrapartida, os mercados de Camada 2 (MATIC e OP) exibem ineficiências exploráveis. A superioridade estatística do Modelo Híbrido Multivariado com Mecanismo de Atenção nestes ativos sugere a existência de padrões não lineares próprios de cada ativo, que escapam à capacidade explicativa dos modelos econométricos tradicionais. Para o Polygon, isto reflete a sua natureza de *sidechain* com dinâmica própria, no caso do Optimism, reflete a complexidade da sua interação com a rede principal. Esta dicotomia sugere que o "alfa" (retorno excedentário financeiro) em estratégias algorítmicas⁹ é mais facilmente extraível nos *tokens* satélites do que no ativo nuclear do ecossistema, corroborando a visão de que modelos não-lineares, como as LSTM, são particularmente eficazes em capturar a volatilidade e interdependências complexas destes ativos.

A ausência de cointegração (Secção 4.1) desafia a premissa intuitiva de que ativos tecnologicamente acoplados deveriam convergir para um equilíbrio de preços de longo prazo, como sugerido por Sguanci et al. (2021) ao descreverem as Camadas 2 como extensões da Camada 1. No entanto, a análise de causalidade e o desempenho dos modelos preditivos esclarecem a natureza desta relação, na qual ligação não é de convergência, mas de dependência informacional.

⁷ Através do algoritmo Auto-ARIMA, referido anteriormente

⁸ Um mercado tem Eficiência Fraca é um conceito central da Hipótese dos Mercados Eficientes de Eugene Fama (Nóbel da Economia em 1970), que significa, quando o preço atual de um ativo já reflete toda a informação contida no histórico de preços passados e volumes de negociação.

⁹ Em trading algorítmico, procurar "Alfa" significa procurar ineficiências de mercado. Por exemplo, antecipar o preço de amanhã com mais precisão do que a média do mercado (e.g. quando se aplica o LSTM)

O caso do Optimism (OP) é paradigmático. A análise preditiva revelou que modelos univariados (que olham apenas para o passado do OP) falham em superar os *benchmarks* lineares. A precisão só aumenta significativamente quando se incorporam os retornos do ETH como variável exógena. Isto valida empiricamente a causalidade de Granger bidirecional detetada, no qual o valor do Optimism é intrinsecamente dependente do fluxo de informação da Camada 1. Este resultado corrobora a tese de que a hierarquia tecnológica (*Rollup* sobre *blockchain* base) se traduz numa hierarquia económica de formação de preços, onde a interação entre tokens introduz dinâmicas complexas. A relação simbiótica observada no Optimism, onde este também influencia o Ethereum, pode refletir a importância crescente das Camadas 2 na geração de atividade e receitas para a rede base.

A análise de volatilidade (Secção 4.3) adiciona uma camada de prudência a estes resultados. A correlação condicional dinâmica (DCC) consistentemente elevada e os efeitos de transmissão de volatilidade (*spillover*) indicam que, apesar das diferenças na previsibilidade, os riscos sistémicos são partilhados. Este facto está em linha com a literatura revista, que aponta para a forte interdependência e complexidade destes mercados (Malladi & Dheeriyaa, 2021) e para a prevalência de contágio de volatilidade e clusters de risco, conforme documentado por Dyhrberg (2016) e Zhang (2019). Estas conclusões têm implicações diretas para a gestão de portefólios e estratégias de investimento, como por exemplo:

- a) Na diversificação limitada. A forte correlação contemporânea sugere que deter um portefólio misto de Camada 1 e Camada 2 oferece poucos benefícios de diversificação real em momentos de crise, uma conclusão também apoiada por Jayawardhana & Colombage (2024), que sugerem a diversificação através de ecossistemas distintos.
- b) Em estratégias ativas. Na previsibilidade superior da Camada 2 através de modelos híbridos abre portas a estratégias de trading ativo (e.g. pairs trading dinâmico), onde a ineficiência do preço do MATIC ou OP pode ser explorada em janelas de curto prazo (3-5 dias), utilizando o ETH como sinal avançado. Esta abordagem tira partido das oportunidades de arbitragem identificadas por Nair (2021) através da análise de cointegração e pares.

Em suma, os resultados desta investigação sugerem que o ecossistema Ethereum não constitui um bloco monolítico. Trata-se de um sistema hierárquico no qual o ativo base dita a tendência macro, caracterizando-se por elevada eficiência e dominância informacional, enquanto os ativos de escalabilidade operam como derivados tecnológicos complexos, cujas ineficiências de preço requerem ferramentas analíticas mais avançadas, como redes neuronais com mecanismos de atenção, para serem adequadamente compreendidas e antecipadas.

6. CONCLUSÃO

A presente dissertação investigou a dinâmica estrutural e preditiva entre criptoativos de Camada 1 (Ethereum) e de Camada 2 (Polygon e Optimism), desafiando a premissa de que a interdependência tecnológica se traduz numa equivalência de comportamento de mercado. Através de uma abordagem metodológica integrada, que combinou econometria clássica (VAR/VECM, GARCH) com técnicas avançadas de Deep Learning (ARIMA–LSTM com Mecanismo de Atenção), a investigação alcançou resultados que permitem responder cabalmente às questões de investigação formuladas.

Relativamente à estrutura e dependência económica, lançada na primeira questão de investigação (Q1), a análise estrutural revelou que, contrariamente à expectativa inicial, não existe um equilíbrio de longo prazo (cointegração) entre os preços do ETH e os seus tokens de escalabilidade. No entanto, identificaram-se mecanismos de transmissão de preços robustos e distintos, por exemplo, o Polygon (MATIC) comporta-se como um ativo satélite, reagindo unidirecionalmente aos choques do Ethereum, enquanto o Optimism (OP) exhibe uma relação simbiótica de causalidade bidirecional. Estes resultados sugerem que no período estudado, a hierarquia da infraestrutura *blockchain* reflete-se numa hierarquia de fluxos de informação, onde o ativo base lidera a tendência, mas interage complexamente com soluções de *Rollup* mais integradas.

No que respeita à inovação preditiva e à eficiência de mercado, associadas à segunda questão de investigação (Q2), os resultados empíricos obtidos nesta dissertação fornecem evidência consistente de que a previsibilidade dos preços difere de forma sistemática entre a Camada 1 e as soluções de Camada 2 do ecossistema Ethereum. Em linha com a discussão desenvolvida para o Ethereum observa-se que os modelos lineares constituem um benchmark difícil de superar, não sendo identificados ganhos preditivos estatisticamente relevantes com a introdução de arquiteturas não lineares, o que é compatível com um elevado grau de eficiência informacional.

Por contraste, para os tokens de Camada 2, os resultados indicam que a incorporação de informação exógena proveniente do Ethereum, bem como a utilização de mecanismos de ponderação temporal dinâmica, contribuem para uma redução mensurável do erro de previsão no contexto analisado. Tal como discutido anteriormente, esta evidência sugere que a dinâmica de curto prazo destes ativos apresenta dependências mais complexas, que podem ser captadas por modelos híbridos multivariados mais flexíveis, sem que tal implique extrapolar conclusões estruturais para além do âmbito empírico considerado..

Por fim, o risco e gestão de portfólio, implícita na terceira questão de investigação (Q3), a análise de volatilidade (DCC-GARCH) evidenciou correlações condicionais consistentemente elevadas e efeitos de contágio (*spillover*) persistentes. Isto implica que, apesar das diferenças na previsibilidade, os riscos sistémicos são partilhados, limitando os benefícios de diversificação passiva entre Camada 1 e 2. Para investidores em criptomoedas, a conclusão é clara: enquanto o Ethereum deve ser tratado como o ativo "beta" do ecossistema, os tokens de Camada 2 oferecem oportunidades de "alfa" através de estratégias de trading ativo que capitalizem na sua previsibilidade de curto prazo condicionada ao comportamento da rede base.

Limitações da Investigação

Apesar dos contributos apresentados, esta investigação apresenta algumas limitações que devem ser consideradas na interpretação dos resultados. Em primeiro lugar, a análise focou-se apenas em dois tokens de Camada 2: um *Optimistic Rollup* e uma *Sidechain*. Isto significa que

as conclusões sobre a relação hierárquica entre Camada 1 e 2 podem não se aplicar a outras soluções de escalabilidade, como os *ZK-Rollups*, que funcionam de forma diferente e podem gerar dinâmicas próprias.

Adicionalmente, os modelos utilizados basearam-se exclusivamente em dados de mercado, como preços e volatilidade. Não foram incluídas métricas *on-chain*, como TVL (Valor Total Bloqueado) ou receitas de gás dos *sequencers*¹⁰, que poderiam ajudar a explicar parte das variações que a rede neuronal tenta capturar apenas com base em padrões históricos.

A validação preditiva também apresenta limitações. Embora o período total analisado tenha sido de três anos, o desempenho do modelo híbrido foi avaliado numa janela *out-of-sample* relativamente curta, de apenas 30 dias. A extensão desta análise a diferentes regimes de mercado, como períodos de correções abruptas ou fases prolongadas de lateralização, e a janelas de teste mais longas permitiria avaliar de forma mais robusta a estabilidade e generalização dos resultados. Importa ainda notar que modelos de *Deep Learning* aplicados a séries financeiras com baixa relação sinal-ruído estão sujeitos a um risco acrescido de *overfitting* e tendem a apresentar menor interpretabilidade quando comparados com modelos econométricos tradicionais, como o VAR.

Por fim, importa referir uma limitação técnica relacionada com o uso de dois ambientes computacionais distintos. A análise econométrica (como testes de cointegração e modelos VAR) foi realizada em R, enquanto os modelos de *Deep Learning* foram implementados em Python. Embora ambos sejam ambientes fiáveis, pequenas diferenças nos algoritmos, nos critérios de convergência ou nas funções de otimização podem gerar variações ligeiras nos resultados. Assim, as comparações diretas entre estimativas obtidas em fases diferentes devem ser interpretadas com cautela, valorizando mais as tendências gerais do que diferenças numéricas muito precisas.

Direções para Investigação Futura

Em síntese, esta investigação contribui para a literatura financeira ao propor um enquadramento empírico que distingue o comportamento informacional da Camada 1 e das soluções de Camada 2 no ecossistema Ethereum, evidenciando diferenças sistemáticas na previsibilidade dos respetivos preços. Do ponto de vista metodológico, os resultados obtidos sugerem que arquiteturas híbridas multivariadas com mecanismos de atenção constituem uma abordagem promissora para modelar ativos cuja dinâmica depende de variáveis exógenas associadas ao protocolo base. Em particular, no contexto analisado, observa-se que a incorporação dos retornos do ativo subjacente contribui para uma redução mensurável do erro de previsão em tokens associados a *rollups* e *sidechains*, sem que tal implique extrapolar conclusões estruturais para além do âmbito empírico considerado.

Não obstante os resultados robustos obtidos com a rede LSTM, a evidência empírica sugere que a complexidade destes mercados comporta ainda margem para exploração através de arquiteturas de *Deep Learning* de nova geração. Nesse sentido, identificam-se três linhas prioritárias para investigações futuras:

¹⁰ Um *Sequencer* é um nó (ou conjunto de nós) especializado na rede de Camada 2 que atua como um "controlador de tráfego". É a entidade responsável por receber, organizar e executar as transações dos utilizadores fora da cadeia principal (*off-chain*) antes de as enviar para a Camada 1 (Ethereum). Um processo automatizado e computarizado.

- a) Exploração de Modelos *Transformer*: Dada a eficácia do mecanismo de atenção isolado neste estudo, o passo natural seguinte será a implementação de arquiteturas baseadas inteiramente em atenção (como Transformers ou Temporal Fusion Transformers). Estes modelos possuem uma capacidade superior para paralelizar o processamento e capturar dependências de muito longo prazo, superando as limitações de memória sequencial das RNNs clássicas.
- b) Arquiteturas Híbridas CNN-LSTM: A integração de Redes Neurais Convolucionais (CNN) com LSTM apresenta-se como uma via promissora para identificar padrões em múltiplas escalas temporais. Enquanto a CNN pode extrair características locais e padrões gráficos da volatilidade de curto prazo, a LSTM encarregar-se-ia da modelação da dependência temporal sequencial.
- c) Desenvolvimento de Modelos de *Ensemble*: Para mitigar a variância das previsões individuais e aumentar a robustez face a mudanças de regime de mercado, sugere-se a combinação de múltiplas arquiteturas (Lineares, LSTM, GARCH) através de técnicas de Ensemble Learning (e.g., Stacking ou Weighted Averaging).

Estas extensões metodológicas, em particular a incorporação de métricas *onchain* (como volumes em DEX e indicadores de atividade dos sequencers), têm o potencial de enriquecer o espaço de informação disponível ao modelo, contribuindo para uma representação mais completa do processo de formação de preços em ambientes *blockchain*. Paralelamente, o avanço da investigação deverá centrar-se na redução do compromisso entre desempenho preditivo e interpretabilidade económica, que caracteriza grande parte dos modelos de *Deep Learning* aplicados a séries financeiras. Neste contexto, a aplicação de técnicas de Inteligência Artificial Explicável (XAI) a arquiteturas baseadas em LSTM emerge como uma linha de investigação prioritária, ao permitir a decomposição e interpretação das dependências temporais e interativas aprendidas pelo modelo. Ao tornar explícitos os mecanismos latentes subjacentes às previsões, estas abordagens não só reforçam a validação empírica das relações de dependência identificadas, como também aproximam a robustez estatística dos modelos preditivos da sua interpretabilidade económica, em linha com os desenvolvimentos recentes na literatura (Park & Yang, 2022).

7. REFERÊNCIAS

- Acuna-García, J. A. (2022). Stock Market Forecasting Using Continuous Wavelet Transform and Long Short-Term Memory Neural Networks. *International Journal of Advanced Research in Computer Science*, 13(06), 33–39. <https://doi.org/10.26483/ijarcs.v13i6.6919>
- Alvi, J., Nizam, K., Jafri, S. M., Rehan, M., & Ali, M. M. (2025). *Cryptocurrency Predictive Analytics: A Comparative Study of LSTM, CNN, and GRU Models*.
- Arumugam, V., & Natarajan, V. (2023a). Time Series Modeling and Forecasting Using Autoregressive Integrated Moving Average and Seasonal Autoregressive Integrated Moving Average Models. *Instrumentation Mesure Métrologie*, 22(4), 161–168. <https://doi.org/10.18280/i2m.220404>
- Arumugam, V., & Natarajan, V. (2023b). Time Series Modeling and Forecasting Using Autoregressive Integrated Moving Average and Seasonal Autoregressive Integrated Moving Average Models. *Instrumentation Mesure Métrologie*, 22(4), 161–168. <https://doi.org/10.18280/i2m.220404>
- Au, S., & Power, T. (2018). *Tokenomics: The crypto shift of blockchains, ICOs, and tokens*. Packt Publishing Ltd.
- Azari, A. (2019). Bitcoin price prediction: An ARIMA approach. *ArXiv Preprint ArXiv:1904.05315*.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327. [https://doi.org/https://doi.org/10.1016/0304-4076\(86\)90063-1](https://doi.org/https://doi.org/10.1016/0304-4076(86)90063-1)
- Box, G. E. P., & Jenkins, G. M. (1976). Time Series Analysis: Forecasting and Control. In *Holden-Day series in time series analysis and digital processing*. Holden-Day. <https://books.google.pt/books?id=1WVHAAAAMAAJ>
- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.
- Buterin, V. (2014). Ethereum White Paper: A Next Generation Smart Contract & Decentralized Application Platform. *Ethereum*, January.
- Cao, R., & Wang, Q. (2024). An evaluation of standard statistical models and llms on time series forecasting. *ArXiv Preprint ArXiv:2408.04867*.
- Catalini, C., & Gans, J. S. (2018). Initial Coin Offerings and the Value of Crypto Tokens. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3137213>
- Caton, J. L. (2020). Cryptoliquidity: the blockchain and monetary stability. *Journal of Entrepreneurship and Public Policy*, 9(2), 227–252.
- Chen, K., Fan, Y., & Liao, S. S. (2023). Token incentives in a volatile crypto market: the effects of token price volatility on user contribution. *Journal of Management Information Systems*, 40(2), 683–711.
- Chen, Y. (2018). Blockchain tokens and the potential democratization of entrepreneurship and innovation. *Business Horizons*, 61(4). <https://doi.org/10.1016/j.bushor.2018.03.006>
- Chen, Y. (2024a). Application of GARCH Model in the Field of Finance. *Advances in Economics, Management and Political Sciences*, 124, 164–169. <https://doi.org/10.54254/2754-1169/2024.17736>
- Chen, Y. (2024b). Application of GARCH Model in the Field of Finance. *Advances in Economics Management and Political Sciences*, 124(1), 164–169. <https://doi.org/10.54254/2754-1169/2024.17736>
- Cong, L. W., Li, Y., & Wang, N. (2021). Tokenomics: Dynamic Adoption and Valuation. *The Review of Financial Studies*, 34(3), 1105–1155. <https://doi.org/10.1093/rfs/hhaa089>

- Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366a), 427–431. <https://doi.org/10.1080/01621459.1979.10482531>
- Diebold, F. X., & Yilmaz, K. (2023). Measuring the connectedness of financial markets: A revised approach. *Journal of Financial Econometrics*, 21(1), 73–96.
- Dyhrberg, A. H. (2016). Bitcoin, gold and the dollar - A GARCH volatility analysis. *Finance Research Letters*, 16. <https://doi.org/10.1016/j.frl.2015.10.008>
- Engle, R. F. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3), 339–350.
- Engle, R. F., Ghysels, E., & Sohn, B. (2013). Stock market volatility and macroeconomic fundamentals. *Review of Economics and Statistics*, 95(3), 776–797.
- Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 55(2), 251–276.
- Fernandez, J. D., Barbereau, T., & Papageorgiou, O. (2024). Agent-based model of initial token allocations: Simulating distributions post fair launch. *ACM Transactions on Management Information Systems*, 15(1), 1–22.
- Fleischer, J., Von Laszewski, G., Theran, C., Jairo, Y., & Bautista, P. (2022). Time Series Analysis of Blockchain-Based Cryptocurrency Price Changes. In *Accepted at the 2022 Emerging Researchers National (ERN) Conference in STEM, Washington, D.C., due to COVID-19, the conference is postponed until 2023* (Vol. 1).
- Freni, P., Ferro, E., & Moncada, R. (2022). Tokenomics and blockchain tokens: A design-oriented morphological framework. *Blockchain: Research and Applications*, 3(1), 100069. <https://doi.org/https://doi.org/10.1016/j.bcr.2022.100069>
- Gans, J. S. (2024). Cryptic Regulation of Crypto-Tokens. *Entrepreneurship and Innovation Policy and the Economy*, 3(1), 139–163.
- Gers, F. A., Schmidhuber, J., & Cummins, F. (1999). Learning to forget: continual prediction with LSTM. *1999 Ninth International Conference on Artificial Neural Networks ICANN 99. (Conf. Publ. No. 470)*, 2, 850–855 vol.2. <https://doi.org/10.1049/cp:19991218>
- Giantsidi, S., & Tarantola, C. (2025). Deep learning for financial forecasting: A review of recent trends. *International Review of Economics & Finance*, 104, 104719. <https://doi.org/https://doi.org/10.1016/j.iref.2025.104719>
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative Adversarial Nets. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems* (Vol. 27). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2014/file/5ca3e9b122f61f8f06494c97b1afcef3-Paper.pdf
- Granger, C. W. J. (1981). Some properties of time series data and their use in econometric model specification. *Journal of Econometrics*, 16(1), 121–130.
- Gudgeon, L., Moreno-Sanchez, P., Roos, S., McCorry, P., & Gervais, A. (2020). SoK: Layer-Two Blockchain Protocols. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12059 LNCS. https://doi.org/10.1007/978-3-030-51280-4_12
- Gunawan, D., & Febrianti, I. (2023). Ethereum Value Forecasting Model Using Autoregressive Integrated Moving Average (ARIMA). *International Journal of Advances in Social Sciences and Humanities*, 2(1), 29–35. <https://doi.org/10.56225/ijassh.v2i1.151>

- Gupta, B. B., Gaurav, A., Piñeiro-Chousa, J., López-Cabarcos, M. Á., & López, I. G. (2025). Predicting the variation of decentralised finance cryptocurrency prices using deep learning and a BiLSTM-LSTM based approach. *Enterprise Information Systems*, 19(7–8), 2483456. <https://doi.org/10.1080/17517575.2025.2483456>
- Gupta, H., & Chaudhary, R. (2022). An Empirical Study of Volatility in Cryptocurrency Market. *Journal of Risk and Financial Management*, 15(11). <https://doi.org/10.3390/jrfm15110513>
- Hafid, A., Hafid, A. S., & Samih, M. (2020). Scaling blockchains: A comprehensive survey. *IEEE Access*, 8, 125244–125262.
- Han, J., Lee, J., & Li, T. (2025). A review of DAO governance: Recent literature and emerging trends. *Journal of Corporate Finance*, 102734.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition (Springer Series in Statistics)*.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780.
- Hyndman, R. J., & Athanasopoulos, G. (2018a). *Forecasting: principles and practice*. OTexts.
- Hyndman, R. J., & Athanasopoulos, G. (2018b). *Forecasting: principles and practice*. OTexts.
- Hyndman, R. J., & Athanasopoulos, G. (2018c). *Forecasting: principles and practice*. OTexts.
- Jayawardhana, A., & Colombage, S. R. N. (2024). Portfolio diversification possibilities of cryptocurrency: global evidence. *Applied Economics*, 56(47), 5618–5633. <https://doi.org/10.1080/00036846.2023.2257928>
- Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12(2–3), 231–254.
- Ju, J., & Liu, F.-A. (2021). Multivariate Time Series Data Prediction Based on ATT-LSTM Network. *Applied Sciences*, 11(20). <https://doi.org/10.3390/app11209373>
- Kardakis, S., Perikos, I., Grivokostopoulou, F., & Hatzilygeroudis, I. (2021). Examining attention mechanisms in deep learning models for sentiment analysis. *Applied Sciences*, 11(9), 3883.
- Katsika, A., Negka, L., Spathoulas, G., & Plagianakos, V. (2024). A Critical View on Blockchain Rollups. In *Security and Privacy in Smart Environments* (pp. 204–239). Springer.
- Kulik, R., & Soulier, P. (2015). Heavy tailed time series with extremal independence. *Extremes*, 18(2), 273–299. <https://doi.org/10.1007/s10687-014-0213-x>
- Lara-Benítez, P., Carranza-García, M., & Riquelme, J. C. (2021). An Experimental Review on Deep Learning Architectures for Time Series Forecasting. *International Journal of Neural Systems*, 31(03), 2130001. <https://doi.org/10.1142/s0129065721300011>
- Lesche, P., Sandner, P., & Treiblmaier, H. (2022). Implications of the token economy: A taxonomy and research agenda. *Blockchains and the Token Economy: Theory and Practice*, 1–30.
- Levakova, M., & Ditlevsen, S. (2024). Penalisation Methods in Fitting High-Dimensional Cointegrated Vector Autoregressive Models: A Review. *International Statistical Review*, 92(2), 160–193. <https://doi.org/https://doi.org/10.1111/insr.12553>
- Li, G., & Dai, Y. (2024). *Time Series Anomaly Detection Using LSTM and Attention*. 88. <https://doi.org/10.1117/12.3035015>
- Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer Berlin Heidelberg. <https://books.google.pt/books?id=muorJ6FHliEC>
- Lütkepohl, H. (2022). *New introduction to multiple time series analysis (3rd ed.)*. Springer.

- Malladi, R. K., & Dheeriyaa, P. L. (2021). Time series analysis of Cryptocurrency returns and volatilities. *Journal of Economics and Finance*, 45(1), 75–94.
<https://doi.org/10.1007/s12197-020-09526-4>
- Mardjo, A., & Choksuchat, C. (2024). HyBiLSTM: Multivariate Bitcoin Price Forecasting Using Hybrid Time-Series Models With Bidirectional LSTM. *IEEE Access*, 12, 50792–50808. <https://doi.org/10.1109/ACCESS.2024.3386029>
- Mohamed, H. (2022). Decentralizing Finance: Cryptocurrencies, ICOs, STOs and Tokenization of Assets. In *Blockchain Technology* (pp. 221–234). CRC Press.
- Nair, S. T. G. (2021). Pairs trading in cryptocurrency market: A long-short story. *Investment Management and Financial Innovations*, 18(3), 127–141.
[https://doi.org/10.21511/imfi.18\(3\).2021.12](https://doi.org/10.21511/imfi.18(3).2021.12)
- Nakamoto, S. (2008). Bitcoin: a peer-to-peer electronic cash system, October 2008. *Cited On*.
- Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., ... Yarkoni, T. (2015). Promoting an open research culture. *Science*, 348(6242), 1422–1425. <https://doi.org/10.1126/science.aab2374>
- Ormaniec, W., Pitera, M., Safarveisi, S., & Schmidt, T. (2022). *Estimating Value at Risk: LSTM vs. GARCH*. <https://doi.org/10.48550/arxiv.2207.10539>
- Park, S., & Yang, J.-S. (2022). Interpretable deep learning LSTM model for intelligent economic decision-making. *Knowledge-Based Systems*, 248, 108907.
<https://doi.org/https://doi.org/10.1016/j.knosys.2022.108907>
- Peng, R. D. (2011). Reproducible Research in Computational Science. *Science*, 334(6060), 1226–1227. <https://doi.org/10.1126/science.1213847>
- Poon, J., & Dryja, T. (2015). The Bitcoin Lightning Network: Scalable Off-Chain Instant Payments. In *Scalable o-chain instant payments*.
- Qin, K., Zhou, L., Gamito, P., Jovanovic, P., & Gervais, A. (2021). An empirical study of defi liquidations: Incentives, risks, and instabilities. *Proceedings of the 21st ACM Internet Measurement Conference*, 336–350.
- Raut, S. (2024). Analysis & Stock Price Prediction and Forecasting Using Different LSTM Models. *Interantional Journal of Scientific Research in Engineering and Management*, 08(04), 1–5. <https://doi.org/10.55041/ijrsrem30115>
- Rizvi, M. F. (2024). ARIMA Model Time Series Forecasting. *International Journal for Research in Applied Science and Engineering Technology*, 12(5), 3782–3785.
<https://doi.org/10.22214/ijraset.2024.62416>
- Saleh, F. (2021). Blockchain without waste: Proof-of-stake. *The Review of Financial Studies*, 34(3), 1156–1190.
- Sandve, G. K., Nekrutenko, A., Taylor, J., & Hovig, E. (2013). Ten Simple Rules for Reproducible Computational Research. *PLOS Computational Biology*, 9(10), e1003285-
<https://doi.org/10.1371/journal.pcbi.1003285>
- Santoso, A. B., & Widodo, T. (2024). Predicting the Number of Forest and Land Fire Hotspot Occurrences Using the ARIMA and SARIMA Methods. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, 13(1), 119–129.
<https://doi.org/10.32736/sisfokom.v13i1.2018>
- Schueffel, P. (2021). Defi: Decentralized finance-an introduction and overview. *Journal of Innovation Management*, 9(3), I–XI.
- Sezgin, F. H., Algorabi, Ö., Sart, G., & Güler, M. (2025). Hyperparameter-Optimized RNN, LSTM, and GRU Models for Airline Stock Price Prediction: A Comparative Study on THYAO and PGSUS. *Symmetry*, 17(11), 1905.

- Sguanci, C., Spatafora, R., & Vergani, A. M. (2021). Layer 2 blockchain scaling: A survey. *ArXiv Preprint ArXiv:2107.10881*.
- Sherstinsky, A. (2020). Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network. *Physica D: Nonlinear Phenomena*, 404, 132306. <https://doi.org/https://doi.org/10.1016/j.physd.2019.132306>
- Song, H., Qu, Z., & Wei, Y. (2024). Advancing blockchain scalability: An introduction to layer 1 and layer 2 solutions. *2024 IEEE 2nd International Conference on Sensors, Electronics and Computer Engineering (ICSECE)*, 71–76.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1), 1929–1958.
- Tunncliffe Wilson, G. (2016). Time Series Analysis: Forecasting and Control, 5th Edition, by George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel and Greta M. Ljung, 2015. Published by John Wiley and Sons Inc., Hoboken, New Jersey, pp. 712. ISBN: 978-1-118-67502-1. *Journal of Time Series Analysis*, 37, n/a-n/a. <https://doi.org/10.1111/jtsa.12194>
- Upadhyay, H. P., & Pradhan, B. L. (2023). Forecasting GDP of Nepal Using Autoregressive Integrated Moving Average (ARIMA) Model. *International Journal of Silkroad Institute of Research and Training*, 1(1), 2–7. <https://doi.org/10.3126/ijst.v1i1.55923>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30.
- Vuković, D. B., Radenković, S. D., Simeunović, I., Zinovev, V., & Radovanović, M. (2024). Predictive patterns and market efficiency: A deep learning approach to financial time series forecasting. *Mathematics*, 12(19), 3066.
- Zhang, X., Liang, X., Zhiyuli, A., Zhang, S., Xu, R., & Wu, B. (2019). At-lstm: An attention-based lstm model for financial time series prediction. *IOP Conference Series: Materials Science and Engineering*, 569(5), 052037.
- Zhang, Z. (2019). Engineering token economy with system modeling. *ArXiv Preprint ArXiv:1907.00899*.
- Zheng, Z., Xie, S., Dai, H. N., Chen, X., & Wang, H. (2018). Blockchain challenges and opportunities: A survey. *International Journal of Web and Grid Services*, 14(4). <https://doi.org/10.1504/IJWGS.2018.095647>

8. ANEXOS

Anexo A - Procedimento do Modelo Híbrido ARIMA–LSTM com *Self-Attention*

Passo 1 — Estimação da Componente Linear (ARIMA)

1. Para cada série y_t (ETH, MATIC, OP):
 - Aplicar Auto-ARIMA para selecionar automaticamente (p, d, q) .
 - Estimar o modelo e obter os resíduos:

$$\hat{\varepsilon}_t = y_t - \hat{y}_t^{(ARIMA)}$$

2. Guardar:
 - Previsões lineares $\hat{y}_{t+h}^{(ARIMA)}$
 - Resíduos $\hat{\varepsilon}_t$ para alimentar a LSTM.

Passo 2 — Preparação dos Dados para a Componente Não Linear (LSTM)

Construir um vetor multivariado de entrada:

$$X_t = \begin{bmatrix} \hat{\varepsilon}_t \\ \sigma_t^{(7)} \\ r_{ETH,t-1} \end{bmatrix}$$

Onde:

- $\hat{\varepsilon}_t$ = resíduo do ARIMA
- $\sigma_t^{(7)}$ = volatilidade rolante de 7 dias
- $r_{ETH,t-1}$ = retorno defasado do ETH, incorporando causalidade de Granger
- Para ETH, utiliza-se apenas univariado (sem exógenas)

Criar janelas temporais (sliding windows):

$$(X_{t-L+1}, \dots, X_t) \rightarrow \hat{\varepsilon}_{t+1}$$

com tamanho de janela L selecionado por validação.

Passo 3 — Arquitetura da Componente Não Linear (LSTM + Self-Attention)

3.1. Camada LSTM

Processa a sequência temporal e produz uma série de estados ocultos:

$$(h_1, h_2, \dots, h_T)$$

3.2. Mecanismo de Self-Attention

Aplicado sobre os estados ocultos h_t :

1. Calcular vetores de consulta, chave e valor:

$$Q = W_Q h_t, K = W_K h_t, V = W_V h_t$$

2. Calcular pesos de atenção:

$$\alpha_{ij} = \text{softmax} \left(\frac{Q_i K_j^T}{\sqrt{d_k}} \right)$$

3. Combinar a informação temporal relevante:

$$z_t = \sum_j \alpha_{ij} V_j$$

O vetor z_t representa a sequência “reponderada”, realçando dias com maior relevância preditiva.

3.3. Camada Densa de Saída

Produz a previsão do componente não linear:

$$\hat{\varepsilon}_{t+1}^{(LSTM)}$$

Passo 4 — Combinação das Componentes Linear + Não Linear

A previsão híbrida final é dada por:

$$\hat{y}_{t+1} = \hat{y}_{t+1}^{(ARIMA)} + \hat{\varepsilon}_{t+1}^{(LSTM)}$$

Passo 5 — Treino e Validação

- Otimizador **Adam**, com *learning rate* adaptativo
- Função de perda: MSE dos resíduos
- Regularização:
 - Dropout = 20%
 - Early Stopping baseado na perda de validação
- Grid search para:
 - Número de unidades LSTM
 - Tamanho da janela temporal
 - Número de cabeças de atenção (se multi-head)

Passo 6 — Avaliação Out-of-Sample

Para os últimos 30 dias:

1. Gerar previsões híbridas $\hat{y}_{t+h}$
2. Comparar com benchmarks (ARIMA, VAR, LSTM sem atenção) usando RMSE e MAE.

Anexo B – Prompts utilizadas no Google® Gravity

Dinâmica de preços

Role: Act as an Expert in Financial Econometrics and Crypto Markets.

Context: I am writing the "Analysis of Results" chapter (specifically Section 4.1) for a Master's dissertation comparing Layer 1 (ETH) and Layer 2 (MATIC, OP) tokens.

Data Scope: Use the ENTIRE dataset (Daily Closing Prices from June 30, 2022, to June 30, 2025). Do not split into train/test sets. We are performing a structural analysis of the whole period.

Objective: Generate the statistical output and economic interpretation to rewrite Chapter : Analysis of Dynamic Price Relationships. The goal is to understand how these tokens interact structurally over the long term.

Required Analysis Steps (Full Sample): Stationarity & Cointegration (The Foundation): ADF Test: Confirm stationarity of Log>Returns for the full 3-year period. Johansen Test: Run on Log-Prices (Levels).

Confirm if there is any long-term equilibrium (Cointegration) now that we are looking at the full sample. (Expected: Likely None, but verify). VAR Model Specification

(Structural): Estimate a Vector Autoregression (VAR) model for the pairs ETH-MATIC and ETH-OP using the full sample of log-returns. Lag Selection: Use Information Criteria

(AIC/BIC) to determine the optimal lag order (\$p\$) for the whole period. Report the selected lags. Granger Causality Tests (Crucial): Perform Granger Causality tests (Wald test) for both directions: Does ETH Granger-cause MATIC? (and vice-versa) Does ETH Granger-cause OP? (and vice-versa) Interpretation: Interpret the F-statistics and p-values.

Discuss if the relationship is unidirectional (Hierarchical) or bidirectional

(Interdependent). Structural Dynamics (Plots): Impulse Response Functions (IRF): Generate and plot the orthogonalized IRFs for 10-15 periods.

How does a shock in ETH affect MATIC/OP over the next 2 weeks? Is the reaction immediate or delayed? Variance Decomposition (FEVD): Calculate the FEVD table.

What percentage of MATIC/OP variance is explained by ETH shocks vs. their own idiosyncratic shocks over the long term? Output Requirements: Provide the Python code to run this full-sample analysis. Generate the Tables for Granger Causality and Cointegration results. Write a detailed interpretation of the results, specifically focusing on the economic implications of the causal links found (e.g., "ETH acts as a market leader for...").

Machine learning

Prompt: Comprehensive Crypto Price Dynamics & Hybrid Forecasting Analysis

****Context:****

I have a dataset (`Price_ETH_MATIC_OP.xlsx`) containing daily closing prices for Ethereum (ETH), Polygon (MATIC), and Optimism (OP). I need a Python script to perform a comprehensive econometric and deep learning analysis of their dynamic price relationships and forecasting performance.

****Objective:****

Write a robust, single-file Python script (`main_analysis.py`) to execute the following analysis pipeline. The script should save all plots and print statistical results to the console.

****Requirements:****

1. **Data Preprocessing & Visualization**
 - * Load the Excel file.
 - * Calculate **Logarithmic Returns** for all tokens ($\ln(P_t / P_{t-1})$).
 - * Generate and save plots for:
 - * Daily Closing Prices.
 - * Daily Log Returns.
2. **Statistical Testing**
 - * **Stationarity**: Perform the **Augmented Dickey-Fuller (ADF)** test on the Log Returns of each token. Interpret the p-values (threshold 0.05).
 - * **Cointegration**: Perform the **Johansen Cointegration Test** on the *Price Levels* (not returns) to check for long-term equilibrium relationships. Report Trace and Max-Eigenvalue statistics.
3. **Vector Autoregression (VAR) Analysis**
 - * Fit a VAR model on the stationary Log Returns.
 - * **Lag Selection**: Automatically select the optimal lag order based on the **AIC** (Akaike Information Criterion).
 - * **Impulse Response Functions (IRF)**: Generate and save the orthogonalized IRF plot (10 steps) to show how shocks propagate.
 - * **Variance Decomposition (FEVD)**: Generate and save the FEVD plot (10 steps) to analyze variance contributions.
4. **Forecasting Model Comparison**
 - * For each token, compare the forecasting performance (RMSE, MAE, MAPE) of three models over a **30-day test set**:
 - * **VAR**: Multivariate linear baseline.
 - * **ARIMA**: Univariate linear baseline (Auto-ARIMA or fixed order).
 - * **Basic Hybrid ARIMA-LSTM**:
 - * Fit ARIMA.
 - * Train a simple LSTM (using PyTorch) on the ARIMA residuals.
 - * Final Forecast = ARIMA Forecast + LSTM Residual Forecast.
5. **Advanced Model Tuning & Implementation**
 - * **Tuned Hybrid Model**:
 - * Implement a grid search (or similar tuning) for the LSTM hyperparameters (Look-back window, Hidden Units) and ARIMA order (Auto-ARIMA).
 - * Use a validation split for model selection.
 - * **Multivariate Attention Hybrid Model**:
 - * Implement a complex LSTM with a **Self-Attention Mechanism**.
 - * **Features**: Input not just past residuals, but also **Lagged ETH Returns** (as an exogenous driver) and **Rolling Volatility** of residuals.
 - * Evaluate if this multivariate attention approach outperforms the univariate baselines for the Layer 2 tokens (MATIC, OP).

Technical Stack

- * **Libraries**: `pandas`, `numpy`, `matplotlib`, `seaborn`, `statsmodels`, `pmdarima`, `scikit-learn`, `torch` (PyTorch).
- * **Best Practices**: Ensure reproducibility (set random seeds), handle missing data, and provide clear console output for all statistical metrics.

Anexo C – Resultados dos Modelos ARIMA

Este anexo apresenta os resultados completos da selecção dos modelos ARIMA utilizados como componente linear na abordagem híbrida ARIMA–LSTM. A estimativa foi realizada para os retornos diários do Ethereum (ETH), Polygon (MATIC) e Optimism (OP), recorrendo ao critério de informação de Akaike (AIC) e ao critério de informação bayesiano (BIC). Em todos os casos, as séries se encontram estacionárias em níveis ($d = 0$), conforme evidenciado nos testes ADF descritos no Capítulo 4.

A2.1 Resultados ARIMA para ETH

Tabela C-1 Comparação de modelos ARIMA para ETH

Ordem (p,d,q)	AIC	BIC
(1, 0, 0)	-4252.4132	-4237.4149
(0, 0, 0)	-4252.0945	-4242.0956
(0, 0, 1)	-4252.3375	-4237.3392
(2, 0, 0)	-4250.6227	-4230.6250

O modelo ARIMA(1,0,0) apresenta o menor AIC, sendo ligeiramente preferível para fins de ajuste. Contudo, as diferenças entre modelos são marginais, e o ARIMA(0,0,0), equivalente a um passeio aleatório, apresenta o melhor BIC, sugerindo que o ETH exhibe uma estrutura extremamente próxima de um comportamento eficiente sem memória. A presença de um coeficiente AR(1) positivo no modelo selecionado aponta para uma fraca reversão média, ainda que estatisticamente pouco relevante.

A2.2 Resultados ARIMA para MATIC

Tabela C-2 Comparação de modelos ARIMA para MATIC

Ordem (p,d,q)	AIC	BIC
(2, 0, 0)	-3674.3249	-3654.3272
(0, 0, 0)	-3674.2650	-3664.2661
(0, 0, 1)	-3674.2752	-3659.2769
(1, 0, 0)	-3674.1058	-3659.1075

Para o MATIC, os valores de AIC e BIC são praticamente idênticos entre modelos, revelando ausência de estrutura autoregressiva relevante. Embora o ARIMA(2,0,0) minimize o AIC, a melhoria relativamente ao ARIMA(0,0,0) é negligenciável. O BIC favorece o modelo mais parcimonioso (0,0,0), sugerindo que o retorno do MATIC se comporta, na prática, como um passeio aleatório, sem evidência robusta de dependências lineares de curto prazo.

A2.3 Resultados ARIMA para OP

Tabela C-3 Comparação de modelos ARIMA para OP

Ordem (p,d,q)	AIC	BIC
(2, 0, 0)	-3006.1356	-2986.1379
(0, 0, 0)	-3005.6431	-2995.6442
(0, 0, 2)	-3005.8519	-2985.8542
(1, 0, 0)	-3005.4090	-2990.4107

Tal como no caso do MATIC, o OP apresenta diferenças mínimas entre os modelos testados. O ARIMA(2,0,0) surge como melhor modelo segundo o AIC, mas o BIC, que penaliza a complexidade, identifica o ARIMA(0,0,0) como a melhor opção. Este padrão sugere que o OP, enquanto ativo de Camada 2, apresenta um comportamento predominantemente ruidoso, sem estrutura linear significativa que permita capturar previsibilidade de curto prazo. A informação econométrica suportou assim a escolha de um modelo parcimonioso na componente ARIMA da arquitetura híbrida.

Anexo D – Resultados dos Testes Augmented Dickey–Fuller

O presente anexo apresenta os resultados dos testes Augmented Dickey–Fuller (ADF) aplicados aos log-preços e aos retornos logarítmicos diários de ETH, MATIC e OP. O objetivo consiste em avaliar a presença de raízes unitárias e determinar a ordem de integração das séries utilizadas na análise econométrica.

No teste ADF, a hipótese nula estabelece que a série possui uma raiz unitária e, consequentemente, é não estacionária. A hipótese alternativa estabelece que a série é estacionária. A decisão estatística é efetuada através da comparação entre a estatística do teste e os respectivos valores críticos. A hipótese nula é rejeitada quando a estatística ADF é inferior, isto é, mais negativa, do que o valor crítico correspondente ao nível de significância considerado.

Para os log-preços, o teste foi especificado com constante e tendência determinística, dado que as séries apresentam alterações de nível e trajetórias persistentes ao longo do tempo. Para os retornos logarítmicos, utilizou-se uma especificação com constante, uma vez que estas séries oscilam em torno de um nível aproximadamente constante e não apresentam uma tendência determinística evidente. O número de defasamentos foi selecionado através do critério de informação de Akaike.

Table D-1 Resultados dos testes ADF para os log-preços e retornos logarítmicos

Série	Especificação	Estatística ADF	Valor crítico 1%	Valor crítico 5%	Valor crítico 10%	Decisão a 1%
Log-preço ETH	Constante e tendência	-2,1738	-3,96	-3,41	-3,12	Não rejeitar (H ₀)
Log-preço MATIC	Constante e tendência	-2,0916	-3,96	-3,41	-3,12	Não rejeitar (H ₀)
Log-preço OP	Constante e tendência	-1,4762	-3,96	-3,41	-3,12	Não rejeitar (H ₀)
Retorno logarítmico ETH	Constante	-23,4284	-3,43	-2,86	-2,57	Rejeitar (H ₀)

Série	Especificação	Estatística ADF	Valor crítico 1%	Valor crítico 5%	Valor crítico 10%	Decisão a 1%
Retorno logarítmico MATIC	Constante	-24,3924	-3,43	-2,86	-2,57	Rejeitar (H ₀)
Retorno logarítmico OP	Constante	-24,4161	-3,43	-2,86	-2,57	Rejeitar (H ₀)

Nota: (H₀) corresponde à hipótese nula de existência de raiz unitária. Nos log-preços, a estatística relevante é (τ_3), correspondente à especificação com constante e tendência. Nos retornos logarítmicos, a estatística relevante é (τ_2), correspondente à especificação com constante.

Os resultados mostram que, para os log-preços de ETH, MATIC e OP, as estatísticas ADF são superiores aos respectivos valores críticos em valor algébrico, não permitindo rejeitar a hipótese nula de existência de raiz unitária. Conclui-se, portanto, que os log-preços não são estacionários em níveis.

Em contraste, para os retornos logarítmicos, as estatísticas ADF são substancialmente inferiores aos valores críticos de 1%. A hipótese nula é, assim, rejeitada para os três tokens, indicando que os retornos logarítmicos são estacionários.

Uma vez que os retornos logarítmicos correspondem às primeiras diferenças dos log-preços,

$$\log(P_{i,t}) - \log(P_{i,t-1})$$

os resultados são compatíveis com a classificação dos log-preços como séries integradas de ordem um, ($I(1)$), e dos retornos logarítmicos como séries integradas de ordem zero, ($I(0)$).

Esta distinção fundamenta a estratégia econométrica adotada na dissertação. Os log-preços são utilizados nos testes de cointegração, destinados a avaliar a existência de relações de equilíbrio de longo prazo, enquanto os retornos logarítmicos são utilizados na estimação de modelos VAR e de modelos de volatilidade condicional, por apresentarem estacionariedade.

Ver código em <https://github.com/ricabreu/Uab>