

UNIVERSIDADE FEDERAL DE OURO PRETO
INSTITUTO DE CIÊNCIAS SOCIAIS APLICADAS
DEPARTAMENTO DE CIÊNCIAS ECONÔMICAS

JOSÉ GILBERTO GOBBI MERIJ

PREVISÃO DE OSCILAÇÕES DE RETORNOS DE VALE3 POR MODELOS
REGRESSORES, AUTORREGRESSORES E MACHINE LEARNING

MARIANA

2026

JOSÉ GILBERTO GOBBI MERIJ

PREVISÃO DE OSCILAÇÕES DE RETORNOS DE VALE3 POR MODELOS
REGRESSORES, AUTORREGRESSORES E MACHINE LEARNING

Trabalho de Conclusão de Curso
apresentado ao Curso de
Bacharelado em Ciências
Econômicas da Universidade Federal
de Ouro Preto, como requisito parcial
para a obtenção do grau de Bacharel
em Ciências Econômicas.

Área de concentração: Econometria e
Finanças Quantitativas.

Orientador: Prof. Dr. Victor Maia
Senna Delgado

MARIANA

2026

SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

M561p Merij, José Gilberto Gobbi.

Previsão de oscilações de retornos de Vale3 por modelos regressores, autorregressores e machine learning. [manuscrito] / José Gilberto Gobbi Merij. - 2026.
88 f.

Orientador: Prof. Dr. Victor Maia Senna Delgado.
Monografia (Bacharelado). Universidade Federal de Ouro Preto.
Instituto de Ciências Sociais Aplicadas. Graduação em Ciências Econômicas .

1. Ações (Finanças). 2. Análise de séries temporais. 3. Aprendizado do computador. 4. Avaliação de riscos. 5. Bolsa de valores. 6. Minérios de ferro. I. Delgado, Victor Maia Senna. II. Universidade Federal de Ouro Preto. III. Título.

CDU 336.7

Bibliotecário(a) Responsável: Essevalter De Sousa - Bibliotecário CRB6 1407



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE OURO PRETO
REITORIA
INSTITUTO DE CIÊNCIAS SOCIAIS E APLICADAS
COLEGIADO DO CURSO DE CIÊNCIAS ECONÔMICAS



FOLHA DE APROVAÇÃO

José Gilberto Gobbi Merij

Previsão de oscilações de retornos de VALE3 por modelos regressores, autorregressores e machine learning

Monografia apresentada ao Curso de Ciências Econômicas da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de Bacharel em ciências econômicas

Aprovada em 27 de fevereiro de 2026

Membros da banca

Dr. - Victor Maia Senna Delgado - Orientador(a) (Universidade Federal de Ouro Preto)
Dr. - Alan André Borges da Costa - (Universidade Federal de Ouro Preto)
Dr. - Lucas Assis Atílio - (Universidade Federal de Ouro Preto)

Victor Maia Senna Delgado, orientador do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 10/03/2026



Documento assinado eletronicamente por **Victor Maia Senna Delgado, COORDENADOR(A) DE CURSO DE CIÊNCIAS ECONÔMICAS**, em 10/03/2026, às 22:46, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **1073109** e o código CRC **ED09E2BA**.

AGRADECIMENTOS

Agradeço à Marcel Sampy e Daniel Carneiro, meus gestores na XP, cujo suporte, compreensão e oportunidade foram essenciais para que eu chegasse a esta fase final da minha conturbada formação acadêmica. É clichê, mas verba volant, scripta manent. Seria cretino de minha parte perder esta oportunidade de registrar, pela eternidade, minha sincera e mais profunda gratidão.

Agradeço também ao meu orientador Victor Senna, que assumiu esse papel por um acidente do acaso, mas o executou com maestria ímpar, peculiar a um sujeito de excepcional nível. Em sua figura, agradeço à Universidade Federal de Ouro Preto, instituição que resgatou não apenas o meu percurso acadêmico, mas reacendeu minha alma e o meu amor ao saber.

Por último, aos meus grandes amigos Felipe, Pedro e Victor, o Porthos, Aramis e Athos deste ignóbil D'Artagnan. Deixo também meus agradecimentos aos outros companheiros “mosqueteiros do Sr. Tréville”: Antonio, Arthur, Filippo e Pedro Lotti.

“Most, probably, of our decisions to do something positive, the full consequences of which will be drawn out over many days to come, can only be taken as a result of animal spirits — of a spontaneous urge to action rather than inaction, and not as the outcome of a weighted average of quantitative benefits multiplied by quantitative probabilities”.

(John Maynard Keynes)

RESUMO

Este trabalho investiga a previsibilidade e a utilização econômica dos retornos da ação da Vale S.A. (VALE3), com ênfase na relação dinâmica entre o preço internacional do minério de ferro e a volatilidade do ativo. Partindo da premissa amplamente documentada de que a previsão direcional de retornos é limitada, enquanto a dinâmica do risco apresenta maior persistência e estrutura temporal, o estudo combina modelos econométricos tradicionais e métodos de aprendizado de máquina em um arcabouço integrado de análise e trading. São empregados modelos Vetoriais Autorregressivos (VAR) para capturar relações dinâmicas entre os retornos da VALE3 e do minério de ferro, modelos GARCH para estimação e previsão da volatilidade condicional, e algoritmos de Random Forest e redes neurais recorrentes do tipo LSTM para explorar não linearidades e padrões complexos nos dados. As previsões são avaliadas estritamente fora da amostra, por meio de esquemas walk-forward, e convertidas em estratégias quantitativas que incorporam explicitamente custos de transação, turnover e métricas de risco. Um elemento central da análise é a separação entre sinal direcional e controle de regime de volatilidade. Modelos direcionais são utilizados para indicar a direção esperada do retorno, enquanto modelos de volatilidade atuam como filtros de regime (risk-on/risk-off), regulando a exposição ao ativo conforme o ambiente de risco previsto. Os resultados mostram que estratégias puramente direcionais ou puramente baseadas em volatilidade apresentam limitações econômicas quando aplicadas isoladamente. Em contrapartida, abordagens combinadas, sob parametrizações parcimoniosas e limiares de volatilidade intermediários, produzem trajetórias de retorno mais estáveis e drawdowns significativamente menores, ainda que nem sempre superem o benchmark buy-and-hold em retorno absoluto. O trabalho enfatiza que os ganhos econômicos observados decorrem menos da complexidade individual dos modelos e mais da forma como previsões de retorno e risco são integradas de maneira disciplinada. Ao focar na volatilidade recorrente e mediana do ativo — e não em eventos extremos —, o estudo oferece uma avaliação realista do potencial e das limitações de estratégias quantitativas aplicadas à VALE3, deixando aberta a possibilidade de extensões futuras voltadas à modelagem explícita de regimes extremos de mercado.

Palavras-chave: séries temporais financeiras; volatilidade; GARCH; VAR; aprendizado de máquina; estratégias quantitativas; minério de ferro; VALE3.

ABSTRACT

This study investigates the predictability and economic usability of returns of Vale S.A. (VALE3), with particular emphasis on the dynamic relationship between international iron ore prices and the stock's volatility. Starting from the well-established premise that directional return forecasting is inherently limited, while risk dynamics exhibit persistence and temporal structure, the analysis combines traditional econometric models and machine learning techniques within an integrated trading framework. Vector Autoregressive (VAR) models are employed to capture dynamic relationships between VALE3 and iron ore returns, GARCH models are used to estimate and forecast conditional volatility, and machine learning methods — namely Random Forests and Long Short-Term Memory (LSTM) networks — are applied to capture nonlinearities and complex patterns in the data. All forecasts are evaluated strictly out-of-sample using walk-forward procedures and translated into quantitative trading strategies that explicitly account for transaction costs, turnover, and risk metrics. A key contribution of the paper lies in the explicit separation between directional signals and volatility regime control. Directional models generate return signals, while volatility models operate as regime filters (risk-on/risk-off), regulating market exposure according to the predicted risk environment. The results indicate that purely directional or purely volatility-based strategies display limited economic performance when used in isolation. Conversely, combined approaches — under parsimonious specifications and intermediate volatility thresholds — yield more stable return paths and substantially lower drawdowns, although not necessarily higher absolute returns than a buy-and-hold benchmark. The findings suggest that economic gains arise less from model complexity per se and more from the disciplined integration of return and risk forecasts. By focusing on recurrent and median volatility regimes rather than extreme market events, the study provides a realistic assessment of the strengths and limitations of quantitative strategies applied to VALE3, while leaving open avenues for future research on the explicit modeling of extreme volatility regimes.

Keywords: financial time series; volatility modeling; GARCH; VAR; machine learning; quantitative trading strategies; iron ore; VALE3.

SUMÁRIO

1 INTRODUÇÃO	10
2 REVISÃO DE LITERATURA E METODOLOGIA	12
2.1 Retornos financeiros: fatos estilizados, não-normalidade e previsibilidade limitada	12
2.2 Séries temporais e dependência dinâmica: da modelagem univariada ao VAR.....	13
2.3 Commodities e ações: minério de ferro como variável explicativa de mineradoras.....	14
2.4 Volatilidade condicional e modelos ARCH/GARCH: estrutura do risco e regimes	15
2.5 Estratégias de regime: filtros de volatilidade e timing de risco	16
2.6 Aprendizado de máquina em finanças: Random Forest, LSTM e armadilhas empíricas.....	16
2.6.1 Random Forest: robustez e não linearidade via ensembles	17
2.6.2 LSTM: dependência temporal longa e estrutura recorrente	17
2.6.3 Armadilhas: leakage, overfitting, não estacionariedade e custo de timing.....	18
2.7 Síntese e implicações para a metodologia: direcional + regime como arquitetura robusta.....	18
3 METODOLOGIA	19
3.1 Delineamento da pesquisa e estratégia empírica	19
3.2 Base de dados, variáveis e construção das séries	20
3.2.1 Fonte de dados e recorte temporal.....	20
3.2.2 Retornos e variáveis derivadas	20
3.3 Testes preliminares e diagnóstico das séries	21
3.3.1 Estacionariedade: teste ADF	21
3.3.2 Autocorrelação e dependência temporal: Ljung-Box	21
3.3.3 Heterocedasticidade condicional: ARCH-LM.....	21
3.4 Modelos econométricos: VAR e GARCH	21
3.4.1 Modelo VAR: especificação e seleção de defasagens	21

3.4.2 Causalidade temporal: teste de Granger	22
3.4.3 Modelo GARCH: volatilidade condicional e regimes	22
3.5 Modelos de aprendizado de máquina: Random Forest e LSTM	23
3.5.1 Random Forest (RF).....	23
3.5.2 LSTM.....	23
3.6 Validação fora da amostra: walk-forward e expanding window	23
3.6.1 Motivação: evitar overfitting e replicar decisão real	23
3.6.2 Esquema expanding window	24
3.6.3 Métricas de avaliação preditiva	24
3.7 Construção das estratégias de negociação	24
3.7.1 Princípio geral: modelos como geradores de sinal	24
3.7.2 Regra de regime por quantis ($q=0,30$ e $q=0,60$).....	25
3.7.3 Estratégias combinadas: direcional + regime	25
3.8 Custos de transação, turnover e fricções operacionais.....	25
3.9 Métricas econômicas e risco: desempenho, drawdown e Sharpe.....	26
3.10 Ambiente computacional e bibliotecas utilizadas	26
3.12 Síntese metodológica e organização dos resultados	28
4 Resultados e discussão	28
4.1 Modelagem VAR (Vetores Autorregressivos)	28
4.1.1 Funções Impulso–Resposta (IRF).....	29
4.1.2 Decomposição da variância do erro de previsão (FEVD).....	31
4.1.3 Testes de causalidade de Granger	33
4.1.4 VAR como ferramenta preditiva: previsões fora da amostra (OOS).....	35
4.1.5 Síntese e implicações para as próximas etapas.....	37
4.2 Modelagem da volatilidade condicional: GARCH.....	38
4.2.1 Testes preliminares pré-GARCH (VALE3 e Minério).....	38
4.2.2 Estimação in-sample do modelo AR(1)–GARCH(1,1)	40
4.2.3 Robustez: comparação entre distribuições Normal e Student-t.....	41
4.2.4 Avaliação fora da amostra: previsão de volatilidade via expanding window	41
4.2.5 Conexão com as estratégias	44

4.3 Modelos de Machine Learning: Random Forest, LSTM e custo computacional.....	44
4.3.1 Random Forest com reestimação diária (refit a cada t).....	45
4.3.2 Random Forest com reestimação periódica (refit a cada 200 observações).....	48
4.3.3 Síntese e conexão com a etapa de estratégias.....	49
4.3.4 Extensão com LSTM (previsão de volatilidade).....	50
4.4 Etapa de estratégias e avaliação econômica dos modelos.....	51
4.4.1 Estratégias baseadas em modelos VAR.....	53
4.4.2 Estratégias baseadas em regimes de volatilidade (GARCH)	57
4.4.3 Estratégias baseadas em Random Forest	63
4.4.4 Estratégias de Trading com LSTM (regimes de volatilidade).....	71
4.4.5 Estratégias híbridas: integração entre sinal direcional e regime de volatilidade	75
4.5 Encerramento da seção de Resultados e Discussão	83
5 Conclusão.....	84
6 Referências	86

1 INTRODUÇÃO

A modelagem de retornos financeiros ocupa posição central na literatura de finanças empíricas e quantitativas (TSAY, 2010), tanto do ponto de vista acadêmico quanto aplicado. Desde os trabalhos clássicos sobre eficiência de mercado (FAMA, 1970), sabe-se que a previsibilidade do retorno médio em horizontes curtos é limitada, sobretudo para ativos líquidos e amplamente acompanhados pelo mercado (MALKIEL, 2019). Em contrapartida, há evidência robusta de que o risco — entendido como variância condicional dos retornos — apresenta estrutura, persistência e alternância de regimes ao longo do tempo (CONT, 2005). Essa assimetria entre dificuldade de prever retorno e maior previsibilidade do risco constitui um ponto de partida fundamental para abordagens quantitativas modernas.

No caso de empresas intensivamente expostas a commodities, como a Vale S.A., esse desafio assume contornos ainda mais interessantes. A VALE3 reflete simultaneamente fatores idiossincráticos do mercado acionário, condições macroeconômicas globais e, de forma particularmente relevante, a dinâmica do preço internacional do minério de ferro. Existe, portanto, uma relação econômica intuitiva entre a commodity e o desempenho da ação, que representou em 2024 e 2023, respectivamente, 65% e 66% da receita operacional (VALE S.A., 2024). No entanto, essa relação não é estável nem mecânica: choques financeiros, mudanças de regime e períodos de elevada volatilidade podem dominar a precificação no curto prazo, reduzindo a utilidade de modelos puramente direcionais ou excessivamente estruturais.

Nesse contexto, este trabalho investiga a possibilidade de previsão e a utilização econômica de oscilações dos retornos diários da VALE3, utilizando como horizonte de tempo da base dados o período de janeiro de 2011 à dezembro de 2024 a partir de uma abordagem integrada, que combina econometria de séries temporais e métodos de aprendizado de máquina. Em vez de buscar previsões pontuais altamente precisas do retorno — objetivo reconhecidamente difícil e, muitas vezes, irrealista — a análise concentra-se em três eixos complementares: identificar relações dinâmicas economicamente plausíveis entre os retornos da VALE3 e do minério de ferro; modelar a dinâmica da volatilidade condicional e explorar regimes de risco; e avaliar se essas previsões podem ser transformadas em estratégias quantitativas

economicamente viáveis, considerando custos de transação, drawdowns e estabilidade fora da amostra.

Do ponto de vista metodológico, o trabalho utiliza modelos econométricos clássicos, como Vetores Autorregressivos (SIMS, 1980) e GARCH (ENGLE, 1982), ao lado de técnicas de aprendizado de máquina, em especial Random Forest (BREIMAN, 2001) e redes neurais recorrentes do tipo LSTM (HOCHREITER; SCHMIDHUBER, 1997). Essa escolha não se fundamenta em uma oposição entre “modelos tradicionais” e “modelos modernos”, mas sim na compreensão de que diferentes ferramentas capturam aspectos distintos do problema. Enquanto o VAR permite analisar relações dinâmicas multivariadas de forma disciplinada e interpretável, e o GARCH modela explicitamente a heterocedasticidade condicional, métodos de ML são capazes de capturar não linearidades e padrões complexos que escapam a estruturas lineares simples.

Um elemento central do trabalho é a separação explícita entre sinal direcional e controle de regime de risco. Modelos direcionais (econométricos ou de ML) são utilizados para indicar a direção esperada do retorno, enquanto modelos de volatilidade atuam como filtros de regime, regulando a exposição da estratégia conforme o ambiente de risco. Essa arquitetura reconhece uma limitação fundamental dos mercados financeiros: mesmo quando há algum poder preditivo direcional, erros de timing em períodos de alta volatilidade tendem a gerar perdas desproporcionais. Assim, controlar quando operar pode ser tão ou mais importante do que tentar prever corretamente o sinal do retorno.

A avaliação empírica é conduzida de forma rigorosa fora da amostra, utilizando esquemas de validação walk-forward com janelas expansivas, de modo a reproduzir o processo real de decisão em tempo de mercado. Todas as estratégias são avaliadas não apenas em termos de retorno bruto, mas também considerando retorno líquido, volatilidade, índice de Sharpe, drawdown máximo, turnover e custos médios de execução. O benchmark adotado ao longo do trabalho é a estratégia Long 100% (buy-and-hold), permitindo comparar diretamente os ganhos e perdas associados à introdução de regras quantitativas.

É importante destacar que o foco deste trabalho não está na modelagem de eventos extremos ou crises raras. Embora esses episódios sejam relevantes do ponto de vista econômico, o objetivo central é analisar a volatilidade recorrente e “mediana”

da VALE3, bem como sua interação com o preço do minério de ferro em regimes normais de mercado. A utilização de limiares de volatilidade baseados em quantis intermediários reflete essa escolha metodológica: busca-se compreender e explorar regimes de risco frequentes e persistentes, e não ancorar a análise exclusivamente em caudas extremas. Ainda assim, a possibilidade de extensão do arcabouço para limiares mais elevados e para o estudo explícito de eventos extremos é reconhecida como uma direção natural para pesquisas futuras.

Em síntese, o trabalho propõe uma abordagem integrada e empiricamente disciplinada para analisar a previsibilidade limitada dos retornos da VALE3, combinando fundamentação econômica, modelagem estatística e avaliação econômica realista. Ao fazê-lo, contribui para aplicabilidade e papel de modelos econométricos e de aprendizado de máquina em finanças, destacando as limitações amplamente destacadas na literatura.

2 REVISÃO DE LITERATURA E METODOLOGIA

2.1 Retornos financeiros: fatos estilizados, não-normalidade e previsibilidade limitada

A modelagem econométrica e quantitativa de retornos financeiros parte de um ponto quase consensual: retornos de ativos não se comportam como processos gaussianos simples. Embora modelos lineares e hipóteses de normalidade sejam úteis como aproximações iniciais, séries financeiras exibem propriedades empíricas recorrentes, conhecidas como fatos estilizados, que desafiam suposições clássicas de independência, homocedasticidade e distribuição Normal. Entre essas regularidades, destacam-se: distribuição com caudas pesadas e excesso de curtose, indicando maior frequência de eventos extremos do que aquela prevista por uma Normal; (assimetria e não-linearidades na dinâmica de risco; e, principalmente, volatility clustering, isto é, a tendência de que períodos de alta volatilidade sejam seguidos por períodos igualmente voláteis, e períodos tranquilos persistam por algum tempo (CONT, 2005).

Do ponto de vista econômico, essa estrutura tem implicações importantes. A previsibilidade do retorno médio é tipicamente limitada, sobretudo em horizontes curtos e para ativos líquidos, o que é compatível com a noção de eficiência de

mercado — ao menos em sua forma fraca, que sugere que a informação contida nos preços passados não deveria gerar lucros sistemáticos após custos. Em contrapartida, a volatilidade apresenta dependência temporal clara: mesmo que não seja trivial prever o sinal do retorno do dia seguinte, muitas vezes é possível prever algo sobre o “estado” do mercado, em termos de risco e instabilidade. Em outras palavras: o retorno é difícil, o risco é estruturado. Esse é o terreno onde modelos de volatilidade condicional e filtros de regime se tornam ferramentas relevantes.

Esse diagnóstico orienta a estratégia metodológica deste trabalho: em vez de insistir em um ideal irrealista de “prever retornos” com alta precisão, a análise busca identificar relações dinâmicas e mecanismos de transmissão entre variáveis (como no VAR), ao mesmo tempo em que utiliza instrumentos estatísticos e computacionais para capturar a dinâmica da volatilidade e explorar regimes de risco (como no GARCH e nos modelos de aprendizado de máquina).

2.2 Séries temporais e dependência dinâmica: da modelagem univariada ao VAR

A literatura de séries temporais oferece uma base formal sólida para a modelagem de dependência dinâmica em dados financeiros. Modelos autorregressivos e de média móvel, AR, MA e ARMA, são frequentemente adotados como ponto de partida por sua simplicidade, interpretabilidade e estrutura probabilística bem definida. Em mercados financeiros, tais modelos tendem a apresentar capacidade limitada para capturar retorno médio, mas continuam relevantes por dois motivos: podem remover autocorrelação residual na equação da média e servem como referência metodológica para extensões multivariadas e para a obtenção de resíduos adequados em testes (ENDERS, 2010).

Quando o objetivo é tratar mais de uma série simultaneamente e permitir que choques em uma variável afetem a trajetória da outra ao longo do tempo, modelos multivariados tornam-se mais apropriados. Nesse contexto, o modelo de Vetores Autorregressivos (VAR) constitui uma abordagem central por permitir que cada variável dependa de suas próprias defasagens e das defasagens das demais variáveis do sistema, sem impor restrições estruturais fortes. Trata-se de um modelo particularmente adequado para estudos em finanças e macrofinanças, nos quais a

interdependência entre preços, retornos e fatores externos (como commodities) é frequentemente bidirecional e dinâmica.

Além disso, o VAR permite formalizar hipóteses sobre precedência temporal e capacidade preditiva de variáveis via testes de causalidade no sentido de Granger. Importa ressaltar que “causalidade de Granger” não implica causalidade estrutural no sentido estrito; trata-se de um conceito estatístico: uma variável X “Granger-causa” Y se a inclusão de defasagens de X melhora a previsão de Y relativamente a um modelo que utiliza apenas defasagens de Y (ENDERS, 2010). Apesar dessa limitação, o teste é extremamente útil como instrumento empírico para detectar relações dinâmicas e orientar especificações, principalmente quando combinado com critérios de seleção de defasagens (AIC/BIC) e diagnóstico de estabilidade.

Nesse trabalho, o VAR é utilizado não como uma promessa de previsibilidade perfeita, mas como um mecanismo econométrico disciplinado para avaliar relações temporais entre retornos do minério de ferro e retornos da VALE3, oferecendo uma ponte entre fundamentação econômica e modelagem estatística.

2.3 Commodities e ações: minério de ferro como variável explicativa de mineradoras

O comportamento de empresas intensivamente expostas ao ciclo de commodities tende a refletir de forma significativa os preços internacionais de seus produtos. No caso de companhias mineradoras, a relação entre preço do minério de ferro, ciclo global e desempenho de ações costuma ser economicamente intuitiva: variações persistentes na commodity afetam receitas esperadas, margens operacionais, fluxo de caixa e, por consequência, a precificação do ativo via expectativas dos agentes. Em termos gerais, trata-se de um canal clássico de transmissão do preço do bem primário para o valor presente de lucros futuros.

Ainda assim, esse vínculo não é mecânico nem estável no tempo. A precificação de ações incorpora não apenas o preço spot da commodity, mas também expectativas sobre demanda global (principalmente China), taxas de câmbio relevantes, estrutura de custos, endividamento e fatores de risco próprios do mercado acionário. Em períodos de estresse, regimes de risco podem dominar a dinâmica de preços, fazendo com que choques financeiros “sobreponham” fundamentos de curto prazo. Logo, mesmo quando existe relação econômica plausível entre minério e

VALE3, é razoável supor que essa relação é parcialmente mediada por regimes de volatilidade e condições de mercado.

A modelagem multivariada via VAR, aliada a testes de estacionariedade e seleção criteriosa de defasagens, fornece uma estratégia empiricamente defensável para investigar essa conexão sem impor relações estruturais rígidas. Essa fundamentação é crucial para que, na etapa de estratégias, sinais econométricos sejam gerados de maneira compatível com a literatura e com o arcabouço estatístico clássico.

2.4 Volatilidade condicional e modelos ARCH/GARCH: estrutura do risco e regimes

A principal contribuição da família ARCH/GARCH é capturar uma das propriedades mais robustas de séries financeiras: a variância condicional não é constante e apresenta forte dependência temporal. Engle (1982) introduziu o modelo ARCH como uma solução para heterocedasticidade condicional, permitindo que a variância dependa de choques passados. Posteriormente, Bollerslev (1986) generalizou esse conceito por meio do modelo GARCH, no qual a variância condicional depende tanto de choques anteriores (componente ARCH) quanto de sua própria trajetória passada (componente GARCH). Essa estrutura simples, porém poderosa, produz previsões de volatilidade que tendem a acompanhar períodos de turbulência e calma, aproximando-se bem do fenômeno de *volatility clustering* documentado empiricamente (ENGLE, 1982; BOLLERSLEV, 1986; CONT, 2001).

Em aplicações financeiras, GARCH não deve ser interpretado apenas como “um modelo estatístico”. Ele possui leitura econômica direta: a volatilidade condicional é uma proxy operacional para risco percebido. Quanto maior a volatilidade prevista, maior a incerteza e, frequentemente, maior o prêmio de risco exigido pelo mercado. Desse modo, previsões de σ podem ser utilizadas para filtros de regime, no sentido de reduzir exposição em estados de maior risco e ampliar exposição em estados de menor risco.

Um ponto adicional relevante é que, embora retornos frequentemente exibam caudas pesadas, nem sempre a substituição da Normal por distribuições mais flexíveis (como Student-t) melhora de forma relevante o modelo operacional. A distribuição Student-t incorpora um parâmetro de graus de liberdade v que controla a espessura

das caudas; quando v é elevado, a distribuição aproxima-se da Normal, e a suposta vantagem de capturar extremos se dissolve. Essa verificação de robustez é importante: ela evita justificar complexidade adicional sem ganho empírico correspondente — algo particularmente valorizado em trabalhos acadêmicos, em que parcimônia e estabilidade paramétrica são critérios centrais.

Nesse trabalho, GARCH(1,1) com distribuição Normal é adotado como especificação de referência por sua estabilidade e ampla aceitação como baseline em finanças empíricas, ao passo que alternativas são tratadas como robustez, não como regra.

2.5 Estratégias de regime: filtros de volatilidade e timing de risco

Modelos de volatilidade condicional são especialmente úteis quando incorporados à lógica de decisão de carteira e trading. A intuição é simples e, ao mesmo tempo, sofisticada: não é necessário acertar o retorno para controlar risco. Estratégias de volatility timing e filtros de regime exploram o fato de que drawdowns severos e perdas persistentes tendem a ocorrer em períodos de volatilidade elevada, quando o mercado opera com elevada incerteza e maior densidade de eventos extremos.

Nesse contexto, a utilização de quantis da volatilidade prevista (por exemplo, limiares associados a $q = 0,30$ e $q = 0,60$) cria regras sistemáticas de exposição: períodos de baixa volatilidade implicam risk-on e maior disposição a manter posição; períodos de maior volatilidade implicam risk-off, com exposição reduzida ou até neutralizada. Essa lógica não é um improviso, ela é, na verdade, um modo de formalizar o que agentes reais fazem na prática: diminuir risco quando o mercado está instável e aumentar exposição quando o regime é mais benigno.

Ao articular essa ideia com modelos direcionalmente orientados (econométricos ou de ML), obtém-se uma estrutura metodológica particularmente defensável: o modelo direcional tenta orientar a direção; o filtro de regime tenta mitigar a volatilidade. Essa separação de papéis melhora a consistência e reduz a dependência de previsibilidade perfeita, que raramente existe em séries financeiras.

2.6 Aprendizado de máquina em finanças: Random Forest, LSTM e armadilhas empíricas

A aplicação de aprendizado de máquina em previsão financeira ganhou destaque pela capacidade desses métodos de capturar não linearidades, interações

e padrões complexos em bases de dados extensas. Diferentemente de modelos lineares tradicionais, algoritmos como Random Forest e redes neurais recorrentes podem, ao menos em tese, extrair sinais não triviais de variáveis explicativas e de estruturas temporais longas. Contudo, finanças é um domínio adversarial para ML: não estacionariedade, mudanças de regime, ruído elevado e custos de transação tornam o problema substancialmente mais difícil do que aplicações típicas em visão computacional ou processamento de linguagem.

2.6.1 Random Forest: robustez e não linearidade via ensembles

O Random Forest, proposto por Breiman (2001), é um método de *ensemble* baseado em bagging e árvores de decisão aleatorizadas. Sua principal vantagem prática é combinar flexibilidade (captura relações não lineares e interações) com robustez contra overfitting, sobretudo quando comparado a modelos altamente parametrizados. Além disso, trata-se de uma ferramenta relativamente transparente do ponto de vista empírico: é possível avaliar importância de variáveis e compreender, em termos gerais, quais fatores exercem maior influência preditiva.

Em finanças, Random Forest costuma performar bem como gerador de sinal quando existem variáveis explicativas com alguma estrutura informacional. Ainda assim, o desempenho deve ser avaliado com cautela: uma alta correlação preditiva ou acurácia direcional não implica, automaticamente, um resultado econômico superior, uma vez que custos, turnover e atraso na incorporação de informação podem corroer ganhos. Logo, Random Forest é tratado aqui como componente de sinal — não como garantia de alfa persistente.

2.6.2 LSTM: dependência temporal longa e estrutura recorrente

Redes neurais recorrentes com unidades LSTM (Long Short-Term Memory), introduzidas por Hochreiter e Schmidhuber (1997), foram desenvolvidas para resolver limitações de memória de redes recorrentes clássicas, permitindo capturar dependências temporais longas. Isso é particularmente relevante em séries financeiras, nas quais padrões podem depender de estruturas de longo prazo, sazonalidades e alternâncias de regime.

Entretanto, LSTM é notoriamente sensível a overfitting, escolha de hiperparâmetros e instabilidade quando aplicado a séries com baixo sinal-ruído. Por esse motivo, neste trabalho, LSTM é utilizado de forma metodologicamente

conservadora: não como “previsor direto de retorno”, mas como previsor de uma proxy de volatilidade/instabilidade ($|r|$), operando como filtro de regime. Essa escolha privilegia robustez e coerência com o restante da metodologia.

2.6.3 Armadilhas: leakage, overfitting, não estacionariedade e custo de timing

A literatura aplicada a quantitative trading enfatiza que resultados preditivos não garantem resultados econômicos. Há armadilhas recorrentes: data leakage, viés de seleção, hiperparametrização, validação inadequada e, sobretudo, ausência de backtesting realista. Um modelo pode apresentar bom desempenho estatístico dentro da amostra e falhar ao ser exposto a novos regimes. Além disso, mesmo quando acerta o sinal, pode fazê-lo com atraso: em mercados financeiros, acertar depois não é o mesmo que acertar — timing é parte do sinal. Essa distinção é crucial em abordagens de ML, especialmente quando maior complexidade computacional implica maior latência operacional ou decisões mais “suavizadas”, o que pode gerar custo de oportunidade e perda de pontos de entrada/saída.

Por esse motivo, o trabalho adota validação fora da amostra em esquema walk-forward e expanding window, aproximando o processo real de decisão e minimizando o risco de conclusões artificiais. Assim, aprendizado de máquina aqui é tratado com disciplina empírica e prudência metodológica: uma ferramenta útil quando bem delimitada, não um “milagre computacional”.

2.7 Síntese e implicações para a metodologia: direcional + regime como arquitetura robusta

A revisão apresentada sugere uma conclusão lógica: em finanças, insistir em previsibilidade perfeita de retornos é um erro metodológico recorrente. Retornos são ruidosos, regimes mudam, e custos operacionais importam. Em contrapartida, volatilidade é estruturada e persistente, podendo ser modelada com ferramentas estatísticas consolidadas.

Diante disso, a arquitetura mais defensável — tanto do ponto de vista teórico quanto empírico — é combinar modelos direcionais (econométricos ou de ML) com filtros de regime baseados em volatilidade. Essa combinação respeita a natureza do problema: (i) o sinal direcional tenta identificar a direção provável do movimento; (ii) o componente de regime regula exposição em função do estado do mercado, reduzindo vulnerabilidade em períodos de stress.

Essa lógica orienta o desenho metodológico e a seção de estratégias do trabalho: primeiro avaliam-se modelos individualmente (VAR, GARCH, Random Forest e LSTM), e posteriormente constroem-se estratégias puras e combinações (como RF+GARCH e VAR+GARCH) sob um framework comparável e replicável. Dessa forma, o trabalho mantém consistência entre fundamentação teórica, modelagem estatística e avaliação econômica, permitindo que os resultados sejam interpretados com clareza e defendidos com firmeza metodológica.

3 METODOLOGIA

3.1 Delineamento da pesquisa e estratégia empírica

Este trabalho adota uma abordagem quantitativa aplicada, combinando econometria de séries temporais e métodos de aprendizado de máquina (ML) para análise preditiva e avaliação econômica de estratégias de negociação envolvendo a ação VALE3 e o preço internacional do minério de ferro. A escolha desse desenho metodológico responde a duas características fundamentais do problema: retornos financeiros exibem elevado ruído e baixa previsibilidade direta na média; em contrapartida, a volatilidade apresenta dependência temporal, persistência e alternância de regimes, podendo ser modelada de forma mais estruturada.

A estratégia empírica foi construída em três etapas principais, organizadas de maneira cumulativa:

1. Modelagem e testes econométricos iniciais: identificação das propriedades das séries (estacionariedade, autocorrelação e heterocedasticidade condicional), testes de causalidade temporal e estimação de modelos econométricos baseline (VAR e GARCH).
2. Modelos de previsão e validação fora da amostra: implementação de previsões em esquema walk-forward com expanding window, reproduzindo o processo real de aprendizagem em tempo.
3. Construção e avaliação de estratégias de negociação: (a) estratégias puras baseadas nos modelos individualmente; e (b) estratégias combinadas do tipo “direcional + regime de volatilidade”, com custos de transação e métricas de risco-retorno.

Esse desenho assegura coerência entre fundamentação teórica, especificação estatística e avaliação econômica, evitando tanto (i) o erro comum de avaliar modelos apenas por métricas estatísticas (ignorando custos e dinâmica de execução) quanto (ii) o risco de construir estratégias com regras ad hoc sem sustentação metodológica.

3.2 Base de dados, variáveis e construção das séries

3.2.1 Fonte de dados e recorte temporal

Os dados utilizados foram obtidos a partir do Yahoo Finance, contemplando preços diários ajustados, do período de janeiro de 2011 à dezembro de 2024 (quando aplicável) para:

- VALE3: série de preços e retornos da ação negociada na B3;
- Minério de ferro (USD): proxy internacional do preço do minério de ferro, em dólares;
- Taxa de câmbio BRL/USD: com objetivo de averiguação de transmissão de efeitos cambiais e conversões.

A análise utiliza frequência diária e horizonte temporal amplo, permitindo capturar diferentes regimes de mercado: períodos de alta estabilidade, choques de commodities, estresse financeiro e choques exógenos relevantes. O recorte final é determinado pela interseção de datas disponíveis entre as séries e pelos tratamentos necessários para construção de retornos.

3.2.2 Retornos e variáveis derivadas

A variável central de modelagem é o retorno logarítmico diário, definido por:

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

onde P_t representa o preço do ativo no tempo t . Essa transformação é padrão em finanças empíricas por oferecer melhor comportamento estatístico do que retornos simples em muitas aplicações, além de facilitar agregação temporal.

Além dos retornos, são construídas variáveis derivadas relevantes para a modelagem da volatilidade e para o backtesting:

- Volatilidade realizada (proxy):

$$|r_{t+1}|$$

utilizada como aproximação prática da volatilidade efetiva observada no período subsequente, conforme prática recorrente na literatura empírica.

- Séries previstas pelos modelos:

- previsão de retorno (quando aplicável);
- previsão de volatilidade condicional (GARCH);
- previsão de proxy de risco/regime (ML, especialmente LSTM).

3.3 Testes preliminares e diagnóstico das séries

Antes da estimação, são aplicados testes estatísticos para caracterização das séries e validação das hipóteses subjacentes aos modelos.

3.3.1 Estacionariedade: teste ADF

A estacionariedade dos retornos é avaliada por meio do teste Augmented Dickey-Fuller (ADF). A hipótese nula do teste é a presença de raiz unitária (não estacionariedade). Em geral, espera-se que séries de preços não sejam estacionárias, enquanto retornos costumam ser estacionários — condição essencial para aplicação de VAR em retornos.

3.3.2 Autocorrelação e dependência temporal: Ljung-Box

A dependência temporal da série é analisada via teste Ljung-Box aplicado aos retornos e aos retornos ao quadrado, com foco em:

- autocorrelação na média: relevante para especificações AR/ARMA;
- autocorrelação na variância: indicativa de clustering e heterocedasticidade condicional, motivando modelos ARCH/GARCH.

3.3.3 Heterocedasticidade condicional: ARCH-LM

A presença de efeito ARCH é testada por ARCH-LM. A rejeição da hipótese nula sugere que a variância condicional depende de choques passados, justificando uso de modelos GARCH para volatilidade.

Esses testes são fundamentais não por “provar” o modelo, mas por disciplinar a especificação: retornos podem ser imprevisíveis na média, mas a variância não é puramente aleatória.

3.4 Modelos econométricos: VAR e GARCH

3.4.1 Modelo VAR: especificação e seleção de defasagens

A interação dinâmica entre retorno do minério e retorno da VALE3 é modelada via Vetores Autorregressivos (VAR). Seja y_t um vetor com duas variáveis:

$$y_t = \begin{bmatrix} r_t^{iron} \\ r_t^{vale} \end{bmatrix}$$

O VAR(p) é especificado como:

$$y_t = c + A_1 y_{t-1} + A_2 y_{t-2} + \dots + A_p y_{t-p} + u_t$$

A seleção do número ótimo de defasagens p é realizada com base em critérios de informação, em especial AIC e/ou BIC, buscando balancear ajuste e parcimônia. Em aplicações financeiras, defasagens excessivas frequentemente geram modelos instáveis e pouca robustez fora da amostra, de modo que a seleção criteriosa é central.

3.4.2 Causalidade temporal: teste de Granger

O teste de causalidade no sentido de Granger avalia se o retorno do minério contém informação incremental para prever o retorno da VALE3. A interpretação é estatística e não estrutural: rejeitar a hipótese nula sugere precedência temporal informacional, fornecendo base para motivar o uso do minério como variável explicativa no VAR.

3.4.3 Modelo GARCH: volatilidade condicional e regimes

A volatilidade condicional dos retornos da VALE3 é modelada via GARCH(1,1), partindo da equação da média com termo AR(1):

$$r_t = \mu + \phi r_{t-1} + \epsilon_t, \quad \epsilon_t = \sigma_t z_t$$

com $z_t \sim N(0, 1)$ na especificação principal.

A equação da variância condicional é dada por:

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

A interpretação econômica dos parâmetros é direta:

- α : impacto imediato de choques recentes (componente ARCH);
- β : persistência da volatilidade (componente GARCH);
- $\alpha + \beta$: persistência total e intensidade de clustering.

Como verificação de robustez, estima-se também a versão com distribuição Student-t para os resíduos. Contudo, a especificação baseline Normal é adotada como principal devido à maior estabilidade e melhor desempenho segundo critérios de informação no contexto analisado.

3.5 Modelos de aprendizado de máquina: Random Forest e LSTM

A etapa de ML tem como objetivo explorar estruturas preditivas não lineares e dependência temporal longa, de maneira complementar aos modelos econométricos.

3.5.1 Random Forest (RF)

O Random Forest é utilizado como método de ensemble baseado em árvores de decisão. Sua vantagem está na capacidade de capturar não linearidades e interações entre variáveis explicativas, com mitigação parcial de overfitting via bagging e aleatorização.

No presente trabalho, o RF foi implementado em duas versões:

- RF principal (refit contínuo): reestimação recorrente, simulando atualização constante do modelo conforme chegada de novas observações.
- RF periódico (*refit at 200*): reestimação a cada 200 observações, reduzindo custo computacional e avaliando robustez a menor frequência de atualização.

As previsões do RF são utilizadas de forma operacional em estratégias (principalmente como sinal direcional), mas com disciplina empírica: um resultado estatístico elevado não é tratado como suficiente sem validação econômica.

3.5.2 LSTM

A LSTM é utilizada como rede neural recorrente capaz de capturar dependência temporal de longo prazo. A aplicação neste trabalho é deliberadamente conservadora: em vez de utilizar LSTM como previsor direto do retorno, ela é utilizada prioritariamente como modelo de regime, prevendo proxy de risco (ligada à magnitude do retorno/volatilidade).

Essa escolha reduz a sensibilidade a ruído, melhora estabilidade fora da amostra e aumenta coerência metodológica com o objetivo central de filtrar períodos de maior instabilidade.

3.6 Validação fora da amostra: walk-forward e expanding window

3.6.1 Motivação: evitar overfitting e replicar decisão real

A validação fora da amostra é um elemento central deste trabalho, pois resultados in-sample em finanças frequentemente refletem ajuste ao ruído (*overfitting*) e não capacidade genuína de generalização. Para mitigar esse risco, adota-se um

esquema de previsão em tempo real do tipo *walk-forward*, com janela expandida (*expanding window*).

3.6.2 Esquema expanding window

No esquema expanding window:

1. define-se um tamanho mínimo de amostra inicial para estimação;
2. estima-se o modelo;
3. gera-se previsão para o período seguinte ($t+1$);
4. adiciona-se a nova observação ao conjunto de treino;
5. repete-se o processo até o final da amostra.

Esse procedimento é aplicado consistentemente nos modelos econométricos e de ML, garantindo comparabilidade e evitando uso indevido de informação futura.

3.6.3 Métricas de avaliação preditiva

A avaliação estatística dos modelos inclui:

- Correlação entre previsão de volatilidade e proxy realizada:

$$\text{Corr}(\hat{\sigma}_{t+1|t}, |r_{t+1}|)$$

- Erros RMSE e MAE para previsões de volatilidade (quando aplicável);
- Acurácia direcional (em modelos direcionais), definida como percentual de acerto do sinal do retorno.

Contudo, ressalta-se que o foco do trabalho não é maximizar métricas estatísticas isoladas, mas avaliar implicações econômicas por meio de estratégias e backtesting com custos.

3.7 Construção das estratégias de negociação

3.7.1 Princípio geral: modelos como geradores de sinal

A lógica de estratégias é estruturada como transformação de previsões em regras sistemáticas de posição. São utilizadas três classes de estratégias:

1. Estratégias puras: sinal derivado diretamente do modelo (ex.: regime de volatilidade por GARCH).
2. Estratégias de regime: exposição controlada por quantis de volatilidade prevista.
3. Estratégias combinadas: arquitetura “direcional + regime”, separando previsão de direção (retorno) e controle de risco (volatilidade).

3.7.2 Regra de regime por quantis (q=0,30 e q=0,60)

A definição operacional de regime utiliza quantis da volatilidade prevista:

$$\sigma_{th}(q) = Q_q(\hat{\sigma}_{t+1|t})$$

onde Q_q é o quantil q. A regra base:

- se $\hat{\sigma}_{t+1|t} \leq \sigma_{th}(q)$: mercado em regime benigno (risk-on), maior exposição;
- caso contrário: regime adverso (risk-off), redução de exposição.

Foram utilizados principalmente dois limiares:

- q=0,3: regime mais conservador (menor fração investida);
- q=0,6: regime mais permissivo (maior fração investida).

3.7.3 Estratégias combinadas: direcional + regime

A arquitetura central proposta é:

- componente direcional (VAR ou RF) gera sinal de compra/venda/posição;
- componente regime (GARCH ou LSTM) filtra períodos de risco.

Formalmente, a posição é definida como:

$$pos_t = signal_t \times I(\hat{\sigma}_{t+1|t} \leq \sigma_{th})$$

onde $I(\cdot)$ é indicador do regime.

Essa separação é metodologicamente sólida, pois reconhece que:

- prever direção é difícil, mas pode haver algum sinal;
- controlar risco via regime é mais consistente e estatisticamente defensável.

3.8 Custos de transação, turnover e fricções operacionais

Uma contribuição essencial do trabalho é incorporar fricções reais no backtest. Em aplicações financeiras, performance bruta sem custos tende a ser ilusória, especialmente em estratégias com alto turnover.

Define-se turnover diário como:

$$TO_t = |pos_t - pos_{t-1}|$$

e o custo proporcional:

$$cost_t = c \cdot TO_t$$

onde c representa custo médio por ajuste de posição (incluindo bid-ask, corretagem e slippage médio aproximado). O retorno líquido é:

$$r_t^{net} = r_t^{gross} - cost_t$$

Além disso, discute-se a ideia de custo de timing: mesmo quando o modelo acerta direção, o acerto pode ocorrer com atraso, diluindo o ganho econômico. Esse problema é particularmente relevante para modelos complexos e suavizados, como redes neurais, nos quais o sinal tende a reagir de forma menos instantânea aos choques.

3.9 Métricas econômicas e risco: desempenho, drawdown e Sharpe

A avaliação das estratégias inclui métricas econômicas tradicionais:

- retorno acumulado bruto e líquido;
- volatilidade anualizada:

$$\sigma = std(r)\sqrt{252}$$

- Índice de Sharpe:

$$Sharpe = \frac{E[r]}{std(r)}\sqrt{252}$$

- máximo drawdown, definido como pior queda relativa em relação ao pico do NAV.

Adicionalmente, reporta-se:

- percentual de tempo investido;
- percentual de períodos em short (quando aplicável);
- turnover médio;
- custo médio diário.

Esse conjunto de métricas permite avaliar simultaneamente desempenho, risco e viabilidade operacional.

3.10 Ambiente computacional e bibliotecas utilizadas

Toda a implementação empírica do presente trabalho foi conduzida em Python, com execução em ambiente Google Colab, utilizando armazenamento e leitura de dados via Google Drive. A escolha por esse ecossistema deve-se à sua ampla adoção na literatura aplicada em finanças quantitativas e econometria computacional, além de viabilizar reprodutibilidade e escalabilidade na estimação dos modelos e na simulação das estratégias.

O pipeline de modelagem foi estruturado em quatro blocos: coleta e tratamento de dados, (estimação econométrica e modelos de volatilidade, modelos de

aprendizado de máquina, e backtesting e geração de outputs (tabelas e gráficos). A seguir, descrevem-se as principais bibliotecas utilizadas e suas funções no processo:

- pandas e numpy: utilizadas para organização e manipulação de séries temporais, criação de variáveis (ex.: retornos logarítmicos), alinhamento temporal entre ativos e consolidação de bases in-sample e out-of-sample. Constituem o núcleo de processamento de dados do trabalho.
- yfinance: utilizada como fonte operacional para download automatizado de preços históricos e séries financeiras necessárias à estimação dos modelos e construção das estratégias.
- statsmodels: empregada na estimação econométrica, incluindo o ajuste de modelos VAR, seleção de defasagens por critérios de informação (AIC/BIC), geração de IRFs (Impulse Response Functions), FEVD (Forecast Error Variance Decomposition) e execução de testes auxiliares como causalidade de Granger e diagnósticos estatísticos.
- arch: utilizada para estimação do modelo AR(1)–GARCH(1,1), cálculo da variância condicional, diagnóstico dos resíduos e execução do procedimento de previsão fora da amostra (walk-forward/expanding window) aplicado à volatilidade.
- scikit-learn (sklearn): aplicada na modelagem de aprendizado de máquina com Random Forest, incluindo etapas de treino, validação e previsão fora da amostra, além de apoio à organização de métricas de desempenho e serialização de outputs em arquivos estruturados.
- TensorFlow / Keras: utilizada para implementação e treinamento da rede neural recorrente LSTM, bem como para geração das previsões fora da amostra empregadas como proxy operacional para regimes de volatilidade/risco e posterior integração com estratégias combinadas.
- matplotlib (e suporte complementar com seaborn em alguns trechos): utilizadas para visualização dos resultados, incluindo construção de gráficos temporais, dispersões e curvas de retorno acumulado das estratégias em comparação ao benchmark.
- joblib e json: empregados para persistência de objetos e outputs, de forma a garantir reprodutibilidade dos resultados sem necessidade de reestimação completa dos modelos, reduzindo riscos de alterações

numéricas decorrentes de reexecuções (especialmente relevantes em modelos estocásticos ou sensíveis a condições iniciais).

- os: utilizada para padronizar a leitura/escrita de outputs e organizar a estrutura de diretórios no Drive, permitindo que as rotinas funcionem de forma consistente em diferentes notebooks.
- tqdm: aplicada para acompanhamento do progresso computacional em rotinas de walk-forward, sobretudo nos exercícios fora da amostra que exigem reestimação iterativa em janelas móveis/expansíveis.

Em conjunto, o arcabouço computacional empregado permite que todo o conjunto de resultados do trabalho — modelos, previsões e estratégias — seja reproduzido de forma transparente, com rastreabilidade de parâmetros, outputs e critérios operacionais, mantendo consistência com o objetivo de avaliação comparativa entre abordagens econométricas e de aprendizado de máquina em contexto de previsão e aplicação em estratégias de negociação

3.12 Síntese metodológica e organização dos resultados

Em síntese, a metodologia adota uma estrutura escalonada:

1. diagnóstico e modelagem econométrica (VAR e GARCH);
2. implementação de ML com disciplina empírica (RF e LSTM);
3. validação fora da amostra com expanding window;
4. transformação de previsões em estratégias comparáveis;
5. avaliação econômica com custos e risco;
6. combinações metodologicamente sólidas (direcional + regime).

A seção de resultados (Capítulo 4) apresenta inicialmente os testes e modelos individualmente, e em seguida compara estratégias puras e estratégias combinadas, permitindo avaliar ganhos marginais de complexidade e robustez.

4 Resultados e discussão

4.1 Modelagem VAR (Vetores Autorregressivos)

Nesta etapa, estima-se um VAR bivariado com base nos retornos logarítmicos diários da VALE3 e do minério de ferro cotado em dólares (Iron_USD). A utilização de retornos, e não de preços, decorre tanto de considerações econométricas (redução de problemas de tendência e não estacionariedade, comprovada na seção seguinte

via teste ADF, mesmo sendo característica de retornos) quanto do foco do trabalho: estratégias quantitativas direcionais e avaliação preditiva em alta frequência. A seleção do número de defasagens é conduzida por critérios de informação, explorando-se duas especificações complementares: uma mais flexível, selecionada pelo critério de Akaike (AIC), e outra mais parcimoniosa, selecionada pelo critério Bayesiano (BIC). Em termos práticos, tal comparação é relevante: o AIC tende a favorecer modelos mais ricos (maior número de parâmetros), enquanto o BIC impõe penalização mais forte e tende a privilegiar estabilidade e generalização.

4.1.1 Funções Impulso–Resposta (IRF)

A análise das funções impulso–resposta tem como objetivo quantificar como um choque exógeno (inovação) no retorno do minério se propaga nos retornos da VALE3 ao longo dos períodos subsequentes. Os resultados indicam resposta positiva inicial da VALE3 a choques positivos no retorno do minério, concentrada sobretudo no curto prazo, o que é consistente com a intuição econômica: alterações inesperadas no preço internacional da commodity afetam expectativas de receita e margem de uma empresa altamente exposta ao minério, sendo rapidamente precificadas pelo mercado acionário.

No entanto, a persistência e a estabilidade dessa resposta variam conforme a especificação do VAR. No modelo selecionado pelo BIC ($p = 1$), a resposta impulsiva apresenta padrão concentrado nos primeiros períodos, com dissipação rápida e comportamento predominantemente monotônico. Já no modelo selecionado pelo AIC ($p = 7$), observa-se uma trajetória mais oscilatória ao longo do horizonte, refletindo maior flexibilidade do modelo em capturar dinâmicas complexas, mas também maior incerteza inferencial (bandas de confiança mais amplas). Em termos interpretativos, isso sugere que, embora exista um efeito dinâmico detectável, sua magnitude não é elevada e sua persistência é limitada, especialmente quando se impõe parcimônia na estrutura dinâmica.

Gráfico 1: IRF (AIC) — Choque no minério (USD) → Retornos da VALE3 | p=7

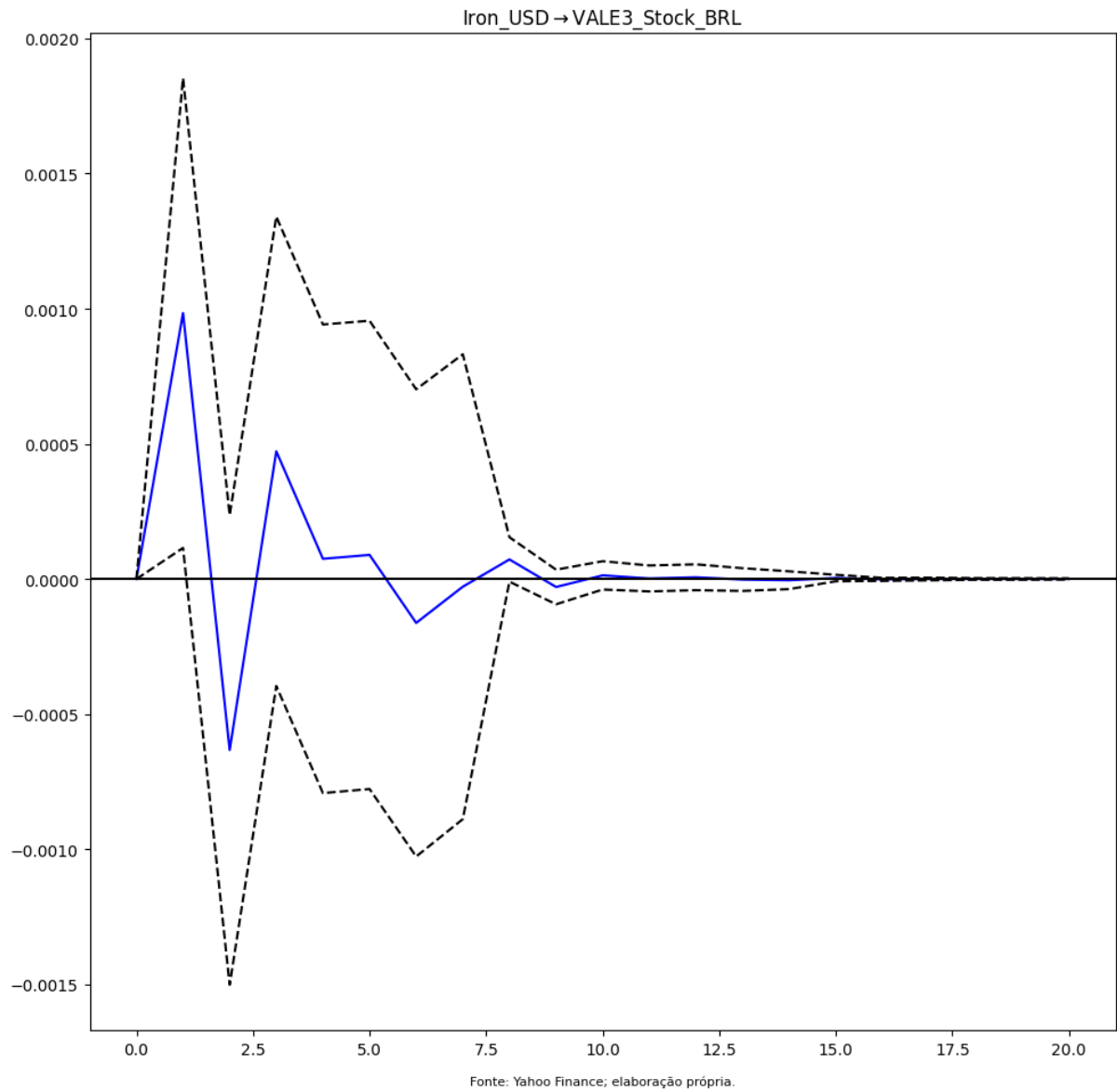
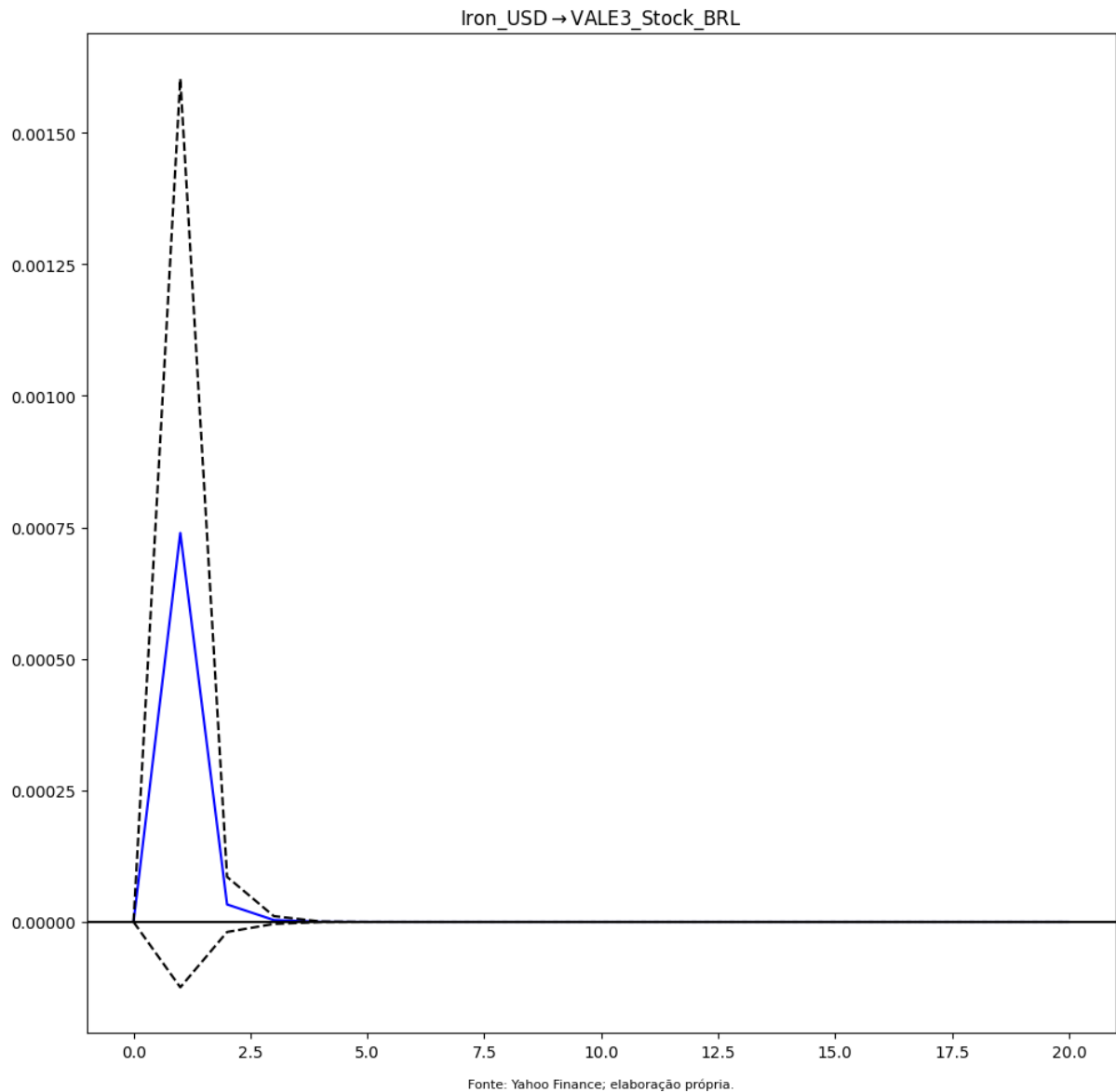


Gráfico 2: IRF (BIC) — Choque no minério (USD) → Retornos da VALE3 | $p=1$



4.1.2 Decomposição da variância do erro de previsão (FEVD)

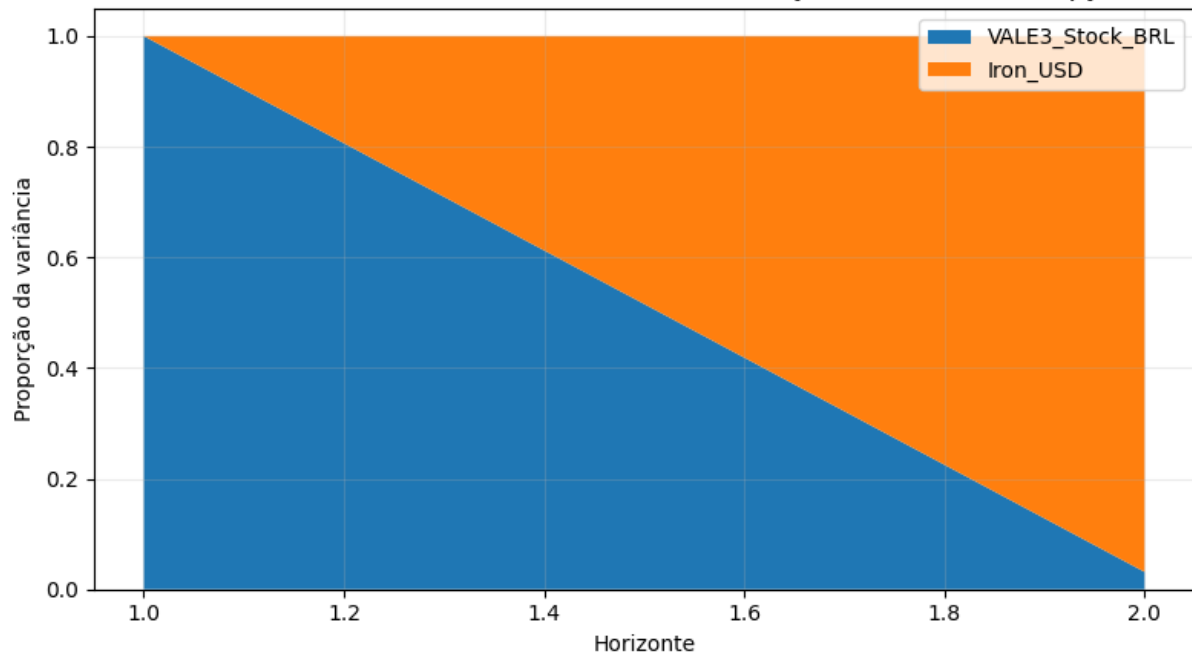
Como complemento às IRFs, aplica-se a decomposição da variância do erro de previsão (Forecast Error Variance Decomposition — FEVD) para avaliar em que medida choques no minério contribuem para explicar a variância dos erros de previsão dos retornos da VALE3 ao longo de diferentes horizontes. Em outras palavras, busca-se medir se o minério é apenas um fator contemporâneo ou se representa componente substantivo da dinâmica de incerteza da ação.

Os resultados mostram que, mesmo no curto prazo, a maior parte da variância do erro de previsão da VALE3 é atribuída a inovações da própria VALE3, enquanto

choques no minério explicam uma fração complementar. Esse achado é importante: ele é consistente com dois fatos estilizados bem documentados em finanças. Primeiro, retornos acionários em alta frequência são fortemente dominados por ruído, microestrutura e choques idiossincráticos (eventos corporativos, fluxos, sentimento, notícias), reduzindo substancialmente o poder explicativo de fundamentos únicos. Segundo, mesmo quando um fundamento macro/commodity é relevante, seu impacto tende a ocorrer via ajustes rápidos e curtos, o que limita seu peso em decomposições de variância.

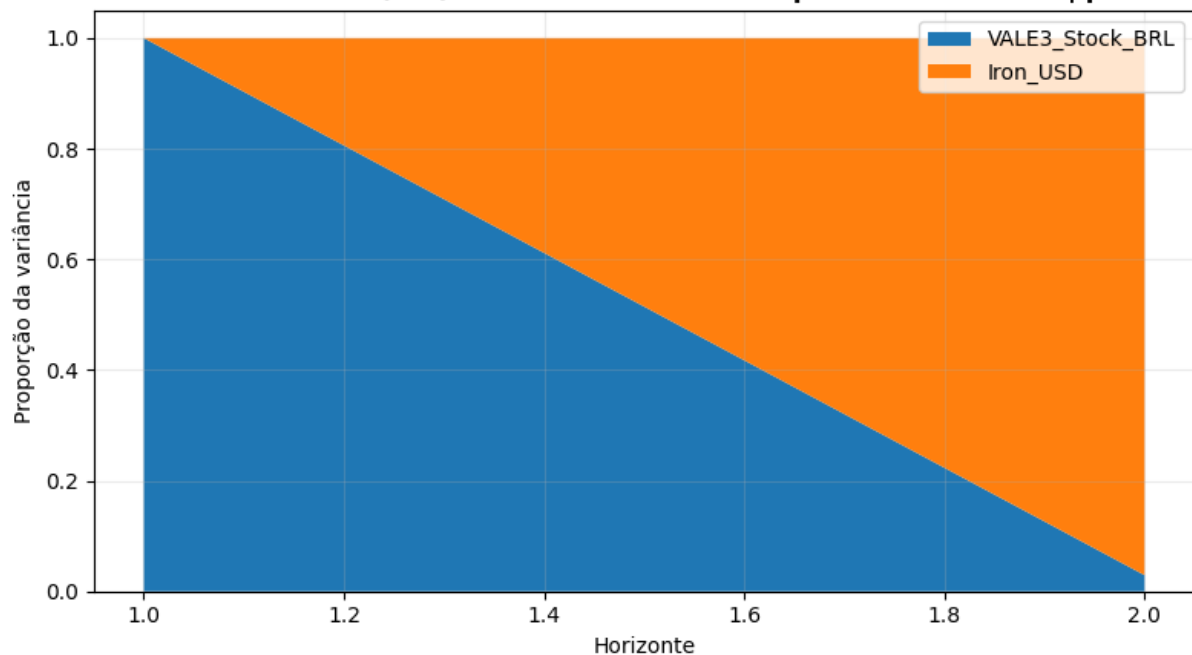
Assim, a FEVD reforça a leitura de que o minério é um determinante relevante e economicamente plausível, mas não dominante na variabilidade diária dos retornos da VALE3, o que impõe limites naturais à previsibilidade linear em horizontes muito curtos.

Gráfico 3: FEVD (AIC) — Variância do erro de previsão da VALE3 | $p=7$



Fonte: Yahoo Finance; elaboração própria.

Gráfico 4: FEVD (BIC) — Variância do erro de previsão da VALE3 | $p=1$



Fonte: Yahoo Finance; elaboração própria.

4.1.3 Testes de causalidade de Granger

Para avaliar formalmente se o minério contém conteúdo preditivo incremental sobre os retornos da VALE3 em uma estrutura linear, conduzem-se testes de causalidade de Granger. Sob essa abordagem, diz-se que o minério “Granger-causa”

VALE3 se suas defasagens contribuem para reduzir o erro de previsão da VALE3, além do que já é explicado pelas próprias defasagens de VALE3.

Os resultados variam de acordo com a estrutura de defasagem considerada. Na especificação baseada no AIC ($p = 7$), o teste não rejeita a hipótese nula de ausência de causalidade de Granger ($p\text{-valor} \approx 0,27$). Na especificação baseada no BIC ($p = 1$), observa-se evidência mais próxima do limiar de significância, com $p\text{-valor} \approx 0,09$, indicando suporte apenas marginal à causalidade ao nível de 10%. Esse padrão — ausência de evidência robusta sob modelos mais ricos e evidência fraca sob modelo parcimonioso — sugere que o conteúdo preditivo intertemporal do minério é, quando presente, de curtíssimo prazo e de magnitude modesta.

É crucial destacar que tal resultado não contradiz as regressões contemporâneas estimadas anteriormente. Pelo contrário: ele é compatível com a hipótese de que o minério afeta os retornos da VALE3 principalmente no mesmo período, com pouca persistência dinâmica detectável por modelos lineares. Em mercados líquidos, informações relevantes tendem a ser rapidamente incorporadas aos preços, o que naturalmente limita a capacidade de se obter previsibilidade em defasagens longas.

Tabela 1 — Testes de causalidade de Granger

	AIC	BIC
x	Ferro (USD)	Ferro (USD)
y	VALE3	VALE3
lags	7	1
Teste-F	1.248072	2.816588
Valor-p	0.272502	0.093387

Fonte: Yahoo Finance; elaboração própria.

Tabela 2 — Acurácia direcional

	VAR-BIC	VAR-AIC
Obs. OOS	2610	2610
Lags	1	7
Dir. Accuracy (%)	51.15	50.27
RMSE	0.027304	0.027636
MAE	0.019183	0.019508
Vol. previsão (std)	0.003070	0.005271
Viés direcional (mean sign)	0.171648	0.111877

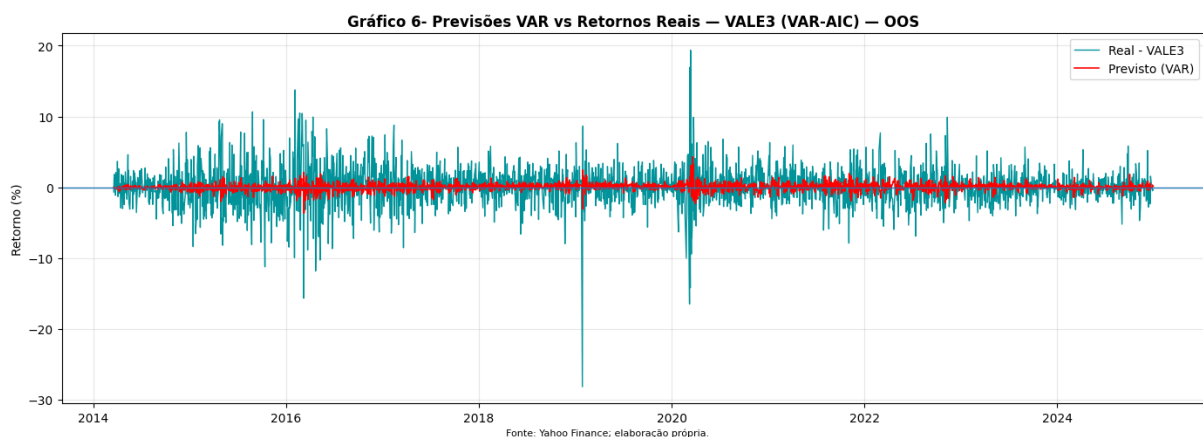
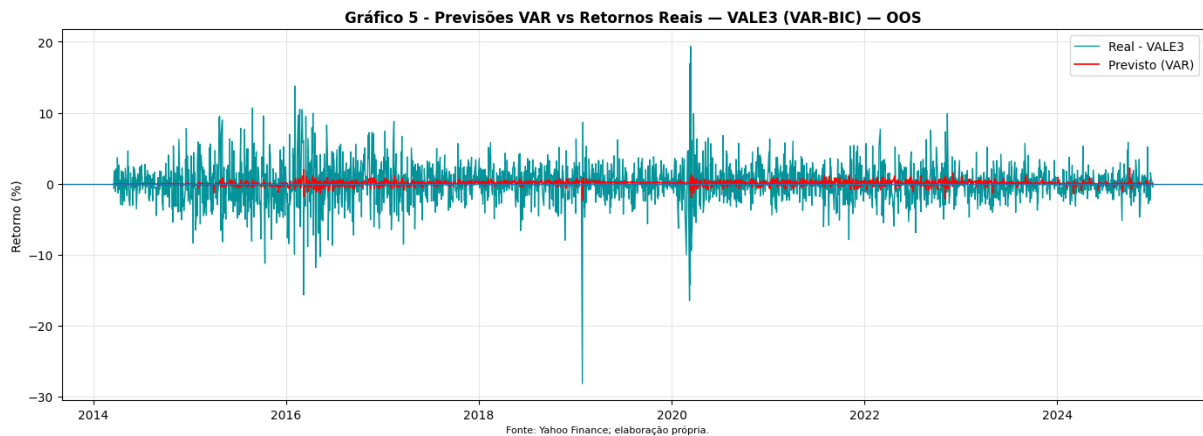
Fonte: Yahoo Finance; elaboração própria.

4.1.4 VAR como ferramenta preditiva: previsões fora da amostra (OOS)

Ainda que o VAR forneça um arcabouço consistente para análise dinâmica e identificação de respostas a choques, o objetivo central deste trabalho envolve avaliar utilidade prática na construção de estratégias quantitativas. Assim, estima-se a capacidade preditiva fora da amostra (out-of-sample) a partir de um procedimento walk-forward/rolling-expanding, no qual o VAR é reestimado iterativamente até o instante t e utilizado para prever o retorno em $t+1$. Esse procedimento aproxima o uso real do modelo em tempo de mercado, evitando contaminação por informação futura.

As previsões foram geradas considerando duas especificações: VAR-BIC ($p = 1$), mais parcimonioso, e VAR-AIC ($p = 7$), mais flexível. A avaliação concentra-se não apenas em métricas de erro (que tendem a penalizar excessivamente modelos lineares em ambientes altamente voláteis), mas sobretudo na dimensão mais útil para estratégias direcionais: a taxa de acerto do sinal do retorno. Os resultados indicam desempenho modesto, com taxa direcional próxima do limite aleatório (50%), sendo o VAR-BIC levemente superior e mais estável que o VAR-AIC. Isso sugere que a maior complexidade dinâmica do modelo com mais defasagens não se traduz em ganho preditivo, possivelmente em razão de maior variância de estimação e risco de sobreajuste.

Os Gráficos 5 e 6 ilustram a comparação entre retornos reais e retornos previstos. Em ambos os casos, observa-se que as previsões possuem amplitude significativamente inferior à dos retornos observados — um resultado típico de modelos lineares aplicados a séries financeiras, que tendem a “encolher” previsões em direção à média e não conseguem reproduzir os extremos e saltos observados no mercado. Isso reforça a interpretação de que, embora o VAR seja metodologicamente sólido como instrumento econométrico, sua utilização isolada como mecanismo de previsão direcional apresenta limitações naturais no contexto estudado.



4.1.5 Síntese e implicações para as próximas etapas

Em síntese, os resultados do VAR corroboram as evidências obtidas nas regressões lineares e múltiplas: existe associação economicamente plausível entre o minério e os retornos da VALE3, com efeitos concentrados no curto prazo e baixa persistência dinâmica. As funções impulso–resposta sugerem resposta inicial positiva, enquanto a FEVD e os testes de Granger indicam que o minério responde por parcela limitada da variabilidade e do conteúdo preditivo intertemporal.

Sob a ótica de modelagem quantitativa, esse conjunto de evidências aponta para uma conclusão importante: o VAR pode ser interpretado como gerador de um sinal direcional fraco, porém razoavelmente estável e consistente, cuja utilidade prática aumenta quando combinado com mecanismos de controle de risco. Essa constatação motiva a transição para os modelos de volatilidade condicional (ARCH/GARCH), que capturam de forma explícita a característica dominante das séries financeiras analisadas — o agrupamento de volatilidade — e são mais diretamente adequados à construção de estratégias baseadas em regimes, como desenvolvido nas seções subsequentes.

4.2 Modelagem da volatilidade condicional: GARCH

A análise descritiva e os exercícios lineares realizados anteriormente indicaram que a informação estatisticamente relevante nas séries financeiras analisadas está concentrada, sobretudo, na dinâmica da variância, e não na média condicional. Em particular, a evidência de caudas pesadas e de heterocedasticidade condicional (volatility clustering) motiva a adoção de modelos da família ARCH/GARCH como ferramenta central para caracterizar, prever e explorar regimes de risco na VALE3. Assim, antes de avançar para estratégias quantitativas, esta seção consolida os testes preliminares de adequação, a estimação in-sample do GARCH(1,1), o exercício de robustez com distribuição Student-t e a avaliação fora da amostra via expanding window.

4.2.1 Testes preliminares pré-GARCH (VALE3 e Minério)

Como etapa de validação econométrica, foram aplicados testes para verificar se as séries de retornos apresentam as propriedades mínimas para justificar uma modelagem de volatilidade condicional. Três conjuntos de diagnósticos foram conduzidos: estacionariedade via ADF, autocorrelação serial via Ljung–Box, e heterocedasticidade condicional via ARCH-LM, tanto sobre os retornos quanto, adicionalmente, sobre os resíduos de uma equação simples para a média. O objetivo é separar um caso em que há apenas ruído na média de outro em que existe dependência temporal na variância — precisamente o alvo de modelos GARCH.

Tabela 3 — Testes pré-GARCH (VALE3)

Teste	Estatística	Descrição	Conclusão
VALE3 ret - ADF stat	-17.724654	ADF: estatística	Estacionária (OK)
VALE3 ret - ADF p	0.000000	ADF: p-valor	Estacionária (OK)
VALE3 ret - LB Q (ret)	26.873691	Ljung-Box Q (retornos), lags=12	Há autocorrelação na média (avaliar ARMA/AR)
VALE3 ret - LB p (ret)	0.008057	Ljung-Box p-valor (retornos), lags=12	Há autocorrelação na média (avaliar ARMA/AR)
VALE3 ret - LB Q (ret ²)	798.477850	Ljung-Box Q (retornos ²), lags=12	Evidência de clustering/efeitos na variância (OK p/ GARCH)
VALE3 ret - LB p (ret ²)	0.000000	Ljung-Box p-valor (retornos ²), lags=12	Evidência de clustering/efeitos na variância (OK p/ GARCH)
VALE3 ret - ARCH LM stat	353.685936	ARCH-LM (LM), nlags=12	Evidência de efeito ARCH (OK p/ GARCH)
VALE3 ret - ARCH LM p	0.000000	ARCH-LM (LM) p-valor, nlags=12	Evidência de efeito ARCH (OK p/ GARCH)
VALE3 ret - ARCH F stat	32.820584	ARCH-LM (F), nlags=12	Evidência de efeito ARCH (OK p/ GARCH)
VALE3 ret - ARCH F p	0.000000	ARCH-LM (F) p-valor, nlags=12	Evidência de efeito ARCH (OK p/ GARCH)
VALE3 ret - ARCH(resid) LM stat	355.357905	ARCH-LM nos resíduos AR(1) (LM), nlags=12	Evidência de ARCH nos resíduos (OK p/ GARCH)
VALE3 ret - ARCH(resid) LM p	0.000000	ARCH-LM nos resíduos AR(1) (LM) p-valor, nlags=12	Evidência de ARCH nos resíduos (OK p/ GARCH)
VALE3 ret - ARCH(resid) F stat	32.995250	ARCH-LM nos resíduos AR(1) (F), nlags=12	Evidência de ARCH nos resíduos (OK p/ GARCH)
VALE3 ret - ARCH(resid) F p	0.000000	ARCH-LM nos resíduos AR(1) (F) p-valor, nlags=12	Evidência de ARCH nos resíduos (OK p/ GARCH)

Fonte: Yahoo Finance; elaboração própria.

Tabela 4 — Testes pré-GARCH (Minério)

Teste	Estatística	Descrição	Conclusão
Minério - ADF stat	-22.713401	ADF: estatística	Estacionária (OK)
Minério - ADF p	0.000000	ADF: p-valor	Estacionária (OK)
Minério - LB Q (ret)	56.065765	Ljung-Box Q (retornos), lags=12	Há autocorrelação na média (avaliar ARMA/AR)
Minério - LB p (ret)	0.000000	Ljung-Box p-valor (retornos), lags=12	Há autocorrelação na média (avaliar ARMA/AR)
Minério - LB Q (ret ²)	10.239244	Ljung-Box Q (retornos ²), lags=12	Evidência fraca de clustering na variância
Minério - LB p (ret ²)	0.594981	Ljung-Box p-valor (retornos ²), lags=12	Evidência fraca de clustering na variância
Minério - ARCH LM stat	9.764097	ARCH-LM (LM), nlags=12	Evidência fraca de efeito ARCH
Minério - ARCH LM p	0.636648	ARCH-LM (LM) p-valor, nlags=12	Evidência fraca de efeito ARCH
Minério - ARCH F stat	0.812887	ARCH-LM (F), nlags=12	Evidência fraca de efeito ARCH
Minério - ARCH F p	0.637444	ARCH-LM (F) p-valor, nlags=12	Evidência fraca de efeito ARCH
Minério - ARCH(resid) LM stat	9.806365	ARCH-LM nos resíduos AR(1) (LM), nlags=12	Evidência fraca de ARCH nos resíduos
Minério - ARCH(resid) LM p	0.632944	ARCH-LM nos resíduos AR(1) (LM) p-valor, nlags=12	Evidência fraca de ARCH nos resíduos
Minério - ARCH(resid) F stat	0.816416	ARCH-LM nos resíduos AR(1) (F), nlags=12	Evidência fraca de ARCH nos resíduos
Minério - ARCH(resid) F p	0.633739	ARCH-LM nos resíduos AR(1) (F) p-valor, nlags=12	Evidência fraca de ARCH nos resíduos

Fonte: Yahoo Finance; elaboração própria.

Os resultados para a VALE3 (Tabela 3) são inequívocos. O ADF rejeita a presença de raiz unitária, corroborando a estacionariedade dos retornos. Ao mesmo tempo, observa-se forte evidência de dependência na segunda ordem: os testes aplicados aos retornos ao quadrado e aos resíduos ao quadrado rejeitam a hipótese nula de ausência de autocorrelação e, sobretudo, os testes ARCH-LM rejeitam de forma clara a homocedasticidade. Em termos econômicos, isso significa que choques de volatilidade não são episódios isolados: eles se agrupam e apresentam persistência, padrão típico de mercados acionários líquidos e consistente com a literatura clássica de Engle e Bollerslev. Em termos econométricos, trata-se exatamente do cenário em que um GARCH(1,1) tende a ser parcimonioso, estável e eficaz.

Em contraste, para os retornos do minério de ferro (Tabela 4), embora a estacionariedade também seja confirmada, a evidência de heterocedasticidade condicional é bem menos pronunciada. Os testes em retornos ao quadrado e ARCH-LM não apontam, de forma robusta, a mesma persistência de variância observada na VALE3. Isso não implica ausência de risco na commodity, mas sugere que, no recorte

diário e no período analisado, a volatilidade do minério pode estar mais associada a episódios específicos e a mecanismos distintos de formação de preço do que a um processo condicional persistente do tipo GARCH simples.

Essa assimetria entre as séries é relevante para a estratégia do trabalho: reforça o minério como variável fundamental (na dinâmica média e na comovimentação contemporânea com VALE3), enquanto consolida a VALE3 como o ativo para o qual a modelagem explícita da variância condicional é mais justificada e operacionalmente útil. Em outras palavras: o minério entra como motor econômico; o GARCH entra como motor de risco.

4.2.2 Estimação in-sample do modelo AR(1)–GARCH(1,1)

Confirmada a adequação do arcabouço, estimou-se um modelo AR(1)–GARCH(1,1) para os retornos diários da VALE3, com resíduos assumidos inicialmente como Normais. Essa especificação funciona como baseline por duas razões: é a parametrização mais utilizada como referência na literatura empírica e oferece um equilíbrio muito favorável entre parcimônia e capacidade de capturar persistência na volatilidade.

Os parâmetros estimados indicam dinâmica típica de ativos financeiros. O componente ARCH (α) assume magnitude moderada, refletindo que choques recentes influenciam imediatamente o risco, enquanto o componente GARCH (β) é elevado, evidenciando persistência — isto é, a volatilidade tende a demorar a retornar ao seu patamar de longo prazo após episódios de stress. A soma $\alpha + \beta$ próxima de 1 caracteriza o conhecido comportamento de clustering: choques aumentam a variância, e esse efeito se dissipa lentamente. Em paralelo, a equação da média apresenta baixo conteúdo informacional (constante e termo AR pouco relevantes), resultado coerente com a dificuldade de prever a média de retornos diários sob eficiência informacional.”

Além disso, os diagnósticos pós-estimação corroboram a adequação do modelo: a ausência de autocorrelação relevante nos resíduos padronizados ao quadrado (avaliada via Ljung–Box) sugere que o GARCH(1,1) absorveu a principal dependência temporal presente na variância. Assim, o papel do modelo não é “acertar retornos”, mas construir uma trajetória coerente para $\hat{\sigma}_t$, que será utilizada como instrumento de filtragem de regime e controle de risco nas estratégias subsequentes.

4.2.3 Robustez: comparação entre distribuições Normal e Student-t

Como verificação de robustez, estimou-se também uma versão do GARCH(1,1) com resíduos seguindo uma distribuição Student-t, frequentemente empregada quando se suspeita de caudas mais pesadas do que a Normal. Em séries financeiras, essa extensão é teoricamente atraente porque permite acomodar maior probabilidade de eventos extremos sem “forçar” a variância condicional a absorver toda a não normalidade.

Tabela 5 — Comparação entre modelos GARCH(1,1)

Dist	α	β	$\alpha+\beta$	AIC	BIC	ν	Convergência	Dist. efetiva
Normal	0.049998	0.929963	0.979961	-15756.242315	-15725.636398		OK	Normal
Student-t	0.561077	0.382708	0.943785	28170.349684	28207.076783	179.762028	OK	Quase Normal

Fonte: Yahoo Finance; elaboração própria.

A Tabela 5 resume a comparação. Embora a especificação Student-t tenha convergido adequadamente, o parâmetro de graus de liberdade (ν) é muito elevado, indicando que a Student-t estimada se aproxima de uma Normal — ou seja, no período analisado, a evidência empírica de caudas substancialmente mais pesadas não foi suficiente para justificar, em termos de ajuste, a parametrização adicional. Esse resultado é consistente com o fato de que, quando ν cresce, a Student-t perde a vantagem relativa sobre a Normal, mantendo apenas o custo de complexidade.

Os critérios de informação (AIC/BIC) também reforçam essa leitura: não há ganho de ajuste que compense a maior flexibilidade do modelo Student-t no contexto desta aplicação. Por esse motivo, a especificação Normal é mantida como baseline operacional para as previsões e para a construção dos filtros de regime. A versão Student-t permanece como exercício complementar, demonstrando que as conclusões principais não dependem criticamente da escolha da distribuição dos resíduos.

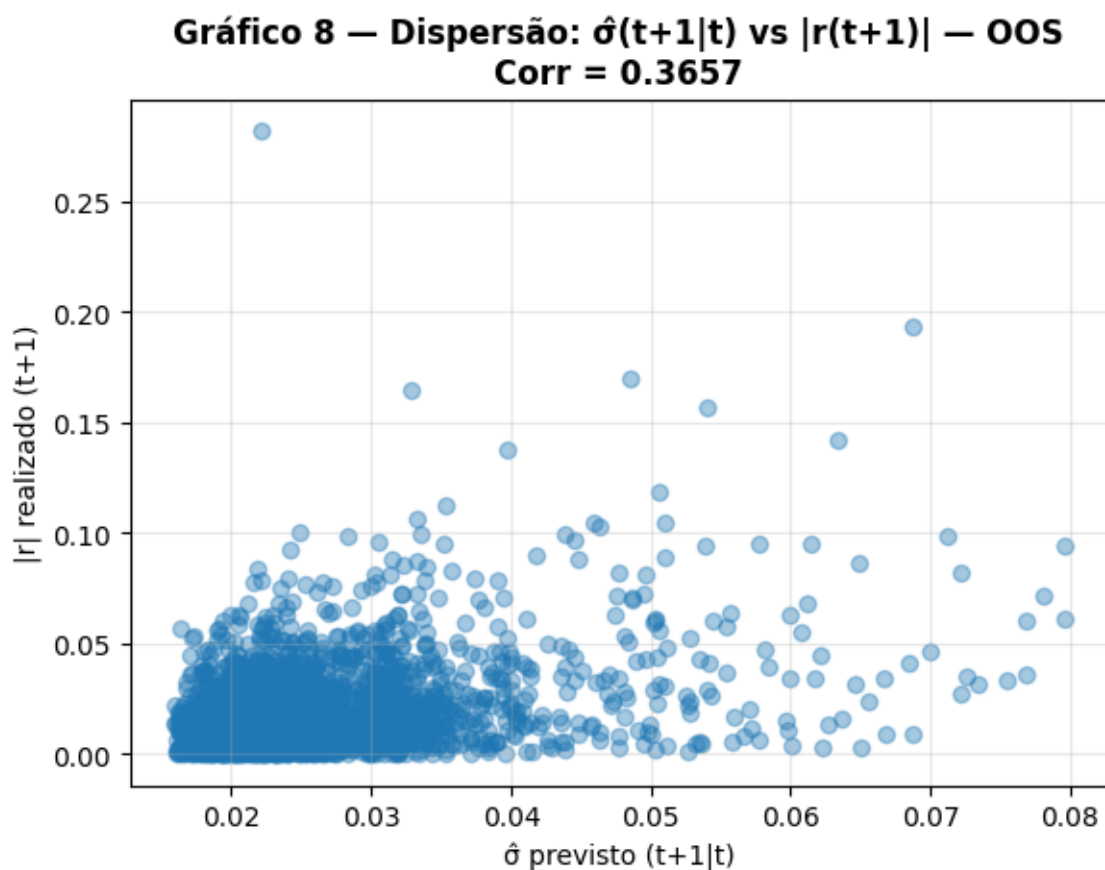
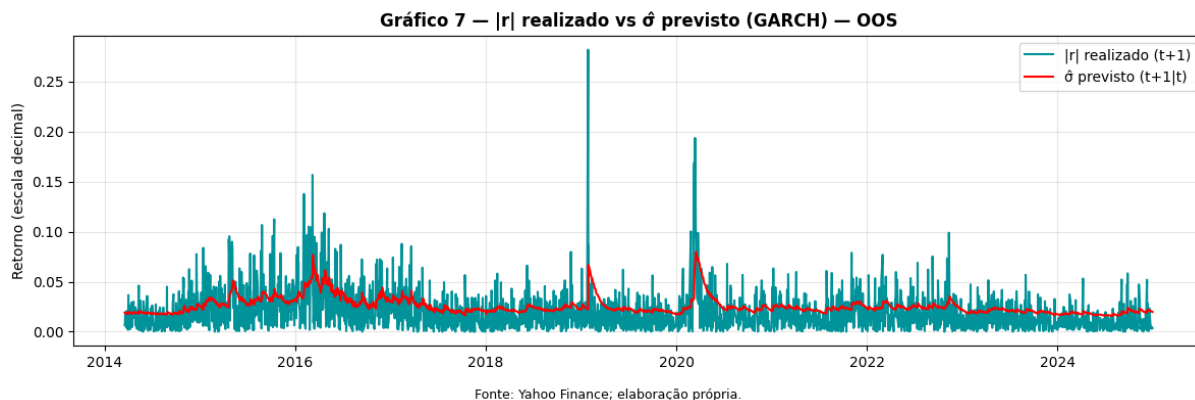
4.2.4 Avaliação fora da amostra: previsão de volatilidade via expanding window

Com foco aplicado, o modelo foi então avaliado em previsão fora da amostra (OOS) sob um esquema de expanding window, análogo ao utilizado na etapa do VAR. A cada instante t , o modelo é reestimado com toda a informação disponível até t e gera a previsão $\sigma_{t+1|t}^2$ para o período seguinte. Essa estrutura mimetiza uma implementação realista: parâmetros e volatilidade condicional são atualizados à medida que novas observações chegam.

Tabela 6 — Avaliação fora da amostra (previsão de volatilidade) — GARCH

GARCH	
Obs. OOS	2610
Modelo	AR(1)-GARCH(1,1) (Normal)
Corr($\hat{\sigma}$, $ r $)	0.365702
RMSE ($\hat{\sigma}$ vs $ r $)	0.019305
MAE ($\hat{\sigma}$ vs $ r $)	0.014789
Média $\hat{\sigma}$	0.025597
Vol. $\hat{\sigma}$ (std)	0.008702
Convergência (%)	100.00

Fonte: Yahoo Finance; elaboração própria.



Fonte: Yahoo Finance; elaboração própria.

A métrica central desta etapa é comparar a volatilidade prevista com uma proxy de volatilidade realizada, adotando-se o valor absoluto do retorno do próximo período, $|r_{t+1}|$, prática usual em aplicações empíricas. Os resultados (Tabela 6) mostram desempenho estatisticamente consistente: a correlação entre $\hat{\sigma}_{t+1|t}$ e $|r_{t+1}|$ é positiva e economicamente relevante ($\approx 0,37$), indicando que o modelo captura parte substantiva da variação temporal do risco, ainda que — como esperado — esteja longe de prever a magnitude exata de eventos extremos.

As métricas de erro (RMSE e MAE) reforçam que a previsão condicional do GARCH acompanha a escala típica dos retornos observados. Há também uma característica estrutural importante: a trajetória de $\hat{\sigma}_{t+1|t}$ é mais suave do que a série de $|r_{t+1}|$, o que não é defeito, mas consequência direta do mecanismo de atualização do GARCH, que reage a choques e os internaliza com persistência, sem “replicar” cada pico extremo da série realizada.

O Gráfico 7 evidencia visualmente esse comportamento: em períodos de maior turbulência, a volatilidade prevista sobe de maneira clara, permanecendo elevada por algum tempo — exatamente o padrão esperado sob clustering. Já o Gráfico 8, de dispersão, confirma a relação positiva entre previsão e realização, embora com dispersão considerável, especialmente nos extremos: choques de cauda permanecem, por natureza, dificilmente antecipáveis em magnitude.

Do ponto de vista prático, a principal conclusão é que o GARCH se comporta como um instrumento de regime: ele não “adivinha” retornos, mas fornece um sinal operacional de risco que tende a aumentar quando o mercado entra em períodos adversos e a reduzir quando o ambiente se normaliza. Essa é precisamente a variável que será explorada adiante como filtro para reduzir exposição em regimes de alta volatilidade e, conseqüentemente, melhorar a relação risco–retorno das estratégias quantitativas baseadas em sinais direcionais.

4.2.5 Conexão com as estratégias

Em síntese, a seção confirma três pontos: a VALE3 apresenta evidência robusta de heterocedasticidade condicional e persistência de volatilidade; o GARCH(1,1) baseline é estável e suficientemente parcimonioso; e em OOS, o modelo entrega um sinal de risco com conteúdo informacional consistente. Com isso, a volatilidade prevista passa a ser tratada como componente central de gestão de risco, servindo de base para filtros de regime e para a construção das estratégias combinadas avaliadas nas subseções seguintes.

4.3 Modelos de Machine Learning: Random Forest, LSTM e custo computacional

Como extensão aos modelos econométricos de volatilidade discutidos anteriormente (em particular, o GARCH), esta seção incorpora uma abordagem não paramétrica baseada em *Machine Learning* para previsão de volatilidade realizada

dos retornos da VALE3. O objetivo é verificar se um algoritmo flexível, capaz de capturar não linearidades e interações entre preditores, melhora a aderência fora da amostra na previsão do risco diário, medido por $|r_{t+1}|$ (valor absoluto do retorno no período $t+1$).

A especificação empregada utiliza Random Forest em regressão, estimada em ambiente *walk-forward* com janela expansiva (*expanding window*), de modo plenamente comparável aos exercícios OOS realizados para VAR e GARCH. Em todas as estimações, as variáveis explicativas são construídas exclusivamente com informações disponíveis até o instante t , evitando vazamento de informação (*look-ahead bias*). O conjunto de *features* é formado por defasagens de $|r_{t+1}|$, defasagens de r_t^2 e estatísticas móveis (médias e desvios-padrão) baseadas em $|r|$, isto é, proxies clássicas de volatilidade histórica.

4.3.1 Random Forest com reestimação diária (refit a cada t)

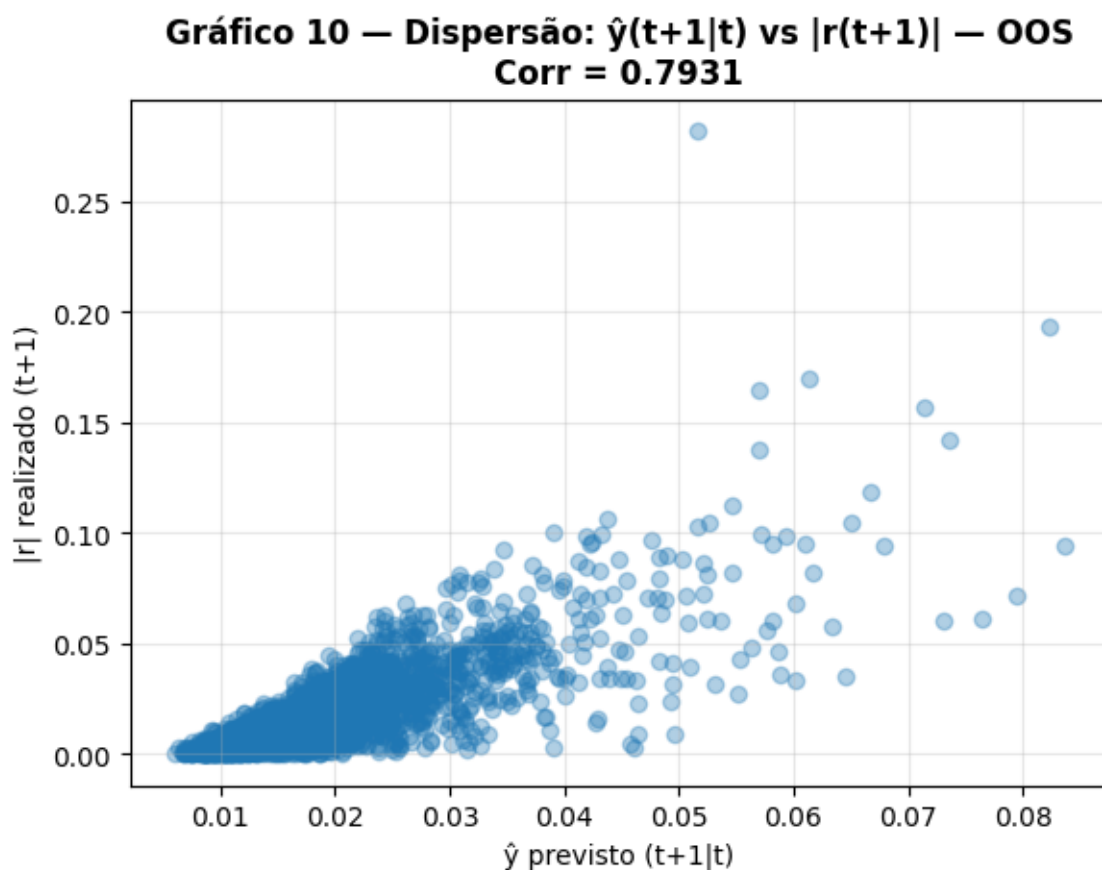
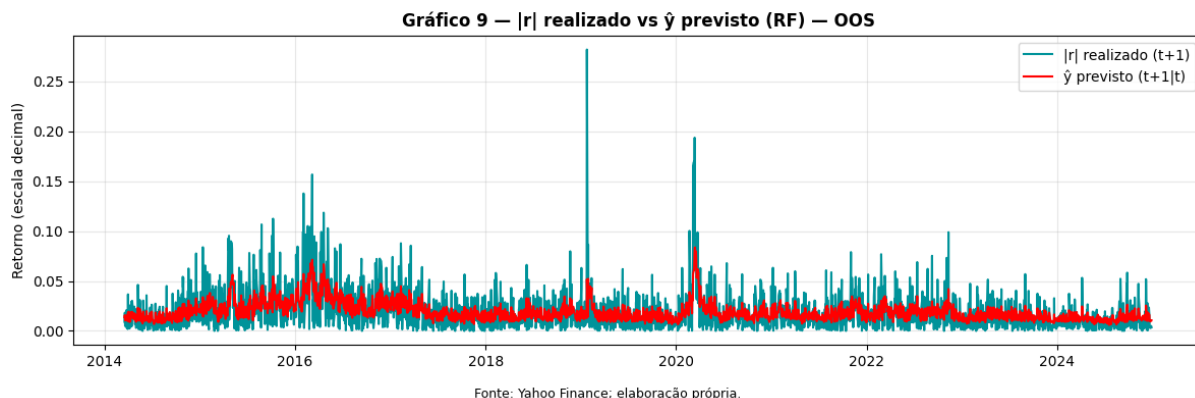
A primeira especificação adota reestimação do Random Forest a cada nova observação fora da amostra, simulando um cenário em que o agente atualiza o modelo diariamente conforme chega nova informação. Os resultados OOS mostram desempenho elevado do ponto de vista estatístico: a correlação entre a volatilidade prevista $\hat{y}_{t+1|t}$ e a volatilidade realizada r_{t+1} atinge patamar substancialmente superior ao observado no GARCH, ainda que as previsões permaneçam naturalmente suavizadas em relação aos picos extremos (o que é típico quando o alvo é $|r|$, uma proxy ruidosa de volatilidade “realizada”).

Tabela 7 — Avaliação OOS (previsão de volatilidade)

	Random Forest
Obs. OOS	2609
Modelo	Random Forest (regressão)
Corr($\hat{y}$, $ r $)	0.793108
RMSE ($\hat{y}$ vs $ r $)	0.013252
MAE ($\hat{y}$ vs $ r $)	0.008668
Acerto Direcional (%)	82.32
Média $\hat{y}$	0.019132
Vol. $\hat{y}$ (std)	0.009482

Fonte: Yahoo Finance; elaboração própria.

A leitura gráfica reforça esse resultado: a série prevista acompanha com consistência os regimes de volatilidade (períodos de calma versus turbulência), respondendo de modo claro às mudanças no nível de risco, embora com compressão dos eventos mais raros e extremos. O diagrama de dispersão evidencia relação positiva bem definida, compatível com a alta correlação reportada.



Fonte: Yahoo Finance; elaboração própria.

Além disso, reporta-se acerto direcional elevado. Como o alvo é volatilidade (e não retorno com sinal), a interpretação adequada é que o modelo consegue identificar, com alta frequência, movimentos relativos de alta/queda no risco (por exemplo, aumento versus redução de $|r|$ em relação ao período anterior, conforme sua definição operacional no código).

O ponto crítico dessa estratégia é custo computacional: reestimar Random Forest diariamente ao longo de milhares de observações OOS pode ser inviável em

contextos práticos, sobretudo quando se exige baixa latência, múltiplos ativos, ou rotinas de *backtesting* extensas.

4.3.2 Random Forest com reestimação periódica (refit a cada 200 observações)

Para explicitar o *trade-off* entre desempenho e viabilidade operacional, estima-se uma segunda especificação em que o Random Forest é reestimado apenas a cada 200 observações fora da amostra. Essa abordagem reduz drasticamente o custo computacional, ao custo de menor adaptabilidade a mudanças de regime.

Tabela 8 — Avaliação OOS (previsão de volatilidade)

Random Forest	
Obs. OOS	2608
Modelo	Random Forest (regressão)
Corr($\hat{y}$, $ r $)	0.339455
RMSE ($\hat{y}$ vs $ r $)	0.018299
MAE ($\hat{y}$ vs $ r $)	0.012606
Acerto Direcional (%)	73.07
Média $\hat{y}$	0.018698
Vol. $\hat{y}$ (std)	0.006813

Fonte: Yahoo Finance; elaboração própria.

Os resultados indicam queda na correlação e piora nas métricas de erro, o que é esperado: com atualização menos frequente, o modelo reage mais lentamente a mudanças estruturais e a quebras no padrão recente de volatilidade. Ainda assim, o desempenho permanece qualitativamente consistente: o modelo preserva capacidade de acompanhar regimes de risco e mantém acerto direcional elevado, sugerindo que parte relevante do padrão de volatilidade diária é relativamente estável e capturável por estatísticas históricas simples, mesmo com reestimações esparsas.

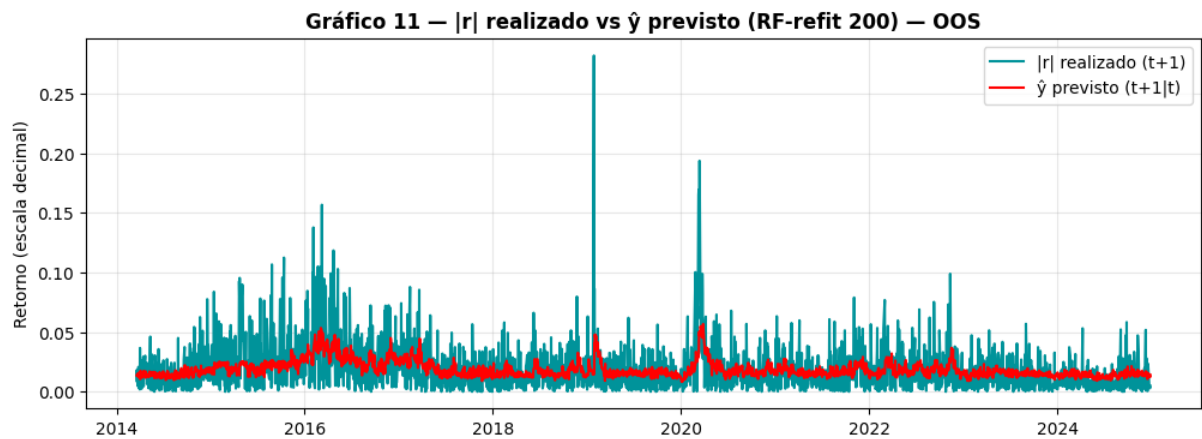
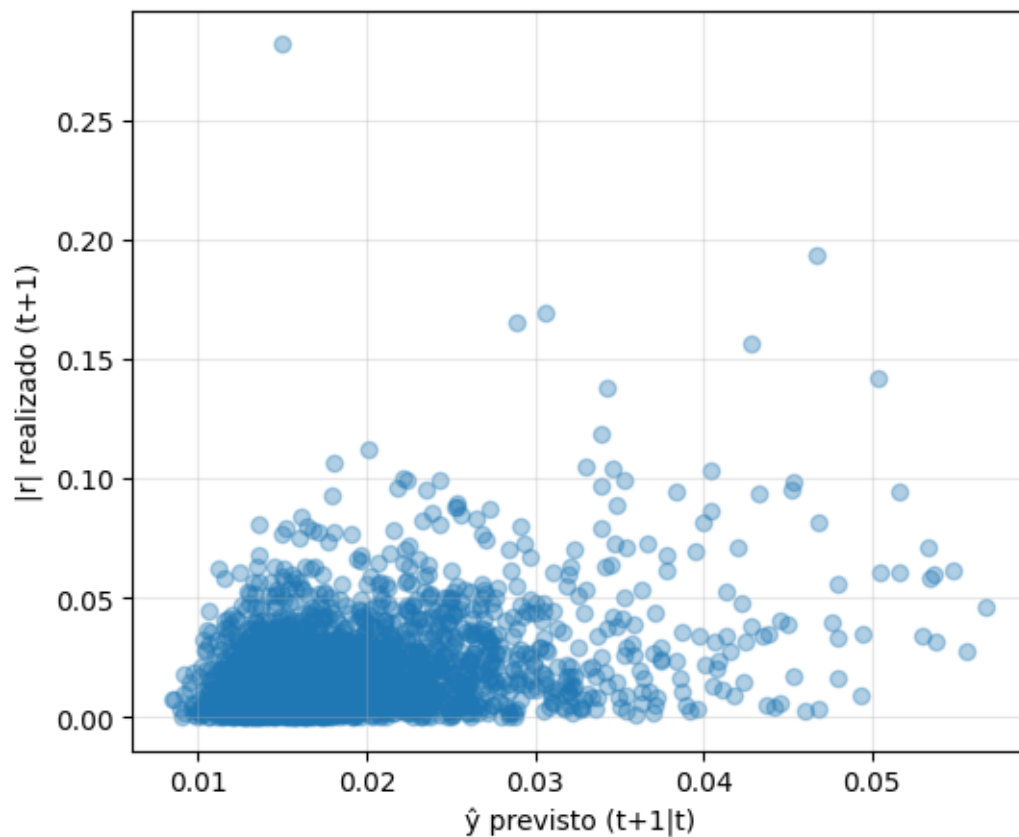


Gráfico 12 — Dispersão: $\hat{y}(t+1|t)$ vs $|r(t+1)|$ (RF-refit 200) — OOS
Corr = 0.3395



O diagrama de dispersão mostra maior dispersão dos pontos em torno da relação esperada, refletindo justamente a perda de ajuste fino quando o modelo deixa de ser recalibrado continuamente.

4.3.3 Síntese e conexão com a etapa de estratégias

Em conjunto, os resultados desta seção apontam que o Random Forest é altamente competitivo para previsão de volatilidade diária (via $|r_{t+1}|$), com ganhos

estatísticos relevantes quando comparado a modelos paramétricos tradicionais, especialmente sob reestimação frequente. Ao mesmo tempo, a análise evidencia que a utilidade prática depende de decisões operacionais: frequência de refit é um determinante central do equilíbrio entre performance e custo computacional.

Essas evidências são diretamente úteis para a próxima etapa do trabalho, em que o foco deixa de ser apenas acurácia estatística e passa a ser relevância econômica: como sinais de volatilidade (GARCH e ML) podem ser convertidos em regras de exposição, controle de risco e filtros de regime, e como essas regras se comportam sob custos de transação e restrições realistas de operação.

4.3.4 Extensão com LSTM (previsão de volatilidade)

Como exercício complementar aos modelos de Machine Learning apresentados anteriormente, estimou-se também uma arquitetura recorrente do tipo LSTM (*Long Short-Term Memory*), com o objetivo de verificar se uma abordagem de aprendizado profundo seria capaz de capturar dependências temporais não lineares na dinâmica da volatilidade realizada dos retornos da VALE3. Assim como nos demais exercícios preditivos fora da amostra, o alvo de previsão foi o valor absoluto do retorno no período seguinte, $|r_{t+1}|$, e a estimação respeitou estritamente o conjunto informacional disponível até o instante t , evitando-se vazamento de informação.

Os resultados do modelo LSTM indicam desempenho similar ao Random Forest com reestimação periódica, sugerindo que, dado o conjunto informacional utilizado (features baseadas em valores defasados e estatísticas móveis da própria série), arquiteturas mais complexas não geram ganhos substanciais em acurácia no presente problema. Em particular, observa-se correlação entre valores previstos e realizados de aproximadamente 0,36, com RMSE $\approx 0,0189$ e MAE $\approx 0,0126$, além de acerto direcional em torno de 73,17%. Esses resultados apontam que o modelo possui capacidade razoável de acompanhar regimes médios de volatilidade, mas com limitação natural na antecipação de movimentos extremos — característica recorrente em previsões diárias de volatilidade.

Em síntese, o exercício com LSTM reforça a principal conclusão da seção de Machine Learning: há previsibilidade estatística na volatilidade realizada, porém o ganho marginal associado a modelos de maior complexidade é limitado quando a informação disponível se restringe à própria trajetória passada do ativo. Assim, o

LSTM é mantido como extensão empírica de robustez, corroborando a consistência dos resultados obtidos com o Random Forest.

4.4 Etapa de estratégias e avaliação econômica dos modelos

A seção anterior concentrou-se na avaliação estatística e econométrica dos modelos propostos para capturar a dinâmica dos retornos e, sobretudo, da volatilidade da ação VALE3. Os resultados construíram um diagnóstico consistente: os retornos apresentam propriedades típicas de séries financeiras — média próxima de zero, caudas pesadas, assimetria e heterocedasticidade condicional — o que justifica metodologias que tratem explicitamente a volatilidade como componente central do processo gerador de dados.

Do ponto de vista de relações contemporâneas, as regressões lineares indicaram de forma robusta que o minério de ferro exerce influência positiva e estatisticamente relevante sobre os retornos diários da VALE3, com sensibilidade maior ao minério cotado em USD do que ao minério convertido para BRL. Em contraste, o câmbio BRL/USD não apresentou contribuição marginal relevante no horizonte diário, resultado consistente com a análise de correlação e com a interpretação econômica de que o impacto cambial se manifesta predominantemente de maneira indireta e em frequências mais baixas. Apesar do baixo poder explicativo (R^2), típico em modelos de retornos de alta frequência, a evidência de relação sistemática entre commodity e ação foi clara e serviu como base para o desenho das especificações multivariadas posteriores.

Ao avançar para modelos dinâmicos, a estimação VAR corroborou a existência de um canal de transmissão no curto prazo entre choques do minério e os retornos da VALE3, conforme evidenciado pelas funções impulso-resposta. Contudo, a decomposição de variância e os testes de causalidade de Granger mostraram que a maior parte da imprevisibilidade permanece idiossincrática ao ativo, e que a capacidade preditiva do VAR isoladamente é limitada. A avaliação fora da amostra confirmou esse diagnóstico: o modelo apresentou acerto direcional apenas marginalmente superior ao acaso quando parametrizado de forma parcimoniosa (BIC), sugerindo que sinais lineares na média são fracos e rapidamente dissipados.

Esse resultado reforçou a motivação para deslocar o foco da previsão da média condicional para a previsão da volatilidade. Os modelos GARCH forneceram evidência sólida de persistência e clustering de volatilidade nos retornos da VALE3, com parâmetros estáveis e desempenho fora da amostra estatisticamente consistente na previsão do risco (correlação positiva entre $\hat{\sigma}_{t+1|t}$ e $|r_{t+1}|$). A comparação com distribuições alternativas indicou ausência de ganho material com Student-t no período analisado, mantendo a especificação Normal como baseline mais parcimonioso e estável.

Por fim, as extensões via Machine Learning (Random Forest e LSTM) mostraram que a volatilidade realizada é previsível com maior acurácia por abordagens flexíveis e não paramétricas. O Random Forest com reestimação frequente apresentou desempenho elevado, enquanto configurações mais viáveis operacionalmente (refit periódico) preservaram parte relevante da performance com custo computacional inferior. O LSTM, por sua vez, apresentou desempenho semelhante ao RF com refit espaçado, funcionando como exercício de robustez: modelos mais complexos não geraram ganhos proporcionais diante do conjunto informacional disponível (proxies de volatilidade baseadas no próprio histórico do ativo).

Em conjunto, esses resultados indicam que, embora haja evidência econômica e estatística de conexão entre minério e VALE3, a previsibilidade direta da média condicional dos retornos é limitada — enquanto a previsibilidade da volatilidade se mostra substancialmente mais relevante e operacionalizável. Entretanto, desempenho estatístico não é sinônimo de desempenho econômico: acerto direcional, RMSE ou correlação podem não se traduzir em retorno ajustado ao risco quando custos, turnover e regimes de mercado são considerados.

Dessa forma, esta seção avança para o objetivo central do trabalho: avaliar a relevância econômica dos sinais gerados pelos modelos, implementando estratégias quantitativas com regras operacionais claras e comparáveis entre metodologias. As estratégias serão testadas fora da amostra, com comparação direta contra um benchmark passivo (buy-and-hold), e avaliadas em termos de retorno, volatilidade, índice de Sharpe, drawdown máximo, tempo investido e impacto de custos de transação. A primeira estratégia considerada é baseada diretamente nos sinais

obtidos pelas regressões lineares, servindo como baseline interpretável para as extensões econométricas e de Machine Learning analisadas em sequência.

4.4.1 Estratégias baseadas em modelos VAR

Como extensão natural da análise econométrica realizada na seção anterior, foram implementadas estratégias de negociação fundamentadas em modelos Vetoriais Autorregressivos (VAR), com o objetivo de avaliar se a dinâmica intertemporal capturada pelos modelos possui valor econômico quando convertida em regras operacionais de alocação. Diferentemente das regressões lineares contemporâneas, o VAR permite explorar a dependência temporal entre retornos, produzindo previsões condicionais que, em princípio, poderiam antecipar movimentos direcionais de curtíssimo prazo.

A lógica da estratégia é propositalmente simples e transparente: em cada instante t , o VAR é reestimado com as informações disponíveis até aquele ponto e utilizado para prever o retorno da VALE3 em $t+1$. A estratégia assume posição comprada quando o retorno previsto é positivo e mantém-se fora do ativo (posição zerada) quando a previsão é negativa. Assim, o VAR é utilizado como um mecanismo de *timing*, buscando reduzir exposição em períodos desfavoráveis e elevar a exposição quando há expectativa direcional positiva.

Os resultados são avaliados fora da amostra (OOS), com comparação direta contra o benchmark Long 100% (buy-and-hold). Além disso, reportam-se resultados em termos brutos (gross) e líquidos (net), incorporando custos de transação estimados a partir do turnover observado — etapa crucial para verificar se o desempenho permanece plausível em condições mais realistas de implementação.

Foram avaliadas duas especificações do VAR, definidas pelos critérios de informação utilizados na seleção de defasagens: uma versão parcimoniosa determinada pelo critério BIC e uma versão mais flexível determinada pelo critério AIC. Essa comparação é particularmente relevante, pois permite avaliar empiricamente o impacto do *trade-off* entre ajuste e robustez — isto é, entre maior flexibilidade dinâmica e maior risco de sobreajuste e instabilidade operacional.

4.4.1.1 Estratégia VAR (BIC)

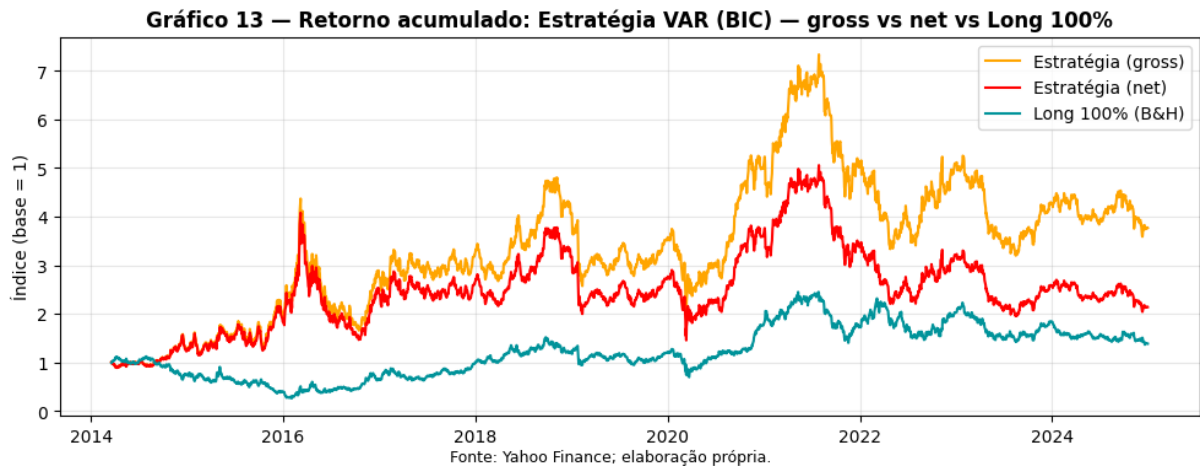
A estratégia baseada no VAR com defasagens selecionadas via BIC apresenta desempenho econômico superior ao benchmark no período analisado. Conforme

sumarizado na Tabela 9 e ilustrado no Gráfico 13, o retorno acumulado bruto atinge aproximadamente 276,8%, superando com ampla margem o retorno da estratégia passiva (38,6%). Após custos de transação, o retorno líquido permanece positivo e elevado (113,8%), o que sugere que o sinal direcional do VAR-BIC possui relevância econômica mesmo após fricções operacionais.

Tabela 9 — Retorno acumulado: VAR (BIC lags) - gross vs net vs Long 100%

	Estratégia VAR (BIC)
Obs. OOS	2610
Benchmark	Long 100% (buy-and-hold)
Retorno acumulado (Gross)	276.82%
Retorno acumulado (Net)	113.84%
Retorno acumulado (B&H)	38.58%
Volatilidade (Gross)	43.22%
Volatilidade (Net)	43.21%
Volatilidade (B&H)	43.23%
Sharpe (Gross)	0.515000
Sharpe (Net)	0.389000
Sharpe (B&H)	0.291000
Max Drawdown (Gross)	-62.08%
Max Drawdown (Net)	-64.31%
Max Drawdown (B&H)	-76.82%
% períodos investido	100.0%
Turnover médio	0.724000
Custo médio diário	0.02%

Fonte: Yahoo Finance; elaboração própria.



Em termos de risco, a volatilidade anualizada permanece praticamente idêntica à do benchmark (em torno de 43%), de modo que o ganho relativo se manifesta principalmente como melhoria de eficiência risco–retorno. Essa melhora é refletida no índice de Sharpe, que se eleva para 0,515 (gross) e 0,389 (net), contra 0,291 no buy-and-hold. Adicionalmente, o drawdown máximo, embora ainda expressivo, é inferior ao observado no benchmark, sugerindo alguma melhora na distribuição intertemporal dos retornos.

Cabe notar que a estratégia VAR (BIC) permanece 100% do tempo investida, o que implica alternância frequente de sinais e turnover elevado (0,724). Ainda assim, os resultados líquidos indicam que o modelo parcimonioso conseguiu produzir um padrão suficientemente consistente para compensar o custo de transação associado, reforçando a evidência discutida na etapa preditiva: o VAR-BIC parece funcionar como um sinal direcional fraco, porém mais estável.

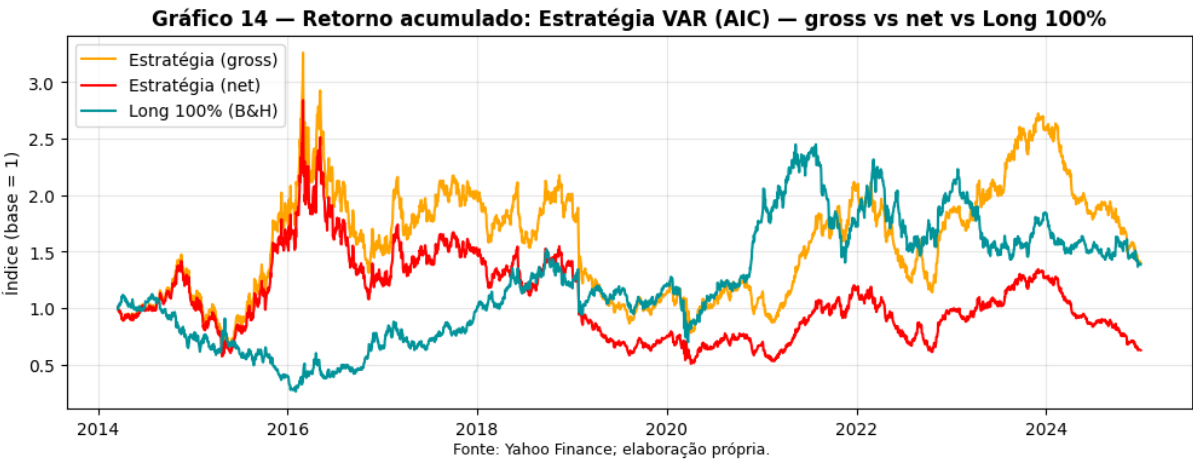
4.4.1.2 Estratégia VAR (AIC)

Em contraste, a estratégia baseada no VAR com defasagens selecionadas pelo critério AIC apresenta desempenho substancialmente inferior. Conforme indicado na Tabela 10 e no Gráfico 14, o retorno acumulado bruto (40,1%) é apenas marginalmente superior ao benchmark, e torna-se negativo quando descontados os custos de transação (−37,3%). Esse resultado evidencia que a maior flexibilidade do modelo não se traduz em desempenho robusto fora da amostra, especialmente quando se leva em consideração a fricção operacional causada pelo aumento do turnover.

Tabela 10 — Retorno acumulado: VAR (AIC lags) - gross vs net vs Long 100%

Estratégia VAR (AIC)	
Obs. OOS	2610
Benchmark	Long 100% (buy-and-hold)
Retorno acumulado (Gross)	40.12%
Retorno acumulado (Net)	-37.34%
Retorno acumulado (B&H)	38.58%
Volatilidade (Gross)	43.23%
Volatilidade (Net)	43.22%
Volatilidade (B&H)	43.23%
Sharpe (Gross)	0.294000
Sharpe (Net)	0.114000
Sharpe (B&H)	0.291000
Max Drawdown (Gross)	-76.05%
Max Drawdown (Net)	-82.18%
Max Drawdown (B&H)	-76.82%
% períodos investido	100.0%
Turnover médio	1.028000
Custo médio diário	0.03%

Fonte: Yahoo Finance; elaboração própria.



O turnover médio da estratégia é significativamente maior (1,028), refletindo elevada instabilidade nas previsões e maior sensibilidade ao ruído de curtíssimo

prazo. Consequentemente, o índice de Sharpe líquido se reduz de forma acentuada (0,114), tornando-se inferior tanto ao VAR (BIC) quanto ao buy-and-hold. Além disso, o drawdown máximo na versão líquida ultrapassa -82%, reforçando a interpretação de que a estratégia sofre com sinais erráticos e com maior vulnerabilidade em regimes adversos.

4.4.1.3 Discussão e implicações

Tomados em conjunto, os resultados indicam que estratégias baseadas em VAR podem apresentar valor econômico, desde que a especificação adotada preserve parcimônia. A superioridade do VAR (BIC) em relação ao VAR (AIC) é consistente com o argumento de que modelos com muitos parâmetros tendem a apresentar menor estabilidade preditiva fora da amostra e maior custo de implementação via turnover — um ponto particularmente relevante em aplicações quantitativas.

Enquanto o critério BIC seleciona estruturas mais enxutas e robustas, reduzindo o risco de sobreajuste, o critério AIC privilegia modelos mais complexos, que podem capturar ruído específico da amostra e gerar sinais excessivamente voláteis. Assim, a evidência obtida reforça que, em estratégias quantitativas aplicadas, a qualidade econômica do modelo depende não apenas da capacidade estatística de ajuste, mas também da sua estabilidade operacional e do impacto de custos de transação.

Dessa forma, o VAR (BIC) posiciona-se como uma alternativa intermediária: mais sofisticado do que o modelo linear contemporâneo, por incorporar dinâmica temporal, mas ainda dentro de um arcabouço interpretável e econometricamente estruturado. Nas subseções seguintes, avançamos para estratégias baseadas na modelagem explícita de volatilidade condicional e em modelos não lineares (Machine Learning), avaliando se tais abordagens geram ganhos adicionais de desempenho econômico e controle de risco.

4.4.2 Estratégias baseadas em regimes de volatilidade (GARCH)

A etapa anterior demonstrou que o modelo AR(1)–GARCH(1,1) é capaz de prever a dinâmica da volatilidade condicional dos retornos da VALE3 de forma estatisticamente consistente no período fora da amostra. Em particular, observou-se correlação positiva relevante entre a volatilidade prevista e a magnitude do retorno realizado, reforçando a utilidade do GARCH como instrumento de mensuração e

previsão de risco. Dando continuidade, esta subseção avalia se essa informação de volatilidade pode ser convertida em valor econômico por meio de regras simples de alocação dinâmica.

A estratégia proposta explora a hipótese — amplamente documentada em finanças empíricas — de que regimes de baixa volatilidade tendem a estar associados a condições mais favoráveis de investimento, enquanto regimes de alta volatilidade concentram choques adversos e episódios de estresse. Assim, adota-se uma regra de market timing baseada exclusivamente na volatilidade prevista: o investidor permanece comprado (long 100%) na VALE3 quando a volatilidade condicional prevista para o período seguinte, $\hat{\sigma}_{t+1|t}$, encontra-se abaixo de um limiar, e permanece fora do mercado quando essa volatilidade supera o limiar.

Formalmente, o limiar é definido como um quantil q da distribuição histórica de $\hat{\sigma}$. Foram avaliadas duas parametrizações: uma estratégia mais restritiva ($q=0,3$), que busca operar apenas em ambientes de risco baixo, e uma versão mais permissiva ($q=0,6$), que amplia o conjunto de períodos em que o investidor permanece alocado. Os resultados são apresentados nos Gráficos 15 e 16 e nas Tabelas 11 e 12, respectivamente, com comparação direta ao benchmark buy-and-hold (long 100%), tanto em retorno bruto quanto líquido de custos.

Gráfico 15 — Retorno acumulado: Estratégia GARCH ($q=0.3$) vs Long 100%



Gráfico 16 — Retorno acumulado: Estratégia GARCH ($q=0.6$) vs Long 100%

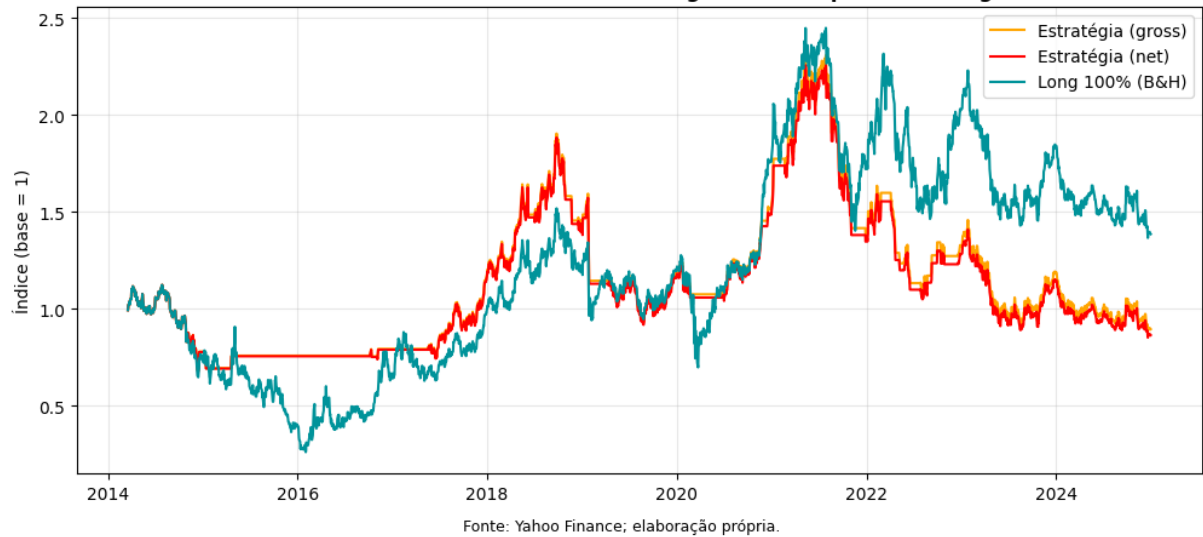


Tabela 11 — Retorno acumulado: GARCH (q=0.3) - gross vs net vs Long 100%

	GARCH (q=0.3)
Obs. OOS	2610
Benchmark	Long 100% (buy-and-hold)
Retorno acumulado (Gross)	13.07%
Retorno acumulado (Net)	9.08%
Retorno acumulado (B&H)	38.58%
Volatilidade (Gross)	15.34%
Volatilidade (Net)	15.34%
Volatilidade (B&H)	43.23%
Sharpe (Gross)	0.154
Sharpe (Net)	0.131
Sharpe (B&H)	0.291
Max Drawdown (Gross)	-33.02%
Max Drawdown (Net)	-33.71%
Max Drawdown (B&H)	-76.82%
% períodos investido	30.00%
Turnover médio	0.046
Custo médio diário	0.000014

Fonte: Yahoo Finance; elaboração própria.

Tabela 12 — Retorno acumulado: GARCH ($q=0.6$) - gross vs net vs Long 100%

	GARCH ($q=0.6$)
Obs. OOS	2610
Benchmark	Long 100% (buy-and-hold)
Retorno acumulado (Gross)	-10.57%
Retorno acumulado (Net)	-13.63%
Retorno acumulado (B&H)	38.58%
Volatilidade (Gross)	25.91%
Volatilidade (Net)	25.91%
Volatilidade (B&H)	43.23%
Sharpe (Gross)	0.091
Sharpe (Net)	0.078
Sharpe (B&H)	0.291
Max Drawdown (Gross)	-61.81%
Max Drawdown (Net)	-62.28%
Max Drawdown (B&H)	-76.82%
% períodos investido	60.00%
Turnover médio	0.044
Custo médio diário	0.000013

Fonte: Yahoo Finance; elaboração própria.

4.4.2.1 Estratégia GARCH conservadora ($q=0,3$)

No cenário conservador, a estratégia mantém exposição ao ativo em aproximadamente 30% dos períodos fora da amostra, refletindo elevada seletividade. Essa característica resulta em um perfil de risco significativamente inferior ao buy-and-hold: a volatilidade anualizada cai de cerca de 43% no benchmark para aproximadamente 15%, e o drawdown máximo reduz-se substancialmente (de aproximadamente -76,8% no buy-and-hold para cerca de -33% na estratégia). Assim, sob a ótica de controle de risco, a regra baseada em baixa volatilidade se mostra eficiente ao evitar exposição prolongada em regimes turbulentos.

Entretanto, os ganhos em proteção não se traduzem em desempenho econômico superior no período analisado. O retorno acumulado permanece inferior

ao benchmark, com retorno líquido de cerca de 9% (versus 38,6% no buy-and-hold). O índice de Sharpe também fica abaixo do benchmark, indicando que, apesar da redução expressiva do risco, o custo de oportunidade associado à permanência fora do mercado durante longos períodos reduz a eficiência risco-retorno no horizonte considerado.

4.4.2.2 Estratégia GARCH permissiva ($q=0,6$)

Ao elevar o limiar para $q=0,6$, a estratégia aumenta sua exposição para aproximadamente 60% do tempo, tornando-se menos defensiva. Contudo, essa flexibilização não melhora o desempenho: o retorno acumulado passa a ser negativo, inclusive antes dos custos em alguns trechos, com deterioração adicional na versão líquida.

Em termos de risco, embora a volatilidade permaneça inferior à do buy-and-hold, a estratégia passa a se expor com maior frequência a períodos de instabilidade intermediária — regimes nos quais a volatilidade já se elevou, mas ainda não aciona o filtro com rapidez suficiente para evitar a realização de perdas relevantes. Como consequência, o drawdown se aprofunda de forma importante e o Sharpe permanece abaixo do benchmark, sugerindo que o ganho informacional trazido por $\hat{\sigma}_{t+1|t}$, isoladamente, não é suficiente para operar como filtro direcional eficiente.

4.4.2.3 Interpretação dos resultados

Os resultados confirmam um ponto central da literatura: modelos GARCH são particularmente adequados para previsão e gestão de risco, mas não foram concebidos como geradores diretos de sinais de retorno. O modelo captura a persistência da variância e o agrupamento de volatilidade (volatility clustering), mas sua estrutura é essencialmente voltada à dinâmica de segunda ordem, isto é, à incerteza condicional, e não à média condicional dos retornos.

Além disso, em ativos com prêmio de risco positivo no longo prazo, como ações, estratégias excessivamente seletivas tendem a reduzir a exposição justamente em momentos de forte recuperação pós-choques, sacrificando parte do retorno acumulado. Em termos econômicos, isso sugere que o GARCH, quando utilizado isoladamente, pode funcionar como mecanismo de proteção (redução de drawdowns), mas apresenta limitação como estratégia de retorno.

4.4.3 Estratégias baseadas em Random Forest

Após a avaliação do desempenho preditivo fora da amostra, os outputs do modelo Random Forest foram empregados diretamente na construção de estratégias de trading baseadas em regimes de risco, seguindo a mesma lógica conceitual adotada nas estratégias baseadas em GARCH. Diferentemente das abordagens lineares e VAR, o Random Forest não é utilizado aqui como gerador de sinal direcional de retorno, mas como filtro de risco, explorando sua capacidade de capturar relações não lineares e interações complexas entre variáveis explicativas.

O modelo é treinado para prever a magnitude do retorno absoluto futuro, $|\widehat{r_{t+1}}|$, interpretada como proxy de risco ou volatilidade realizada no período seguinte. A regra de decisão é, portanto, do tipo risk-on/risk-off: o investidor permanece comprado na ação da VALE3 apenas quando o risco previsto é relativamente baixo, permanecendo fora do mercado nos demais períodos.

Formalmente, define-se um limiar com base no quantil q da distribuição fora da amostra de $|\widehat{r_{t+1}}|$. A estratégia assume posição comprada quando:

$|\widehat{r_{t+1}}| \leq Q_q$, e posição nula caso contrário. Essa construção permite comparabilidade direta com as estratégias GARCH e possibilita avaliar diferentes níveis de seletividade por meio do parâmetro qq . Foram analisados dois cenários: $q=0,30$ e $q=0,60$.

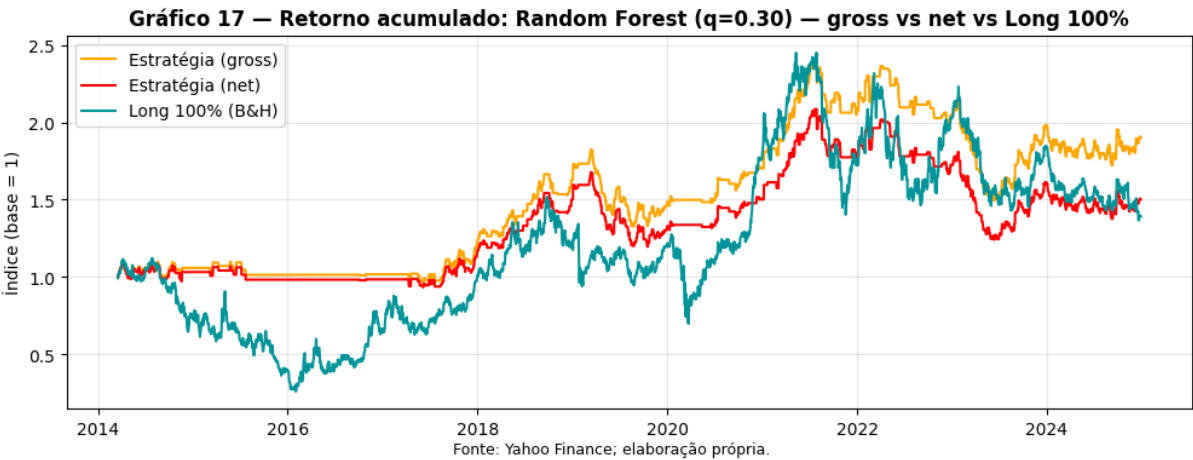
4.4.3.1 Random Forest — estratégia com $q=0,30$

A versão mais restritiva da estratégia, com $q=0,30$, mantém exposição ao ativo em aproximadamente 30% dos períodos fora da amostra, concentrando-se em janelas classificadas pelo modelo como de menor risco esperado. Os resultados são apresentados na Tabela 13 e no Gráfico 17.

Tabela 13 — Retorno acumulado: RF (q=0.3) - gross vs net vs Long 100%

Random Forest (q=0.30)	
Obs. OOS	2609
Benchmark	Long 100% (buy-and-hold)
Retorno acumulado (Gross)	89.91%
Retorno acumulado (Net)	49.98%
Retorno acumulado (B&H)	39.06%
Volatilidade (Gross)	16.35%
Volatilidade (Net)	16.34%
Volatilidade (B&H)	43.23%
Sharpe (Gross)	0.460
Sharpe (Net)	0.321
Sharpe (B&H)	0.292
Max Drawdown (Gross)	-37.78%
Max Drawdown (Net)	-40.48%
Max Drawdown (B&H)	-76.82%
% períodos investido	30.01%
Turnover médio	0.302
Custo médio diário	0.000090

Fonte: Yahoo Finance; elaboração própria.



Do ponto de vista econômico, a estratégia apresenta desempenho robusto. O retorno acumulado bruto atinge 89,9%, superando o benchmark de buy-and-hold

(39,1%). Mesmo após a incorporação de custos de transação, o retorno líquido permanece superior ao benchmark, em torno de 50,0%, indicando que o ganho não é puramente ilusório ou fruto de sobreajuste.

Em termos de risco, observa-se melhora substancial: a volatilidade anualizada da estratégia (cerca de 16,3%) é significativamente inferior à do buy-and-hold (43,2%), e o máximo drawdown é reduzido de aproximadamente -76,8% no benchmark para cerca de -40,5% na versão líquida da estratégia. Essa redução é coerente com o mecanismo de funcionamento do modelo, que evita exposição em períodos associados a picos de instabilidade.

O índice de Sharpe líquido (0,321) supera o do benchmark (0,292), reforçando que a estratégia não apenas reduz risco absoluto, mas também melhora a eficiência risco-retorno. Em síntese, com $q=0,30$, o Random Forest atua como um filtro de risco eficaz, gerando ganhos econômicos relevantes com menor exposição contínua ao mercado.

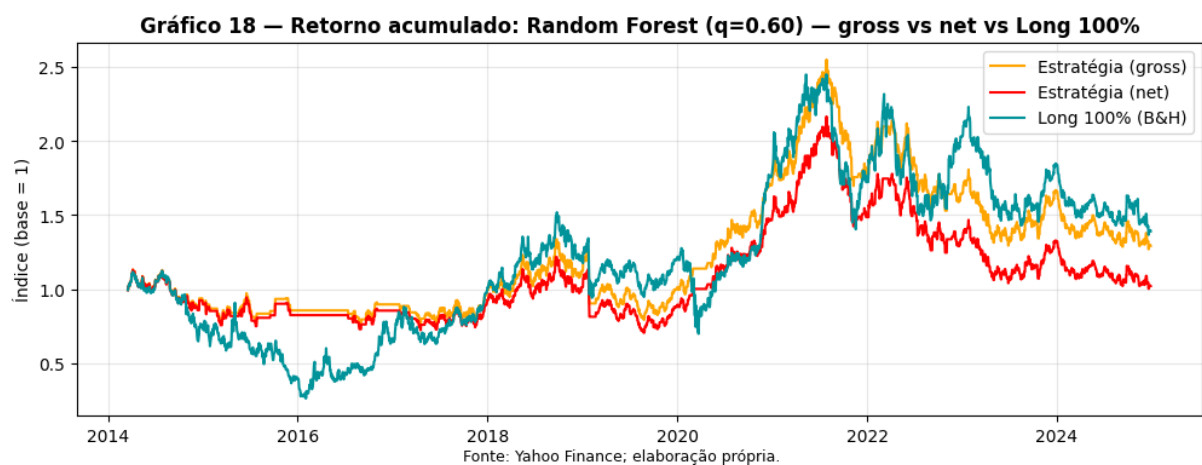
4.4.3.2 Random Forest — estratégia com $q=0,60$

Ao elevar o limiar para $q=0,60$, a estratégia torna-se mais permissiva, permanecendo investida em aproximadamente 60% dos períodos fora da amostra. Os resultados correspondentes estão apresentados na Tabela 14 e no Gráfico 18.

Tabela 14 — Retorno acumulado: RF (q=0.6) - gross vs net vs Long 100%

	Random Forest (q=0.60)
Obs. OOS	2609
Benchmark	Long 100% (buy-and-hold)
Retorno acumulado (Gross)	29.13%
Retorno acumulado (Net)	1.98%
Retorno acumulado (B&H)	39.06%
Volatilidade (Gross)	26.94%
Volatilidade (Net)	26.93%
Volatilidade (B&H)	43.23%
Sharpe (Gross)	0.229
Sharpe (Net)	0.145
Sharpe (B&H)	0.292
Max Drawdown (Gross)	-50.23%
Max Drawdown (Net)	-53.72%
Max Drawdown (B&H)	-76.82%
% períodos investido	59.98%
Turnover médio	0.302
Custo médio diário	0.000090

Fonte: Yahoo Finance; elaboração própria.



Nesse cenário, o desempenho econômico se deteriora de forma significativa. O retorno acumulado bruto é de 29,1%, já inferior ao benchmark, e o retorno líquido

cai para apenas 2,0%, ficando muito aquém do buy-and-hold. O índice de Sharpe líquido (0,145) também se mostra inferior ao do benchmark, indicando perda clara de eficiência risco–retorno.

Embora a estratégia ainda reduza volatilidade e drawdown em relação à exposição passiva, essa melhoria de risco não é suficiente para compensar o efeito dos custos de transação e da maior frequência de alternância de posição. O aumento do tempo investido eleva o impacto do turnover e amplifica o custo econômico associado a entradas e saídas mal sincronizadas.

Esses resultados ilustram um ponto central: boa capacidade preditiva do risco não garante, por si só, uma estratégia lucrativa, especialmente quando o critério de entrada é excessivamente relaxado. No caso do Random Forest, o ganho econômico está fortemente condicionado à seletividade do regime adotado.

4.4.3.3 Random Forest com reestimação periódica (RF refit=200)

Como etapa adicional de robustez e maior aproximação do ambiente operacional, a mesma lógica de estratégia por regimes foi aplicada a um Random Forest com reestimação periódica a cada 200 observações, doravante denominado RF refit=200. O objetivo é reduzir o risco de envelhecimento paramétrico, permitindo que o modelo incorpore mudanças estruturais e transições de regime ao longo do tempo.

Mantém-se o mesmo arcabouço de decisão: o modelo prevê $\widehat{r_{t+1}}$, e a estratégia assume posição comprada apenas quando o risco previsto se encontra abaixo do quantil q . Foram novamente avaliadas as configurações $q=0,30$ e $q=0,60$.

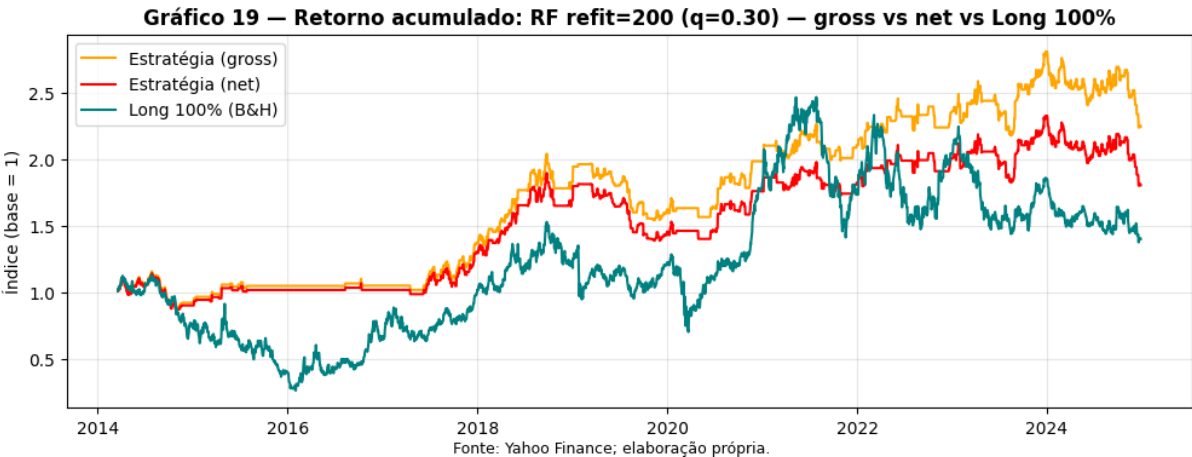
4.6.3.4 RF refit=200 — estratégia com $q=0,30$

No caso mais seletivo, com $q=0,30$, a estratégia permanece investida em aproximadamente 30% dos períodos, priorizando janelas de menor risco previsto. Os resultados estão apresentados na Tabela 15 e no Gráfico 19.

Tabela 15 — Retorno acumulado: RF refit=200 (q=0.3) - gross vs net vs Long 100%

Random Forest q=0.30	
Obs. OOS	2608
Retorno acumulado (Gross)	124.32%
Retorno acumulado (Net)	80.50%
Retorno acumulado (B&H)	40.00%
Volatilidade (Gross)	16.31%
Volatilidade (Net)	16.31%
Volatilidade (B&H)	43.24%
Sharpe (Gross)	0.560
Sharpe (Net)	0.431
Sharpe (B&H)	0.293
Max Drawdown (Gross)	-24.46%
Max Drawdown (Net)	-26.78%
Max Drawdown (B&H)	-76.82%
% períodos investido	30.02%
Turnover médio	0.278
Custo médio diário	0.000083
q	0.30
threshold	0.014690

Fonte: Yahoo Finance; elaboração própria.



O desempenho econômico desta configuração é o melhor entre todas as estratégias baseadas em Random Forest. O retorno acumulado bruto atinge 124,3%, e o retorno líquido permanece elevado, em torno de 80,5%, praticamente o dobro do retorno do buy-and-hold (40,0%). O índice de Sharpe líquido (0,431) supera com folga o benchmark (0,293), indicando ganho claro em eficiência risco–retorno.

Além disso, o perfil de risco é significativamente mais controlado: a volatilidade permanece baixa (16,3%) e o máximo drawdown líquido é reduzido para cerca de -26,8%, um valor substancialmente inferior ao observado nas estratégias passivas e mesmo nas versões sem refit.

Esses resultados sugerem que a reestimação periódica melhora a capacidade do Random Forest de se adaptar a mudanças estruturais, reduzindo o custo de obsolescência do modelo e aumentando a qualidade econômica do filtro de regimes quando operado de forma seletiva.

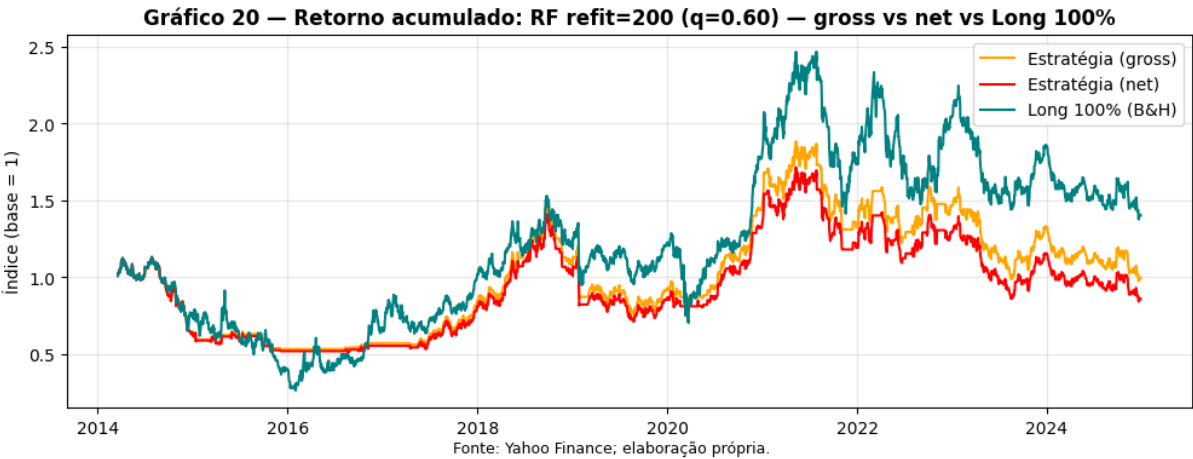
4.6.3.5 RF refit=200 — estratégia com $q=0,60$

Ao elevar o limiar para $q=0,60$, a estratégia passa a permanecer investida em cerca de 60% dos períodos fora da amostra. Os resultados correspondentes são apresentados na Tabela 16 e no Gráfico 20.

Tabela 16 — Retorno acumulado: RF refit=200 (q=0.6) - gross vs net vs Long 100%

Random Forest q=0.60	
Obs. OOS	2608
Retorno acumulado (Gross)	-0.93%
Retorno acumulado (Net)	-14.34%
Retorno acumulado (B&H)	40.00%
Volatilidade (Gross)	26.35%
Volatilidade (Net)	26.35%
Volatilidade (B&H)	43.24%
Sharpe (Gross)	0.132
Sharpe (Net)	0.078
Sharpe (B&H)	0.293
Max Drawdown (Gross)	-53.97%
Max Drawdown (Net)	-55.12%
Max Drawdown (B&H)	-76.82%
% períodos investido	60.01%
Turnover médio	0.186
Custo médio diário	0.000056
q	0.60
threshold	0.017990

Fonte: Yahoo Finance; elaboração própria.



Nesse cenário, o desempenho se deteriora novamente. O retorno acumulado bruto torna-se praticamente nulo ($-0,9\%$), e o retorno líquido é negativo ($-14,3\%$), ficando bem abaixo do benchmark. O índice de Sharpe líquido ($0,078$) confirma a baixa eficiência econômica da estratégia nessa configuração.

Embora a reestimação periódica reduza parcialmente o risco de envelhecimento do modelo, o problema estrutural permanece: ao relaxar excessivamente o critério de entrada, a estratégia se expõe com maior frequência a períodos de transição de regime, ampliando o impacto dos custos operacionais e do chamado *timing cost* — isto é, o custo associado ao atraso inevitável na identificação de mudanças abruptas de risco.

4.4.4 Estratégias de Trading com LSTM (regimes de volatilidade)

A partir das previsões fora da amostra geradas pelo modelo LSTM, foram construídas estratégias de negociação baseadas em regimes de risco, seguindo a mesma lógica adotada para os modelos GARCH e Random Forest. Em vez de explorar sinais direcionais de retorno, as estratégias utilizam a magnitude prevista do retorno absoluto no período seguinte, $|\widehat{r_{t+1}}|$, interpretada como proxy de volatilidade e instabilidade esperada.

A regra operacional é do tipo *risk-on/risk-off*: o investidor permanece comprado na ação da VALE3 apenas quando o risco previsto pelo modelo encontra-se abaixo de um limiar definido por um quantil qqq da distribuição das previsões fora da amostra. Caso contrário, a posição é zerada. Essa construção permite avaliar se o LSTM é capaz de atuar como um filtro eficiente de regimes de risco, reduzindo exposição em períodos adversos sem comprometer excessivamente a captura de retorno.

Foram analisadas duas configurações:

- $q=0,30$, estratégia mais restritiva, investida em aproximadamente 30% dos períodos;
- $q=0,60$, estratégia mais permissiva, investida em cerca de 60% dos períodos.

Os resultados são apresentados nas Tabelas 17 e 18 e ilustrados nos Gráficos 21 e 22.

Tabela 17 — Retorno acumulado: LSTM (q=0.30) — gross vs net vs Long 100%

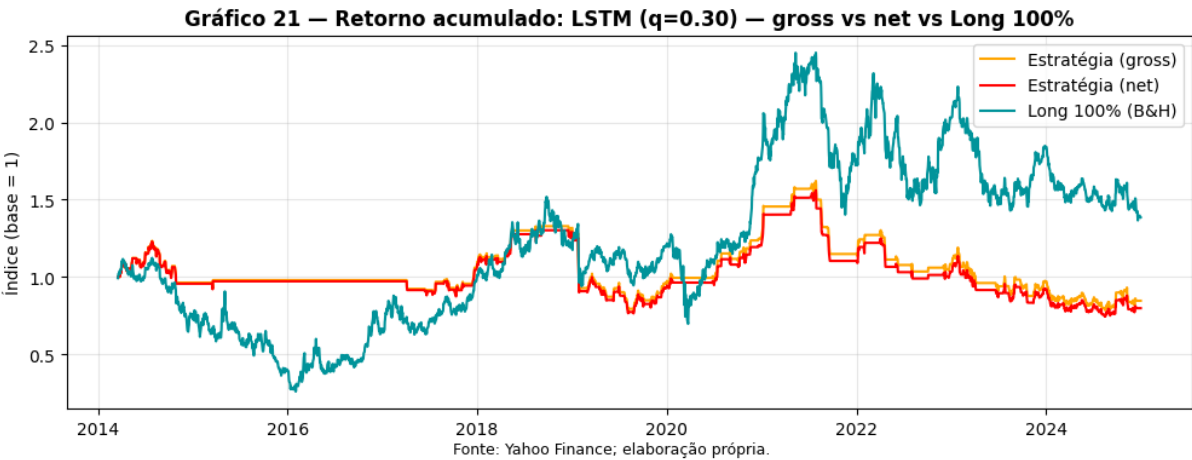
	LSTM (q=0.30)
Obs. OOS	2610
Benchmark	Long 100% (buy-and-hold)
Retorno acumulado (Gross)	-15.30%
Retorno acumulado (Net)	-20.10%
Retorno acumulado (B&H)	38.58%
Volatilidade (Gross)	18.41%
Volatilidade (Net)	18.42%
Volatilidade (B&H)	43.22%
Sharpe (Gross)	0.010
Sharpe (Net)	-0.021
Sharpe (B&H)	0.291
Max Drawdown (Gross)	-51.43%
Max Drawdown (Net)	-52.22%
Max Drawdown (B&H)	-76.82%
% períodos investido	30.00%
Turnover médio	0.074
Custo médio diário	0.000022
Threshold (quantil)	0.015944

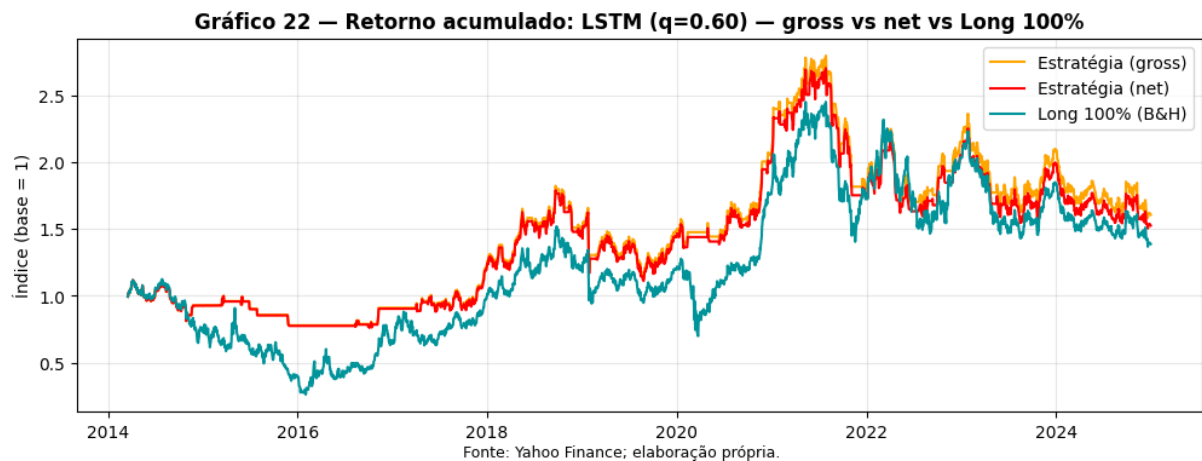
Fonte: Yahoo Finance; elaboração própria.

Tabela 18 — Retorno acumulado: LSTM (q=0.60) — gross vs net vs Long 100%

	LSTM (q=0.60)
Obs. OOS	2610
Benchmark	Long 100% (buy-and-hold)
Retorno acumulado (Gross)	60.55%
Retorno acumulado (Net)	52.44%
Retorno acumulado (B&H)	38.58%
Volatilidade (Gross)	25.65%
Volatilidade (Net)	25.65%
Volatilidade (B&H)	43.22%
Sharpe (Gross)	0.310
Sharpe (Net)	0.290
Sharpe (B&H)	0.291
Max Drawdown (Gross)	-43.37%
Max Drawdown (Net)	-44.38%
Max Drawdown (B&H)	-76.82%
% períodos investido	60.00%
Turnover médio	0.066
Custo médio diário	0.000020
Threshold (quantil)	0.0169

Fonte: Yahoo Finance; elaboração própria.





4.4.4.1 Estratégia LSTM (q = 0,30): filtro rígido de risco e perda econômica

No cenário mais conservador, com $q=0,30$, a estratégia permanece investida em apenas 30% dos períodos, com threshold em torno de 0,0159 (Tabela 17). Embora esse filtro reduza substancialmente a exposição ao risco, os resultados indicam desempenho econômico insatisfatório ao longo da amostra fora da amostra.

O retorno acumulado bruto foi de $-15,3\%$, deteriorando-se para $-20,1\%$ após a inclusão de custos de transação, enquanto o benchmark buy-and-hold apresentou retorno positivo de $38,6\%$. O Gráfico 21 evidencia que a estratégia participa pouco dos principais ciclos de alta do ativo, sobretudo no período pós-2020, permanecendo fora do mercado em momentos cruciais de recuperação.

Do ponto de vista de risco, observa-se redução expressiva da volatilidade anualizada ($18,4\%$ contra $43,2\%$ do benchmark) e melhora no drawdown máximo (cerca de -52% frente a $-76,8\%$). No entanto, esses ganhos defensivos não se traduzem em eficiência econômica: o índice de Sharpe líquido é negativo ($-0,021$), indicando que a estratégia não remunera adequadamente o risco assumido.

Em síntese, o caso $q=0,30$ ilustra um trade-off clássico em estratégias baseadas em regimes: o modelo evita risco, mas evita também retorno. O excesso de cautela gera erro de timing e impede a captura do drift positivo do ativo, resultando em desempenho líquido inferior ao benchmark.

4.4.4.2 Estratégia LSTM (q = 0,60): desempenho econômico competitivo com menor risco

Ao elevar o limiar para $q=0,60$, a estratégia torna-se significativamente mais permissiva, permanecendo investida em aproximadamente 60% dos períodos, com

threshold próximo de 0,0169 (Tabela 18). Essa alteração modifica substancialmente os resultados econômicos.

O retorno acumulado bruto atinge 60,6%, enquanto o retorno líquido permanece elevado em 52,4%, superando de forma clara o buy-and-hold (38,6%). O Gráfico 22 mostra que a estratégia passa a capturar os principais ciclos de alta do ativo, ao mesmo tempo em que reduz parcialmente a exposição em períodos de maior deterioração.

Os indicadores de risco confirmam a melhoria do perfil da estratégia. A volatilidade anualizada cai para cerca de 25,6%, bem abaixo dos 43,2% do benchmark, enquanto o drawdown máximo líquido é reduzido para aproximadamente -44%. O índice de Sharpe líquido (0,29) torna-se praticamente equivalente ao do buy-and-hold (0,291), porém com risco significativamente menor.

Esses resultados sugerem que, quando operado com um filtro menos agressivo, o LSTM consegue cumprir melhor seu papel como moderador de risco: preserva a participação nos períodos favoráveis e reduz parcialmente a exposição em regimes mais instáveis, gerando desempenho econômico competitivo.

4.6.4.3 Interpretação geral: limiar, defasagem e custo de timing

A comparação entre $q=0,30$ e $q=0,60$ evidencia que o desempenho das estratégias baseadas em LSTM é altamente sensível ao nível de seletividade adotado. Filtros excessivamente rígidos tendem a amplificar erros de timing, excluindo o investidor de janelas críticas de retorno — um problema particularmente relevante em mercados acionários, nos quais uma parcela significativa do retorno de longo prazo concentra-se em períodos relativamente curtos.

Além disso, há um custo estrutural associado à defasagem do sinal. Como o LSTM utiliza informação histórica extensa e gera previsões suavizadas, sua resposta a mudanças abruptas de regime é, por construção, gradual. Isso implica entradas tardias em rallies e saídas atrasadas em choques, fenômeno que se intensifica quando a estratégia busca ser excessivamente seletiva.

4.4.5 Estratégias híbridas: integração entre sinal direcional e regime de volatilidade

Como etapa final dos experimentos empíricos, foram avaliadas estratégias híbridas que combinam um modelo direcional com um filtro de regime de volatilidade, buscando integrar previsão de retorno esperado e controle de risco em uma única

regra operacional. Diferentemente das seções anteriores — que exploraram estratégias puras —, o foco aqui é testar se a coordenação entre “quando operar” e “em que direção operar” gera ganhos econômicos adicionais ou melhora o perfil risco–retorno.

Foram consideradas duas combinações:

1. Random Forest (direcional) + GARCH (regime)
2. VAR (direcional) + GARCH (regime)

Em ambos os casos, a lógica é idêntica:

- o modelo direcional indica sinal de posição (long);
- o GARCH atua como filtro de regime, permitindo execução apenas quando a volatilidade condicional prevista se encontra abaixo de um determinado quantil q .

Foram testados dois níveis de seletividade do regime: $q = 0,30$ (restritivo) e $q = 0,60$ (permissivo).

4.4.5.1 Random Forest (direcional) + GARCH (regime)

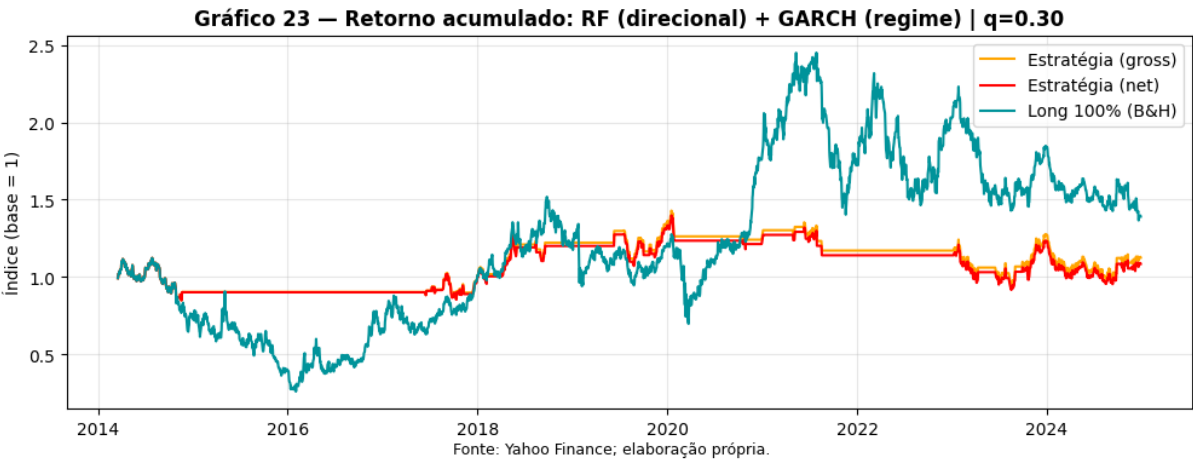
Os resultados da estratégia RF + GARCH indicam que a combinação é funcional como mecanismo de controle de risco, mas não gera ganhos econômicos adicionais consistentes em termos de retorno absoluto.

No caso $q = 0,30$ (Tabela 19 e Gráfico 23), a estratégia opera em aproximadamente 30% dos períodos e apresenta retorno acumulado positivo, embora inferior ao benchmark buy-and-hold. O principal benefício aparece no perfil de risco: a volatilidade e o drawdown são substancialmente reduzidos, resultando em uma trajetória de retorno mais estável. Trata-se, portanto, de uma estratégia com caráter defensivo, adequada para mitigação de risco, mas não para maximização de retorno.

Tabela 19 — Estratégia combinada (RF + GARCH): q=0.30

	RF+GARCH (q=0.30)
Obs. OOS	2609
Retorno acumulado (Gross)	12.51%
Retorno acumulado (Net)	8.50%
Retorno acumulado (B&H)	39.06%
Volatilidade (Gross)	15.34%
Volatilidade (Net)	15.34%
Volatilidade (B&H)	43.23%
Sharpe (Gross)	0.151
Sharpe (Net)	0.128
Sharpe (B&H)	0.292
Max Drawdown (Gross)	-33.59%
Max Drawdown (Net)	-34.27%
Max Drawdown (B&H)	-76.82%
% períodos investido	30.01%
% períodos short	0.0000
Turnover médio	0.046
Custo médio diário	0.000014
q (regime)	0.3000
Threshold sigma	0.0209

Fonte: Yahoo Finance; elaboração própria.



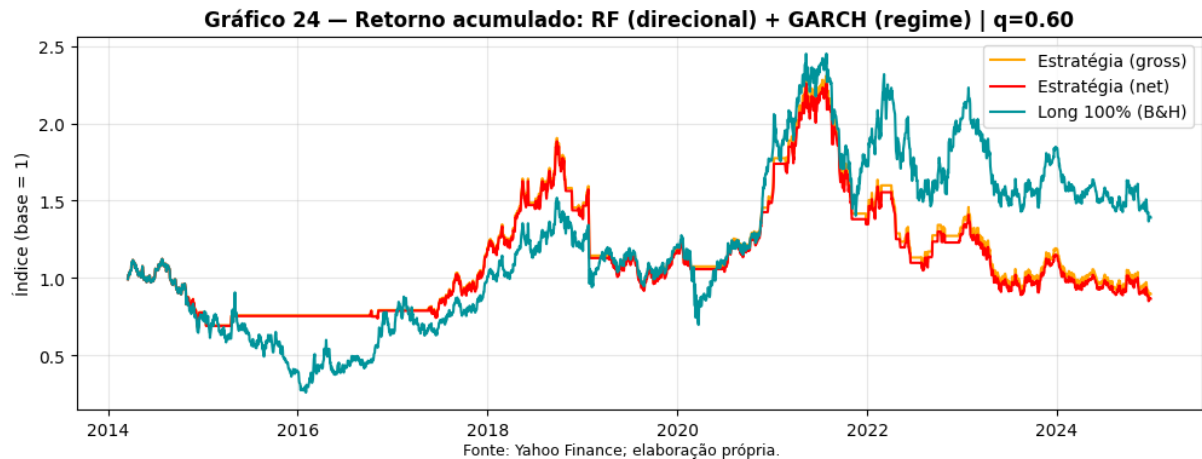
Quando o regime é relaxado para $q = 0,60$ (Tabela 20 e Gráfico 24), o desempenho econômico se deteriora. A maior exposição ao mercado faz com que o

filtro de volatilidade perca eficácia, permitindo entradas em ambientes ruidosos nos quais o sinal direcional do Random Forest não é suficiente para compensar erros de timing e fricções operacionais. O retorno líquido torna-se negativo e o drawdown volta a se aproximar do benchmark, indicando que a combinação deixa de cumprir seu papel disciplinador.

Tabela 20— Estratégia combinada (RF + GARCH): $q=0.60$

	RF+GARCH ($q=0.60$)
Obs. OOS	2609
Retorno acumulado (Gross)	-10.25%
Retorno acumulado (Net)	-13.35%
Retorno acumulado (B&H)	39.06%
Volatilidade (Gross)	25.91%
Volatilidade (Net)	25.91%
Volatilidade (B&H)	43.23%
Sharpe (Gross)	0.093
Sharpe (Net)	0.080
Sharpe (B&H)	0.292
Max Drawdown (Gross)	-61.81%
Max Drawdown (Net)	-62.28%
Max Drawdown (B&H)	-76.82%
% períodos investido	59.98%
% períodos short	0.0000
Turnover médio	0.045
Custo médio diário	0.000013
q (regime)	0.6000
Threshold sigma	0.0243

Fonte: Yahoo Finance; elaboração própria.



Em síntese, o RF + GARCH funciona apenas sob regime restritivo, atuando mais como mecanismo de proteção do que como estratégia geradora de alfa.

4.4.5.2 VAR (direcional) + GARCH (regime)

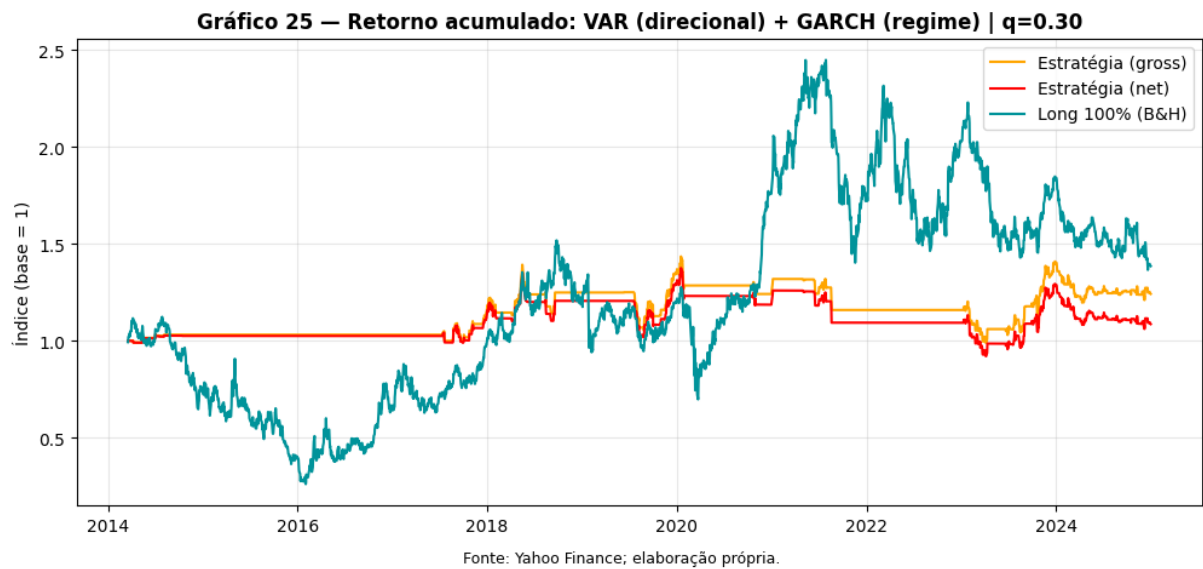
A estratégia VAR + GARCH apresenta um contraste ainda mais claro entre os dois níveis de seletividade do regime.

Sob $q = 0,30$ (Tabela 21 e Gráfico 25), o filtro de volatilidade é extremamente rigoroso, resultando em participação muito limitada no mercado. Ainda assim, a estratégia alcança retorno acumulado positivo e, principalmente, uma redução expressiva de volatilidade e drawdown em relação ao buy-and-hold. O VAR atua apenas em janelas de mercado particularmente “limpas”, o que confirma o papel do GARCH como componente disciplinador essencial para tornar o uso do VAR operacionalmente viável.

Tabela 21 — Estratégia combinada (VAR + GARCH): $q=0.30$

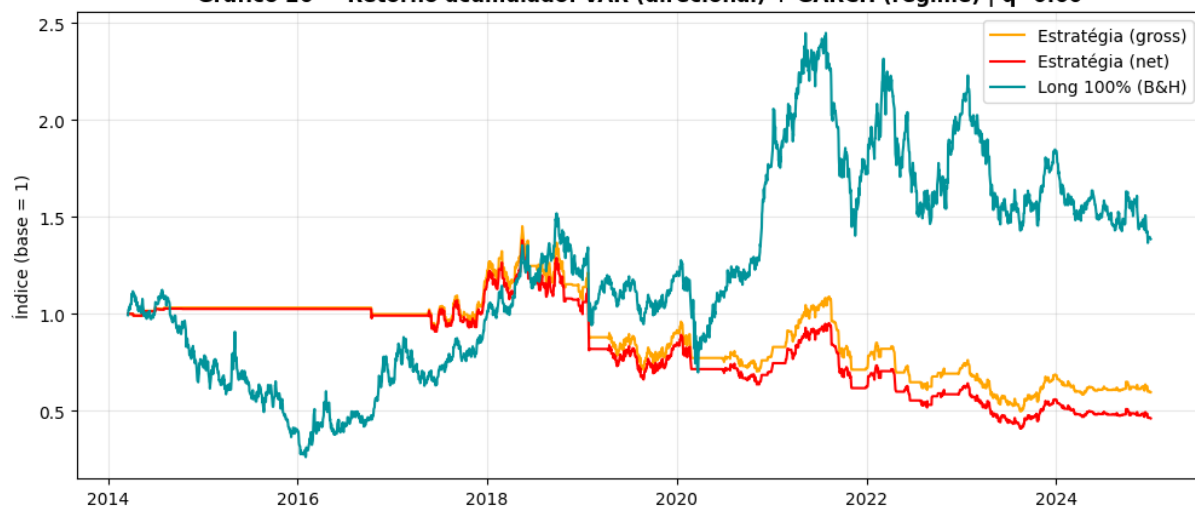
	VAR+GARCH
Obs. OOS	2610
Benchmark	Long 100% (buy-and-hold)
Retorno acumulado (Gross)	24.22%
Retorno acumulado (Net)	8.64%
Retorno acumulado (B&H)	38.58%
Volatilidade (Gross)	11.33%
Volatilidade (Net)	11.33%
Volatilidade (B&H)	43.22%
Sharpe (Gross)	0.242
Sharpe (Net)	0.127
Sharpe (B&H)	0.291
Max Drawdown (Gross)	-31.05%
Max Drawdown (Net)	-33.09%
Max Drawdown (B&H)	-76.82%
% períodos investido	0.17%
% períodos short	0.000000
Turnover médio	0.103
Custo médio diário	0.000051
q (regime)	0.300000
Threshold sigma	0.020919

Fonte: Yahoo Finance; elaboração própria.



Por outro lado, quando o regime é relaxado para $q = 0,60$ (Tabela 22 e Gráfico 26), o desempenho torna-se fortemente negativo. Mesmo com baixa frequência de operação, a estratégia passa a entrar em momentos inadequados, caracterizados por maior instabilidade e menor previsibilidade direcional. O resultado é uma combinação de perdas econômicas, drawdowns elevados e forte degradação do desempenho líquido, evidenciando que o VAR é particularmente sensível a regimes de alta volatilidade quando não adequadamente filtrado.

Gráfico 26 — Retorno acumulado: VAR (direcional) + GARCH (regime) | $q=0.60$



Fonte: Yahoo Finance; elaboração própria.

Tabela 22 — Estratégia combinada (VAR + GARCH): $q=0.60$

	VAR+GARCH
Obs. OOS	2610
Benchmark	Long 100% (buy-and-hold)
Retorno acumulado (Gross)	-40.48%
Retorno acumulado (Net)	-53.98%
Retorno acumulado (B&H)	38.58%
Volatilidade (Gross)	21.39%
Volatilidade (Net)	21.40%
Volatilidade (B&H)	43.22%
Sharpe (Gross)	-0.123
Sharpe (Net)	-0.239
Sharpe (B&H)	0.291
Max Drawdown (Gross)	-65.75%
Max Drawdown (Net)	-70.38%
Max Drawdown (B&H)	-76.82%
% períodos investido	0.38%
% períodos short	0.000000
Turnover médio	0.197
Custo médio diário	0.000098
q (regime)	0.600000
Threshold sigma	0.024285

Fonte: Yahoo Finance; elaboração própria.

Assim, a evidência empírica sugere que o VAR só funciona como componente direcional quando fortemente condicionado por um filtro de regime rigoroso; fora disso, torna-se economicamente inviável.

4.5 Encerramento da seção de Resultados e Discussão

O conjunto de resultados apresentados ao longo desta seção permite extrair conclusões claras e consistentes sobre o uso de modelos econométricos e de aprendizado de máquina em estratégias quantitativas aplicadas à VALE3.

Primeiro, observa-se que bons resultados preditivos não se traduzem automaticamente em bons resultados econômicos. Estratégias excessivamente seletivas podem reduzir risco, mas eliminar retorno; estratégias permissivas, por sua vez, tendem a sofrer com custos de timing, fricções operacionais e erros de regime.

Segundo, os experimentos mostram que modelos de volatilidade desempenham melhor o papel de filtros de risco do que de geradores diretos de retorno, enquanto modelos direcionais exigem forte controle de regime para não se tornarem reféns de choques e instabilidade.

Terceiro, as estratégias híbridas confirmam que a combinação entre previsão de direção e controle de risco é conceitualmente correta, mas sua eficácia depende criticamente da parametrização do regime. Em todos os casos analisados, filtros mais rígidos ($q = 0,30$) produziram resultados mais estáveis e economicamente defensáveis do que versões permissivas ($q = 0,60$).

Por fim, o trabalho evidencia que o maior ganho prático dessas abordagens não está em “bater o mercado” de forma sistemática, mas em modular exposição, reduzir drawdowns e suavizar a trajetória do portfólio, especialmente em ativos sujeitos a ciclos longos, choques exógenos e elevada volatilidade, como é o caso da VALE3.

Esses achados reforçam a visão de que estratégias quantitativas eficazes não dependem de um único modelo “ótimo”, mas sim de arquiteturas simples, bem disciplinadas e economicamente interpretáveis, nas quais previsão e gestão de risco são tratadas como problemas distintos e complementares.

5 Conclusão

Este trabalho teve como objetivo investigar a relação entre o preço internacional do minério de ferro e os retornos da ação da VALE3, avaliando em que medida essa relação pode ser explorada de forma sistemática por meio de modelos econométricos tradicionais e métodos de aprendizado de máquina, tanto em termos preditivos quanto na construção de estratégias quantitativas de negociação. Ao longo da análise, buscou-se manter um equilíbrio entre rigor estatístico, interpretação econômica e relevância prática, evitando soluções excessivamente complexas ou ancoradas exclusivamente em eventos extremos.

Os resultados obtidos ao longo das seções empíricas permitem extrair algumas conclusões centrais. Em primeiro lugar, confirmou-se que o minério de ferro possui papel relevante como variável explicativa da dinâmica dos retornos da VALE3, ainda que essa relação seja instável ao longo do tempo e sujeita a não linearidades. Modelos lineares simples mostraram-se úteis como benchmark interpretável, enquanto abordagens multivariadas e não lineares — como VAR e Random Forest — apresentaram maior capacidade de capturar padrões direcionais, especialmente quando avaliadas fora da amostra.

Em segundo lugar, os experimentos com modelos de volatilidade e regimes de risco reforçaram um ponto amplamente discutido na literatura financeira: prever risco e prever retorno são problemas distintos. O modelo GARCH demonstrou elevada eficácia na identificação de períodos de maior instabilidade, reduzindo volatilidade e drawdowns quando utilizado como filtro de exposição, mas mostrou limitações claras como gerador autônomo de sinais direcionais. De forma semelhante, o LSTM aplicado à previsão da magnitude dos retornos apresentou desempenho sensível à escolha do limiar de ativação, funcionando melhor como moderador parcial de risco do que como filtro excessivamente restritivo.

A etapa final do trabalho — dedicada às estratégias híbridas — sintetiza de forma clara essa separação conceitual entre direção e regime. As combinações entre modelos direcionais (VAR e Random Forest) e filtros de volatilidade (GARCH) evidenciaram que a integração entre sinais pode ser economicamente plausível, desde que o regime de risco seja utilizado de maneira disciplinada. Em particular, estratégias com limiares mais restritivos mostraram-se eficazes na redução de drawdowns e volatilidade, ainda que à custa de menor participação nos movimentos de alta. Por outro lado, a flexibilização excessiva do filtro de regime deteriorou o desempenho econômico, expondo as estratégias a custos de timing e instabilidade justamente nos períodos em que o controle de risco deveria ser mais rigoroso.

Um aspecto importante deste trabalho é a escolha deliberada de não ancorar a análise em eventos extremos de volatilidade. Embora seja plausível — e potencialmente promissor — explorar limiares mais elevados, focados em episódios de estresse agudo, crises sistêmicas ou caudas extremas da distribuição, essa não foi a proposta central deste estudo. O foco adotado concentrou-se na volatilidade “mediana” da ação da VALE3 e em sua relação estrutural com o mercado de minério

de ferro, buscando entender como variações recorrentes e economicamente relevantes do risco afetam o desempenho das estratégias ao longo do tempo.

Essa escolha metodológica não representa uma limitação conceitual, mas sim uma delimitação consciente do escopo. Trabalhos futuros podem expandir essa abordagem, investigando regimes extremos de volatilidade, thresholds mais altos, combinações com medidas assimétricas de risco ou ainda modelos que incorporem explicitamente eventos raros, como choques macroeconômicos, rupturas geopolíticas ou mudanças abruptas no ciclo de commodities. A evidência empírica aqui apresentada sugere que tais extensões são plausíveis e merecem investigação, especialmente em contextos nos quais o objetivo seja proteção em cenários de cauda ou estratégias defensivas mais agressivas.

Por fim, os resultados reforçam uma mensagem central: não existe um modelo universalmente dominante. Estratégias quantitativas bem-sucedidas dependem menos da sofisticação isolada do algoritmo e mais da coerência entre hipótese econômica, desenho da regra de decisão e controle de risco. Nesse sentido, o trabalho contribui ao mostrar que a relação entre VALE3 e minério de ferro pode ser explorada de forma sistemática, mas apenas quando se respeitam os limites da previsibilidade, da instabilidade do mercado e dos custos implícitos de implementação. O avanço não está em “prever melhor o futuro”, mas em estruturar decisões mais consistentes diante da incerteza — objetivo que permanece central tanto para a pesquisa acadêmica quanto para a prática quantitativa em mercados financeiros.

6 Referências

ENDERS, Walter. *Applied econometric time series*. 3. ed. Hoboken: Wiley, 2010.

ANG, Andrew; BOUDOUKH, Jacob; RICHARDSON, Matthew. The cross-section of volatility and expected returns. *Journal of Finance*, v. 61, n. 1, p. 259–299, 2006.

BOLLERSLEV, Tim. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, v. 31, n. 3, p. 307–327, 1986.

BOX, George E. P.; JENKINS, Gwilym M.; REINSEL, Gregory C.; LJUNG, Greta M. *Time series analysis: forecasting and control*. 5. ed. Hoboken: Wiley, 2015.

BREIMAN, Leo. Random forests. *Machine Learning*, v. 45, n. 1, p. 5–32, 2001.

CAMPBELL, John Y.; LO, Andrew W.; MACKINLAY, A. Craig. *The econometrics of financial markets*. Princeton: Princeton University Press, 1997.

HOCHREITER, Sepp; SCHMIDHUBER, Jürgen. Long short-term memory. *Neural Computation*, Cambridge, v. 9, n. 8, p. 1735–1780, 1997.

CHEN, Jiaxuan. Exploration of quant trading talents' training mode of university education based on artificial intelligence. *Journal of Physics: Conference Series*, v. 1881, 022057, 2021. DOI: 10.1088/1742-6596/1881/2/022057.

CHRISTOFFERSEN, Peter F. *Elements of financial risk management*. 2. ed. San Diego: Academic Press, 2012.

ENGLE, Robert F. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, v. 50, n. 4, p. 987–1007, 1982.

ENGLE, Robert F.; RANGEL, José G. The spline-GARCH model for low-frequency volatility and its global macroeconomic causes. *Review of Financial Studies*, v. 21, n. 3, p. 1187–1222, 2008.

FAMA, Eugene F. Efficient capital markets: a review of theory and empirical work. *Journal of Finance*, v. 25, n. 2, p. 383–417, 1970.

FAMA, Eugene F.; FRENCH, Kenneth R. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, v. 33, n. 1, p. 3–56, 1993.

HAMILTON, James D. *Time series analysis*. Princeton: Princeton University Press, 1994.

HANSEN, Peter R.; LUNDE, Asger. A forecast comparison of volatility models: does anything beat a GARCH(1,1)? *Journal of Applied Econometrics*, v. 20, n. 7, p. 873–889, 2005.

HARVEY, Andrew C. *Forecasting, structural time series models and the Kalman filter*. Cambridge: Cambridge University Press, 1990.

HASTIE, Trevor; TIBSHIRANI, Robert; FRIEDMAN, Jerome. *The elements of statistical learning: data mining, inference, and prediction*. 2. ed. New York: Springer, 2009.

HULL, John C. *Options, futures, and other derivatives*. 10. ed. Boston: Pearson, 2018.

KAHNEMAN, Daniel; TVERSKY, Amos. Prospect theory: an analysis of decision under risk. *Econometrica*, v. 47, n. 2, p. 263–291, 1979.

KYLE, Albert S.; XIONG, Wei. Contagion as a wealth effect. *Journal of Finance*, v. 56, n. 4, p. 1401–1440, 2001.

LUCAS, Robert E. Asset prices in an exchange economy. *Econometrica*, v. 46, n. 6, p. 1429–1445, 1978.

MALKIEL, Burton G. *A random walk down Wall Street*. New York: W. W. Norton & Company, 2019.

MARKOWITZ, Harry. Portfolio selection. *Journal of Finance*, v. 7, n. 1, p. 77–91, 1952.

MERTON, Robert C. An intertemporal capital asset pricing model. *Econometrica*, v. 41, n. 5, p. 867–887, 1973.

NELSON, Daniel B. Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, v. 59, n. 2, p. 347–370, 1991.

PAGAN, Adrian; SCHWERT, G. William. Alternative models for conditional stock volatility. *Journal of Econometrics*, v. 45, n. 1–2, p. 267–290, 1990.

TAYLOR, Stephen J. *Modelling financial time series*. Chichester: Wiley, 1986.

TSAY, Ruey S. *Analysis of financial time series*. 3. ed. Hoboken: Wiley, 2010.

VARIAN, Hal R. *Intermediate microeconomics: a modern approach*. 9. ed. New York: W. W. Norton & Company, 2014.

CONT, Rama. Volatility clustering in financial markets: empirical facts and agent-based models. In: KIRMAN, Alan; TEYSSIERRE, Gilles (org.). *Long memory in economics*. Berlin: Springer, 2005.

SIMS, Christopher A. Macroeconomics and reality. *Econometrica*, v. 48, n. 1, p. 1–48, 1980.

YAHOO FINANCE. *VALE3, TIO=F, USDBRL historical data*. Disponível em: <https://finance.yahoo.com>. Acesso em: 10 jan. 2026.

VALE S.A. Formulário de Referência 2024. Rio de Janeiro: Vale S.A., 2024. Disponível em: <https://www.vale.com>. Acesso em: 3 mar. 2026.