



CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL

**MODELO MATEMÁTICO-COMPUTACIONAL
PARA PREVISÃO DE TENDÊNCIA DE PREÇOS:
ABORDAGEM COM MÚLTIPLOS ATIVOS
BUSCANDO OPORTUNIDADES EM DESVIOS
DA HIPÓTESE DE MERCADO EFICIENTE**

CHARLENE CÁSSIA DE RESENDE

Orientador: Arthur Rodrigo Bosco de Magalhães
CEFET-MG

BELO HORIZONTE
FEVEREIRO DE 2021

CHARLENE CÁSSIA DE RESENDE

**MODELO MATEMÁTICO-COMPUTACIONAL PARA
PREVISÃO DE TENDÊNCIA DE PREÇOS: ABORDAGEM
COM MÚLTIPLOS ATIVOS BUSCANDO
OPORTUNIDADES EM DESVIOS DA HIPÓTESE DE
MERCADO EFICIENTE**

Tese apresentada ao Programa de Pós-graduação em Modelagem Matemática e Computacional do Centro Federal de Educação Tecnológica de Minas Gerais, como requisito parcial para a obtenção do título de Doutor em Modelagem Matemática e Computacional.

Área de concentração: Modelagem Matemática e Computacional

Linha de pesquisa: Métodos Matemáticos Aplicados

Orientador: Arthur Rodrigo Bosco de Magalhães
CEFET-MG

CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL
BELO HORIZONTE
FEVEREIRO DE 2021

R433m Resende, Charlene Cássia de
Modelo matemático-computacional para previsão de tendência de preços :
abordagem com múltiplos ativos buscando oportunidades em desvios da
hipótese de mercado eficiente / Charlene Cássia de Resende. – 2021.
122 f.

Tese de doutorado apresentada ao Programa de Pós-Graduação em
Modelagem Matemática e Computacional.

Orientador: Arthur Rodrigo Bosco de Magalhães.

Tese (doutorado) – Centro Federal de Educação Tecnológica de Minas
Gerais.

1. Mercado de ações – Modelos matemáticos – Teses. 2. Equações
diferenciais – Teses. 4. Finanças – Teses. 5. Futuros financeiros – Teses.
6. Probabilidades – Teses. 7. Simulação (Computadores) – Teses. I. Magalhães,
Arthur Rodrigo Bosco de. II. Centro Federal de Educação Tecnológica de
Minas Gerais. III. Título.

CDD 530.13



SERVIÇO PÚBLICO FEDERAL
MINISTÉRIO DA EDUCAÇÃO
CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
COORDENAÇÃO DO PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL

**MODELO MATEMÁTICO-COMPUTACIONAL PARA PREVISÃO DE
TENDÊNCIA DE PREÇOS: ABORDAGEM COM MÚLTIPLOS ATIVOS
BUSCANDO OPORTUNIDADES EM DESVIOS DA HIPÓTESE DE
MERCADO EFICIENTE.**

Tese de Doutorado apresentada por **Charlene Cássia de Resende**, em 26 de fevereiro de 2021, ao Programa de Pós-Graduação em Modelagem Matemática e Computacional do CEFET-MG, e aprovada pela banca examinadora constituída pelos professores:

Prof. Dr. Arthur Rodrigo Bosco de Magalhães
Centro Federal de Educação Tecnológica de Minas Gerais

Prof. Dr. Cristiano Arbex Valle
Universidade Federal de Minas Gerais

Prof. Dr. Allbens Atman Picardi Faria
Centro Federal de Educação Tecnológica de Minas Gerais

Prof. Dr. Felipe Dias Paiva
Centro Federal de Educação Tecnológica de Minas Gerais

Prof. Dr. Adriano César Machado Pereira
Universidade Federal de Minas Gerais

Prof. Dr. José Luiz Acebal Fernandes
Centro Federal de Educação Tecnológica de Minas Gerais

Visto e permitida a impressão,

Prof. Dr. Thiago de Souza Rodrigues
Coordenador do Programa de Pós-Graduação Stricto Sensu em
Modelagem Matemática e Computacional

Agradecimentos

Agradeço a Deus por ter abençoado todos os dias da minha vida, por iluminar meu caminho e me dar forças para seguir sempre em frente.

Agradeço aos meus pais que muitas vezes se doaram e renunciaram aos seus sonhos para que eu pudesse realizar os meus. Quero dizer que essa conquista não é só minha, mas nossa. Tudo que consegui só foi possível graças ao amor, apoio e dedicação que vocês sempre tiveram por mim. Sempre me ensinaram agir com respeito, simplicidade, dignidade, honestidade e amor ao próximo. E graças ao trabalho e perseverança, os obstáculos foram ultrapassados e vitórias foram conquistadas.

Meu filho Miguel, a quem dedico em especial este trabalho, que embora pequeno em idade, foi grande em pensamento, sabendo apoiar, compreender e permitir a minha ausência em momentos importantes de sua vida. Mesmo diante das dificuldades, nós conseguimos!

A toda minha família pelo apoio, torcida e confiança que sempre depositaram em mim.

Ao meu orientador, Dr. Arthur Rodrigo Bosco de Magalhães, que me acompanhou ao longo de toda minha formação acadêmica. Tenho muito orgulho de citá-lo como um dos maiores responsáveis pela conclusão deste trabalho. Agradeço pela confiança, motivação, inspiração, amizade e conselhos. Um exemplo de dedicação, compreensão e competência. Preocupado não só com a realização do trabalho, mas principalmente com o ser humano. Enfim, cumpriu o seu papel de forma brilhante. Muitíssimo obrigada!!!

Meu agradecimento aos docentes do curso pelos conhecimentos e competências que me transmitiram ao longo deste percurso acadêmico. Em especial, agradeço aos professores Adriano, Rodrigo, José Geraldo, Albens, Acebal e Giovani pelas valiosas sugestões.

Aos amigos, colegas e funcionários que fizeram parte desta minha longa passagem pelo CEFET. Em especial, Dr. Nilmar, o grande amigo que ganhei nesta caminhada.

Ao Programa de Pós Graduação em Modelagem Matemática Computacional.

Ao CEFET-MG e à FAPEMIG, pelo apoio financeiro.

Os sonhos não determinam o lugar onde vocês vão chegar, mas produzem a força necessária para tirá-los do lugar em que vocês estão. Sonhem com as estrelas para que vocês possam pisar pelo menos na Lua. Sonhem com a Lua para que vocês possam pisar pelo menos nos altos montes. Sonhem com os altos montes para que vocês possam ter dignidade quando atravessarem os vales das perdas e das frustrações. Bons alunos aprendem a matemática numérica, alunos fascinantes vão além, aprendem a matemática da emoção, que não tem conta exata e que rompe a regra da lógica. Nessa matemática você só aprende a multiplicar quando aprende a dividir, só consegue ganhar quando aprende a perder, só consegue receber, quando aprende a se doar.

Sem sonhos, a vida não tem brilho. Sem metas, os sonhos não têm alicerces. Sem prioridades, os sonhos não se tornam reais. Sonhe, trace metas, estabeleça prioridades e corra riscos para executar seus sonhos. Melhor é errar por tentar do que errar por se omitir!

Augusto Cury

Resumo

A teoria econômica clássica defende que os mercados são eficientes, que os preços caminham em direção a um ponto de equilíbrio no longo prazo e que não há possibilidades de arbitragens. Para investigar esta Hipótese de Mercado Eficiente (HME), definimos um modelo de previsão de tendências de retornos de ativos financeiros, baseado em um sistema de equações diferenciais lineares acopladas discretizado. Neste modelo, consideramos que as relações entre os desvios dos preços das ações em relação a um preço considerado justo traz informação sobre a dinâmica dos preços das ações. Tais desvios são calculados utilizando-se a diferença logarítmica entre os preços reais e os preços justos, modelados por uma média móvel exponencial. Visando a comparação entre um mercado emergente e um desenvolvido, aplicamos o modelo para a bolsa de valores brasileira e norte-americana. A consistência do modelo foi investigada comparando-se resultados em quadrimestres diferentes. A relação entre a acurácia do modelo e o expoente de Hurst foi estudada. Parâmetros de execução do modelo, tais como, janela de ajuste, número de ações e frequência de amostragem dos dados, são explorados com o objetivo de encontrar uma combinação que apresente taxas de acurácia relevantes. Testes de hipóteses foram aplicados aos resultados da previsão para investigar se o modelo tem uma taxa de acerto de tendência de preços diferente de um processo totalmente aleatório. Diante dos parâmetros explorados, escolhemos um conjunto que apresentou resultados satisfatórios e aplicamos o modelo em simulação realística do mercado utilizando uma plataforma de *algotrading*. Na simulação, definimos dois modelos de operação para as estratégias de negociação: MO1 e MO2. No MO1, as operações são abertas de acordo com o sinal enviado pelo preditor e encerradas no fechamento do mesmo candle. Para MO2, as operações são encerradas por meio de *stops*. Os resultados foram comparados levando em consideração indicadores de desempenho, tais como: risco de exposição ao mercado, retorno financeiro, taxas operacionais e capital alocado.

Palavras-chave: Equações Diferenciais. Mercado de Ações. Séries Financeiras. Hipótese de Mercado Eficiente. Previsão de Tendências. Simulação Realística. Expoente de Hurst. Modelo de previsão.

Abstract

The classical economic theory argues that markets are efficient, that prices move towards a long-term equilibrium and that there is no possibility of arbitrage. To investigate this Efficient Market Hypothesis (EMH), we defined a model for forecasting trends in financial asset returns, based on a discretized coupled linear differential equation system. In this model, we consider that the relationship between stock price deviations from a price considered fair provides information on the dynamics of stocks prices. Such deviations are calculated using the logarithmic difference between actual and fair prices, modeled by an exponential moving average. To compare an emerging and a developed market, we applied the model to the Brazilian and North American stock exchanges. The consistency of the model was investigated by comparing results in different periods of four months. The relationship between model accuracy and Hurst exponent was studied. The execution parameters of the model, such as the fitting window, number of assets, and frequency of data sampling, are explored in order to find a combination that presents relevant accuracy rates. Hypothesis tests were applied to the forecast results to investigate whether the model has a success rate different than a completely random process. Given the parameters explored, we chose a set that presented satisfactory results and applied the model in a realistic simulation of the market using an algotrading platform. In the simulation, we defined two operating models for the trading strategies: MO1 and MO2. In MO1, the operations are opened according to the signal sent by the predictor and closed at the closing of the same candle. For MO2, operations are closed using stops. The results were compared taking into account performance indicators, such as: risk of exposure to the market, financial return, operating rates and allocated capital.

Keywords: Differential Equations. Stock market. Financial Series. Efficient Market Hypothesis. Trend Forecasting. Realistic Simulation. Hurst exponent. Forecasting Model.

Lista de Figuras

Figura 1 – Histórico de perfil dos investidores pessoas físicas na Bovespa . . .	4
Figura 2 – Organização do trabalho	12
Figura 3 – Candles do WINZ20	23
Figura 4 – Comparação de preços brutos, médias móveis e preços tratados .	29
Figura 5 – Distribuição das distâncias entre as matrizes A em tempos adjacentes	33
Figura 6 – Acurácia média para todas as n ações que compõem os índices Dow Jones e Ibovespa	39
Figura 7 – Acurácia média para as 10 ações com maiores participações no Ibovespa E Dow Jones	41
Figura 8 – Exemplo de um formato de um boxplot	43
Figura 9 – Boxplot da Acurácia para todas as ações que compõem os índices Dow Jones e Ibovespa	44
Figura 10 – Acurácia média para diferentes valores de n ações que compõem as carteiras do Ibovespa e Dow Jones	46
Figura 11 – Expoente de Hurst para todas ações que compõem os índices Dow Jones e Ibovespa	49
Figura 12 – Acurácia média e expoente de Hurst para todas as ações que compõem os índices Dow Jones e Ibovespa Q1	50
Figura 13 – Acurácia média e expoente de Hurst para todas as ações que compõem os índices Dow Jones e Ibovespa Q2	51
Figura 14 – Acurácias obtidas no primeiro e segundo quadrimestres de 2017 .	55
Figura 15 – Acurácias obtidas no primeiro e segundo quadrimestres de 2017 .	58
Figura 16 – Acurácias obtidas no primeiro quadrimestre de 2017 utilizando votação, média e sem votação	63
Figura 17 – Acurácias obtidas no segundo quadrimestre de 2017 utilizando votação, média e sem votação	64
Figura 18 – Fluxograma das estratégias	73
Figura 19 – Rendimentos esperados e os efetivamente obtidos para o ano de 2017	84
Figura 20 – Rendimentos esperados e os efetivamente obtidos para o ano de 2018	86
Figura 21 – Retornos e riscos do modelo de operação 2	91
Figura 22 – Taxas anuais de crescimento do Ibovespa	122

Lista de Tabelas

Tabela 1 – Perfil dos investidores pessoa física na Bovespa	3
Tabela 2 – Coeficientes angulares (β) e os correspondentes coeficientes de determinação (r^2)	40
Tabela 3 – Acurácia do modelo aplicado a diferentes granularidades de tempo para $R = 500$	47
Tabela 4 – Comparação das acurácias obtidas no primeiro e segundo quadrimestres para as ações dos índices Ibovespa e Dow Jones	54
Tabela 5 – Acurácias do modelo no primeiro e segundo quadrimestres para a combinação das 10 ações com maiores pesos	57
Tabela 6 – Comparação das acurácias para todas as ações dos índices Ibovespa e Dow Jones	60
Tabela 7 – Acurácias do modelo para as 10 ações dos índices Ibovespa e Dow Jones com os melhores α	61
Tabela 8 – Taxas de negociação e de liquidação da BOVESPA	71
Tabela 9 – Acurácias canônicas calculadas para as 10 ações mais líquidas do Ibovespa no ano de 2017	76
Tabela 10 – Acurácias canônicas calculadas para as 10 ações mais líquidas do Ibovespa no ano de 2018	76
Tabela 11 – Acurácias efetivas calculadas para as 10 ações mais líquidas do Ibovespa no ano de 2018	76
Tabela 12 – Acurácias efetivas calculadas para as 10 ações mais líquidas do Ibovespa no ano de 2018	77
Tabela 13 – Resultados da simulação utilizando MO1 sem taxas ano de 2017	78
Tabela 14 – Resultados da simulação utilizando MO1 sem taxas no ano de 2018	78
Tabela 15 – Resultados da simulação utilizando MO1 aplicando taxas operacionais para o ano de 2017	78
Tabela 16 – Resultados da simulação utilizando MO1 aplicando taxas operacionais para o ano de 2018	79
Tabela 17 – Rendimentos do MO1 no ano de 2017	83
Tabela 18 – Rendimentos do MO1 no ano de 2018	85
Tabela 19 – Resultados da simulação utilizando stops, sem taxas, para o ano de 2017	92
Tabela 20 – Resultados da simulação utilizando stop gain S_g (equação 49) e loss S_l (equação 48) sem taxas para o ano de 2018	93
Tabela 21 – Resultados da simulação utilizando stops, com taxas, para o ano de 2017	94

Tabela 22 – Resultados da simulação utilizando stops, sem taxas, para o ano de 2018	95
Tabela 23 – Média de preços e retornos das ações	97
Tabela 24 – Lotes padrão no mercado real e na plataforma de simulação no ano de 2017	115
Tabela 25 – Lotes padrão no mercado real e na plataforma de simulação no ano de 2017	116
Tabela 26 – Composição da carteira teórica do Ibovespa	119
Tabela 27 – Composição do índice Dow Jones	120

Lista de Abreviaturas e Siglas

HME	Hipótese de mercado eficiente
MO1	Modelo de operação 1
MO2	Modelo de operação 2
<i>DFA</i>	Detrended fluctuation analysis
<i>DMA</i>	Detrending moving average

Lista de Símbolos e Variáveis

a	Acurácia do modelo.
ρ	Coeficiente de correlação de Pearson
σ	Desvio padrão
ϵ	Erro associado
R	Quantidade de pontos utilizados no ajuste
\bar{a}	Acurácia média do modelo
σ_a	Desvio padrão da acurácia média
H	Expoente de Hurst
Q	Tamanho da janela para calculo de quartis
\bar{u}_4	Representa os dados do quarto quartil
n	Número de ações.
n_r	Número de retornos.
N	Número de execuções (previsões)
N_c	Número de execuções (previsões) corretas
d	Distância
β	Coeficiente angular da reta
r^2	Coeficiente de determinação
S	Tamanho da janela para calcular média móvel
x	Preço bruto
q_o	Aproximação para o preço de abertura do candle
q_c	Aproximação para o preço de fechamento do candle
\bar{x}	Preço justo
X	Preço tratado

T	Tendência
r_a	Retorno (rendimento) acumulado
r_a	Retorno acumulado para o conjunto de ações
λ	Quantidade de ações negociadas
Δx	Retorno bruto da ação
ΔX	Retorno tratado da ação
$\hat{\Delta}x$	Retorno previsto bruto da ação
$\hat{\Delta}X$	Retorno previsto tratado da ação
Sg	Stop gain
Sl	Stop loss

Sumário

1 – Introdução	1
1.1 Objetivos	6
1.2 Organização do trabalho	7
2 – Delineamento Metodológico	8
2.1 Levantamento bibliográfico	8
2.2 Definição do modelo	8
2.3 Dados utilizados	8
2.4 Tratamento dos dados	9
2.5 Exploração do modelo	10
2.6 Simulação	11
2.7 Avaliação dos resultados do modelo	11
3 – Trabalhos relacionados	13
3.1 Hipótese de Mercado Eficiente	13
3.2 Finanças comportamentais	15
3.3 Trabalhos aplicados em previsão	16
3.4 Expoente de Hurst	18
4 – Fundamentação teórica	20
4.1 Fundamentação teórica em finanças	20
4.1.1 Indicadores utilizados nas estratégias	24
5 – Modelos	27
5.1 Tratamento dos dados	27
5.2 Ajuste linear do dados	28
5.3 Modelo de previsão	32
6 – Exploração do modelo de previsão	37
6.1 Variações do tamanho de R	37
6.2 Variações da quantidade de ações	45
6.3 Variação da granularidade	47
6.4 Análise da acurácia em função do expoente de Hurst das séries	48
6.5 Acurácias em quadrimestres diferentes	52
6.5.1 Comparação das acurácias obtidas no primeiro e segundo quadri- mestres de 2017	52

6.5.2	Aplicação do modelo utilizando a combinação das 10 ações com maiores pesos nos índices Ibovespa e Dow Jones	56
6.5.3	Análise do desempenho do modelo aplicado às ações de melhor acurácia no quadrimestre precedente	59
6.6	Votação	62
6.7	Considerações gerais	65
7	Simulação	67
7.1	Visão geral dos modelos de operação	68
7.2	Modelo de operação 1	74
7.2.1	Acurácias	75
7.2.2	Simulação	77
7.2.3	Análise estatística de resultados financeiros relativos ao MO1	81
7.2.4	Considerações sobre MO1	87
7.3	Modelo de operação 2	87
7.4	Considerações finais	96
8	Conclusão	98
8.1	Trabalhos Futuros	102
	Referências	103
	 Apêndices	 109
	APÊNDICE A – Estimação da matriz de coeficientes pelo método dos mínimos quadrados	110
	APÊNDICE B – Tabelas lotes padrão das ações	115
	 Anexos	 117
	ANEXO A – Tabela de ações que participaram da carteira teórica do Ibovespa no primeiro quadrimestre de 2017	118
	ANEXO B – Tabela de ações que participaram da carteira teórica do Dow Jones no primeiro quadrimestre de 2017	120
	ANEXO C – Tabela das taxas anuais de crescimento do Ibovespa	121

Capítulo 1

Introdução

A teoria econômica clássica defende que não é possível obter informações de dados baseando-se em informações passadas que levem a ganhos fora do padrão do mercado (SAMUELSON, 1965). Seus teóricos defendem que os mercados seriam, assim, eficientes, e que os preços caminham em direção a um ponto de equilíbrio no longo prazo. Esta discussão inclui a definição da hipótese de mercado eficiente (HME) (MALKIEL; FAMA, 1970; SHILLER, 1980). Tal hipótese considera que não há possibilidades de arbitragem, que o mercado *descobre* o preço correto dos ativos, não sendo possível ganhos acima da média sem correr riscos acima da média. Considera, também, que os investidores são racionais e que toda informação relevante é disponível e absorvida imediatamente pelo mercado (MALKIEL; FAMA, 1970). No entanto, um grande número de analistas do mercado, especuladores e investidores, fundamentalistas ou técnicos, estão em busca de tais informações (FENG et al., 2012).

Investigações empíricas sobre a eficiência do mercado têm sido realizadas em muitos contextos, utilizando diversas ferramentas estatísticas (BARKOULAS; BAUM, 1996; GRECH; MAZUR, 2004; CAJUEIRO; TABAK, 2004; ZUNINO et al., 2007; CHORDIA; ROLL; SUBRAHMANYAM, 2008; BELAIRE-FRANCH; OPONG, 2010; CAGINALP; DESANTIS, 2011; CHARLES; DARNÉ; KIM, 2012; AL-KHAZALI; PYUN; KIM, 2012; LAZĂR; TODEA; FILIP, 2012; CHOUDHRY; JAYASEKERA, 2012; CHOUDHRY; JAYASEKERA, 2014; KATUSIIME; SHAMSUDDIN; AGBOLA, 2015; NTIM et al., 2015; CURME; STANLEY; VODENSKA, 2015; PODOBNIK et al., 2006; RESENDE et al., 2017). Pesquisas mostram que o mercado não é tão eficiente em termos de absorção de informação: nem toda informação relevante está disponível aos investidores (ROSS, 1989; HEALY; PALEPU, 2001; BRUNNERMEIER, 2005). No estudo realizado por (CAJUEIRO; TABAK, 2004), os expoentes de Hurst das séries de países desenvolvidos estão bem próximos de 0,5. Já os valores obtidos para países emergentes se distanciam mais deste valor, indicando que mercados emergentes são menos eficientes que mercados desenvolvidos. A ineficiência dos mercados emergentes, dentre outros fatores, pode estar relacionada à baixa liquidez do mercado (CHORDIA; ROLL;

SUBRAHMANYAM, 2008; CHUNG; HRAZDIL, 2010) e ao fluxo de informação (ROSS, 1989; HEALY; PALEPU, 2001; BRUNNERMEIER, 2005). Também há pesquisas no campo de finanças comportamentais onde o comportamento não racional é estudado em detalhes (KAHNEMAN; TVERSKY, 1979; SHILLER; FISCHER; FRIEDMAN, 1984). Em (CAGINALP; ERMENTROUT, 1990), os autores desenvolveram um modelo utilizando equações diferenciais, que incorpora o aspecto emocional dos investidores em um mercado financeiro teórico. Indicam que fatores psicológicos são capazes de influenciar a estabilidade dos preços, evidenciando desvios da hipótese de mercado eficiente.

Algoritmos de automatização de estratégias e operações também têm sido foco de várias pesquisas na área. Muitos desses trabalhos utilizam redes neurais artificiais para a construção dos modelos de previsão. Em (NAMETALA et al., 2016), os autores desenvolveram uma estratégia automatizada que combina redes neurais artificiais e preditores econométricos para gerar sinais de compra ou venda por meio de um modelo de negociação incorporado ao algoritmo.

Pesquisas nestas áreas de automatização, caracterização e modelagem de mercados constituem-se relativamente novas e promissoras, pois, com o surgimento de novas tecnologias e diversas plataformas de negociação, o número de pessoas atuando na bolsa tem aumentado significativamente. O cadastro de CPFs na Bolsa brasileira teve um grande avanço, sendo que em maio de 2020 este número chegou à marca de 2.483.286. O aumento pode ainda estar relacionado, dentre outros fatores, à ampla divulgação e também à redução da taxa de juros do país, tornando a renda variável uma atrativa fonte de investimento. As pessoas físicas representam 98,86% do total de investidores, sendo destes 75,05% homens e 23,81% mulheres. O valor investido pelas pessoas físicas é de 303,63 bilhões de reais. Ainda de acordo com o relatório disponibilizado pela Bolsa, a maior parte dos investimentos está concentrada no estado de São Paulo. Os dados com os números de investidores, bem como os valores por estado, são mostrados na tabela 1, que foi adaptada de (BM&FBOVESPA, 2020a).

A figura 1 apresenta a evolução da distribuição da participação de homens e mulheres no total de investidores pessoa física na bolsa brasileira. A curva azul corresponde aos CPFs relacionados ao sexo masculino, que em todo o período apresentado se manteve acima da curva vermelha, que representa a participação das mulheres na bolsa. Para ambos os perfis, percebemos um aumento significativo dos investidores a partir do ano de 2018. Segundo dados da Bovespa (BM&FBOVESPA, 2020a), o número de CPFs cadastrados na bolsa brasileira dobrou em 2019, saltando de 813.291 para 1.681.033. O critério de contagem considera o CPF cadastrado em cada agente de custódia, ou seja, pode contabilizar o mesmo investidor caso ele possua conta em mais de um corretora. Entretanto, este aumento no número de pessoas é bem relevante.

Tabela 1 – Perfil dos investidores pessoa física na Bovespa

Estado	Contas			Valor (R\$ bilhões)			%
	Homens	Mulheres	Total	Homens	Mulheres	Total	
SP	725.848	245.265	971.113	116,15	34,26	150,42	49,54%
RJ	204.690	70.602	275.292	35,36	12,12	47,48	15,64%
MG	182.502	54.644	237.146	25,38	4,24	29,62	9,75%
RS	108.117	30.292	138.409	12,46	3,60	16,06	5,29%
SC	91.315	25.268	116.583	7,89	2,06	9,95	3,28%
PR	118.032	33.968	152.000	10,71	2,84	13,55	4,46%
BA	59.552	18.373	77.925	5,11	1,10	6,22	2,05%
DF	65.687	23.372	89.059	6,20	1,19	7,39	2,43%
ES	39.356	11.092	50.448	2,95	0,58	3,52	1,16%
PE	42.806	12.570	55.376	2,80	0,75	3,55	1,17%
CE	36.110	10.321	46.431	2,24	0,69	2,92	0,96%
GO	44.555	13.562	58.117	2,22	0,46	2,68	0,88%
PB	16.386	4.761	21.147	0,75	0,17	0,93	0,31%
MT	20.172	5.841	26.013	1,13	0,20	1,33	0,44%
MS	17.871	4.995	22.866	1,28	0,16	1,44	0,47%
PA	19.571	6.147	25.718	1,01	0,22	1,23	0,41%
RN	17.016	4.454	21.470	0,90	0,17	1,07	0,35%
AM	12.893	4.137	17.030	0,75	0,14	0,88	0,29%
MA	13.847	4.135	17.982	0,73	0,12	0,85	0,28%
AL	10.212	2.863	13.075	0,47	0,08	0,55	0,18%
SE	9.966	2.801	12.767	0,50	0,10	0,60	0,20%
PI	8.512	2.456	10.968	0,39	0,07	0,46	0,15%
RO	7.805	2.339	10.144	0,34	0,08	0,42	0,14%
AP	1.921	616	2.537	0,08	0,01	0,09	0,03%
AC	2.557	802	3.359	0,11	0,03	0,14	0,04%
RR	2.204	712	2.916	0,08	0,01	0,09	0,03%
TO	5.693	1.702	7.395	0,17	0,03	0,20	0,06%
Total	1.885.196	598.090	2.483.286	238,17	65,46	303,63	

Nota: Os dados apresentados são referentes ao perfil dos investidores pessoa física na Bovespa em 29/05/2020. Os resultados foram segmentados por estados e mostram os números de contas de pessoas físicas na Bovespa, bem como o saldo financeiro. A contagem dos CPFs de investidores são por agente de custódia. Adaptado do site de serviços da Bovespa: Histórico de perfil dos investidores pessoas físicas de Maio de 2020, disponível em: http://www.b3.com.br/pt_br/market-data-e-indices/servicos-de-dados/market-data/consultas/mercado-a-vista/historico-pessoas-fisicas/.

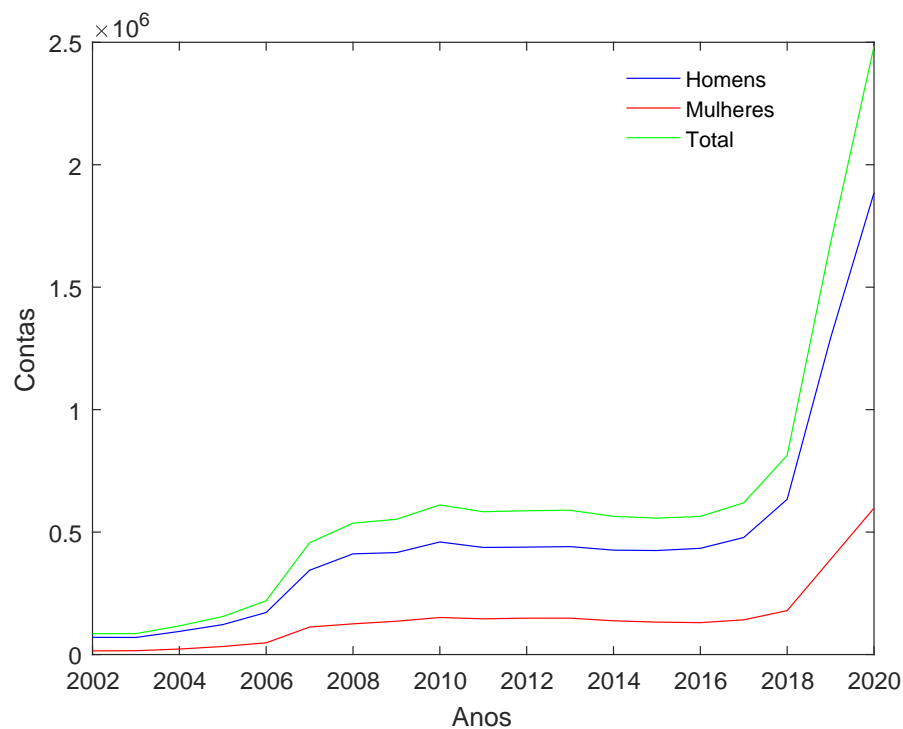


Figura 1 – **Histórico de perfil dos investidores pessoas físicas na Bovespa.** Representa a evolução do número de investidores na Bovespa entre os anos de 2002 a 2020. A curva na cor azul corresponde ao número de CPFs de investidores do sexo masculino, a curva na cor vermelha representa a evolução da participação de mulheres na Bolsa e a curva na cor verde mostra o total de pessoas físicas ao longo do período.

Nem todos os agentes atuantes no mercado têm retornos positivos e muitos acabam fracassando e desistindo do mercado, uma vez que a renda variável depende de diversos fatores, tais como gerenciamento de risco, definição de estratégia e até mesmo controle emocional. Desta forma, as estratégias automatizadas podem facilitar o ingresso neste cenário de grandes oscilações e incertezas com mais assertividade e precisão nas operações. Isto nos incentiva a pesquisar e desenvolver ferramentas que auxiliem na tomada de decisão.

Neste contexto, motivados pela lacuna deixada pela hipótese de mercado eficiente, sugerida por modelos interessados na previsão de dinâmicas de mercados (ATSALAKIS; K, 2013; ATSALAKIS; VALAVANIS, 2009), e por modelos empenhados na previsão do sinal de retornos de ativos (NYBERG, 2011; ANATOLYEV; GOSPODINOV, 2012; CHEVAPATRAKUL, 2013; PÖNKÄ, 2014), propomos também um modelo de previsão. Este é baseado em um sistema de equações diferenciais lineares acopladas. Discretizando o sistema, obtemos um modelo que se assemelha a um modelo de regressão linear multivariado.

Neste modelo, consideramos que os desvios dos preços brutos das ações em relação a um preço justo são os fatores importantes na nossa análise. Tais desvios serão denominados

preços tratados. De fato, sabemos que existem modelos para precificação de ativos, porém, neste trabalho, vamos utilizar médias móveis exponenciais como aproximação para o preço justo. Assim, o preço considerado justo será denominado preço de referência. O preço tratado será dado pela diferença logarítmica entre o preço bruto e o preço de referência.

Na investigação, utilizamos dados de todas as ações que compõem os índices Dow Jones e Ibovespa no ano de 2017. O Dow Jones é um índice do mercado americano que representa a negociação das 30 maiores empresas de capital aberto negociadas neste mercado. O Ibovespa é o índice composto, em média, por 58 ações negociadas na bolsa de valores de São Paulo. Para compor a carteira teórica do Ibovespa, as ações têm que atender a critérios estabelecidos pela BM&FBOVESPA¹, relativos à liquidez, volume de negociação, participação nos pregões de negociação e outros. As séries utilizadas são compostas por dados de *candles intraday*², onde maiores correlações são esperadas (ALLEZ; BOUCHAUD, 2011). O índice Ibovespa foi escolhido por ser de um mercado de país emergente, onde a ineficiência tende a ser mais evidenciada, possibilitando, assim, maiores oportunidades de sucesso na previsão de tendências de retornos (CAJUEIRO; TABAK, 2004; ZUNINO et al., 2007). O mercado Americano foi escolhido para contrastar com este cenário e também devido à sua importância.

Os parâmetros do modelo são explorados de forma a definir-se uma combinação ideal, que apresente resultados relevantes de acurácia na previsão de tendência de retornos das ações. Nosso objetivo é mostrar que o modelo de previsão de tendência apresenta uma taxa de acurácia superior a um processo totalmente aleatório. Para validar estes resultados realizaremos análises estatísticas. A hipótese de pesquisa a ser testada é que a acurácia do modelo é diferente de 0,5, evidenciando desvio da hipótese de mercado eficiente.

Diante dos parâmetros definidos durante a exploração do modelo, vamos desenvolver duas estratégias de negociação e aplicá-las no mercado real através de uma plataforma de simulação. Os resultados serão comparados levando em consideração as variáveis envolvidas na simulação, tais como: risco de exposição ao mercado, retorno financeiro, taxas operacionais e capital alocado.

Um problema desafiador que surge no âmbito das pesquisas em mercados financeiros é a possibilidade de se utilizar técnicas e modelos matemáticos para representar a dinâmica dos movimentos dos preços. O problema torna-se mais complexo quando envolve técnicas de previsão de tais valores, porém um grande número de agentes do mercado estão em busca destas informações. Assim, nosso problema neste trabalho é desenvolver um modelo matemático-computacional para a previsão de tendência de preços no mercado de ações brasileiro e no mercado americano.

¹Bolsa de Valores, Mercadorias e Futuros de São Paulo.

²Dados dentro de um período regular da bolsa (dia).

Nesta perspectiva, este trabalho Justifica-se como complementação de outros estudos que envolvem a combinação de ferramentas estatísticas, redes neurais artificiais e outras abordagens aplicadas para a modelagem de mercado de ações, podendo, assim, ser utilizado em pesquisas futuras.

1.1 Objetivos

Este trabalho tem como objetivo geral apresentar um modelo de previsão de tendências de preços das ações que compõem os índices Dow Jones e Ibovespa buscando oportunidades em desvios da hipótese de mercado eficiente. A pesquisa irá contribuir para a compreensão da dinâmica dos mercados de ações.

Na decomposição do objetivo geral, destacam-se alguns dentre os principais objetivos específicos:

- Desenvolver um modelo capaz de prever tendências de preços com taxas de acertos superiores a um processo totalmente aleatório;
- Testar o nível de eficiência de dois mercados com características diferentes. A investigação será por meio da análise de resultados obtidos para Dow Jones e Ibovespa;
- Explorar o modelo de previsão de tendência com o objetivo de encontrar parâmetros que retornem os melhores resultados;
- Desenvolver estratégias de negociação aplicadas ao mercado real em períodos diferentes utilizando tais parâmetros.

De forma específica, a principal contribuição deste trabalho é favorecer a compreensão da dinâmica do mercado de ações. Já sob o ponto de vista social, diante do aumento do número de agentes atuando na Bolsa de valores e da queda na taxa de juros, podemos dizer que o mercado de renda variável tem se tornado atrativo para novos investimentos, impactando, assim, a forma como as pessoas aplicam as suas economias. Por isso, é importante verificar a Hipótese de Mercado Eficiente neste cenário.

Neste contexto, o modelo de operação implementado, depois de adaptado e validado, também pode contribuir para operações em outros cenários, tais como mercados futuros(dólar e índices) e *commodities*³ (por exemplo: milho, feijão, café, soja).

³São produtos de qualidade e características uniformes que funcionam como matéria-prima, produzidos em escala e que podem ser estocados sem perda de qualidade. Seu preço é definido de acordo com a oferta e demanda americana.

1.2 Organização do trabalho

A tese está estruturada em 8 capítulos, dois apêndices e três anexos.

No capítulo 1, que é o presente capítulo, apresentamos a introdução do trabalho. Aqui são introduzidas as etapas e características do desenvolvimento da pesquisa.

No capítulo 2, descrevemos o delineamento da pesquisa, o planejamento do modelo, a exploração do modelo e a definição dos dados utilizados nas análises.

No capítulo 3, apresentamos uma breve revisão bibliográfica dos trabalhos referentes ao tema da pesquisa.

No capítulo 4, descrevemos alguns conceitos necessários para a compreensão do trabalho e as ferramentas computacionais utilizadas.

No capítulo 5, apresentamos a formulação do tratamento dos dados, a apresentação de um modelo baseado em um sistema de equações diferenciais lineares e a sua adaptação para um novo modelo de ajuste de pontos, bem como a definição de um modelo de previsão de tendência de preços.

No capítulo 6, é apresentada a exploração do modelo. Neste, analisamos os parâmetros de entrada com o objetivo de encontrar a combinação que apresente as melhores acurácias do modelo na previsão de tendência dos preços.

No capítulo 7, mostramos as estratégias de negociação aplicadas no mercado real. Para a simulação, utilizamos o modelo que retorna um sinal de compra, venda ou não opera apresentado no capítulo 5. Para calibrar o modelo de previsão, utilizamos os parâmetros explorados no Capítulo 6.

No capítulo 8, descrevemos a conclusão, bem como os passos futuros do trabalho.

O apêndice A, complementa a tese com a descrição da forma de estimar a matriz de coeficientes por meio da técnica de otimização matemática dos mínimos quadrados. Já o apêndice B contém a tabela de lotes padrão de negociação das ações na Bovespa.

Os anexos A, B e C mostram, respectivamente, a tabela das ações que participaram da carteira teórica do Ibovespa no primeiro quadrimestre de 2017, a tabela das ações que participaram da carteira teórica do índice Dow Jones também no primeiro quadrimestre de 2017 e a tabela das taxas anuais de crescimento do Ibovespa.

Capítulo 2

Delineamento Metodológico

Neste capítulo, vamos apresentar o delineamento da pesquisa, o planejamento, a exploração do modelo, simulação, bem como a definição dos dados utilizados nas análises.

2.1 Levantamento bibliográfico

A etapa de levantamento bibliográfico tem como objetivo contextualizar a pesquisa dentro do conjunto de atividades científicas relacionadas a modelos de previsão desenvolvidos para o mercado de ações.

2.2 Definição do modelo

A definição do modelo será apresentada no capítulo 5. O modelo de previsão foi definido baseando-se em um sistema de equações diferenciais lineares acopladas (RESENDE et al., 2017). Este sistema foi discretizado de forma a se aproximar de um modelo de regressão linear multivariado. De fato, não esperamos que um modelo linear seja capaz de representar a dinâmica do mercado financeiro. Porém, a violação ou não das premissas de adequação a um modelo de regressão não inviabiliza o trabalho, uma vez que o objetivo principal deste é prever tendências de preços de ações. A matriz de coeficientes do modelo, obtida até o tempo t , será determinada com base em manipulações algébricas, expostas por (AGUIRRE, 2007), e será utilizada como instrumento para a previsão da tendência dos preços no tempo $t + 1$. Trabalhamos aqui com a hipótese de linearidade local na vizinhança de t .

2.3 Dados utilizados

Para a definição e exploração do modelo, utilizamos preços de fechamento das ações que participaram da composição da carteira teórica dos índices Dow Jones e Ibovespa, com frequências de 5, 10, 15, 20 e 30 minutos no primeiro e segundo quadrimestres de 2017.

Na seção 6.3, vamos analisar todas estas granularidades de tempo e definir a amostra que apresentar melhores resultados. As séries de dados utilizadas foram obtidas na plataforma Bloomberg ¹.

Alterações na composição da carteira do Ibovespa, denominadas viradas, são realizadas a cada quadrimestre. As inclusões e exclusões de ações no índice são realizadas nos meses de Janeiro, Maio e Setembro, utilizando como base o último dia de pregão do mês anterior. A tabela 26, anexo A, mostra as ações que participaram da carteira teórica do Ibovespa no primeiro quadrimestre de 2017, 02/01/2017 a 28/04/2017. As ações estão ordenadas de acordo com a maior porcentagem de participação no Ibovespa. No quadrimestre 2, uma nova ação foi incluída na carteira, alterando, assim, as porcentagens de participação das ações. Porém, por simetria e comparação de resultados, decidimos manter as mesmas ações e as porcentagens baseados no quadrimestre 1.

O índice Dow Jones é um dos principais indicadores do mercado americano. A sua composição é definida pela média das 30 maiores e mais importantes empresas do mercado. As ações que compõem o índice são apresentadas na tabela 27, anexo B. Alterações na composição do índice Dow Jones são pouco frequentes. As alterações que aconteceram no período foram em relação à ordem de peso de participação na carteira. Por isso, decidimos manter a ordem de participação baseada no primeiro quadrimestre de 2017. A Tabela 27, anexo B apresenta as 30 ações que compõem o índice Dow Jones, ordenadas de acordo com a sua porcentagem de participação na carteira.

Para a simulação realística apresentada no capítulo 7, utilizamos dados com frequência de cotação de 5 minutos das 10 ações com maior liquidez na Bovespa nos anos de 2017 e 2018. A escolha do mercado brasileiro foi baseada no resultado da exploração do modelo (6), sendo este cenário o que apresentou melhores resultados.

2.4 Tratamento dos dados

Na seção 5.1, é apresentado de forma detalhada o tratamento dos dados. No modelo analisado, consideramos que os desvios dos preços em relação a um preço considerado justo traz informações sobre a dinâmica dos preços das ações. Tais desvios são calculados utilizando-se a diferença logarítmica entre o preço bruto e o preço de referência. Porém, calcular o preço referência de cada ação envolve diversas análises e variáveis que não estão no foco do trabalho. Por isto, utilizamos como alternativa para tal preço uma média móvel exponencial que será considerada o preço de referência. No cálculo da aproximação para o valor, definimos uma janela móvel de tamanho 12, composta pelos dados observados

¹A Bloomberg é uma empresa privada que oferece serviços para profissionais do mercado com foco em software, dados e mídia financeira privada. Foi fundada por Michael Bloomberg em 1981.

da amostra analisada. A escolha do valor de $S = 12$ foi baseada em aspectos experimentais de trabalhos anteriores (RESENDE et al., 2017), e também considerando que em uma hora de negociação temos 12 candles com frequência de cotação de 5 minutos.

2.5 Exploração do modelo

No capítulo 6, vamos explorar os diversos parâmetros do modelo. Porém, o objetivo principal não é encontrar parâmetros *ótimos*, mas sim selecionar uma combinação que apresente bons resultados de acurácia² na previsão de tendência. Para isto, exploramos as seguintes combinações de parâmetros do modelo:

- **Variação do tamanho da janela de pontos utilizada no ajuste**

Aplicamos quantidades diferentes de pontos, R , ao ajuste e definimos, baseados nos resultados, um valor considerado *satisfatório* como parâmetro para ajustes futuros.

- **Variação da quantidade de ações** Utilizando a quantidade de pontos definida no item anterior, variamos a quantidade n de ações envolvidas na análise. O valor inicial de n foi a combinação mínima de 2 ações, incrementando 1 ação na ordem do seu peso de participação nas carteiras, até o valor máximo de $n = 30$ para ações do Dow Jones e $n = 58$ para as ações que compõem o Ibovespa.

- **Variação da granularidade dos dados**

Analizamos as frequências de cotações de 5, 10, 15, 20 e 30 minutos para avaliar em qual granularidade o modelo obtém as melhores taxas de previsão de tendência.

Também exploramos o modelo tendo em vista as seguintes análises:

- **Taxas de acerto de tendências em função do expoente de Hurst**

Nesta análise, estamos interessados em estudar o comportamento de persistência e anti-persistência das séries de preços brutos das ações e investigar se há uma relação entre os resultados e o poder de previsibilidade do modelo.

- **Taxas de acerto de tendências em quadrimestres diferentes**

O intuito desta análise é investigar se as taxas de acertos obtidas no primeiro quadrimestre apresentam correlação com as taxas relativas ao segundo quadrimestre de 2017. Isto nos possibilita verificar a consistência do modelo em diferentes períodos.

- **Votação**

Uma votação entre tendências apontadas pelos ajustes utilizando diferentes quantidades de pontos R também deve ser elaborada. Nesta votação consideraremos como a tendência prevista a direção apontada pela maioria dos votantes e empate entre os votantes como sinal para não abrir novas posições no momento.

- **Validação estatística dos resultados**

Testes de hipóteses serão aplicados aos resultados de taxas de acertos obtidas pelo

²Capítulo 5 equação (30).

modelo. A hipótese a ser testada é que a acurácia do modelo é estatisticamente diferente de 0,5.

2.6 Simulação

Depois de explorados e definidos parâmetros satisfatórios para o modelo, no capítulo de simulação realística (capítulo 7), desenvolveremos duas estratégias de negociação utilizando tais parâmetros. Estas estratégias serão denominadas modelo de operação 1 (MO1) e modelo de operação 2 (MO2). As ações utilizadas nos modelos de operação foram as 10 ações com maior liquidez no ano de 2017 e 2018. As séries de dados utilizadas foram de candles de 5 minutos. Em ambos os modelos de operação, as posições são abertas quando o modelo de previsão de tendência envia um sinal não neutro. De acordo com este sinal, as operações de compra ou venda descoberta são abertas. No modelo de operação 1 (MO1), as operações são fechadas a cada passo de tempo (fechamento do candle). Já no modelo de operação 2 (MO2), as posições são fechadas de acordo com valores de *stop gain* e *stop loss* definidos.

2.7 Avaliação dos resultados do modelo

A análise dos resultados obtidos com os modelos de operação nos permite avaliar as características de cada estratégia, apontando as vantagens e desvantagens de cada uma delas. As variáveis levadas em consideração serão baseadas em risco de exposição ao mercado, retorno financeiro, taxas operacionais e capital alocado. A figura 2 mostra as etapas e organização do trabalho, descritas anteriormente.

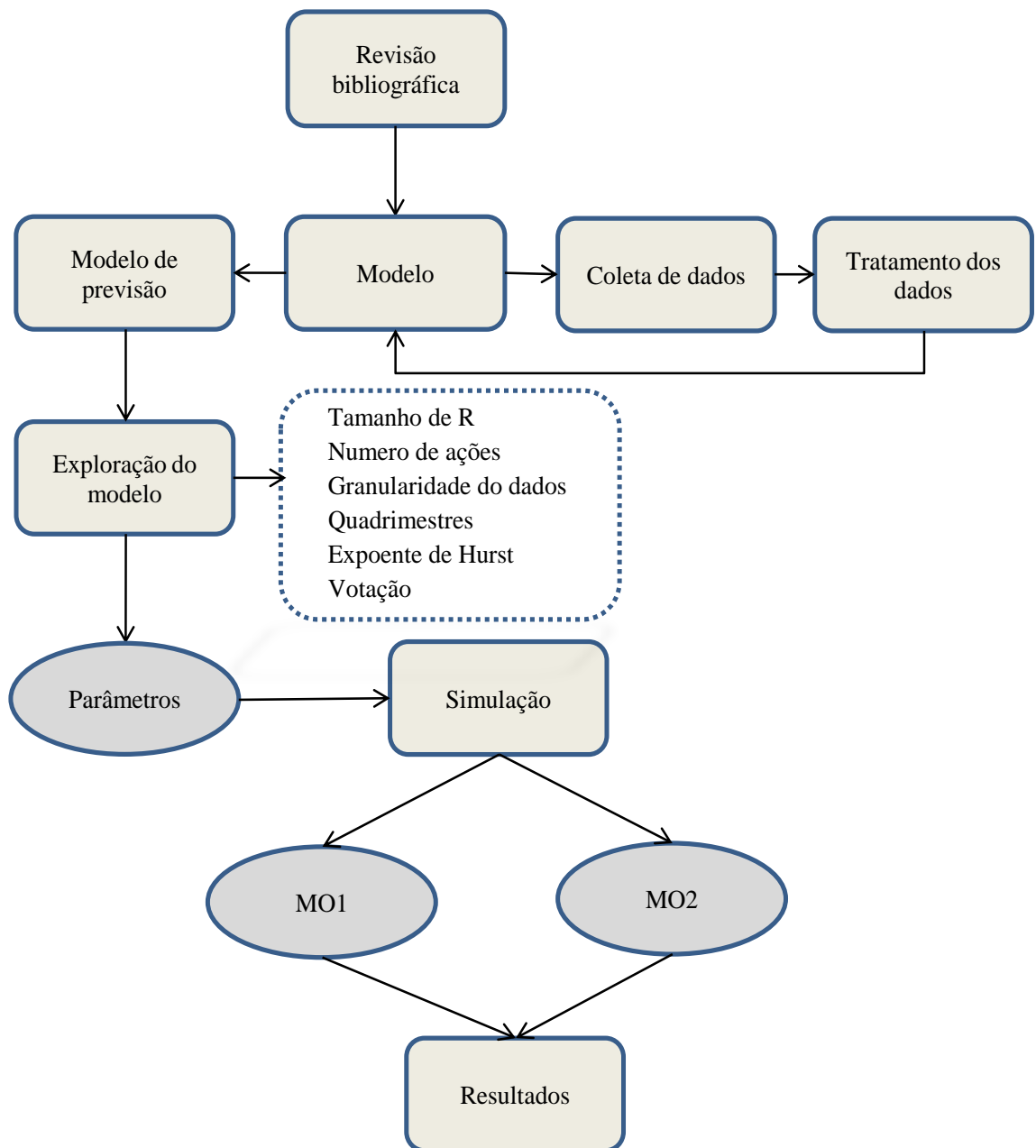


Figura 2 – **Organização geral do trabalho.** Na etapa de revisão bibliográfica, uma pesquisa é realizada para identificar o estado da arte do tema abordado, bem como as técnicas de previsão utilizadas no mercado de ações. Na parte do modelo, apresentamos a formulação do modelo baseado em um sistema de equações diferenciais lineares e a sua adaptação para um novo modelo de ajuste de pontos, a definição e o tratamento dos dados, bem como a definição de um modelo de previsão de tendência de preços. Na etapa de exploração, o modelo foi testado em diferentes cenários, investigando a influência do tamanho de R , número de ações, granularidade dos dados e outros parâmetros nos resultados do modelo. Na simulação, desenvolvemos duas estratégias de negociação (MO1 e MO2) baseadas no modelo de previsão e utilizando os parâmetros definidos. Na fase de resultados, apresentamos a análise dos resultados obtidos com os modelos de operação.

Capítulo 3

Trabalhos relacionados

Neste capítulo, vamos apresentar uma breve revisão bibliográfica dos trabalhos relativos ao tema da pesquisa. Na seção 3.1, vamos relacionar os principais trabalhos sobre Hipótese de Mercado Eficiente (HME). Na seção 3.2, apresentaremos trabalhos que exploram a racionalidade dos agentes do mercado por meio de estudos em finanças comportamentais. Nas seções 3.3 e 3.4, descreveremos alguns trabalhos que investigam desvios da HME por meio de modelos de previsão e por análises utilizando algumas ferramentas estatísticas.

3.1 Hipótese de Mercado Eficiente

Em um mercado eficiente, os preços dos ativos deveriam assimilar e refletir imediatamente toda informação relevante disponível. Esta é a denominada Hipótese de Mercado Eficiente (HME).

De acordo com a HME, o comportamento dos investidores é racional e não existe memória nas séries de preços que possibilite oportunidade de arbitragem. A Hipótese de Mercado Eficiente foi apresentada matematicamente por (SAMUELSON, 1965), mostrando que as flutuações dos preços acontecem de forma aleatória: a expectativa para o preço em $t + 1$ é o preço no tempo t . Seus teóricos defendem que os preços caminham para um ponto de equilíbrio e, com isso, não é possível obter ganhos demasiados sem correr riscos proporcionais. Defendem que toda informação relevante está disponível e que logo é absorvida e processada pelo mercado, sendo, desta forma, refletida nos valores dos ativos (MANTEGNA; STANLEY, 1999). Os mercados seriam, assim, capazes de reunir a sabedoria coletiva e precificar os ativos com precisão.

Três formas diferentes de HME são consideradas: fraca, semi-forte e forte. Um mercado é dito eficiente na forma fraca quando toda as informações públicas disponíveis são absorvidas imediatamente.

No nível semi-forte, considera-se que o mercado reflete toda a informação publicamente disponível, incluindo informações obtidas através de demonstrações contábeis publicadas e informações históricas. A diferença entre a eficiência semi-forte e a eficiência fraca é que a semi-forte exige dos agentes uma análise completa de dados fundamentalistas das empresas. Esse tipo de análise requer um maior embasamento teórico para avaliar com propriedade empresas ou setores da economia. Isso implica conhecimento, tempo e custo, o que pode limitar a forma de eficiência do mercado.

Um mercado é eficiente no sentido forte quando reflete, além das informações publicamente disponíveis e das históricas de preços, as informações privadas. Este nível engloba as hipóteses fraca e semi-forte e considera que os preços refletem imediatamente toda informação pública e privada inerente ao ativo.

No entanto, considerando todas as características, um mercado dito eficiente na forma mais estrita é um sistema idealizado. Em mercados reais, resíduos de ineficiência são observados. "Mercados reais são apenas uma aproximação do mercado eficiente"(MANTEGNA; STANLEY, 1999). A eficiência de mercado tem sido investigada em muitos contextos, utilizando uma variedade de ferramentas estatísticas (BARKOULAS; BAUM, 1996; GRECH; MAZUR, 2004; CAJUEIRO; TABAK, 2004; ZUNINO et al., 2007; CHORDIA; ROLL; SUBRAHMANYAM, 2008; BELAIRE-FRANCH; OPONG, 2010; CAGINALP; DESANTIS, 2011; CHARLES; DARNÉ; KIM, 2012; AL-KHAZALI; PYUN; KIM, 2012; LAZĂR; TODEA; FILIP, 2012; CHOUDHRY; JAYASEKERA, 2012; CHOUDHRY; JAYASEKERA, 2014; KATUSIIME; SHAMSUDDIN; AGBOLA, 2015; NTIM et al., 2015; CURME; STANLEY; VODENSKA, 2015; PODOBNIK et al., 2006). Analistas técnicos também acreditam que é possível obter informação sobre os preços baseados em dados históricos. Em finanças comportamentais, o comportamento irracional dos investidores também tem sido investigado (KAHNEMAN; TVERSKY, 1979; SHILLER; FISCHER; FRIEDMAN, 1984).

Desvios da Hipótese de Mercado Eficiente também têm sido associados a diversos fatores, tais como: custos de operação (LIU, 2010), baixa liquidez (CHORDIA; ROLL; SUBRAHMANYAM, 2008; CHUNG; HRAZDIL, 2010), limitações no fluxo de informação (ROSS, 1989; HEALY; PALEPU, 2001; BRUNNERMEIER, 2005), intervenções do banco central (RAFFERTY; TOMLJANOVICH, 2002; BORDO; LANDON-LANE, 2013; ALONSO-RIVERA; CRUZ-AKÉ; VENEGAS-MARTÍNEZ, 2014; CHEN; HUANG; SU, 2014), incertezas políticas (PASQUARIELLO; ZAFEIRIDOU, 2014) e crises financeiras (LIM; BROOKS; KIM, 2008).

No estudo realizado por (SHARMA; BANERJEE, 2015) nos variados setores da Bolsa de valores de Bombaim¹ (BSE), em um período de 8 anos (abril 2006 a março 2014), os resultados indicam que os dados não satisfazem à suposição de normalidade. Mostram que os retornos dos ativos analisados são guiados por curvas assimétricas e com caudas

¹Foi fundada em 1875 e é a principal bolsa de valores da Índia.

pesadas, indicando, também, que o mercado não segue a hipótese de *random walk*.

Nas próximas seções vamos apresentar de forma sucinta alguns trabalhos no contexto de investigação de desvios da HME por meio de estudos em finanças comportamentais, modelos de previsão e expoente de Hurst.

3.2 Finanças comportamentais

No trabalho (KAHNEMAN; TVERSKY, 1979), Kahneman e Tversky se uniram e seus estudos foram essenciais para quebrar os paradigmas que existiam em teoria de finanças. São considerados os precursores da chamada economia comportamental. Eles estudaram e exploraram os desvios e falhas de racionalidade na tomada de decisões por meio do comportamento humano no mercado. Em 1979 desenvolveram a Teoria da Perspectiva, também conhecida como Teoria do Prospecto (*Prospect Theory*). As teorias racionais consideram que agentes idealizados tomam decisões de forma racional, já a Teoria da Perspectiva surge como um questionamento a esta completa racionalidade dos agentes, o que denominaram de racionalidade limitada.

Na teoria apresentada, os autores observaram que os agentes tomam suas decisões baseados nos valores de perdas e ganhos e não no resultado final. Desta forma, tentam evitar perdas, sendo este o foco principal. De forma geral, são mais impactados pelas perdas do que pelos ganhos. Esta atitude está associada à aversão ao risco, prolongando posições em ativos perdedores para evitar assumir prejuízo. Diante disso, formularam a Teoria da Perspectiva, que é uma crítica à Teoria da Utilidade (KAHNEMAN; TVERSKY, 1979). Em resumo, a Teoria da Perspectiva tem como objetivo estudar o comportamento de tomada de decisão humana no mercado. Já a Teoria da Utilidade tem como objetivo mostrar qual a melhor decisão a ser tomada. A Teoria da Perspectiva concebe que os agentes têm mais aversão ao risco, sendo que tendem a evitar perder algo a ganhar, sendo assimétrica a relação ganho e perda. Já a Teoria da Utilidade considera que perdas e ganhos têm o mesmo efeito e magnitude.

A análise sugere que um agente que não aceita a perda e tem uma falha de adaptação está sujeito a aceitar apostas mais arriscadas que provavelmente não seriam aceitas em outras circunstâncias (KAHNEMAN; TVERSKY, 1979). Na análise também realizada por (TVERSKY; KAHNEMAN, 1991), a aversão às perdas (*loss aversion*) é investigada e sugere que o impacto das desvantagens é maior que vantagens auferidas, implicando aversão dos agentes a perdas e realçando a assimetria da função de utilidade de ganhos e perdas. Outra observação foi que, quando o preço está aumentando, o impacto na decisão de consumo é maior. A relutância para aceitar perdas foi observada nos agentes. Um estudo no mercado de ações mostrou que o número de operações aumenta quando o mercado

está em tendência de alta, sendo, assim, a propensão a compras é maior quando o preço do ativo está subindo (TVERSKY; KAHNEMAN, 1991).

Na mesma linha, os autores de (MERKLE; MÜLLER-DETHARD; WEBER, 2020) desenvolveram experimentos para analisar a exposição ao risco depois de realizadas e não realizadas perdas e também para investimentos com diferentes assimetrias. O estudo explorava os desvios e falhas de racionalidade na tomada de decisões. Os resultados de tais experimentos evidenciaram que os agentes se expõem a riscos menores depois de perdas realizadas e mais riscos depois de perdas ainda não realizadas, definido como efeito de realização.

No trabalho publicado por (KAHNEMAN et al., 2016), os autores exploram o conceito de ruídos como fonte de inconsistência e irracionalidade em resultados. Argumentam que algoritmos estão menos propensos a apresentarem tais ruídos e vieses nas tomadas de decisões.

No estudo (HOMMES; KOPÁNYI-PEUKER; SONNEMANS, 2020) desenvolveram um experimento com grupos de agentes do mercado e investigou-se a formação de bolhas nos preços. No experimento observaram que o contágio através de notícias afeta o comportamento de tais agentes. Por outro lado, no trabalho (FAN et al., 2009), os autores desenvolveram um modelo baseado em autômatos celulares para descrever a dinâmica do mercado de ações e investigar se o grau de imitação entre os agentes influencia a complexidade e estabilidade do mercado. Os resultados obtidos com o mercado simulado sugerem que há uma correlação entre tais fatores. Também observaram que quando o comportamento dos agentes é distribuído de forma aleatória o mercado tende a ser mais estável.

Em (CAGINALP; ERMENTROUT, 1990), os autores desenvolveram um modelo utilizando equações diferenciais que incorpora o aspecto emocional dos investidores em um mercado financeiro teórico. Resultados indicaram que fatores psicológicos são capazes de influenciar a estabilidade dos preços, evidenciando desvios da hipótese de mercado eficiente.

3.3 Trabalhos aplicados em previsão

Há na literatura uma abundância de artigos interessados na previsão da dinâmica do mercado. Alguns dos trabalhos estudados utilizam técnicas computacionais interessados na possibilidade de explorar possíveis desvios de eficiência do mercado e, com isso, obter estimativas razoáveis para preços futuros.

No trabalho (CARBONE; CASTELLI; STANLEY, 2004), os autores argumentam que a modelagem para previsão de retorno e volatilidade de preços é a principal tarefa da pesquisa financeira. A atividade neste sentido tem sido mantida pela suposição de que muitas séries financeiras temporais apresentam algum padrão de correlação e pela oportunidade

de arbitragem. Se correlação de longo prazo está presente nas séries de retornos de ativos financeiros, então a hipótese de que o preço segue um caminho completamente aleatório não é válida (CAJUEIRO; TABAK, 2004). Podemos destacar algumas contribuições relacionadas à área de previsão realizada por (RECHENTHIN; STREET, 2013) e (OLIVEIRA; NOBRE; ZARATE, 2013). Na primeira contribuição, os autores realizam um estudo utilizando probabilidade condicional para identificar continuidade e reversão de tendências de preços em dados de negociações em alta frequência. O estudo é realizado para os dados dos preços do fundo de ações do S&P 500 para as frequências de cotações de 1, 3, 5, 10, e 20 segundos, e 1, 5, e 30 minutos. Os autores concluem que o mercado apresenta ineficiência até aproximadamente 30 minutos. Para intervalos maiores o mercado analisado caminha em direção a se tornar mais eficiente. Já na contribuição (OLIVEIRA; NOBRE; ZARATE, 2013), os autores desenvolveram um modelo baseado em redes neurais artificiais. As entradas da rede são alguns indicadores técnicos e fundamentalistas, juntamente com a série de preços da ação preferencial da Petrobrás (PETR4), e a saída da rede é o preço previsto. As taxas de acerto do preço e de sua tendência, para as configurações das redes e os períodos apresentados, são satisfatórias.

Na referência (ATSALAKIS; VALAVANIS, 2009), os autores realizam uma revisão de artigos baseados em modelos que utilizam técnicas derivadas de redes neurais e *fuzzy* aplicadas na previsão de valores de mercados. Nesta contribuição, são apresentados diversos trabalhos que são classificados de acordo com o tipo de processamento de dados, as variáveis de entrada, tamanho da amostra, tipo e característica da técnica de implementação, bem como método de treinamento e validação. Os artigos aplicam os diferentes modelos na previsão de ações específicas, portfólios e índices de mercados desenvolvidos e emergentes. Os autores encontraram dificuldades na definição dos modelos em relação à estrutura das redes no que se refere à quantidade de camadas, neurônios e outras parâmetros.

Na referência (ATSALAKIS; K, 2013), o autor analisa 150 artigos científicos que utilizam técnicas de previsão de mercados de ações baseadas em modelos autoregressivos (AR), média móvel (MA), autoregressivo com média móvel (ARMA), autoregressivo integrado com média móvel (ARIMA), autoregressivo generalizado condicional (GARCH), dentre outros. Na revisão são observadas as técnicas de processamento e tratamento dos dados, bem como os resultados e análise estatística dos modelos.

No trabalho (GARCIA et al., 2020), os autores não argumentam que o mercado pode ser representado por uma dinâmica linear, mas baseiam-se na suposição de que possa existir componentes lineares na dinâmica das séries. Baseados nisso, propõem um modelo para capturar tal comportamento através de um oscilador harmônico amortecido, em que tentam descrever a dinâmica das oscilações de preços comumente observadas no mercado pela suposição de que os preços retornam a um preço justo, assimilando a restauração das

forças. Os parâmetros obtidos na janela analisada são utilizados como estimativas para fazer a previsão dos preços futuros adjacentes.

Na referência (MEI et al., 2017), os autores realizam um estudo para investigar a melhora na previsão de volatilidade do mercado de ações utilizando dados de alta frequência e incorporando assimetria e curtose como variáveis adicionais aos modelos de volatilidade comuns. Argumentam que o uso de dados de alta frequência permite melhor compreensão da dinâmica dos preços, facilitando a previsão de volatilidade e, com isso, melhorando a alocação de portfólio e o gerenciamento de risco.

No trabalho apresentado em (NELSON; PEREIRA; OLIVEIRA, 2017), os autores estudam o uso de redes neurais recorrentes, em particular a rede LSTM (*Long-Short term memory*), para a previsão de tendências de preços de ações utilizando dados históricos e também a combinação de indicadores da análise técnica. A investigação foi realizada no mercado brasileiro para dados com frequência de cotação de 15 minutos para as ações que compõem o índice Ibovespa no período avaliado. As acurácias obtidas são avaliadas de acordo com métricas estabelecidas e comparadas com outros modelos de previsão tradicionais que também utilizam redes neurais. De acordo com os autores, os resultados são promissores.

Algoritmos de automatização de estratégias e operações também têm sido foco de várias pesquisas na área. Muitos destes trabalhos utilizam redes neurais artificiais para a construção dos modelos de previsão. Em (NAMETALA et al., 2016), os autores desenvolveram uma estratégia automatizada que combina redes neurais artificiais e preditores econométricos para gerar sinais de compra ou venda por meio de um modelo de negociação incorporado ao algoritmo.

A previsão da direção dos preços de ativos em mercados de ações é uma tarefa complicada e tem sido um grande desafio, mas que tem chamado a atenção de pesquisadores e agentes de mercado, tornando-se um campo de pesquisas promissoras. Assim, as pesquisas não se limitam apenas a estes trabalhos citados aqui. Um amplo conjunto de técnicas e modelos de várias áreas tem sido empregado neste cenário com o objetivo de entender e modelar o movimento deste sistema dinâmico e complexo que é o mercado.

3.4 Expoente de Hurst

O expoente de Hurst foi resultado de estudos realizados pelo hidrólogo britânico Harold Edwin Hurst para o projeto de uma represa, em meados 1951. Neste, ele verificou que alguns fenômenos da natureza não se comportavam de maneira aleatória. Tais fenômenos observados possuíam uma dependência, dentre eles, as flutuações do rio Nilo e, a partir disso, ele criou o método para verificar a dependência de longo alcance.

Os modelos econômicos apresentados nos artigos (CARBONE; CASTELLI; STANLEY, 2004) e (CAJUEIRO; TABAK, 2004) testam a hipótese de mercado eficiente utilizando a metodologia do expoente de Hurst. O expoente de Hurst é calculado utilizando dados de séries temporais com o objetivo de observar o seu comportamento. O resultado assume valores entre 0 e 1. Valores próximos de 1 indicam persistência, valores próximos de zero indicam anti-persistência e valores próximos de 0,5 indicam que a série não possui tendência. Uma vez que o método DFA objetiva retirar tendências das séries, quando ele é usado, a interpretação sobre persistência e anti persistência vale para as séries sem tendência. No artigo (CARBONE; CASTELLI; STANLEY, 2004), os autores calcularam o expoente de Hurst para séries artificiais de fractais e em séries dos log retornos do DAX² e BOBL³ utilizando o método DFA (*detrended fluctuation analysis*) e DMA (*detrending moving average*). Para as séries de preços reais de ativos encontraram valores fora da faixa de variabilidade local, indicando que o mercado não segue o movimento browniano. Este resultado contraria a Hipótese de Mercado Eficiente, onde os expoentes devem ser próximos de 0,5. Já no artigo (CAJUEIRO; TABAK, 2004), com o objetivo de mostrar que mercados emergentes são menos eficientes que mercados desenvolvidos, os autores calculam o expoente de Hurst das séries dos índices de 11 mercados emergentes e de 2 países desenvolvidos. Como resultado, encontraram que a dinâmica do expoente de Hurst, para países desenvolvidos, se aproxima de passeio aleatório sem viés, flutuando em torno do valor 0,5. Nos países emergentes, os valores dos expoentes apresentam uma maior variação, indicando certos níveis de ineficiência, com predominância de persistência. Ainda de acordo com (CAJUEIRO; TABAK, 2004), a maior ineficiência em mercados emergentes pode estar relacionada a algumas características que os diferenciam dos mercados desenvolvidos, relacionadas à velocidade de propagação de informações, ao fluxo de capitais e à eficiência no processo de negociação. Na mesma linha, os autores do artigo (CAGINALP; DESANTIS, 2011) afirmam que, em um país em desenvolvimento, em um mercado de ações em início de implantação, pode ser esperada ineficiência, devida à dificuldade de capitalização dos ativos, ao menor número e à falta de experiência dos investidores. Quanto à ineficiência atribuída à quantidade de capital disponível, resultados experimentais mostram que um desbalanceamento entre a quantidade de capital e de ativos pode levar o mercado a preços incorretos (CAGINALP, 2002). Esse resultado sugere que o preço negociado no mercado é reflexo também do fluxo de recursos e não apenas do preço verdadeiro do ativo. Outros resultados experimentais que indicam a ineficiência de mercado podem ser encontrados em (CAGINALP; PORTER; SMITH, 2000).

²Índice que representa as 30 maiores empresas de capital aberto da Alemanha.

³Título do governo federal alemão.

Capítulo 4

Fundamentação teórica

Neste capítulo, vamos descrever de forma sucinta alguns conceitos de finanças necessários para a compreensão do trabalho. O mercado de capitais é um amplo sistema financeiro que envolve diversas variáveis que não serão contempladas neste trabalho. Vamos concentrar nossos estudos no mercado de ações, por isso, vamos estender as discussões para definições relacionados a este mercado.

4.1 Fundamentação teórica em finanças

O mercado financeiro é um extenso sistema que envolve diversos tipos de mercados, dentre eles, o mercado de renda variável, que é aquele em que não se pode dimensionar os resultados do investimento no momento da aplicação. As ações são títulos específicos deste mercado de renda variável. São frações do capital social de uma empresa de capital aberto (sociedade anônima). Cada empresa tem o seu código de negociação na bolsa que, no Brasil, geralmente, é dado por quatro letras seguidas de um número. Podem ser negociadas em lotes padrão ou fracionário. Elas podem ser do tipo ordinárias (ON) ou preferenciais (PN). As principais diferenças entre os dois tipos é que a preferencial tem preferência no recebimento de proventos e também preferência e reembolso de capital em caso de liquidação da empresa, já a ordinária concede direito a voto nas assembleias de acordo com o volume de participação em ativos. A remuneração aos acionistas é dada por meio de proventos como dividendos, juros sobre capital próprio, bonificações e direito de subscrição. Os dividendos correspondem ao valor de parte do lucro obtido pela empresa que é distribuído aos acionistas. Juros sobre capital próprio são valores pagos em dinheiro, porém deduzido do lucro tributável da empresa. Do ponto de vista do acionista, a diferença do recebimento entre os dois é que os dividendos não tem incidência de imposto de renda, já os juros sobre capital próprio tem uma alíquota de 15% sobre o valor recebido. A bonificação é quando a empresa distribui a cada acionista uma quantidade de ações proporcional às possuídas em custódia. A subscrição confere ao acionista o direito de adquirir uma

quantidade proporcional às suas a um preço preestabelecido pela empresa.

As negociações em ações no Brasil são realizadas na B3, comumente conhecida como Bovespa, que é a bolsa de Valores, Mercadorias e Futuros de São Paulo que atua no segmento BM&FBOVESPA. A B3 administra uma única plataforma de negociação multiativos – o PUMA Trading System. Através deste sistema são realizadas as operações com ativos de renda variável, renda fixa, derivativos e commodities (BM&FBOVESPA, 2020b). Os custos transacionais para operar ações na B3 são taxa de corretagem e emolumentos que são divididos em taxa de liquidação, negociação e registro. O imposto de renda é cobrado sobre o valor do ganho nas operações sendo 20% para *day trade* e 15% em operações comuns.

O Ibovespa é o principal índice de ações da B3. É composto por uma carteira teórica contendo, em média, sessenta principais ações listadas na bolsa, sendo um importante indicador de desempenho destas empresas. A cada quatro meses a composição é reavaliada e as ações são incluídas (excluídas) de acordo com os critérios descritos na sua metodologia de seleção, levando em consideração o número de negócios e o volume financeiro (BM&FBOVESPA, 2015). A composição do Ibovespa em janeiro de 2017 é apresentada na tabela 26, anexo A. Já o Dow Jones, também referenciado como *Dow Jones Industrial Average*, é um índice do mercado americano que representa a negociação das 30 maiores e mais importantes empresas de capital aberto negociadas. A sua composição em janeiro de 2017 é apresentada na tabela 27, anexo B. O Dow Jones é um dos principais indicadores do movimento do mercado americano e tem grande influência sobre outros mercados, tal como o mercado brasileiro.

As negociações no mercado de ações são realizadas por meio de ordens lançadas na corretora e que, posteriormente, são enviadas à bolsa. As ordens podem ser enviadas através do *home broker*, que é uma plataforma oferecida pelas corretoras para possibilitar a conexão e o envio de ordens à bolsa, ou através de outras plataformas de negociação contratadas pelos agentes. As ordens enviadas podem ser do tipo a mercado, ordem limite ou ordem stop. Na ordem a mercado, a operação é realizada no preço atual do ativo. Na ordem limite o preço é definido e a operação só ocorrerá se o valor for atingido ou, em casos específicos, em um preço melhor do que o definido. Já na ordem stop, é possível enviar ordens com valores de ganhos (perdas) pré definidos para uma operação. Nesta opção, as ordens só serão disparadas quando o preço for atingido. Os stops são ferramentas de segurança utilizadas para controlar a perda máxima aceitável (stop loss) e o ganho almejado (stop gain) definido para uma estratégia. O perfil do investidor pode ser definido como conservador, moderado ou agressivo (arrojado), de acordo com a sua tolerância a riscos.

As operações no mercado de ações brasileiro podem ser classificadas com base na

frequência em que realizam-se as transações. Day trade é uma estratégia de curta duração em que as posições em ativos são abertas e fechadas em um mesmo dia. *Swing trade* é modalidade em que as operações são fechadas em intervalos menores de dias, geralmente de dois a cinco dias. Já em operações de *position trade*, as transações são realizadas em frequências maiores, pensando em longo prazo.

No mercado de ações, o *retorno* é dado pela diferença percentual entre o preço no tempo t e $t - 1$. É usual considerar-se retornos logarítmicos que são dados pela diferença dos logaritmos dos preços no tempo t e $t - 1$. A volatilidade está relacionada à amplitude de variação destes retornos. Quanto mais os preços de um ativo variam em um intervalo de tempo, maiores as possibilidades de ganhos ou perdas em operações. Em alguns modelos, a volatilidade é utilizada como medida de risco da estratégia. Já a liquidez de um ativo, refere-se à capacidade deste converter-se em dinheiro sem perdas em relação a isso. Quanto maior a liquidez, menor a diferença (*spread*) do preço entre as ofertas de compra e venda de um ativo.

Série temporal é um conjunto de observações (dados) armazenadas ao longo de um intervalo de tempo. No mercado de ações, as séries podem ser compostas por valores de fechamento, abertura, máximo, mínimo, volume e outros dados de ativos. São importantes para a realização de *backtests*, análises empíricas e caracterização de movimento e oscilação do mercado. Também são utilizadas por analistas técnicos na composição de gráficos.

No mercado de ações existem duas metodologias de análises amplamente utilizadas e que são comumente chamada de *escolas* de análises técnicas e fundamentalistas. Análise técnica baseia-se na suposição de que preços passados podem ajudar a prever o movimento futuro e, para tanto, utilizam-se gráficos para definir a dinâmica dos preços dos ativos. Não leva em consideração informações fundamentalistas das empresas, pois acredita que o gráfico já está precificado com tais informações. O candle (*candlestick*) é a ferramenta da análise técnica mais utilizada no mercado de ações. Representa graficamente a variação do preço em um determinado intervalo de tempo. É formado por um corpo e sombras inferior e superior, sendo possível representar o preço de abertura, fechamento, máximo e mínimo na frequência analisada. Na figura 3 mostramos os candles diários para o contrato futuro de mini índice no mês de outubro de 2020. Os candles na cor vermelha indicam que a abertura foi maior que o fechamento, sendo um candle de baixa. Os candles na cor verde indicam que a abertura foi menor que o fechamento, sendo assim, de alta. As sombras inferiores (superiores), também denominadas pavios, representam os valores mínimos (máximos) na frequência analisada.

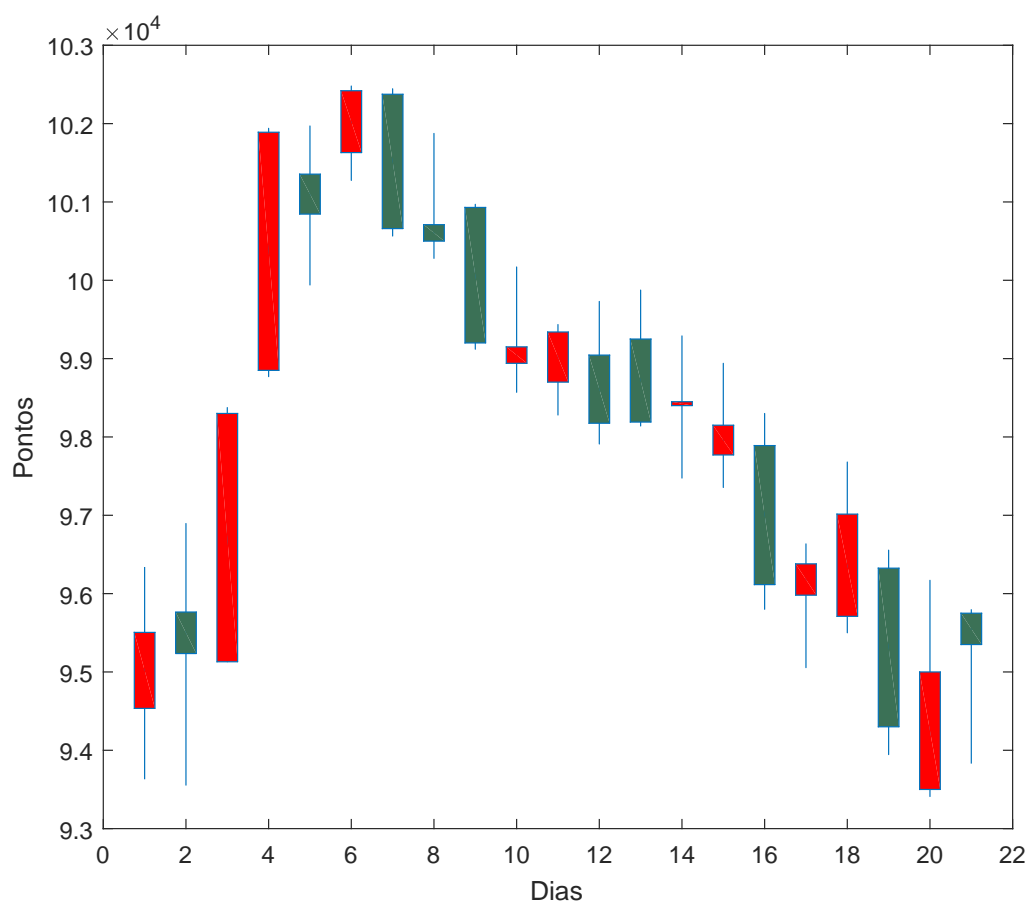


Figura 3 – **Candles do WINZ20**. Representa os dados da série de pontos diários do contrato futuro do mini índice no mês de outubro de 2020. Os candles na cor vermelha indicam que a abertura foi maior que o fechamento. Candles na cor verde indicam que a abertura foi menor que o fechamento. As sombras inferiores (superiores) indicam o valor mínimo (máximo) do intervalo.

A análise fundamentalista leva em consideração várias informações financeiras para projetar resultados futuros no médio e longo prazo. As projeções são baseadas principalmente em interpretações de dados e indicadores disponibilizados através de balanços da empresa.

Em ambas as análises, é possível desenvolver um modelo e realizar um backtest utilizando dados históricos relevantes. O backtest serve para testar a estratégia e verificar como ela se comportou em dados passados e, com base nestes resultados, tentar aplicar em dados futuros com a perspectiva de que os resultados serão aproximadamente replicados. Algumas plataformas que possibilitam o backtest apresentam certas limitações, tais como não considerar execução de ordem no preço pior do que o esperado (*slippage*), não simular a fila do livro de ofertas, não replicar momentos de grande volatilidade e liquidez do ativo, bem como não considerar a diferença vigente entre o menor preço de venda e o maior preço de compra (*spread*). Outro problema do backtest é a super otimização (*overfit*), que

consiste em trabalhar em um certo período e otimizar todos os parâmetros da estratégia para obter uma boa condição naquele intervalo, porém em períodos diferentes a estratégia não apresenta bons resultados.

Na nossa simulação, utilizamos alguns indicadores retornados pela plataforma de simulação, que geralmente são utilizados para avaliar o desempenho dos investimentos. Estes serão descritos na seção seguinte.

4.1.1 Indicadores utilizados nas estratégias

Nesta seção, vamos descrever os indicadores utilizados para medir o desempenho das estratégias que serão apresentadas no capítulo 7. Os cálculos dos indicadores econômicos financeiros são baseados principalmente nos valores de retorno e de capital inicial (C_i) e final (C_f). Os indicadores, aqui apresentados, são os que a plataforma de simulação retorna e, alguns casos, são ligeiramente diferentes dos usualmente utilizados no mercado.

- **Perdas e ganhos** - PNL

Mostra o resultado financeiro (lucro ou prejuízo) após a execução da estratégia. É a diferença entre o capital possuído ao final da execução da estratégia e o capital inicial investido,

$$PNL = C_f - C_i. \quad (1)$$

- **Retorno acumulado** - r_a

O retorno acumulado obtido com a estratégia ao final da execução é dado pelo razão entre o capital final possuído C_f e o capital inicial investido C_i ,

$$r_a = \frac{C_f}{C_i}. \quad (2)$$

- **Retorno anual esperado** - EAR

Retorno estimado que a estratégia alcançará em um período de um ano mantendo a mesma performance. É medido utilizando a inclinação da regressão linear da curva de retorno acumulado por dia ($slope$),

$$EAR = 252 \times slope. \quad (3)$$

- **Estabilidade** - R^2

Mostra a estabilidade da estratégia através do coeficiente de determinação R^2 da série de retornos acumulados. O valor do coeficiente varia entre 0 e 1. Quanto maior o valor de R^2 , maior a capacidade do modelo se ajustar à amostra.

- **Volatilidade** - *Volatilidade*

Mostra a volatilidade anualizada da estratégia, que é baseada no desvio padrão dos retornos diários σ_r extrapolado para o período de um ano. Neste cálculo, são considerados apenas os dias úteis (252),

$$Volatilidade = \sigma_r \sqrt{252}. \quad (4)$$

- **Sharpe** - *IS*

É um indicador utilizado para avaliar a relação entre o risco e o retorno de um investimento. Mostra o retorno r acima de uma taxa livre de risco r_f ajustado pela volatilidade (risco) σ . Quanto maior o Sharpe, maior a expectativa de retorno frente ao risco associado. É dado por

$$IS = \frac{r - r_f}{\sigma}. \quad (5)$$

- **Sortino** - *ISO*

É um indicador que, como o índice Sharpe, avalia a relação entre o risco e o retorno de um investimento. Porém, o Sortino leva em consideração apenas a volatilidade negativa σ_d , anulando os desvios dos retornos positivos,

$$ISO = \frac{r - r_f}{\sigma_d}. \quad (6)$$

- **Max Drawdown** - *MDD*

Retrata a maior perda percentual que a estratégia teve durante sua execução.

- **Beta** - β

É uma medida de volatilidade da estratégia em relação a um *benchmark*¹. Mede o quanto o desempenho da estratégia segue o da economia em geral, é dado por

$$\beta = \frac{cov(r, r_b)}{var(r_b)}, \quad (7)$$

onde r é o retorno da estratégia e r_b o retorno do benchmark.

- **Alpha** - α

Este parâmetro estima o quanto a estratégia se destacou positivamente dos retornos retratados pelo benchmark selecionado, ou seja, tenta mensurar quanto retorno a estratégia gerou acima do crescimento/decrescimento natural da economia.

$$\alpha = \bar{r}_d - \beta r_b, \quad (8)$$

onde \bar{r}_d é a média dos retornos diários da estratégia, r_b é o retorno do benchmark e β o coeficiente calculado.

¹Parâmetro utilizado como orientação para avaliar o desempenho de um investimento.

- **Value at Risk** - VaR

Este parâmetro estima o risco potencial de uma perda máxima de um ativo ou portfólio em um horizonte de tempo definido e com um determinado nível de confiança. Na plataforma, seu cálculo pelo método paramétrico é feito assumindo a distribuição normal nos retornos diários de toda a carteira e considerando um nível de confiança de 95%,

$$VaR = ppf(0,05, \bar{r}_d, \sigma r_d), \quad (9)$$

onde ppf é a função de densidade de probabilidade gaussiana, \bar{r}_d é a média dos retornos diários e σr_d é o desvio padrão dos retornos diários. O ponto de 5% corresponde aproximadamente a 1.645 desvios à esquerda da média dos retornos.

Capítulo 5

Modelos

Neste capítulo vamos mostrar: a formulação do tratamento dos dados que serão utilizados nas etapas seguintes do trabalho, a apresentação de um modelo baseado em um sistema de equações diferenciais lineares e a sua adaptação para um novo modelo de ajuste de pontos, bem como a definição de um modelo de previsão de tendência de preços.

5.1 Tratamento dos dados

O preço bruto de cada uma das n ações no tempo $t \in \mathbb{Z}$ ¹ é representado por $x_k(t)$ ($k = 1, 2, \dots, n$). Neste trabalho, vamos utilizar dados com *preços tratados*, que consistem nos desvios destes preços brutos $x_k(t)$ em relação a preços considerados justos. Contudo, determinar o preço justo de cada ação envolve diversas variáveis que não estão no foco do estudo. Uma forma que estabelecemos como aproximação para o cálculo foi utilizar uma média móvel exponencial $\bar{x}_k(t)$, calculada em um período precedente ao preço bruto analisado, definida como:

$$\bar{x}_k(t) = (x_k(t-1) - \bar{x}_k(t-1)) \times \frac{2}{(S+1)} + \bar{x}_k(t-1), \quad (10)$$

onde S é um parâmetro relacionado ao peso dado ao último valor $x_k(t)$ no cálculo e $\bar{x}_k(t-1)$ é a média móvel exponencial calculada no período anterior². Quanto maior o tamanho de S , menor o peso dado ao último preço. Neste trabalho, utilizamos $S = 12$

¹Embora t seja usualmente utilizado para definir tempo contínuo, no trabalho vamos utilizar para representar tempo discreto.

²A Eq. (10) tem característica recursiva, ou seja, o cálculo de $\bar{x}_k(t)$ depende de um valor de $\bar{x}_k(t_0)$ inicial. Desta forma, para o primeiro cálculo foi utilizada uma média simples, calculada em uma janela de pontos precedentes ao preço analisado, escolhida com tamanho S ,

$$\bar{x}_k(0) = \frac{1}{S} \sum_{t=t_0-S}^{t_0-1} x_k(t). \quad (11)$$

para o cálculo das médias móveis, que foram utilizadas como aproximação para os preços justos. Com isto, a nova série de preços tratados será dada pela diferença entre o logaritmo dos preços brutos e o logaritmo dos preços considerados justos, definida como

$$X_k(t) = \ln x_k(t) - \ln \bar{x}_k(t). \quad (12)$$

Consideramos que os desvios dos preços brutos em relação aos preços justos é que são importantes para a representação da dinâmica das séries. Desta forma, a partir deste ponto, em nossas análises utilizaremos as novas séries de *preços tratados*.

Para ilustrar o comportamento das séries, construímos a figura 4, que mostra o logaritmo da série de preços brutos x_1 e da série de médias móveis \bar{x}_1 (em (a)), bem como o logaritmo da série de preços tratados X_1 para a ação ITUB4 (em (b)). A escolha da série foi justificada por ser ITUB4 a ação que tem a maior porcentagem de participação na carteira teórica do Ibovespa do primeiro quadrimestre de 2017. Podemos observar que a curva das médias móveis, representada pela linha pontilhada, é mais suave que a curva dos valores brutos. Também, observamos que a curva de preços tratados X_1 tende a se aproximar da origem após as variações observadas.

5.2 Ajuste linear do dados

O modelo de equações diferenciais lineares, que propusemos em (RESENDE et al., 2017), é dado na forma

$$\frac{d}{dt} \mathbf{X}_a(t) = \mathbf{B} \mathbf{X}_a(t), \quad (13)$$

onde

$$\mathbf{X}_a(t) = \begin{bmatrix} X_1(t-n+1) & \cdots & X_1(t-1) & X_1(t) \\ X_2(t-n+1) & \cdots & X_2(t-1) & X_2(t) \\ \vdots & \ddots & \vdots & \vdots \\ X_n(t-n+1) & \cdots & X_n(t-1) & X_n(t) \end{bmatrix}$$

e

$$\mathbf{B} = \begin{bmatrix} B_{1,1} & B_{1,2} & \cdots & B_{1,n} \\ B_{2,1} & B_{2,2} & \cdots & B_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ B_{n,1} & B_{n,2} & \cdots & B_{n,n} \end{bmatrix}.$$

A equação (13) equivale a n equações do tipo

$$\frac{d}{dt} \begin{bmatrix} X_1(\tau) \\ X_2(\tau) \\ \vdots \\ X_n(\tau) \end{bmatrix} = \mathbf{B} \begin{bmatrix} X_1(\tau) \\ X_2(\tau) \\ \vdots \\ X_n(\tau) \end{bmatrix}, \quad (14)$$

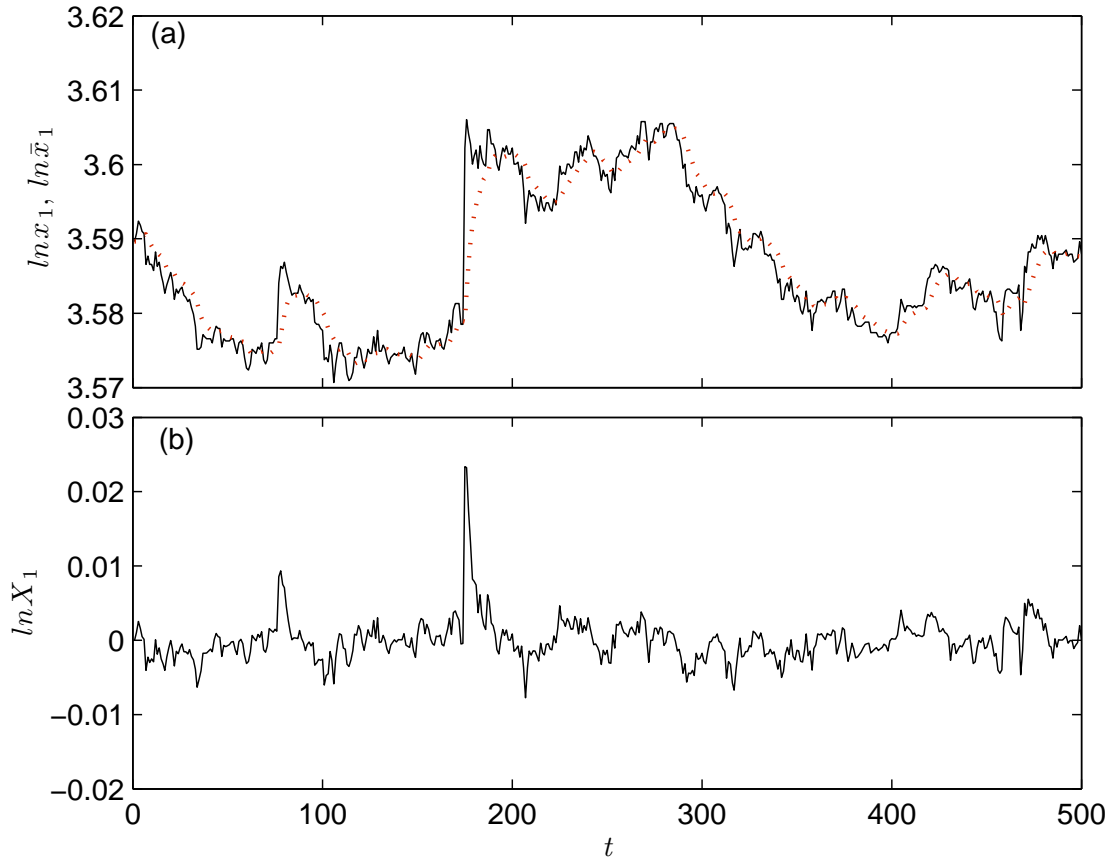


Figura 4 – **Comparação de preços brutos, médias móveis e preços tratados da série da ITUB4 ($k = 1$).** A figura mostra 500 pontos com granularidade de 5 minutos para a ação com maior participação na carteira teórica do Ibovespa no primeiro quadrimestre de 2017. (a) A linha contínua representa a evolução do logaritmo dos preços brutos $\ln x_1$ e a linha pontilhada descreve as médias móveis $\ln \bar{x}_1$. (b) Representa os preços tratados $\ln X_1$.

com $t - n + 1 \leq \tau \leq t$.

Neste modelo, consideramos que os desvios dos preços brutos em relação aos preços considerados justos são os elementos importantes para a caracterização das dinâmicas das séries.

Discretizando o modelo de equações diferenciais, representado pela equação (13), obtemos

$$\frac{\mathbf{X}_a(t) - \mathbf{X}_a(t-1)}{\Delta t} = \mathbf{B}\mathbf{X}_a(t), \quad (15)$$

equivalente a

$$\begin{aligned} \mathbf{X}_a(t-1) &= \mathbf{X}_a(t) - \Delta t \mathbf{B} \mathbf{X}_a(t), \\ &= (\mathbf{I} - \Delta t \mathbf{B}) \mathbf{X}_a(t), \end{aligned} \quad (16)$$

onde \mathbf{I} é a matriz identidade $n \times n$. Assumindo que $\mathbf{A}_a = (\mathbf{I} - \Delta t \mathbf{B})^{-1}$ existe, podemos escrever

$$\mathbf{X}_a(t) = \mathbf{A}_a \mathbf{X}_a(t-1). \quad (17)$$

No modelo apresentado acima, se existir a inversa de $\mathbf{X}_a(t-1)$, podemos encontrar as matrizes \mathbf{A}_a perfeitamente ajustadas para a equação (17), isto é, o sistema definido desta maneira apresenta soluções que passam em todos os pontos no intervalo $[t-n; t]$. Agora, adaptaremos o modelo buscando matrizes \mathbf{A}_a que são aproximadamente ajustadas em intervalos maiores. Para a definição, suponha que a evolução dos preços tratados ainda possa ser expressa pela combinação linear das n ações analisadas, sendo

$$\mathbf{X}(t) = \mathbf{A} \mathbf{X}(t-1) + \epsilon, \quad (18)$$

onde

$$\begin{aligned} \mathbf{X}(t) &= \begin{bmatrix} X_1(t-R+1) & \dots & X_1(t-1) & X_1(t) \\ X_2(t-R+1) & \dots & X_2(t-1) & X_2(t) \\ \vdots & \ddots & \vdots & \vdots \\ X_n(t-R+1) & \dots & X_n(t-1) & X_n(t) \end{bmatrix}, \\ \mathbf{X}(t-1) &= \begin{bmatrix} X_1(t-R) & \dots & X_1(t-2) & X_1(t-1) \\ X_2(t-R) & \dots & X_2(t-2) & X_2(t-1) \\ \vdots & \ddots & \vdots & \vdots \\ X_n(t-R) & \dots & X_n(t-2) & X_n(t-1) \end{bmatrix}, \\ \mathbf{A} &= \begin{bmatrix} A(1,1) & \dots & A(1,n-1) & A(1,n) \\ A(2,1) & \dots & A(2,n-1) & A(2,n) \\ \vdots & \ddots & \vdots & \vdots \\ A(n,1) & \dots & A(n,n-1) & A(n,n) \end{bmatrix}, \\ \epsilon &= \begin{bmatrix} \varepsilon_1(t-R+1) & \dots & \varepsilon_1(t-1) & \varepsilon_1(t) \\ \varepsilon_2(t-R+1) & \dots & \varepsilon_2(t-1) & \varepsilon_2(t) \\ \vdots & \ddots & \vdots & \vdots \\ \varepsilon_n(t-R+1) & \dots & \varepsilon_n(t-1) & \varepsilon_n(t) \end{bmatrix}, \end{aligned}$$

ϵ é o erro associado ao ajuste, R é a quantidade de pontos utilizados no ajuste e n a quantidade de ações analisadas. Os diferentes valores que R pode assumir serão testados e apresentados mais detalhadamente na seção 6.1.

Observe que a equação (18) é diferente da equação (17), uma vez que existe um erro associado ao ajuste aproximado e também pelo fato de que $\mathbf{X}_a(t)(\mathbf{X}_a(t-1))$ é $n \times n$ e $\mathbf{X}(t)(\mathbf{X}(t-1))$ é da ordem $n \times R$ e, conseqüentemente, a matriz de coeficientes \mathbf{A} é diferente de \mathbf{B} .

Expandindo a equação (18), para $t - R + 1 \leq \tau \leq t$, o sistema fica na forma

$$\begin{cases} X_1(\tau) = A(1,1)X_1(\tau-1) + A(1,2)X_2(\tau-1) + \dots A(1,n)X_n(\tau-1) + \varepsilon_1(\tau) \\ X_2(\tau) = A(2,1)X_1(\tau-1) + A(2,2)X_2(\tau-1) + \dots A(2,n)X_n(\tau-1) + \varepsilon_2(\tau) \\ \vdots \\ X_n(\tau) = A(n,1)X_1(\tau-1) + A(n,2)X_2(\tau-1) + \dots A(n,n)X_n(\tau-1) + \varepsilon_n(\tau) \end{cases} \quad (19)$$

A equação (18) se assemelha a um *modelo de regressão linear multivariado*. Entretanto, a sua validação como tal modelo depende da capacidade do mesmo em representar a dinâmica que produziu os dados. A possível violação das premissas de um modelo de regressão não invalida o nosso modelo. Como veremos adiante, estamos interessados nos resultados de acerto de tendência, que serão analisados com rigor estatístico. De fato, não esperamos que um modelo linear seja capaz de representar a dinâmica do mercado financeiro. O que é refletido na distribuição dos erros que não atendem as premissas de independência e normalidade. Contudo, as matrizes apresentaram melhores resultados, em termos computacionais e de qualidade das soluções, que os obtidos com o método do gradiente, explorado em trabalhos anteriores (RESENDE et al., 2017).

Na equação (18), desconsiderando o erro, a partir dos $\mathbf{X}(t)$ e $\mathbf{X}(t-1)$, podemos definir a matriz de coeficientes \mathbf{A} de forma análoga à descrita por (AGUIRRE, 2007). Como a matriz $\mathbf{X}(t-1)$ não é quadrada, ela não pode ser invertida. Entretanto, podemos multiplicar a equação (18) em ambos os lados por $\mathbf{X}(t-1)^T$ e obter \mathbf{A} fazendo

$$\mathbf{X}(t)\mathbf{X}(t-1)^T = \mathbf{A}\mathbf{X}(t-1)\mathbf{X}(t-1)^T$$

$$\mathbf{A}(\mathbf{X}(t-1)\mathbf{X}(t-1)^T)(\mathbf{X}(t-1)\mathbf{X}(t-1)^T)^{-1} = \mathbf{X}(t)\mathbf{X}(t-1)^T(\mathbf{X}(t-1)\mathbf{X}(t-1)^T)^{-1}$$

$$\mathbf{A} = (\mathbf{X}(t)\mathbf{X}(t-1)^T)(\mathbf{X}(t-1)\mathbf{X}(t-1)^T)^{-1}. \quad (20)$$

Para as manipulações algébricas, consideramos a existência da inversa de $(\mathbf{X}(t-1)\mathbf{X}(t-1)^T)$. Outra forma para estimar a matriz \mathbf{A} é através do método dos mínimos quadrados, conforme descrito por (AGUIRRE, 2007). A técnica da otimização matemática é apresentada no apêndice A para complementar o trabalho.

A matriz \mathbf{A} pode variar de acordo com a escolha de t , conforme a equação (20). Deste modo, usaremos a notação $\mathbf{A}(t)$ para denotar a matriz \mathbf{A} , calculada utilizando pontos do tempo $t - R$ até o tempo t .

Com a matriz $\mathbf{X}(t)$ e $\mathbf{X}(t - 1)$, podemos calcular a variação nos preços das ações, que é a diferença do preço da ação no período t em relação ao período anterior $t - 1$. Esta variação do preço $\Delta\mathbf{X}(t)$ é denominada *retorno* do ativo, definida por

$$\Delta\mathbf{X}(t) = \begin{bmatrix} \Delta X_1(t) \\ \Delta X_2(t) \\ \vdots \\ \Delta X_n(t) \end{bmatrix}, \quad (21)$$

sendo

$$\begin{aligned} \Delta X_k(t) &= X_k(t) - X_k(t - 1), \\ &= \log \left(\frac{x_k(t)}{\bar{x}_k(t)} \right) - \log \left(\frac{x_k(t - 1)}{\bar{x}_k(t - 1)} \right), \end{aligned} \quad (22)$$

onde usamos no passo final a equação (12).

5.3 Modelo de previsão

Para tentar prever o preço no tempo $t + 1$, vamos utilizar como parâmetro a matriz $\mathbf{A}(t)$. Para tanto, é desejável que $\mathbf{A}(t)$ varie pouco no tempo, sendo $\mathbf{A}(t)$ próxima a $\mathbf{A}(t - 1)$. Trabalhamos com a hipótese de linearidade local na vizinhança de t . Para analisar esta variação da matriz $\mathbf{A}(t)$ no tempo, definimos a distância

$$d(\mathbf{A}(t), \mathbf{A}(t - 1)) = \sqrt{\sum_{i=1}^n \sum_{j=1}^n |A(t)_{i,j} - A(t - 1)_{i,j}|^2}, \quad (23)$$

onde $A(t)_{i,j}$ ($A(t - 1)_{i,j}$) é o elemento da matriz $\mathbf{A}(t)$ ($\mathbf{A}(t - 1)$) na linha i e coluna j , entre as matrizes adjacentes $\mathbf{A}(t)$ e $\mathbf{A}(t - 1)$ para vários valores de R e mostramos os resultados na figura 5.

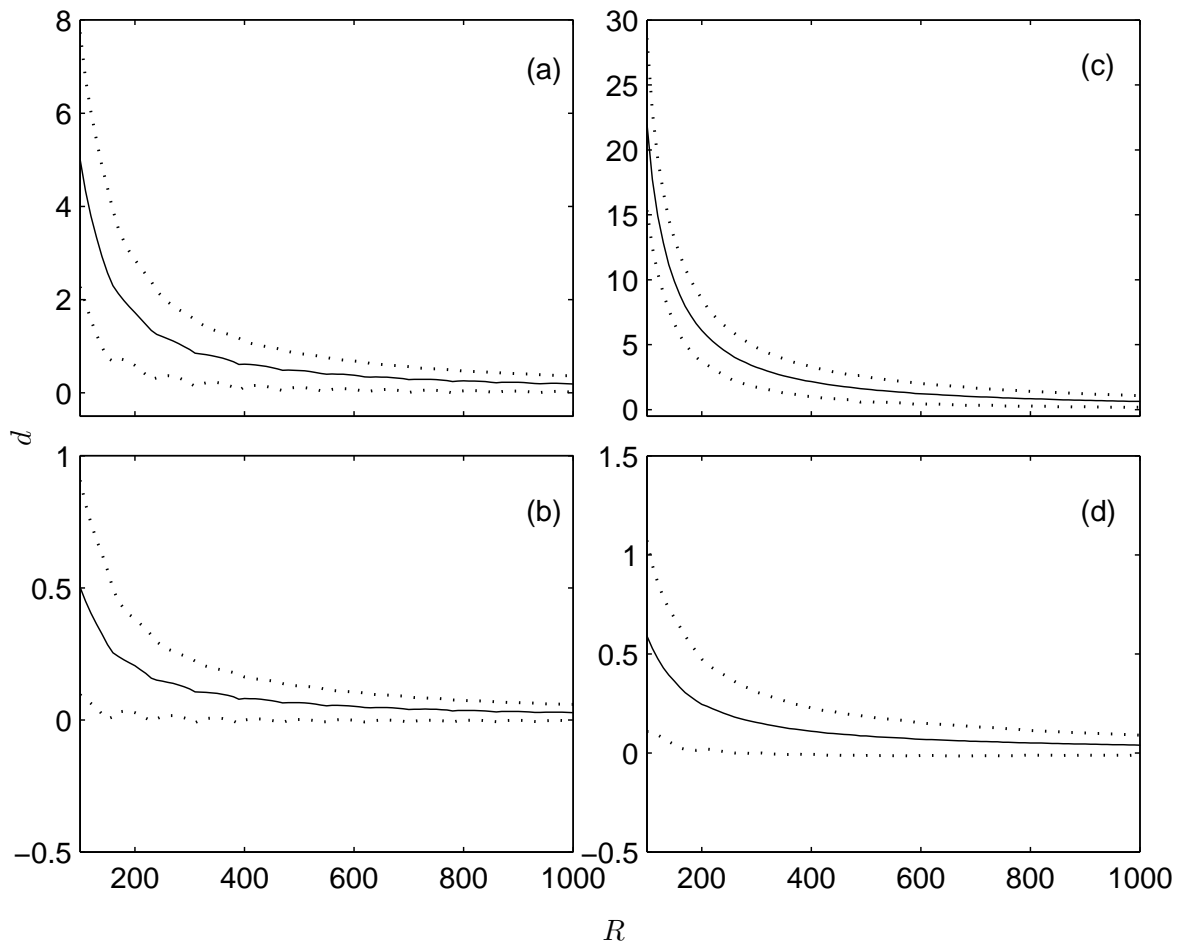


Figura 5 – **Distribuição das distâncias entre as matrizes A em tempos adjacentes para valores de R variando de 100 a 1.000.** Foram realizadas 5, 000 execuções para as ações do Ibovespa e 3, 900 para as ações que compõem o Dow Jones. Utilizamos séries de preços com granularidade de 5 minutos, no período de 02/01/17 a 28/04/17 (primeiro quadrimestre de 2017). A linha contínua representa a média das distâncias e a linha pontilhada mostra a média de $d \pm$ os desvios. (a) Calculamos a distância entre as matrizes adjacentes obtidas para todas as ações do índice Dow Jones. (b) Distância das matrizes adjacentes obtidas utilizando as 10 ações com maiores participações no índice Dow Jones. (c) Distância entre as matrizes adjacentes calculadas utilizando todas as ações que compõem a carteira teórica do Ibovespa. (d) Distância entre as matrizes obtidas utilizando as 10 ações com maiores pesos de participação na composição do Ibovespa.

Analisando os resultados da figura 5, obtidos com a Eq. (20), percebemos que, à medida em que aumentamos o valor de R , as distâncias entre as matrizes calculadas em tempos adjacentes diminuirão. Análises da quantidade de ações, tamanho de R e outras, serão exploradas no Cap. 6.

Com a perspectiva de que a matriz $\mathbf{A}(t)$, calculada utilizando-se $R + 1$ preços passados, varie pouco, e, com isso, seja uma boa estimativa para $\mathbf{A}(t + 1)$, a empregamos como instrumento para o modelo de previsão. Este modelo será baseado na equação (18), definido por

$$\hat{\mathbf{X}}(t + 1) = \mathbf{A}(t)\mathbf{X}_c(t), \quad (24)$$

onde $\mathbf{A}(t)$ foi obtida com os $R + 1$ preços até o tempo t ,

$$\hat{\mathbf{X}}(t + 1) = \begin{bmatrix} \hat{X}_1(t + 1) \\ \hat{X}_2(t + 1) \\ \vdots \\ \hat{X}_n(t + 1) \end{bmatrix}$$

é a matriz com preços previstos das ações no tempo $t + 1$ e

$$\mathbf{X}_c = \begin{bmatrix} X_1(t) \\ X_2(t) \\ \vdots \\ X_n(t) \end{bmatrix}$$

é uma matriz coluna com os valores dos preços tratados no tempo t .

Mesmo que haja pouca variação entre as matrizes $\mathbf{A}(t)$ e $\mathbf{A}(t + 1)$, é natural obtermos um erro associado à previsão.

O retorno previsto das ações é dado por

$$\hat{\Delta}\mathbf{X}(t + 1) = \begin{bmatrix} \hat{\Delta}_1 X(t + 1) \\ \hat{\Delta}_2 X(t + 1) \\ \vdots \\ \hat{\Delta}_n X(t + 1) \end{bmatrix}, \quad (25)$$

onde

$$\hat{\Delta}X_k(t + 1) = \hat{X}_k(t + 1) - X_k(t). \quad (26)$$

Os retornos reais $\Delta\mathbf{X}(t + 1)$ e os previstos $\hat{\Delta}\mathbf{X}(t)$ dos ativos são referentes a retornos de preços tratados. A partir destes valores, podemos definir tendências, porém, para analisar a direção destes preços, é preciso analisar dados de preços brutos $\mathbf{x}(t)$ e não mais os preços tratados $\mathbf{X}(t)$. Para o cálculo com estes dados brutos é preciso fazer a operação para reverter os dados tratados, que foram utilizados até este ponto do

trabalho. Para o tratamento inicial dos dados, utilizamos a equação (12), reproduzida abaixo:

$$X_k(t) = \ln x_k(t) - \ln \bar{x}_k(t), \quad (27)$$

onde $X_k(t)$ é o preço tratado de fechamento da ação k no tempo t , $x_k(t)$ é o preço bruto de fechamento da ação k no tempo t e \bar{x}_k é a média móvel calculada utilizando dados precedentes a $x_k(t)$.

Para transformar os preços tratados em preços brutos partimos da equação (12) e obtemos

$$x_k(t) = \bar{x}_k(t) * e^{X_k(t)}. \quad (28)$$

Evidentemente, os preços brutos reais são conhecidos, e este procedimento é importante para calcularmos preços brutos previstos. Assim, podemos definir os retornos de preços brutos reais $\Delta \mathbf{x}(t)$ e os previstos $\hat{\Delta} \mathbf{x}(t)$ utilizando as equações 22 e 26, respectivamente.

A partir da variação do preço no período, podemos definir o vetor de tendência

$$\mathbf{T} = \begin{bmatrix} T_1(t) \\ T_2(t) \\ \vdots \\ T_n(t) \end{bmatrix}, \quad (29)$$

onde T assume os valores -1 , 0 e 1 , sendo -1 associado à tendência de baixa, 0 sinal neutro e 1 à tendência de alta do preço. Deste ponto em diante do trabalho, estamos interessados apenas na tendência e não no valor do preço previsto. Assim, o bom resultado do modelo ocorrerá quando a tendência do preço previsto for igual a tendência observada do preço real. Deste modo, o desempenho do modelo será dado pela acurácia

$$a = \frac{N_c}{N}, \quad (30)$$

onde N_c é o número de previsões de tendências corretas e N é o número total de previsões executadas pelo modelo de previsão.

Nos resultados que apresentamos em (RESENDE et al., 2017), mostramos que o tamanho do retorno previsto traz informação sobre a tendência. Logo, dividimos o conjunto de retornos previstos U em 4 subconjuntos U_1 , U_2 , U_3 e U_4 , de forma que cada subconjunto contenha um quarto dos retornos previstos. Os maiores retornos (em módulo) estão em U_4 , seguidos pelos retornos em U_3 , U_2 e U_1 , nesta ordem. Os melhores resultados em (RESENDE et al., 2017) foram obtidos para o subconjunto U_4 . Portanto, neste trabalho, vamos analisar apenas as tendências previstas obtidas com retornos previstos em U_4 . Para verificar se os retornos previstos estão em U_4 , considerado o quarto quartil, utilizamos uma janela de tamanho $Q = 100$ pontos. A

janela Q é ordenada e 25% dos maiores valores são armazenados em U_4 . Desta forma, a primeira operação só poderá ocorrer a partir do $(Q + 1)$ -ésimo sinal produzido. Antes disso, o modelo sempre retornará sinais neutros. Para verificar se o $(Q + 1)$ -ésimo sinal está em U_4 , comparamos se o seu valor é maior que o sinal armazenado na posição 75 da janela Q de retornos precedentes dispostos em ordem crescente. Se a condição é satisfeita, o modelo envia um sinal válido, caso contrário, um sinal neutro é enviado.

Capítulo 6

Exploração do modelo de previsão

O modelo, apresentado na seção 5.2, conforme a equação (18), foi utilizado para desenvolver o modelo de previsão, representado pela equação (24). Neste capítulo, vamos explorar algumas das possibilidades de combinações de parâmetros para ajustar este modelo de forma a obter melhores acurácias. Entretanto, encontrar a combinação *ótima* de parâmetros envolve análises que não estão no foco do trabalho. O capítulo será organizado no seguinte modo: na seção 6.1, investigaremos se há uma relação entre a acurácia a e a variação da quantidade de pontos R , utilizados para calcular a matriz de coeficientes \mathbf{A} . Na seção 6.2, vamos investigar se a quantidade de ações interfere na acurácia do modelo. Na seção 6.3, analisaremos a relação da acurácia do modelo com a granularidade dos dados. Na seção 6.4, calcularemos os expoentes de Hurst das séries de preços das ações e analisaremos se há uma relação entre as acurácias obtidas e os seus expoentes de Hurst. Já na seção 6.5, vamos investigar se as taxas de acurácia calculadas em um quadrimestre se mantêm no quadrimestre seguinte. Com esta análise, estamos interessados em verificar se o modelo se mantém para outro período. Também vamos selecionar quais as ações que apresentam melhores acurácias no primeiro quadrimestre e investigar as acurácias do modelo para estas ações no segundo quadrimestre. Na seção 6.6, vamos elaborar uma votação em relação à tendência entre diferentes tamanhos de R .

6.1 Variações do tamanho de R

O objetivo do modelo é prever a tendência do preço no tempo $t + 1$. Para isto, ele utiliza a matriz de coeficientes \mathbf{A} que minimiza a diferença quadrática entre a matriz de preços reais $\mathbf{X}(t + 1)$ e a matriz de preços previstos $\hat{\mathbf{X}}(t + 1)$. Para calcular esta matriz \mathbf{A} é preciso $R + 1$ pontos. Não sabemos o tamanho ideal de R , logo, definimos a quantidade de pontos baseados em alguns testes básicos de acurácia do modelo na previsão de tendência de preços. No Capítulo 5, Eq. (30), reproduzida abaixo,

definimos a acurácia para cada ação como

$$a = \frac{N_c}{N}, \quad (31)$$

onde N_c é o número de previsões corretas e N é o número total de previsões.

Como estamos analisando diferentes tamanhos de R para várias ações, vamos definir a acurácia média

$$\bar{a} = \frac{\sum_k^n a_k}{n} \quad (32)$$

e o desvio padrão

$$\sigma_a = \sqrt{\frac{\sum_k^n (a_k - \bar{a})^2}{n - 1}}, \quad (33)$$

onde a_k é a acurácia a para cada ação k e n é o número total de ações analisadas.

Nesta análise, uma acurácia média \bar{a} , envolvendo todas as k ações, é calculada para cada tamanho de R . A figura 6 mostra a acurácia média \bar{a} e $\pm\sigma_a$ do modelo para todas as ações que compõem os índices Dow Jones e Ibovespa. Os dados analisados são referentes ao primeiro quadrimestre de 2017. Utilizamos diferentes tamanhos de R variando de 100 a 1.100, com incremento de 10 pontos a cada nova execução do programa. O objetivo desta figura é sinalizar o momento em que não há mais ganhos significativos na acurácia à medida que aumentamos o tamanho de R .

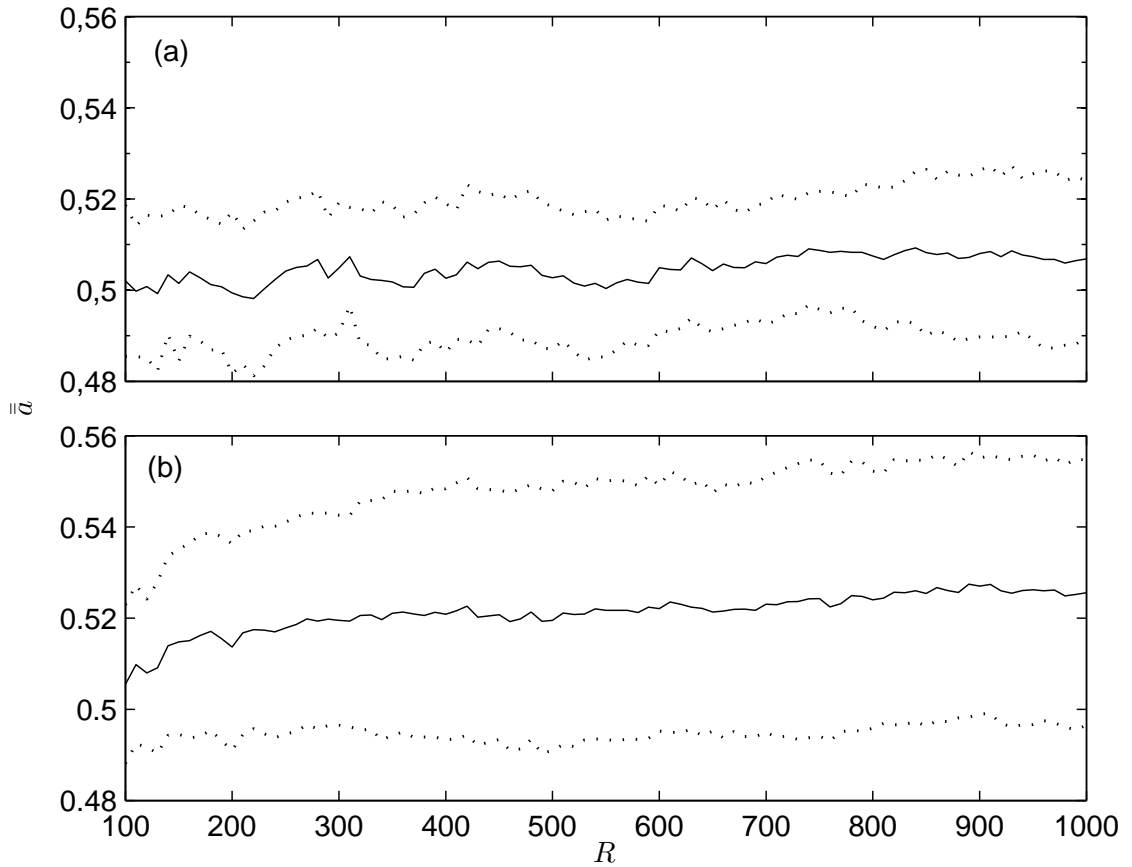


Figura 6 – **Acurácia média para todas as n ações que compõem os índices Dow Jones e Ibovespa.** Os dados analisados são referentes às séries de preços com granularidade de 5 minutos, no período de 02/01/17 a 28/04/17 (primeiro quadrimestre de 2017). Foram utilizados diferentes tamanhos de R . Com valor inicial de $R = 100$, com incremento de 10. Para cada valor de R , foram realizadas 5.000 execuções para a série do Ibovespa e 3.900 para Dow Jones. (a) A linha contínua representa a acurácia média \bar{a} obtida para todas as ações do índice Dow Jones. A linha pontilhada representa a acurácia média $\bar{a} \pm \sigma_a$ das acurácias obtidas. (b) A linha contínua mostra a acurácia média \bar{a} de todas as ações que compõem a carteira teórica do índice Ibovespa. A linha pontilhada apresenta a acurácia média $\bar{a} \pm \sigma_a$.

Podemos observar na figura 6 que para valores maiores de R a curva da acurácia média \bar{a} tende a quase se estabilizar, porém ainda podemos perceber um leve crescimento. Para analisar se ainda há ganhos na acurácia para valores de $R > 500$, calculamos o coeficiente angular β e os coeficientes de determinação r^2 da reta de ajuste aos pontos com valores de R superiores a 500. Os coeficientes angulares β para todas as ações dos índices Dow Jones e Ibovespa, calculados para o intervalo $500 \leq R \leq 1.000$, foram $\beta = 5,63 \times 10^{-5}$ e $\beta = 1,24 \times 10^{-5}$, respectivamente. Embora pequenos, não são irrelevantes: observe que numa variação de 0,02 na acurácia (o que é significativo) entre $R = 500$ e $R = 1.000$, temos coeficiente angular 4×10^{-5} , ou seja, na ordem dos observados. Os coeficientes de determinação foram $r^2 = 0,2905$ para todas as ações do índice Dow Jones e $r^2 = 0,6295$ para todas as ações do Ibovespa. Eles indicam que o ajuste linear não é satisfatório; Contudo, isto não é relevante aqui, uma vez que estamos interessados sobretudo no caráter crescente das curvas. Os valores de β e r^2 obtidos são apresentados na tabela 2. Valores dos coeficientes angulares β , calculados para o intervalo $500 \leq R \leq 1.000$, pequenos, porém positivos, indicam crescimento da curva e nos motivam a investigar o comportamento do gráfico para valores de R maiores que 1.000.

Tabela 2 – Coeficientes angulares (β) e os correspondentes coeficientes de determinação (r^2) para o ajuste linear das curvas de acurácias médias \bar{a} para diferentes valores de R das Fig. 6 e 7.

		Todas Ações (Fig. 6)	10 Ações + (Fig. 7)
Dow Jones	β	$5,63 \times 10^{-5}$	$9,79 \times 10^{-5}$
	r^2	0.2905	0.62953
Ibovespa	β	$1,24 \times 10^{-4}$	$9,98 \times 10^{-5}$
	r^2	0,6970	0,7629

Nota: Para a Fig. 6, que analisa o conjunto de todas as ações que compõem os índices Dow Jones e Ibovespa, os coeficientes β e r^2 foram calculados para o intervalo de $500 \leq R \leq 1000$. Já na Fig. 7, os coeficientes β e r^2 foram calculados para o intervalo de $1000 \leq R \leq 2000$. Todos os coeficientes β são positivos, indicando que as evoluções das curvas ainda são positivas.

Para cada tamanho de R , são realizadas 5.000 execuções para as ações do Ibovespa e 3.900 para as ações do Dow Jones. A diferença na quantidade de execuções entre Dow Jones e Ibovespa é devida ao menor tempo diário de permanência do pregão de negociação do Dow Jones em relação ao Ibovespa. Com isso, as quantidades de dados disponíveis nas amostras são diferentes. Ainda assim, o número de execuções para cada valor de R é grande. Como estamos interessados em analisar mais tamanhos diferentes de R , vamos diminuir o número de ações envolvidas nos testes. O critério para escolha destas ações é selecionar as 10 ações com maiores pesos nos índices Ibovespa e Dow Jones. As composições e as ordens de pesos das participações nos índices Ibovespa e Dow Jones são apresentadas nas tabelas 26 e 27, respectivamente.

Nesta análise, utilizando valores maiores de pontos nos ajustes, vamos explorar R variando de 10 a 2.000, com incremento de 10 pontos a cada nova execução do programa. Para cada execução, foram realizadas 3.000 e 2.000 previsões para as 10 ações com maiores pesos de participação nos índices Ibovespa e Dow Jones, respectivamente. A figura 7 apresenta as acurácias médias \bar{a} , linha contínua, e $\bar{a} \pm \sigma_a$, linha pontilhada, obtidos na análise.

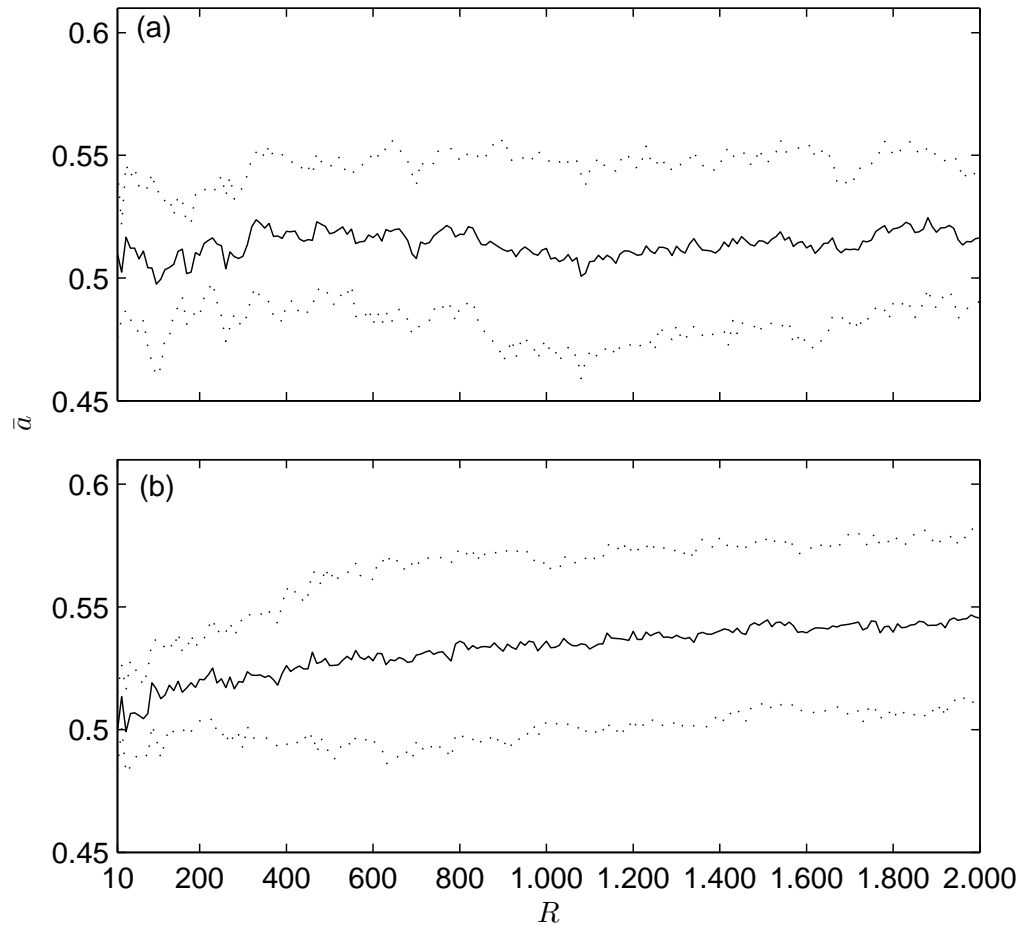


Figura 7 – **Acurácia média para as 10 ações com maiores participações na carteira teórica do Ibovespa E Dow Jones.** Os dados analisados são referentes às séries de preços de fechamento com granularidade de 5 minutos, no período de 02/01/17 a 28/04/17 (primeiro quadrimestre de 2017). Foram utilizados 200 diferentes tamanhos de R . Com o valor inicial de $R = 10$, com incremento de 10. Para cada valor de R , foram realizadas 3.000 execuções com as séries das ações do Ibovespa e 2.000 com a séries do Dow Jones. (a) A linha contínua representa a acurácia média obtida para as 10 ações com maior participação do índice Dow Jones. A linha pontilhada representa a média $\pm \sigma_a$, que é o desvio padrão das acurácias obtidas. (b) A linha contínua mostra a acurácia média para as 10 as ações com maior peso de participação na carteira teórica do índice Ibovespa. A linha pontilhada apresenta a acurácia média $\pm \sigma_a$.

Analisando a figura 7, percebemos que à medida que aumentamos os valores de R , a curva da acurácia média tende a oscilar menos. Isso acontece devido ao incremento de

10 novos pontos serem cada vez menos significativos à medida em que aumentamos a quantidade de pontos na janela. O efeito é mais evidente ao compararmos os valores de R próximos a 10 com aqueles no final da série.

A figura 7 sugere que a acurácia para as ações do índice Dow Jones tende a se estabilizar para valores de $R \geq 400$ pontos. Porém, a curva ainda mostra um leve crescimento dos resultados. Para as ações do índice Ibovespa, é evidente que a curva de acurácia apresenta ganhos significativos à medida em que aumentamos o tamanho de R . Estas observações são evidenciadas pelos valores dos coeficientes angulares apresentados na tabela 2. Desta forma, na aplicação do modelo, seria ideal utilizar tamanhos maiores de R , mas esta quantidade de pontos definida para o ajuste afeta diretamente o tamanho da amostra restante para fazer as N previsões a cada execução do programa. Por isto, definimos, baseados nas figuras 6 e 7, um valor de R que não comprometa a acurácia e nem o tamanho da amostra. Uma escolha coerente da quantidade de pontos para utilizar no ajuste é $R = 1.000$.

As figuras 6 e 7, nos mostram as acurácias médias \bar{a} . Como estes resultados se referem às médias das acurácias a de *todas* as ações para cada valor de R , não podemos visualizar o comportamento *individual* das ações para os diferentes tamanhos de R . Com isto, para mostrar a dispersão das acurácias das ações para cada valor de R , construímos os boxplots¹ dos resultados. A dispersão, localização e simetria das acurácias podem ser representadas pela figura 9. As caixas contêm 50% dos dados, sendo o primeiro quartil (Q1), a mediana e o terceiro quartil (Q3). A altura da caixa é a amplitude interquartil (AIQ), que é dada pela distância entre Q1 e Q3. A linha superior é traçada iniciando em Q3 até o maior valor observado dentro da faixa com o conjunto de pontos que vai de Q3 até o valor do limite superior (Ls), calculado utilizando 1,5 vezes o valor do interquartil AIQ,

$$Ls = Q3 + 1,5AIQ. \quad (34)$$

A linha inferior é traçada iniciando-se em Q1 até o menor valor que é encontrado dentro da faixa que começa em Q1 e se estende até o valor do limite inferior (Li), calculado utilizando 1,5 do valor do interquartil AIQ,

$$Li = Q1 - 1,5AIQ. \quad (35)$$

Os pontos fora destas linhas traçadas são chamados *outliers* (MONTGOMERY, 2008). A figura 8 mostra as características do boxplot. Como esperado, a figura 9 mostra que as ações que compõem o índice Ibovespa apresentam mais outliers com altas taxas de acurácia quando comparadas às ações do índice Dow Jones.

¹Diagrama de caixa. "É uma apresentação gráfica que descreve simultaneamente várias características de um conjunto de dados observados."(MONTGOMERY, 2008)

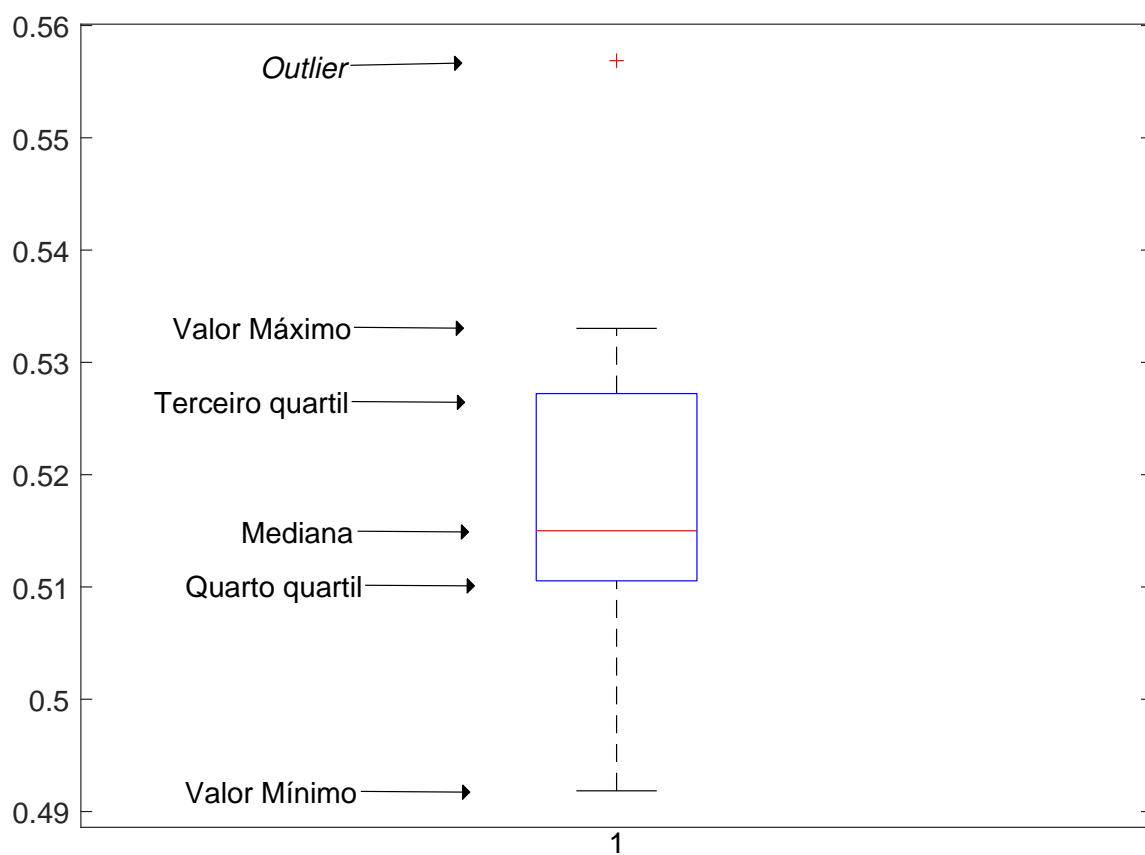


Figura 8 – Exemplo de um formato de um boxplot.

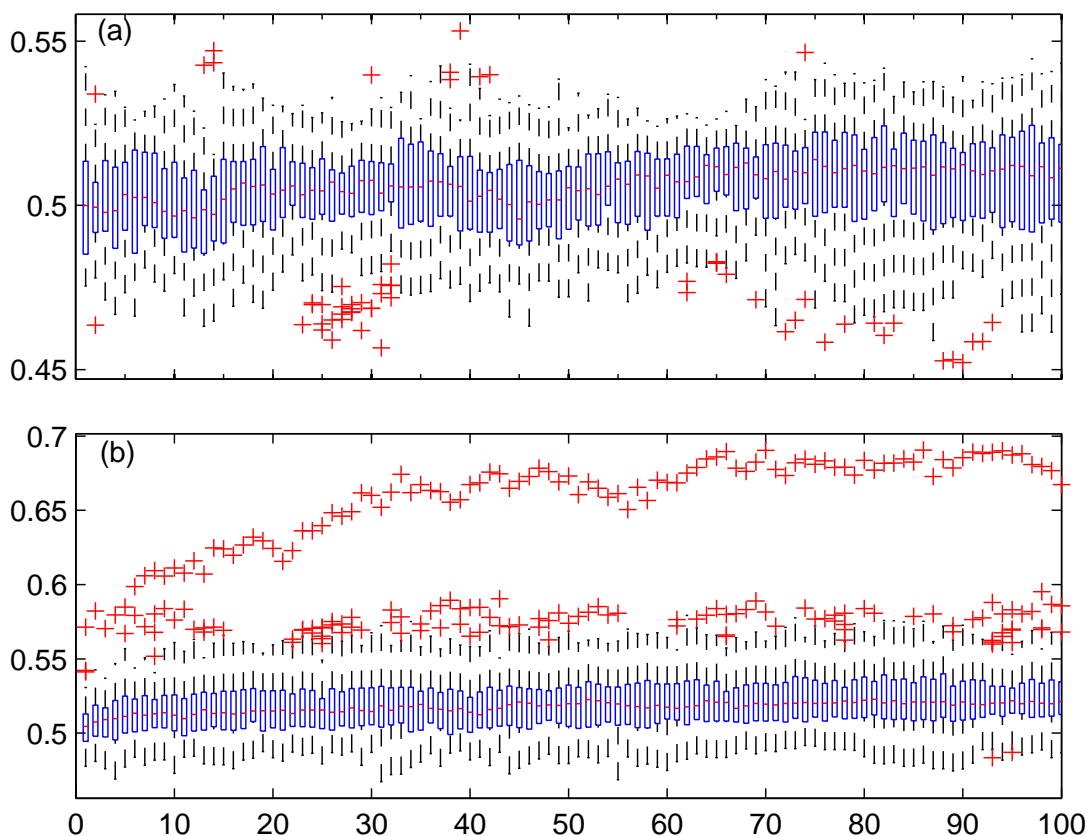


Figura 9 – **Boxplot da Acurácia para todas as ações que compõem os índices Dow Jones e Ibovespa.** Os dados utilizados foram preços de fechamento das ações que compõem os índices Dow Jones e Ibovespa com granularidade de 5 minutos, no período de 02/01/17 a 28/04/17 (primeiro quadrimestre de 2017). Foram analisados 100 diferentes tamanhos de R . Para cada valor de R , foram realizadas 3.900 execuções para as ações do índice Dow Jones e 5.000 execuções para as ações que compõem a carteira teórica do Ibovespa. (a) Boxplots com as acurácias para todas as ações do Dow Jones. (b) Acurácias para todas as ações do índice Ibovespa.

6.2 Variações da quantidade de ações

Um ponto a ser questionado refere-se à quantidade de ações n que devemos usar para calibrar o modelo de forma a obter a melhor acurácia. Para investigar a influência deste valor n nos resultados do modelo, desenvolvemos uma análise que consiste em executar o modelo com o número mínimo de combinações, $n = 2$, e expandir até o número máximo de ações que compõem os índices Ibovespa e Dow Jones. O valor de n aumenta seguindo a ordem decrescente do peso de participação das ações nos índices. Para esta análise utilizamos os dados de preços de fechamento das ações com granularidade de 5 minutos no período de 02/01/17 a 28/04/17, que se refere ao primeiro quadrimestre de 2017. A quantidade de pontos utilizada no ajuste foi $R = 1.000$ pontos. Para cada combinação de n ações, realizamos $N = 3.000$ previsões para as séries do Ibovespa e $N = 2.000$ previsões para as séries do Dow Jones. A figura 10 mostra a acurácia média \bar{a} e $\pm\sigma_a$ para diferentes combinações de quantidades de ações envolvidas no modelo.

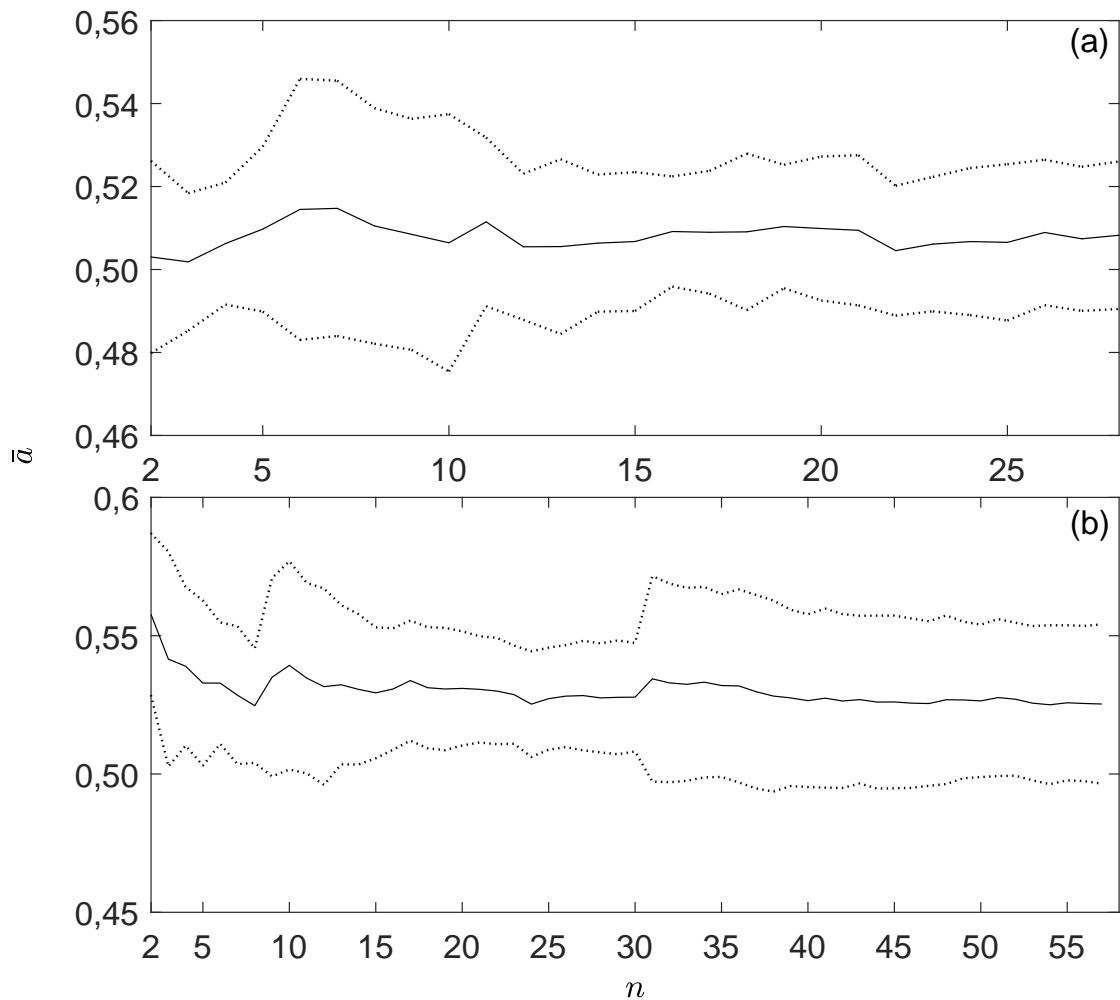


Figura 10 – **Acurácia média para diferentes valores de n ações que compõem as carteiras do Ibovespa e Dow Jones.** O valor de n aumenta seguindo a ordem decrescente do peso de participação das ações nos índices. Os dados analisados são referentes às séries de preços de fechamento com granularidade de 5 minutos, no período de 02/01/17 a 28/04/17 (primeiro quadrimestre de 2017). A quantidade de pontos utilizados no ajuste foi $R = 1.000$. Foram realizadas 3.000 previsões com as séries das ações do Ibovespa e 2.000 com as séries do Dow Jones. (a) A linha contínua representa a acurácia média obtida para n variando de 2 a 30 ações que compõem o índice Dow Jones. A linha pontilhada representa a média $\pm\sigma_a$, que é o desvio padrão das acurácias obtidas. (b) A linha contínua mostra a acurácia média para n variando de 2 a 58 ações que compõem a carteira teórica do índice Ibovespa. A linha pontilhada apresenta a acurácia média $\pm\sigma_a$.

Na figura 10, percebemos uma variação maior nas acurácias médias \bar{a} para as combinações de n ações do Ibovespa, quando comparadas às acurácias obtidas com as ações do índice Dow Jones. Este resultado é mais evidente para valores menores de n . As combinações de n ações do índice Dow Jones são mais estáveis e, como esperado, apresentam valores de acurácias médias \bar{a} menores quando comparados aos resultados das combinações das ações do Ibovespa, uma vez que a bolsa americana é mais eficiente que a brasileira. Para as combinações das ações do Ibovespa, observamos um pico maior em torno de $n = 10$ e depois uma certa estabilização de \bar{a} . Isto nos ajuda a definir a quantidade de ações que devemos usar para calibrar o modelo.

6.3 Variação da granularidade

Nesta seção, vamos analisar as acurácias para dados com granularidade de 5, 10, 15, 20 e 30 minutos. Para simetria das análises, definimos que o modelo deve ser aplicado em um mesmo intervalo de tempo. Com isto, à medida em que aumentamos a granularidade das séries, diminuimos a quantidade de dados da amostra. Assim, nesta análise, não utilizamos o valor de R definido da seção 6.1. O intervalo de negociação também foi alterado para um período referente a 2 quadrimestres. A quantidade de dados utilizada no ajuste foi $R = 500$ e o período de negociação foi de 02/01/17 a 31/08/17 (primeiro e segundo quadrimestres de 2017). Baseados na seção 6.2, além da composição total do índice, uma nova combinação quantidade de ações, $n = 10$, também foi analisada. A tabela 3 mostra as médias das acurácias \bar{a} para os parâmetros definidos.

Tabela 3 – Acurácia do modelo aplicado a diferentes granularidades de tempo para $R = 500$

Índice	n	5'	10'	15'	20'	30'
Dow Jones	30	0,5079 (0,0092)	0,5065 (0,0090)	0,5012 (0,0162)	0,5052 (0,0255)	0,5022 (0,0309)
	10	0,5132 (0,0116)	0,5062 (0,0136)	0,5143 (0,0178)	0,5061 (0,0201)	0,4976 (0,0337)
Ibovespa	58	0,5216 (0,0260)	0,5222 (0,0287)	0,5185 (0,0271)	0,5215 (0,0276)	0,5143 (0,0338)
	10	0,5316 (0,0294)	0,5196 (0,0200)	0,5187 (0,0186)	0,5033 (0,0268)	0,5021 (0,0251)

Nota: Os resultados são referentes às médias e os desvios das acurácias do modelo de todas as ações que compõem os índices Dow Jones (30 ações) e Ibovespa (58 ações). Também analisamos as 10 ações com maiores participações nesses índices. Os dados analisados são referentes às séries de preços de fechamento com granularidades de 5, 10, 15, 20 e 30 minutos, no período de 02-01-17 a 31-08-17 (primeiro e segundo quadrimestres de 2017). A quantidade de pontos utilizada no ajuste foi $R = 500$.

Analisando a tabela 3, percebemos que não há grandes oscilações nas acurácias médias \bar{a} quando variamos a granularidade da base de dados. Menores variações da acurácia média \bar{a} são observadas quando aplicamos o modelo para todas as ações do índice Dow Jones. Para a combinação das 10 ações com maiores pesos, um moderado aumento nas acurácias foi obtido para as granularidades de 5 e 15 minutos.

Para as ações do Ibovespa, maiores oscilações foram observadas, principalmente, para a combinação das 10 ações com maiores participações no índice. A maior acurácia média obtida foi para a granularidade de 5 minutos, $\bar{a} = 0,5316$. Em linhas gerais, observamos uma discreta tendência que a acurácia diminua com o aumento da granularidade dos dados para as ações do Ibovespa. Para as ações do índice Dow Jones, não observamos tendência alguma.

6.4 Análise da acurácia em função do expoente de Hurst das séries

O expoente de Hurst (H), aplicado a séries de preços de ações, nos permite visualizar o comportamento dos dados em relação à memória de longo prazo. Nas nossas análises, vamos utilizar o método alternativo DFA introduzido por (PENG et al., 1994), que também foi explorado por (HU et al., 2001) e (CHEN et al., 2002). Para realizar o cálculo, consideramos uma série de retornos $\Delta x(i)$ ($i = 1, 2, \dots, N_{max}$). Integrando tal série, obtemos

$$y(j) = \sum_{i=1}^j (\Delta x(i) - \langle \Delta x \rangle), \quad (36)$$

onde $\langle \Delta x \rangle$ é a média

$$\langle \Delta x \rangle = \frac{1}{N_{max}} \sum_{j=1}^{N_{max}} \Delta x(i). \quad (37)$$

A nova série é dividida em janelas de tamanho igual m . Em cada uma dessas janelas, ajustamos a série integrada, encontrando a função polinomial Y_{fit} . Esta função indica a tendência local que há em cada intervalo. O grau do polinômio utilizado pode, em princípio, ser qualquer. O método DFA-I é aquele em que o grau 1 é escolhido.

Na série temporal dividida em janelas de tamanho m , subtraímos a tendência local Y_{fit} em cada das janelas. Em seguida, calculamos a função de flutuação.

$$F(m) = \sqrt{\frac{1}{N_{max}} \sum_{i=1}^{N_{max}} (y(i) - Y_{fit})^2}. \quad (38)$$

Todo o cálculo acima é repetido para janelas de diferentes tamanhos m para obter-se uma relação entre $F(m)$ e m e os dados são dispostos em um gráfico $\log - \log$. Uma relação de lei de potência entre $F(m)$ e a janela de tamanho m indica a presença de escala: $\log F(m) \cong \log m^H$ ($F(m) \cong m^H$). Neste caso, o parâmetro H , dado pelo coeficiente angular da reta que ajusta os pontos $(m, F(m))$, é chamado de expoente de escala ou expoente de correlação, e é a generalização do expoente de Hurst. Valores de $H < 0,5$ indicam anti-persistência da série, valores de $H > 0,5$ indicam que a

série tem persistência, e valores de $H = 0,5$ indicam que as séries não possuem tendência. Com o intuito de analisar se os dados passados fornecem informações sobre dados futuros, calculamos os expoentes de Hurst das séries com preços brutos x de *todas* as ações que compõem os índices Dow Jones e Ibovespa. Os valores de H obtidos são mostrados na figura 11.

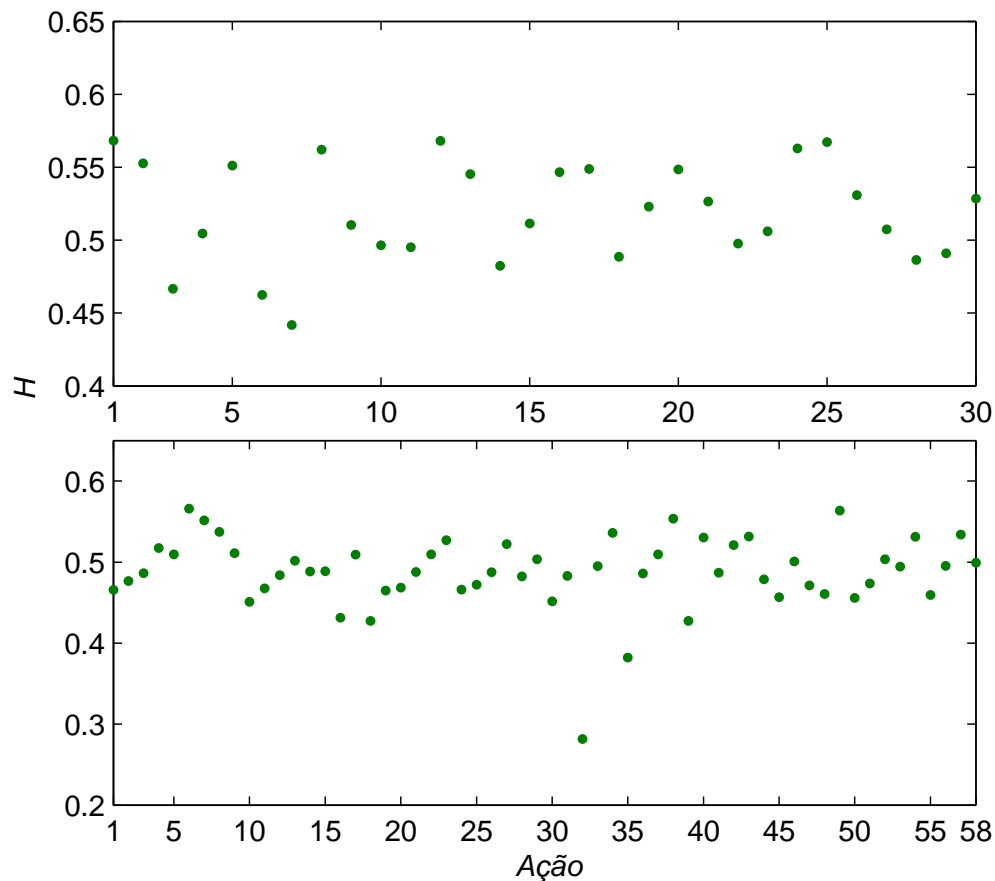


Figura 11 – **Expoente de Hurst para todas ações que compõem os índices Dow Jones e Ibovespa.** Os dados analisados são referentes às séries de preços brutos com granularidade de 5 minutos, no período de 02/01/17 a 28/04/17 (primeiro quadrimestre de 2017). (a) Representa os coeficientes de Hurst para as ações do índice Dow Jones. Foram utilizados 6276 pontos para o cálculo. (b) Mostra os coeficientes de Hurst utilizando $l = 1$ para as ações que compõem o índice Ibovespa. Foram utilizados 7370 pontos para o cálculo.

Analisando a figura 11, percebemos que a maioria dos resultados se concentram na faixa de $0,45 \leq H \leq 0,55$. Assim, não percebemos forte tendência a persistência ou anti-persistência das séries analisadas.

Na análise seguinte, vamos investigar a relação entre os resultados dos expoentes de Hurst e as acurácias obtidas pelo modelo. As séries de preços ainda são as referentes aos preços de fechamento de todas as ações que compõem os índices Dow Jones e Ibovespa. A Fig. 12 mostra a acurácia média \bar{a} de cada uma destas ações que

compõem os índices para 100 diferentes valores de R e os expoentes de Hurst das séries de preços brutos x destas ações.

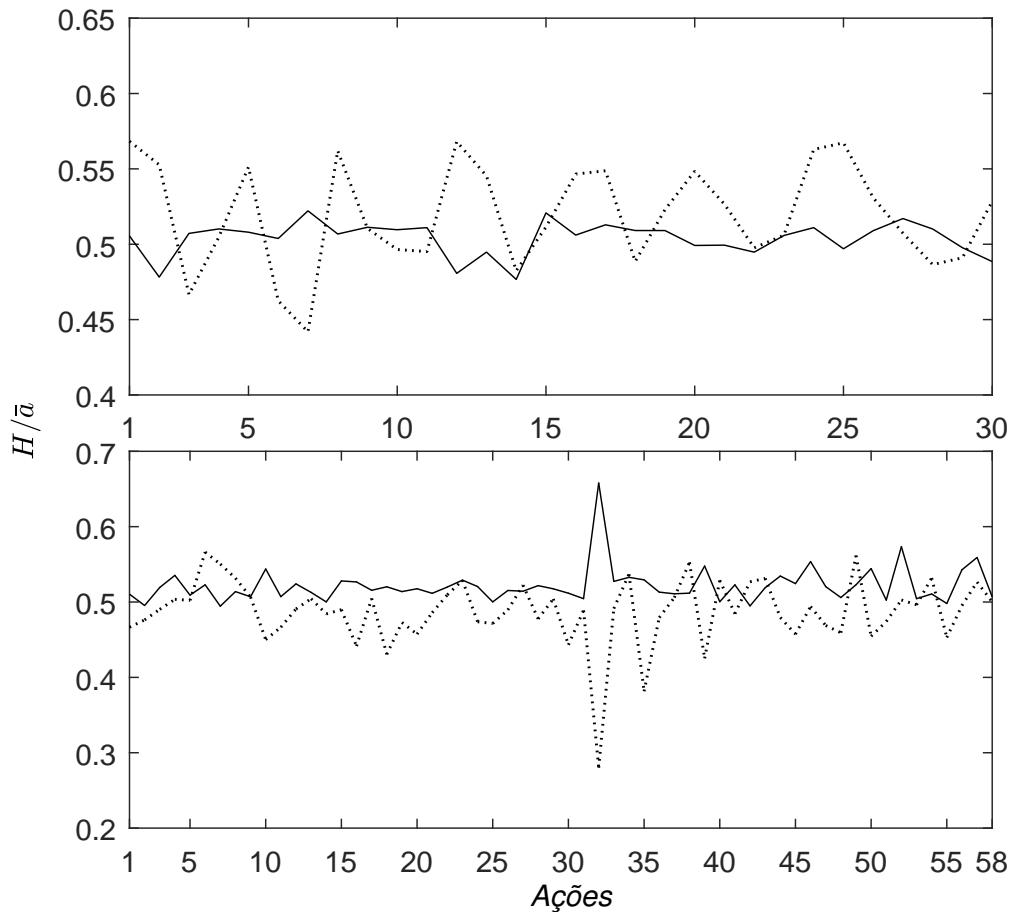


Figura 12 – **Acurácia média e expoente de Hurst para todas as ações que compõem os índices Dow Jones e Ibovespa.** Os dados analisados são referentes às séries de preços brutos com granularidade de 5 minutos, no período de 02/01/17 a 28/04/17 (primeiro quadrimestre de 2017). Para o cálculo da acurácia média foram utilizados 100 tamanhos diferentes de R para cada ação. A linha tracejada representa os coeficientes de Hurst H e a linha contínua a acurácia média \bar{a} para cada ação, utilizando 100 valores diferentes de R no ajuste. O valor de R inicial foi 100, com incremento de 10 pontos a cada nova execução. (a) Representa a acurácia média e os coeficientes de Hurst para as ações do índice Dow Jones. Foram utilizados 6276 pontos para o cálculo do expoente de Hurst. (b) Mostra a acurácia média e os coeficientes de Hurst para as ações que compõem o índice Ibovespa. Foram utilizados 7370 pontos para o cálculo dos coeficientes.

Analisando visualmente a figura 12, não percebemos padrões para as ações que compõem o índice Dow Jones. As suas taxas de acurácia média \bar{a} , bem como os valores de H , estão próximos de 0,5. Já as ações que compõem o índice Ibovespa apresentaram maiores taxas de acurácia média \bar{a} e também maiores oscilações nos valores dos expoentes de Hurst H . A figura nos sugere, ainda, que os valores menores de H resultam em maiores valores de \bar{a} , indicando que o modelo tem maior capacidade

de previsão para séries de preços de ações que apresentam anti-persistência. Os resultados mais discrepantes foram obtidos para a ação 32, com $\bar{a} = 0,68$ e $H = 0,28$. Para melhor visualização destes resultados, construímos a figura 13, que mostra a concentração dos expoentes de Hurst H em função da acurácia média \bar{a} , obtidos para cada uma das ações que compõem os índices Dow Jones e Ibovespa, representados pelos pontos nas cores verde e vermelha, respectivamente.

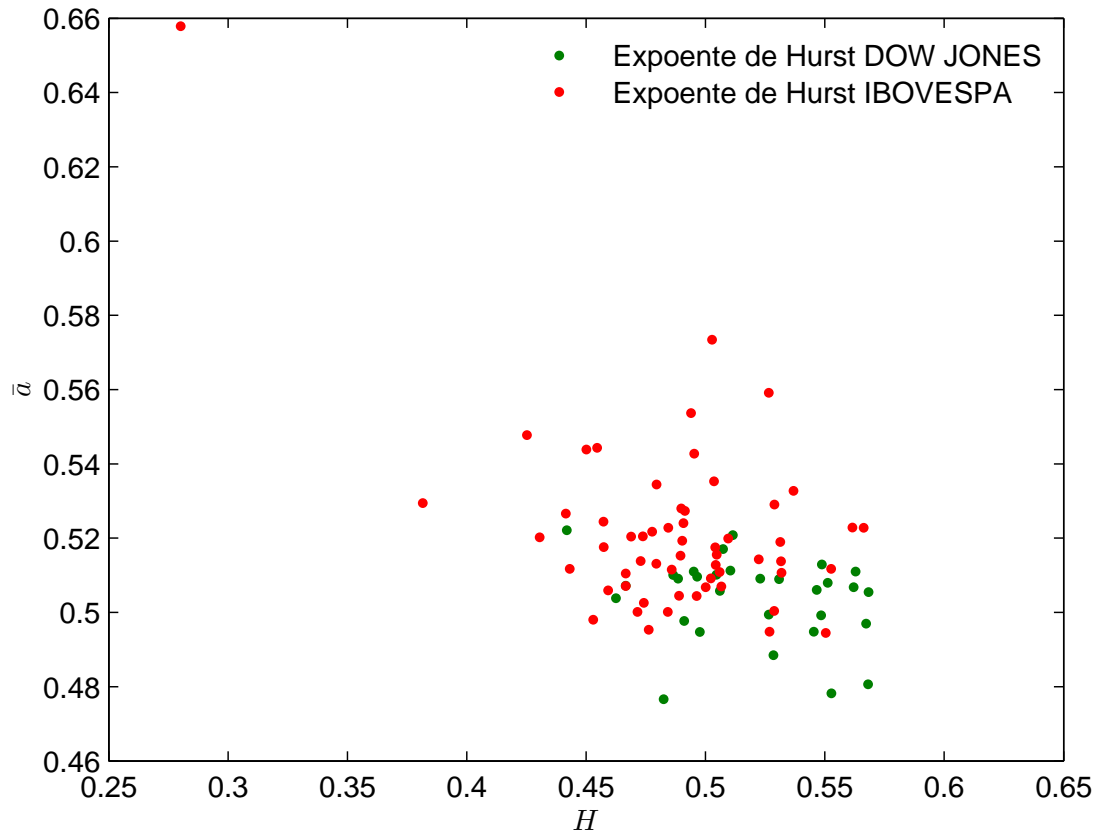


Figura 13 – **Acurácia média e expoente de Hurst para todas as ações que compõem os índices Dow Jones e Ibovespa.** Os dados analisados são referentes às séries de preços brutos com granularidade de 5 minutos, no período de 02/01/17 a 28/04/17 (primeiro quadrimestre de 2017). Para o cálculo da acurácia média \bar{a} foram utilizados 100 tamanhos diferentes de R para cada ação. Com valor inicial de $R = 100$, com incremento de 10 pontos a cada nova execução. Os pontos na cor vermelha representam os expoentes H em função da acurácia \bar{a} para as ações do Ibovespa. Os pontos na cor verde representam os coeficientes de Hurst H em função da acurácia média \bar{a} para cada ação do índice Dow Jones. Para as ações do índice Dow Jones, utilizamos 6276 pontos para o cálculo de H . Para as ações que compõem o índice Ibovespa, foram utilizados 7370 pontos para o cálculo dos coeficientes de Hurst.

Pela figura 13, fica mais evidente que as taxas de acurácia média \bar{a} são maiores para as ações que compõem o índice Ibovespa quando comparadas às ações do índice Dow Jones. Observamos, também, que os maiores desvios de H foram obtidos para

a série do Ibovespa. Estes resultados vão ao encontro dos estudos de (CAJUEIRO; TABAK, 2004), onde os países emergentes apresentam valores de H mais distantes de 0,5, sugerindo que mercado de países emergentes são mais ineficientes que os de países desenvolvidos.

Outra investigação relevante foi baseada no cálculo da correlação linear das acurácias obtidas em cada quadrimestre com os expoentes de Hurst calculados nos respectivos quadrimestres. Para tanto, calculamos o coeficiente de correlação linear de Pearson ρ^2 . O coeficiente ρ obtido quando analisada a correlação das acurácias e os expoentes de Hurst calculados no quadrimestre 1 foi $\rho = -0,3791$ ($p - valor = 0,0388$) para as ações que compõem o índice Dow Jones e para as ações do Ibovespa, no mesmo quadrimestre, o resultado foi $\rho = -0,4622$ ($p - valor = 2,6 \times 10^{-4}$). A correlação entre as acurácias e os coeficientes de Hurst das séries de preços calculados no segundo quadrimestre de 2017 foi $\rho = -0,6966$ ($p - valor = 1,2571 \times 10^{-9}$) para as ações do Ibovespa e $\rho = -0,1743$ ($p - valor = 0,3568$) para as ações do Dow Jones. Estes resultados indicam que há uma anti-correlação entre as acurácias obtidas e os expoentes H das séries analisadas (COHEN, 1977).

Baseados nestes resultados, em etapas futuras da tese, vamos aprofundar mais as análises dos valores de H e \bar{a} , explorando as questões de memória das séries de preços de ações e investigando até que ponto informações dos dados passados interferem na acurácia do modelo de previsão.

6.5 Acurácias em quadrimestres diferentes

Nas subseções seguintes vamos aplicar o modelo em dados de quadrimestres diferentes com o objetivo de investigar a consistência das previsões em diferentes períodos.

6.5.1 Comparação das acurácias obtidas no primeiro e segundo quadrimestres de 2017

Nesta análise, aplicamos o modelo utilizando as séries de preços tratados com granularidade de 5 minutos do primeiro e segundo quadrimestres de 2017 de todas as ações que integraram os índices Ibovespa e Dow Jones no primeiro quadrimestre de 2017. A tabela 4 mostra as acurácias obtidas nos quadrimestres 1 e 2 do ano de 2017. Para a execução do modelo, foi utilizado como tamanho de ajuste $R = 1.000$. A coluna 'Acurácia Q1' refere-se às acurácias obtidas no quadrimestre 1. Analogamente, a coluna denominada 'Acurácia Q2' refere-se às acurácias no quadrimestre 2. Os resultados estão dispostos na ordem da porcentagem de peso de participação de

²Mede o grau da correlação ($-1 \leq \rho \leq 1$) e a direção dessa correlação (+ ou -) entre duas variáveis

cada ação nos seus respectivos índices. Os valores em **negrito** indicam que a acurácia da ação obtida no quadrimestre 2 foi maior em relação à mesma no quadrimestre 1. Este aumento de ganhos do quadrimestre 2 em relação ao quadrimestre 1 foi em 53,30% das ações que compõem o Dow Jones e 60,30% das ações do Ibovespa.

Tabela 4 – Comparação das acurácias obtidas no primeiro e segundo quadrimestres para as ações dos índices Ibovespa e Dow Jones.

Ibovespa			Dow Jones		
Ação	Acurácia Q1	Acurácia Q2	Ação	Acurácia Q1	Acurácia Q2
1	0,5183	0,5223	1	0,4912	0,4948
2	0,4833	0,5329	2	0,4577	0,5340
3	0,5054	0,5427	3	0,5195	0,5126
4	0,5363	0,5510	4	0,5207	0,5042
5	0,5130	0,5296	5	0,5016	0,5079
6	0,5363	0,5268	6	0,5208	0,4813
7	0,4965	0,4974	7	0,5341	0,5117
8	0,5188	0,5203	8	0,5011	0,5230
9	0,5226	0,4840	9	0,5097	0,4995
10	0,5541	0,6053	10	0,5103	0,5159
11	0,4862	0,5597	11	0,5087	0,5124
12	0,5362	0,5578	12	0,5000	0,4893
13	0,5147	0,5493	13	0,4894	0,5011
14	0,5097	0,5070	14	0,4839	0,4989
15	0,5449	0,5349	15	0,5178	0,4907
16	0,5072	0,5479	16	0,5138	0,5058
17	0,5320	0,4852	17	0,5218	0,4793
18	0,5260	0,5283	18	0,5021	0,4952
19	0,5350	0,4968	19	0,5175	0,5142
20	0,5266	0,5262	20	0,5049	0,5318
21	0,5208	0,5166	21	0,4843	0,5128
22	0,5270	0,5138	22	0,4807	0,4964
23	0,5100	0,5180	23	0,5089	0,5217
24	0,5208	0,5210	24	0,5256	0,4829
25	0,5057	0,5028	25	0,4904	0,4821
26	0,5142	0,5167	26	0,5023	0,5084
27	0,5018	0,5120	27	0,5222	0,5224
28	0,5138	0,5169	28	0,5230	0,5204
29	0,5142	0,5465	29	0,5061	0,5272
30	0,5043	0,5020	30	0,4710	0,5113
31	0,5201	0,5353			
32	0,6810	0,7418			
33	0,5414	0,5379			
34	0,5458	0,5098			
35	0,5362	0,5964			
36	0,5053	0,5165			
37	0,4968	0,5522			
38	0,5055	0,5431			
39	0,5494	0,5165			
40	0,4969	0,5090			
41	0,5230	0,5125			
42	0,5018	0,5230			
43	0,5288	0,5035			
44	0,5396	0,5189			
45	0,5048	0,5318			
46	0,5467	0,5304			
47	0,5181	0,5213			
48	0,5122	0,5051			
49	0,5385	0,5225			
50	0,5457	0,5565			
51	0,5227	0,5494			
52	0,5845	0,5802			
53	0,5214	0,5225			
54	0,5197	0,5304			
55	0,5229	0,5339			
56	0,5651	0,5514			
57	0,5612	0,6110			
58	0,5121	0,5205			

Nota: As acurácias estão representadas de acordo com a ordem decrescente da porcentagem de participação das ações nos índices Dow Jones e Ibovespa. Os resultados das 10 ações com maiores pesos estão delimitados na tabela. Os valores em negrito indicam que a acurácia aumentou no segundo quadrimestre em relação ao primeiro, 60,3% dos casos para o índice Ibovespa e 53,3% dos casos para o índice Dow Jones.

A figura 14 mostra o panorama das acurácias apresentadas na tabela 4. A linha traçada no gráfico delimita as 10 ações que apresentaram as melhores acurácias em cada quadrimestre. Estas ações são esboçadas na cor vermelha. Os pontos na cor vermelha e circulados, indicam que, além de estarem entre as 10 ações que apresentaram as melhores acurácias no quadrimestre 1, também estão entre as 10 ações que apresentaram as melhores taxas no quadrimestre 2. O número destas ações foi maior para as ações do Ibovespa.

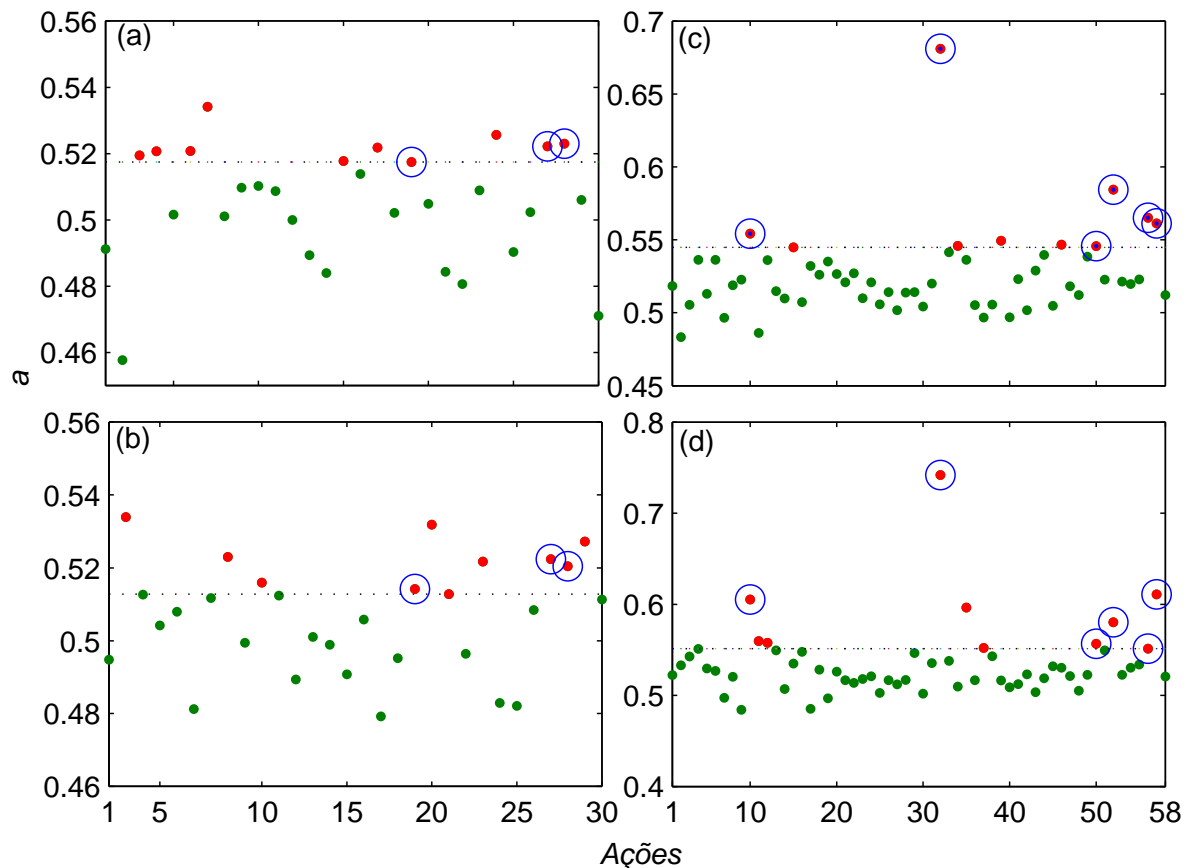


Figura 14 – **Acurácias obtidas no primeiro e segundo quadrimestres de 2017.** Os resultados são referentes às acurácias do modelo de todas as ações que compõem os índices Dow Jones e Ibovespa. Os dados analisados são referentes às séries de preços de fechamento com granularidades de 5 minutos, no período de 02/01/17 a 31/08/17 (primeiro e segundo quadrimestres de 2017). O tamanho do ajuste utilizado foi $R = 1.000$. A linha traçada sobre a figura delimita as 10 ações que apresentaram as melhores acurácias no quadrimestre em questão, mostradas na cor vermelha. Os pontos circulados indicam que as ações apresentaram acurácias entre as 10 melhores no quadrimestre 1 e também no quadrimestre 2. (a) Representa as acurácias de todas as ações que compõem o índice Dow Jones no primeiro quadrimestre. (b) Mostra as acurácias obtidas no segundo quadrimestre para todas as ações do índice Dow Jones. (c) e (d) Representam as acurácias obtidas no primeiro e segundo quadrimestres, respectivamente, para todas as ações que compõem o índice Ibovespa.

Se a presença entre as ações com 10 maiores acurácias for completamente aleatória, a chance de uma ação que está nesta situação repetir o comportamento no quadrimestre seguinte é $\frac{10}{n}$, onde n é o número total de ações no grupo analisado. Assim, o número médio de ações que aparecem em dois quadrimestres consecutivos entre aqueles com as 10 maiores acurácias é $10 \times \frac{10}{n}$. Para as ações que compõem o índice Dow Jones, esse número é $\frac{10}{3}$, para as ações que compõem o Ibovespa, é $\frac{100}{58}$. Na figura 14, observamos 3 ações nessa situação para o Dow Jones, sugerindo a completa aleatoriedade. Já para o Ibovespa, a figura 14 mostra 6 ações nessa situação, um número 3,48 vezes maior que o esperado, indicando que o processo não deve ser completamente aleatório. Estudos envolvendo um maior número de quadrimestres seriam interessantes para validação estatística.

O coeficiente de correlação linear de Pearson ρ foi calculado para mostrar a relação entre as acurácias obtidas no quadrimestre 1 e 2 para ambos mercados. O coeficiente ρ obtido quando analisada a correlação das acurácias no quadrimestre 1 e 2 para as ações que compõem o Dow Jones foi $\rho = -0.1493$ ($p - valor = 0,4311$). Já para as ações que compõem o Ibovespa, a correlação entre as acurácias obtidas nos quadrimestres 1 e 2 foi $\rho = 0.7156$ ($p - valor = 2,7521 \times 10^{-10}$). O resultado indica que há uma correlação significativa entre as acurácias obtidas para as ações do Ibovespa.

6.5.2 Aplicação do modelo utilizando a combinação das 10 ações com maiores pesos nos índices Ibovespa e Dow Jones

Na subseção 6.5.1, apresentamos as acurácias obtidas com a combinação de todas as ações dos índices no modelo. Aqui, estamos interessados em investigar o resultado da combinação linear apenas das 10 ações com maiores pesos de participação nos índices. Para tanto, vamos executar o programa utilizando os mesmos parâmetros da análise anterior. A tabela 5, mostra os resultados de acurácias obtidos. A coluna denominada 'Acurácia Q1' mostra as acurácias obtidas no quadrimestre 1 e a coluna 'Acurácia Q2' mostra as acurácias do quadrimestre 2. Os valores em negrito indicam que a acurácia aumentou no quadrimestre 2 em relação ao quadrimestre 1, 70% dos casos do Ibovespa e 60% para o Dow Jones.

Tabela 5 – Acurácias do modelo no primeiro e segundo quadrimestres para a combinação das 10 ações com maiores participações nos índices Ibovespa e Dow Jones.

Ibovespa			Dow Jones		
Ação	Acurácia Q1	Acurácia Q2	Ação	Acurácia Q1	Acurácia Q2
1	0,5417	0,5630	1	0,4857	0,5189
2	0,5142	0,5293	2	0,4706	0,4969
3	0,5604	0,5496	3	0,5142	0,5054
4	0,5445	0,5770	4	0,5096	0,5207
5	0,5197	0,5336	5	0,4867	0,5018
6	0,5299	0,5226	6	0,5369	0,5050
7	0,5089	0,5354	7	0,5547	0,5302
8	0,4897	0,5430	8	0,4853	0,4981
9	0,5258	0,5212	9	0,5322	0,5405
10	0,6155	0,6689	10	0,5113	0,5022

Nota: Os valores em negrito indicam que a acurácia aumentou no segundo quadrimestre em relação ao primeiro. Percebemos que os resultados são melhores no segundo quadrimestre para as ações dos dois índices analisados. A relação de melhoria da acurácia foi de 70% para as ações do Ibovespa e 60% para as ações do Dow Jones.

Analisando, de forma geral, as tabelas 4 e 5, não é perceptível se houve ou não ganhos quando combinadas todas as ações que compõem os índices em relação a quando combinamos apenas as 10 ações com maiores pesos nestes mesmos índices. Para facilitar a análise, construímos a figura 15, que mostra as acurácias obtidas no primeiro e segundo quadrimestres de 2017 quando combinamos todas as ações dos índices Dow Jones e Ibovespa, e, também, quando combinamos apenas a 10 ações com maiores participações nestes índices. O eixo das ordenadas do gráfico se refere às acurácias obtidas no quadrimestre 2, já o eixo das abscissas se refere às acurácias no quadrimestre 1.

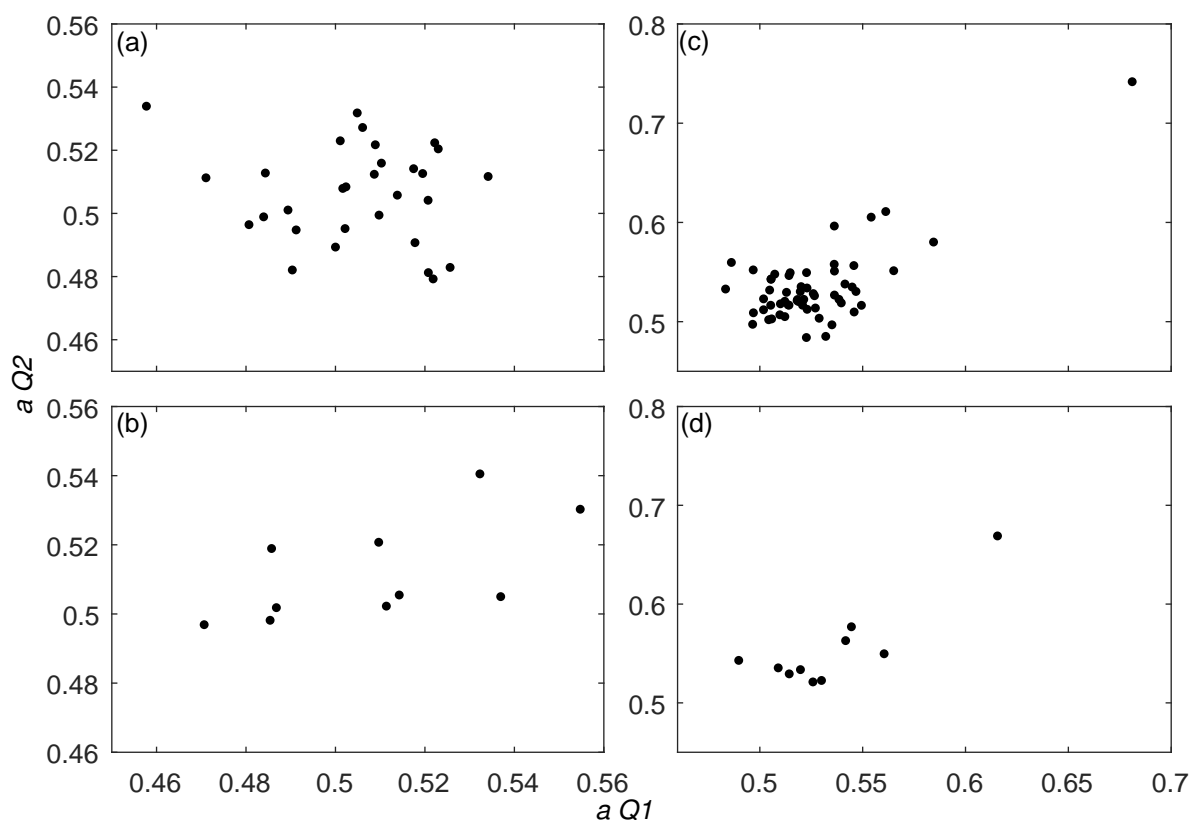


Figura 15 – **Acurácias obtidas no primeiro e segundo quadrimestres de 2017.** Os resultados são referentes às acurácias do modelo para o primeiro e segundo quadrimestres de 2017 aplicado a todas as ações que compõem os índices Dow Jones e Ibovespa e também para as 10 ações com maiores pesos nestes índices. Os dados analisados são referentes às séries de preços de fechamento com granularidades de 5 minutos, no período de 02/01/17 a 31/08/17 (primeiro e segundo quadrimestres de 2017). O tamanho do ajuste utilizado foi $R = 1.000$. (a) Representa as acurácias para a combinação de todas ações que compõem o índice Dow Jones no primeiro e segundo quadrimestres. (b) mostra as acurácias obtidas no primeiro e segundo quadrimestres para a combinação das 10 ações do índice Dow Jones com maiores participações no índice. (c) Representa as acurácias obtidas no primeiro e segundo quadrimestres para a combinação de todas as ações que compõem o índice Ibovespa. (d) Mostra as acurácias do primeiro e segundo quadrimestres para a combinação das 10 ações do Ibovespa com maiores pesos no índice.

Analisando os pontos obtidos na figura 15, não podemos concluir que a combinação das 10 ações com maiores pesos nos índices melhora as acurácias obtidas quando comparadas com as taxas obtidas combinando todas as ações que compõem os índices.

6.5.3 Análise do desempenho do modelo aplicado às ações de melhor acurácia no quadrimestre precedente

A tabela 6 mostra as acurácias obtidas para todas as ações que compõem os índices Ibovespa e Dow Jones. As acurácias obtidas são para $R = 1.000$, e estão ordenadas do maior para a menor valor. Destacamos os 10 melhores resultados de a . As cores semelhantes mostram as ações que participaram do conjunto com os melhores resultados no quadrimestre 1 e também apareceram entre as 10 ações com os melhores resultados no quadrimestre 2. Os valores em negrito indicam que a ação apresentou maior acurácia no quadrimestre 2 em relação ao quadrimestre 1. Pela tabela 6 e figura 14 podemos observar que as 10 melhores acurácias obtidas conjuntamente em ambos os quadrimestres não são necessariamente as ações com maiores participações nos índices, exceto a ação 10 do Ibovespa. Isso nos motiva a investigar se obtemos melhores acurácias combinando no modelo, a cada quadrimestre, apenas as 10 ações que apresentaram melhores resultados no quadrimestre anterior.

Os resultados da análise são mostrados na tabela 7, e comparados a resultados de acurácias obtidas também no mesmo quadrimestre, porém utilizando dados de todas as ações dos respectivos índices ($n = 30$ Dow Jones e $n = 58$ Ibovespa). Os valores em negrito indicam que a combinação das 10 ações é melhor quando comparada às acurácias obtidas utilizando todas as ações dos índices. Utilizando o primeiro quadrimestre para selecionar ações a investir no segundo quadrimestre não utilizamos dados futuros na análise. Para as ações do Ibovespa, a média das acurácias obtida para a combinação de todas as ações foi 0,5737 e para a combinação das 10 melhores ações do quadrimestre anterior foi 0,5746. A média para o índice Dow Jones foi 0,5019 para a combinação de todas as ações e 0,5017 para a combinação das 10 melhores ações do quadrimestre anterior. Estes valores se referem à média das acurácias das 10 ações analisadas. Pela proporção dos valores obtida não podemos concluir que há ganhos significativos aplicando esta metodologia.

Tabela 6 – Comparação das acurácias para todas as ações dos índices Ibovespa e Dow Jones no primeiro e segundo quadrimestres.

Ibovespa				Dow Jones			
Ação	Acurácia Q1	Ação	Acurácia Q2	Ação	Acurácia Q1	Ação	Acurácia Q2
32	0,6810	32	0,7417	7	0,5341	2	0,5339
52	0,5844	57	0,6109	24	0,5256	20	0,5318
56	0,5650	10	0,6053	28	0,5229	29	0,5272
57	0,5612	35	0,5963	27	0,5221	8	0,5229
10	0,5541	52	0,5802	17	0,5218	27	0,5223
39	0,5493	11	0,5596	6	0,5207	23	0,5217
46	0,5467	12	0,5578	4	0,5207	28	0,5204
34	0,5458	50	0,5565	3	0,5194	10	0,5159
50	0,5457	37	0,5521	15	0,5177	19	0,5141
15	0,5448	56	0,5513	19	0,5174	21	0,5127
33	0,5413	4	0,5510	16	0,5138	3	0,5126
44	0,5395	51	0,5494	10	0,5102	11	0,5123
49	0,5384	13	0,5493	9	0,5097	7	0,5116
4	0,5362	16	0,5478	23	0,5089	30	0,5112
6	0,5362	29	0,5465	11	0,5086	26	0,5084
35	0,5362	38	0,5430	29	0,5060	5	0,5079
12	0,5361	3	0,5427	20	0,5048	16	0,5057
19	0,5350	33	0,5379	26	0,5023	4	0,5041
17	0,5320	31	0,5353	18	0,5021	13	0,5010
43	0,5288	15	0,5349	5	0,5016	9	0,4994
22	0,5270	55	0,5339	8	0,5010	14	0,4989
20	0,5265	2	0,5329	12	0,5000	22	0,4964
18	0,5259	45	0,5318	1	0,4912	18	0,4951
41	0,5229	46	0,5304	25	0,4903	1	0,4947
55	0,5229	54	0,5303	13	0,4893	15	0,4907
51	0,5227	5	0,5295	21	0,4843	12	0,4893
9	0,5226	18	0,5282	14	0,4839	24	0,4829
53	0,5213	6	0,5268	22	0,4806	25	0,4820
24	0,5208	20	0,5261	30	0,4710	6	0,4812
21	0,5207	42	0,5229	2	0,4577	17	0,4792
31	0,5200	49	0,5225				
54	0,5196	53	0,5224				
8	0,5187	1	0,5223				
1	0,5183	47	0,5212				
47	0,5181	24	0,5210				
13	0,5147	58	0,5205				
29	0,5141	8	0,5203				
26	0,5141	44	0,5189				
28	0,5137	23	0,5180				
5	0,5129	28	0,5168				
48	0,5121	26	0,5166				
58	0,5121	21	0,5166				
23	0,5100	36	0,5165				
14	0,5097	39	0,5164				
16	0,5072	22	0,5137				
25	0,5057	41	0,5124				
38	0,5055	27	0,5120				
3	0,5053	34	0,5097				
36	0,5052	40	0,5089				
45	0,5047	14	0,5069				
30	0,5042	48	0,5050				
42	0,5018	43	0,5034				
27	0,5017	25	0,5027				
40	0,4968	30	0,5019				
37	0,4967	7	0,4973				
7	0,4964	19	0,4968				
11	0,4861	17	0,4851				
2	0,4833	9	0,4840				

Nota: As acurácias estão em ordem decrescente. Os 10 melhores resultados estão delimitados na tabela. As cores equivalentes indicam que a ação tem o valor de acurácia dentre os 10 melhores resultados no quadrimestre 1 e também se manteve entre as 10 melhores no quadrimestre 2. Os valores em negrito indicam que, além da ação estar entre as 10 melhores nos 2 quadrimestres, a acurácia aumentou no segundo quadrimestre.

Tabela 7 – Acurácias do modelo para a combinação das 10 ações dos índices Ibovespa e Dow Jones com os melhores resultados no primeiro quadrimestre.

Ibovespa			Dow Jones		
Ação	Acurácia Q2 <i>n</i> = 58	Acurácia Q2 <i>n</i> = 10	Ação	Acurácia Q2 <i>n</i> = 30	Acurácia Q2 <i>n</i> = 10
32	0,7418	0,7707	7	0,5117	0,4996
52	0,5802	0,5800	24	0,4829	0,5032
56	0,5514	0,5438	28	0,5204	0,5174
57	0,6110	0,6050	27	0,5224	0,5243
10	0,6053	0,5811	17	0,4793	0,4960
39	0,5165	0,5174	6	0,4813	0,4941
46	0,5304	0,5326	4	0,5042	0,4966
34	0,5098	0,5253	3	0,5126	0,4842
50	0,5565	0,5545	15	0,4907	0,5034
15	0,5349	0,5363	19	0,5142	0,4986

Nota: Para Ibovespa, na coluna com *n* = 58, todas as ações do índice foram utilizadas pelo modelo; na coluna com *n* = 10, somente as ações com as 10 melhores acurácias no quadrimestre 1 foram empregadas pelo modelo. Analogamente para Dow Jones. Os valores em negrito indicam que a acurácia aumentou no segundo quadrimestre utilizando a combinação das 10 melhores ações no quadrimestre 1 em relação ao segundo quadrimestre utilizando a combinação de todas as ações dos índices. Não observamos ganhos nos resultados quando utilizamos apenas a combinação das 10 melhores ações em relação ao modelo que utiliza a combinação de todas as ações dos índices.

Diante das acurácias obtidas na tabela 6, investigamos se há uma correlação entre estes valores de a . Para isso, calculamos o coeficiente de correlação linear de Pearson ρ . O resultado do cálculo do coeficiente de correlação entre as acurácias das ações do índice Ibovespa obtidas no primeiro e segundo quadrimestre foi $\rho = 0,7156$ ($p - valor = 2,7521 \times 10^{-10}$) e para as ações do índice Dow Jones foi $\rho = -0,1493$. Este resultado indica que, para as ações do Ibovespa, quando o modelo teve um bom resultado no quadrimestre ele tende a ter bons resultados no quadrimestre seguinte. Como esperado, o mesmo não acontece para as ações do índice Dow Jones, em que o coeficiente de correlação $\rho = -0,1493$ ($p - valor = 0,4311$) indica uma pequena anti-correlação dos acertos de tendências. A ação 32 do índice Ibovespa apresentou uma acurácia alta em ambos quadrimestres. Assim, para evitar resultados tendenciosos, ela foi retirada do cálculo da correlação. Com a retirada destes valores o coeficiente de correlação das acurácias das ações do índice Ibovespa foi menor, $\rho = 0,3894$ ($p - valor = 0,0028$), mas ainda significativa (COHEN, 1977).

6.6 Votação

Uma votação entre tendências previstas foi utilizada com o intuito de melhorar a consistência da acurácia do modelo e aumentar o tamanho da amostra. A metodologia da técnica consiste em fazer várias previsões utilizando diferentes tamanhos de R e, posteriormente, fazer a votação, levando em consideração a direção de tendência de preços apontada por cada modelo e o tamanho em módulo do retorno previsto. Apenas os ajustes que têm previsões no conjunto U_4 , que se referem aos maiores retornos previstos em módulo, entram na votação. Nesta votação, consideramos como a decisão de tendência prevista a direção apontada pela maioria dos votantes e empate entre os votantes como sinal para manter a posição no tempo $t + 1$ (não operar).

Podemos variar diversos parâmetros de entrada na votação, tais como: quantidade de votantes, tamanho inicial de R , incremento da quantidade de pontos a cada nova iteração, ações que participam da combinação linear e quantidade de votos que decidem a previsão. Nas figuras 16 e 17, mostramos os resultados das acurácias obtidas para todas as ações dos índices Dow Jones e Ibovespa no primeiro e segundo quadrimestres de 2017. Nesta análise, utilizamos 10 votantes, um valor inicial de $R = 1.000$, incremento de 100 pontos a cada nova execução e a diferença acima de 1 voto já decide direção da tendência prevista que será considerada.

A figura 16 mostra os resultados obtidos para a amostra com os dados referentes ao primeiro quadrimestre de 2017. A curva na cor vermelha é referente às acurácias a obtidas sem a votação, utilizando apenas um valor de $R = 1.000$. A curva na cor azul refere-se à média das acurácias \bar{a} obtidas utilizando 10 diferentes tamanhos de R ,

com o valor inicial de $R = 1.000$ com incremento de 100 pontos a cada nova execução do modelo. A curva na cor verde mostra a acurácia a obtida com a votação entre os 10 diferentes tamanhos de R .

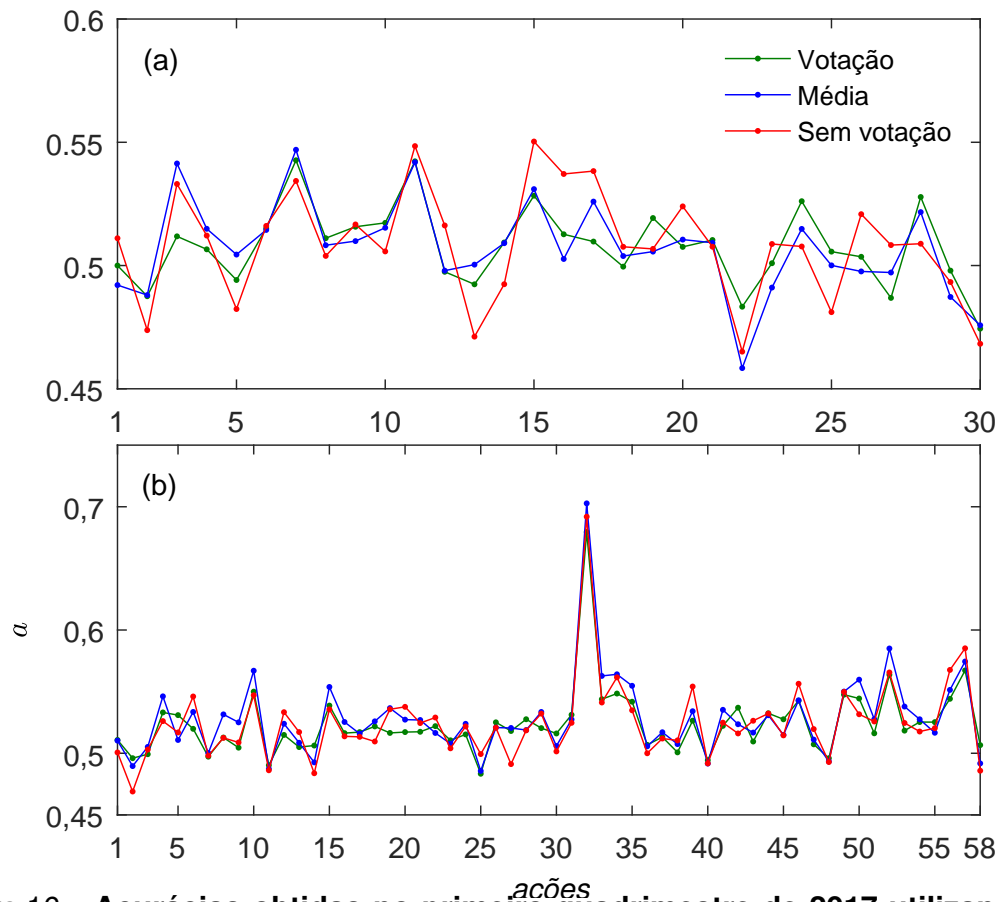


Figura 16 – **Acurácias obtidas no primeiro quadrimestre de 2017 utilizando *Voting*, média e sem votação.** A curva na cor vermelha representa a acurácia a utilizando apenas um valor de $R = 1.000$. A curva na cor verde mostra a acurácia a utilizando a votação de 10 diferentes valores de R . O valor inicial é $R = 1.000$, com incremento de 100 pontos a cada nova execução. A curva azul define a média da acurácia \bar{a} de todos os 10 tamanhos de R utilizados na votação. Os dados analisados são referentes às séries de preços de fechamento com granularidades de 5 minutos, no período de 02/01/17 a 28/04/17 (primeiro quadrimestre de 2017). (a) Acurácias de todas as ações que compõem o índice Dow Jones. Foram realizadas 2.200 execuções para cada tamanho de R . (b) Acurácias obtidas para todas as ações que compõem o índice Ibovespa. Para cada tamanho de R , realizamos 3.300 execuções.

A figura 17 mostra os resultados obtidos para a amostra composta pelos dados referentes ao segundo quadrimestre de 2017. Análoga à figura 16, os parâmetros utilizados são os mesmos estabelecidos anteriormente.

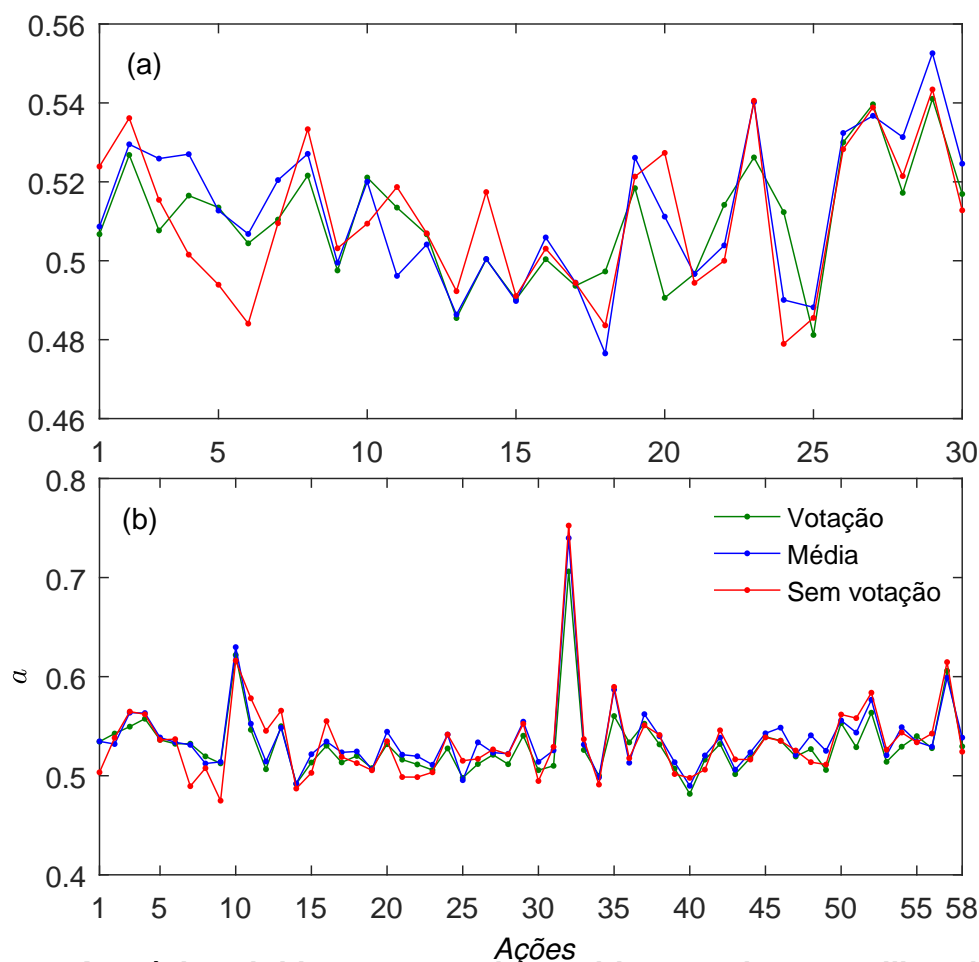


Figura 17 – **Acurácias obtidas no segundo quadrimestre de 2017 utilizando votação, média e sem votação.** A curva na cor verde mostra a acurácia a utilizando a votação entre os 10 diferentes valores de R . O valor inicial é $R = 1.000$, com incremento de 100 pontos a cada nova execução. A curva azul representa a média da acurácia \bar{a} para todos os tamanhos de R utilizados na votação. A curva em vermelho descreve a acurácia a obtida para um valor de $R = 1.000$. Os dados analisados são referentes às séries de preços de fechamento com granularidades de 5 minutos, no período de 02/05/17 a 31/08/17 (segundo quadrimestre de 2017). (a) Acurácias de todas as ações que compõem o índice Dow Jones. Foram realizadas 2.700 execuções para cada tamanho de R . (b) Representam as acurácias obtidas para todas as ações que compõem o índice Ibovespa. Para cada valor de R foram realizadas 3.300 execuções.

Analisando as figuras 16 e 17, visualmente não percebemos ganhos de acurácia a quando comparada à técnica de votação com os outros métodos aplicados. Como mencionado anteriormente, neste trabalho utilizamos apenas dados com maiores retornos previstos em módulo, que se referem às tendências previstas que estão no conjunto do quarto quartil U_4 . Assim, a vantagem da votação em relação à utilização de médias ou de um valor fixo de R é a possibilidade de aumento na quantidade de previsões, uma vez que sem a votação utilizamos, em média, apenas $\frac{1}{4}$ dos dados para as previsões no quarto quartil. Por outro lado, a presença de um ou mais votantes em U_4 já é suficiente para que a votação ocorra, levando o número de previsões a crescer substancialmente.

6.7 Considerações gerais

Na seção 6.1, investigamos a relação da variação do tamanho de R no resultado final da acurácia a . Nesta análise, observamos que para valores maiores de R os valores de a tendem a aumentar. Porém, utilizar valores grandes de R está diretamente ligado a perda de dados da amostra. Assim, é preciso definir uma quantidade de pontos para o ajuste que apresente uma boa acurácia sem comprometer o volume de dados. Baseados nas figuras 7 e 6, definimos que nas etapas de aplicação e simulação a quantidade de pontos utilizados no ajuste será $R = 1.000$.

Na seção 6.2, investigamos se a quantidade de ações interfere na acurácia do modelo. A figura 10 sugere uma moderada melhora nos resultados das acurácias em torno da combinação de 10 ações.

Na seção 6.3, analisamos a relação da acurácia do modelo com a granularidade dos dados. Para as ações do índice Dow Jones, não notamos diferenças quando combinamos todas as ações que compõem o índice. Um leve ganho no resultado foi observado quando combinamos as 10 ações com maiores participações no índice. Para as ações que compõem a carteira teórica do Ibovespa, a melhor acurácia obtida foi para os dados com frequência de cotação de 5 minutos quando combinadas as 10 ações com maiores participações no índice. Desta forma, a amostra com granularidade de 5 minutos é a que será utilizada em aplicações futuras do modelo de previsão.

Na seção 6.4, investigamos se há uma relação entre as acurácias obtidas e os expoentes de Hurst das séries analisadas. Em geral, não encontramos padrões relacionados às acurácias e os expoentes H . Um caso atípico foi a acurácia e o expoente de Hurst relacionado à ação 32 do índice Ibovespa. Para esta ação o expoente foi $H = 0,2816$ e a taxa de acurácia $a = 68,10$. Este resultado sugere que para a série com anti-persistência o modelo obtém melhores taxas de acerto. Os

outros expoentes ficaram na faixa de variabilidade de $0,45 \leq H \leq 0,55$. Com isso, não podemos afirmar que há um padrão relacionando as acurácias e os expoentes de Hurst das séries analisadas.

Na seção 6.5, investigamos se as taxas de acurácia calculadas em um quadrimestre se mantêm no quadrimestre seguinte. Nas análises, investigamos qual o número das ações que apareceram entre as 10 melhores acurácias no quadrimestre 1 e também no quadrimestre 2. Na figura 14, observamos que para as ações do índice Dow Jones o número de ações que estavam presentes entre as 10 que apresentaram melhores acurácias nos quadrimestres 1 e 2 foi 3. Este resultado sugere a completa aleatoriedade. Já para as ações que compõem o Ibovespa, a figura 14 mostra 6 ações nessa condição, indicando que o processo não deve ser completamente aleatório.

Ainda na seção 6.5, outra análise foi baseada na quantidade e em quais ações que devemos combinar para aplicar o modelo. Baseados nas médias das acurácias obtidas, não encontramos grandes variações nos valores quando utilizamos no modelo todas as ações, quando combinamos as 10 ações com maiores pesos e, ainda, quando utilizamos a combinação das 10 ações que apresentaram as melhores acurácias no quadrimestre anterior. Em geral, as análises desta seção nos mostraram que o modelo tende a manter as taxas de acurácia quando utilizamos diferentes ações e quando variamos a sua quantidade. As taxas também se mantêm para diferentes quadrimestres, indicando consistência das previsões.

Na seção 6.6, desenvolvemos uma técnica de votação em relação à direção da tendência prevista. Nessa análise, 10 diferentes valores de R participaram da votação. Em geral, os resultados obtidos com a técnica não foram melhores quando comparados com a média das acurácias obtida utilizando os 10 valores de R . Em média, os resultados da votação também se equipararam aos resultados obtidos utilizando um valor único de $R = 1.000$. A vantagem do uso da técnica está na possibilidade de aumento do tamanho da amostra.

Na próxima etapa do trabalho, vamos fazer simulações do modelo combinando os melhores parâmetros explorados neste capítulo.

Capítulo 7

Simulação

No Capítulo 6, exploramos algumas das possibilidades de combinações de parâmetros que podem ser utilizadas para ajustar o modelo de previsão de forma a obter uma melhor acurácia. Na seção 6.1, investigamos a relação entre a variação do tamanho de R e o resultado final da acurácia para definirmos um valor razoável para ser utilizado na etapa de aplicação do modelo. Na seção 6.2, investigamos o quanto o número de ações utilizadas como entrada para o modelo impactava as acurácias. Na seção 6.3, analisamos a relação da acurácia do modelo com a granularidade dos dados e percebemos uma moderada melhora nos resultados para os dados com frequência de cotação de 5 minutos. Na seção 6.4, investigamos se há uma relação entre as acurácias obtidas e os expoentes de Hurst das séries. Diante destas possibilidades, ajustamos o modelo de previsão utilizando as combinações de parâmetros mais adequadas a cada análise. Não podemos afirmar que a combinação escolhida é a combinação ótima, porém, no Capítulo 6 foram os parâmetros que apresentaram as melhores acurácias. Assim, nesta etapa, estas são as entradas do modelo que serão utilizadas na simulação.

Neste capítulo, vamos mostrar as estratégias de negociação baseadas no modelo que retorna um sinal de compra, venda ou não opera apresentado no capítulo 5. Para calibrar o modelo de previsão, utilizamos os parâmetros explorados no Capítulo 6. Os critérios para definir uma compra ou venda dependem da tendência prevista do preço de fechamento da ação. Os dados utilizados para a execução do modelo de previsão para gerar o sinal da operação são os preços tratados $\mathbf{X}(t)$, definidos na seção 5.1. Na etapa de aplicação e análise de ganhos financeiros, vamos utilizar os preços brutos de fechamento $\mathbf{x}(t)$, definidos como

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix}, \quad (39)$$

onde $x_k(t)$ é o preço bruto da ação k no tempo t .

Na próxima seção, vamos descrever o fluxo das estratégias, as bases de dados e as configurações aplicadas nas simulações. Nas seções 7.2 e 7.3, vamos apresentar as características de cada uma das estratégias, bem como os seus resultados obtidos com a simulação.

7.1 Visão geral dos modelos de operação

Para a simulação desenvolvemos duas estratégias com características diferentes: uma que, de acordo com o sinal enviado pelo modelo, executa a operação e desfaz no passo seguinte (modelo de operação 1 - MO1) e outra que executa a operação e desfaz somente quando atingir os valores dos stops definidos (modelo de operação 2 - MO2). Nesta seção, vamos apresentar o fluxo geral das estratégias e nas seções 7.2 e 7.3 cada uma delas será apresentada de forma detalhada.

A simulação da negociação no mercado real foi baseada em backtest¹ disponibilizado pelo portal dedicado ao desenvolvimento de algoritmos de investimento (algotrading) *Stratsphera BETA*. Na plataforma de algotrading as estratégias foram desenvolvidas utilizando a linguagem de programação Python, numa estrutura baseada em eventos. A plataforma foi descontinuada em Julho de 2019.

Durante a simulação, o modelo de previsão utiliza os candles disponíveis até o tempo t para gerar um sinal indicando a tendência do preço de fechamento no tempo $t + 1$. Baseados neste sinal, podemos realizar operações de compra ou venda a mercado. De posse deste sinal, que utiliza o último preço de fechamento do candle, o ideal seria executar imediatamente a ordem na abertura do próximo candle, porém, por limitações da plataforma de simulação, não conseguimos realizar as operações exatamente no preço de abertura, mas sim em um preço aproximado de abertura

$$q_o = p_o + (p_c - p_o)\epsilon_1, \quad (40)$$

onde p_o é o preço de abertura do candle, p_c é o preço de fechamento do candle e ϵ_1 foi escolhido $\epsilon_1 = \frac{1}{60}$.

Para a estratégia que utiliza fechamento de candle para zerar a posição (MO1), consideramos a aproximação do preço de fechamento

$$q_c = p_o + (p_c - p_o)\epsilon_2, \quad (41)$$

com $\epsilon_2 = \frac{14}{15}$. Valores menores de ϵ_1 e maiores de ϵ_2 levavam a operações que não aconteciam dentro do candle esperado. Nas etapas seguintes do trabalho, quando

¹Execução de uma estratégia com uso de dados históricos.

referirmos a preços de abertura e fechamento de candle utilizados em operações, consideraremos estas aproximações aqui definidas.

Analogamente aos capítulos anteriores, somente retornos previstos que estão no quarto quartil (ver seção 5.3) são considerados para a operação. Deste modo, se o preço de fechamento previsto no tempo $t + 1$ for maior (menor) que o preço real no tempo t , a tendência prevista é de alta (baixa). Assim, durante a simulação, a cada novo evento (que corresponde a um novo candle), analisamos se o retorno previsto, para cada valor de R , está no quarto quartil e, se esta condição for satisfeita, o sinal correspondente é considerado em uma votação (ver seção 6.6). De acordo com o sinal da votação, realizamos uma nova operação de compra (*long*) ou venda descoberta (*short*) a mercado. As posições de cada uma das operações são fechadas de forma independente.

Para fechar as posições de compra ou venda descoberta de cada operação, definimos dois modelos de operação: o modelo de operação 1 (MO1) e o modelo de operação 2 (MO2). O MO1 sempre desfaz a operação no candle seguinte através de uma contra ordem. O MO2 desfaz as posições baseado em stops. As características de cada uma das estratégias serão detalhadas nas seções 7.2 e 7.3.

Nas simulações, utilizamos 3 tamanhos diferentes de R . Para cada valor de R , o modelo gera um sinal de tendência prevista, sendo possível realizar a votação entre estes sinais gerados. Nas execuções seguintes, a cada chegada de um novo candle, a janela é atualizada com o novo valor e descarta-se o dado mais antigo. Os dados utilizados para a simulação das estratégias foram preços de abertura e fechamento das 10 ações com maior liquidez na Bovespa nos anos de 2017 e 2018. Informações de volume também foram utilizadas, para análise dos experimentos. Esta redução no número de ações analisadas é devida a limitações da plataforma em relação ao custo de execução e também pelo fato de diminuir execuções parciais devido ao volume reduzido de candles de ações com menor liquidez. A frequência dos candles utilizados é de 5 minutos. Os tamanhos das janelas de dados utilizadas foram $R = 500$, $R = 1000$ e $R = 1500$.

Diante de cada novo candle disponível, o modelo utiliza a janela de dados para tratar o preço, conforme a seção 5.1, e calcular o novo sinal naquele instante para cada valor de R . Uma votação é realizada entre os sinais retornados para cada R . Conforme o resultado obtido, a ordem é enviada pela plataforma de acordo com o sinal de compra, venda ou não opera retornado pelo modelo (ver seção 6.6). O modelo prevê o preço de fechamento no tempo $t + 1$, por isso, o ideal seria realizar a operação no mesmo preço que foi utilizado para fazer a previsão, porém, não é possível executar a ordem neste preço de fechamento $p_c(t)$ ². Desta forma, configuramos as funções agendadas

² $p_c(t)$ coincide com os preços brutos $x(t)$, definidos em 39

da plataforma para que o preço de execução fique o mais próximo possível do preço de abertura do candle seguinte $p_o(t + 1)$. Como discutido anteriormente, a plataforma não permite a negociação exatamente com esse preço, sendo realizada no preço $q_0(t + 1)$, conforme a equação 40. Ou seja, utilizamos o preço de fechamento no tempo t para gerar o sinal e realizamos a operação num preço próximo ao de abertura $p_o(t + 1)$ do candle seguinte.

Para uma dada ação, a estratégia só recebe um sinal de compra/venda quando o modelo prevê retornos que estão no quarto quartil (veja seção 5.3). Em caso contrário, o modelo envia um sinal neutro para que não execute operações naquele momento. Para verificar se os retornos previstos estão no quarto quartil, utilizamos uma janela de tamanho $Q = 100$ pontos. Desta forma, a primeira operação só poderá ocorrer a partir do $(Q + 1)$ -ésimo sinal produzido. Antes disso, o modelo sempre retornará sinais neutros. Dado o sinal não neutro, uma ordem de compra ou venda descoberta é submetida. A venda descoberta é feita quando o modelo retorna um sinal indicando que o preço poderá cair. Neste caso, acreditando que o preço da ação irá baixar, vendemos a ação a descoberto, e, de acordo com o critério de cada estratégia, compramos a ação para repor a venda em tempos seguintes. A sequência da simulação é apresentada na figura 18.

Para padronizar, utilizamos os mesmos parâmetros na configuração da plataforma para as duas estratégias. O capital inicial utilizado em ambas foi de R\$1.000.000,00. A quantidade de cada uma das ações negociadas a cada ordem foi 10 vezes o tamanho do lote padrão. Na plataforma, as ações só são negociadas em lotes padrão, não sendo permitidas negociações no mercado fracionário³. A quantidade de ações no lote padrão definido na plataforma, em alguns casos, difere da quantidade definida no mercado real, conforme apresentado nas tabelas 24 e 25, disponíveis no apêndice B. Como exemplo, temos a ação do Bradesco (BBDC4), que tem na composição de seu lote padrão 145 ações, enquanto no mercado real o lote padrão desta ação é composto por 100 ações.

As configurações de custos são utilizadas para incluir custos transacionais durante a execução da estratégia. As configurações de custo aplicadas na simulação foram baseadas no cenário brasileiro. Aplicamos as taxas cobradas pela Bovespa, conforme a tabela 8, e consideramos as operações através de corretoras que oferecem taxas zero para negociações no mercado à vista. Os custos transacionais foram aplicados de acordo com o volume financeiro negociado. Para comparação, as simulações foram realizadas com taxas e sem taxas.

No mercado real, é possível que as ordens lançadas sejam executadas parcialmente. Isso acontece devido ao volume da ordem e à liquidez do ativo. Na simulação aqui

³No mercado fracionário são negociadas frações do lote padrão das ações.

Tabela 8 – Taxas de negociação e de liquidação da BOVESPA

Volume daytrade (R\$ milhões)		Negociação	Liquidação	Total
Pessoa física	Pessoa jurídica			
Ate 4	Até 20	0,004245%	0,020000%	0,024245%
De 4 até 12,5	De 20 até 50	0,003000%	0,020000%	0,023000%
De 12,5 até 25	De 50 até 250	0,000500%	0,019500%	0,020000%
De 25 até 50	De 250 até 500	0,000500%	0,017500%	0,018000%
Mais de 50	Mais de 500	0,000500%	0,015500%	0,016000%

Nota: Os valores são para operações de *day trade* e foram obtidos no site da BOVESPA, disponível em: http://www.b3.com.br/pt_br/produtos-e-servicos/tarifas/listados-a-vista-e-derivativos/renda-variavel/tarifas-de-acoes-e-fundos-de-investimento/a-vista/

descrita, configuramos para que o volume da ordem lançada seja executado totalmente na simulação. Porém, em alguns casos, em ações com menor liquidez, o volume do candle disponível será menor que o tamanho da ordem submetida e, neste caso, a ordem será executada parcialmente até o limite máximo do candle disponível. Por isso, para evitarmos estas execuções parciais, utilizamos as ações com maior liquidez na Bovespa. De fato, as ordens que submetemos têm volume bem menor do que o dos candles. Portanto, acreditamos que a aproximação que usamos (volume total da ordem negociado) não leva a importantes distorções.

A plataforma permite alocar o capital parado durante a estratégia em uma aplicação, tal como CDI⁴. Nas simulações, definimos que o capital parado não será remunerado. Com isso, evitamos interferência, em nossos resultados, de remunerações não advindas da estratégia.

Nas próximas seções, vamos apresentar as duas estratégias diferentes. No modelo de operação 1 (MO1), vamos abrir a posição de compra ou venda no instante de tempo analisado e fechar a posição no instante seguinte. Já no modelo de operação 2 (MO2), uma ordem de compra ou venda é enviada e executada e a posição referente a esta ordem só será zerada quando o stop definido for atingido.

Para ilustrar as etapas das estratégias, desenvolvemos a figura 18, que apresenta as principais fases da execução. Na etapa de configuração dos parâmetros, definimos os dados que serão utilizados na simulação, tais como a frequência dos candles e as ações escolhidas. Nesta etapa, também configuramos os custos operacionais, capital que será alocado para a execução e as variáveis de controle de risco da estratégia, dentre outras configurações possíveis.

Na coleta inicial de dados, uma janela de dados históricos de tamanho R é armazenada para o cálculo da primeira matriz inicial. Estes dados são enviados para o

⁴Certificados de Depósitos Interbancários são títulos emitidos para a realização de operações de empréstimo ou captação de recursos financeiros entre os bancos.

modelo de previsão (MP), onde são tratados. Com a matriz de dados tratados X é possível executar o modelo e gerar um sinal de previsão. Este sinal é aplicado de acordo com cada estratégia. No modelo de operação MO1, ao receber o sinal, uma posição de compra ou venda é aberta, o *status* da ordem é atualizado e a ordem anterior, se existir posição aberta, é zerada através de uma contra ordem. Já na estratégia MO2, os stops são calculados e a operação é aberta. Nesta estratégia, as operações são zeradas quando atingem os valores de stop loss ou stop gain. Todo o procedimento é repetido a cada chegada de um novo candle.

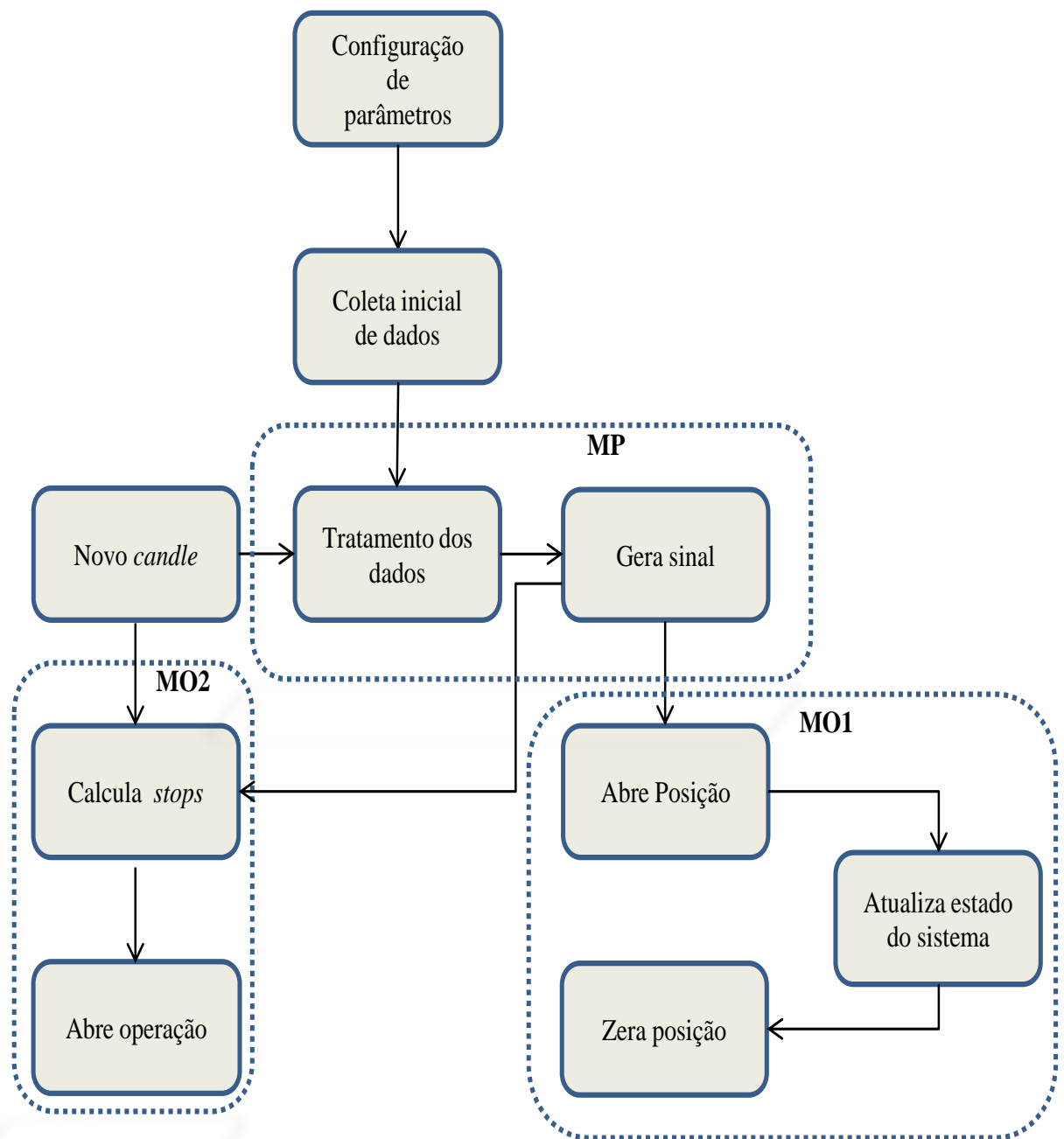


Figura 18 – **Fluxograma das estratégias.** Principais fases de execução do modelo de operação 1 (MO1) e do modelo de operação 2 (MO2). Na etapa de configuração de parâmetros são definidas as características gerais das simulações. Na etapa de coleta inicial de dados, uma sequência de tamanho R de dados históricos das ações é armazenada para o cálculo da primeira matriz. No modelo de previsão (MP), os dados são tratados e enviados para que o modelo gere um sinal de previsão de tendência. Para MO1, dado o sinal de previsão não neutro, uma posição de compra/venda é aberta, o estado do sistema é atualizado e a posição anterior é zerada. Já em MO2, o valor do stop é calculado e a operação é aberta, associada aos valores de stop gain e stop loss. A cada novo dado disponível (novo candle), o modelo de previsão é computado e todo o processo é repetido.

Para quantificar o desempenho das estratégias, utilizamos os indicadores calculados pela plataforma durante a execução das simulações. A descrição completa de cada uma destes indicadores foi apresentada na seção 4.1.1. Os principais indicadores utilizados foram:

Retorno acumulado, que mostra o retorno obtido com a estratégia ao final da execução. É dado pela razão entre o capital final possuído e o inicial investido.

Volatilidade é o indicador retornado pela plataforma que representa a volatilidade da estratégia anualizada considerando os retornos diários da mesma.

Sharpe é o índice que indica a relação do risco e retorno de um investimento. Quanto maior o Sharpe, maior a expectativa de retorno frente ao risco associado. Utiliza como benchmark um investimento livre de risco. Na simulação, utilizamos como benchmark o CDI.

Sortino é semelhante ao índice Sharpe e avalia a relação entre risco e retorno de uma estratégia, porém, leva em consideração apenas a volatilidade dos retornos negativos.

Max drawdown mostra a maior perda que a estratégia obteve durante a execução.

Beta é o indicador que mostra a correlação entre a estratégia e um benchmark, medindo o quanto o crescimento/decrescimento da estratégia segue o benchmark.

Alpha é o parâmetro que mostra o quanto a estratégia se destacou positivamente dos retornos retratados pelo benchmark.

Retorno anual esperado é o indicador que estima o retorno que a estratégia alcançará no período de um ano, mantendo a mesma performance.

Estabilidade mostra a estabilidade da estratégia baseada nos retornos acumulados.

PNL indica o ganho ou perda financeira ao final da execução da estratégia.

VAR é utilizado para calcular o risco da estratégia. Na plataforma, este parâmetro estima o pior caso esperado de retorno diário da simulação.

Taxas representa os custos operacionais da estratégia, tais como corretagem, emolumentos e outras tarifas de serviços aplicadas ao mercado acionário brasileiro. Taxas de aluguel não foram consideradas, uma vez que a estratégia é desenvolvida no *intraday*⁵ e as posições são zeradas no mesmo dia.

7.2 Modelo de operação 1

A estratégia MO1 consiste em comprar ou vender a descoberto conforme o sinal retornado pelo modelo e desfazer (zerar) a posição no passo seguinte de acordo com

⁵Operações realizadas dentro de um mesmo dia

a frequência do candle analisado. Para esta estratégia, desenvolvemos três funções de agendamento⁶: uma para manipulação dos dados e execução do modelo, outra para desfazer as posições abertas no mercado, e por fim, uma função de agendamento para emitir ordens de posicionamento (compra ou venda) de acordo com o sinal gerado pelo modelo.

Neste modelo de operação, além da simulação, é possível também calcular as acurácias da forma que foram apresentadas no capítulo 6. Deste modo, na próxima subseção, vamos apresentar as acurácias obtidas para os mesmos dados e períodos utilizados na simulação do modelo.

7.2.1 Acurácias

As acurácias obtidas com o modelo apresentadas no capítulo 6 são referentes a preços de fechamento dos candles. Os resultados que lá apresentamos utilizam a matriz de preços até o tempo t para gerar um sinal e encontrar os preços previstos $\hat{X}(t + 1)$. Consideramos para o cálculo de tendência prevista a diferença do preço previsto $\hat{X}(t + 1)$ e o preço real $X(t)$. Aqui, analisamos a acurácia de duas formas: uma é idêntica à anterior, comparando a tendência prevista àquela observada através da diferença entre os preços de fechamento $X(t)$ e $X(t + 1)$, chamaremos esta de *acurácia canônica*. A outra é de acordo com a simulação, mais realística, em que, de posse do preço de fechamento do candle $X(t)$ é possível calcular um sinal utilizando este valor e imediatamente lançar uma ordem de acordo com o sinal, porém, o preço de execução será $q_o(t + 1)$, que é um valor próximo da abertura do candle seguinte. Desta forma, a acurácia deverá ser calculada comparando-se a tendência prevista com o movimento observado entre $q_o(t + 1)$ e $X(t + 1)$. Estes resultados serão denominados *acurácias efetivas*.

Nas tabelas 9 e 10 mostramos as acurácias obtidas para $R = 500$, $R = 1000$, $R = 1500$ e também para a votação unânime entre estes valores de R para o ano de 2017 e 2018, respectivamente. Esses valores são referentes a acurácias canônicas.

Os valores apresentados nas tabelas 11 e 12 são referentes às acurácias efetivas obtidas nos anos de 2017 e 2018, respectivamente. Como na simulação anterior, aqui também utilizamos os seguintes valores de R como parâmetros de entrada para o modelo: $R = 1500$, $R = 1000$ e $R = 500$. Também apresentamos a votação unânime entre esses valores de R .

Analisando os resultados, percebemos que as acurácias efetivas geralmente são menores quando comparadas às acurácias canônicas.

⁶Funções programadas para serem executadas em intervalos pré-determinados a cada novo evento.

Tabela 9 – Acurácias canônicas calculadas para as 10 ações mais líquidas do Ibovespa no ano de 2017

	R=500	R=1000	R=1500	Votação
BBDC4	0,55176541	0,54323459	0,53798481	0,5496368
VALE5	0,53415179	0,54329485	0,53480018	0,5384045
CIEL3	0,53849399	0,54848801	0,54806465	0,5534425
VALE3	0,52433407	0,52550201	0,51997615	0,5267777
ITUB4	0,55874726	0,5566846	0,54563297	0,5615032
PETR4	0,58868605	0,57385475	0,57413249	0,5831128
ABEV3	0,56385379	0,55507608	0,55918367	0,5668191
BOVA11	0,54358257	0,54148385	0,53963893	0,5485689
BBAS3	0,53731644	0,52841136	0,53358886	0,5343546
PETR3	0,54584775	0,54169374	0,53303011	0,5469478

Tabela 10 – Acurácias canônicas calculadas para as 10 ações mais líquidas do Ibovespa no ano de 2018

	R=500	R=1000	R=1500	Votação
BBDC4	0,5248441	0,5222335	0,5313253	0,5270455
ITUB4	0,53228346	0,53066878	0,52536956	0,5364042
ITUB3	0,7219697	0,71792261	0,69898219	0,7209546
ABEV3	0,52848723	0,52844997	0,52051169	0,5241737
PETR4	0,52412925	0,50901468	0,51598078	0,5126012
BBAS3	0,52238209	0,52155772	0,53148111	0,5242525
BOVA11	0,54383243	0,53680922	0,53006939	0,5438518
B3SA3	0,54016106	0,53573623	0,53500714	0,5414768
VALE3	0,49833301	0,49735864	0,49960427	0,4964061
MGLU3	0,52293755	0,52034095	0,51935298	0,5264673

Tabela 11 – Acurácias efetivas calculadas para as 10 ações mais líquidas do Ibovespa no ano de 2018

	R=500	R=1000	R=1500	Votação
BBDC4	0,5011914	0,5024945	0,5049741	0,5296443
VALE5	0,5531525	0,5955037	0,6322361	0,5145472
CIEL3	0,4749896	0,4713191	0,4787458	0,5386266
VALE3	0,4914488	0,5031373	0,4957248	0,5077800
ITUB4	0,4987131	0,5083433	0,5004837	0,5428696
PETR4	0,4947859	0,4978012	0,4975523	0,5332230
ABEV3	0,4814815	0,4783004	0,4835861	0,5285027
BOVA11	0,5052568	0,5092074	0,5058312	0,5167801
BBAS3	0,4971075	0,4975282	0,4996000	0,5235765
PETR3	0,4963441	0,5061648	0,5031515	0,5144113

Tabela 12 – Acurácias efetivas calculadas para as 10 ações mais líquidas do Ibovespa no ano de 2018

	R=500	R=1000	R=1500	Votação
BBDC4	0,5101960	0,5039117	0,5081028	0,5101540
ITUB4	0,5006803	0,5002930	0,5037490	0,5204274
ITUB3	0,5900638	0,5841450	0,5702323	0,5326117
ABEV3	0,4933448	0,4994650	0,4991361	0,5103865
PETR4	0,5015291	0,4919239	0,4924980	0,4957188
BBAS3	0,4973493	0,4922655	0,4955735	0,5166485
BOVA11	0,5018457	0,5005853	0,4940293	0,5096904
B3SA3	0,4980726	0,5030328	0,5012077	0,5233708
VALE3	0,4958454	0,4913661	0,4944735	0,4910852
MGLU3	0,4899507	0,4966571	0,4910714	0,5047889

As acurácias obtidas indicam, na maioria das vezes, que o modelo é diferente de um processo totalmente aleatório. Mas somente as acurácias não indicam retorno financeiro positivo pelo modelo. Outras variáveis devem ser consideradas, tais como: taxas, volatilidade e liquidez. Um exemplo de ação em que o modelo apresentou uma acurácia considerável foi a CPFE3, tanto no capítulo 6 quanto para os dados utilizados neste capítulo para a simulação (anos de 2017 e 2018). Porém, no momento da simulação este ativo apresentou problemas nas execuções das ordens e contra ordens devido à sua baixa liquidez. Quando lançamos a ordem de compra de 1000 ações e em seguida a contra ordem de venda da mesma quantidade, em alguns casos, as ordens não foram totalmente executadas (execução parcial). Percebemos que isso acontece quando o volume do candle naquele instante é menor que o tamanho definido da ordem. Este resultado sugere que apenas as acurácias não são as melhores variáveis para representar o sucesso do modelo. Informações como liquidez, volume de negociação e volatilidade do ativo também devem ser levadas em consideração. Deste modo, para evitarmos ordens parcialmente executadas, utilizamos na simulação apenas as 10 ações com maior liquidez na Bovespa. Os resultados da simulação para estas ações serão apresentados na próxima subseção.

7.2.2 Simulação

Nesta subseção, vamos apresentar a simulação do modelo de operação 1 (MO1) utilizando as ações com maior liquidez na Bovespa. Para analisar o retorno financeiro do MO1, realizamos esta simulação em dois cenários: um que não considera taxas operacionais e outro em que utilizam-se os custos praticados pela Bovespa, apresentados na tabela 8. Para ambas as simulações, consideramos taxa de corretagem⁷ zero. Este benefício é oferecido por algumas corretoras de valores que isentam desta taxa

⁷A taxa de corretagem é um valor cobrado pelas corretoras de valores para cada operação de compra ou venda de ações e outros ativos na Bolsa.

de compra e venda de ações. Os resultados da simulação sem custos operacionais para todo o ano de 2017 e também no ano de 2018 são apresentados nas tabelas 13 e 14, respectivamente. Os indicadores obtidos quando consideramos as taxas praticadas pela Bovespa nos mesmos períodos são apresentados nas tabelas 15 e 16.

Tabela 13 – Resultados da simulação utilizando MO1 sem taxas ano de 2017

Sharpe	Sortino	Retorno acumulado	Volatilidade	Max drawdown
10,15	25,06	11,17%	1,02%	-0,46%
Alpha	Beta	Retorno anual esperado	Estabilidade	PNL
11,34%	0	10,36%	0,97	R\$ 111688,8
Taxa	VaR	Média caixa estratégia (R\$)	Taxa de custódia	Taxa de aluguel
R\$ 0,0	-0,08%	1.021.215,47 (14.018,87)	R\$ 0,0	R\$ 0,0

Nota: Indicadores resultantes para a simulação MO1 sem taxas para o período de 2017 para as 10 ações com maior liquidez no mercado brasileiro. Os sinais foram gerados utilizando-se as 10 ações e as operações foram realizadas com as quais apresentaram sinais não neutros.

Tabela 14 – Resultados da simulação utilizando MO1 sem taxas no ano de 2018

Sharpe	Sortino	Retorno acumulado	Volatilidade	Max drawdown
6,93	15,93	15,81%	2,38%	-1,47%
Alpha	Beta	Retorno anual esperado	Estabilidade	PNL
16,46%	-0,01	16,48%	0,98	R\$ 158,066,09
Taxa	VaR	Média Caixa estratégia (R\$)	Taxa de custódia	Taxa de aluguel
R\$0,00	-0,21%	1.031.232,46 (32.176,83)	R\$0,00	R\$0,00

Nota: Indicadores resultantes para a simulação MO1 sem taxas para o período de 2018 para as 10 ações com maior liquidez no mercado brasileiro. Os sinais foram gerados utilizando-se as 10 ações e as operações foram realizadas com as quais apresentaram sinais não neutros.

Tabela 15 – Resultados da simulação utilizando MO1 aplicando taxas operacionais para o ano de 2017

Sharpe	Sortino	Retorno acumulado	Volatilidade	Max drawdown
-19,28	-19,28	-43,26%	2,25%	-43,26%
Alpha	Beta	Retorno anual esperado	Estabilidade	PNL
-44,20%	0,01	-43,34%	1	R\$ -432608,2
Taxa	VaR	Média caixa estratégia (R\$)	Taxa de custódia	Taxa de aluguel
R\$ 544297,0	-0,39%	889.146,84 (63.762,6)	R\$ 0,0	R\$ 0,0

Nota: Indicadores resultantes para a simulação MO1 com taxas para o período de 2017 para as 10 ações com maior liquidez no mercado brasileiro. Os sinais foram gerados utilizando-se as 10 ações e as operações foram realizadas com as quais apresentaram sinais não neutros.

Os valores dos indicadores apresentados nas tabelas 13, 14, 15 e 16 são obtidos com a simulação que utiliza a combinação de todas as 10 ações para gerar o sinal e,

Tabela 16 – Resultados da simulação utilizando MO1 aplicando taxas operacionais para o ano de 2018

Sharpe	Sortino	Retorno acumulado	Volatilidade	Max drawdown
-9,92	-10,15	-70,30%	6,87%	-70,31%
Alpha	Beta	Retorno anual esperado	Estabilidade	PNL
-71,16%	-0,01	-68,15%	1	R\$ -702,954,05
Taxa	VaR	Média Caixa estratégia	Taxa de custódia	Taxa de aluguel
R\$ 861,020,15	-1,04%	833709,19 (94645,73)	R\$0,00	R\$0,00

Nota: Indicadores resultantes para a simulação MO1 com taxas para o período de 2018 para as 10 ações com maior liquidez no mercado brasileiro. Os sinais foram gerados utilizando-se as 10 ações e as operações foram realizadas com as quais apresentaram sinais não neutros.

posteriormente, realiza operações com as ações nas quais o modelo envia um sinal não neutro. O retorno obtido nesta situação é o retorno acumulado r_a entre todas estas ações.

Na simulação sem aplicar taxas, o retorno acumulado r_a foi positivo no ano de 2017 e 2018. Já na simulação em que aplicamos taxas transacionais, o retorno acumulado r_a foi negativo em ambos os períodos analisados. Isso exemplifica o quanto as taxas impactam os retornos financeiros.

O resultado do indicador sharpe, que mostra a relação entre risco e retorno de um investimento quando comparado a um investimento sem risco, foi negativo em ambos os períodos analisados para a simulação aplicando taxas transacionais. Já as estratégias sem taxas os índices apresentaram valores positivos, indicando que há uma relação positiva entre risco e retorno da estratégia.

Quando avaliada a relação risco e retorno da estratégia levando em consideração apenas a volatilidade negativa, os resultados obtidos para o indicador sortino foram positivos para as simulações sem taxas e, quando inseridos os custos transacionais, os resultados do indicador foram negativos em ambos os períodos.

Os maiores percentuais de perdas da estratégia durante as execuções, retratada pelo indicador Max drawdown, foram obtidos para as estratégias simuladas com taxas, sendo $-43,26\%$ no ano de 2017 e $-70,31\%$ no ano de 2018. Já para as estratégias sem taxas, os Max drawdown obtidos foram $-0,46\%$ no ano de 2017 e $-1,47\%$ no ano de 2018.

Os piores casos esperados de retornos diários da estratégia, estimado pelo indicador VAR, foram obtidos para as simulações com taxas. No anos de 2017 o valor obtido foi $-0,39\%$ e no ano de 2018 o valor foi $-1,04\%$. Para as simulações sem taxas, o valor foi $-0,08\%$ no ano de 2017 e $-0,21\%$ no ano de 2018. Analisando as simulações com taxas e sem taxas, percebemos que os piores casos para ambas as estratégias foram obtidos no ano de 2018. Neste mesmo período, a volatilidade anualizada da

estratégia, considerando os desvios dos retornos diários, foi maior que em 2017. Isto nos sugere que, em períodos de maior volatilidade diária, os piores casos dos retornos diários tendem a ser maiores do que em períodos de menor volatilidade.

O resultado do indicador Beta, que mostra a correlação do desempenho negativo e positivo da estratégia em relação a um benchmark, no ano de 2018 foi negativo para a estratégia executada sem taxas operacionais e também para a estratégia com taxas. No ano de 2017, o resultado de Beta para a estratégia com taxas foi positivo e para a simulação sem taxas o valor de Beta foi zero.

Os resultados do indicador Alpha, que retorna o valor que representa o quanto a estratégia se destacou em relação ao desempenho da economia (representado por um benchmark), para as simulações referentes aos anos de 2017 e 2018, foram positivos para as estratégias em que não aplicamos taxas operacionais. Já para as estratégias com taxas inseridas, os resultados foram negativos em ambos os períodos.

O indicador de retorno anual esperado mostra o resultado estimado que a estratégia alcançará no período de 1 ano. Os resultados para as simulações sem taxas foram positivos para ambos os períodos. Já para as simulações com taxas, os resultados dos retornos estimados foram negativos para os anos de 2017 e 2018.

Para as simulações com taxas, o resultado do indicador de estabilidade da estratégia, baseado no coeficiente de determinação R^2 dos retornos acumulados, foi igual a 1 para os anos de 2017 e 2018. Para as simulações sem taxas, os resultados foram 0,97 e 0,98 para os anos de 2017 e 2018, respectivamente.

Como esperado, o indicador PNL, que mostra o resultado financeiro (lucro ou prejuízo) após a execução da estratégia, para ambas as simulações aplicando taxas foram negativos. Já para as simulações sem considerar taxas operacionais, os resultados para este indicador foram positivos em ambos os períodos.

A média de caixa da estratégia ficou próxima do valor total de capital inicial configurado. A pouca alocação de capital foi proporcionada pelo fato de que as posições nesta simulação são zeradas no instante (candle) seguinte.

Os custos referentes à taxa de custódia⁸ e de aluguel de ações não se aplicam nas simulações, uma vez que as posições são zeradas no mesmo dia.

Na maioria das vezes o modelo apresentou acurácias com valores superiores a um modelo aleatório, mas ainda assim, quando inserimos taxas, o mesmo não apresentou um retorno financeiro considerável. Isso nos motivou a investigar as estatísticas dos preços e custos de execução da estratégia. As taxas de negociação, liquidação e corretagem são responsáveis por diminuir o retorno. Para explorarmos as carac-

⁸É uma taxa cobrada pela manutenção de posição de ações na bolsa de valores.

terísticas dos dados, uma análise estatística dos preços na amostra completa e nos tempos com negociação foi realizada e será apresentada na próxima subseção.

7.2.3 Análise estatística de resultados financeiros relativos ao MO1

Para o modelo de operação 1, que zera a posição no fechamento do candle, é possível verificar estatisticamente se os resultados obtidos estão de acordo com o que podemos esperar da estratégia. Até aqui, utilizamos a combinação de todas as 10 ações para gerar o sinal e depois executamos a simulação operando com todas as ações de acordo com o sinal retornado pelo modelo. Como um dos resultados da simulação, é obtido um rendimento total agregado r_a (referido como retorno acumulado na seção (7.1)). Aqui, para analisarmos as estatísticas do rendimento de cada ação, continuamos usando a combinação de todas as ações para gerar o sinal, porém empregamos apenas uma ação por vez para fazer as operações e calcular o rendimento da estratégia \mathfrak{R}_k , obtido para cada ação k . Tal rendimento será comparado com o rendimento esperado das operações, cujo cálculo é descrito a seguir.

Para cada ação, durante a simulação é possível observar η retornos⁹ $\Delta x(t)$ no intervalo analisado, com $(t = 1, \dots, \eta)$. Como o modelo não opera em todos os passos de tempo, uma vez que a operação depende de um sinal não neutro enviado, ele irá atuar em um subconjunto dos η retornos. Para cada valor de t , o modelo apresenta uma probabilidade p_o de operar naquele dado candle. Assumimos aqui que a probabilidade que haja sinal não neutro em tal candle não depende de t . Dado um sinal não neutro, a probabilidade de que o modelo acerte a previsão é p_a e de que erre é $p_e = 1 - p_a$.

Consideremos que na simulação foram realizadas $m \leq \eta$ operações. Tomando aleatoriamente uma dessas operações, ela pode se referir a qualquer um dos retornos $\Delta x(t)$, com probabilidade uniforme $1/\eta$. Já a probabilidade de qual tal operação seja referente a $\Delta x(t)$ e que a previsão a ela relacionada seja acertada (errada) é p_a/η (p_e/η).

Durante a simulação, se o modelo acerta a previsão de tendência, o resultado financeiro é acrescido do valor do retorno. Se ele erra, é decrescido de forma equivalente. Assim, se o retorno $\Delta x(t)$ de módulo $|\Delta x(t)|$ for escolhido para operar, ele pode produzir uma variação $|\Delta x(t)|$ ou $-|\Delta x(t)|$ no resultado (considerando-se, neste ponto, que a operação se refere a uma ação apenas e não a um lote de ações). Portanto, para cada operação escolhida aleatoriamente no conjunto de todas as operações, a probabilidade de que o retorno $\Delta x(t)$ tenha sido escolhido e que tenha levado a uma variação $|\Delta x(t)|$ ($-|\Delta x(t)|$) é p_a/η (p_e/η).

⁹Retornos reais entre cada um dos candles no período da simulação.

Definamos uma variável aleatória M que pode assumir valores no conjunto $\{|\Delta x(t)|, -|\Delta x(t)|\}$ com probabilidade p_a/η associada aos valores $|\Delta x(t)|$ e p_e/η aos valores $-|\Delta x(t)|$. A variável M corresponde à variação do resultado financeiro de uma operação escolhida aleatoriamente no conjunto das operações.

Suponhamos que tomamos, uma por uma, as operações disponíveis no conjunto até esgotá-lo, e que somamos as variações correspondentes. Tal soma coincide com a variação total das m operações, ou seja, o resultado final da estratégia aplicada àquela ação específica. A variável aleatória a ela associada é $S_M = \sum_{j=1}^m M_j$. Consideraremos, no que segue, que as M_j são independentes e identicamente distribuídas, com distribuição igual à de M . De fato, isto é uma aproximação do cenário real, uma vez que ao retirarmos um retorno do conjunto, a distribuição do mesmo muda.

A média e a variância de M são, respectivamente,

$$\mu_M = \sum_{i=1}^{\eta} \frac{p_a}{\eta} |\Delta x(t)| + \sum_{i=1}^{\eta} \frac{p_e}{\eta} (-|\Delta x(t)|) = \left(\frac{2p_a - 1}{\eta} \right) \sum_{i=1}^{\eta} |\Delta x(t)| \quad (42)$$

e

$$\begin{aligned} \sigma_M^2 &= \sum_{i=1}^{\eta} \frac{p_a}{\eta} (|\Delta x(t)| - \mu_M)^2 + \sum_{i=1}^{\eta} \frac{p_e}{\eta} (-|\Delta x(t)| - \mu_M)^2 \\ &= \left(\frac{p_a + p_e}{\eta} \right) \left(\sum_{i=1}^{\eta} |\Delta x(t)|^2 + \sum_{i=1}^{\eta} \mu_M^2 \right) - 2\mu_M \left(\frac{p_a - p_e}{\eta} \right) \sum_{i=1}^{\eta} |\Delta x(t)| \\ &= \frac{1}{\eta} \sum_{i=1}^{\eta} |\Delta x(t)|^2 - \mu_M^2, \end{aligned} \quad (43)$$

onde usamos $p_a + p_e = 1$. Assumindo a independência entre os diferentes M_j , a média e a variância de S_M são, respectivamente,

$$\mu_{S_M} = m\mu_M \quad (44)$$

e

$$\sigma_{S_M}^2 = m\sigma_M^2. \quad (45)$$

Consideremos que as operações são realizadas em lotes de tamanho λ de ações. Desta forma, o valor esperado do rendimento financeiro referente às m operações pode ser escrito como

$$\mu_M = \left(\lambda \left(\frac{2p_a - 1}{\eta} \right) \sum_{i=1}^{\eta} |\Delta x(t)| \right) m. \quad (46)$$

A variância, por sua vez:

$$\sigma_M^2 = \left(\frac{\lambda^2}{\eta} \sum_{i=1}^{\eta} |\Delta x(t)|^2 - \mu_M^2 \right) m. \quad (47)$$

Para investigarmos se estatisticamente os rendimentos \mathcal{R}_k efetivamente obtidos para cada ação k estão de acordo com o que poderíamos esperar do modelo, analisamos a partir da acurácia e do retorno $\Delta x(t)$, o valor esperado de tais rendimentos. As médias são dadas por μ_M , seguidas de seus desvios σ_M . Para o cálculo da variação das operações, utilizamos duas amostras: a amostra que contém o conjunto com todos retornos $\Delta x(t)$ do período analisado e a amostra que contém somente os retornos $\Delta x(t)$ em que t coincide com os tempos em que o modelo operou. Os resultados com as médias (desvios) calculadas (os) nas 2 amostras e os dados dos rendimentos \mathcal{R}_k obtidos com o modelo de operação 1 nos anos de 2017 e 2018 são apresentados nas tabelas 17 e 18. Os valores obtidos são também apresentados nas figuras 19 e 20.

Tabela 17 – Rendimentos do modelo de operação 1 no ano de 2017

2017	\mathcal{R}_k	Amostra completa $\mu_M (\sigma_M)$	Amostra com negociação $\mu_M (\sigma_M)$
BBDC4	17093,17	21821,92 (5437,64)	26327,16 (6343,21)
VALE5	14007,52	9371,55 (5079,52)	12177,75 (6472,43)
CIEL3	12194,86	20887,78 (4234,95)	27174,77 (5453,93)
VALE3	4722,09	9439,61 (8807,60)	11418,90 (10467,72)
ITUB4	23449,4	30644,58 (5040,63)	36807,01 (5952,52)
PETR4	10198,21	11904,26 (2576,48)	13447,99 (2823,12)
ABEV3	9668,56	9620,73 (2505,78)	11792,46 (3102,27)
BOVA11	1108,68	1497,47 (677,75)	1717,38 (747,92)
BBAS3	16429,88	16885,65 (5345,99)	20248,18 (6377,86)
PETR3	2816,42	5774,13 (2864,12)	6798,61 (3345,79)

Nota: Na variável \mathcal{R}_k apresentamos o rendimento referente a cada ação k . Em $\mu_M (\sigma_M)$ estão os valores referentes à média (desvio) do retorno esperado de cada ação. Os valores foram obtidos utilizando o cálculo de retornos esperados na amostra completa e na amostra em tempos com negociação.

A figura 19 mostra os rendimentos esperados e os efetivamente obtidos com a simulação no ano de 2017 para a amostra completa e para os tempos com negociação. Já a figura 20 mostra resultados análogos para o ano de 2018.

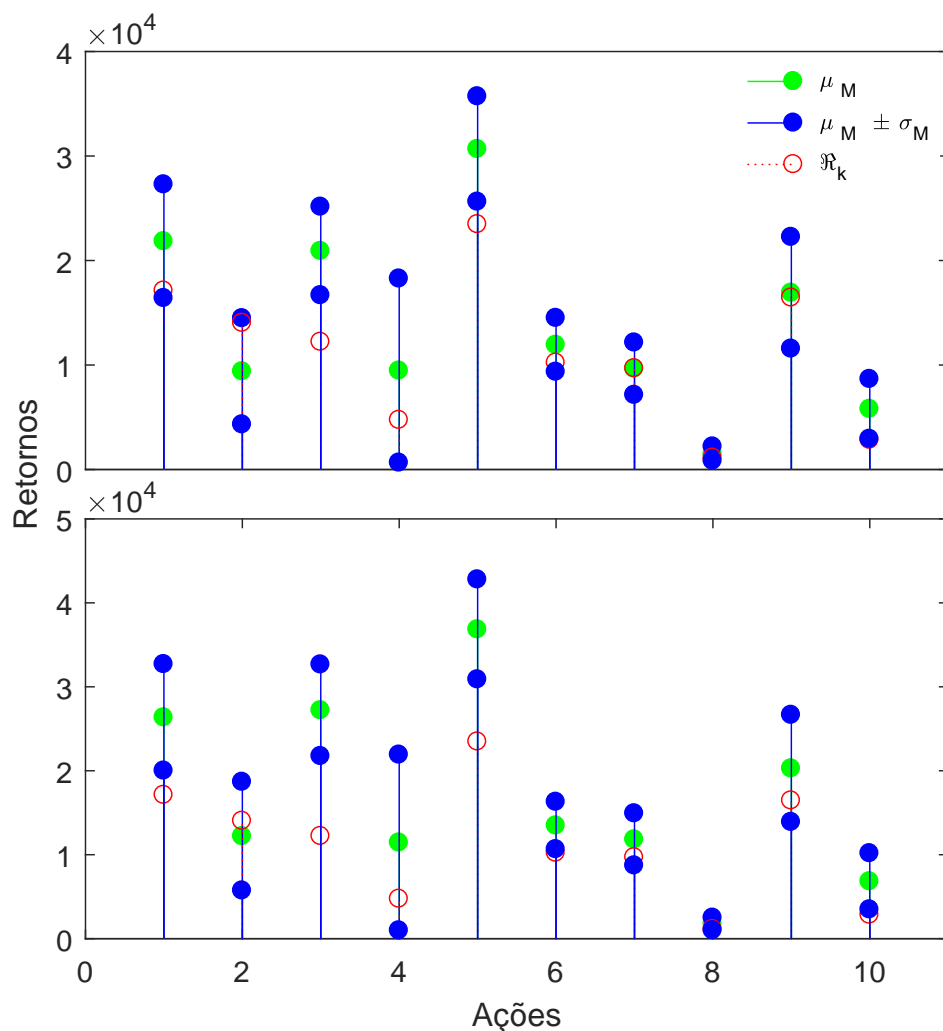


Figura 19 – **Rendimentos esperados e os efetivamente obtidos \mathcal{R}_k para o ano de 2017 para a amostra completa e a amostra com negociação.** Os resultados apresentados são referentes aos rendimentos esperados e os desvios calculados nos tempos com negociação durante a simulação no ano de 2017. No quadro 1, mostramos os resultados referentes à amostra completa, no quadro 2 estão plotados os resultados obtidos com a amostra referente a períodos com negociação. Os pontos na cor vermelha representam a média dos rendimentos efetivamente obtidos na simulação, os pontos verdes mostram as médias dos rendimentos esperados e os pontos em azul representam o rendimento esperado \pm os desvios.

Tabela 18 – Retornos do modelo de operação 1 no ano de 2018

2018	\mathcal{R}_k	Amostra completa $\mu_M (\sigma_M)$	Amostra com negociação $\mu_M (\sigma_M)$
ABEV3	6108,98	4353,44 (2999,79)	5078,60 (3486,69)
BOVA11	-17,08	1282,45 (955,94)	1447,97 (1067,73)
MGLU3	42455,92	16740,92 (29154,45)	22892,44 (38701,77)
ITUB4	25870,75	22474,39 (7655,28)	26135,96 (8938,25)
PETR4	2178,76	-2763,2 (4881,72)	-3171,57 (5692,06)
BBAS3	14730,77	16582,55 (7279,62)	19998,88 (8744,14)
ITUB3	37863,26	12456,87 (4726,40)	16325,59 (5654,28)
VALE3	-6255,68	-10611,41 (8749,29)	-12517,94 (10520,79)
BBDC4	17860,22	9439,03 (6635,68)	10977,92 (7927,68)
B3SA3	17270,16	16981,16 (5371,50)	19971,59 (6278,90)

Nota: Na variável \mathcal{R}_k apresentamos o rendimento referente a cada ação k . Em $\mu_M (\sigma_M)$ estão os valores referentes à média (desvio) do retorno esperado de cada ação. Os valores foram obtidos utilizando o cálculo de retornos esperados na amostra completa e na amostra em tempos com negociação.

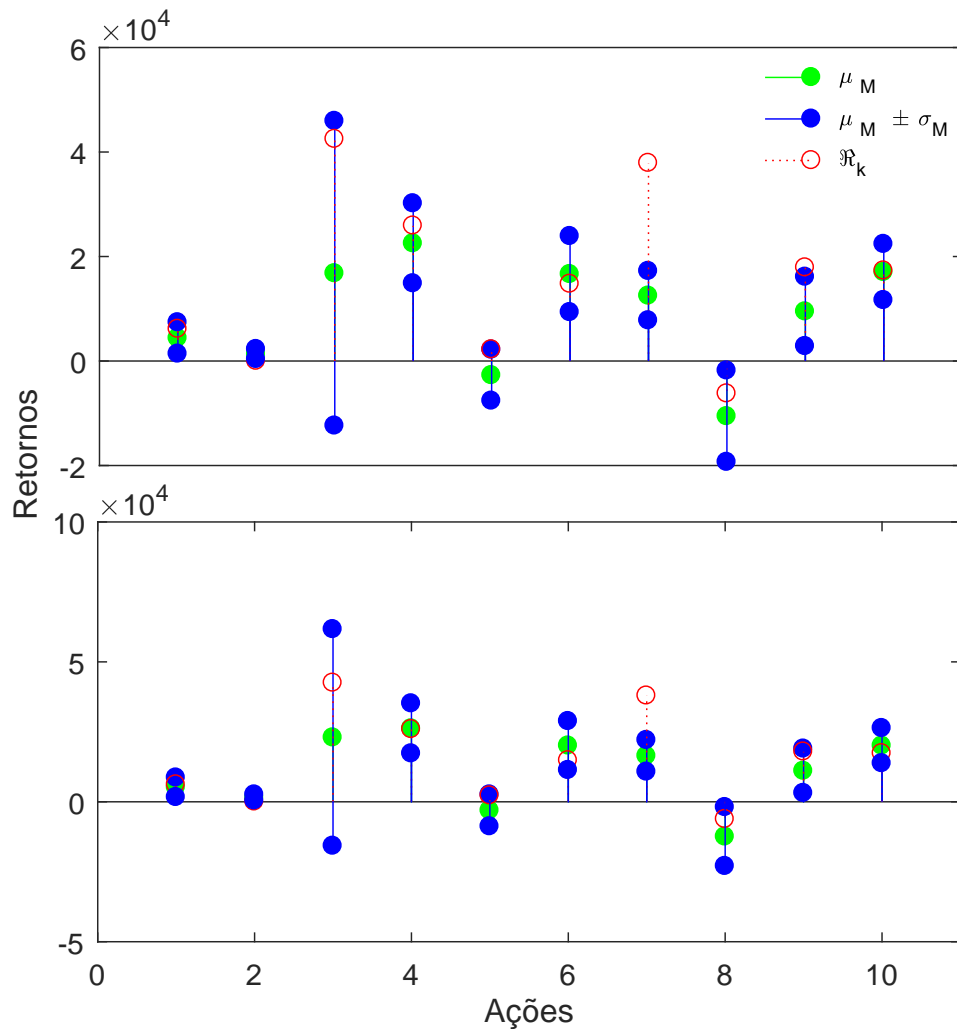


Figura 20 – **Rendimentos esperados e os efetivamente obtidos \mathcal{R}_k para o ano de 2018 para a amostra completa e a amostra com negociação.** Os resultados apresentados são referentes aos rendimentos esperados e os desvios calculados nos tempos com negociação durante a simulação no ano de 2018. No quadro 1, mostramos os resultados referentes à amostra completa, no quadro 2 estão plotados os resultados obtidos com a amostra referente a períodos com negociação. Os pontos na cor vermelha representam a média dos rendimentos efetivamente obtidos na simulação, os pontos verdes mostram as médias dos rendimentos esperados e os pontos em azul representam o rendimentos esperado \pm os desvios.

De acordo com os resultados apresentados nas tabelas 17 e 18 e também nas figuras 19 e 20, podemos observar que 82,5% dos rendimentos financeiros \mathcal{R}_k das ações estão dentro faixa de variação da média do retorno esperado $\mu_M \pm$ os desvios σ_M (33 casos em 40 observações). Tal resultado indica que é possível ter-se uma ideia do resultado da aplicação de um modelo de operação a partir do conhecimento de sua acurácia e dos retornos absolutos médios.

7.2.4 Considerações sobre MO1

Neste modelo, a cada passo de tempo analisamos se há posições em aberto e fechamos a operação com ordens a mercado. Em alguns casos, a ordem desfeita era imediatamente refeita devido ao sinal enviado pelo modelo. Este processo gerou um número de operações maior que o necessário para a aplicação do modelo. Assim, a simulação gerou no ano de 2017 um número de 97.209 operações e no ano de 2018 o total de 91.733 operações. Esta quantidade elevada de operações está relacionado à ordens e contra ordens que podem ser executadas em um mesmo instante de tempo, mas não apenas a isso. Sendo um modelo de operação baseado em negociações de curta duração, o ganho proveniente da acurácia alcançada é pequeno em relação às taxas. Por outro lado, o MO1 é mais controlável, pois conhecendo as acurácias, o retorno médio do ativo e as taxas praticadas no mercado, é possível definir a partir de qual nível o modelo se torna rentável. A partir deste ponto, neste tipo de estratégia o risco é menor, uma vez que desfazemos rapidamente a posição, ficando pouco tempo alocados no ativo, diminuindo, assim, o risco de exposição a grandes oscilações de mercado. Isto acontece pelo fato de que toda operação é desfeita no passo seguinte antes de abrir novas posições. Outra vantagem é em relação ao valor reduzido de capital necessário para se manter alocado nas operações. Desta forma, concluímos que o MO1 não gera bons resultados no cenário de taxas e acurácias que apresentamos. Contudo, para taxas menores e acurácias maiores, pode constituir-se um modelo interessante devido ao baixo risco que apresenta, bem como ao reduzido capital necessário.

7.3 Modelo de operação 2

Nesta estratégia, quando o modelo de previsão gera um sinal não neutro, operações de compra ou venda são abertas no preço $q_o(t+1)$, que é o preço mais próximo ao de abertura $p_o(t+1)$ do candle seguinte disponível (ver equação 40). Estas operações são zeradas apenas quando o preço atinge um valor de stop gain (Sg) ou stop loss (Sl). Neste modelo de operação, é possível limitar as perdas e ganhos das operações de forma que as posições são zeradas no momento que o preço atinge um ganho ou perda determinados, independente do tempo. No modelo definido, se o preço não

atingir nenhum stop no período da simulação, a posição é fechada no último tempo da amostra.

Para esta estratégia, desenvolvemos na plataforma apenas uma função de agendamento que manipula os dados, executa o modelo e lança as ordens ao mercado de acordo com o sinal de compra ou venda.

Os valores dos stops foram inspirados na metodologia apresentada em (GARCIA et al., 2020). Aqui, adaptamos para que os valores de uma média relativa sejam calculados em uma janela deslizante precedente ao dado analisado. Os stops são definidos com base no módulo do retorno médio de cada ação calculado na janela de dados Q , que também é utilizada para cálculo do quartil no período precedente ao instante analisado. Esta metodologia é utilizada para evitar o uso de dados futuros no cálculo. Os valores dos stops foram definidos como

$$Sl = q_o(t + 1) - \kappa_l \left(\frac{|\overline{\Delta x}|(t)}{x(t)} \right) q_o(t + 1) \quad (48)$$

e

$$Sg = q_o(t + 1) + \kappa_g \left(\frac{|\overline{\Delta x}|(t)}{x(t)} \right) q_o(t + 1), \quad (49)$$

onde $x(t)$ é o preço de fechamento no tempo t , $q_o(t + 1)$ é o preço de abertura da posição, $\kappa_l = 10$ e $\kappa_g = 20$ são parâmetros escolhidos. Já $|\overline{\Delta x}|(t)$ é a média dos módulos dos retornos na janela de tamanho Q , dada por

$$|\overline{\Delta x}|(t) = \frac{1}{Q} \sum_{\tau=t-Q}^{t-1} \Delta x(\tau). \quad (50)$$

Estes valores foram definidos considerando a variação de cada ação. Assim, a cada instante de tempo o valor do stop varia tanto pela mudança da janela de dados, pelo preço $x(t)$ e também pelo preço de abertura de cada operação $q_o(t + 1)$.

Nesta estratégia, como as operações levam mais passos de tempo para serem zeradas, aqui a alocação de capital durante a execução foi muito maior que no modelo de operação 1. Em todos os períodos analisados, o valor financeiro em caixa ficou em algum momento da execução totalmente alocado.

Analogamente à seção anterior, para mostrar o efeito das taxas no retorno financeiro, aqui também simulamos em dois cenários. Em um deles não aplicamos taxas de negociação. Na outra estratégia utilizamos as taxas praticadas pela Bovespa, que foram apresentadas na tabela 8.

Os resultados das simulações sem taxas para o ano de 2017 e 2018 são apresentados nas tabelas 19 e 20, respectivamente. Nestes resultados, não consideramos a remuneração do capital parado no decorrer da estratégia.

Para mostrar a consistência do modelo em diferentes períodos, executamos as estratégias em todos os quadrimestres do ano de 2017 e 2018. Esta simulação em diferentes períodos é para evitar a super otimização (*overfitting*) do modelo. A super otimização acontece quando a estratégia, depois de vários ajustes, funciona muito bem no período analisado, porém, em outros períodos, os resultados não se repetem, tendo um resultado pior. Os parâmetros perfeitos são encontrados para aquele período e só funcionam bem naquela amostra. Uma forma de testar a estratégia é simulando em períodos diferentes para o teste de validação. Por isso, aplicamos a estratégia em 6 diferentes quadrimestres.

Os resultados da simulação para o ano de 2017 com taxas praticadas no mercado são apresentados na tabela 21. Para o ano de 2018, os resultados são mostrados na tabela 22. Nestes resultados, também não consideramos a remuneração do capital parado no decorrer da estratégia. Como taxa de negociação, utilizamos o maior valor de taxa de negociação da Bovespa para day trade apresentado na tabela 8. Como analisamos vários períodos, o número de tabelas e consequentemente o número de indicadores resultantes das simulações é grande, portanto vamos fazer uma análise geral dos resultados.

Os valores dos indicadores apresentados nas tabelas 19, 20, 21 e 22 são obtidos com a simulação da estratégia que utiliza a combinação de todas as 10 ações para gerar o sinal e negocia apenas com as ações em que o modelo retorna um sinal não neutro, indicando uma operação naquele instante. O retorno obtido nesta situação é o retorno acumulado r_a .

Nas simulações, quando aplicamos a estratégia sem inserir taxas, os resultados dos retornos acumulados r_a foram positivos em todos os períodos dos anos de 2017 e 2018 analisados. Para a simulação utilizando taxas aplicadas ao cenário brasileiro, todos os períodos analisados também apresentaram retorno acumulado positivo, com exceção do quadrimestre 1 de 2018. Considerando-se todo o ano de 2017, a estratégia rendeu 28,5% e o desempenho do Ibovespa foi 20,09%. Já no ano de 2018 inteiro, o desempenho da estratégia foi 28,94% e o Ibovespa apresentou um rendimento de 15,03%. Todas as taxas anuais de crescimento do Ibovespa desde o ano de 1968 são apresentadas na tabela 22, anexada ao capítulo C. A taxa de crescimento apresentada refere-se à taxa média de variação do índice composta anualmente, entre dois períodos (BM&FBOVESPA, 2019).

Os resultados do indicador alpha, que representa o quanto a estratégia se destacou em relação ao desempenho da economia (representado por um benchmark), foram

positivos para as simulações em todos os períodos dos anos de 2017 e 2018.

O indicador de retorno anual esperado, que mostra o resultado estimado que a estratégia alcançará no período de 1 ano, foi positivo em todos os períodos analisados.

O indicador PNL, que mostra o resultado financeiro (lucro ou prejuízo) após a execução da estratégia, foi positivo para os períodos analisados, exceto no primeiro quadrimestre de 2018.

A média de caixa desta estratégia é menor que a estratégia anterior. No decorrer da simulação, em alguns instantes o capital ficou completamente alocado nas operações. A maior alocação de capital foi proporcionada pelo fato de que as posições nesta simulação são zeradas apenas quando os stops são alcançados.

Os indicadores sharpe e sortino foram positivos em todos os períodos analisados. Os maiores percentuais de perdas da estratégia durante as execuções, retratados pelo indicador Max Drawdown, foram observados para as estratégias simuladas com taxas no ano de 2018. O primeiro quadrimestre de 2018 apresentou uma perda de $-21,68\%$, corroborando o fato de ser o único período de retorno negativo. De acordo com o indicador de volatilidade, este período também apresentou grandes variações nos retornos diários.

Para mostrar os retornos obtidos com MO2, elaboramos a figura 21. Nela apresentamos os retornos obtidos com MO2 na simulação com taxas e sem taxas operacionais e também o índice sharpe associado a cada período.

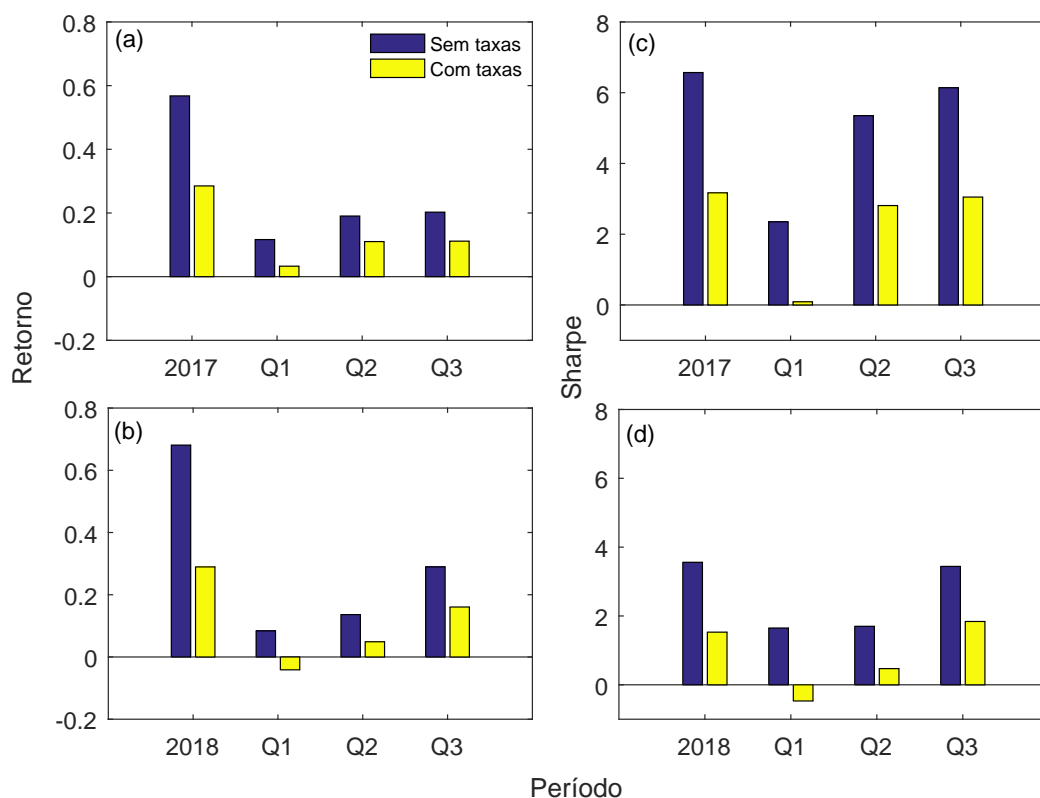


Figura 21 – **Retornos e riscos do modelo de operação 2.** Os resultados são referentes aos retornos da simulação com MO2. Os períodos analisados foram os anos de 2017 e 2018 e também os seus quadrimestres. As barras nas cores rochas representam os resultados para as simulações sem considerar taxas operacionais. Já as barras amarelas correspondem aos resultados obtidos aplicando as taxas operacionais. Os valores referentes ao ano de 2017 e 2018 indicam que a simulação foi executada em todo o período. Os valores representados por Q1, Q2 e Q3 correspondem à simulação executada nos quadrimestres 1, 2 e 3, respectivamente. (a) Retornos acumulados r_a para os dois cenários (com taxas e sem taxas) no ano de 2017. (b) Retornos acumulados r_a no período do ano de 2018. (c) Valores dos indicadores de risco Sharpe associados a cada período da simulação no ano de 2017. (d) Índices Sharpe obtidos para as execuções no período do ano de 2018.

Tabela 19 – Resultados da simulação utilizando stops, sem taxas, para o ano de 2017

2017				
Sharpe	Sortino	Retorno acumulado	Volatilidade	Max drawdown
6,57	12,8	56,78%	8,76%	-6,05%
Alpha	Beta	Retorno anual esperado	Estabilidade	PNL
61,68%	-0,05	57,57%	0,97	R\$ 567824,0
Taxa	VaR	Média caixa (R\$)	Taxa de custódia	Taxa de aluguel
R\$ 0,0	-0,78%	643.935,16 (306.240,23)	R\$ 0,0	R\$ 0,0
Quadrimestre 1				
Sharpe	Sortino	Retorno acumulado	Volatilidade	Max drawdown
2,35	4,27	11,65%	11,59%	-6,05%
Alpha	Beta	Retorno anual esperado	Estabilidade	PNL
45,76%	-0,07	27,2%	0,59	R\$ 116481,1
Taxa	VaR	Média caixa (R\$)	Taxa de custódia	Taxa de aluguel
R\$ 0,0	-1,1%	486.417,98 (230.307,91)	R\$ 0,0	R\$ 0,0
Quadrimestre 2				
Sharpe	Sortino	Retorno acumulado	Volatilidade	Max drawdown
5,35	16,04	19,04%	9,2%	-2,58%
Alpha	Beta	Retorno anual esperado	Estabilidade	PNL
78,27%	-0,26	49,21%	0,87	R\$ 190412,8
Taxa	VaR	Média caixa (R\$)	Taxa de custódia	Taxa de aluguel
R\$ 0,0	-0,81%	578.812,34 (237.587,55)	R\$ 0,0	R\$ 0,0
Quadrimestre 3				
Sharpe	Sortino	Retorno acumulado	Volatilidade	Max drawdown
6,14	11,45	20,26%	9,64%	-3,13%
Alpha	Beta	Retorno anual esperado	Estabilidade	PNL
81,4%	0	59,21%	0,92	R\$ 202628,2
Taxa	VaR	Média caixa (R\$)	Taxa de custódia	Taxa de aluguel
R\$ 0,0	-0,84%	518.760,29 (254.837,95)	R\$ 0,0	R\$ 0,0

Nota: Indicadores resultantes para a simulação empregando stops gain S_g (equação 49) e loss S_l (equação 48) e sem taxas para o período de 2017 para as 10 ações com maior liquidez no mercado brasileiro. Os sinais foram gerados utilizando-se as 10 ações e as operações foram realizadas com as quais apresentaram sinais não neutros.

Tabela 20 – Resultados da simulação utilizando stop gain Sg (equação 49) e loss Sl (equação 48) sem taxas para o ano de 2018

2018				
Sharpe	Sortino	Retorno acumulado	Volatilidade	Max drawdown
3,56	6,22	68,09%	20,60%	-16,95%
Alpha	Beta	Retorno anual esperado	Estabilidade	PNL
80,77%	-0,16	73,35%	0,84	R\$ 680851,4
Taxa	VaR	Média caixa (R\$)	Taxa de custódia	Taxa de aluguel
R\$ 0,0	-1,98%	499.078,73 (358.427,2)	R\$ 0,0	R\$ 0,0
Quadrimestre 1				
Sharpe	Sortino	Retorno acumulado	Volatilidade	Max drawdown
1,65	2,64	8,38%	18,28%	-16,05%
Alpha	Beta	Retorno anual esperado	Estabilidade	PNL
46,58%	-0,28	30,15%	0,16	R\$ 83822,8
Taxa	VaR	Média caixa (R\$)	Taxa de custódia	Taxa de aluguel
R\$ 0,0	-1,81%	340.228,38 (251.160,93)	R\$ 0,0	R\$ 0,0
Quadrimestre 2				
Sharpe	Sortino	Retorno acumulado	Volatilidade	Max drawdown
1,7	2,57	13,62%	21,02%	-9,98%
Alpha	Beta	Retorno anual esperado	Estabilidade	PNL
42,35%	-0,17	35,66%	0,7	R\$ 136166,8
Taxa	VaR	Média caixa (R\$)	Taxa de custódia	Taxa de aluguel
R\$ 0,0	-2,06%	470.881,96 (244.125,98)	R\$ 0,0	R\$ 0,0
Quadrimestre 3				
Sharpe	Sortino	Retorno acumulado	Volatilidade	Max drawdown
3,44	6,25	28,94%	25,07%	-14,19%
Alpha	Beta	Retorno anual esperado	Estabilidade	PNL
156,73%	-0,16	86,18%	0,53	R\$ 289358,1
Taxa	VaR	Média caixa (R\$)	Taxa de custódia	Taxa de aluguel
R\$ 0,0	-2,36%	429.438,42 (312.006,91)	R\$ 0,0	R\$ 0,0

Nota: Indicadores resultantes para a simulação empregando stops gain Sg (equação 49) e loss Sl (equação 48) e sem taxas para o período de 2018 para as 10 ações com maior liquidez no mercado brasileiro. Os sinais foram gerados utilizando-se as 10 ações e as operações foram realizadas com as quais apresentaram sinais não neutros.

Tabela 21 – Resultados da simulação utilizando stops, com taxas, para o ano de 2017

2017				
Sharpe	Sortino	Retorno acumulado	Volatilidade	Max drawdown
3,17	5,57	28,5%	9,25%	-8,76%
Alpha	Beta	Retorno anual esperado	Estabilidade	PNL
32,09%	-0,06	29,31%	0,92	R\$ 285008,0
Taxa	VaR	Média caixa (R\$)	Taxa de custódia	Taxa de aluguel
R\$ 265579,2	-0,89%	522.518,56 (257.945,12)	R\$ 0,0	R\$ 0,0
Quadrimestre 1				
Sharpe	Sortino	Retorno acumulado	Volatilidade	Max drawdown
0,09	0,16	3,29%	11,94%	-8,76%
Alpha	Beta	Retorno anual esperado	Estabilidade	PNL
14,46%	-0,08	1,13%	0	R\$ 32909,4
Taxa	VaR	Média caixa (R\$)	Taxa de custódia	Taxa de aluguel
R\$ 83796,2	-1,2%	451.074,09 (223.386,08)	R\$ 0,0	R\$ 0,0
Quadrimestre 2				
Sharpe	Sortino	Retorno acumulado	Volatilidade	Max drawdown
2,81	7,1	11,02%	9,15%	-3,61%
Alpha	Beta	Retorno anual esperado	Estabilidade	PNL
45,52%	-0,26	25,7%	0,66	R\$ 110208,1
Taxa	VaR	Média caixa (R\$)	Taxa de custódia	Taxa de aluguel
R\$ 78614,1	-0,86%	545.237,54 (230.772,41)	R\$ 0,0	R\$ 0,0
Quadrimestre 3				
Sharpe	Sortino	Retorno acumulado	Volatilidade	Max drawdown
3,05	5,09	11,15%	9,74%	-3,98%
Alpha	Beta	Retorno anual esperado	Estabilidade	PNL
41,09%	0	29,73%	0,74	R\$ 111510,1
Taxa	VaR	Média caixa (R\$)	Taxa de custódia	Taxa de aluguel
R\$ 89283,3	-0,91%	476.858,63 (243.342,32)	R\$ 0,0	R\$ 0,0

Nota: Indicadores resultantes para a simulação empregando stops gain S_g (equação 49) e loss S_l (equação 48) e aplicando taxas para o período de 2017 para as 10 ações com maior liquidez no mercado brasileiro. Os sinais foram gerados utilizando-se as 10 ações e as operações foram realizadas com as quais apresentaram sinais não neutros.

Tabela 22 – Resultados da simulação utilizando stops, sem taxas, para o ano de 2018

2018				
Sharpe	Sortino	Retorno acumulado	Volatilidade	Max drawdown
1,53	2,57	28,94%	22,80%	-22,86%
Alpha	Beta	Retorno anual esperado	Estabilidade	PNL
40,15%	-0,22	34,94%	0,57	R\$ 289406,8
Taxa	VaR	Média caixa (R\$)	Taxa de custódia	Taxa de aluguel
R\$ 398446,9	-2,28%	359.300,05 (281.505,24)	R\$ 0,0	R\$ 0,0
Quadrimestre 1				
Sharpe	Sortino	Retorno acumulado	Volatilidade	Max drawdown
-0,47	-0,7	-4,13%	19,23%	-21,68%
Alpha	Beta	Retorno anual esperado	Estabilidade	PNL
1,6%	-0,31	-9,13%	0,02	R\$ -41263,3
Taxa	VaR	Média caixa (R\$)	Taxa de custódia	Taxa de aluguel
R\$ 125425,5	-2,01%	299.465,91 (238.614,08)	R\$ 0,0	R\$ 0,0
Quadrimestre 2				
Sharpe	Sortino	Retorno acumulado	Volatilidade	Max drawdown
0,47	0,69	4,89%	21,43%	-11,07%
Alpha	Beta	Retorno anual esperado	Estabilidade	PNL
12,63%	-0,18	10%	0,16	R\$ 48917,8
Taxa	VaR	Média caixa (R\$)	Taxa de custódia	Taxa de aluguel
R\$ 94889,9	-2,16%	434.344,58 (231.679,73)	R\$ 0,0	R\$ 0,0
Quadrimestre 3				
Sharpe	Sortino	Retorno acumulado	Volatilidade	Max drawdown
1,84	3,15	16,06%	25,86%	-17,6%
Alpha	Beta	Retorno anual esperado	Estabilidade	PNL
86,1%	-0,19	47,7%	0,27	R\$ 160619,0
Taxa	VaR	Média caixa (R\$)	Taxa de custódia	Taxa de aluguel
R\$ 127013,4	-2,53%	382.076,94 (290.015,82)	R\$ 0,0	R\$ 0,0

Nota: Indicadores resultantes para a simulação empregando stops gain S_g (equação 49) e loss S_l (equação 48) e aplicando taxas para o período de 2018 para as 10 ações com maior liquidez no mercado brasileiro. Os sinais foram gerados utilizando-se as 10 ações e as operações foram realizadas com as quais apresentaram sinais não neutros.

7.4 Considerações finais

Os dois modelos de operação descritos têm características diferentes e apresentam vantagens e desvantagens em relação ao outro. O modelo de operação 1, que zera a posição no tempo seguinte antes de abrir novas operações, tem a vantagem de que o tempo de permanência na posição é muito pequeno, diminuindo o risco de exposição a grandes oscilações do mercado. Porém, o excessivo número de operações leva a taxas operacionais elevadas, inviabilizando a execução em mercados reais. Por outro lado, o MO1 é mais controlável e nos dá uma ideia de qual o patamar de taxas a partir do qual torna-se rentável. Já o modelo de operação 2 tem um número menor de operações, em contrapartida, o risco associado à estratégia é maior, uma vez que a operação só é zerada quando atinge um valor de stop. O número de operações em MO1 foi 97.209 em 2017 e 91.733 em 2018, enquanto o número de operações de MO2 em todo o ano de 2017 foi 9.534; no ano de 2018 a quantidade de operações de MO2 foi 9.238. O número de execuções em MO2 ficou em torno de 10% em relação ao número de operações realizadas em MO1.

Analisando os resultados das simulações, percebemos que os melhores resultados dos rendimentos foram obtidos com a estratégia 2. Embora MO2 tenha se mostrado mais lucrativo, o modelo de operação 1 praticamente não apresenta risco. Com relação a MO2, como mitigação do risco temos o fato de que os valores de stop loss são pequenos. Percebemos isso analisando a equação 48. Uma vez que $x(t)$ é o preço de fechamento e $q_o(t+1)$ é o preço de abertura de posição no candle seguinte, podemos observar que $\frac{q_o(t+1)}{x(t)} \cong 1$. Assim, o stop loss é aproximadamente

$$Sl \cong q_o(t+1) - \kappa_l \left(\overline{|\Delta x|} \right) (t). \quad (51)$$

Uma vez que escolhemos $\kappa_l = 10$, o valor da perda máxima por operação são significativamente menores que o valor da ação. Durante a simulação, para o cálculo do stop utilizamos a janela de dados Q precedente ao tempo investigado para calcular uma média dos módulos dos retornos naquela janela, definida como $\left(\overline{|\Delta x|} \right) (t)$. Aqui, para analisar relação entre os valores dos stops e os preços das ações, utilizamos as séries de fechamento referentes a todo o período da simulação e calculamos a média dos módulos dos retornos, definida por $\overline{|\Delta p_c|}$, e a média dos preços de cada ação, definida por $\overline{p_c}$. Assim, supondo $\overline{|\Delta p_c|} \cong \left(\overline{|\Delta x|} \right) (t)$, podemos considerar a aproximação apresentada na equação 51, com isso, percebemos que os valores dos stops são substancialmente menores que os valores das ações. A tabela 23 mostra as médias dos módulos dos retornos $\overline{|\Delta p_c|}$, médias de preços $\overline{p_c}$ e as razões $\frac{\overline{p_c}}{\overline{|\Delta p_c|}}$.

Por outro lado, como mitigação do risco, também podemos considerar o fato de trabalharmos com uma carteira de ações diversificada. Em uma evolução do trabalho, seria

interessante desenvolver um modelo de controle de risco utilizando as correlações entre as ações.

Em relação à alocação de capital, o risco no modelo de operação 2 é maior, pois em todas as simulações, em algum momento, o capital disponível zerou. Já no modelo de operação 1, durante a simulação utilizamos, em média, apenas 20% do capital disponível. Porém, em todas as simulações o modelo de operação 2 apresentou melhores resultados que o modelo de operação 1.

Tabela 23 – Média de preços e retornos das ações

2017			
	$ \Delta p_c $	\bar{p}_c	$\frac{\bar{p}_c}{ \Delta p_c }$
BBDC4	0,02558665	21,87134354	0,00117
VALE5	0,03627768	28,87922779	0,001256
CIEL3	0,02340959	19,53467112	0,001198
VALE3	0,04188494	29,50070746	0,00142
ITUB4	0,02380873	23,14603021	0,001029
PETR4	0,01766138	13,99935784	0,001262
ABEV3	0,01726891	18,14834573	0,000952
BOVA11	0,04623921	65,87925137	0,000702
BBAS3	0,03640741	28,98630173	0,001256
PETR3	0,02058712	15,23066685	0,001352
2018			
	$ \Delta p_c $	\bar{p}_c	$\frac{\bar{p}_c}{ \Delta p_c }$
ABEV3	0,0216501	19,24430907	0,001125
BOVA11	0,0701715	78,94486398	0,000889
MGLU3	0,21140572	119,69484777	0,001766
ITUB4	0,03648753	29,54197693	0,001235
PETR4	0,03357403	20,18010619	0,001664
BBAS3	0,0510998	34,43680947	0,001484
ITUB3	0,02432655	25,47002471	0,000955
VALE3	0,06316666	49,01690744	0,001289
BBDC4	0,0367708	26,28898962	0,001399
B3SA3	0,03739952	23,53833938	0,001589

Nota: Os valores são referentes à média dos módulos dos retornos $|\Delta p_c|$, média dos preços \bar{p}_c de cada ação e as razões $\frac{\bar{p}_c}{|\Delta p_c|}$. Séries de preços de fechamentos relativos aos períodos em que executamos a simulação (2017 e 2018) foram utilizadas.

Capítulo 8

Conclusão

O trabalho teve o objetivo de desenvolver um modelo baseado em um sistema de equações diferenciais lineares aplicado ao mercado acionário. O intuito foi contribuir para a compreensão da dinâmica dos preços no mercado financeiro e auxiliar na decisão de compra e venda das ações listadas na Bovespa. Para tal, nos apoiamos na lacuna deixada pela hipótese de mercado eficiente.

Como parte do desenvolvimento do trabalho, inicialmente realizamos uma pesquisa bibliográfica. Nesta, autores importantes defendem que os mercados são eficientes e que as séries de preços passados não trazem informações de preços futuros. Na contramão da teoria clássica, outros autores mostram que os mercados não são tão eficientes em termos de equilíbrio de preços, e que as séries apresentam alguma memória e um certo nível de correlação. Tendo em vista estas possibilidades e acreditando que o mercado de ações brasileiro também apresenta resíduos de ineficiência, elaboramos um modelo para ajudar na compreensão da dinâmica do mercado de ações. Para tal, desenvolvemos um modelo baseado em um sistema de equações diferenciais lineares com o objetivo de prever tendências de preços das ações.

No capítulo 5, apresentamos a formulação do modelo baseado em um sistema de equações diferenciais lineares e a sua adaptação para um novo modelo de ajuste de pontos, bem como a definição de um modelo de previsão de tendência de preços. Diante das séries de dados dos ativos, calculamos as matrizes de coeficientes $\mathbf{A}(t)$. Para verificar a variação de tais matrizes em t , calculamos a distância entre as matrizes obtidas em tempos adjacentes. O resultado da análise foi que, à medida em que aumentamos o valor de R , as distâncias das matrizes calculadas em tempos seguintes diminuíam. Este resultado nos sugere que as matrizes $\mathbf{A}(t)$ são boas estimativas para $\mathbf{A}(t + 1)$, podendo, assim, serem utilizadas como parâmetro no modelo de previsão.

No capítulo 6 exploramos algumas das possibilidades de combinação de parâmetros e dados para ajustar o modelo de previsão de forma a obter melhores resultados de

acurácia. As análises foram realizadas utilizando diferentes valores de R , diferentes números de ações, várias granularidades das séries de dados, expoente de Hurst e também a votação entre os valores de R . Nesta parte do trabalho, também investigamos o nível de eficiência em dois tipos de mercados com características diferentes por meio da análise dos resultados de acurácias. As análises foram realizadas utilizando as ações dos índices Dow Jones e Ibovespa e os resultados apontaram que as melhores taxas de acurácias foram obtidas no mercado brasileiro, indicando uma certa ineficiência.

Como resultado da exploração do modelo, definimos os parâmetros satisfatórios dentro do cenário explorado. Como mencionamos anteriormente, não podemos afirmar que os parâmetros definidos são os ótimos. Para tal afirmação seria necessário um estudo mais elaborado utilizando técnicas de otimização.

Na investigação foram considerados os parâmetros que retornaram as melhores taxas de acurácias do modelo na previsão de tendência dos preços das ações. Quando analisamos os resultados dos diferentes valores de R , observamos que para valores maiores de R as acurácias a tendem a aumentar. Assim, para as análises que utilizaram apenas uma janela para o ajuste, o valor definido foi $R = 1000$. Já para análises aplicando a votação, os valores dos parâmetros definidos foram $R = 500$, $R = 1000$ e $R = 1500$.

Quando analisamos os resultados obtidos de a em relação à quantidade de ações utilizadas na combinação das séries de entrada para o modelo, percebemos uma moderada melhora nos valores em torno da combinação de 10 ações (ver figura 10).

Na análise da relação entre a acurácia do modelo e a granularidade dos dados, os melhores resultados foram obtidos para a frequência de cotação de 5 minutos.

Quando analisamos a associação entre as acurácias obtidas e os expoentes de Hurst H das séries, em geral não encontramos padrões relevantes relacionados às acurácias e tais expoentes. Um padrão atípico foi encontrado para a ação CPFE3, em que o expoente de Hurst da série correspondente tem o valor baixo ($H = 0,28$), enquanto que o valor da acurácia para esta ação é relevante ($a = 68, 10\%$). Calculando os coeficientes de correlação linear de Pearson ρ entre as acurácias obtidas em cada quadrimestre e os expoentes de Hurst calculados nos respectivos períodos, os valores obtidos de ρ foram negativos em todos os casos. Os valores mais significativos foram obtidos para as ações do Ibovespa. Estes resultados indicam que há uma anti-correlação entre as acurácias obtidas e os expoentes H das séries analisadas.

A outra exploração do modelo foi baseada na votação em relação à direção da tendência prevista utilizando 10 diferentes valores de R . Os resultados não apontaram ganhos significativos em relação à utilização de apenas o valor $R = 1000$. A vantagem do uso da técnica de votação está na possibilidade de aumento do tamanho da

amostra, uma vez que teremos mais sinais não neutros retornados pelo modelo (retornos previstos no quarto quartil).

Nesta parte do trabalho também investigamos a consistência dos resultados, aplicando o modelo de previsão em diferentes períodos. Os resultados sugeriram que o modelo mantém consistência nos diferentes quadrimestres em relação às ações que apresentaram melhores acurácias e também em relação aos seus valores de acurácias.

Dessa maneira, os parâmetros definidos para a simulação foram baseados nesse capítulo de exploração do modelo. Estas análises foram importantes para ajudar na definição da quantidade de ações, frequência de cotação e tamanho da janela de dados que foram utilizadas na simulação. Como resultado, as séries utilizadas foram as com frequência de cotação de 5 minutos das 10 ações com maior liquidez na Bovespa. A votação entre os sinais obtidos com diferentes tamanhos de R foi utilizada como sinal de tendência. Nesta exploração também observamos que os melhores resultados foram obtidos no mercado brasileiro, sendo este o definido para a simulação. O mercado emergente escolhido para a simulação foi devido à expectativa de que o modelo tenha melhor desempenho neste cenário.

No capítulo 7, desenvolvemos duas estratégias de negociação baseadas no modelo apresentado no capítulo 5. Em ambas as estratégias, se o modelo envia um sinal não neutro, as operações são abertas de acordo com a tendência do preço prevista (alta ou baixa). No modelo de operação 1 (MO1), as operações são fechadas a cada passo de tempo (cada candle) com ordens a mercado. No modelo de operação 2 (MO2), as operações são fechadas de acordo com os valores dos stops definidos nas equações 48 e 49. Ambos os modelos de negociação apresentaram custos de processamento baixos, sendo possível a execução em frequências de dados menores. Para quantificar o efeito dos custos operacionais, as estratégias foram simuladas sem considerar taxas e também aplicando taxas operacionais.

O modelo de operação 1, quando executado aplicando taxas, não obteve um retorno financeiro satisfatório. Sendo um modelo de operação baseado em negociações de curta duração, o ganho proveniente da acurácia alcançada é pequeno em relação às taxas. Porém, este modelo foi importante para a análise estatística dos dados nos períodos utilizados na simulação. Também foi um modelo importante para a análise das acurácias e o retorno acumulado. De fato, que as acurácias por si só não são bons indicadores de sucesso de um modelo. Outras variáveis devem ser levadas em consideração, tais como liquidez, volume e volatilidade do ativo.

O MO1 é mais controlável, pois conhecendo as acurácias, o retorno médio do ativo e as taxas praticadas no mercado, é possível definir em qual nível o modelo se torna rentável. A partir deste ponto o risco é menor, uma vez que nesta estratégia

desfazemos as posições a cada passo de tempo, não ficando expostos a grandes oscilações do mercado. O valor financeiro investido para se manter alocado nas posições também é menor. A desvantagem é em relação ao elevado número de operações realizadas, gerando altas taxas operacionais.

Dessa maneira, concluímos que o MO1 não gera bons resultados no cenário de taxas e acurácias que apresentamos. Contudo, para taxas menores e acurácias maiores, pode constituir-se um modelo interessante devido ao baixo risco que apresenta, bem como ao reduzido capital necessário.

O modelo de operação 2 apresentou um retorno financeiro interessante mesmo no contexto real, considerando os custos operacionais da Bovespa, o modelo ainda apresentou resultados consistentes nos períodos analisados (resultados positivos em cinco quadrimestres e negativo em apenas um). Estes resultados nos sugerem indícios em direção a desvios da hipótese de mercado eficiente. A vantagem do MO2 é o número reduzido de operações. Como mitigação do risco deste modelo, podemos considerar o fato de trabalharmos com uma carteira diversificada de ações (10 ações mais negociadas na Bovespa nos anos de 2017 e 2018). Também, como amenização de risco, podemos considerar a escolha de valores de stop loss pequenos em relação aos preços das ações. Assim, os valores das perdas máximas por operação são significativamente menores que os valores das ações, limitando as perdas caso a operação seja fechada com stop loss (ver tabela 23).

Diante dos resultados, o modelo de operação 2 apresentou maiores ganhos e se apresentou como uma alternativa promissora como estratégia de operação em mercados reais. Porém, apesar de MO2 ter se mostrado mais rentável durante a simulação, não podemos afirmar que pode ser utilizado como ferramenta de decisão para operações em mercado real. O modelo deve ser submetido em outros intervalos, em outras situações de mercado e também levando em consideração outras variáveis para que se possa obter consistência, estabilidade e, com isso, aumentar os retornos financeiros. Assim, este trabalho não tem por finalidade apresentar uma ferramenta de negociação finalizada para ser aplicada no mercado real.

Durante o trabalho, detectamos resíduos de ineficiência nas séries de preços analisadas. Estes resíduos possibilitam oportunidades de arbitragem. Isto nos levou a acreditar que é possível fazer inferência do preço futuro baseando-se em dados passados. Assim, os resultados aqui apresentados apontam na direção de explicitar desvios da Hipótese de Mercado Eficiente no mercado brasileiro. Dadas estas oportunidades e a importância do tema estudado, considera-se que muito há ainda que percorrer no campo da investigação nesta área, sendo, portanto, um campo fértil para trabalhos futuros.

8.1 Trabalhos Futuros

Dentre as inúmeras possibilidades de futuras investigações, destacamos as seguintes propostas:

- Até este ponto do trabalho, sugerimos que nossos dados podem ser descritos por um modelo de regressão multidimensional. Porém, é preciso testar a adequação do modelo como tal com rigor estatístico. Lembremos que a violação das premissas não invalida o modelo, pois até o momento estamos interessados principalmente nos resultados de acerto de previsão de tendência.
- Utilizar técnicas de otimização para auxiliar na escolha de parâmetros ótimos para aplicar ao modelo de previsão.
- Calcular valores de stop gain e stop loss por meio de técnicas de otimização.
- Desenvolver um modelo de controle de risco com critérios de otimização de carteiras para minimizar o risco das estratégias.
- Utilizar o modelo em combinação com outros indicadores.
- Testar o modelo em negociação de contratos em ambiente de mercados futuros, tais como commodities (agrícola¹, mineral²), moedas (dólar e euro) e índices (Bovespa e S&P 500).

¹Soja, milho, trigo e café.

²Petróleo, minério de ferro e ouro.

Referências

AGUIRRE, L. **Introdução à Identificação de Sistemas – Técnicas Lineares e Não-Lineares Aplicadas a Sistemas Reais**. Editora UFMG, 2007. ISBN 9788570415844. Disponível em: <<https://books.google.com.br/books?id=f9lwE7Ph0fYC>>. Citado 3 vezes nas páginas 8, 31 e 113.

AL-KHAZALI, O. M.; PYUN, C. S.; KIM, D. Are exchange rate movements predictable in asia-pacific markets? evidence of random walk and martingale difference processes. **International Review of Economics & Finance**, Elsevier, v. 21, n. 1, p. 221–231, 2012. Citado 2 vezes nas páginas 1 e 14.

ALLEZ, R.; BOUCHAUD, J.-P. Individual and collective stock dynamics: intra-day seasonalities. **New Journal of Physics**, IOP Publishing, v. 13, n. 2, p. 025010, 2011. Citado na página 5.

ALONSO-RIVERA, A.; CRUZ-AKÉ, S.; VENEGAS-MARTÍNEZ, F. Impact of monetary policy on financial markets efficiency and speculative bubbles: A non-linear entropy-based approach. 2014. Citado na página 14.

ANATOLYEV, S.; GOSPODINOV, N. Modeling financial return dynamics via decomposition. **Journal of Business & Economic Statistics**, Taylor & Francis, 2012. Citado na página 4.

ATSALAKIS, G.; K, V. Surveying stock market forecasting techniques - part i: Conventional methods. **Computation optimization in economics and finance research compendium**, p. 49–104, 01 2013. Citado 2 vezes nas páginas 4 e 17.

ATSALAKIS, G. S.; VALAVANIS, K. P. Surveying stock market forecasting techniques—part ii: Soft computing methods. **Expert Systems with Applications**, Elsevier, v. 36, n. 3, p. 5932–5941, 2009. Citado 2 vezes nas páginas 4 e 17.

BARKOULAS, J. T.; BAUM, C. F. Long-term dependence in stock returns. **Economics Letters**, Elsevier, v. 53, n. 3, p. 253–259, 1996. Citado 2 vezes nas páginas 1 e 14.

BELAIRE-FRANCH, J.; OPONG, K. K. Testing for random walk in euro exchange rates using the subsampling approach. **Applied Economics Letters**, Taylor & Francis, v. 17, n. 12, p. 1145–1151, 2010. Citado 2 vezes nas páginas 1 e 14.

BM&FBOVESPA. **Metodologia do índice Bovespa**. 2015. Disponível em: <http://www.b3.com.br/pt_br/market-data-e-indices/indices/indices-amplos/ibovespa.htm>. Citado na página 21.

BM&FBOVESPA. **Taxa média de crescimento Ibovespa**. 2019. Disponível em: <http://www.b3.com.br/pt_br/market-data-e-indices/indices/indices-amplos/indice-ibovespa-ibovespa-estatisticas-historicas.htm>. Citado 2 vezes nas páginas 89 e 121.

BM&FBOVESPA. **Histórico pessoas físicas**. 2020. Disponível em: <http://www.b3.com.br/pt_br/market-data-e-indices/servicos-de-dados/market-data/consultas/mercado-a-vista/historico-pessoas-fisicas/>. Citado na página 2.

BM&FBOVESPA. **Negociação- Segmento BM&FBOVESPA**. 2020. Disponível em: http://www.b3.com.br/pt_br/regulacao/estrutura-normativa/regulamentos-e-manuais/negociacao.htm. Citado na página 21.

BORDO, M. D.; LANDON-LANE, J. **Does expansionary monetary policy cause asset price booms; some historical and empirical evidence**. [S.l.], 2013. Citado na página 14.

BRUNNERMEIER, M. K. Information leakage and market efficiency. **Review of Financial Studies**, Soc Financial Studies, v. 18, n. 2, p. 417–457, 2005. Citado 3 vezes nas páginas 1, 2 e 14.

CAGINALP, G. Does the market have a mind of its own, and does it get carried away with excess cash? **The Journal of Psychology and Financial Markets**, Taylor & Francis, v. 3, n. 2, p. 72–75, 2002. Citado na página 19.

CAGINALP, G.; DESANTIS, M. A paradigm for quantitative behavioral finance. **American Behavioral Scientist**, Sage Publications, v. 55, n. 8, p. 1014–1034, 2011. Citado 3 vezes nas páginas 1, 14 e 19.

CAGINALP, G.; ERMENTROUT, G. A kinetic thermodynamics approach to the psychology of fluctuations in financial markets. **Applied Mathematics Letters**, v. 3, n. 4, p. 17 – 19, 1990. ISSN 0893-9659. Citado 2 vezes nas páginas 2 e 16.

CAGINALP, G.; PORTER, D.; SMITH, V. Momentum and overreaction in experimental asset markets. **International Journal of Industrial Organization**, Elsevier, v. 18, n. 1, p. 187–204, 2000. Citado na página 19.

CAJUEIRO, D. O.; TABAK, B. M. The hurst exponent over time: testing the assertion that emerging markets are becoming more efficient. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 336, n. 3, p. 521–537, 2004. Citado 6 vezes nas páginas 1, 5, 14, 17, 19 e 52.

CARBONE, A.; CASTELLI, G.; STANLEY, H. E. Time-dependent hurst exponent in financial time series. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 344, n. 1-2, p. 267–271, 2004. Citado 2 vezes nas páginas 16 e 19.

CHARLES, A.; DARNÉ, O.; KIM, J. H. Exchange-rate return predictability and the adaptive markets hypothesis: Evidence from major foreign exchange rates. **Journal of International Money and Finance**, Elsevier, v. 31, n. 6, p. 1607–1626, 2012. Citado 2 vezes nas páginas 1 e 14.

CHEN, P.-w.; HUANG, H.-c.; SU, Y.-c. The central bank in market efficiency: The case of taiwan. **Pacific-Basin Finance Journal**, Elsevier, v. 29, p. 239–260, 2014. Citado na página 14.

CHEN, Z. et al. Effect of nonstationarities on detrended fluctuation analysis. **Physical review. E, Statistical, nonlinear, and soft matter physics**, v. 65, p. 041107, 05 2002. Citado na página 48.

CHEVAPATRAKUL, T. Return sign forecasts based on conditional risk: Evidence from the uk stock market index. **Journal of Banking & Finance**, Elsevier, v. 37, n. 7, p. 2342–2353, 2013. Citado na página 4.

CHORDIA, T.; ROLL, R.; SUBRAHMANYAM, A. Liquidity and market efficiency. **Journal of Financial Economics**, Elsevier, v. 87, n. 2, p. 249–268, 2008. Citado 3 vezes nas páginas 1, 2 e 14.

CHOUDHRY, T.; JAYASEKERA, R. Comparison of efficiency characteristics between the banking sectors of us and uk during the global financial crisis of 2007–2011. **International Review of Financial Analysis**, Elsevier, v. 25, p. 106–116, 2012. Citado 2 vezes nas páginas 1 e 14.

CHOUDHRY, T.; JAYASEKERA, R. Market efficiency during the global financial crisis: Empirical evidence from european banks. **Journal of international Money and Finance**, Elsevier, v. 49, p. 299–318, 2014. Citado 2 vezes nas páginas 1 e 14.

CHUNG, D.; HRAZDIL, K. Liquidity and market efficiency: A large sample study. **Journal of Banking & Finance**, Elsevier, v. 34, n. 10, p. 2346–2357, 2010. Citado 2 vezes nas páginas 2 e 14.

COHEN, J. **Statistical power analysis for the behavioral sciences**. [S.l.]: New York: Academic Press, 1977. Citado 2 vezes nas páginas 52 e 62.

CURME, C.; STANLEY, H. E.; VODENSKA, I. Coupled network approach to predictability of financial market returns and news sentiments. **International Journal of Theoretical and Applied Finance**, World Scientific, v. 18, n. 07, p. 1550043, 2015. Citado 2 vezes nas páginas 1 e 14.

FAN, Y. et al. The effect of investor psychology on the complexity of stock market: An analysis based on cellular automaton model. **Computers & Industrial Engineering**, Elsevier, v. 56, n. 1, p. 63–69, 2009. Citado na página 16.

FENG, L. et al. Linking agent-based models and stochastic models of financial markets. **Proceedings of the National Academy of Sciences**, National Acad Sciences, v. 109, n. 22, p. 8388–8393, 2012. Citado na página 1.

GARCIA, M. et al. Forecast model for financial time series: An approach based on harmonic oscillators. **Physica A: Statistical Mechanics and its Applications**, Elsevier, p. 124365, 2020. Citado 2 vezes nas páginas 17 e 88.

GRECH, D.; MAZUR, Z. Can one make any crash prediction in finance using the local hurst exponent idea? **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 336, n. 1, p. 133–145, 2004. Citado 2 vezes nas páginas 1 e 14.

HEALY, P. M.; PALEPU, K. G. Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. **Journal of accounting and economics**, Elsevier, v. 31, n. 1, p. 405–440, 2001. Citado 3 vezes nas páginas 1, 2 e 14.

HOMMES, C.; KOPÁNYI-PEUKER, A.; SONNEMANS, J. Bubbles, crashes and information contagion in large-group asset market experiments. **Experimental Economics**, Springer, p. 1–20, 2020. Citado na página 16.

HU, K. et al. Effect of trends on detrended fluctuation analysis. **Phys. Rev. E**, American Physical Society, v. 64, p. 011114, Jun 2001. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevE.64.011114>>. Citado na página 48.

KAHNEMAN, D. et al. Noise: How to overcome the high, hidden cost of inconsistent decision making. **Harvard Business Review**, v. 94, p. 38–46, 2016. Citado na página 16.

KAHNEMAN, D.; TVERSKY, A. Prospect theory: An analysis of decision under risk. **Econometrica: Journal of the econometric society**, JSTOR, v. 47, n. 2, p. 263–291, 1979. Citado 3 vezes nas páginas 2, 14 e 15.

KATUSIIME, L.; SHAMSUDDIN, A.; AGBOLA, F. W. Foreign exchange market efficiency and profitability of trading rules: Evidence from a developing country. **International Review of Economics & Finance**, Elsevier, v. 35, p. 315–332, 2015. Citado 2 vezes nas páginas 1 e 14.

LAZĂR, D.; TODEA, A.; FILIP, D. Martingale difference hypothesis and financial crisis: Empirical evidence from european emerging foreign exchange markets. **Economic Systems**, Elsevier, v. 36, n. 3, p. 338–350, 2012. Citado 2 vezes nas páginas 1 e 14.

LIM, K.-P.; BROOKS, R. D.; KIM, J. H. Financial crisis and stock market efficiency: Empirical evidence from asian countries. **International Review of Financial Analysis**, Elsevier, v. 17, n. 3, p. 571–591, 2008. Citado na página 14.

LIU, S. Transaction costs and market efficiency: Evidence from commission deregulation. **The Quarterly Review of Economics and Finance**, Elsevier, v. 50, n. 3, p. 352–360, 2010. Citado na página 14.

MALKIEL, B. G.; FAMA, E. F. Efficient capital markets: A review of theory and empirical work. **The journal of Finance**, Wiley Online Library, v. 25, n. 2, p. 383–417, 1970. Citado na página 1.

MANTEGNA, R. N.; STANLEY, H. E. **Introduction to econophysics: correlations and complexity in finance**. [S.l.]: Cambridge university press, 1999. Citado 2 vezes nas páginas 13 e 14.

MEI, D. et al. Forecasting stock market volatility: Do realized skewness and kurtosis help? **Physica A: Statistical Mechanics and its Applications**, v. 481, p. 153–159, 2017. Citado na página 18.

MERKLE, C.; MÜLLER-DETHARD, J.; WEBER, M. Closing a mental account: The realization effect for gains and losses. **Experimental Economics**, Springer, p. 1–27, 2020. Citado na página 16.

MONTGOMERY, D. C. **Design and analysis of experiments**. [S.l.]: John Wiley & Sons, 2008. Citado na página 42.

NAMETALA, C. et al. Uma estratégia automatizada de investimento por meio de redes neurais artificiais e preditores econométricos. In: **Anais Principais do XII Simpósio Brasileiro de Sistemas de Informação**. Porto Alegre, RS, Brasil: SBC, 2016. p. 152–159. ISSN 0000-0000. Disponível em: <<https://sol.sbc.org.br/index.php/sbsi/article/view/5957>>. Citado 2 vezes nas páginas 2 e 18.

NELSON, D. M.; PEREIRA, A. C.; OLIVEIRA, R. A. de. Stock market's price movement prediction with lstm neural networks. In: IEEE. **2017 International joint conference on neural networks (IJCNN)**. [S.l.], 2017. p. 1419–1426. Citado na página 18.

NTIM, C. G. et al. On the efficiency of the global gold markets. **International Review of Financial Analysis**, Elsevier, v. 41, p. 218–236, 2015. Citado 2 vezes nas páginas 1 e 14.

NYBERG, H. Forecasting the direction of the us stock market with dynamic binary probit models. **International Journal of Forecasting**, Elsevier, v. 27, n. 2, p. 561–578, 2011. Citado na página 4.

OLIVEIRA, F. A. de; NOBRE, C. N.; ZARATE, L. E. Applying artificial neural networks to prediction of stock price and improvement of the directional prediction index—case study of petr4, petrobras, brazil. **Expert Systems with Applications**, Elsevier, v. 40, n. 18, p. 7596–7606, 2013. Citado na página 17.

PASQUARIELLO, P.; ZAFEIRIDOU, C. Political uncertainty and financial market quality. **Ross School of Business Paper**, n. 1232, 2014. Citado na página 14.

PENG, C.-K. et al. Mosaic organization of dna nucleotides. **Phys. Rev. E**, American Physical Society, v. 49, p. 1685–1689, Feb 1994. Citado na página 48.

PODOBNIK, B. et al. Fractionally integrated process for transition economics. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 362, n. 2, p. 465–470, 2006. Citado 2 vezes nas páginas 1 e 14.

PÖNKÄ, H. Predicting the direction of us stock markets using industry returns. **Empirical Economics**, Springer, p. 1–30, 2014. Citado na página 4.

RAFFERTY, M.; TOMLJANOVICH, M. Central bank transparency and market efficiency: An econometric analysis. **Journal of Economics and Finance**, Springer, v. 26, n. 2, p. 150–161, 2002. Citado na página 14.

RECHENTHIN, M.; STREET, W. N. Using conditional probability to identify trends in intra-day high-frequency equity pricing. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 392, n. 24, p. 6169–6188, 2013. Citado na página 17.

RESENDE, C. C. de et al. Investigating market efficiency through a forecasting model based on differential equations. **Physica A: Statistical Mechanics and its Applications**, v. 474, n. C, p. 199–212, 2017. Citado 6 vezes nas páginas 1, 8, 10, 28, 31 e 35.

ROSS, S. A. Information and volatility: The no-arbitrage martingale approach to timing and resolution irrelevancy. **The Journal of Finance**, Wiley Online Library, v. 44, n. 1, p. 1–17, 1989. Citado 3 vezes nas páginas 1, 2 e 14.

SAMUELSON, P. A. **Proof that properly anticipated prices fluctuate randomly**. [S.l.: s.n.], 1965. 41–45 p. Citado 2 vezes nas páginas 1 e 13.

SHARMA, C.; BANERJEE, K. A study of correlations in the stock market. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 432, p. 321–330, 2015. Citado na página 14.

SHILLER, R. J. **Do stock prices move too much to be justified by subsequent changes in dividends?** [S.l.]: National Bureau of Economic Research Cambridge, Mass., USA, 1980. 421–436 p. Citado na página 1.

SHILLER, R. J.; FISCHER, S.; FRIEDMAN, B. M. Stock prices and social dynamics. **Brookings papers on economic activity**, JSTOR, v. 1984, n. 2, p. 457–510, 1984. Citado 2 vezes nas páginas 2 e 14.

TVERSKY, A.; KAHNEMAN, D. Loss Aversion in Riskless Choice: A Reference-Dependent Model*. **The Quarterly Journal of Economics**, v. 106, n. 4, p. 1039–1061, 11 1991. ISSN 0033-5533. Disponível em: <<https://doi.org/10.2307/2937956>>. Citado 2 vezes nas páginas 15 e 16.

ZUNINO, L. et al. Inefficiency in latin-american market indices. **The European Physical Journal B**, Springer, v. 60, n. 1, p. 111–121, 2007. Citado 3 vezes nas páginas 1, 5 e 14.

Apêndices

APÊNDICE A – Estimação da matriz de coeficientes pelo método dos mínimos quadrados

Seja a relação

$$\mathbf{x}(t) = \mathbf{A}\mathbf{x}(t-1), \quad (52)$$

onde

$$\mathbf{x}(t) = \begin{pmatrix} X_1(t) \\ X_2(t) \\ \vdots \\ X_n(t) \end{pmatrix} \quad \text{e} \quad \mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}. \quad (53)$$

Esta equação pode ser escrita na forma

$$X_i(t) = \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix} \begin{pmatrix} X_1(t-1) \\ X_2(t-1) \\ \vdots \\ X_n(t-1) \end{pmatrix}, \quad i = 1, 2, \dots, n, \quad (54)$$

ou, equivalentemente,

$$X_i(t) = \begin{pmatrix} X_1(t-1) & X_2(t-1) & \cdots & X_n(t-1) \end{pmatrix} \begin{pmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{in} \end{pmatrix}. \quad (55)$$

Consideremos múltiplas observações para cada $X_i(t)$:

$$\begin{aligned}
 X_i(t) &= \begin{pmatrix} X_1(t-1) & X_2(t-1) & \cdots & X_n(t-1) \end{pmatrix} \begin{pmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{in} \end{pmatrix}, \\
 X_i(t-1) &= \begin{pmatrix} X_1(t-2) & X_2(t-2) & \cdots & X_n(t-2) \end{pmatrix} \begin{pmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{in} \end{pmatrix}, \\
 &\vdots \\
 X_i(t-N+1) &= \begin{pmatrix} X_1(t-N) & X_2(t-N) & \cdots & X_n(t-N) \end{pmatrix} \begin{pmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{in} \end{pmatrix} \quad (56)
 \end{aligned}$$

Observe que estamos assumindo que o vetor de parâmetros não muda dentro da faixa de tempos analisada. Estas equações podem ser escritas na forma

$$\begin{pmatrix} X_i(t-N+1) \\ \vdots \\ X_i(t-1) \\ X_i(t) \end{pmatrix} = \begin{pmatrix} X_1(t-N) & X_2(t-N) & \cdots & X_n(t-N) \\ \vdots & \vdots & \cdots & \vdots \\ X_1(t-2) & X_2(t-2) & \cdots & X_n(t-2) \\ X_1(t-1) & X_2(t-1) & \cdots & X_n(t-1) \end{pmatrix} \begin{pmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{in} \end{pmatrix}. \quad (57)$$

Se $N > n$, temos, para cada valor fixo de i , mais equações que valores de a_{ij} ($j = 1, 2, \dots, n$) para estimar. É natural, então, que ao estimarmos os parâmetros a_{ij} cometamos erros. Definindo os vetores de parâmetros estimados e de erros (resíduos) na forma

$$\hat{\mathbf{a}}_i = \begin{pmatrix} \hat{a}_{i1} \\ \hat{a}_{i2} \\ \vdots \\ \hat{a}_{in} \end{pmatrix} \quad \text{e} \quad \boldsymbol{\xi}_i = \begin{pmatrix} \xi_i(t) \\ \xi_i(t-1) \\ \vdots \\ \xi_i(t-N+1) \end{pmatrix}, \quad i = 1, 2, \dots, n, \quad (58)$$

respectivamente, a relação (57) torna-se

$$\begin{pmatrix} X_i(t-N+1) \\ \vdots \\ X_i(t-1) \\ X_i(t) \end{pmatrix} = \begin{pmatrix} X_1(t-N) & X_2(t-N) & \cdots & X_n(t-N) \\ \vdots & \vdots & \cdots & \vdots \\ X_1(t-2) & X_2(t-2) & \ddots & X_n(t-2) \\ X_1(t-1) & X_2(t-1) & \cdots & X_n(t-1) \end{pmatrix} \begin{pmatrix} \hat{a}_{i1} \\ \hat{a}_{i2} \\ \vdots \\ \hat{a}_{in} \end{pmatrix} + \begin{pmatrix} \xi_i(t-N+1) \\ \vdots \\ \xi_i(t-1) \\ \xi_i(t) \end{pmatrix}. \quad (59)$$

Numa forma compacta:

$$\mathbf{y}_i = \mathbf{X}\hat{\mathbf{a}}_i + \xi_i, \quad (60)$$

onde

$$\mathbf{y}_i = \begin{pmatrix} X_i(t-N+1) \\ \vdots \\ X_i(t-1) \\ X_i(t) \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} X_1(t-N) & X_2(t-N) & \cdots & X_n(t-N) \\ \vdots & \vdots & \cdots & \vdots \\ X_1(t-2) & X_2(t-2) & \ddots & X_n(t-2) \\ X_1(t-1) & X_2(t-1) & \cdots & X_n(t-1) \end{pmatrix}$$

e $i = 1, 2, \dots, n$.

Calculando a transposta em ambos os lados da equação (59), encontramos

$$\begin{pmatrix} X_i(t-N+1) & \cdots & X_i(t) \end{pmatrix} = \begin{pmatrix} \hat{a}_{i1} & \hat{a}_{i2} & \cdots & \hat{a}_{in} \end{pmatrix} \begin{pmatrix} X_1(t-N) & \cdots & X_1(t-1) \\ X_2(t-N) & \cdots & X_2(t-1) \\ \vdots & \ddots & \vdots \\ X_n(t-N) & \cdots & X_n(t-1) \end{pmatrix} + \begin{pmatrix} \xi_i(t-N+1) & \cdots & \xi_i(t) \end{pmatrix}. \quad (61)$$

Agregando as equações acima, válidas para $i = 1, 2, \dots, n$, chegamos a

$$\begin{pmatrix} X_1(t-N+1) & \cdots & X_1(t) \\ X_2(t-N+1) & \cdots & X_2(t) \\ \vdots & \ddots & \vdots \\ X_n(t-N+1) & \cdots & X_n(t) \end{pmatrix} = \begin{pmatrix} \hat{a}_{11} & \cdots & \hat{a}_{1n} \\ \hat{a}_{21} & \cdots & \hat{a}_{2n} \\ \vdots & \ddots & \vdots \\ \hat{a}_{n1} & \cdots & \hat{a}_{nn} \end{pmatrix} \begin{pmatrix} X_1(t-N) & \cdots & X_1(t-1) \\ X_2(t-N) & \cdots & X_2(t-1) \\ \vdots & \ddots & \vdots \\ X_n(t-N) & \cdots & X_n(t-1) \end{pmatrix} + \begin{pmatrix} \xi_1(t-N+1) & \cdots & \xi_1(t) \\ \xi_2(t-N+1) & \cdots & \xi_2(t) \\ \vdots & \ddots & \vdots \\ \xi_n(t-N+1) & \cdots & \xi_n(t) \end{pmatrix}. \quad (62)$$

Explicitando a matriz de erros:

$$\begin{pmatrix} \xi_1(t-N+1) & \cdots & \xi_1(t) \\ \xi_2(t-N+1) & \cdots & \xi_2(t) \\ \vdots & \ddots & \vdots \\ \xi_n(t-N+1) & \cdots & \xi_n(t) \end{pmatrix} = \begin{pmatrix} X_1(t-N+1) & \cdots & X_1(t) \\ X_2(t-N+1) & \cdots & X_2(t) \\ \vdots & \ddots & \vdots \\ X_n(t-N+1) & \cdots & X_n(t) \end{pmatrix} - \begin{pmatrix} \hat{a}_{11} & \cdots & \hat{a}_{1n} \\ \hat{a}_{21} & \cdots & \hat{a}_{2n} \\ \vdots & \ddots & \vdots \\ \hat{a}_{n1} & \cdots & \hat{a}_{nn} \end{pmatrix} \begin{pmatrix} X_1(t-N) & \cdots & X_1(t-1) \\ X_2(t-N) & \cdots & X_2(t-1) \\ \vdots & \ddots & \vdots \\ X_n(t-N) & \cdots & X_n(t-1) \end{pmatrix}. \quad (63)$$

Notemos que, para i fixo, os resíduos $\xi_i(\tau)$ ($\tau = t, t-1, \dots, t-N+1$) só dependem dos coeficientes \hat{a}_{ij} ($j = 1, 2, \dots, n$). Assim, podemos encontrar os valores dos \hat{a}_{ij} que minimizam a soma dos quadrados dos $\xi_i(\tau)$,

$$S_i = \sum_{\tau=t-N+1}^t (\xi_i(\tau))^2, \quad (64)$$

para cada valor de i , de forma independente. A soma dos quadrados de todos os resíduos,

$$S = \sum_{i=1}^n S_i, \quad (65)$$

será minimizada utilizando-se os \hat{a}_{ij} calculados para todos os valores de i . Buscaremos, então, a minimização dos S_i , separadamente. Conforme descrito em (AGUIRRE, 2007) a matriz que realiza essa minimização é dada pela equação

$$\hat{\mathbf{a}}_i = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}_i. \quad (66)$$

Definindo a matriz

$$\mathbf{M} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T = \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1N} \\ M_{21} & M_{22} & \cdots & M_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n1} & M_{n2} & \cdots & M_{nN} \end{pmatrix}, \quad (67)$$

podemos escrever

$$\begin{pmatrix} \hat{a}_{i1} \\ \hat{a}_{i2} \\ \vdots \\ \hat{a}_{in} \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1N} \\ M_{21} & M_{22} & \cdots & M_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n1} & M_{n2} & \cdots & M_{nN} \end{pmatrix} \begin{pmatrix} X_i(t-N+1) \\ \vdots \\ X_i(t-1) \\ X_i(t) \end{pmatrix}, \quad i = 1, 2, \dots, n, \quad (68)$$

relação que pode ser colocada na forma

$$\begin{pmatrix} \hat{a}_{11} & \cdots & \hat{a}_{n1} \\ \hat{a}_{12} & \cdots & \hat{a}_{n2} \\ \vdots & \ddots & \vdots \\ \hat{a}_{1n} & \cdots & \hat{a}_{nn} \end{pmatrix} = \begin{pmatrix} M_{11} & \cdots & M_{1N} \\ M_{21} & \cdots & M_{2N} \\ \vdots & \ddots & \vdots \\ M_{n1} & \cdots & M_{nN} \end{pmatrix} \begin{pmatrix} X_1(t-N+1) & \cdots & X_n(t-N+1) \\ \vdots & \cdots & \vdots \\ X_1(t-1) & \cdots & X_n(t-1) \\ X_1(t) & \cdots & X_n(t) \end{pmatrix}. \quad (69)$$

Observe que a matriz representada no lado esquerdo da equação acima é a matriz transposta da matriz $\hat{\mathbf{A}}$, correspondente ao valor estimado para \mathbf{A} que minimiza a soma dos quadrados dos resíduos. Definamos

$$\mathbf{X}_N(t) = \begin{pmatrix} X_1(t-N+1) & \cdots & X_1(t-1) & X_1(t) \\ X_2(t-N+1) & \cdots & X_2(t-1) & X_2(t) \\ \vdots & \ddots & \vdots & \vdots \\ X_n(t-N+1) & \cdots & X_n(t-1) & X_n(t) \end{pmatrix}, \quad (70)$$

para escrevermos a equação (69) como

$$\hat{\mathbf{A}}^T = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X}_N^T(t). \quad (71)$$

Observando que

$$\mathbf{X} = \mathbf{X}_N^T(t-1), \quad (72)$$

obtemos a seguinte expressão para $\hat{\mathbf{A}}$:

$$\hat{\mathbf{A}} = (\mathbf{X}_N(t) \mathbf{X}_N^T(t-1)) (\mathbf{X}_N(t-1) \mathbf{X}_N^T(t-1))^{-1}. \quad (73)$$

APÊNDICE B – Tabelas lotes padrão das ações

Tabela 24 – Tabela da quantidade de ações que compõem o lote padrão de negociação no mercado real e na plataforma de simulação no ano de 2017

Empresa	Código de negociação	Lote na plataforma	Lote Bovespa
Vale on	VALE3	100	100
Itaú pn	ITUB4	150	100
Vale pn	VALE5	100	100
Bradesco pn	BBDC4	145	100
Fundo Ibovespa	BOVA11	10	10
Ambev on	ABEV3	100	100
Petrobras pn	PETR4	100	100
Cielo on	CIEL3	120	100
Petrobras on	PETR3	100	100
Banco do Brasil on	BBAS3	100	100

Nota: Apresentamos o lote padrão de negociação definido no mercado real e também na plataforma de simulação. As informações são referentes às 10 ações com maior liquidez no ano de 2017. Os dados destacados na cor azul referem-se às ações que não compõem a lista das 10 ações com maior liquidez no ano de 2018. Já os dados em destaque na cor verde referem-se às ações que têm o lote padrão de negociação no mercado real diferente do lote de negociação definido na plataforma. O código pn é referente as ações preferencias e on as ações ordinárias das empresas. Os dados de negociação das ações estão disponíveis em: http://www.b3.com.br/pt_br/market-data-e-indices/servicos-de-dados/market-data/consultas/mercado-a-vista/lotes-de-negociacao/

Tabela 25 – Tabela da quantidade de ações que compõem o lote padrão de negociação no mercado real e na plataforma de simulação no ano de 2018

Empresa	Código de negociação	Lote na plataforma	Lote Bovespa
Vale on	VALE3	100	100
Itaú pn	ITUB4	150	100
B3 on	B3SA3	100	100
Bradesco pn	BBDC4	145	100
Fundo Ibovespa	BOVA11	10	10
Ambev on	ABEV3	100	100
Petrobras pn	PETR4	100	100
Magazine Luiza on	MGLU3	100	100
Itaú on	ITUB3	150	100
Banco do Brasil on	BBAS3	100	100

Nota: Apresentamos o lote padrão de negociação definido no mercado real e também na plataforma de simulação. As informações são referentes às 10 ações com maior liquidez no ano de 2018. Os dados destacados na cor azul referem-se às ações que não compõem a lista das 10 ações com maior liquidez no ano de 2017. Já os dados em destaque na cor verde referem-se às ações que têm o lote padrão de negociação no mercado real diferente do lote de negociação definido na plataforma. O código pn é referente as ações preferencias e on as ações ordinárias das empresas. Os dados de negociação das ações estão disponíveis em: <http://www.b3.com.br/pt_br/market-data-e-indices/servicos-de-dados/market-data/consultas/mercado-a-vista/lotos-de-negociacao/>

Anexos

**ANEXO A – Tabela de ações que
participaram da carteira teórica do
Ibovespa no primeiro quadrimestre de
2017**

Tabela 26 – Composição da carteira teórica do Ibovespa no primeiro quadrimestre de 2017

Virada do Ibovespa para Janeiro a Abril de 2017			
CÓDIGO	AÇÃO	TIPO	PARTICIPAÇÃO (%)
ITUB4	ITAUNIBANCO	PN EDJ N1	10,681
BBDC4	BRADESCO	PN EJ N1	7,7402
ABEV3	AMBEV S/A	ON EJ	7,1825
PETR4	PETROBRAS	PN	6,0083
PETR3	PETROBRAS	ON	4,5918
VALE5	VALE	PNA N1	4,5442
VALE3	VALE	ON N1	3,776
BRFS3	BRF SA	ON NM	3,7215
BBAS3	BRASIL	ON NM	3,4462
ITSA4	ITAUSA	PN EJ N1	3,1492
BVMF3	BMFBOVESPA	ON EJ NM	2,9414
UGPA3	ULTRAPAR	ON NM	2,7619
CIEL3	CIELO	ON NM	2,5919
BBSE3	BBSEGURIDADE	ON NM	1,9019
KROT3	KROTON	ON NM	1,8728
VIVT4	TELEF BRASIL	PN EJ	1,8181
JBSS3	JBS	ON NM	1,7712
BBDC3	BRADESCO	ON EJ N1	1,6499
LREN3	LOJAS RENNEN	ON NM	1,4737
CCRO3	CCR SA	ON NM	1,3755
RADL3	RAIADROGASIL	ON NM	1,2199
EMBR3	EMBRAER	ON NM	1,1756
CTIP3	CETIP	ON EJ NM	1,1596
SANB11	SANTANDER BR	UNT	1,1399
EQTL3	EQUATORIAL	ON EJ NM	1,0802
HYPE3	HYPERMARCAS	ON NM	1,0534
SBSP3	SABESP	ON NM	0,9795
GGBR4	GERDAU	PN N1	0,9264
BRKM5	BRASKEM	PNA N1	0,9068
LAME4	LOJAS AMERIC	PN	0,893
WEGE3	WEG	ON NM	0,884
PCAR4	P.ACUCAR-CBD	PN N1	0,8504
CPFE3	CPFL ENERGIA	ON NM	0,8192
KLBN11	KLABIN S/A	UNT N2	0,7869
FIBR3	FIBRIA	ON NM	0,7324
EGIE3	ENGIE BRASIL	ON NM	0,7139
BRML3	BR MALLS PAR	ON NM	0,6907
CSNA3	SID NACIONAL	ON	0,6527
CMIG4	CEMIG	PN EJ N1	0,6394
TIMP3	TIM PART S/A	ON NM	0,6333
ELET3	ELETROBRAS	ON N1	0,5955
CSAN3	COSAN	ON NM	0,5843
SUZB5	SUZANO PAPEL	PNA N1	0,5522
RENT3	LOCALIZA	ON EJ NM	0,5095
ESTC3	ESTACIO PART	ON ED NM	0,4868
MULT3	MULTIPLAN	ON N2	0,4651
RUMO3	RUMO LOG	ON NM	0,4619
QUAL3	QUALICORP	ON NM	0,4229
NATU3	NATURA	ON NM	0,3964
BRAP4	BRADESPAR	PN EJ N1	0,3292
ENBR3	ENERGIAS BR	ON EJ NM	0,3119
MRVE3	MRV	ON NM	0,3022
GOAU4	GERDAU MET	PN N1	0,2877
CPL6	COPEL	PNB N1	0,2765
SMLE3	SMILES	ON EJ NM	0,2552
CYRE3	CYRELA REALT	ON NM	0,2516
MRFG3	MARFRIG	ON NM	0,21
USIM5	USIMINAS	PNA N1	0,1975
ECOR3	ECORODOVIAS	ON NM	0,164

Adaptado do site BM&FBOVESPA.

ANEXO B – Tabela de ações que participaram da carteira teórica do Dow Jones no primeiro quadrimestre de 2017

Tabela 27 – Composição do índice Dow Jones no primeiro quadrimestre de 2017

NÚMERO	AÇÃO	CÓDIGO	PARTICIPAÇÃO (%)
1	Boeing	BA	7,28
2	Goldman Sachs	GS	7,18
3	3M	MMM	6,41
4	UnitedHealth Group	UNH	6,05
5	Apple	AAPL	5,00
6	Home Depot	HD	4,82
7	McDonalds	MCD	4,81
8	IBM	IBM	4,40
9	Johnson & Johnson	JNJ	4,14
10	Travelers	TRV	4,04
11	United Technologies	UTX	3,64
12	Caterpillar	CAT	3,55
13	Chevron	CVX	3,44
14	Disney	DIS	3,19
15	Visa A	V	3,14
16	JPMorgan Chase	JPM	2,91
17	Procter & Gamble	PG	2,86
18	American Express	AXP	2,66
19	Wal Mart	WMT	2,53
20	Dupont	DD	2,53
21	ExxonMobil	XOM	2,49
22	Microsoft	MSFT	2,25
23	Merck	MRK	1,94
24	NIKE B	NKE	1,86
25	Verizon Communications	VZ	1,50
26	Coca-Cola	KO	1,42
27	Intel	INTC	1,14
28	Pfizer	PFE	1,04
29	Cisco Systems	CSCO	0,98
30	General Electric	GE	0,80

Adaptado do site www.indexarb.com.

ANEXO C – Tabela das taxas anuais de crescimento do Ibovespa

Reproduzimos neste anexo a tabela das taxas anuais de crescimento do Ibovespa desde o ano de 1968. A taxa de crescimento refere-se à taxa média de variação do índice composta anualmente, entre dois períodos, disponível em ([BM&FBOVESPA, 2019](#)).

