

# ANÁLISE TEMPORAL DO ÍNDICE SMALL CAP (SMLL) USANDO OS MODELOS ARMA-ARCH

André Gondim Carlucci – andre\_carlucci@hotmail.com

Matheus Castro de Paula – matheuscastrodepaula@gmail.com

Paulo Henrique de Araujo Lima – paulo.h\_lima@hotmail.com

Orlando Yesid Esparza Albarracin (orientador) – orlando.albarracin@mackenzie.br

## RESUMO

Este artigo estuda o índice *Small Cap* da B3 durante o período de janeiro de 2016 a junho de 2021, com base na teoria de séries temporais. Esse índice, composto por empresas de baixa capitalização da bolsa de valores brasileira, é pouco abordado na literatura, e são poucos os trabalhos que desenvolveram modelos estatísticos que fossem adequados à série. Nesse estudo, foram calculados os log-retornos do índice e em seguida proposto um modelo ARMA-ARCH que fosse adequado na modelagem da média e da volatilidade da série transformada. A análise de resíduos do modelo comprovou sua eficácia, mostrando que ele foi devidamente ajustado à série e que capturou o comportamento dos log-retornos. Além da análise de resíduos, foi feita a utilização do modelo para estimar valores do índice que pudessem ser comparados com dados observados, que mostraram, para um período de curto prazo, resultados satisfatórios.

Palavras-chave: *Small Cap*. Modelagem. ARMA. ARCH.

## 1 INTRODUÇÃO

O *Small Cap* (SMLL) é um índice feito pela B3 que existe desde 2005 e funciona como indicador de desempenho médio de empresas de menor capitalização da bolsa brasileira (B3, 2021). Os ativos presentes no índice são selecionados seguindo uma metodologia, de forma que não façam parte das empresas com os maiores valores de mercado negociadas na bolsa, portanto, o índice tem como objetivo o auxílio no acompanhamento do comportamento de *small caps* brasileiras (B3, 2021). Esse índice se mostra relevante ao representar o comportamento dessas empresas, uma vez que em sua composição estão empresas de setores diversos e relevantes para a economia, como energia elétrica, serviços médicos e exploração de imóveis. Além disso, o índice SMLL é usado como lastro por um dos fundos de índice que está entre os mais negociados no último ano (2020) na bolsa de valores, o fundo SMAL11 (MorningStar e Economatica, 2021).

Apesar de sua relevância, na literatura são poucos os trabalhos que têm modelado o comportamento do índice nos últimos anos com o intuito de realizar previsões. Em contrapartida, diversos trabalhos têm analisado o comportamento e a volatilidade do índice Bovespa (IBOVESPA) objetivando a construção de modelos probabilísticos para fazer previsões a curto prazo. Destacam-se

nesse contexto os trabalhos de Nogueira Júnior e Kobunda (2019), Filho *et al.* (2011) e Morais e Portugal (1999).

A implementação de modelos estatísticos para analisar do comportamento do SMLL se mostra cada vez mais relevante na atualidade. Destaca-se que a quantidade de pessoas operando ativos da bolsa presentes nesse índice é crescente. Segundo dados informados pela B3 (2021), entre 2016 e o primeiro semestre de 2021, o número de investidores no mercado de capitais foi de 500 mil para 3,2 milhões, o que representa um aumento de 540%. Isso pode ser explicado pelas recorrentes reduções na taxa SELIC desde o final de 2016 e atingimento da mínima histórica em 2020, no patamar de 2% ao ano (Banco Central, 2021), o que pode ter levado os investidores a buscarem novas opções de investimentos para alocarem seus recursos.

Como o índice SMLL possui um valor que muda ao longo do tempo, o estudo a ser desenvolvido irá utilizar a teoria de séries temporais. Este trabalho tem como objetivo analisar o comportamento do índice SMLL no período entre janeiro de 2016 e junho de 2021, visando entender seu comportamento, modelá-lo e fazer previsões a curto prazo. Neste estudo será modelada a série dos log-retornos do índice *Small Cap* usando os modelos ARMA-ARCH, que permitem modelar a volatilidade da série.

## **2 REVISÃO DA LITERATURA**

Nesta seção serão abordados os principais trabalhos em que o índice *Small Cap* foi modelado.

Recentemente o índice *Small Cap* foi modelado por meio de duas abordagens: regressão linear multivariada e Redes Neurais Artificiais (RNAs). Neste estudo, os autores Kaczorowski *et al.* (2021), implementaram algumas variáveis explicativas nos modelos para melhorar as previsões, destacam-se o Índice Nacional de Preços ao Consumidor – Amplo (IPCA), o Índice Geral de Preços (IGP) e o Índice de Volatilidade (VIX). De acordo com os autores, ambos os métodos tiveram resultados satisfatórios para estimar o comportamento do índice. Ao compararem ambas as abordagens, os autores identificaram que o modelo de RNAs foi o que teve o menor erro quadrático médio (46,72 contra 79,33 o modelo multivariado). Destaca-se que no modelo regressão foi necessário retirar observações atípicas. Isso não aconteceu para a abordagem da rede neural, em que o modelo aprendeu com as observações atípicas e apresentou melhores resultados na predição.

Segundo o estudo realizado por Almeida Filho (2015), ativos caracterizados como *small caps* pela B3 podem ser precificados adequadamente por modelos que utilizam a técnica *Bayesian Model Averaging* (BMA). O autor verificou que essa abordagem gerava resultados mais satisfatórios na precificação de *small caps* do que a abordagem do *Capital Asset Pricing Model* (CAPM), mais comumente utilizada. Ao analisar o uso de ambos os modelos, BMA e CAPM, o autor verificou que o primeiro era mais adequado para a precificação de 80% das *small caps* estudadas, segundo métricas

de probabilidade associada aos possíveis modelos. Com isso, verificou que a implementação de variáveis ligadas às expectativas macroeconômicas, como a expectativa da taxa de câmbio, é adequada para uso na precificação desses ativos segundo uma abordagem BMA.

No estudo de Mól *et al.* (2014), em que os autores investigaram a existência de assimetria na estrutura da volatilidade do índice *Small Cap* utilizando a abordagem de modelagem GARCH, verificou-se que o índice apresenta evidências de assimetria. Os autores constataram que a volatilidade do SMLL é maior na ocorrência de choques negativos do que em choques positivos no valor do índice. Além disso, foi observado que a volatilidade de tais choques é persistente e pode levar longos períodos para se dissipar.

Por fim, no trabalho de Araujo e Filho (2012), em que os autores analisaram o comportamento da série dos retornos diários do índice SMLL, verificou-se que após uma grande variação no valor do índice, ocorre forte volatilidade dos retornos. Nesse estudo foi também constatado que períodos de queda nos valores do índice são seguidos de intensa volatilidade, enquanto períodos de elevação são seguidos de volatilidade moderada. Os autores, por meio da aplicação da metodologia ARCH na modelagem do SMLL, atestaram que a variância após um choque demora a convergir para a média. Além disso, o comportamento dessa volatilidade se dá de forma assimétrica, de forma que choques positivos e negativos de mesma magnitude geram impactos diferentes na série.

### 3 METODOLOGIA

#### 3.1 DADOS

Neste trabalho será realizada uma análise exploratória dos log-retornos do índice *Small Cap*. Foram utilizados os valores diários de fechamento do SMLL do período de janeiro de 2016 até junho de 2021, sendo que os dados para estudo foram obtidos na B3 e analisados por meio do software “RStudio”.

O valor do índice é calculado pela B3 ao longo do período regular de negociação a cada intervalo de 30 segundos, considerando o preço das últimas negociações efetuadas no mercado a vista com os ativos componentes da carteira. Segundo a B3, a fórmula utilizada para calcular o valor do índice é descrita como demonstra a Equação 1.

$$\text{Índice}_t = \frac{\sum_{i=1}^n P_{i_t} \times Q_{i_t}}{\alpha} \quad (1)$$

sendo  $n$  o número total de ativos da carteira teórica do índice,  $P_{i_t}$  o preço do ativo  $i$  no momento  $t$ ,  $Q_{i_t}$  a quantidade na carteira do ativo  $i$  no momento  $t$  e  $\alpha$  um fator utilizado para adequar o valor total da carteira ao valor de divulgação do índice (utilizado para que o valor seja de uma grandeza adequada e que não ocorram flutuações bruscas no valor do índice entre as revisões da carteira).

A carteira teórica possui vigência de 4 meses, para os períodos de janeiro a abril, maio a agosto e setembro a dezembro, entrando em vigor na primeira segunda-feira do mês inicial de vigência e tendo seus ativos e pesos revisitados a cada período.

Os ativos elegíveis para participação no SMLL são ações e *units* listadas na B3, excluindo *Brazilian Depositary Receipts* (BDRs) e ativos de companhias em recuperação judicial ou extrajudicial, regime especial de administração temporária ou intervenção. Para fazerem parte do índice, precisam ainda atender aos seguintes critérios de inclusão:

- a) estar entre os ativos que, em ordem decrescente, estejam classificados fora da lista dos que representam 85% do valor de mercado de todas as empresas listadas no mercado a vista da B3;
- b) estar entre os ativos elegíveis que, no período de vigência das 3 carteiras anteriores, em ordem decrescente de ativos mais negociados segundo o Índice de Negociabilidade da B3, façam parte do conjunto dos ativos que representam 99% das negociações;
- c) ter presença em pregão de 95% no período de vigência das 3 carteiras anteriores;
- d) não ser classificado como *Penny Stock* (ativos com cotação média abaixo de R\$ 1,00 durante o período de vigência da carteira anterior).

Um ativo que passa a ser negociado em bolsa durante o período de vigência das 3 carteiras anteriores podem ser elegíveis mesmo sem estar listados durante todo o período, desde que atenda aos seguintes critérios:

- a) a oferta pública de distribuição de ações ou *units*, conforme o caso, tenha sido realizada antes do rebalanceamento imediatamente anterior;
- b) possua 95% de presença em pregão desde seu início de negociação;
- c) atenda cumulativamente aos critérios de inclusão a, b e c.

A ponderação dos ativos pertencentes ao índice é feita considerando o valor de mercado total das ações em *free float* (ativos em livre circulação).

### 3.2 MODELOS

A seguir apresentam-se os principais conceitos da teoria de séries temporais que serão utilizados durante o estudo.

Segundo Morettin e Toloi (2020), uma série temporal é um conjunto de observações ordenadas no tempo, como exemplos podem-se citar o número diário de casos do novo coronavírus e a cotação de uma ação da bolsa de valores no período de um ano.

Em séries financeiras, é preferível trabalhar com retornos, dado que são livres de escalas e possuem propriedades estatísticas mais interessantes, como estacionariedade e ergodicidade (MORETTIN, 2011). Os retornos podem ser calculados conforme descrito na Equação 2.

$$R_t = \frac{Z_t - Z_{t-1}}{Z_{t-1}} \quad (2)$$

sendo  $Z_t$  o valor do índice no momento  $t$ .

Se tratando de séries financeiras, é preferível trabalhar com os log-retornos, dado que existe a vantagem de seguirem aproximadamente uma distribuição Normal (JORION, 1997). Além disso, os log-retornos raramente demonstram tendências ou sazonalidades, salvo algumas exceções eventuais (MORETTIN; TOLOI, 2020). Assim, neste trabalho será modelada a série log-retorno do índice *Small Cap*, que serão calculados conforme a Equação 3.

$$r_t = \ln(Z_t) - \ln(Z_{t-1}) = \ln\left(\frac{Z_t}{Z_{t-1}}\right) \quad (3)$$

No processo de identificação de modelos para séries temporais, a análise da estrutura de dependência serial dos dados é fundamental. Em algumas séries, observa-se que os valores atuais estão correlacionados com os valores passados e a força dessa dependência diminui quando considerados valores mais distantes no tempo. Assim, para a modelagem de uma série temporal, é possível utilizar os modelos auto-regressivos de ordem  $p$ , AR( $p$ ), nos quais o valor da série no instante  $t$  está relacionado com valores defasados. Existem ainda os modelos de médias móveis (MA), que são menos intuitivos que os modelos auto-regressivos, porém são amplamente utilizados na literatura uma vez que requerem, na sua estrutura, menos parâmetros que os modelos AR. Nestes modelos, o valor da série no instante  $t$  depende de erros aleatórios defasados.

Finalmente, os modelos auto-regressivos e de médias móveis é a união de um processo AR( $p$ ) e MA( $q$ ) denominado por ARMA( $p,q$ ), e pode ser descrito conforme a Equação 4.

$$r_t = \beta_0 + \beta_1 r_{t-1} + \dots + \beta_p r_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (4)$$

sendo  $r_t$ , por exemplo, o valor do log-retorno no instante  $t$  e  $\beta_0, \dots, \beta_p, \theta_1, \dots, \theta_q$  parâmetros a serem estimados e  $a_t$  um ruído branco (sequência de valores independentes com média zero e variância constante).

Uma outra abordagem em séries temporais é a implementação dos modelos de heteroscedasticidade condicional auto-regressiva (ARCH( $m$ )), proposto por Engle (1982), para a modelagem da volatilidade condicional de séries. Nesta abordagem, a variância condicional, por exemplo, dos log-retornos,  $h_t = Var(r_t | F_{t-1})$ , pode ser modelada usando valores passados ao quadrado da série, conforme a Equação 5.

$$r_t = \sqrt{h_t} \varepsilon_t, \quad (5.1)$$

$$h_t = a_0 + \sum_{i=1}^m \alpha_i r_{t-1}^2, \quad (5.2)$$

em que  $F_{t-1}$  denota a informação até o instante  $t$ , isto é,  $\{r_{t-1}, r_{t-2}, \dots, r_1\}$  e  $\{\varepsilon_t, t = 1, \dots\}$  são variáveis aleatórias independentes e identicamente distribuídas. Na prática, usualmente, se assume que os erros  $\varepsilon_t$  seguem uma distribuição normal com média zero e variância um.

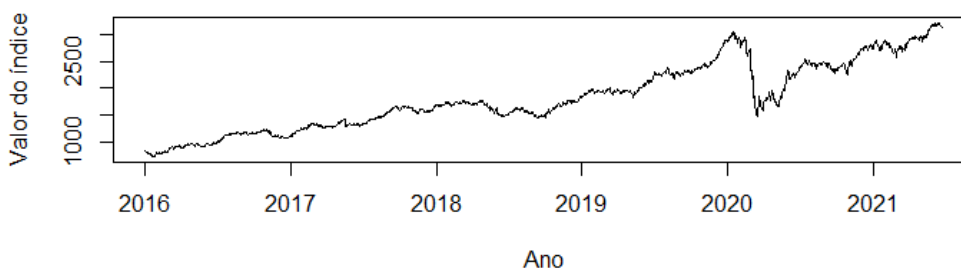
Neste trabalho será utilizado o modelo ARMA-ARCH para modelar a média e a volatilidade condicional dos log-retornos do índice SMLL. O modelo estimado para a série temporal dos log-retornos irá possibilitar a realização de previsões dos valores do índice.

## 4 RESULTADOS E DISCUSSÃO

### 4.1 ANÁLISE DESCRITIVA

Nesta seção será realizada uma análise descritiva do índice diário *Small Cap*, cujas cotações diárias representam um indicador médio do desempenho de empresas brasileiras negociadas na B3 que cumprem certos requisitos, mencionados na seção de Metodologia. Os dados foram coletados no *site* da Brasil, Bolsa, Balcão (B3, 2021). Na Figura 1 apresenta-se o comportamento da série no período considerado.

Figura 1 – Série diária do índice *Small Cap* (janeiro de 2016 a junho de 2021)



Fonte: autoria própria (2021).

Observa-se uma tendência crescente do índice no período de janeiro de 2016 até janeiro de 2020 com valores variando entre 722,92 e 3224,79. Ressalta-se que durante o período de 2016 a 2018 houve um crescimento do lucro e das cotações das empresas dos setores de saúde, imobiliário e elétrico que representavam, conjuntamente, entre 30% e 40% do total de setores que faziam parte do índice no período mencionado (Economática, 2021). Neste período de três anos, por exemplo, as empresas Odontoprev, do setor de saúde, MRV, do setor imobiliário e Alupar, do setor de energia elétrica, tiveram um aumento de 32%, 30% e 9% no lucro, respectivamente, e um aumento na cotação de 53%, 48% e 51%, respectivamente (StatusInvest e Google Finance, 2021). Por fim, observam-se algumas quedas no valor do índice nos períodos que foram marcados por eleições nos Estados Unidos (novembro de 2016, quando o índice caiu 8%) e o período de pré-eleições no Brasil (entre os meses de maio e setembro 2018, período em que o índice diminuiu 14%).

No ano de 2019 o crescimento do índice pode estar relacionado com as reduções na taxa básica de juros (SELIC) pelo Banco Central, que resultaram numa queda de 30% da taxa. Essas diminuições geralmente tendem a ser benéficas para as empresas ligadas a setores como transporte aéreo, turismo, imobiliário e vestuário uma vez que com crédito mais barato, a tendência é que as pessoas adquiram mais empréstimos para o consumo de itens não essenciais como lazer, viagens, entretenimento e compra de imóveis (Banco Central, 2021), beneficiando as empresas desses setores. Em 2019, a participação de empresas desses setores no SMLL era entre 20% e 30% (Economatica, 2021).

No início de 2020, o índice apresentou uma tendência de alta durante o mês de janeiro. Porém, em fevereiro e março de 2020 o valor do índice caiu 41% quando comparados ao início e ao fim do primeiro trimestre desse ano. Esse comportamento está possivelmente relacionado com as restrições tomadas devido à pandemia do COVID-19 que afetaram setores de comércio, transporte e turismo. Segundo o Ministério da Economia, os setores mais impactados pela pandemia no Brasil foram os de entretenimento, transporte, fabricação de veículos e vestuário, setores que representavam no primeiro semestre de 2020 entre 15% e 23% do total de setores que faziam parte do índice nesse período. Destacam-se as empresas Locamerica, do setor de transporte, e Arezzo do setor de vestuário, que tiveram quedas de 53% e 39% nas suas cotações, respectivamente (Google Finance, 2021). Por fim, observa-se que os valores do índice começaram a aumentar a partir de abril de 2020 e seguiram com uma tendência de alta até junho de 2021. Nesse período, o valor do SMLL teve um aumento de 93%, refletindo provavelmente a retomada dos setores mais impactados citados anteriormente.

A Tabela 1 apresenta as estatísticas descritivas do índice SMLL por semestre no período estudado.

Tabela 1 – Estatísticas descritivas do índice SMLL por semestre

Semestre	Média	Desvio Padrão	Mediana	Mínimo	Máximo	1º quartil	3º quartil
1º 2016	879,68	77,09	909,71	722,92	993,11	796,37	942,43
2º 2016	1.130,02	51,23	1.139,49	993,07	1.221,86	1.095,99	1.162,14
1º 2017	1.290,48	58,72	1.299,99	1.107,31	1.424,57	1.267,02	1.321,78
2º 2017	1.547,04	93,67	1.576,22	1.334,52	1.666,52	1.470,28	1.624,46
1º 2018	1.670,14	84,68	1.697,29	1.463,06	1.768,72	1.637,11	1.727,92
2º 2018	1.608,12	96,64	1.598,09	1.435,95	1.795,84	1.536,44	1.693,41
1º 2019	1.952,00	67,83	1.943,52	1.817,92	2.141,66	1.912,01	1.973,71
2º 2019	2.374,13	153,99	2.322,28	2.129,42	2.848,03	2.269,48	2.419,82
1º 2020	2.262,75	508,16	2.188,94	1.480,15	3.064,07	1.799,16	2.878,32
2º 2020	2.489,48	140,21	2.457,94	2.250,59	2.822,39	2.408,12	2.535,04
1º 2021	2.900,04	163,66	2.872,93	2.586,49	3.224,79	2.779,21	2.972,53

Fonte: autoria própria (2021).

Observa-se na Tabela 1, através do aumento da média no valor do índice semestre a semestre, que o SMLL esteve numa tendência de alta durante o período analisado. No entanto, essa tendência

de alta foi interrompida em dois momentos: no 2º semestre de 2018, em que a média caiu de 1.670,14 no semestre anterior para 1608,12, uma diminuição de 3,7%, e no 1º semestre de 2020, em que a média caiu de 2.374,13 no semestre anterior para 2.262,75, diminuição de 4,7%. No primeiro semestre de 2020, período em que tiveram início as restrições devido à pandemia do COVID-19, o índice apresentou uma alta volatilidade, como se observa no aumento do desvio padrão. Nota-se também uma diferença grande entre os valores máximos e mínimos do índice nesse período de pandemia, com o menor valor sendo 1.480,15 e o maior 3.064,07.

Na Tabela 2 apresenta-se a participação de cada setor no índice, durante o período estudado. As porcentagens são referentes à participação do setor no índice no primeiro dia de negociação da carteira vigente do período. Cada período (P1 até P17) é referente aos 4 meses em que a carteira esteve em vigência, com início em janeiro de 2016. Assim, por exemplo, P2 representa o período de maio de 2016 a julho de 2016. As empresas do SMLL foram agrupadas nos setores econômicos da B3, consumo cíclico (empresas de viagens, entretenimento, construção civil e vestuário), consumo não cíclico (alimentos), utilidade pública (energia e saneamento), financeiro, bens industriais (transporte e fabricação de veículos), saúde, materiais básicos (siderurgia, madeira e mineração), tecnologia da informação (TI) e petróleo gás e biocombustíveis (B3, 2021).

Tabela 2 – Participação setorial nas carteiras do índice SMLL (janeiro de 2016 a junho de 2021)

Período	Cons. cíclico	Cons. não cíclico	Util. pública	Financeiro	Bens industriais	Saúde	Materiais básicos	TI	Petróleo gás e biocombust.
P1	32,6%	11,7%	12,2%	10,9%	11,1%	11,7%	6,2%	3,0%	0,6%
P2	29,0%	9,9%	16,7%	13,6%	9,8%	10,1%	8,3%	2,2%	0,4%
P3	31,1%	7,7%	12,4%	12,3%	10,7%	15,2%	7,9%	2,3%	0,4%
P4	26,3%	6,8%	18,3%	17,8%	7,6%	11,9%	8,9%	2,0%	0,4%
P5	29,0%	7,3%	16,0%	10,2%	8,7%	17,2%	8,9%	2,2%	0,4%
P6	36,3%	6,8%	18,0%	12,5%	8,7%	5,0%	10,2%	2,1%	0,4%
P7	28,4%	6,3%	18,1%	11,0%	17,0%	10,2%	6,2%	2,2%	0,5%
P8	32,5%	6,3%	18,0%	14,5%	8,8%	12,1%	5,1%	1,7%	1,1%
P9	25,8%	6,1%	17,4%	14,5%	11,8%	9,2%	10,2%	1,3%	3,6%
P10	28,7%	4,7%	21,2%	10,3%	14,6%	7,6%	8,6%	2,1%	2,2%
P11	28,2%	4,6%	21,7%	14,8%	12,1%	7,6%	7,5%	2,0%	1,5%
P12	30,8%	4,8%	24,6%	8,2%	9,4%	8,6%	9,1%	2,3%	2,1%
P13	27,3%	6,0%	21,7%	12,9%	10,0%	9,1%	9,0%	2,0%	2,2%
P14	26,3%	9,0%	22,7%	12,4%	10,3%	7,7%	7,8%	2,1%	1,8%
P15	26,9%	6,4%	18,2%	16,1%	10,8%	7,2%	8,5%	3,8%	2,0%
P16	27,0%	5,4%	19,4%	14,7%	12,5%	6,6%	6,7%	4,6%	3,2%
P17	27,2%	7,7%	12,0%	14,4%	15,3%	5,8%	5,9%	6,1%	5,6%

Fonte: elaborado a partir dos dados da Economatica (2021).

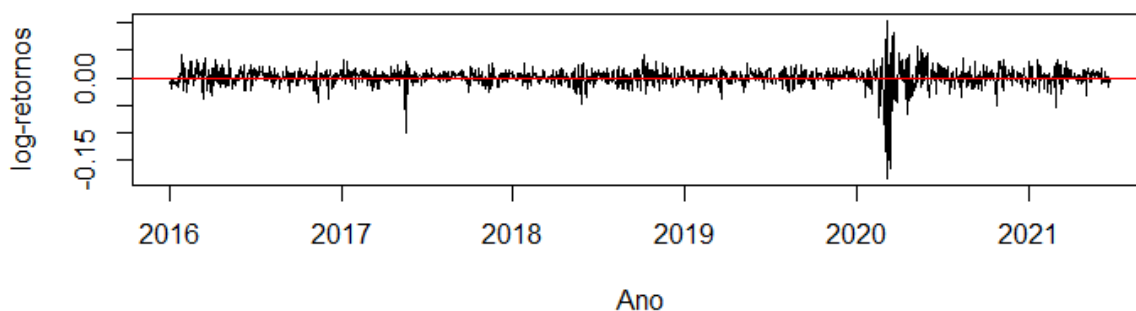


Destaca-se que durante o período total analisado as carteiras do SMLL vigentes possuíram uma participação relevante de empresas do setor econômico de consumo cíclico, em torno de 30% em todos os períodos. Nota-se também um aumento da participação no índice de empresas de setores de tecnologia da informação e de petróleo gás e biocombustíveis, entre o período P1 e o período P17. Destaca-se um aumento de participação do setor de utilidade pública (empresas de energia e saneamento) de 12,2% no período P1 (janeiro de 2016) para 24,6% no período P12 (setembro de 2019) e posterior queda até o patamar de 12,0% no período P17 (maio de 2021). Com relação aos demais setores, as participações nas carteiras nos períodos estudado se mantiveram sem grandes mudanças.

## 4.2 MODELAGEM

Nesta seção será feita a modelagem dos log-retornos do índice SMLL, no período janeiro de 2016 a junho de 2021, usando os modelos  $ARMA(p,q)+ARCH(m)$  comentados na seção de Metodologia. Na Figura 2 apresentam-se os log-retornos, calculados conforme a Equação 3.

Figura 2 – Log-retornos da série diária do índice *Small Cap*



Fonte: autoria própria (2021).

Nota-se que os log-retornos variam em torno de média zero, destacada pela linha vermelha. Um valor aparentemente atípico do log-retorno do índice SMLL é observado no dia 18/05/2017. É importante destacar que nesse dia o índice teve uma queda em seu valor de 9,6%, marcado por notícias ligadas à possível corrupção no governo, juntamente com quedas relevantes nas cotações de ações da bolsa de forma generalizada (G1, 2017). Também é possível observar que no primeiro semestre de 2020 ocorreu um aumento na volatilidade do índice. Como o cálculo dos log-retornos utiliza a diferença entre o valor do índice em  $t$  e  $t-1$  na base do log, grandes variações positivas ou negativas do índice em dias consecutivos geram valores do log-retorno altos, como acontece nesse período. Esse período foi marcado pelo início das restrições ao comércio e transporte devido ao avanço da pandemia do coronavírus (COVID-19) no Brasil. A incerteza a respeito do futuro da economia nacional e mundial aparenta ter levado aos movimentos de queda mais acentuada nas cotações das

ações presentes no índice, conforme comentado na seção 4.1, o que se reflete na maior volatilidade representada no gráfico.

Em razão dessa volatilidade atípica apresentada apenas nesse período e para dar sequência ao estudo, foi decidido modelar os log-retornos até 4 de março de 2020, ou seja, antes do período marcado pelo avanço da pandemia no Brasil. Dessa forma, as próximas análises serão referentes aos log-retornos do índice sem considerar esse período.

Inicialmente, ajustou-se o modelo ARMA(1,0)+ARCH(1), uma vez que é o mais simples. Neste modelo considera-se um termo defasado do log-retorno (valor anterior  $r_{t-1}$ ) para modelar a sua média e um termo defasado ao quadrado para modelar a volatilidade condicional ( $r_{t-1}^2$ ). O modelo ajustado é dado pela Equação 6.

$$r_t = 0,001667 + 0,062189 r_{t-1} + a_t \quad (6.1)$$

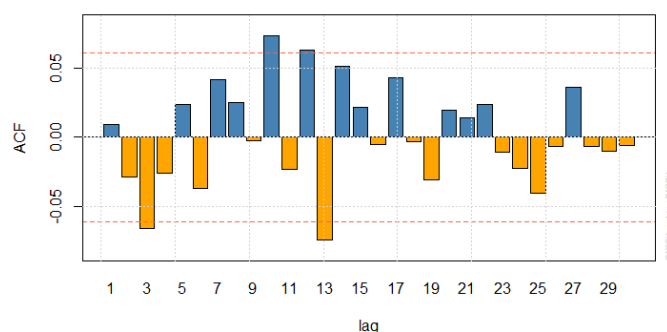
$$a_t = \sqrt{h_t} \varepsilon_t \quad (6.2)$$

$$h_t = 0,000129 + 0,14828 r_{t-1}^2, \quad (6.3)$$

em que  $r_t$  é o log-retorno da série no momento  $t$ ,  $h_t$  é a variância condicional no momento  $t$  e  $\varepsilon_t$  são variáveis aleatórias independentes e identicamente distribuídas que seguem uma distribuição normal com média zero e variância um.

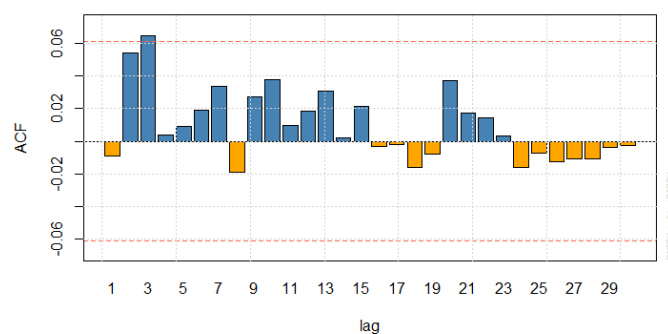
Na análise de resíduos deste modelo, verificou-se que os resíduos são independentes. Primeiramente foram plotados dois gráficos: autocorrelação de resíduos padronizados e autocorrelação de resíduos padronizados ao quadrado. Observam-se 4 *lags* com autocorrelação significativa na Figura 3 e 1 *lag* com autocorrelação significativa na Figura 4. Assim, para confirmar ou não a independência realizou-se o teste de hipótese de Ljung-Box.

Figura 3 – Autocorrelação de resíduos padronizados



Fonte: autoria própria (2021).

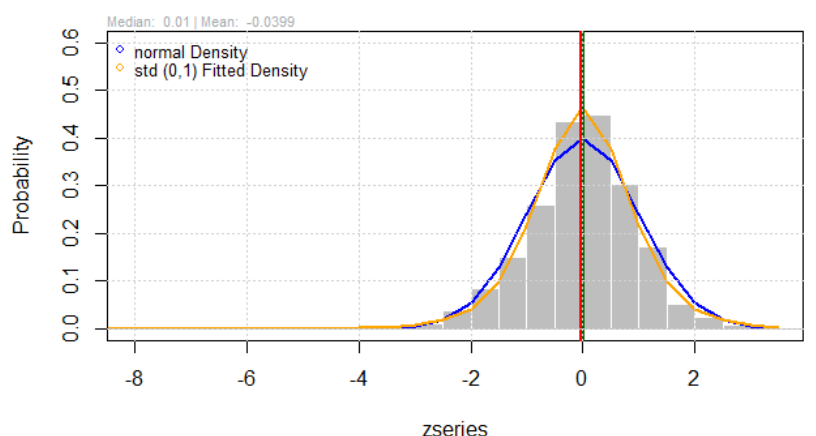
Figura 4 – Autocorrelação de resíduos padronizados ao quadrado



Fonte: autoria própria (2021).

Para verificar a independência dos resíduos, aplicaram-se dois testes: o teste ponderado de Ljung-Box em resíduos padronizados e o teste ponderado de Ljung-Box em resíduos padronizados ao quadrado, considerando 5 defasagens (BOX; LJUNG, 1978), cuja hipótese nula, em ambos os testes, são que os dados são independentes. Obtiveram-se, respectivamente, valores-p de 0,2415 e 0,1435, sendo assim, como são valores-p maiores que 0,05, pode-se concluir, ao nível de significância de 5%, que os resíduos e os resíduos ao quadrado são independentes. Por fim, apesar de os resíduos dos log-retornos geralmente seguirem uma distribuição normal, verificou-se que a distribuição t-student apresentou um melhor resultado para essa série, conforme indicado na Figura 5. Estes resultados confirmam que o modelo é adequado para modelar a dependência linear entre sucessivos retornos e que a heteroscedasticidade condicional foi adequadamente modelada.

Figura 5 – Distribuições normal e t-student dos resíduos da série



Fonte: autoria própria (2021).

Neste trabalho considerou-se somente este modelo, uma vez que a análise de resíduos apresentou resultados satisfatórios. Utilizando o modelo descrito na Equação 6, foram realizadas previsões para os log-retornos do índice em 7 dias a partir da data 04/03/2020. Na sequência, foi feito

o cálculo dos valores previstos do índice a partir dos valores dos log-retornos encontrados, conforme descrito na Equação 7.

$$\hat{r}_{T+h} = \ln\left(\frac{\hat{Z}_{T+h}}{Z_T}\right) \rightarrow \hat{Z}_{T+h} = Z_T \exp(\hat{r}^{T+h}) \quad (7)$$

em que  $\hat{Z}_{T+h}$  e  $\hat{r}^{T+h}$  são, respectivamente, os valores previstos do índice e do log-retorno  $h$  passos à frente e  $T$  o valor da série no dia 04/03/2020. Para analisar a acurácia das previsões foram comparados os valores previstos com os observados calculando-se o erro relativo para cada dia, conforme descrito na Equação 8.

$$E_{T+h} = \left| \frac{Z_{T+h} - \hat{Z}_{T+h}}{Z_{T+h}} \right| \quad (8)$$

Na Tabela 3, apresentam-se os valores reais do índice nos 7 dias, os resultados obtidos do cálculo dos valores previstos e os erros relativos calculados.

Tabela 3 – Comparação entre os valores reais do índice SMLL e os previstos pelo modelo

Período	Valor observado do índice	Valor previsto do índice	Erro
T+0	2742,07	-	-
T+1	2574,98	2780,84	8,0%
T+2	2480,53	2607,55	5,1%
T+3	2170,90	2511,33	15,7%
T+4	2326,53	2197,77	5,5%
T+5	2182,38	2355,32	7,9%
T+6	1818,87	2209,37	21,5%
T+7	2017,60	1841,36	8,7%

Fonte: autoria própria (2021).

Nesses resultados é possível observar erros relativos altos, afetados provavelmente pela maior volatilidade da série nesse período. Finalmente, foi calculada a Raiz do Erro Quadrático Médio (REQM), cuja fórmula está descrita na Equação 9.

$$REQM = \sqrt{\sum_{h=1}^7 \frac{(Z_{T+h} - \hat{Z}_{T+h})^2}{7}} \quad (9)$$

O valor obtido da *REQM* para esses períodos foi de 240,37. Esse valor, representa aproximadamente o erro médio que se estaria cometendo ao substituir cada valor observado pelo previsto pelo modelo descrito em (6). Seu valor alto deve-se provavelmente à alta volatilidade da série SMLL nesses dias de março de 2020, período em que se iniciaram as restrições adotadas pelo governo brasileiro para combater o COVID-19.

Assumindo que fazer previsões para o índice SMLL em 2020 é difícil, uma vez que o mercado apresentou grandes variações no seu comportamento e o como modelo ARMA+ARCH utiliza os valores passados para tentar entender a série, é esperado não ter uma acurácia nas previsões. No entanto, surgiu a hipótese: se for ignorado o período de alta volatilidade da série em 2020, como foi feito na modelagem, o modelo resultante consegue fazer boas previsões para o ano de 2021?

Com a hipótese de que o modelo apresentaria melhores resultados caso fosse desconsiderado da série um ano após o impacto inicial dessas restrições, ele foi testado também nas observações do índice SMLL após esse período, com as previsões feitas para os 7 dias a partir da data 04/03/2021. Os resultados foram consolidados na Tabela 4.

Tabela 4 – Comparação entre os valores reais do índice SMLL e os previstos pelo modelo após um ano

Período	Valor real do índice	Valor previsto do índice	Erro
T+0	2676,85	-	-
T+1	2735,00	2714,70	0,7%
T+2	2590,51	2769,60	6,9%
T+3	2586,49	2622,67	1,4%
T+4	2643,60	2618,50	0,9%
T+5	2728,16	2676,32	1,9%
T+6	2736,80	2761,90	0,9%
T+7	2766,62	2770,64	0,1%

Fonte: autoria própria (2021).

Notam-se erros relativos baixos para os períodos, em contraste aos erros apresentados na Tabela 3. Para esses períodos, o valor da REQM foi de 73,44. Destaca-se que as previsões para o ano de 2021 ficaram mais próximas dos valores observados. Isso se deve, provavelmente, ao fato de que o valor do índice no começo de 2021 começou a atingir valores esperados, sem o impacto da alta volatilidade que ocorreu no começo de 2020, devido ao início da pandemia no Brasil.

## 5 CONSIDERAÇÕES FINAIS

Neste trabalho foi testada também a modelagem da série utilizando os dados do período afetado pela pandemia do COVID-19, porém não se ajustou um modelo adequado à série. Isso se deve provavelmente à uma volatilidade atípica observada nesse período, o que fez com que os resíduos dos modelos testados não apresentassem boas propriedades.

No início do trabalho foi modelado os log-retornos do índice SMLL utilizando a abordagem ARMA e ARCH, com o intuito de validar a eficiência desses modelos e utiliza-los para fazer previsões a curto prazo. O modelo ARMA(1,0)-ARCH(1) foi ajustado à série de forma que o resultado da análise de resíduos confirmou que ele é adequado e modelou a dependência linear entre sucessivos retornos para o período considerado. Com isso, foi utilizado para fazer previsões a curto

prazo e também considerando dados da série depois de um ano, após a influência inicial da pandemia do COVID-19 nas cotações das ações do índice. E com esses resultados conclui-se que o modelo ajustado foi capaz de realizar previsões de forma satisfatória.

Para pesquisas futuras sugere-se que os modelos ARMA e ARCH sejam utilizados incluindo variáveis explicativas, de forma a encontrar um melhor modelo que possa englobar também o período afetado pela pandemia, ou ainda a utilização de outros modelos para a série.

## REFERÊNCIAS

ARAÚJO, J. M. de; FILHO, P. A. M. **Modelagem da volatilidade apresentada pelos índices IVBX-1 e SMLL em 2008 usando modelos da família ARCH**. 2012. 22 f. TCC (Graduação) - Curso de Administração, Universidade Presbiteriana Mackenzie, São Paulo, 2012. Disponível em: <https://www.scielo.br/j/ram/a/FKsBqP45b3pH4FYMymcpjHv/?lang=pt>. Acesso em: 21 nov. 2021.

B3. **Classificação setorial**. Disponível em: [https://www.b3.com.br/pt\\_br/produtos-e-servicos/negociacao/renda-variavel/acoes/consultas/classificacao-setorial/](https://www.b3.com.br/pt_br/produtos-e-servicos/negociacao/renda-variavel/acoes/consultas/classificacao-setorial/). Acesso em: 21 nov. 2021.

B3. **Índice Small Cap (SMLL B3)**. Disponível em: [http://www.b3.com.br/pt\\_br/market-data-e-indices/indices/indices-de-segmentos-e-setoriais/indice-small-cap-sml-estatisticas-historicas.htm](http://www.b3.com.br/pt_br/market-data-e-indices/indices/indices-de-segmentos-e-setoriais/indice-small-cap-sml-estatisticas-historicas.htm). Acesso em: 21 nov. 2021.

B3. **Manual de definições e procedimentos dos índices da B3**. Novembro de 2018. Disponível em: < <http://www.b3.com.br/data/files/AF/83/C4/BA/25CB7610F157B776AC094EA8/Conceitos-Procedimentos-nov2018.pdf> >. Acesso em: 21 nov. 2021.

B3. **Metodologia do índice Small Cap (SMLL)**. 2020. Disponível em: [http://www.b3.com.br/data/files/4C/70/8E/79/75B217102C51CE07DC0D8AA8/MetodologiaSMLL\\_PT.pdf](http://www.b3.com.br/data/files/4C/70/8E/79/75B217102C51CE07DC0D8AA8/MetodologiaSMLL_PT.pdf). Acesso em: 21 nov. 2021.

B3. **Uma análise da evolução dos investidores na B3**. 2021. Disponível em: [https://www.b3.com.br/data/files/EC/B5/B4/6F/6C63B71027085EA7AC094EA8/Book\\_PF-Agosto2021.pdf](https://www.b3.com.br/data/files/EC/B5/B4/6F/6C63B71027085EA7AC094EA8/Book_PF-Agosto2021.pdf). Acesso em: 21 nov. 2021.

BANCO CENTRAL DO BRASIL. **Taxa Selic**. 2021. Disponível em: <https://www.bcb.gov.br/controleinflacao/taxaselic>. Acesso em: 21 nov. 2021.

BOLLERSLEV, Tim. Generalized autoregressive conditional heteroskedasticity. **Journal of Econometrics**, [S.L.], v. 31. N. 3, p.307-327, abr. 1986.

BUENO, Rodrigo de Losso da Silveira. **Econometria de Séries Temporais**. 2 ed. São Paulo: Cengage Learning, 2018.

ECONOMATICA. **Banco de dados**. Disponível em: [www.economatica.com.br](http://www.economatica.com.br). Acesso em: 21 nov. 2021.

ENGLE, Robert F. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. **Econometrica**, [S.L.], v. 50, n. 4, p. 987, jul. 1982.

FILHO, D. B. de A. **Modelos de precificação forward-looking aplicados à blue chips e small caps do mercado brasileiro**. 2015. 35 f. TCC (graduação em Finanças) - Universidade Federal do Ceará, Faculdade de Economia, Administração, Atuária e Contabilidade, Fortaleza/CE, 2015. Disponível em: <http://www.repositorio.ufc.br/handle/riufc/34000>. Acesso em: 21 nov. 2021.

FILHO, E. L. dos S.; DA SILVA, W. V.; DA VEIGA, C. P.; TORTATO, U. Previsão dos retornos do IBOVESPA utilizando redes neurais artificiais feedforward evolutivas. **Revista Produção Online**, [S. l.], v. 11, n. 4, p. 1114–1140, 2011. Disponível em: <https://producaoonline.org.br/rpo/article/view/784>. Acesso em: 21 nov. 2021.

GOOGLE FINANCE. **Banco de dados**. Disponível em: <https://www.google.com/finance/>. Acesso em: 21 nov. 2021.

GOVERNO FEDERAL. **Ministério da Economia divulga lista dos setores mais afetados pela pandemia da Covid-19 no Brasil**. Disponível em: <https://www.gov.br/economia/pt-br/assuntos/noticias/2020/setembro/ministerio-da-economia-divulga-lista-dos-setores-mais-afetados-pela-pandemia-da-covid-19-no-brasil>. Acesso em: 21 nov. 2021.

JORION, Philippe. **Value at Risk: The New Benchmark for Controlling Market Risk**. Illinois: Irwin Professional, 1997.

KACZOROWSKI, B.; KLEINA, M.; AUGUSTO MENDES MARQUES, M.; DE ASSIS SILVA, W. Artificial Intelligence And The Multivariate Approach In Predictive Analysis Of The Small Cap Index Of The Brazilian Stock Exchange. **IEEE Latin America Transactions**, [S. l.], v. 19, n. 11, p. 1924–1932, 2021. Disponível em: <https://latamt.ieee9.org/index.php/transactions/article/view/4808>. Acesso em: 21 nov. 2021.

KOBUNDA, C. N.; JÚNIOR E. C. N. Análise da volatilidade do Ibovespa entre 2001 e 2016: uma estimação por meio de modelos ARCH. **Estudos Econômicos**, São Paulo, v.29, n. 3, p. 303-341, jul./set., 1999. Disponível em: <https://revistas.ufpr.br/economia/article/view/63969/39490>. Acesso em: 21 nov. 2021.

LJUNG, G. M.; BOX, G. E. P. On a Measure of a Lack of Fit in Time Series Models. **Biometrika**. 65 (2): 297–303. 1978.

MÓL, A. L. R.; FELIPE, I. J. dos S.; JÚNIOR, F. M. G. Volatilidade dos índices de ações Mid-large Cap e Small Cap: uma investigação a partir de modelos ARIMA/GARCH. **Revista de Gestão, Finanças e Contabilidade**, UNEB, Salvador, v. 4, n. 1, p. 04-29, jan./abr., 2014. Disponível em: <https://www.revistas.uneb.br/index.php/financ/article/view/522/508>. Acesso em: 21 nov. 2021.

MORAIS I. A. C.; PORTUGAL M. S. Modelagem e Previsão de Volatilidade Determinística e Estocástica para a Série do Ibovespa. **Revista de Economia**, v. 40, n. 72, p. 1. Disponível em: <https://www.revistas.usp.br/ee/article/view/161468/155418>. Acesso em: 21 nov. 2021.

MORETTIN, P. A. **Econometria Financeira: um curso em séries temporais financeiras**. 2. ed. São Paulo: Blucher, 2011. 398 p.

MORETTIN, P. A., TOLOI, Clélia M. C. **Análise das Séries Temporais**. 2. ed. São Paulo. Editora Universidade de São Paulo, 2020.

SHUMWAY, Robert H.; STOFFER, David S.; STOFFER, David S. **Time series analysis and its applications**. New York: springer, 2000.

STATUSINVEST. **Banco de dados**. Disponível em: <https://statusinvest.com.br/>. Acesso em: 21 nov. 2021.

TREVIZAN, Karina. Bovespa fecha na maior queda em quase 9 anos após denúncias da JBS. **G1**, 18 mai. 2017. Disponível em: <https://g1.globo.com/economia/mercados/noticia/bovespa-fecha-em-forte-queda-de-olho-em-denuncias-sobre-temer.ghtml>. Acesso em 21 nov. 2021.