

Séries Temporais: Análise do Índice Bovespa 2020

Clovis Rodrigues Cavalcanti Filho
clovis.rodrigues@ufpe.br

Universidade Federal de Pernambuco

Lucas Daniel Gomes Ribeiro
lucas.gois@ufpe.br

Universidade Federal de Pernambuco

Maria Gabriela Bezerra Nascimento Gonçalves
mariagabriela.goncalves@ufpe.br

Universidade Federal de Pernambuco

Resumo

Utilizando-se de séries temporais este artigo realiza uma análise do Índice Bovespa o principal indicador de desempenho das ações negociadas na Bolsa de Valores brasileira IBovespa (B3) do ano de 2020, o conjunto de dados correspondeu aos meses de janeiro a setembro o que se totalizou 180 dados levando em consideração as variações do índice ao longo do tempo, essa pesquisa possibilita entender o comportamento dinâmico do ativo ou mesmo de um mercado de ações. A natureza da pesquisa enquadra-se como tecnológica uma vez que conhecimentos básicos são aplicados e novos conhecimentos são gerados como resultado do processo de pesquisa. Levando em consideração que a teoria do mercado é eficiente, concluiu-se que, não se mostra adequada a utilização de modelos *Autoregressive Moving Average* (ARMA) ou *Autoregressive Integrated Moving Average* (ARIMA) para previsão dos dados.

Palavras-chave: Série temporal. ARIMA. Índice Bovespa.

Abstract

Using time series, this article analyzes the Bovespa Index, the main performance indicator of shares traded on the Brazilian Stock Exchange IBovespa (B3) in 2020, the data set corresponded to the months from January to September, which totaled 180 data, taking into account the variations of the index over time, this research makes it possible to understand the dynamic behavior of the asset or even of a stock market. The nature of the research fits technological as basic knowledge is applied and new knowledge is generated as a result of the research process. Taking into account that the theory of the market is efficient, it was concluded that it is not appropriate to use *Autoregressive Moving Average* (ARMA) or *Autoregressive Integrated Moving Average* (ARIMA) models for data prediction.

Keywords: Time series. ARIMA. Bovespa Index.

1 INTRODUÇÃO

Séries Temporais podem ser observadas quando os dados são contabilizados em diferentes instantes do tempo, seja diariamente (preço de ações, relatórios meteorológicos), mensalmente (taxa de desemprego, Índice de Preço ao Consumidor), trimestralmente (Produto Interno Bruto), etc. (DAVILA, 2020).

Sendo assim, uma modelagem estatística capaz de quantificar e prever um determinado atributo ao longo do tempo são as análises de Séries Temporais (ALBUQUERQUE, 2015). O objetivo dessas análises compreende a elaboração de um modelo estatístico que descreva adequadamente o comportamento de um determinado atributo ao longo do tempo, de maneira que as implicações teóricas do modelo sejam compatíveis com as amostras observadas na série temporal. A partir do modelo ajustado à série temporal é possível prever a evolução futura da série ou explicar a relação entre os distintos componentes do modelo (ALBUQUERQUE, 2015; NEPOMUCENO; de CARVALHO; COSTA, 2020).

O Índice Bovespa é o principal indicador de desempenho das ações negociadas na Bolsa de Valores brasileira IBovespa (B3), e reúne as empresas mais importantes do mercado de capitais brasileiro. Foi criado em 1968 e, ao longo desses 50 anos, consolidou-se como referência para investidores ao redor do mundo (B3, 2022).

A análise de Séries Temporais financeiras lida com a avaliação do preço ou do retorno de ativos no decorrer do tempo, possibilitando o entendimento do seu comportamento dinâmico e construindo modelos capazes de prever valores futuros da série (KEIEL; BENDER, 2018). As séries de ativos, por sua vez, geralmente possuem certas propriedades como baixa correlação serial e agrupamento de volatilidade, dificultando a identificação de modelos para os processos (KEIEL; BENDER, 2018). A análise e classificação dessas séries quanto as suas principais características é de suma importância, pois possibilita entender o comportamento dinâmico do ativo ou mesmo de um mercado de ações a fim de usá-la para diversas finalidades, entre as quais a de realizar previsões futuras da série (KEIEL; BENDER, 2018).

Na década de 70, o livro *Time series analysis: forecasting and control* sumariza as metodologias conhecidas até então para a formulação de séries temporais através de equações matemáticas, popularizando o uso do modelo autorregressivo integrador de média móvel, que ficou conhecido como ARIMA (KEIEL; BENDER, 2018).

2 REVISÃO DA LITERATURA

Uma série temporal é um conjunto de valores observados, medidos ao longo de períodos sucessivos de tempo (GAITHER; FRAZIER, 2002, apud KIRSTEN 2009, p. 20).

Prever o futuro, e em especial o comportamento de séries temporais, é fundamental em análises e tomada de decisões, e continua sendo um desafio para a estatística e a computação (ABELÉM, 1994, apud KIRSTEN 2009, p. 15).

Para Corrar e Theóphilo (2004),

Uma série temporal é um conjunto de observações sequenciais de determinada variável, expressas numericamente, obtidas em períodos regulares de tempo. A disponibilidade dos dados é que condiciona a formatação das séries. O analista deve cuidar para que as observações sejam dispostas em pontos equidistantes no tempo, obtendo, dessa forma, séries temporais discretas (CORRAR; THEÓPHILO, 2004 apud KIRSTEN 2009, p. 20).

Em trabalho (Yule, 1927) que contribuiu para o desenvolvimento de ferramentas e métodos para a análise de séries temporais, em que refere uma série temporal que deveria ser vista simplesmente como a realização de um processo estocástico. O autor criou o modelo autorregressivo do inglês *Autoregressive* (AR), em que o valor previsto dependia dos valores passados.

Contudo, os modelos lineares mostram-se insuficientes para a análise de séries temporais, em virtude de a maioria das séries reais apresentarem fortes tendências de não linearidade (FISCHER, 1982).

Quanto aos componentes de uma série temporal, pode-se elencar:

I. Tendência

É caracterizada como aquele movimento regular e contínuo de longo prazo, refletindo um movimento ascendente ou descendente em longo período de tempo...pode ser vista como aquela componente que descreve as variações graduais que se mantêm em um longo período de observação da variável no tempo (FISCHER, 1982, p.20).

II. Variações cíclicas

Consistem em variações que apresentam movimentos regulares em torno da reta ou da curva de tendência. Referem-se às oscilações de longo prazo - períodos de um ano ou mais. Os ciclos podem ser ou não periódicos, isto é, podem ou não seguir padrões análogos, depois de intervalos de tempos iguais. Por exemplo: eleições políticas, guerra, condições econômicas ou pressão psicológica (CORRAR; THEÓPHILO, 2004, apud KIRSTEN 2009, p. 21).

III. Variações sazonais

São aquelas variações periódicas (cíclicas) que ocorrem com certa regularidade dentro de um curto período de tempo. Embora o próprio nome dê a entender que esses movimentos ocorrem por um período anual, de acordo com as estações climáticas, esses movimentos podem ser estendidos a qualquer intervalo de curto prazo, como diário, horário, semanal trimestral, etc. (FISCHER, 1982, p. 20).

IV. Variações irregulares ou aleatórias

Referem-se não só aqueles movimentos esporádicos ocasionados por eventos aleatórios imprevisíveis, tais como as calamidades da natureza, mas também ao conjunto de todos aqueles movimentos da série que não foram possíveis de identificação em seus demais componentes, uma vez que não obedecem a nenhuma lei comportamental capaz de ser descrita de forma determinística, através de relação funcionais exclusivamente matemáticas (FISCHER, 1982, p. 20).

“Além dos quatro itens propostos por Corrar e Theóphilo, Tubino (2000) coloca mais um componente (que é comum a outros autores): a variação aleatória” (KISTEN, 2009, p. 21). A intervenção, muitas vezes, pode estar obscura por três fontes de ‘ruídos’: a tendência, sazonalidade e o erro aleatório (CORRAR, THEOPHILO, 2004). O fato de existir tendência na série pode induzir o pesquisador a tirar falsas conclusões, pois a ocorrência de uma inclinação ou uma mudança de nível na série pode ocorrer devido a uma tendência (CORRAR, THEOPHILO, 2004).

Dentre os modelos estatísticos propostos, os modelos de Box & Jenkins, conhecidos como modelos em inglês *autoregressive integrated moving average* (ARIMA) (1), são os mais difundidos tanto no meio acadêmico como no meio prático, principalmente em aplicações econômicas e financeiras (bolsas de valores, mercado financeiros etc.) (FERREIRA, 2006). Os modelos de Box & Jenkins são muito bem entendidos, porém são modelos lineares, o que pode vir a ser um fator limitante se o fenômeno em estudo for governado por leis de relacionamentos não-lineares (FERREIRA, 2006). Utiliza-se também teste Jarque Bera para homocedasticidade e teste Ljung-Box para independência dos resíduos.

Braga (2006) verificou a correlação do índice Bovespa e o impacto disso na previsão do preço de fechamento utilizando redes neurais e o modelo GARCH (*Generalized Autoregressive Conditional Heteroskedasticity*). Concluiu-se que os resultados utilizando redes neurais são motivadores e indicam a possibilidade de se utilizar a solução. A previsão gerada pelo modelo GARCH serviu como entrada para a previsão da rede neural, porém não houve ganho substancial. Em relação a redes neurais e sua importância na tomada de decisão Carneiro e Costa (2020) realizam trabalhos sobre a relação que as ferramentas de apoio a decisão possuem sobre as ações de gestores em diversos meios, inclusive no meio financeiro.

3 METODOLOGIA

Segundo JUNG (2004), as pesquisas científicas podem ser classificadas quanto à natureza, objetivos, procedimentos, fonte de referências, além de ambiente e tempo de aplicação. Em relação à natureza, esta pesquisa enquadra-se como tecnológica uma vez que

conhecimentos básicos são aplicados e novos conhecimentos são gerados como resultado do processo de pesquisa (JUNG, 2004).

Para análise de estimativas de previsões com maior praticidade e rapidez, pode-se utilizar os métodos de previsão, pela adoção de um pacote estatístico. Dentre os softwares estatísticos utilizáveis destaca-se o R (R Studio, 2020), um software livre e de recomendação do docente da disciplina de séries temporais. O R foi criado por Ross Ihaka e Robert Gentleman e, desenvolvido por esforço colaborativo (MENSURAÇÃO FLORESTAL, 2022). A intenção desse trabalho é mostrar uma aplicação da previsão de uma série temporal, índice BOVESPA, utilizando o software R. Os dados são referentes ao período de janeiro de 2020 a setembro de 2020, num total de 180 observações.

3.1 Modelo autorregressivo

O modelo AR (1) é um modelo autorregressivo de ordem $p = 1$ que assume grande importância, onde a série temporal X_t é definida pela Equação 1, conforme Xavier (2016).

Equação 1 - Série temporal do modelo autorregressivo

$$X_t = \phi_1 X_{t-1} + a_t$$

Fonte: Xavier (2016)

onde:

ϕ_1 – parâmetro autorregressivo de ordem 1;

X_{t-1} – Série de tempo defasado um período;

a_t – termo do erro do modelo;

Generalizando e considerando X_t , $t \in \mathbb{Z}$, pode-se afirmar que se trata de um modelo autorregressivo de ordem p ($X_t, \sim AR(p)$) como mostra a Equação 2.

Equação 2 - Modelo autorregressivo

$$X_t = \phi_0 + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + a_t$$

Fonte: Xavier (2016)

sendo: $\phi_0, \phi_1, \dots, \phi_p$ parâmetros reais.

3.2 Modelo de média móvel

Diz-se que um processo linear $\{X_t, t \in \mathbb{Z}\}$ é um processo de média móvel de ordem q , é nomeado por MA(q), se satisfaz a seguinte equação de diferenças, apontada por Xavier (2016) e apresentada na Equação 3.

Equação 3 - Modelo de média móvel

$$X_t = \alpha_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

Fonte: Xavier (2016)

onde $\mu, \theta_1, \dots, \theta_q$ são constantes reais. e $a_t \sim N(0, \sigma^2)$.

3.3 Modelo ARIMA

Para Xavier (2016) os modelos autorregressivos e de médias móveis denominados por *Autoregressive Moving Average* (ARMA) (q,p) e são a junção dos modelos AR e *Moving Average* (MA) e podem representar-se como mostra a Equação 4.

Equação 4 - Modelo ARMA

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \alpha_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

Fonte: Xavier (2016)

O modelo ARMA (1,1) é um dos que mais se utiliza e é visto na Equação 5.

Equação 5 - Modelo ARMA (1,1)

$$X_t = \phi_1 X_{t-1} + \alpha_t - \theta_1 \alpha_{t-1}$$

Fonte: Xavier (2016)

A função autocorrelação poderá ser calculada como é apresentado na Equação 6.

Equação 6 - Função de autocorrelação

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \dots + \phi_p \gamma_{j-p} + \gamma_{\alpha}(j) - \theta_1 \gamma_{\alpha}(j-1) - \dots - \theta_q \gamma_{\alpha}(j-q)$$

Fonte: Xavier (2016)

sendo $\gamma_{\alpha}(j)$ a covariância cruzada entre X_t e α_t e definida por:

$$\gamma_{\alpha}(j) = \begin{cases} 0, & j > 0 \\ \neq 0, & j \leq 0 \end{cases}$$

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \dots + \phi_p \gamma_{j-p}, j > q$$

$$\rho_j = \gamma_j / \gamma_0 = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \dots + \phi_p \rho_{j-p}, j > q$$

“Se esta for infinita, a decrescer de acordo com exponenciais e/ou senoides amortecidas após ‘lag’ p-q é um bom indicativo para a escolha do modelo ARMA (p,q) para a análise dos dados” (XAVIER, 2016, p. 23).

4 DADOS

Os dados utilizados no presente trabalho correspondem ao Índice BOVESPA do ano 2020, onde o conjunto de dados compreende os meses de janeiro a setembro, abordando as variações do índice ao longo do tempo. Ao todo, 180 dados são utilizados para análise neste trabalho. Na Figura 1 pode-se visualizar a média, mediana, desvio padrão, variância, primeiro e terceiro quartil, além dos valores mínimos e máximos do conjunto de dados analisados.

Figura 1 - Estatística descritiva dos dados

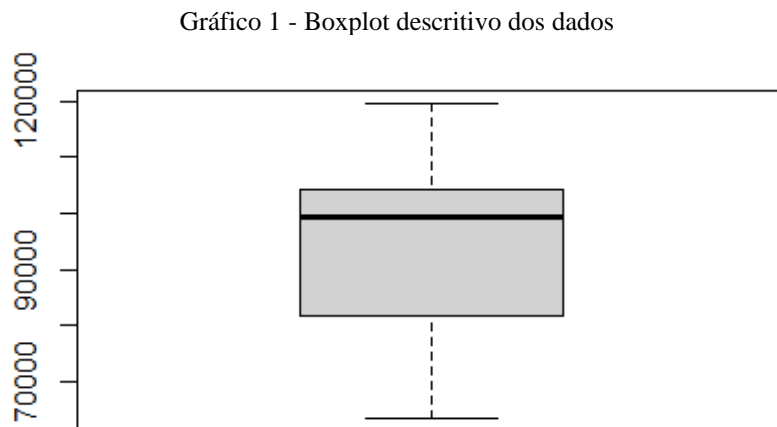
```
summary(Y)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
63570  81960   99482   96248  104339  119528

sd(Y)
14500.15

var(Y)
210254373
```

Fonte: Autores (2022).

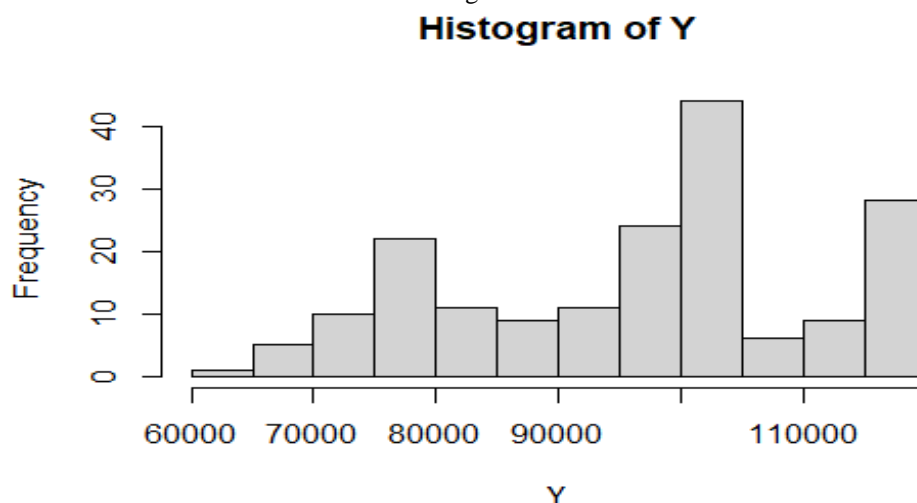
O boxplot do conjunto de dados pode ser visto no Gráfico 1, onde pode-se perceber que os limites máximos e mínimos são de, respectivamente, 63570 e 119528. A mediana localiza-se na faixa de 99482, e o primeiro e terceiro quartil estão localizados nos limites da caixa, tendo valores de 81960 e 104339 respectivamente.



Fonte: Autores (2022).

A utilização de histogramas é útil quando se tem grandes conjuntos de dados, onde é possível resumir os mesmos graficamente para melhor visualização dos dados mais frequente dentro do cenário analisado. Sendo assim, o Gráfico 2 sintetiza o histograma dos dados em estudo, onde pode-se observar que os valores no intervalo 100000-105000 apresentam uma maior frequência quando comparados às outras observações do modelo.

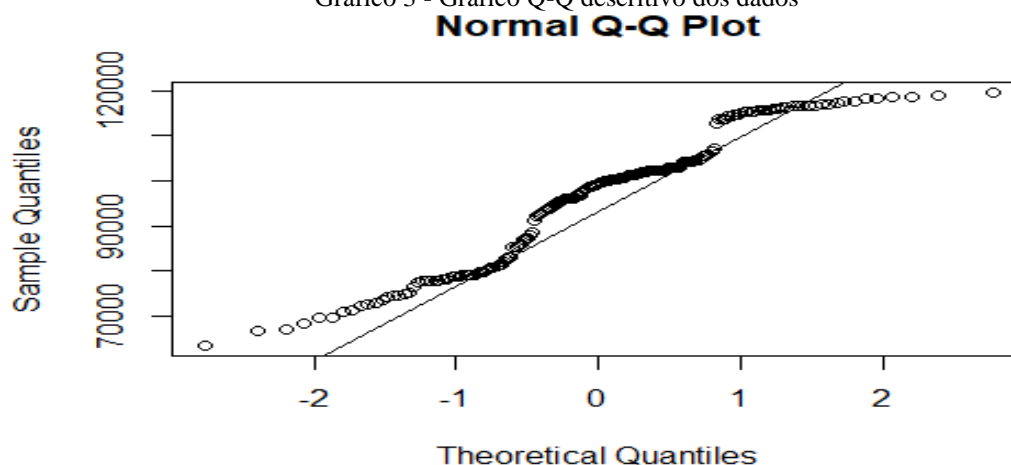
Gráfico 2 - Histograma descritivo dos dados



Fonte: Autores (2022).

Por fim, pode-se usar o Gráfico Q-Q para checar a adequação de distribuição de frequência dos dados à uma distribuição de probabilidade, onde verifica-se se os dados analisados apresentam uma distribuição normal. Sendo assim, pode-se visualizar o Gráfico Q-Q no Gráfico 3.

Gráfico 3 - Gráfico Q-Q descritivo dos dados



Fonte: Autores (2022).

Pelo Gráfico 3 acima, percebe-se que os dados não se adequam à reta de normalidade. Logo, eles não seguem uma distribuição normal.

5 ANÁLISE DOS DADOS

O banco de dados utilizado para aplicação dos algoritmos de suavização exponencial foi o histórico de cotações do Índice BOVESPA durante o ano de 2020. Com os descritos acima, neste capítulo se dará a análise deles, e se subdividir em quatro subseções, e são elas: Suavização

Exponencial Simples, Suavização Exponencial de Holt (com e sem amortecimento), Suavização Exponencial de Holt-Winters e por fim a realização do Modelo ARIMA.

5.1 Suavização Exponencial Simples

Para a Suavização Exponencial Simples, é escolhido um valor do Alfa próximo do otimizado, foi escolhido arbitrariamente o valor 0,2, visto que, pretende-se suavizar mais rápido a série temporal analisada. O intervalo de confiança definido foi (80,95). Os resultados obtidos para a previsão de 10 períodos à frente podem ser vistos na Figura 2.

Figura 2 - Suavização Exponencial Simples previsão de 10 períodos

```

      AIC      AICC      BIC
3969.422 3969.490 3975.808

Error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set -475.5121 4530.475 2690.215 -0.6984766 3.125288 1.489856 0.8005641

Forecasts:
      Point Forecast      Lo 80      Hi 80      Lo 95      Hi 95
181      100107.2 94268.66 105945.8 91177.91 109036.5
182      100107.2 94153.04 106061.4 91001.08 109213.4
183      100107.2 94039.61 106174.8 90827.61 109386.8
184      100107.2 93928.27 106286.2 90657.33 109557.1
185      100107.2 93818.90 106395.6 90490.06 109724.4
186      100107.2 93711.40 106503.1 90325.65 109888.8
187      100107.2 93605.68 106608.8 90163.96 110050.5
188      100107.2 93501.65 106712.8 90004.86 110209.6
189      100107.2 93399.23 106815.2 89848.23 110366.2
190      100107.2 93298.35 106916.1 89693.95 110520.5

```

Fonte: Autores (2022).

Para os valores da previsão ‘um passo à frente’, tem-se os seguintes resultados descritos na Figura 3.

Figura 3 - Suavização Exponencial Simples previsão um passo afrente de 10 períodos

```

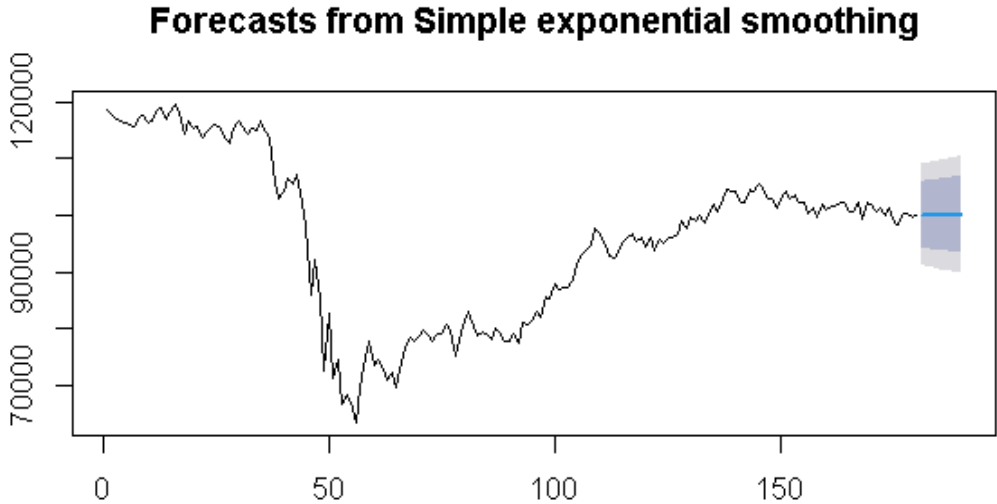
Time Series:
Start = 1
End = 180
Frequency = 1
[1] 118573.10 118573.10 118399.81 118095.43 117808.73 117496.39 117186.54 116849.91 116944.99
[10] 117082.47 116948.85 116899.92 117215.59 117544.80 117441.05 117631.11 118010.42 118083.60
[19] 117363.25 117186.40 116826.09 116566.48 116005.30 115730.08 115695.40 115761.98 115647.58
[28] 115272.12 114731.76 114859.53 115222.45 115310.44 115124.49 115161.41 115124.59 115403.19
[37] 115239.80 114928.12 113086.16 111065.63 109686.82 109074.54 108367.06 108138.49 106957.44
[46] 105165.31 101345.69 99519.44 96649.78 91836.33 90004.65 86237.33 83913.31 80509.64
[55] 78074.07 75873.13 73412.43 72675.80 73131.75 74047.34 73923.62 74066.80 73857.39
[64] 73279.25 73074.09 72366.79 72708.02 73438.04 74475.35 75116.67 75860.50 76672.07
[73] 77103.95 77245.53 77594.48 77870.14 78433.54 78681.49 78011.32 78056.77 78707.86
[82] 79600.45 79781.54 79600.48 79574.54 79472.36 79201.61 79413.95 79344.08 79049.66
[91] 78794.17 78837.49 78581.32 79103.91 79431.60 79809.17 80452.75 80796.85 81770.17
[100] 82509.92 83597.19 84267.57 84894.57 85391.18 86036.96 87038.85 88231.50 89350.93
[109] 90408.15 91855.46 92833.67 93204.14 93122.36 92972.99 93084.63 93577.16 94086.78
[118] 94583.84 94734.27 94982.44 94861.43 95085.76 94835.51 95015.47 95023.54 95259.48
[127] 95454.57 95716.63 96360.73 96640.80 97266.61 97645.36 98122.65 98237.53 98678.07
[136] 99300.57 99551.11 100218.54 101060.10 101710.03 102225.94 102239.41 102267.85 102709.69
[145] 102989.57 103512.69 103811.89 103631.96 103471.56 103020.42 102976.69 103206.48 103120.29
[154] 103185.13 102982.98 102809.95 102340.08 102142.75 101633.28 101719.70 101546.50 101530.77
[163] 101528.88 101682.69 101769.68 101541.21 101357.70 101514.74 101085.63 101302.03 101423.85
[172] 101283.35 101275.03 101030.11 101082.50 100632.92 100178.98 100198.08 100218.05 100109.58

```

Fonte: Autores (2022).

A visualização gráfica da série é vista logo a seguir no Gráfico 4.

Gráfico 4 - Visualização gráfica da Suavização Exponencial Simples previsão de 10 períodos



Fonte: Autores (2022).

5.2 Suavização Exponencial de Holt

5.2.1 Sem amortecimento

Para a Suavização Exponencial de Holt sem amortecimento, foi escolhido os valores de Alfa e Beta próximos do otimizado, com valores de 0,2 e 0,1 respectivamente, visto que se pretende uma suavização mais forte para o nível e tendência apresentados na série temporal analisada. O intervalo de confiança permanece o mesmo (80,95). Os resultados obtidos para a previsão de 10 períodos a frente são sintetizados na Figura 5 e Figura 6.

Figura 5 - Suavização Exponencial de Holt previsão sem amortecimento para 10 períodos

Error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	210.0463	4301.103	2864.954	0.2858058	3.317833	1.586628	0.7727584

Forecasts:

	Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
181		99546.58	94034.50	105058.7	91116.57	107976.6
182		99436.31	93681.52	105191.1	90635.12	108237.5
183		99326.03	93163.33	105488.7	89901.00	108751.1
184		99215.76	92464.86	105966.7	88891.15	109540.4
185		99105.49	91588.00	106623.0	87608.49	110602.5
186		98995.21	90545.35	107445.1	86072.26	111918.2
187		98884.94	89353.65	108416.2	84308.09	113461.8
188		98774.67	88029.63	109519.7	82341.56	115207.8
189		98664.39	86588.02	110740.8	80195.17	117133.6
190		98554.12	85041.08	112067.2	77887.70	119220.5

Fonte: Autores (2022).

Para os valores da previsão ‘um passo à frente’, tem-se os seguintes resultados, conforme a Figura 6.

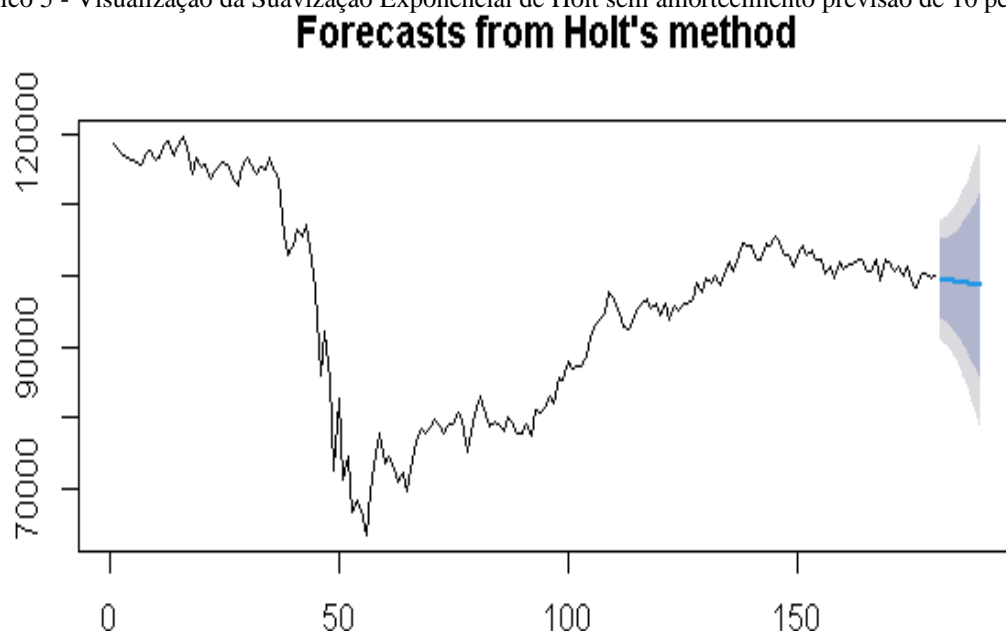
Figura 6 - Suavização Exponencial de Holt previsão um passo à frente sem amortecimento de 10 períodos

```
Time Series:
Start = 1
End = 180
Frequency = 1
[1] 117706.66 117030.84 116330.41 115615.26 115020.89 114486.93 114028.98 113603.37 113701.70
[10] 113920.39 113901.62 114000.63 114524.20 115106.48 115243.57 115689.26 116349.84 116688.58
[19] 116136.53 116101.17 115839.73 115652.98 115112.23 114843.71 114828.65 114934.91 114857.35
[28] 114489.63 113917.07 114048.15 114466.25 114622.29 114485.96 114579.04 114595.10 114954.46
[37] 114848.32 114559.10 112558.28 110219.19 108464.56 107514.85 106497.87 106036.23 104592.66
[46] 102458.59 98037.60 95613.79 92057.23 86304.76 83649.33 78973.38 75835.34 71601.64
[55] 68436.66 65624.84 62634.33 61615.75 62112.95 63373.45 63726.79 64469.85 64911.35
[64] 64975.05 65428.93 65331.03 66334.62 67795.00 69633.19 71076.19 72616.55 74211.39
[73] 75362.28 76128.06 77033.62 77793.34 78801.87 79423.36 78970.15 79174.56 79995.56
[82] 81087.58 81416.58 81303.04 81294.47 81161.58 80805.39 80938.55 80767.85 80334.84
[91] 79917.23 79812.74 79393.18 79821.09 80091.46 80447.73 81125.87 81518.54 82613.64
[100] 83507.91 84807.56 85690.67 86522.11 87199.36 88018.08 89218.88 90646.34 92017.24
[109] 93328.05 95064.56 96307.77 96857.80 96838.42 96649.72 96667.51 97062.57 97475.45
[118] 97877.07 97900.31 98008.24 97702.41 97744.50 97270.25 97240.33 97036.80 97086.77
[127] 97116.07 97238.46 97804.81 98021.79 98632.10 99009.01 99505.29 99619.20 100075.38
[136] 100744.68 101028.85 101760.36 102706.52 103472.19 104097.03 104161.58 104195.27 104646.96
[145] 104923.96 105458.40 105757.66 105520.87 105261.16 104649.67 104440.70 104532.00 104299.89
[154] 104230.88 103880.53 103553.67 102898.88 102522.71 101811.63 101741.82 101425.88 101296.81
[163] 101208.72 101315.36 101380.66 101119.77 100900.40 101053.61 100587.74 100806.33 100952.00
[172] 100825.97 100837.54 100592.79 100659.30 100184.52 99673.99 99659.85 99665.97 99546.61
```

Fonte: Autores (2022).

A visualização gráfica da série é vista conforme o Gráfico 5.

Gráfico 5 - Visualização da Suavização Exponencial de Holt sem amortecimento previsão de 10 períodos



Fonte: Autores (2022).

5.2.2 Com amortecimento

Para a Suavização Exponencial de Holt com amortecimento, será mantido os mesmos valores para os coeficientes Alfa e Beta. Além disso, o valor de *Phi* otimizado foi de 0,8. O

intervalo de confiança permanece o mesmo (80,95), e os resultados obtidos para a previsão de 10 períodos a frente são sintetizados na Figura 7.

Figura 7 - Suavização Exponencial de Holt previsão com amortecimento para 10 períodos

	AIC	AICC	BIC
	3837.277	3837.413	3846.855

Error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	-143.1831	3121.149	1933.262	-0.1808754	2.194172	1.070651	0.5472572

Forecasts:

Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
181	99632.09	95575.44	103688.7	93427.98	105836.2
182	99622.65	95409.97	103835.3	93179.92	106065.4
183	99615.09	95177.30	104052.9	92828.07	106402.1
184	99609.05	94890.55	104327.5	92392.73	106825.4
185	99604.21	94564.88	104643.5	91897.23	107311.2
186	99600.34	94213.93	104986.8	91362.53	107838.2
187	99597.25	93848.51	105346.0	90805.31	108389.2
188	99594.77	93476.64	105712.9	90237.89	108951.7
189	99592.79	93103.97	106081.6	89669.00	109516.6
190	99591.21	92734.36	106448.0	89104.57	110077.8

Fonte: Autores (2022).

Para os valores da previsão ‘um passo à frente’, tem-se os seguintes resultados descritos na Figura 8.

Figura 8 - Suavização Exponencial de Holt previsão um passo a frente sem amortecimento de 10 períodos

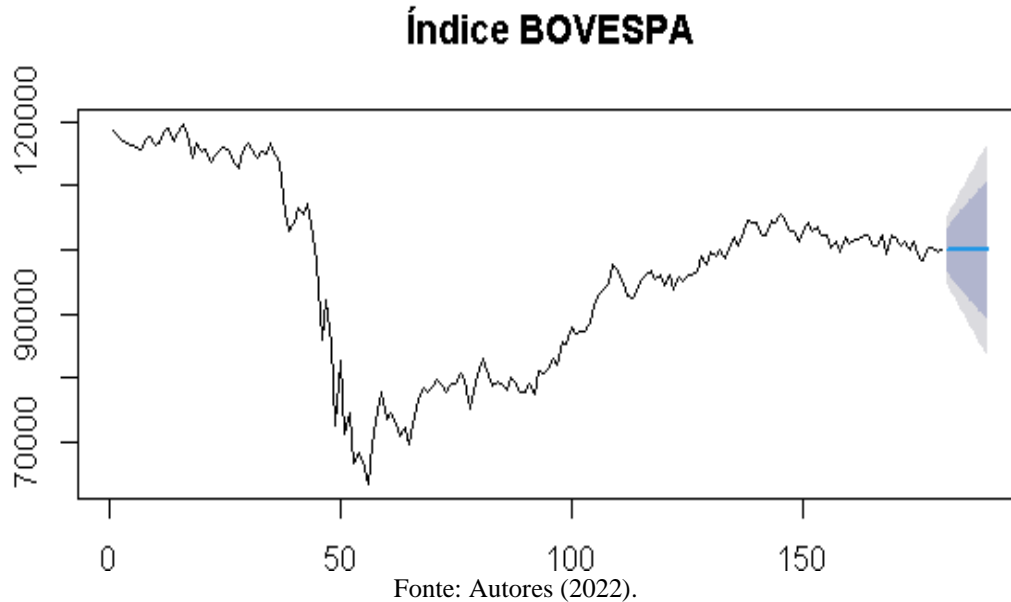
Time Series:
Start = 1
End = 180
Frequency = 1

[1]	117783.14	117457.43	117140.25	116773.18	116490.35	116213.75	115956.75	115666.87	115971.86
[10]	116415.46	116504.34	116631.58	117218.48	117852.61	117865.56	118155.40	118687.37	118806.31
[19]	117740.37	117226.34	116501.32	115943.43	115041.64	114553.98	114510.61	114740.42	114807.32
[28]	114498.54	113877.53	114107.31	114770.95	115140.77	115081.18	115218.93	115225.01	115630.46
[37]	115455.62	114986.05	112299.29	109024.32	106536.09	105346.96	104434.61	104455.39	103403.48
[46]	101403.62	96374.60	93640.38	89747.23	83181.68	80534.53	75874.80	73293.54	69638.13
[55]	67371.81	65683.08	63788.72	64274.74	66703.70	70019.61	71865.95	73574.14	74341.70
[64]	74099.45	73928.77	72858.05	73044.25	73926.69	75417.85	76493.06	77646.94	78831.20
[73]	79415.26	79433.51	79580.55	79598.92	80015.56	80078.93	78854.85	78462.78	79045.56
[82]	80210.94	80565.79	80329.40	80170.21	79870.40	79317.09	79419.69	79250.94	78786.64
[91]	78351.79	78350.75	78022.12	78774.48	79419.83	80153.14	81240.56	81911.80	83350.04
[100]	84493.65	86036.24	86973.50	87697.40	88118.31	88646.72	89661.06	91024.09	92365.21
[109]	93625.50	95395.60	96546.88	96730.70	96073.58	95142.10	94537.54	94594.37	94906.92
[118]	95378.27	95477.08	95702.35	95431.88	95581.84	95124.36	95209.03	95136.10	95401.07
[127]	95675.79	96067.08	97009.49	97514.64	98429.97	99005.91	99637.05	99714.62	100130.25
[136]	100811.53	101018.60	101749.08	102783.96	103610.72	104218.07	104056.41	103766.09	104000.88
[145]	104105.25	104591.41	104857.20	104458.47	103994.70	103106.17	102754.73	102906.13	102771.35
[154]	102872.89	102650.85	102435.71	101795.86	101476.13	100764.52	100860.39	100727.08	100828.92
[163]	100985.73	101367.82	101673.49	101505.13	101291.01	101499.29	100933.16	101166.74	101364.52
[172]	101223.55	101218.77	100884.91	100918.75	100297.11	99591.78	99528.11	99583.47	99530.41

Fonte: Autores (2022).

A visualização gráfica da série é vista no Gráfico 6.

Gráfico 6 - Visualização da Suavização Exponencial de Holt com amortecimento previsão de 10 períodos



5.3 Suavização Exponencial de Holt-Winters

Para a Suavização Exponencial de Holt-Winters, a frequência da sazonalidade adotada foi de 5 dias, já que as cotações do Índice IBOVESPA tendem a sofrer variações frequentes toda semana a partir da segunda-feira, finalizando suas operações nas sextas-feiras.

Os valores otimizados dos coeficientes Alfa e Beta foram mantidos, e o valor otimizado de Gama foi de 0,1, visto que se pretende uma suavização mais forte da sazonalidade apresentada na série temporal. O intervalo de confiança permanece o mesmo (80,95), e os resultados obtidos para a previsão dos 10 períodos subsequentes são sintetizados na Figura 9.

Figura 9 - Suavização Exponencial de Holt-Winters previsão para 5 períodos

	AIC	AICC	BIC				
	3849.829	3850.480	3872.180				
Error measures:							
	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	-137.5574	3160.87	1989.633	-0.1749338	2.254535	0.5727024	0.5285491
Forecasts:							
Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95		
41.00	99350.15	95181.90	103518.4	92975.36	105725.0		
41.20	99710.80	95382.23	104039.4	93090.83	106330.8		
41.40	99448.60	94888.72	104008.5	92474.86	106422.3		
41.60	99753.64	94905.33	104601.9	92338.79	107168.5		
41.80	99698.19	94520.22	104876.2	91779.17	107617.2		
42.00	99314.70	93619.52	105009.9	90604.68	108024.7		
42.20	99682.44	93624.83	105740.1	90418.12	108946.8		
42.40	99425.91	92997.63	105854.2	89594.71	109257.1		
42.60	99735.49	92934.27	106536.7	89333.92	110137.1		
42.80	99683.67	92511.35	106856.0	88714.56	110652.8		

Fonte: Autores (2022).

É demonstrado na Figura 10 os valores da previsão ‘um passo à frente’.

Figura 10 - Suavização Exponencial de Holt-Winters previsão um passo afrente de 5 períodos

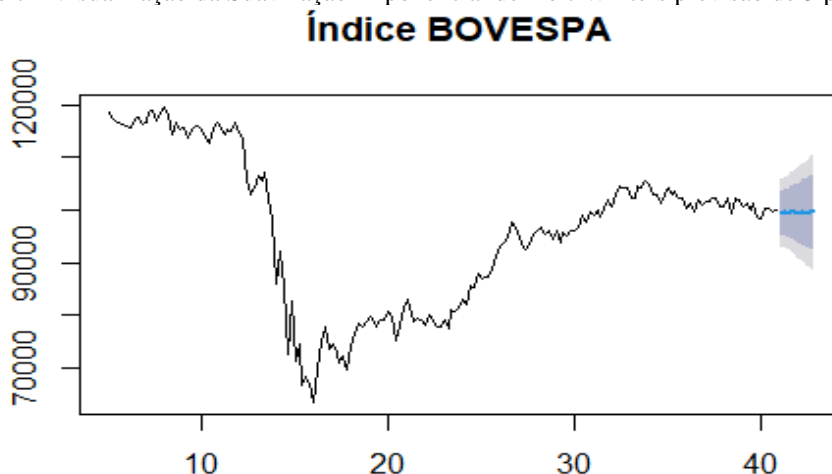
```
Time Series:
Start = c(5, 1)
End = c(40, 5)
Frequency = 5
```

[1]	117834.59	117827.80	116805.23	116455.69	116799.35	116121.52	116171.82	115275.90	115666.56
[10]	116692.26	116453.84	116851.06	117118.80	117761.47	118063.78	118080.72	119015.80	118789.12
[19]	117447.82	117370.18	116499.78	116119.65	114546.41	114259.50	114588.73	114822.90	114933.49
[28]	114228.64	113883.68	114415.24	114891.71	115110.83	114622.39	115392.01	115733.15	115738.12
[37]	115246.61	114525.20	112360.37	109571.55	106485.30	104985.65	103152.57	103888.07	103964.40
[46]	102011.65	96603.89	93153.50	89156.32	83224.66	79717.76	76223.50	72616.35	68109.81
[55]	68401.76	64883.97	64867.19	63728.51	65830.94	71331.27	71201.65	75293.43	74825.61
[64]	73938.35	74632.62	71730.93	73770.68	73496.06	74846.29	76643.38	76957.68	80007.01
[73]	79621.92	79110.46	79819.86	79014.15	80818.15	79918.89	78433.02	78578.73	78646.68
[82]	80891.64	79982.60	80057.24	80725.77	79988.76	79868.41	78629.52	78856.63	79139.84
[91]	78305.98	79025.20	77346.98	78331.03	79700.00	80265.66	81819.26	81695.27	83096.08
[100]	84804.02	86250.94	87356.09	87671.40	87830.79	88988.71	89639.34	91157.45	92129.79
[109]	93314.69	95876.43	96780.13	96989.19	95904.12	95046.84	94811.99	94356.34	94608.31
[118]	94898.14	95414.23	96288.95	95570.85	95593.30	94706.07	95176.12	95471.53	95569.28
[127]	95489.43	95781.90	96967.49	97928.22	98648.06	98911.58	99599.04	99602.87	100573.37
[136]	100902.36	100879.09	101472.97	102674.93	104109.19	104204.29	104083.59	103716.06	103921.42
[145]	104464.02	104253.56	104669.42	104518.69	103966.01	103612.78	102494.21	102520.94	102689.84
[154]	102643.49	103221.91	102490.74	101515.97	101512.87	100485.10	101314.13	100593.99	100599.02
[163]	100893.95	101351.22	102125.09	101497.26	101150.91	101508.17	100909.80	101368.23	101224.35
[172]	101179.06	100995.30	101037.34	101192.91	100107.52	99565.88	99218.51	99793.67	99590.22

Fonte: Autores (2022).

A visualização gráfica da série é vista logo a seguir.

Gráfico 7 - Visualização da Suavização Exponencial de Holt-Winters previsão de 5 períodos



Fonte: Autores (2022).

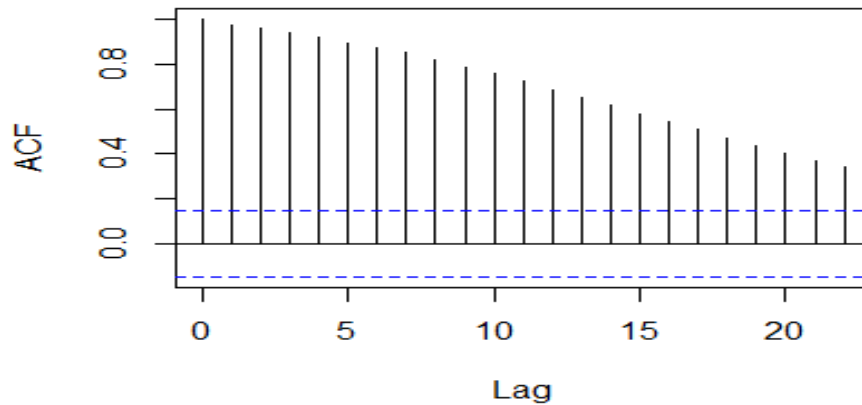
Pelos resultados do AIC, vemos que a Suavização Exponencial de Holt com amortecimento apresenta o menor valor do AIC, sendo o modelo de melhor ajuste para o conjunto de dados. Além disso, ele apresenta as menores medidas de erro em comparação com os demais modelos de previsão, o que reforça a afirmação acima.

5.4 Modelo ARMA

Para avaliar os dados pelo modelo ARMA, começou avaliando a estacionariedade dos dados e realizando testes de autocorrelação e autocorrelação parcial, para conseguir as ordens nas quais o modelo será montado conforme o Gráfico 8 e 9.

Gráfico 8 - Função de Autocorrelação

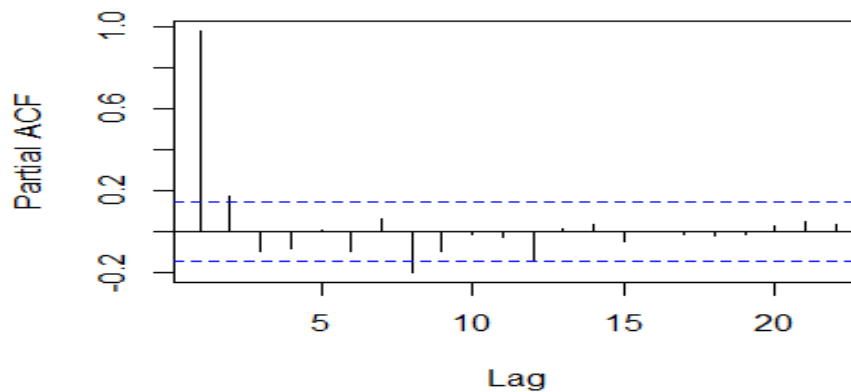
Series dados



Fonte: Autores (2022).

Gráfico 9 - Função de Autocorrelação Parcial

Series dados



Fonte: Autores (2022).

Nota-se que os correlogramas das cotações lembram os correlogramas de um *Random Walk*, apesar de não se tratar de um *Random Walk*, o que diz que primeiramente que os dados não são estacionários. Contudo se seguiu na análise, indo para o *Augmented Dickey-Fuller test* na Figura 11.

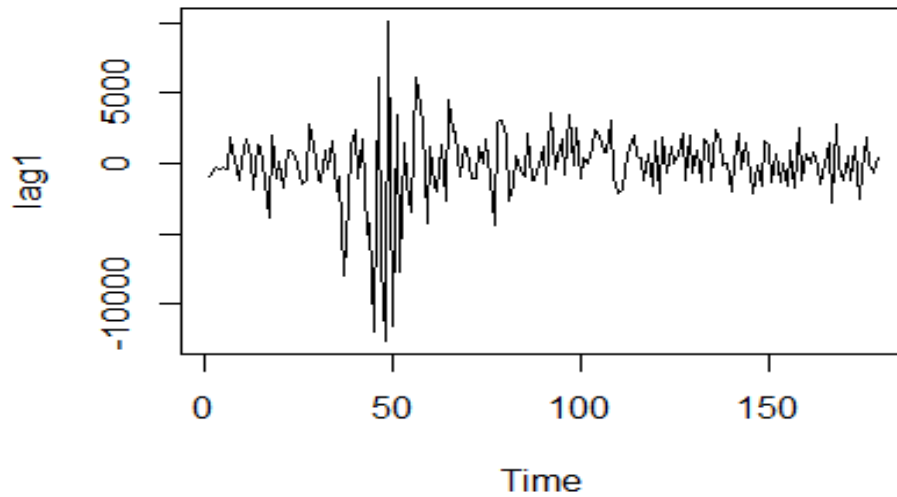
Figura 11 - Visualização *Augmented Dickey-Fuller test* primeira tentativa

```
Augmented Dickey-Fuller Test
data: dados
Dickey-Fuller = -1.6014, Lag order = 5, p-value = 0.7428
alternative hypothesis: stationary
```

Fonte: Autores (2022).

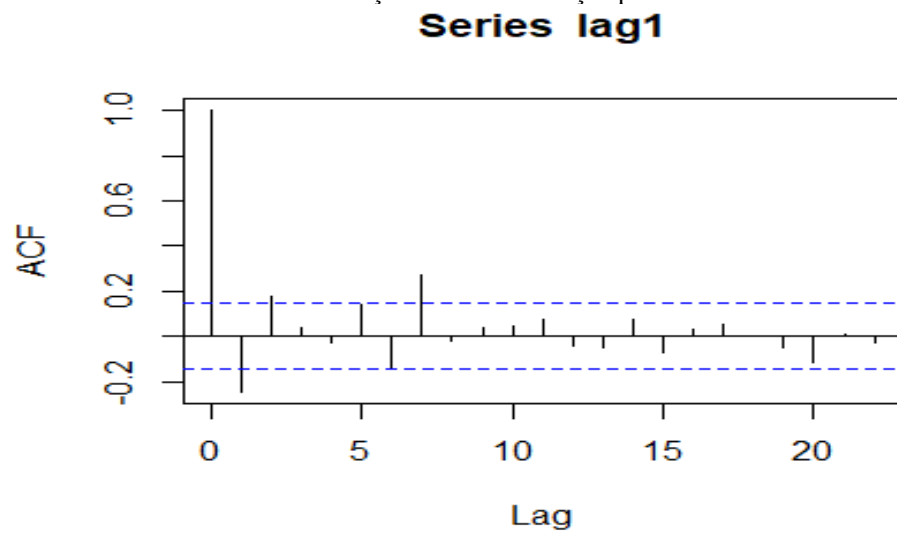
Percebe-se que os dados são não-estacionários. Para encontrar as ordens dos modelos, foi realizado diferenciações até os dados se tornarem estacionários conforme Gráficos 10, 11 e 12., em seguida se realizou o modelo ARMA e simulações.

Gráfico 10 - Primeira diferenciação



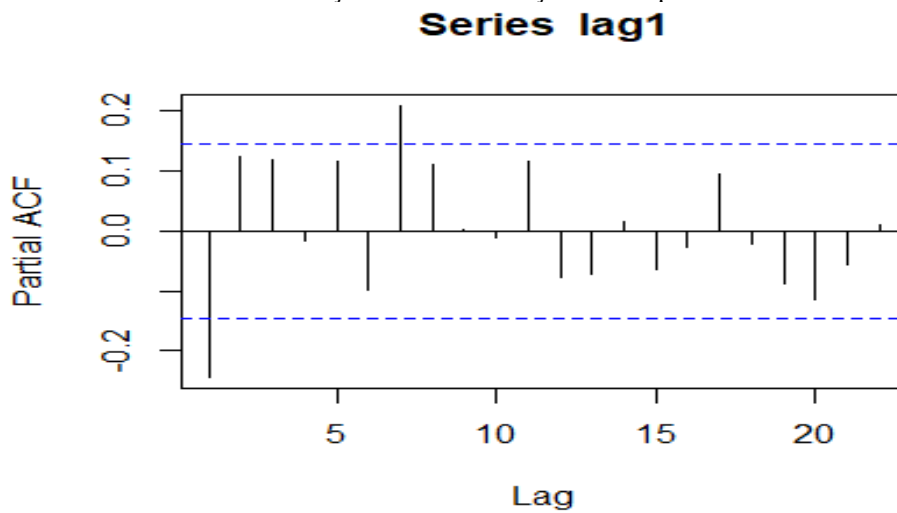
Fonte: Autores (2022).

Gráfico 11 - Função de Autocorrelação primeira tentativa



Fonte: Autores (2022).

Gráfico 12 - Função de Autocorrelação Parcial primeira tentativa



Fonte: Autores (2022).

Figura 12 - Visualização *Augmented Dickey-Fuller test* segunda tentativa

```
Augmented Dickey-Fuller Test

data: lag1
Dickey-Fuller = -5.1153, Lag order = 5, p-value = 0.01
alternative hypothesis: stationary
```

Fonte: Autores (2022).

Se percebe que os *lag* 1, 2 e 7 podem ser utilizados para o modelo. Logo tem-se, para ARMA (1,1), ARMA (2,2), ARMA (7,7) às demonstrações conforme as Figuras 13, 14 e 15 respectivamente.

Figura 13 - Modelo ARMA (1,1)

```
Call:
arima(x = dados, order = c(1, 0, 1))

Coefficients:
      ar1      ma1  intercept
    0.9885 -0.1770 102319.82
s.e. 0.0091  0.0621 10342.25

sigma^2 estimated as 6952365: log likelihood = -1675.04, aic = 3358.07
```

Fonte: Autores (2022).

Figura 14 - Modelo ARMA (2,2)

```
Call:
arima(x = dados, order = c(2, 0, 2))

Coefficients:
      ar1      ar2      ma1      ma2  intercept
    1.1697 -0.1830 -0.4025  0.2383 101156.876
s.e. 0.3373  0.3337  0.3264  0.0848  9551.608

sigma^2 estimated as 6650611: log likelihood = -1671.13, aic = 3354.26
```

Fonte: Autores (2022).

Figura 15 - Modelo ARMA (7,7)

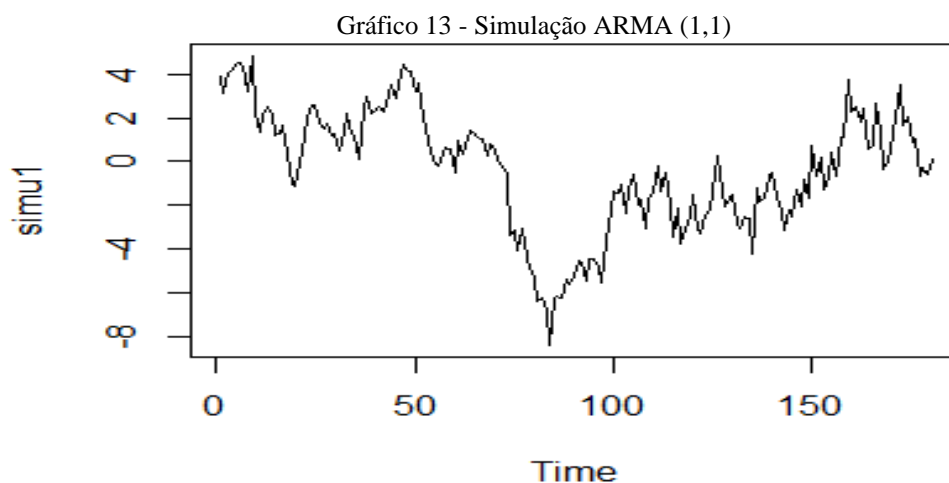
```
Call:
arima(x = dados, order = c(7, 0, 7))

Coefficients:
      ar1      ar2      ar3      ar4      ar5      ar6      ar7      ma1      ma2      ma3      ma4      ma5
    1.0414  0.2033  0.2664 -0.3107 -0.5638  0.2244  0.1256 -0.2689 -0.0726 -0.4866 -0.1522  0.4317
s.e. 0.2393  0.3134  0.3356  0.2936  0.2819  0.3361  0.2300  0.2315  0.2133  0.2260  0.2025  0.2149
      ma6      ma7  intercept
    -0.1124  0.3296 99805.261
s.e. 0.2459  0.1034  7927.238

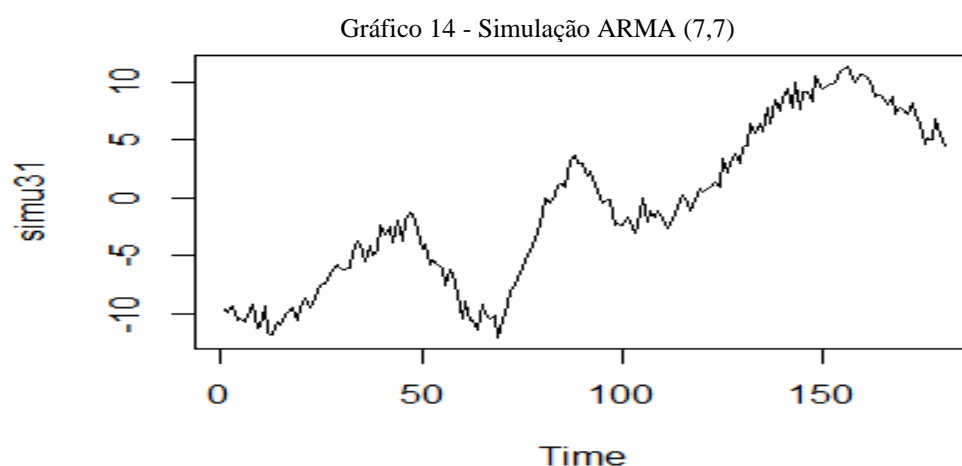
sigma^2 estimated as 5794506: log likelihood = -1659.47, aic = 3350.94
```

Fonte: Autores (2022).

É possível visualizar que o modelo ARMA (7,7) tem menor AIC e menor variância, contudo, para comparar, foi realizado simulações dos dois modelos e a fim de realizar previsões utilizando os dois modelos. Logo, tem-se para ARMA (1,1) e para ARMA (7,7) às configurações descritas nos Gráficos 13 e 14 respectivamente.



Fonte: Autores (2022).



Fonte: Autores (2022).

É passível de observação que a simulação feita com o modelo ARMA (1,1) lembra mais o gráfico do índice IBOVESPA, já a simulação do modelo ARMA (7,7) parece diferente, todavia isso não quer dizer que o modelo ARMA (1,1) é melhor que o ARMA (7,7), para definir isso é necessário materializar as previsões.

As previsões foram iniciadas com os próprios dados, para o período imediatamente depois e para os 10 próximos para ARMA (1,1) descritos na Figura 16 e Gráfico 15.

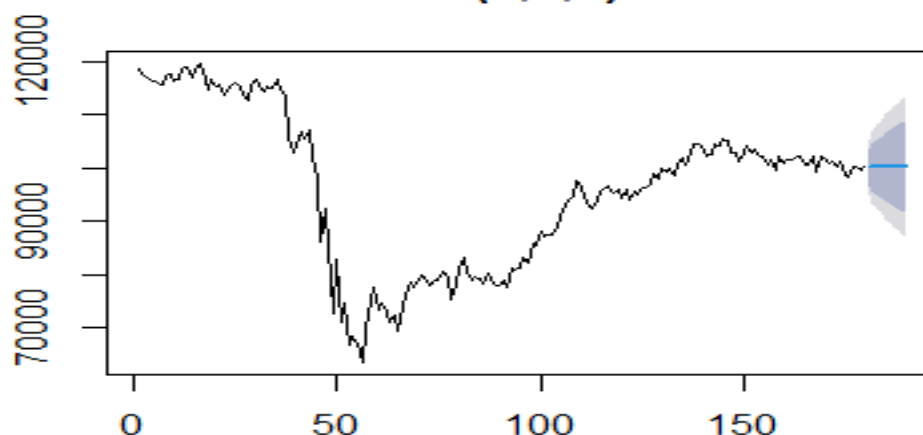
Figura 16 - Previsões para o modelo ARMA (1,1)

\$pred		> forecast(modelao)					
Time Series:		Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Start = 181	181	100072.5	96693.39	103451.6	94904.60	105240.4	
End = 181	182	100098.3	95746.51	104450.0	93442.83	106753.7	
Frequency = 1	183	100123.8	94996.79	105250.7	92282.73	107964.8	
[1] 100072.5	184	100148.9	94363.94	105933.9	91301.55	108996.3	
	185	100173.8	93811.23	106536.4	90443.06	109904.6	
\$se	186	100198.4	93318.08	107078.8	89675.84	110721.1	
Time Series:	187	100222.8	92871.61	107573.9	88980.13	111465.4	
Start = 181	188	100246.8	92463.05	108030.6	88342.57	112151.1	
End = 181	189	100270.6	92086.15	108455.1	87753.56	112787.6	
Frequency = 1	190	100294.1	91736.21	108852.0	87205.94	113382.3	
[1] 2636.734							

Fonte: Autores (2022).

Gráfico 15 – Simulação gráfica da previsão para o modelo ARMA (1,1)

Forecasts from ARIMA(1,0,1) with non-zero me



Fonte: Autores (2022).

O mesmo procedimento foi realizado para ARMA (7,7), e obteve-se como resultado o que está descrito na Figura 17 e Gráfico 16.

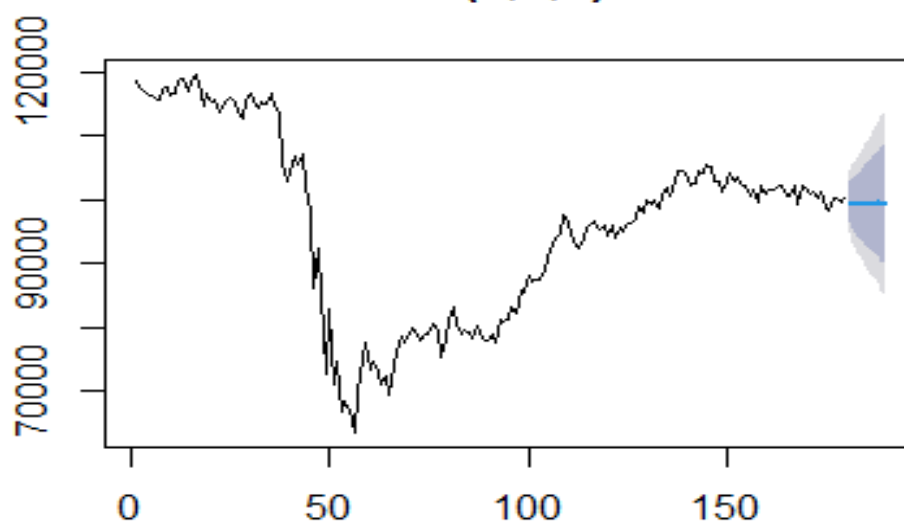
Figura 17 - Previsões para o modelo ARMA (7,7)

\$pred		> forecast(modelao3)				
Time Series:	Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Start = 181	181	99705.24	96620.32	102790.2	94987.26	104423.2
End = 181	182	99355.16	95457.02	103253.3	93393.47	105316.9
Frequency = 1	183	99107.86	94258.29	103957.4	91691.08	106524.6
[1] 99705.24	184	99315.62	93711.01	104920.2	90744.11	107887.1
	185	99294.50	93065.19	105523.8	89767.59	108821.4
\$se	186	99106.02	92181.77	106030.3	88516.29	109695.7
Time Series:	187	99290.16	91909.41	106670.9	88002.28	110578.0
Start = 181	188	99422.37	91279.40	107565.3	86968.77	111876.0
End = 181	189	99337.23	90503.22	108171.2	85826.78	112847.7
Frequency = 1	190	99410.54	89920.35	108900.7	84896.55	113924.5
[1] 2407.178						

Fonte: Autores (2022).

Gráfico 16 - Simulação gráfica da previsão para o modelo ARMA (7,7)

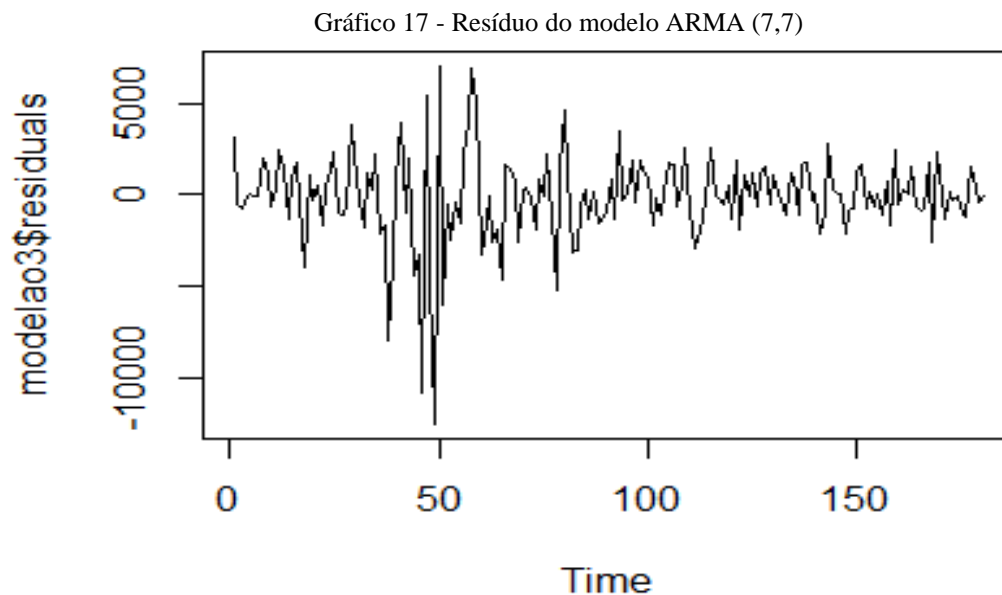
Forecasts from ARIMA(7,0,7) with non-zero me



Fonte: Autores (2022).

De acordo com os dados de previsão de exatamente um passo à frente, observa-se que o ARMA (7,7) apresenta um erro padrão menor, o que colabora com a escolha desse modelo em vez do ARMA (1,1).

Para finalizar, foram analisados os resíduos do modelo ARMA (7,7) que se mostrou o melhor modelo, conforme é possível visualizar no Gráfico 17.



Fonte: Autores (2022).

Aparentemente, os erros são heterocedásticos para confirmar ou negar isso, realizou-se o teste Box-Ljung para analisar se os erros são dependentes, tem-se.

Figura 18 - Teste de Box-Ljung do modelo ARMA (7,7)

```
> boxljung <- Box.test(modelao3$residuals, lag=1, type=c('Ljung-Box'))
> boxljung

Box-Ljung test

data: modelao3$residuals
X-squared = 0.013681, df = 1, p-value = 0.9069
```

Fonte: Autores (2022).

De acordo com o teste Box-Ljung demonstrado na Figura 18, não há evidências suficientes para afirmar que os erros são dependentes. Com isso, acredita-se que os erros, apesar de serem heterocedásticos, não possuem uma heterocedasticidade condicional.

6 CONCLUSÃO

Concluiu-se que, considerando a teoria do mercado eficiente, não se mostra adequada a utilização de modelos ARMA ou ARIMA para previsão dos dados. Ao analisar os resultados gráficos, foi observado que algumas das condições de ajuste do modelo não eram atendidas, e

que o método não estava bem ajustado à série. Sendo, portanto, recomendado a avaliação de uso de outros modelos, o que fica para estudos futuros. Contudo, sem essa consideração, se observa que os dados mais se adequam ao ARMA pois só é necessária uma diferenciação para torná-los estacionários, e tem ajustes razoáveis também com os modelos ARMA testados.

Para próximos trabalhos, poderiam ser realizadas regressões, deixando de ver as cotações somente como uma série temporal, mas como uma variável dependente de outras, como volatilidade, câmbio, Índice de Preço ao Consumidor Amplo e o Sistema Especial de Liquidação de Custódia. Pode-se também utilizar uma maior base de dados, ou diminuir os intervalos para analisar mês a mês, para tentar ver o problema por ângulos diferentes.

Uma outra alternativa seria utilizar índices com menos papéis, determinados portfólios de ações, ou ações de determinado segmento, pois a análise e previsão seria mais precisa e possível. Sempre levando em conta que a existência de um pequeno erro é inevitável, pois é a soma de inúmeros fatores que ocasiona a rentabilidade diária e um modelo estatístico capaz de reproduzir esses efeitos teria valor imensurável.

REFERENCES

1. Abelém, A. J. G. **Redes neurais artificiais na previsão de séries temporais**. 1994. Dissertação (Mestrado em Engenharia Elétrica) - Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 1994.
2. Albuquerque, R. C. De. **Modelagem em séries temporais aplicados a dados climatológicos no sertão paraibano**. 2015. Trabalho de Conclusão de Curso (Bacharelado em Estatística) - Universidade Estadual da Paraíba, Campina Grande, 2015.
3. Braga, P. C. A. S. **Previsão de IBOVESPA utilizando modelos híbridos**. 2006, Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2006.
4. B3. **Índice Ibovespa (Ibovespa B3)**. Disponível em: https://www.b3.com.br/pt_br/market-data-e-indices/indices/indices-amplos/ibovespa.htm. Acesso em: 15 mar. 2022.
5. Corrar, L. J; Theóphilo, C. R. **Pesquisa operacional para decisão em contabilidade e administração: contabilometria**. São Paulo: Atlas, 2004.
6. Davila, V. H. L. **Introdução a séries temporais**. Notas de Aula. Disponível em: https://edisciplinas.usp.br/pluginfile.php/176835/mod_resource/content/1/Slides%20%20Introdução%20às%20series%20temporais.pdf. Acesso em: 10 nov. 2020.
7. Ferreira, T. A. E. **Uma nova metodologia híbrida inteligente para previsão de séries temporais**. 2006. Dissertação (Mestrado em Ciência da Computação) - Universidade Federal de Pernambuco, Recife, 2006.
8. Gaither, N; Frazier, G. **Production and operations management**. 8. ed. São Paulo: Thomson, 200.
9. Jung, C. F. **Metodologia Para a Pesquisa & Desenvolvimento**. Rio de Janeiro: Axcel Books do Brasil Editora, 2004.
10. Keiel, G.; Bender, F. A. **Modelagem de séries temporais financeiras: uma abordagem estatística para a identificação de modelos de média condicional**. *Scientia cum Industria*, v. 6, n. 1, p. 22-28, 2018.
11. Kirsten, H. A. **Comparação entre os modelos hold-winters e redes neurais para previsão de séries temporais financeiras**. 2009. Dissertação (Mestrado em Engenharia de Produção e Sistemas) - Pontifícia Universidade Católica do Paraná, Curitiba, 2009.
12. MENSURAÇÃO FLORESTAL, **sobre o R**. Disponível em: <http://www.mensuracaoflorestal.com.br/sobre-o-r>. Acesso em: 15 mar. 2022.

13. Nepomuceno, T. C. C.; De Carvalho, V. D. H.; Costa, A. P. C. S. Time-series directional efficiency for knowledge benchmarking in service organizations. In: World conference on information systems and technologies. Springer, Cham. p. 333-339, 2020.
14. IBOVESPA. **Índice Bovespa.** 2020. Disponível em: <http://www.bovespa.com.br/Mercado/RendaVariavel/Indices/FormConsultaApresentacaoP.asp?Indice-Ibovespa>. Acesso em: 10 out. 2020.
15. Silva, A. L. C. De L Da; Costa, A. P. C. S. FITradeoff decision support system: an exploratory study with neuroscience tools. In: **Information Systems and Neuroscience**. Springer, Cham, 2020. p. 365-372.
16. Tubino, D. F. **Manual de planejamento e controle da produção.** 2. ed. São Paulo: Atlas, 2000.
17. Xavier, J. M. N. **Análise e previsão de séries temporais com modelos ARIMA e análise espectral singular.** 2016. Dissertação (Mestrado em Bioestatística e Biometria) - Universidade Aberta, 2016.
18. Yule, G. U. V. On a method of investigating periodicities disturbed series, with special reference to Wolfer's sunspot numbers. **Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character**, v. 226, n. 636-646, p. 267-298, 1927.