

UNIVERSIDADE DE CAXIAS DO SUL
ÁREA DO CONHECIMENTO DE CIÊNCIAS EXATAS E ENGENHARIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA MECÂNICA
MESTRADO PROFISSIONAL EM ENGENHARIA MECÂNICA

RAFAELA BOEIRA CECHIN

**ANÁLISE DE PREVISÃO DE PREÇOS DE AÇÕES DE UMA CARTEIRA
OTIMIZADA, UTILIZANDO ANÁLISE ENVOLTÓRIA DE DADOS, REDES
NEURAIS ARTIFICIAIS E MODELO DE BOX-JENKINS**

CAXIAS DO SUL

2018

RAFAELA BOEIRA CECHIN

**ANÁLISE DE PREVISÃO DE PREÇOS DE AÇÕES DE UMA CARTEIRA
OTIMIZADA, UTILIZANDO ANÁLISE ENVOLTÓRIA DE DADOS, REDES
NEURAIS ARTIFICIAIS E MODELO DE BOX-JENKINS**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Mecânica da Universidade de Caxias do Sul, visando à obtenção do título de Mestre em Engenharia Mecânica.

Área de concentração: Processos de Fabricação e Produção de Componentes, Ferramentas e Matrizes.

Orientador: Prof. Dr. Leandro Luís Corso
Coorientador: Prof. Dr. Alexandre Fassini Michels

**CAXIAS DO SUL
2018**

Dados Internacionais de Catalogação na Publicação (CIP)
Universidade de Caxias do Sul
UCS - BICE - Processamento Técnico

C387a Cechin, Rafaela Boeira, 1991-

Análise de previsão de preços de ações de uma carteira otimizada, utilizando análise envoltória de dados, redes neurais artificiais e modelo de Box-Jenkins / Rafaela Boeira Cechin. – 2018.

164 f. : il ; 30 cm

Apresenta bibliografia.

Dissertação (Mestrado) – Universidade de Caxias do Sul, Programa de Pós-Graduação em Engenharia Mecânica, 2018.

Orientação: Prof. Dr. Leandro Luís Corso

Coorientação: Prof. Dr. Alexandre Fassini Michels

1. Pesquisa operacional. 2. Análise envoltória de dados. 3. Redes neurais (Computação). 4. Previsão econômica. 5. Bolsa de valores. I. Corso, Leandro Luís, orient. II. Michels, Alexandre Fassini, coorient. III. Título.

CDU 2. ed.: 519.8

Índice para o catálogo sistemático:

1. Pesquisa operacional	519.8
2. Análise envoltória de dados	519.87
3. Redes neurais (Computação)	007.52
4. Previsão econômica	519.22
5. Bolsa de valores	336.76

Catalogação na fonte elaborada pela bibliotecária
Ana Guimarães Pereira – CRB 10/1460

“ANÁLISE DE PREVISÃO DE PREÇOS DE AÇÕES DE UMA CARTEIRA OTIMIZADA, UTILIZANDO ANÁLISE ENVOLTÓRIA DE DADOS, REDES NEURAIS ARTIFICIAIS E MODELO DE BOX-JENKINS”

Rafaela Boeira Cechin

Dissertação de Mestrado submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Mecânica da Universidade de Caxias do Sul, como parte dos requisitos necessários para a obtenção do título de Mestre em Engenharia Mecânica, Área de Concentração: Projeto e Fabricação.

Caxias do Sul, 16 de março de 2018.

Banca Examinadora:

Prof. Dr. Leandro Luís Corso
Orientador
Universidade de Caxias do Sul

Prof. Dr. Alexandre Fassini Michels
Coorientador
Universidade de Caxias do Sul

Prof. Dr. Guilherme Holsbach Costa
Universidade de Caxias do Sul

Prof. Dr. Marcos Alexandre Luciano
Universidade de Caxias do Sul

Prof. Dr. Sandro Rogério dos Santos
Universidade de Caxias do Sul

À minha mãe, minha melhor
amiga e maior fonte de amor,
que tornou e continua tornando
tudo possível.

AGRADECIMENTOS

Agradeço primeiramente a minha mãe, Eva Elisabete Boeira, por todo amor, ajuda e apoio incondicional durante toda a minha vida. A minha avó, Suely Padilha Boeira, por todos os cuidados e carinhos que só uma avó sabe dar. Sem a educação e o exemplo de vida delas, com certeza não seria a pessoa que sou hoje. O meu eterno agradecimento e admiração, amo vocês.

Ao meu orientador, professor Leandro Corso, por todas explicações, ajuda no texto e no entendimento geral do trabalho, e também por me propor novos assuntos desafiadores, por acreditar na minha capacidade. Além de me inspirar profissionalmente, um exemplo de engenheiro e professor a ser seguido.

Ao meu coorientador, professor Alexandre Fassini Michels, por todas explicações dos conceitos de estatística e das contribuições para o meu trabalho. Além das risadas, tornando as reuniões sempre divertidas.

A todos os professores que contribuíram na minha formação acadêmica e profissional, tanto no curso de Engenharia de Produção quanto no Mestrado em Engenharia Mecânica.

Gostaria de agradecer também a todos meus amigos, por toda ajuda nas aulas e também pelos vários momentos de descontração.

Enfim, a todos que de alguma forma contribuíram na minha formação, o meu muito obrigado.

“Investir em conhecimento rende sempre os melhores juros.”

Benjamin Franklin

RESUMO

O mercado de ações atrai cada vez mais investidores, por sua alta rentabilidade, quando comparado com outras opções de investimentos. Porém, esta comercialização de ativos possui um alto risco, sendo aconselhável utilizar conceitos matemáticos para auxiliar na tomada de decisões. Este trabalho visou estudar alguns desses conceitos aplicados na bolsa de valores, para maximizar o retorno e controlar os riscos. Primeiramente, para a seleção de carteira de ativo foi utilizada a Análise Envolvória de Dados (DEA), para em seguida, prever o preço das ações das empresas selecionadas. Para esta prognose, foi realizada uma comparação de modelos de previsão de séries temporais, com o uso das Redes Neurais Artificiais e da metodologia Box-Jenkins. Para o cálculo da DEA, foram utilizados indicadores como preço/lucro, retorno, entre outros, como variáveis de entrada e saída. Já para a previsão, foi empregado o preço diário de dia comercial para cada uma das empresas previamente selecionadas para a carteira otimizada. Para o estudo com DEA, obteve-se um portfólio otimizado com retorno 15 vezes maior do que o índice Ibovespa; e para o comparativo dos métodos de previsão, concluiu-se que a metodologia Box-Jenkins foi a mais indicada para prever preço de ações, por ter menores médias de erros.

Palavras-chave: Pesquisa Operacional. Otimização. Análise Envolvória de Dados. Redes Neurais Artificiais. Modelo de Box-Jenkins. Mercado de ações.

ABSTRACT

The stock market attracts more and more investors because of its high profitability, when compared to other investment options. However, this commercialization of assets has a high risk, and it is advisable to use mathematical concepts to assist in decision making. This paper aims to study some of these concepts applied in the stock exchange, to maximize the return and to control the risks. Firstly, for asset portfolio selection was used Data Envelopment Analysis (DEA), to then predict the stock price of the selected companies. For this prognosis, a comparison was made of time series forecast models, using the Artificial Neural Networks and Box-Jenkins method. For the DEA's calculus, it was used indicators such as price/profit, return, among others, as input and output variables. For the forecast, the daily commercial day price was used for each of the companies previously selected for the optimized portfolio. For the study with DEA, an optimized portfolio was obtained with a 15 times higher return than the Ibovespa index; and for the comparison of the forecasting methods, it was concluded that the Box-Jenkins methodology was the most adequate to predict share prices, due to lower error rates.

Keywords: Operations Research. Optimization. Data Envelopment Analysis. Artificial Neural Networks. Box-Jenkins method. Stock market.

LISTA DE FIGURAS

Figura 1 – Comparativo das orientações.....	25
Figura 2 – Fronteira de eficiência	26
Figura 3 – Neurônio artificial.....	35
Figura 4 – Redes alimentadas diretamente com múltiplas camadas	37
Figura 5 – Parte de um MLP com aprendizagem por backpropagation	38
Figura 6 – Processo estocástico estacionário e não estacionário	41
Figura 7 – Correlogramas	43
Figura 8 – Correlograma de AR(1)	46
Figura 9 – Correlograma de AR(2)	46
Figura 10 – Correlograma de MA(1).....	47
Figura 11 – Correlograma de MA(2).....	48
Figura 12 – Correlograma de ARMA(1,1).....	49
Figura 13 – Série não estacionária quanto ao nível	50
Figura 14 – Série não estacionária quanto ao nível e à inclinação	50
Figura 15 – Redução da não estacionariedade de uma série temporal após sucessivas diferenciações	51
Figura 16 – Correlograma de um processo não estacionário sazonal	53
Figura 17 – Correlograma de um processo estacionário sazonal	53
Figura 18 – Etapas de modelagem de Box-Jenkins	55
Figura 19 – Construção repetida de um intervalo de confiança de μ	59
Figura 20 – Gráficos de resíduos	60
Figura 21 – Etapas propostas	61
Figura 22 – Representação do modelo DEA.....	64
Figura 23 – Etapas da avaliação dos resultados.....	69
Figura 24 – Possíveis previsões para cada empresa.....	80
Figura 25 – Classificação dos dados para RNA.....	81
Figura 26 – Comparativo dos dados reais com os calculados com RNA	85
Figura 27 – Análise dos resíduos	85
Figura 28 – Histograma dos erros	86
Figura 29 – Correlogramas.....	87
Figura 30 – Correlograma da ACF dos resíduos.....	90
Figura 31 – Correlograma da PACF dos resíduos	91

Figura 32 – Variação percentual do índice Ibovespa	92
Figura 33 – Gráfico comparativo de investimentos	92
Figura 34 – Histograma dos erros de previsão.....	94
Figura 35 – Média dos erros de cada empresa	94
Figura 36 – Resultados das previsões para a empresa Ambev	96
Figura 37 – Erros das previsões do restante da semana da empresa Ambev	96
Figura 38 – Histograma dos erros de previsão do restante da semana	97
Figura 39 – Média de erros por empresa para previsão do restante da semana	98
Figura 40 – Histograma de todos os erros	99
Figura 41 – Média de todos os erros para cada empresa	99

LISTA DE TABELAS

Tabela 1 – Equações do Modelo CCR	33
Tabela 2 – Equações do Modelo BCC	33
Tabela 3 – Equações do método Box-Jenkins	56
Tabela 4 – Lista das empresas escolhidas	72
Tabela 5 – Dados de input das empresas	73
Tabela 6 – Dados de output das empresas	75
Tabela 7 – Dados de eficiência	77
Tabela 8 – Previsão para 12, 24, 36, 48 e 60 meses	82
Tabela 9 – Previsão para 72, 84, 96, 108 e 120 meses	82
Tabela 10 – Erro quadrático médio de cada prognose	83
Tabela 11 – Coeficiente de determinação de cada prognose	84
Tabela 12 – Resultados do cálculo de previsão	88
Tabela 13 – Erro quadrático médio de cada prognose	89
Tabela 14 – Coeficiente de determinação de cada prognose	89
Tabela 15 – Erro de previsão da empresa Ambev	93
Tabela 16 – Classificação das empresas	100
Tabela 17 – Comparativo dos métodos de previsão	100

LISTA DE ABREVIATURAS E SIGLAS

ACF	<i>Autocorrelation Function</i> ou Função de Autocorrelação
AR	Auto regressão
ARIMA	<i>Autoregressive Integrated Moving Average</i> ou Modelos Auto Regressivos Integrados a Média Móvel
BCC	Modelo de Banker, Charnes e Cooper
B-J	Box-Jenkins
BM&FBovespa	Bolsa de Valores, Mercadorias e Futuros de São Paulo
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CCR	Modelo de Charnes, Cooper e Rhodes
Cov	Covariância
CRS	<i>Constant returns to scale</i> ou retornos constantes à escala
DEA	<i>Data Envelopment Analysis</i> ou Análise Envoltória de Dados
DMU	<i>Decision Making Unit</i> ou Unidades Tomadoras de Decisão
FF	<i>Feedforward</i>
HH	Hospital hipotético
I	Integração
LPA	Lucro por ação
MA	Média móvel
MLP	<i>Multilayer perceptron</i>
MSE	<i>Mean Squared Error</i> ou Erro Quadrático Médio
PACF	<i>Partial Autocorrelation Function</i> ou Função de Autocorrelação Parcial
PPGMEC	Programa de Pós-Graduação Mestrado de Engenharia Mecânica
P/L	Preço/lucro
R	Retorno sobre investimento
RNA	Redes Neurais Artificiais
SLP	<i>Single layer perceptron</i>
SQReg	Soma dos quadrados da regressão
STQ	Soma total dos quadrados
UCS	Universidade de Caxias do Sul

V	Volatilidade anualizada
Var	Variância
VH	Volatilidade histórica
VRS	<i>Variable returns to scale</i> ou retornos variáveis à escala

LISTA DE SÍMBOLOS

N	Conjunto de valores que variam de 1 a n
Eff	Eficiência
S	Conjunto de saídas
J	Conjunto de valores que variam de 1 a j
E	Conjunto de entradas
K	Conjunto de valores que variam de 1 a k
u	Peso ou multiplicador de cada valor de saída
v	Peso ou multiplicador de cada valor de entrada
η	Eficiência da Programação Linear, para orientação por <i>input</i>
ε	Variável que impede distorção da solução ótima
h	Eficiência da Programação Linear, para orientação por <i>output</i>
λ	Preço sombra
f	Função de ativação
w	Ponderações
x	Entradas ou insumos
y	Produto
Ω	Espaço amostral
$\mathfrak{S}$	σ -álgebra de subconjuntos de Ω
$P(\bullet)$	Probabilidade de um evento
Z	Série diferenciada para se tornar estacionária
Y	Observação da série temporal
∇	Operador de diferença de retardo
t	Tempo
ρ	Coeficiente de autocorrelação populacional
μ	Média
k	Defasagem ou <i>lag</i>
σ	Desvio padrão
r	Coeficiente de autocorrelação amostral
ϕ	Coeficiente de autocorrelação parcial
p	Ordem máxima de parâmetros de auto regressão simples

d	Número de diferenciações não sazonais
q	Ordem máxima de média móvel simples
P	Ordem máxima dos parâmetros de auto regressão sazonal
D	Número de diferenciações sazonais
Q	Número máximo dos parâmetros de média móvel sazonal
s	Número de períodos por ciclo sazonal
e	Erro aleatório ou resíduo
c	Termo constante
γ	Coeficiente de auto regressão
B	Operador de substituição
θ	Coeficiente de média móvel
Γ	Coeficiente de auto regressão sazonal
θ	Coeficiente de média móvel sazonal
$\hat{Y}$	Valor futuro estimado
n	Quantidade de dados
r^2	Coeficiente de determinação
$\bar{Y}$	Média aritmética para o valor de Y
$\bar{X}$	Limites superior e inferior do intervalo de confiança
z	Probabilidade
d_t	Resíduo padronizado
P	Preço da ação em um determinado período

SUMÁRIO

1	INTRODUÇÃO	19
1.1	JUSTIFICATIVA	20
1.2	OBJETIVOS	21
1.2.1	Objetivo geral.....	21
1.2.2	Objetivos específicos.....	21
1.3	APRESENTAÇÃO DA INSTITUIÇÃO	21
1.4	ABORDAGEM E DELIMITAÇÃO.....	22
2	FUNDAMENTAÇÃO TEÓRICA	24
2.1	ANÁLISE ENVOLTÓRIA DE DADOS.....	24
2.1.1	Modelo CCR	28
2.1.2	Modelo BCC	30
2.1.3	Recapitulação da literatura.....	32
2.1.3.1	Agrupamento das equações.....	32
2.1.3.2	Revisão bibliográfica – DEA.....	34
2.2	REDES NEURAIS ARTIFICIAIS.....	34
2.2.1	Topologia de rede.....	34
2.2.2	Processo de aprendizagem	37
2.2.3	Revisão bibliográfica – RNA.....	39
2.3	MODELOS BOX-JENKINS.....	39
2.3.1	Conceitos básicos para compreensão do método Box-Jenkins.....	39
2.3.1.1	Previsão de séries temporais	39
2.3.1.2	Processo estocástico.....	40
2.3.1.3	Função de autocorrelação.....	42
2.3.2	Metodologia Box-Jenkins e seus modelos matemáticos	44
2.3.2.1	Modelos estacionários.....	45
2.3.2.1.1	<i>Modelo Auto Regressivo (AR)</i>	<i>45</i>
2.3.2.1.2	<i>Modelo de Médias Móveis (MA)</i>	<i>47</i>
2.3.2.1.3	<i>Modelo Auto Regressivo de Médias Móveis (ARMA)</i>	<i>48</i>
2.3.2.2	Modelos não estacionários.....	49
2.3.2.2.1	<i>Modelo Auto Regressivo Integrado de Médias Móveis (ARIMA)</i>	<i>49</i>
2.3.2.3	Modelos sazonais.....	52

2.3.3	Etapas de modelagem.....	53
2.3.4	Recapitulação da literatura.....	55
2.3.4.1	Agrupamento das equações.....	56
2.3.4.2	Revisão bibliográfica – Box-Jenkins	57
2.4	ANÁLISE ESTATÍSTICA	57
2.4.1	Erros de previsão	57
2.4.2	Coeficiente de determinação.....	57
2.4.3	Intervalo de confiança.....	58
2.4.4	Análise de resíduos.....	59
3	MÉTODOS DA DISSERTAÇÃO	61
3.1	ESCOLHA DAS EMPRESAS E COLETA DE SEUS DADOS	62
3.2	CÁLCULO DA EFICIÊNCIA COM A UTILIZAÇÃO DA DEA	64
3.3	DETERMINAÇÃO DOS DADOS PARA A ANÁLISE DE PREVISÃO	65
3.4	UTILIZAÇÃO DAS REDES NEURAIS ARTIFICIAIS	66
3.5	APLICAÇÃO DA METODOLOGIA DE BOX-JENKINS.....	68
3.6	AVALIAÇÃO DOS RESULTADOS	69
3.6.1	Avaliação da DEA	70
3.6.2	Avaliação dos métodos de previsão.....	70
4	APLICAÇÃO DOS MÉTODOS	72
4.1	ESCOLHA DAS EMPRESAS E COLETA DE SEUS DADOS	72
4.2	CÁLCULO DA EFICIÊNCIA COM A UTILIZAÇÃO DA DEA	76
4.3	DETERMINAÇÃO DOS DADOS PARA A ANÁLISE DE PREVISÃO	78
4.4	UTILIZAÇÃO DA RNA.....	80
4.4.1	Previsão com RNA	80
4.4.2	Análise estatística	83
4.4.3	Análise dos resíduos e erro	85
4.5	APLICAÇÃO DA METODOLOGIA DE BOX-JENKINS.....	87
4.5.1	Previsão com Box-Jenkins.....	87
4.5.2	Análise estatística	88
4.5.3	Análise dos resíduos	90
4.6	AVALIAÇÃO DOS RESULTADOS	91
4.6.1	Resultado da DEA	91

4.6.2	Resultados dos métodos de previsão	93
4.6.2.1	Considerações sobre os métodos de previsão	98
5	CONCLUSÃO	103
5.1	CONSIDERAÇÕES FINAIS	103
5.2	SUGESTÕES DE TRABALHOS FUTUROS	105
	REFERÊNCIA BIBLIOGRÁFICA.....	106
	APÊNDICE A – SÍNTESE DOS TRABALHOS ESTUDADOS REFERENTES A ANÁLISE ENVOLTÓRIA DE DADOS	114
	APÊNDICE B – SÍNTESE DOS TRABALHOS ESTUDADOS REFERENTES A RNA: PARÂMETROS ESCOLHIDOS	121
	APÊNDICE C – SÍNTESE DOS TRABALHOS ESTUDADOS REFERENTES A RNA: DADOS UTILIZADOS	126
	APÊNDICE D – SÍNTESE DOS TRABALHOS ESTUDADOS REFERENTES A METODOLOGIA BOX-JENKINS.....	131
	APÊNDICE E – PREÇOS DIÁRIO DAS EMPRESAS QUE FORAM CONSIDERADAS EFICIENTES	138
	APÊNDICE F – RESULTADOS OBTIDOS UTILIZANDO BOX-JENKINS.....	142
	APÊNDICE G – ERRO DOS RESULTADOS	147
	APÊNDICE H – PREVISÕES PARA O RESTANTE DA SEMANA.....	155
	APÊNDICE I – INFORMAÇÕES DAS RNA PARA PREVISÕES PARA O RESTANTE DA SEMANA.....	159
	APÊNDICE J – INFORMAÇÕES DO MÉTODO DE BOX-JENKINS PARA PREVISÕES PARA O RESTANTE DA SEMANA.....	160
	APÊNDICE L – ERRO DOS RESULTADOS DAS PREVISÕES PARA O RESTANTE DA SEMANA.....	161

1 INTRODUÇÃO

Quem deseja investir em bolsa de valores tem como meta a maximização de seus lucros e a minimização de possíveis perdas. Para isso, devem-se considerar os riscos envolvidos em compra de ações, e uma maneira de controlá-los é possuir uma carteira de ativos, isto é, não comprar ações de somente uma empresa. Quanto mais eficiente a carteira, melhor o seu desempenho, em outras palavras, maior o retorno sobre o capital investido. Lopes et al. (2008) comentam sobre a complexidade que é selecionar uma carteira de ativos e que, cada vez mais, técnicas para escolher quais ações deverão compor a carteira estão chamando a atenção dos investidores, devido a seus cálculos computacionalmente rápidos e precisos.

Um método matemático para determinar a carteira de ativos mais eficiente é a Análise Envoltória de Dados (em inglês, *Data Envelopment Analysis* – DEA), a qual foi estudada e utilizada neste trabalho. Esta metodologia, desenvolvida por Charnes, Cooper e Rhodes, em 1978, mensura a eficiência relativa de unidade tomadora de decisão (em inglês, *decision making units* – DMU). A DEA é um método da Pesquisa Operacional e usa técnicas de Programação Linear para encontrar a solução do problema. Neste trabalho, utilizou-se o *software* MatLab para a realização dos cálculos.

Outra maneira de controlar os riscos é prever o valor de uma ação ou como ela variará no dia seguinte. Gaio, Castro Junior e Oliveira (2007) comentam que modelos de previsão de preço futuro podem auxiliar até na decisão de novas estratégias de compra e venda de ações, por contribuir com uma maior precisão dos investimentos e gerenciamento de riscos. Por haver diversas variáveis envolvidas, Freitas (2001) sugere a utilização de modelos de previsão, pela dificuldade em analisar satisfatoriamente a relação entre elas, e obter uma prognose mais adequada. Dentre as várias metodologias de previsão de séries temporais, este trabalho estudou Redes Neurais Artificiais (RNA) e os modelos Box-Jenkins (B-J).

As Redes Neurais Artificiais tiveram seu início na década de 40 com a publicação de um artigo de McCullen e Pitts, e progrediu juntamente com o avanço do computador no decorrer dos anos. Raia Junior (2000) e Souza (2008) definem como sendo um algoritmo computacional inspirado no cérebro humano, por se basear na mesma maneira de funcionamento, ou seja, aprendendo a partir de dados fornecidos. Já a metodologia Box-Jenkins obteve seus primeiros estudos na década de 70,

analisando o comportamento de variáveis e, conseqüentemente, calcula previsões de uma série temporal, utilizando conceitos de estatística. Fischer (1982) conceitua que este método surgiu com a necessidade de simular e quantificar as complexas relações existentes nas atividades econômicas.

A proposta de pesquisa deste trabalho foi aplicar DEA para criar uma carteira otimizada de ações de empresas presentes na Bolsa de Valores, Mercadorias e Futuros de São Paulo (BM&FBovespa). Em seguida, foram utilizadas as Redes Neurais Artificiais e a metodologia Box-Jenkins para prever o valor da ação de cada uma destas empresas.

1.1 JUSTIFICATIVA

Quando comparado com outras ciências, a Pesquisa Operacional é um assunto relativamente novo, já que se iniciou há aproximadamente 70 anos. Muitos estudos já foram feitos, porém existem oportunidades ainda inexploradas.

Ao operar com o mercado de ações, é sugerido que se tenha uma carteira composta de diferentes empresas, ou seja, é aconselhável dividir o dinheiro investido em diversos papéis da BM&FBovespa, e um exemplo de método matemático para auxiliar nesta decisão é a Análise Envolvória de Dados. Assim, foi decidido a sua utilização para poder otimizar a carteira, a fim de ter uma melhor probabilidade de maior rentabilidade nas operações com ações. Em outras palavras, este trabalho buscou o melhor retorno do investimento possível, e isso não seria possível se fossem selecionadas empresas aleatoriamente.

Após realizado esta seleção de empresas, desejou-se prever o preço da ação para a data seguinte, e foi escolhido calcular com Redes Neurais Artificiais e metodologia Box-Jenkins. Devido à complexidade do mercado de ações e todas suas variáveis envolvidas, podendo ser de natureza não-linear e de difícil mensuração, é aconselhável utilizar modelos de previsão de séries temporais, já que compreendem o padrão de comportamento dos dados fornecidos para fins preditivos. Além disso, foi analisado qual a quantidade de dados de entrada mais indicada para cada modelo de previsão, sendo um diferencial da maioria dos trabalhos com esta aplicabilidade.

Em relação aos conceitos da fundamentação teórica, optou-se por aplicar os métodos estudados na bolsa de valores, para estudar conteúdos explorados no Mestrado em Engenharia Mecânica em uma área que não há modelo padrão eficaz

para auxiliar na tomada de decisão. Ainda, almejou-se um estudo que fosse possível utilizar na prática e ter resultados reais.

1.2 OBJETIVOS

A seguir, foram detalhados o objetivo geral e os objetivos específicos.

1.2.1 Objetivo geral

O objetivo deste trabalho foi, a partir de um portfólio otimizado com DEA, determinar qual método de previsão de série temporal é o mais indicado, entre RNA e metodologia Box-Jenkins, para realizar a prognose do preço de ações de empresas presentes na BM&FBovespa.

1.2.2 Objetivos específicos

Este trabalho teve os seguintes objetivos específicos:

- a) escolher as empresas e coletar seus respectivos dados;
- b) otimizar a carteira de ações, com a utilização da DEA;
- c) determinar dados a serem utilizados na análise de previsão, comparando Rede Neural Artificial com metodologia de Box-Jenkins;
- d) utilizar Redes Neurais Artificiais;
- e) aplicar a metodologia de Box-Jenkins; e
- f) avaliar os resultados obtidos.

1.3 APRESENTAÇÃO DA INSTITUIÇÃO

Este trabalho foi realizado juntamente com a Universidade de Caxias do Sul (UCS), que é uma instituição de ensino superior, de caráter comunitário e regional. Sua criação foi em 1967, sendo a Instituição de Ensino Superior mais antiga da região. De acordo com seu endereço eletrônico, a UCS atua no campo do ensino, pesquisa, inovação e extensão, e seu Campus Sede está em Caxias do Sul, um polo de desenvolvimento industrial com reconhecimento nacional. Além de Caxias do Sul, possui unidades universitárias em outras oito cidades, sendo elas: Bento Gonçalves,

Vacaria, Canela, Farroupilha, Guaporé, Nova Prata, Veranópolis e São Sebastião do Caí.

Dentro da universidade, ocorrem diversos tipos de pesquisa, sendo que a Pesquisa Operacional está inserida no curso de Graduação em Engenharia de Produção e também no Programa de Pós-Graduação *stricto sensu* Mestrado de Engenharia Mecânica (PPGMEC). A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) autorizou o funcionamento do PPGMEC no dia 1º de setembro de 2011. Segundo o próprio endereço eletrônico da universidade, o objetivo principal do Mestrado é capacitar os acadêmicos para a prática profissional avançada e transformadora de processos, integrada ao fundamento científico, assim apresentando inovações e aperfeiçoamento tecnológico. Para os acadêmicos que concluírem este Mestrado, é proporcionado o grau de Mestre em Engenharia.

1.4 ABORDAGEM E DELIMITAÇÃO

Como este trabalho foi sobre um tema da Pesquisa Operacional, que envolve modelos matemáticos, sua abordagem se torna quantitativa exploratória. Por este motivo, ao final deste estudo, o resultado foi expresso como um valor numérico a ser discutido. Conforme classificações de Gil (2010), é uma pesquisa exploratória, pois proporciona maior familiaridade com o problema em questão, a fim de torná-lo mais compreensível. Ainda, com base nos procedimentos técnicos, este trabalho pode ser classificado como uma pesquisa bibliográfica *ex post facto*. Ainda conforme o autor, entende-se por pesquisa bibliográfica aquela que é desenvolvida com base em material já elaborado, ou seja, com o auxílio de livros, dissertações e artigos científicos. Já uma pesquisa *ex post facto*, que foi traduzido do latim como sendo a partir do fato passado, significa que o estudo foi realizado após a ocorrência de variações nos valores dos dados utilizados.

Este trabalho apresentou uma opção de resolução para o problema, porém não pode ser considerado como verdade absoluta, já que não se considera todo o sistema real do mercado de ações e outros fatores podem ser relevantes.

Este estudo não abrange uma coleta de dados em qualquer local da universidade, nem na implantação do resultado. As informações das empresas são de conhecimento público, podendo ser consultadas nos endereços eletrônicos da BM&FBovespa e Exame, na seção Mercados. Considerando que os métodos que

previsão empregados analisam oscilações e comportamento de dados e o foco deste trabalho ser a aplicação destes conceitos e análise dos resultados, não foi estudado o quanto o mercado impactou nos dados coletados, nem justificado os comportamentos dos valores devidos a eventos históricos.

Ainda, os modelos matemáticos já são existentes, assim como a implementação da DEA e da RNA no *software* MatLab e o uso da metodologia Box-Jenkins no *software* SPSS. Para os parâmetros necessários para aplicação destes conceitos, foram escolhidos conforme outros trabalhos de autores analisados na fundamentação teórica, pois o objetivo deste estudo foi analisar os resultados obtidos nos cálculos e não se aprofundar na determinação de parâmetros.

1.5 ESTRUTURA DA DISSERTAÇÃO

Este trabalho foi iniciado no ano de 2016, continuou pelo ano de 2017 e foi finalizado em 2018, junto com a Universidade de Caxias do Sul, na área da Pesquisa Operacional, inserida na linha de pesquisa Processos de Fabricação e Produção de Componentes, Ferramentas e Matrizes, do Mestrado de Engenharia Mecânica. O escopo deste trabalho foi estruturado em cinco capítulos, sendo este o primeiro como a Introdução. O segundo capítulo consiste na fundamentação teórica da Análise Envolvente de Dados, das Redes Neurais e da metodologia Box-Jenkins, além do estudo da análise estatística. Já o terceiro apresentou as propostas da dissertação, e o quarto capítulo exhibe as aplicações destas propostas. Para o quinto e último capítulo, foi descrito as conclusões obtidas com este estudo.

2 FUNDAMENTAÇÃO TEÓRICA

Para o capítulo que segue, foram apresentados os conceitos de Análise Envoltória de Dados (DEA), de Redes Neurais Artificiais (RNA) e dos modelos de Box-Jenkins, além do estudo de conteúdos estatísticos, de acordo com as referências bibliográficas.

2.1 ANÁLISE ENVOLTÓRIA DE DADOS

A Análise Envoltória de Dados (DEA) foi descrita por Mello et al. (2005), Mardani et al. (2017) e Guo et al. (2017) como sendo um método matemático que mede a eficiência de unidades produtivas. Esta técnica de medição de desempenho faz parte da Pesquisa Operacional e foi desenvolvida por Charnes, Cooper e Rhodes, em 1978. Lopes et al. (2008) comentam que, como mensura o desempenho de determinada unidade quando comparada com as demais, é uma medida de eficiência relativa. Consequentemente, o resultado mudará se for acrescentado ou retirado qualquer unidade da análise.

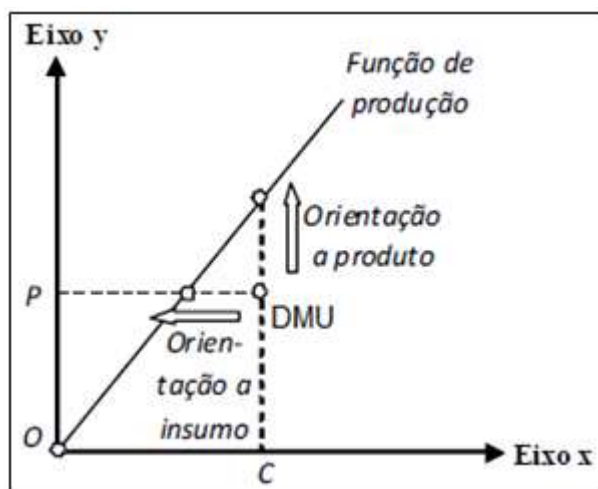
No artigo, Charnes, Cooper e Rhodes (1978) trabalharam com as unidades tomadoras de decisão (DMU), salientando que todas estas medidas devem ter os mesmos insumos e produtos (*inputs* e *outputs*, respectivamente), e que estas informações sejam de valor positivo, devido às restrições impostas nos modelos matemáticos. Ainda, Zhou, Poh e Ang (2007), Rosano-Peña (2012) e Rotela Junior, Pamplona e Salomon (2014) comentam que, a DEA avalia a eficiência relativa de unidades que utilizam múltiplos insumos que geram múltiplos produtos.

Rosano-Peña (2012) entende eficiência como sendo o resultado da otimização, em que se tem a utilização da menor quantidade possível de insumos para um determinado valor de produto, ou como quando há o maior nível de produção para uma determinada quantidade de insumo. Para o autor, uma técnica para quantificar este dado é a atribuição de hipóteses flexíveis sobre o comportamento das variáveis e não necessita especificar a relação funcional entre *inputs* e *outputs*. Desta forma, é possível estabelecer a fronteira de eficiência com a utilização de combinação linear e calcular a eficiência a partir da distância entre a DMU com esta fronteira. Além disso, possibilita o estudo de situações de múltiplos *inputs* e *outputs*.

A DEA pode ser aplicada além dos exemplos clássicos de comparações entre empresas ou unidades internas dentro de uma organização, mas também para analisar problemas globais. Rafaeli (2009) exemplifica isso em seu trabalho ao sugerir uma nova forma de avaliar o desempenho dos países do globo, considerando seu grau de desenvolvimento, e também por analisar as companhias de transporte aéreo de passageiros. Já Martins (2014) utiliza DEA para analisar a eficiência de empresas brasileiras distribuidoras de energia elétrica, já que uma economia neste setor possibilitaria ganhos sociais e ambientais. Lim, Oh e Zhu (2014) reforçam que a DEA é uma medida de desempenho multidisciplinar, por ser possível sua utilização em seleção de portfólio de ações, de projetos, de fornecedores, classificação de inventário ABC, entre diversas outras aplicações.

A DEA pode ser analisada em duas orientações: de *input* ou de *output*. Mello et al. (2005), Rosano-Peña (2012), Gökşen, Doğan e Özkarabacak (2015) e Martins (2014) comentam que é orientação por *input* quando o tomador de decisão tiver controle sobre os dados de entrada, por exemplo, produzir uma quantidade fixada com o mínimo de insumos. Caso contrário, deve ser usada a orientação por *output*, por exemplo, quando se deseja maximizar a produção para uma determinada quantidade de recurso. A Figura 1 demonstra a representação gráfica de ambas orientações. O eixo y representa o produto e o eixo x, o insumo. Além disso, as setas indicam o deslocamento para atingir a fronteira de eficiência, e consequentemente, a orientação.

Figura 1 – Comparativo das orientações

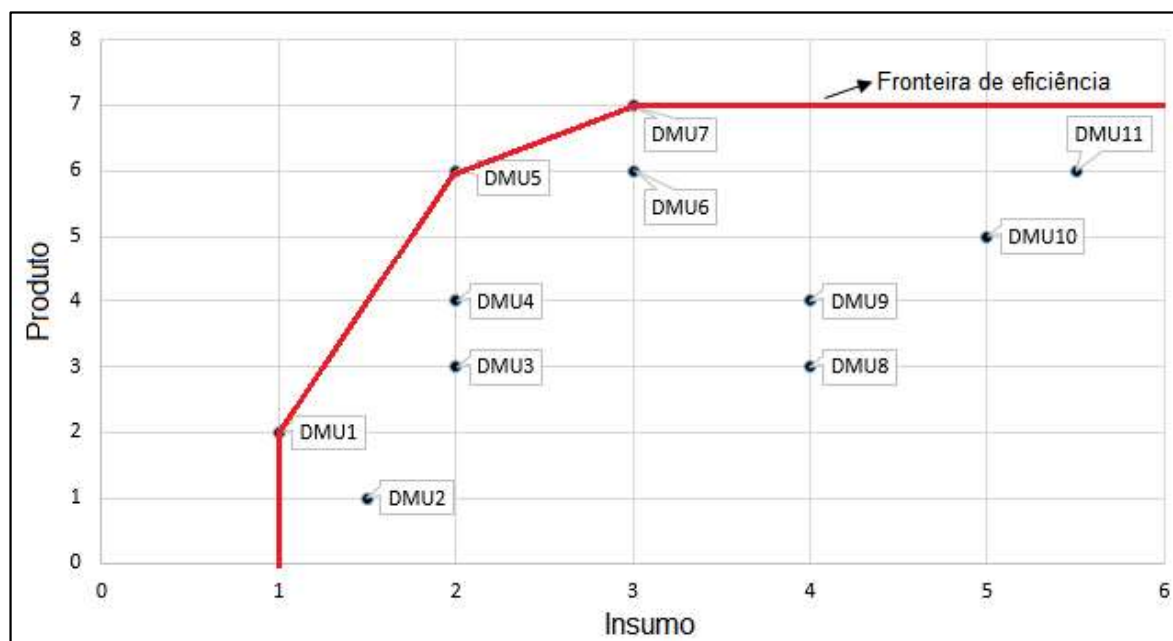


Fonte: Adaptado de Lopes et al. (2008)

Analisando a figura, entende-se que, a orientação a insumo (*input*) significa aumentar a eficiência relativa da DMU em questão, por diminuição de insumos, para a mesma quantidade de produto. Lopes et al. (2008) ainda conceituam que, a orientação a produto (*output*) reflete no aumento da eficiência relativa da DMU, quando se tem um aumento de produção, com uma quantidade de insumos constantes.

Ainda, Lopes et al. (2008), Lim, Oh e Zhu (2014) e Jebali, Essid e Khraief (2017) definem que a DMU que estiver na fronteira de eficiência é considerada eficiente, enquanto as ineficientes são as que estão deslocadas desta fronteira. A Figura 2 apresenta esta fronteira de eficiência.

Figura 2 – Fronteira de eficiência



Fonte: Elaborado pela autora (2016)

Pode-se analisar que, as DMUs com 100% de eficiência relativa são as DMU1, DMU5 e DMU7, pois elas estão na fronteira de eficiência, consequentemente, as demais são consideradas ineficientes. Ainda, pode-se observar que a DMU4 é mais eficiente do que a DMU3, pois produz mais produtos com a mesma quantidade de insumos. Quando se compara as DMUs 4 e 9, entende-se que elas produzem a mesma quantidade de produto, porém a DMU4 utiliza menos insumos, por isso ela é mais eficiente que a DMU9. Sendo assim, conforme citam Martins (2014), Mardani et al. (2017) e Guo et al. (2017), conclui-se que as DMUs mais eficientes são aquelas

que conseguem um maior número possível de produtos combinando com o menor número possível de insumos.

Colin (2007) conceitua que cada DMU seja representada por um índice n , em que $n = 1, 2, \dots, N$, e que a sua eficiência seja Eff_n , para cada unidade n , conforme Equação 1.

$$Eff_n = \frac{\text{valor de mercado das saídas geradas pela unidade } n}{\text{valor de mercado das entradas consumidas pela unidade } n} \quad (1)$$

Ainda, o autor considera um conjunto de saídas S_{nj} para cada unidade n , onde j representa a quantidade de saídas ($j = 1, 2, \dots, J$), e E_{nk} para o conjunto de entradas para cada unidade n , sendo k a quantidade de entradas ($k = 1, 2, \dots, K$). Deste modo, pode-se concluir que, para uma DMU mais eficiente, necessita-se de menos entradas consumidas ou mais saídas geradas.

Charnes, Cooper e Rhodes (1978) e Jebali, Essid e Khraief (2017) argumentam que a DEA possui um critério de ponderação para cada unidade n . Então, cada saída j e cada entrada k possuem um peso, u_j e v_k , respectivamente. Mariano, Almeida e Rebelatto (2006a) entendem estas variáveis como um conceito de economia, representando a importância do determinado insumo ou produto. A variável peso também pode ser denominada multiplicadores, conforme explica Martins (2014). Assim, a eficiência é reescrita de acordo com a Equação 2.

$$Eff_n = \frac{u_1 S_{n1} + u_2 S_{n2} + \dots + u_J S_{nJ}}{v_1 E_{n1} + v_2 E_{n2} + \dots + v_K E_{nK}} = \frac{\sum_{j=1}^J u_j S_{nj}}{\sum_{k=1}^K v_k E_{nk}} \quad (2)$$

Como S_{nj} e E_{nk} são constantes e observações do passado, então u_j e v_k são as variáveis de decisão do problema e devem ser maiores que zero. Para Charnes, Cooper e Rhodes (1978), a eficiência de cada DMU é o valor máximo da razão das saídas ponderadas com as entradas ponderadas, conforme Equações 3 e 4.

$$\text{Max } Eff_0 = \text{Max } \frac{\sum_{j=1}^J u_j S_{0j}}{\sum_{k=1}^K v_k E_{0k}} \quad (3)$$

$$\text{Sujeito às restrições: } \begin{cases} \frac{\sum_{j=1}^J u_j S_{nj}}{\sum_{k=1}^K v_k E_{nk}} \leq 1 \\ u_j, v_k \geq 0 \end{cases} \quad (4)$$

Em que, $n = 1, \dots, N$, $j = 1, \dots, J$ e $k = 1, \dots, K$. Além disso, ressalta-se que Eff_o é a eficiência relativa de DMU_o , que é a DMU sob avaliação. Zhou, Poh e Ang (2007), Batista (2009), Martins (2014) e Rotela Junior, Pamplona e Salomon (2014) comentam que, quando Eff_o for igual a 1, significa que a DMU_o é eficiente, comparada às demais unidades; e quando for menor do que 1, é ineficiente.

A Análise Envoltória de Dados possui dois modelos: CCR e BCC, que são estudados na sequência.

2.1.1 Modelo CCR

O modelo CCR foi desenvolvido por Charnes, Cooper e Rhodes, em 1978, e possui esta sigla por causa dos nomes de seus autores. De acordo com Lopes et al. (2008), ele também é chamado de CRS, abreviatura do inglês *constant returns to scale*, ou seja, retornos constantes à escala. Mello et al. (2005) e Rosano-Peña (2012) explicam que, neste modelo, qualquer variação nas entradas (*inputs*) produz variação proporcional nas saídas (*outputs*). Mariano, Almeida e Rebelatto (2006b) e Galvão (2008) comentam que, para este modelo, a fronteira de eficiência será sempre uma reta com inclinação de 45°, conforme representado anteriormente pela Figura 1.

Colin (2007), Rosano-Peña (2012) e Mardani et al. (2017) ressaltam que, a função objetivo da Equação 3 está associada a uma classe de problemas da Programação Não-Linear, denominada programação fracionária. Então, foi considerado arbitrariamente o número 1 para o denominador da função objetivo, gerando a restrição $\sum_{k=1}^K v_k E_{0k} = 1$. Assim, as Equações 5 e 6 representam a DEA-CCR por orientação por *input*, que utiliza Programação Linear. Esta programação é a primal e também pode ser denominada de Modelo dos Multiplicadores.

$$\text{Max } \eta_o = \text{Max } \sum_{j=1}^J u_j S_{0j} \quad (5)$$

$$\text{Sujeito às restrições: } \begin{cases} \sum_{j=1}^J u_j S_{nj} - \sum_{k=1}^K v_k E_{nk} \leq 0 \\ \sum_{k=1}^K v_k E_{0k} = 1 \\ u_j \geq \varepsilon > 0 \\ v_k \geq \varepsilon > 0 \end{cases} \quad (6)$$

Em que, η é a eficiência para orientação por *input*, $n = 1, \dots, N$, $j = 1, \dots, J$ e $k = 1, \dots, K$. Colin (2007) e Rosano-Peña (2012) comentam que, para não haver uma distorção da solução ótima, as variáveis não podem ser iguais a zero, por isso o ε nas restrições. Por isso, Colin (2007) o definiu como $\varepsilon = 0,001$.

As Equações 7 e 8 representam a DEA-CCR por orientação por *output*, sendo uma programação também primal e também denominada Modelo dos Multiplicadores.

$$\text{Min } h_0 = \text{Min } \sum_{k=1}^K v_k E_{0k} \quad (7)$$

$$\text{Sujeito às restrições: } \begin{cases} \sum_{k=1}^K v_k E_{nk} - \sum_{j=1}^J u_j S_{nj} \leq 0 \\ \sum_{j=1}^J u_j S_{0j} = 1 \\ u_j \geq \varepsilon > 0 \\ v_k \geq \varepsilon > 0 \end{cases} \quad (8)$$

Em que, h é a eficiência para orientação por *output*, $n = 1, \dots, N$, $j = 1, \dots, J$ e $k = 1, \dots, K$. Além, h_0 é a inversa da eficiência, ou seja, $h_0 = 1/_{Eff_0} = \sum_{k=1}^K v_k E_{0k} / \sum_{j=1}^J u_j S_{0j}$, gerando a restrição $\sum_{j=1}^J u_j S_{0j} = 1$ para manter a linearidade da programação.

Mello et al. (2005) e Batista (2009) apresentam a programação dual para ambas orientações. Por serem duais, elas e as equações primais possuem o mesmo valor para a função objetivo. Estas programações também são denominadas como Modelo do Envelope. As Equações 9 e 10 são referentes a orientação por *input*.

$$\text{Min } \eta_0 \quad (9)$$

$$\text{Sujeito às restrições: } \begin{cases} \eta_0 E_{0k} - \sum_{n=1}^N E_{nk} \lambda_n \geq 0, \forall k \\ -S_{0j} + \sum_{n=1}^N S_{nj} \lambda_n \geq 0, \forall j \\ \lambda_n \geq 0, \forall n \end{cases} \quad (10)$$

Já as Equações 11 e 12, são para orientação por *output*.

$$\text{Max } h_o \quad (11)$$

$$\text{Sujeito às restrições: } \begin{cases} E_{0k} - \sum_{n=1}^N E_{nk} \lambda_n \geq 0, \forall k \\ -h_0 S_{0j} + \sum_{n=1}^N S_{nj} \lambda_n \geq 0, \forall j \\ \lambda_n \geq 0, \forall n \end{cases} \quad (12)$$

Mariano, Almeida e Rebelatto (2006a) comentam que, para a função objetivo das programações duais, basta utilizar a eficiência, previamente calculada. De acordo com Andrade (1998), isto é possível, porque programações primais e duais possuem o mesmo valor para função objetivo.

Mariano, Almeida e Rebelatto (2006a) também conceituam o λ_n , também denominados de preço sombra, como a meta para uma DMU ineficiente se tornar eficiente, sendo $\sum_{n=1}^N E_{nk} \lambda_n$ e $\sum_{n=1}^N S_{nj} \lambda_n$ as metas para um *input* E_k e um *output* S_j , respectivamente, da DMU ineficiente sob análise. Quando a eficiência da DMU_n for igual a 1, ou seja, 100%, o seu valor de λ_n será 1 também e os lambdas restantes serão 0. Já para uma DMU ineficiente, o seu λ_n será 0 e os restantes poderão assumir valores iguais ou maiores do que 0, conforme imposto pela restrição na programação dual. Ainda sobre as DMUs ineficientes, Mariano, Almeida e Rebelatto (2006a) comentam que, os preços sombra que tiverem valor diferente de zero farão parte do *benchmark* da DMU ineficiente para se tornar eficiente.

2.1.2 Modelo BCC

O modelo BCC foi desenvolvido por Banker, Charnes e Cooper, em 1984, e as iniciais de seus nomes também originaram o nome deste modelo. Lopes et al. (2008) e Jebali, Essid e Khraief (2017) também o denominam de VRS, abreviatura do inglês *variable returns to scale*, ou seja, retornos variáveis à escala. Mello et al. (2005) explicam que, este modelo possibilita que DMUs tenham retornos crescentes de escala para as que operam com baixos valores de *inputs*, e tenham retornos decrescentes de escala para as que operam com altos valores. Ao contrário do modelo

CCR, Rosano-Peña (2012) comenta que qualquer variação nos *inputs* ocasiona uma variação não proporcional nos *outputs*.

Ainda, Mariano, Almeida e Rebelatto (2006b) comentam que, para este modelo, a fronteira de eficiência é linear por partes, conforme representado anteriormente pela Figura 2.

Mello et al. (2005) equacionam o modelo BCC orientado por *input* conforme Equações 13 e 14. Eles comentam que esta é a programação primal e também a denominam de Modelo dos Multiplicadores.

$$\text{Max } \eta_0 = \text{Max } \sum_{j=1}^J u_j S_{0j} - u_* \quad (13)$$

$$\text{Sujeito às restrições: } \begin{cases} \sum_{k=1}^K v_k E_{nk} = 1 \\ \sum_{j=1}^J u_j S_{nj} - \sum_{k=1}^K v_k E_{nk} - u_* \leq 0, \forall n \\ v_k \geq 0 \\ u_j \geq 0 \\ u_* \in \mathcal{R} \end{cases} \quad (14)$$

Já as Equações 15 e 16, são para orientação por *output*.

$$\text{Min } h_o = \text{Min } \sum_{k=1}^K v_k S_{0k} - v_* \quad (15)$$

$$\text{Sujeito às restrições: } \begin{cases} \sum_{j=1}^J u_j S_{0j} = 1 \\ \sum_{j=1}^J u_j S_{nj} - \sum_{k=1}^K v_k E_{nk} - v_* \leq 0, \forall n \\ v_k \geq 0 \\ u_j \geq 0 \\ v_* \in \mathcal{R} \end{cases} \quad (16)$$

Lopes et al. (2008) e Galvão (2008) comentam que a diferença das equações dos modelos BCC e CCR é a adição das variáveis u_* e v_* , para orientação por *input* e *output*, respectivamente, e que representam os fatores de escala, que podem ser constantes, crescentes ou decrescentes.

Mello et al. (2005) apresentam também a programação dual para ambas orientações e também são denominadas como Modelo do Envelope. As Equações 17 e 18 são referentes a orientação por *input*.

$$\text{Min } \eta_0 \quad (17)$$

$$\text{Sujeito às restrições: } \begin{cases} \eta_0 E_{0k} - \sum_{n=1}^N E_{nk} \lambda_n \geq 0, \forall k \\ -S_{0j} + \sum_{n=1}^N S_{nj} \lambda_n \geq 0, \forall j \\ \sum_{n=1}^N \lambda_n = 1 \\ \lambda_n \geq 0, \forall n \end{cases} \quad (18)$$

As Equações 19 e 20 são a representação da programação dual para orientação por *output*.

$$\text{Max } h_0 \quad (19)$$

$$\text{Sujeito às restrições: } \begin{cases} E_{0k} - \sum_{n=1}^N E_{nk} \lambda_n \geq 0, \forall k \\ -h_0 S_{0j} + \sum_{n=1}^N S_{nj} \lambda_n \geq 0, \forall j \\ \sum_{n=1}^N \lambda_n = 1 \\ \lambda_n \geq 0, \forall n \end{cases} \quad (20)$$

Para este modelo, os valores de λ_n encontrado com a programação dual possuem o mesmo significado explicado anteriormente no modelo CCR, com a diferença que o somatório para todos λ de uma determinada DMU deve ser 1, conforme imposto pela restrição.

2.1.3 Recapitulação da literatura

Para esta subseção, foram reapresentados aspectos importantes para a DEA, utilizando-se tabelas para possibilitar comparações e um melhor entendimento.

2.1.3.1 Agrupamento das equações

As equações do modelo CCR foram apresentadas na Tabela 1, sendo elas as Equações 5, 6, 7, 8, 9, 10, 11 e 12, respectivamente. Enquanto a Tabela 2 apresenta as equações referentes ao modelo BBC, reapresentando as Equações 13, 14, 15, 16, 17, 18, 19 e 20, respectivamente.

Tabela 1 – Equações do Modelo CCR

MODELO CCR	
Programação primal – Modelo dos Multiplicadores	
Orientação à <i>input</i>	Orientação à <i>output</i>
$\text{Max } \eta_o = \text{Max } \sum_{j=1}^J u_j S_{0j}$ $\text{s.a.: } \begin{cases} \sum_{j=1}^J u_j S_{nj} - \sum_{k=1}^K v_k E_{nk} \leq 0 \\ \sum_{k=1}^K v_k E_{0k} = 1 \\ u_j \geq \varepsilon > 0 \\ v_k \geq \varepsilon > 0 \end{cases}$	$\text{Min } h_o = \text{Min } \sum_{k=1}^K v_k E_{0k}$ $\text{s.a.: } \begin{cases} \sum_{k=1}^K v_k E_{nk} - \sum_{j=1}^J u_j S_{nj} \leq 0 \\ \sum_{j=1}^J u_j S_{0j} = 1 \\ u_j \geq \varepsilon > 0 \\ v_k \geq \varepsilon > 0 \end{cases}$
Programação dual – Modelo do Envelope	
Orientação à <i>input</i>	Orientação à <i>output</i>
$\text{Min } \eta_o$ $\text{s.a.: } \begin{cases} \eta_o E_{0k} - \sum_{n=1}^N E_{nk} \lambda_n \geq 0, \forall k \\ -S_{0j} + \sum_{n=1}^N S_{nj} \lambda_n \geq 0, \forall j \\ \lambda_n \geq 0, \forall n \end{cases}$	$\text{Max } h_o$ $\text{s.a.: } \begin{cases} E_{0k} - \sum_{n=1}^N E_{nk} \lambda_n \geq 0, \forall k \\ -h_o S_{0j} + \sum_{n=1}^N S_{nj} \lambda_n \geq 0, \forall j \\ \lambda_n \geq 0, \forall n \end{cases}$

Fonte: Elaborado pela autora (2016)

Tabela 2 – Equações do Modelo BCC

MODELO BCC	
Programação primal – Modelo dos Multiplicadores	
Orientação à <i>input</i>	Orientação à <i>output</i>
$\text{Max } \eta_o = \text{Max } \sum_{j=1}^J u_j S_{0j} - u_*$ $\text{s.a.: } \begin{cases} \sum_{k=1}^K v_k E_{nk} = 1 \\ \sum_{j=1}^J u_j S_{nj} - \sum_{k=1}^K v_k E_{nk} - u_* \leq 0, \forall n \\ v_k \geq 0 \\ u_j \geq 0 \\ u_* \in \Re \end{cases}$	$\text{Min } h_o = \text{Min } \sum_{k=1}^K v_k S_{0k} - v_*$ $\text{s.a.: } \begin{cases} \sum_{j=1}^J u_j S_{0j} = 1 \\ \sum_{j=1}^J u_j S_{nj} - \sum_{k=1}^K v_k E_{nk} - v_* \leq 0, \forall n \\ v_k \geq 0 \\ u_j \geq 0 \\ v_* \in \Re \end{cases}$
Programação dual – Modelo do Envelope	
Orientação à <i>input</i>	Orientação à <i>output</i>
$\text{Min } \eta_o$ $\text{s.a.: } \begin{cases} \eta_o E_{0k} - \sum_{n=1}^N E_{nk} \lambda_n \geq 0, \forall k \\ -S_{0j} + \sum_{n=1}^N S_{nj} \lambda_n \geq 0, \forall j \\ \sum_{n=1}^N \lambda_n = 1 \\ \lambda_n \geq 0, \forall n \end{cases}$	$\text{Max } h_o$ $\text{s.a.: } \begin{cases} E_{0k} - \sum_{n=1}^N E_{nk} \lambda_n \geq 0, \forall k \\ -h_o S_{0j} + \sum_{n=1}^N S_{nj} \lambda_n \geq 0, \forall j \\ \sum_{n=1}^N \lambda_n = 1 \\ \lambda_n \geq 0, \forall n \end{cases}$

Fonte: Elaborado pela autora (2016)

2.1.3.2 Revisão bibliográfica – DEA

Para esta subseção, foi realizada uma síntese das aplicações estudadas nos trabalhos dos autores citados na seção 2.1 do Capítulo 2, apresentados no Apêndice A.

2.2 REDES NEURAIS ARTIFICIAIS

Uma Rede Neural Artificial (RNA) pode ser definida, de acordo com Souza (2008), Andrade (2011) e Devadoss e Ligor (2013), como sendo um algoritmo computacional que possui um modelo matemático e foi inspirado no cérebro humano, em que ele aprende a partir do fornecimento de dados de entrada. Os autores ainda fazem uma comparação ao citar que uma RNA pode ter centenas ou milhares de unidades de processamento, enquanto um cérebro humano pode possuir bilhões de neurônios. Krishnan et al. (2018) entendem que uma RNA é considerada como uma classificação de *Machine Learning* ou aprendizado de máquina.

Raia Junior (2000) e Dartora (2003) comentam que os estudos se iniciaram em 1943, com a publicação de um artigo de McCullen e Pitts. Com o passar dos anos, a implementação da RNA foi facilitada devido ao avanço da capacidade dos computadores. Rich, Knight e Nair (2009) exemplificam algumas aplicações possíveis: percepções de visão e fala, controle robótico, jogo de xadrez, cálculos matemáticos, análises científicas, diagnósticos médicos, previsão de séries temporais, reconhecimento de padrões, análises financeiras, entre outros. Raia Junior (2000) ainda ressalta alguns benefícios no uso da RNA: habilidade de aprender com exemplos, robustez (pela capacidade de tratar com ruídos, ou seja, apesar da existência de alguma falha, a rede continua fornecendo respostas adequadas), velocidade de processamento e apresentam desempenho superior quando comparado com métodos estatísticos, dependendo do problema em questão.

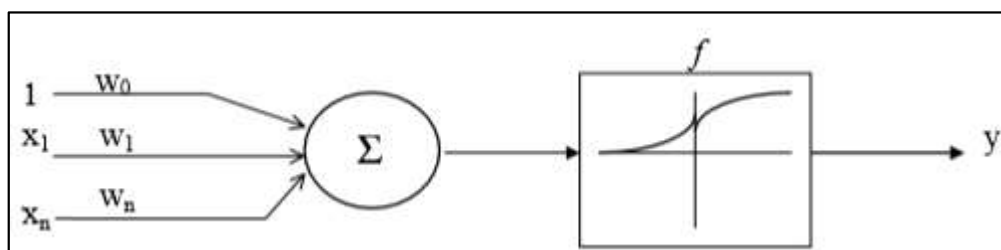
2.2.1 Topologia de rede

Raia Junior (2000), Haykin (2001) e Souza (2008) apresentam como topologia da rede, ou ainda arquitetura de rede, como sendo a sua estrutura, ou seja, a maneira pela qual os neurônios de uma RNA estão organizados, que está ligada ao algoritmo

de aprendizado que esta rede utilizará. Os autores citam duas classificações de arquitetura ou estrutura de rede: *single layer perceptrons* (ou redes alimentadas adiante com camada única) e *multilayer perceptron* (ou redes alimentadas diretamente com múltiplas camadas). Freitas (2001) e Guimarães (2006) comentam que a arquitetura de rede é definida pelas seguintes características: número de camada de rede, número de neurônios em cada camada e tipo de conexão entre os neurônios.

Para auxiliar no entendimento destes conceitos, Guimarães (2006), Thakur, Bhattacharyya e Mondal (2016) e Mulero et al. (2017) conceituam camada de entrada como sendo onde os dados são apresentados a rede, camada intermediária ou oculta como onde é realizado a maior parte do processamento e podem ser consideradas como extratoras de características, e camada de saída que é onde é apresentado o resultado final. Estes autores ainda comentam que, quanto mais camadas de neurônios, melhor será o desempenho da rede, visto que aumenta a capacidade de aprendizado e, conseqüentemente, melhorando a precisão do resultado gerado. Bressan (2001) ilustra o neurônio artificial com a Figura 3.

Figura 3 – Neurônio artificial



Fonte: Adaptado de Bressan (2001)

Conforme apresentado, a estrutura possui três componentes básicos: um elemento somatório Σ , uma função de ativação f e ponderações w . Segundo Bressan (2001), o neurônio artificial soma as entradas ou insumos x_n multiplicadas por w_n , que se desloca pela função de ativação, resultando no produto y . Matematicamente, este sistema está representado na Equação 21.

$$y = f(\sum_{i=0}^N x_i w_i) \quad (21)$$

A Equação 22 apresenta a função de ativação.

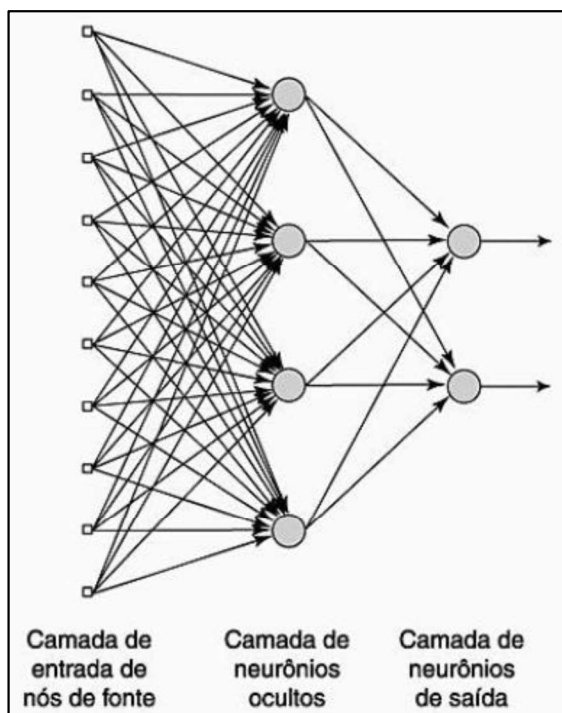
$$f = \frac{1}{1 + e^{-(\sum_{i=0}^N x_i w_i)}} \quad (22)$$

Assim, para as topologias, *single layer perceptron* (SLP) pode ser definida, de acordo com Haykin (2001) e Andrade (2011), como sendo a estrutura que consiste em uma camada de entrada de nós de fonte que se projeta sobre uma camada de saída, mas não vice-versa. Por isso, conforme Bressan (2001), esta rede é do tipo alimentada adiante ou *feedforward* (FF). A classificação de camada única está relacionada à camada de saída de neurônios, pois é onde ocorrem os cálculos de computação, sendo assim, não é contada a camada de entrada. Giacomel (2016) comenta que esta classificação de rede é somente para problemas lineares, sendo necessário utilizar uma rede mais complexa para problemas maiores.

Já a classificação de múltiplas camadas na *multilayer perceptron* (MLP) implica na presença de uma ou mais camadas ocultas. Conforme Haykin (2001), tem a função de intervir entre as camadas de entrada e de saída de maneira útil, como por exemplo, extraindo estatísticas de ordem elevada e ainda, aumentando a capacidade de processamento e aprendizagem, conforme citam Gaio, Castro Junior e Oliveira (2007) e Giacomel (2016). Além disso, estes autores comentam que esta rede é mais avançada do que a de única camada, e propiciando assim a resolução de problemas mais complexos. Freitas (2001) e Andrade (2011) explicam que pode haver mais de um neurônio em quaisquer camadas de entrada ou de saída da rede. A Figura 4 apresenta esta rede com múltiplas camadas e também do tipo *feedforward*.

Conforme pode ser observado na figura, a camada de entrada fornece os sinais de entrada para a segunda camada (a primeira camada oculta). Haykin (2001) e Dumitru e Gligor (2017) seguem a explicação ao citar que os sinais de saída da segunda camada são utilizados como entrada para a terceira camada, e assim por adiante. O conjunto de sinais de saída da camada de saída corresponde a resposta global da rede para os dados fornecidos pelos nós da primeira camada. Ainda, esta rede é considerada como totalmente conectada, por todos os nós de uma camada estarem conectados a todos os nós da camada seguinte. No entanto, se alguma conexão estivesse faltando, seria uma rede parcialmente conectada.

Figura 4 – Redes alimentadas diretamente com múltiplas camadas



Fonte: Haykin (2001, p. 48)

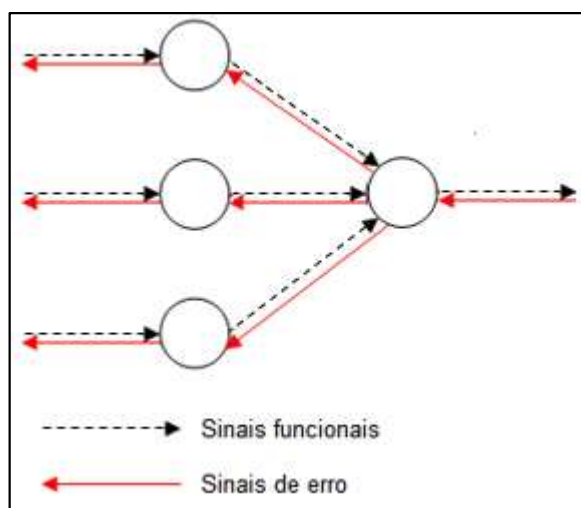
2.2.2 Processo de aprendizagem

Gaio, Castro Junior e Oliveira (2007), Souza (2008) e Devadoss e Ligor (2013) entendem que processo de aprendizagem é a maneira que a RNA aprende com os exemplos fornecidos. Este procedimento pode ser classificado em três categorias: treinamento supervisionado, não supervisionado e híbrida. Raia Junior (2000), Bressan (2001) e Guimarães (2006) definem como treinamento supervisionado quando são dados os valores de entrada e de saída, em outras palavras, o sistema compara a saída da rede com a resposta desejada, previamente dada; não supervisionado como quando não exige valores de saída, então a rede agrupa os valores de maior semelhança; e o treinamento híbrido quando mescla os dois conceitos apresentados anteriormente.

Apesar de haver diversos algoritmos de aprendizagem, Souza (2008) considera o *backpropagation* ou retropropagação como sendo um dos principais processos, e Rich, Knight e Nair (2009), Moghaddam, Moghaddam e Esfandyari (2016) e Mulero et al. (2017) ressaltam que é o mais popular aplicado às redes MLP. Gaio, Castro Junior e Oliveira (2007) e Dumitru e Gligor (2017) comentam que *backpropagation* segue a norma do treinamento supervisionado, tendo o aprendizado

ocorrendo em duas fases: a primeira etapa é a *forward*, no qual os valores de saída são calculados com base nos valores de entrada dados (definidos por Haykin (2001) como sinais funcionais), e a segunda é a *backward*, em que é calculado o erro, que é a diferença dos valores de saída obtidos com os valores desejados, da última camada até a camada de entrada (chamados de sinais de erro, por Haykin (2001)), conforme pode ser observado na Figura 5. De acordo com Guimarães (2006), é nesta segunda etapa que são usados os valores dos erros para ajustar os pesos na camada de saída. Bressan (2001) esclarece que este algoritmo ajusta as ponderações a fim de minimizar os erros de previsão dentro dos testes.

Figura 5 – Parte de um MLP com aprendizagem por *backpropagation*



Fonte: Adaptado de Haykin (2001)

Giacomel (2016) indica a utilização deste algoritmo para problemas com as seguintes características:

- a) há uma grande quantidade de dados de entrada e é desconhecido a relação com a saída esperada;
- b) é possível criar um conjunto de exemplos do comportamento desejado; e
- c) a solução pode variar ao longo do tempo, mesmo para os mesmos valores de entrada.

O autor exemplifica o uso citando problemas como processamento de imagens, reconhecimento de voz, previsão de séries temporais, entre outros.

Bressan (2001) e Guimarães (2006) comentam que, para o aprendizado da rede, os dados devem ser classificados em treino, validação e teste. Thakur,

Bhattacharyya e Mondal (2016) e Mulero et al. (2017) justificam esta divisão ao citar que, durante o treinamento, podem ocorrer erros inaceitavelmente grandes devido a *overfitting*. No treino, a rede entende o padrão de comportamento da série, por isso, ela deve ser suficientemente grande. Nesta etapa, a rede tenta aprender as características desses padrões. O conjunto de validação serve para comparar a performance da rede, e o de teste, para comprovar o desempenho da arquitetura escolhida. Andrade (2011) conceitua *overfitting* como sendo um erro que pode acontecer quando a rede memoriza os padrões de treino, não prevendo corretamente diferentes pontos que não estejam nos dados de entrada; ou quando há número excessivo de camadas.

2.2.3 Revisão bibliográfica – RNA

Para esta subseção, foi realizada uma síntese das aplicações estudadas nos trabalhos dos autores citados na seção 2.2 deste capítulo, apresentados no Apêndice B para os parâmetros escolhidos pelos autores e no Apêndice C para os dados utilizados pelos autores.

2.3 MODELOS BOX-JENKINS

Para esta seção, foram apresentados os conceitos dos Modelos de Box-Jenkins. Para facilitar o entendimento deste método, é necessário a compreensão de alguns conceitos básicos, que serão estudados na próxima subseção.

2.3.1 Conceitos básicos para compreensão do método Box-Jenkins

Nesta seção, foram estudados alguns conceitos básicos dos Modelos de Box-Jenkins, como séries temporais e processos estocásticos.

2.3.1.1 Previsão de séries temporais

De acordo com a definição de Fischer (1982) e de Levine et al. (2012), uma série temporal consiste em um conjunto de dados numéricos coletados em um período de tempo. Estes dados podem ser classificados como determinísticos ou estocásticos.

Fischer (1982) entende que uma série é determinística quando é possível determinar precisamente os valores futuros com uma relação matemática; enquanto é estocástica quando a previsão dos valores futuros só pode ser descrita em termos probabilísticos, já que os dados não são influenciados somente pelo tempo, ou seja, é necessária alguma outra variável para explicar algum ponto irregular da série.

Lapponi (2005), Levine et al. (2012) e Ramos, Santos e Rebelo (2015) citam exemplos de casos onde a previsão é importante para a organização, sendo eles: gestores de indústria que fazem prognose de demanda e de estoque de produtos; executivos que trabalham para o governo necessitam prever índices de desemprego, inflação, entre outros; administradores de universidades podem utilizar da previsão de número de matrículas para ajustar a quantidade de aulas oferecidas, salas disponíveis; e ainda, predizer preço das ações presentes na bolsa de valores. Taneja et al. (2016) ainda explicam que métodos de previsão de séries temporais consiste em utilizar dados históricos para projetar valores futuros.

Werner (2004) e Kirchner (2006) comentam que uma série temporal é constituída por quatro elementos:

- a) tendência: analisa o sentido do deslocamento da série;
- b) ciclo: representa o movimento oscilatório, o qual tem a propensão de ser periódico ao longo do tempo;
- c) sazonalidade: padrão de redução e/ou aumento nos valores observados, com duração de curto prazo, podendo estar associada a, por exemplo, mudanças climáticas; e
- d) ruído aleatório ou erro: variações que ocorrem sem uma causa específica, sem padrão e não pode ser matematicamente modelado.

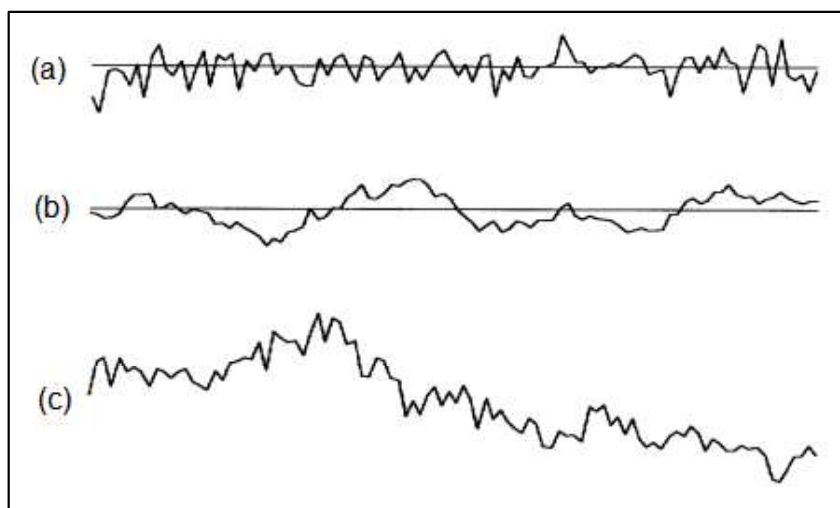
2.3.1.2 Processo estocástico

Conforme Fischer (1982) e Kirchner (2006), um processo estocástico pode ser definido como um modelo que descreve a estrutura probabilística de uma sequência de observações. Em outras palavras, um estudo de alguma série temporal é gerado por um processo estocástico, já que é com a utilização deste que é possível descrever numericamente a função de probabilidade da série e assim, deduzir a probabilidade de um valor futuro.

Um processo estocástico pode ser descrito como a família de funções de probabilidade que foram associadas por cada resultado de um experimento realizado, dentro de um espaço de probabilidade $(\Omega, \mathfrak{F}, P(\bullet))$, definido por Fischer (1982) como sendo Ω o espaço amostral, $\mathfrak{F}$ a σ -álgebra de subconjuntos de Ω , e $P(\bullet)$ a probabilidade do evento, que é elemento de $\mathfrak{F}$ e é caracterizado por Lapponi (2005), como sendo “m subconjunto composto por um ou mais resultados do espaço amostral.

Fischer (1982) e Queirós (2012) classificam um processo estocástico como estacionário se suas distribuições finito-dimensionais não variarem no decorrer do tempo e a média e a variância dos valores da série forem constantes; e como não estacionário se a série de observações for variante com o perpassar do tempo. Pellegrini (2000) apresenta visualmente este conceito, com a Figura 6.

Figura 6 – Processo estocástico estacionário e não estacionário



Fonte: Pellegrini (2000, p. 18)

Pellegrini (2000) explica que os gráficos (a) e (b) demonstram um processo estacionário, pois as séries temporais oscilam de maneira estável no decorrer do tempo, sobre um valor fixo de média. Já o gráfico (c) apresenta um exemplo de processo não estacionário, uma vez que não varia com um valor fixo de média.

Ainda Werner (2004) classifica um processo estocástico estacionário em forte e fraco. Um processo fraco (ou, como Fischer (1982) também denomina, estacionário de segunda ordem) é aquele que sua média e variância permanecem constantes no tempo, e um processo fortemente estacionário se todos os momentos estatísticos forem constantes ao longo do tempo.

Fischer (1982) ainda conceitua um processo evolutivo homogêneo como sendo uma série temporal não estacionária com um comportamento evolutivo ao decorrer do tempo, com exemplos encontrados na economia, sendo eles: preço, renda, vendas, consumo, investimentos, entre outros. O autor e também Sharafi et al. (2017) comentam que, diferenciando uma ou mais vezes, é possível transformar esta série em estacionária, sendo o número de vezes que a série for diferenciada denominado de ordem de homogeneidade. Fischer (1982) e Kim et al. (2017) explicam que, se Y_t for uma série não estacionária homogênea de primeira ordem, então $Z_t = Y_t - Y_{t-1} = \nabla Y_t$ é estacionária, e ∇ é denominado como operador de diferença de retardo. Do mesmo modo, se Y_t for uma série não estacionária homogênea de segunda ordem, então $Z_t = \nabla^2 Y_t = \nabla Y_t - \nabla Y_{t-1}$ será estacionária.

2.3.1.3 Função de autocorrelação

Para auxiliar na determinação se um processo é estacionário ou não estacionário, é necessário utilizar a Função da Autocorrelações, do inglês *Autocorrelation Function* (ACF) e a Função de Autocorrelações Parciais, do inglês *Partial Autocorrelation Function* (PACF), para analisar os seus respectivos coeficientes, pois são com eles que se observa o comportamento da série temporal. Fischer (1982), Pellegrini (2000) e Lemos (2006) explicam que o coeficiente de autocorrelação populacional (ρ_k) caracteriza a correlação da série temporal com ela mesmo, e apresentam a Equação 23.

$$\rho_k = \frac{\sum (Y_t - \mu_t)(Y_{t-k} - \mu_{t-k})}{\sqrt{\text{Var}(Y_t) * \text{Var}(Y_{t-k})}} = \frac{\text{Cov}(Y_t, Y_{t-k})}{\sigma(Y_t) * \sigma(Y_{t-k})} \quad (23)$$

Onde, Y_t é a observação da série no período t , k é a defasagem ou *lag*, Y_{t-k} é a observação da série no período $t - k$, μ é a média das observações (que, de acordo com Lapponi (2005), é o somatório de todas as observações dividido pela quantidade de observações realizadas), Var é a variância (Lapponi (2005) descreve como sendo $[\sum (Y_1 - \mu)^2 + (Y_2 - \mu)^2 + \dots]/t$, onde t representa a quantidade total de observações), Cov significa covariância (Lapponi (2005) define como sendo

$[\sum(Y_t - \mu_t) * (Y_{t-k} - \mu_{t-k})]/t$, onde t representa a quantidade total de observações) e σ é o desvio padrão (que, segundo Lapponi (2005), é a raiz quadrada positiva da variância).

Fischer (1982) e Wei (2005) ainda comentam que, para um processo estacionário, há a igualdade $\text{Var}(Y_t) = \text{Var}(Y_{t-k})$, podendo reescrever a Equação 23 e resultando, assim, a 24.

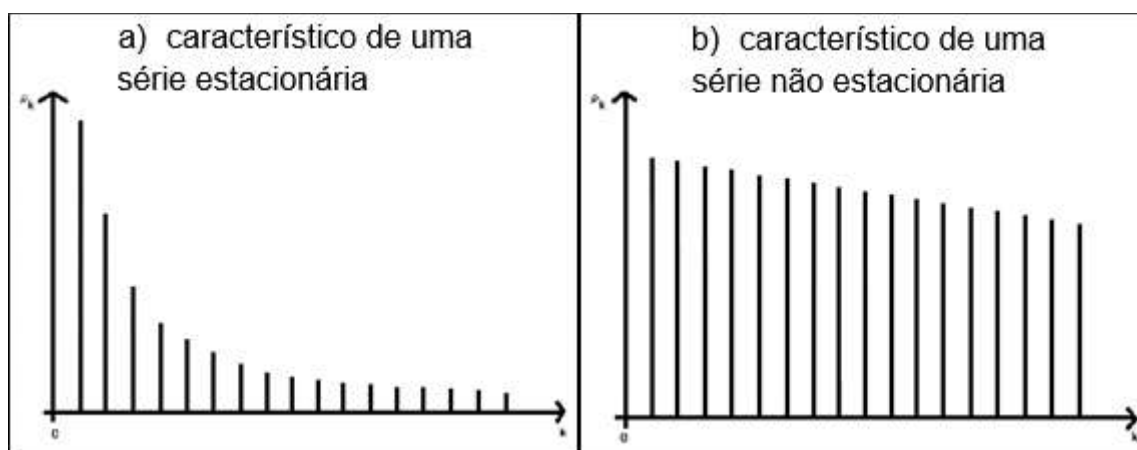
$$\rho_k = \text{Cov}(Y_t, Y_{t-k}) / \text{Var}(Y_t) \quad (24)$$

Ainda, Fischer (1982), Bressan (2001) e Pellegrini (2000) apresentam o coeficiente de autocorrelação amostral r_k , expresso na Equação 25.

$$r_k = \frac{\sum_{t=k+1}^n (Y_t - \mu)(Y_{t-k} - \mu)}{\sum_{t=1}^n (Y_t - \mu)^2} \quad (25)$$

Fischer (1982) resalta uma característica importante para o gráfico da ACF, conhecido por correlograma: a relação entre a defasagem k e o coeficiente ρ_k . Quando a série é estacionária, quanto mais k aumentar, mais rapidamente o coeficiente se aproximará de zero, como ilustra a Figura 7a. Em contrapartida, se a série for não estacionária, o ρ_k não diminuirá tão rapidamente ao passo que aumenta a defasagem, apresentado na Figura 7b.

Figura 7 – Correlogramas



Fonte: Adaptado de Fischer (1982)

O coeficiente de autocorrelação parcial de ordem k é apresentado na Equação 26, conceituado por Wei (2005).

$$\phi_{kk} = \frac{r_k - \phi_1\rho_{k-1} - \phi_2\rho_{k-2} - \dots - \phi_{k-1}\rho_1}{1 - \phi_1\rho_1 - \phi_2\rho_2 - \dots - \phi_{k-1}\rho_{k-1}} \quad (26)$$

Onde, ϕ_{kk} é o coeficiente de autocorrelação parcial.

2.3.2 Metodologia Box-Jenkins e seus modelos matemáticos

A Metodologia de Box-Jenkins, também conhecida como Modelos Auto Regressivos Integrados a Média Móvel, do inglês *Autoregressive Integrated Moving Average* (ARIMA), conforme comentado por Pellegrini (2000) e Queirós (2012), foi apresentada na década de 1970, pelos estatísticos George Box e Gwilym Jenkins. Ramos, Santos e Rebelo (2015) e Sen, Roy e Pal (2016) citam que o ARIMA é um dos modelos lineares mais versáteis para calcular previsão de série temporal.

Pellegrini (2000) ainda explica que, para o ARIMA, os valores de uma série temporal são altamente dependentes, ou seja, podem ser explicados por valores prévios da série. Estes modelos matemáticos estudam o comportamento da autocorrelação existente nos valores da série temporal, além de quantificar a significância da autocorrelação. Werner (2004) comenta que, embasado neste comportamento, é possível realizar prognose, e quão melhor a estrutura de correlação, melhor será a previsão.

Lemos (2006) e Sharafi et al. (2017) explicam que há algumas classificações para este método matemático: auto regressão (AR), que modela a dependência de valores atuais com valores passados; integração (I), que remove a tendência, tornando a série estacionária; e média móvel (MA), que assume que valores são dependentes de erros de previsão de períodos passados. Ainda, é necessário determinar alguns parâmetros, sendo eles:

- a) p – ordem máxima de parâmetros de auto regressão simples;
- b) d – número de diferenciações não sazonais;
- c) q – ordem máxima de média móvel simples;
- d) P – ordem máxima dos parâmetros de auto regressão sazonal;

- e) D – número de diferenciações sazonais;
- f) Q – número máximo dos parâmetros de média móvel sazonal; e
- g) s – número de períodos por ciclo sazonal.

Os modelos serão explicados a seguir, separados em estacionários e não estacionários.

2.3.2.1 Modelos estacionários

Nesta seção, foram estudados os modelos Auto Regressivo (AR), de Médias Móveis (MA) e Auto Regressivo de Médias Móveis (ARMA).

2.3.2.1.1 Modelo Auto Regressivo (AR)

Wei (2005), Lemos (2006), Rossi (2013), Yuan, Liu e Fang (2016) e Kim et al. (2017) definem que, em um modelo Auto Regressivo AR(p), a série histórica Y_t é descrita pela regressão de valores passados e pelo ruído aleatório e_t , conforme descrito na Equação 27. Esta equação pode ser reescrita em termos do operador B , de acordo com a Equação 28, ou ainda, na sua forma simplificada, apresentada na Equação 29.

$$Y_t = c + \gamma_1 Y_{t-1} + \gamma_2 Y_{t-2} + \dots + \gamma_p Y_{t-p} + e_t \quad (27)$$

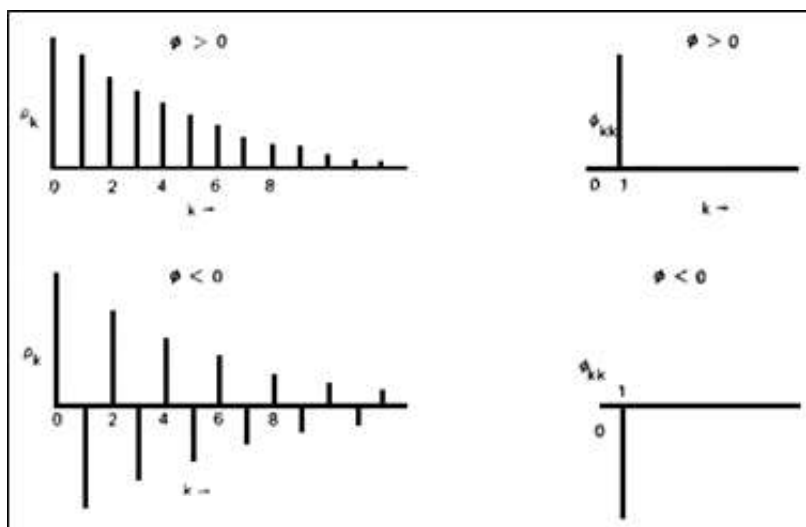
$$(1 - \gamma_1 B - \gamma_2 B^2 - \dots - \gamma_p B^p) Y_t = e_t \quad (28)$$

$$\gamma(B) Y_t = e_t \quad (29)$$

Onde, c é um termo constante, γ_i o coeficiente de auto regressão no período i , e_t é o erro aleatório no período t , e $\gamma(B)$ é um polinômio auto regressivo de ordem p , em que $B^p Y_t = Y_{t-p}$.

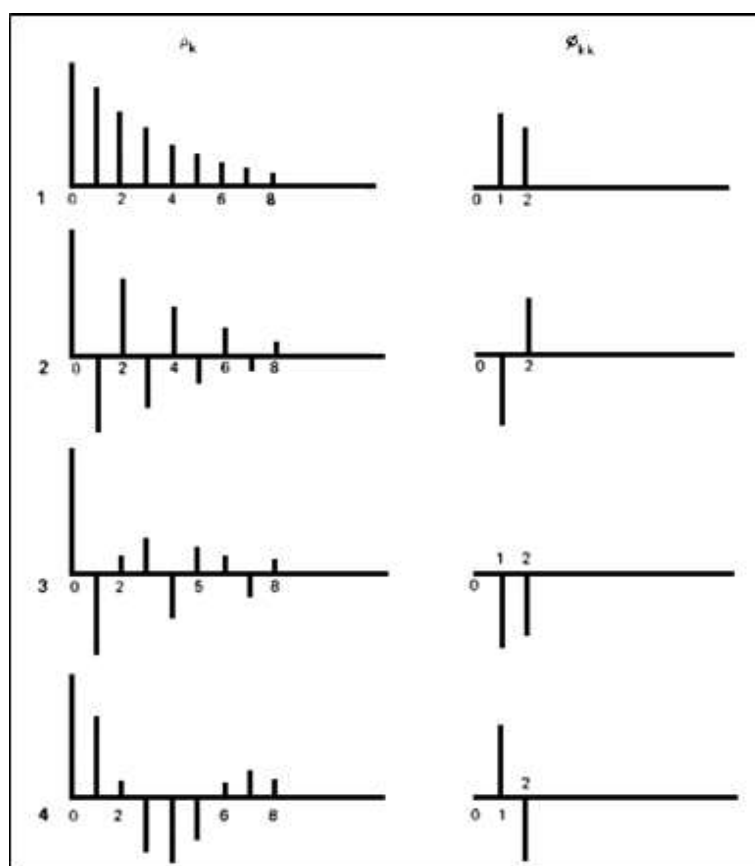
Pellegrini (2000) comenta que, se $p = 1$, então se tem um processo auto regressivo de primeira ordem AR(1). Nas Figuras 8 e 9, Fischer (1982) mostra o correlograma de um modelo AR(1) e do AR(2), respectivamente.

Figura 8 – Correlograma de AR(1)



Fonte: Adaptado de Fischer (1982)

Figura 9 – Correlograma de AR(2)



Fonte: Adaptado de Fischer (1982)

2.3.2.1.2 Modelo de Médias Móveis (MA)

O modelo de média móvel MA(q), conceituado por Wei (2005), Lemos (2006), Rossi (2013), Yuan, Liu e Fang (2016) e Kim et al. (2017), faz a regressão da variável Y_t , resultando da combinação do erro do período atual com os do passado (e_{t-q}), de acordo com a Equação 30, e as Equações 31 e 32 apresentam a sua forma polinomial.

$$Y_t = \mu - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} + e_t \quad (30)$$

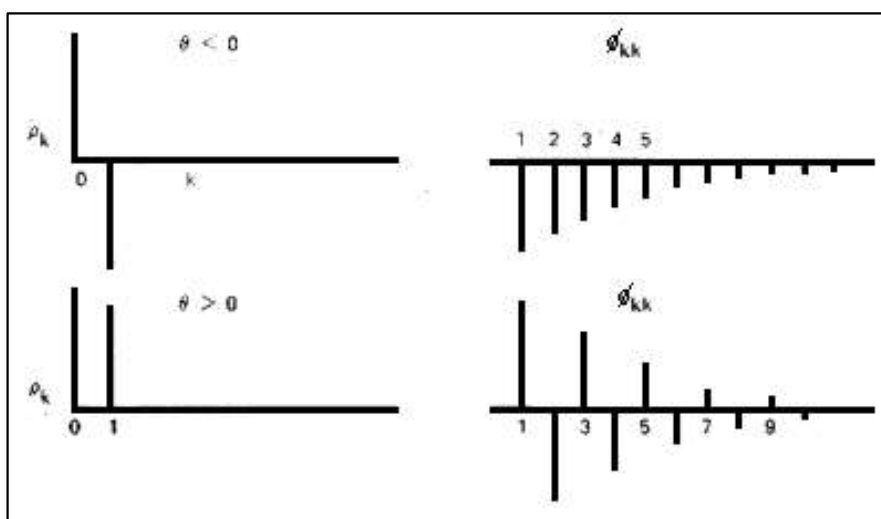
$$Y_t = \mu + (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) e_t \quad (31)$$

$$Y_t = \theta(B) e_t + \mu \quad (32)$$

Onde, θ_i é o coeficiente de média móvel do período i , e_t é o erro aleatório do período t e $\theta(B)$ é um polinômio de média móvel de ordem q .

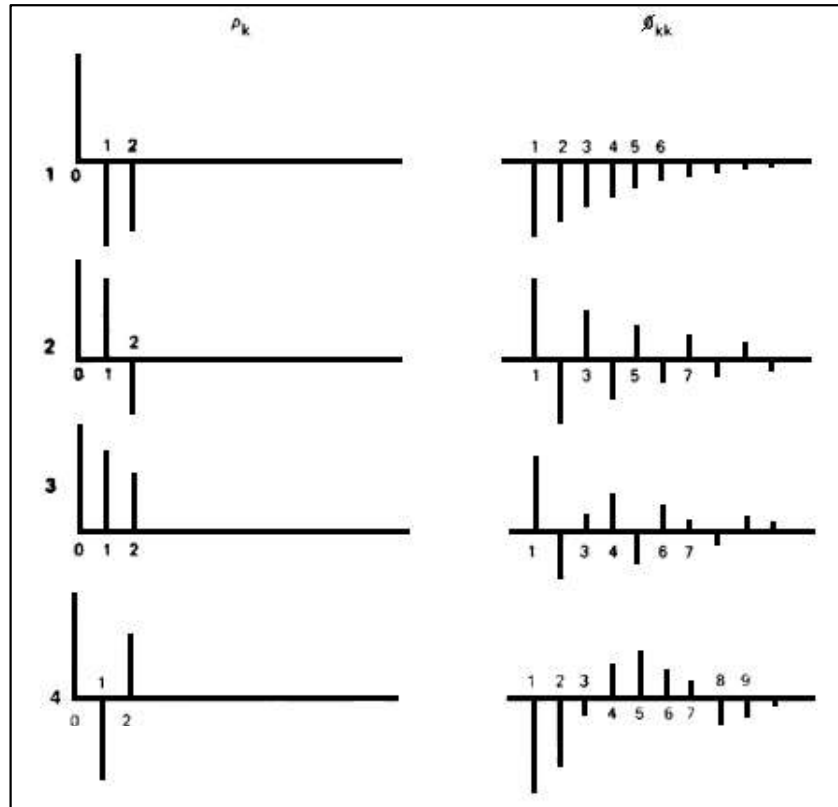
As Figuras 10 e 11 representam os correlogramas dos modelos MA(1) e MA(2) respectivamente, de acordo com Fischer (1982).

Figura 10 – Correlograma de MA(1)



Fonte: Adaptado de Fischer (1982)

Figura 11 – Correlograma de MA(2)



Fonte: Adaptado de Fischer (1982)

2.3.2.1.3 Modelo Auto Regressivo de Médias Móveis (ARMA)

Pellegrini (2000) e Kim et al. (2017) comentam que, quando é combinado as características dos processos AR(p) e MA(q), o modelo então denomina-se auto regressivo de médias móveis, ou seja, ARMA(p,q), conforme as Equações 33 e 34, descritas por Lemos (2006), Rossi (2013), Taneja et al. (2016) e Yuan, Liu e Fang (2016).

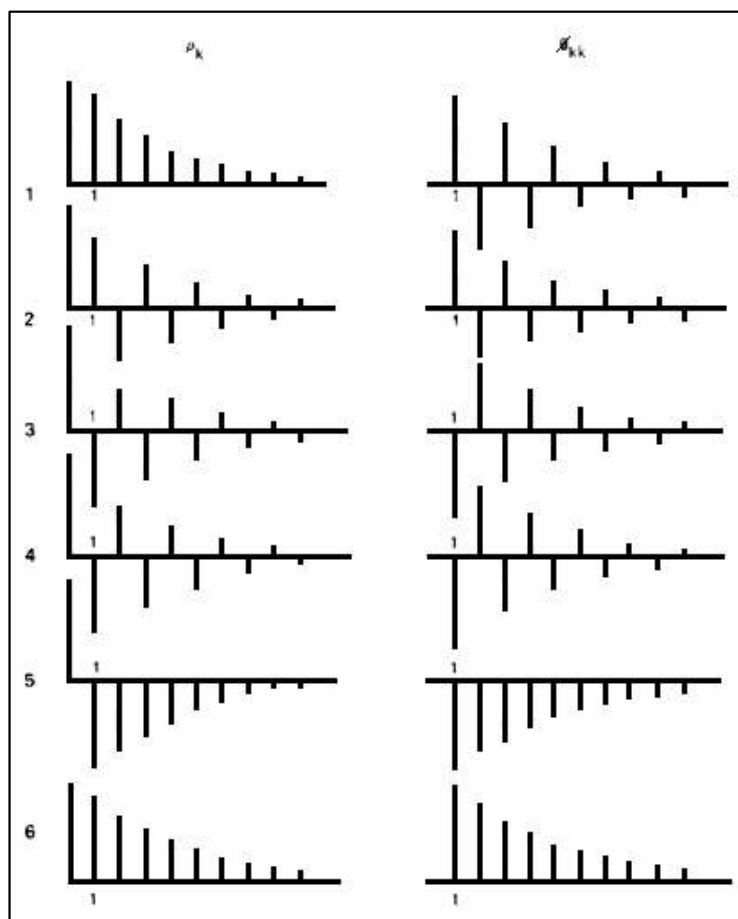
$$Y_t = c + \gamma_1 Y_{t-1} + \dots + \gamma_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (33)$$

$$\gamma(B)Y_t = c + \theta(B)e_t \quad (34)$$

Lemos (2006) faz uma relação entre os modelos: o AR(p) corresponde ao ARMA(p,0) enquanto o modelo MA(q) corresponde ao ARMA(0,q). Fischer (1982) comenta que este modelo não necessita de muitos parâmetros para descrever a série

de tempo, geralmente satisfazendo a condição de $p + q \leq 2$. Ainda, este autor exhibe o correlograma de ARMA(1,1) na Figura 12.

Figura 12 – Correlograma de ARMA(1,1)



Fonte: Adaptado de Fischer (1982)

2.3.2.2 Modelos não estacionários

Para esta seção, foi estudado o modelo Auto Regressivo Integrado de Médias Móveis (ARIMA).

2.3.2.2.1 Modelo Auto Regressivo Integrado de Médias Móveis (ARIMA)

Para o modelo auto regressivo integrado de médias móveis, caracterizado por ARIMA(p, d, q), Lemos (2006), Rossi (2013) e Taneja et al. (2016) apresentam-no na Equação 35 e citam que os valores de p , d e q geralmente variam entre 0, 1 e 2. Fischer (1982) explica que a ordem d representa o operador capaz de transformar um

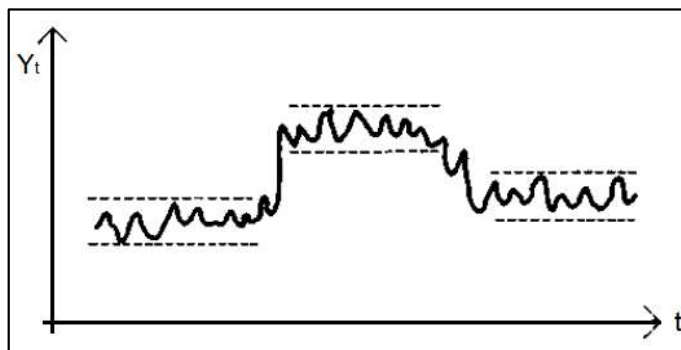
processo evolutivo homogêneo em estacionário. Pellegrini (2000) comenta que a palavra “integrado” deste modelo representa “somado”.

$$\gamma(B)(1 - B)^d Y_t = \theta(B)e_t \quad (35)$$

Em que, $(1 - B)^d$ é igual a ∇^d , sendo o operador de diferenciação, tal que $\nabla Y_t = Y_t - Y_{t-1}$.

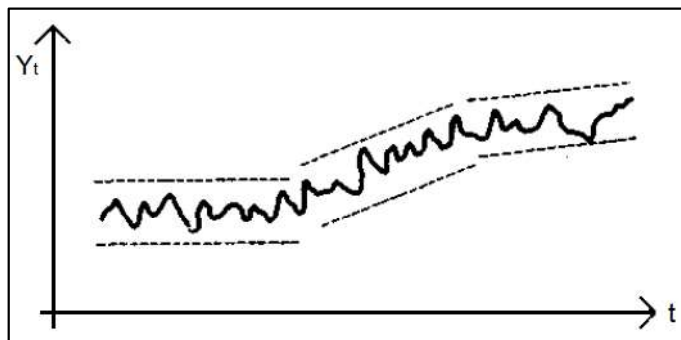
Fischer (1982) ressalta que a ordem máxima de diferenciação é, geralmente, igual a dois, e estes casos de $d = 1$ e $d = 2$ se caracterizam por descrever uma não estacionariedade quanto ao nível e/ou inclinação dos dados no gráfico. Necessita-se de uma diferenciação ($d = 1$), quando há mudança no nível, isto é, o comportamento da série varia ao redor de um nível por um certo período de tempo e depois, muda este nível, mas sem ter uma mudança significativa na direção, como ilustra a Figura 13. De acordo com Fischer (1982), é preciso $d = 2$ quando a série é não estacionária também quanto a inclinação, ou seja, quando no gráfico, a série muda o nível e a direção, conforme Figura 14.

Figura 13 – Série não estacionária quanto ao nível



Fonte: Adaptado de Fischer (1982)

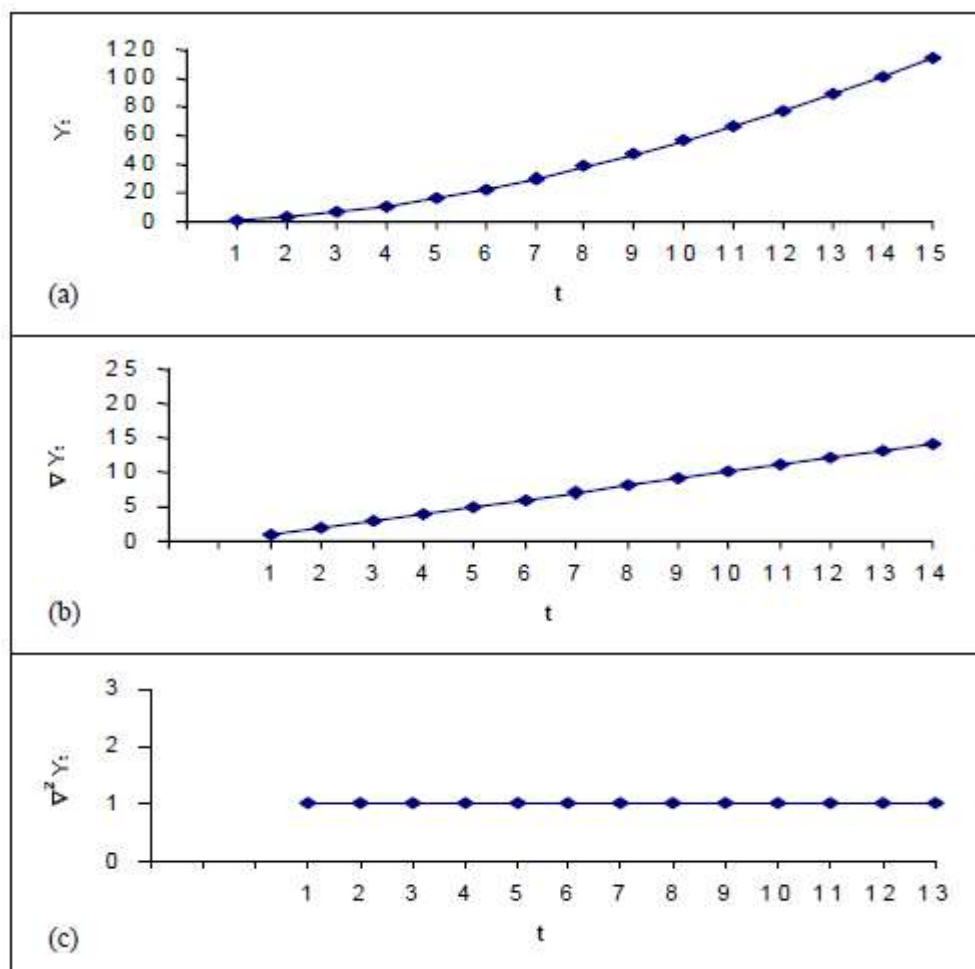
Figura 14 – Série não estacionária quanto ao nível e à inclinação



Fonte: Adaptado de Fischer (1982)

Pellegrini (2000) mostra na Figura 15, o efeito da diferenciação em uma série temporal não estacionária.

Figura 15 – Redução da não estacionariedade de uma série temporal após sucessivas diferenciações



Fonte: Adaptado de Pellegrini (2000)

O primeiro gráfico (letra a) apresenta um comportamento homogêneo não estacionário na média e na declividade. Após uma diferenciação $Y_t - Y_{t-1}$ (letra b), a série passa a ser não estacionária somente na média, conforme cita Pellegrini (2000). Depois de uma segunda diferenciação $(Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) = Y_t - 2Y_{t-1} + Y_{t-2}$, a série torna-se estacionária, conforme mostrado na letra c da figura.

Fischer (1982) sugere que a escolha do valor de d pode ser auxiliada com a análise do comportamento da variância das séries sucessivamente diferenciadas. A variância de uma série não estacionária é maior do que uma transformada em

estacionária através do operador de diferença, e se ocorrer uma operação diferença a mais do que o necessário, este valor de variância aumentará.

2.3.2.3 Modelos sazonais

Lemos (2006), Taneja et al. (2016) e Sharafi et al. (2017) conceituam um modelo sazonal como sendo $ARIMA(P,D,Q)_s$ e entendem que é a versão sazonal da Equação 35, substituindo o operador B pelo operador sazonal B^S , conforme descrito na Equação 36, descrevendo somente o comportamento sazonal e não considerando o erro aleatório. Fischer (1982) e Taneja et al. (2016) ainda conceituam sazonalidade como sendo uma tendência de repetição de um determinado comportamento dos dados, ocorrendo com uma regularidade no tempo.

$$\Gamma(B^S)(1 - B^S)^D Y_t = \theta(B^S) \quad (36)$$

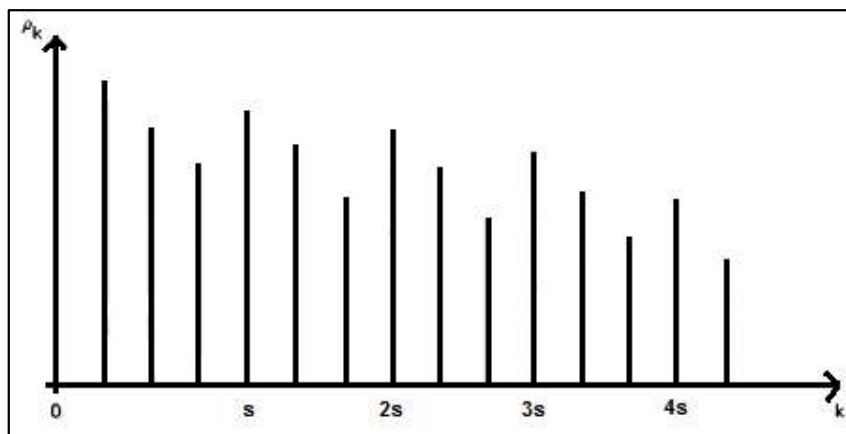
Onde, $\Gamma(B^S)$ é um polinômio auto regressivo sazonal de ordem P , $\theta(B^S)$ é o polinômio de média móvel sazonal de ordem Q , e $(1 - B^S)^D$ que é igual a ∇_S^D e representa o operador de diferenciação sazonal, tal que $\nabla_S Y_t = Y_t - Y_{t-s}$.

Ainda, Pellegrini (2000), Werner (2004), Lemos (2006) e Rossi (2013) apresentam o modelo geral sazonal $ARIMA(p,d,q)(P,D,Q)_s$ na Equação 37, que inclui o erro aleatório do período.

$$\gamma(B)\Gamma(B^S)(1 - B^S)^D Y_t = \theta(B)\theta(B^S)e_t \quad (37)$$

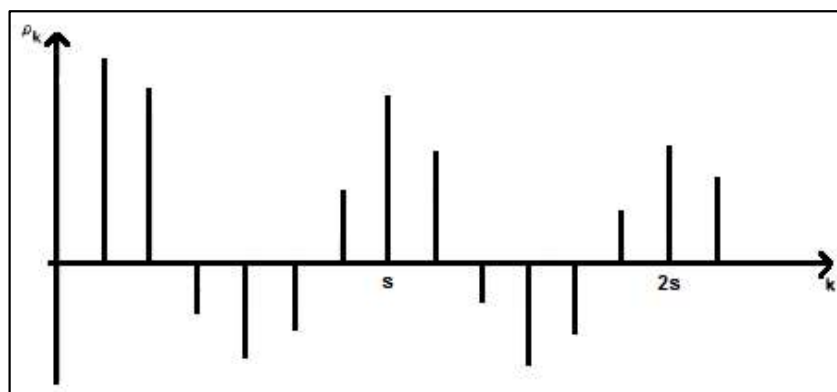
Fischer (1982) apresenta o correlograma de uma série sazonal, com período igual a “s”, não estacionária, na Figura 16, e estacionária, na Figura 17. A presença de picos nos correlogramas é característica de série sazonal; e este autor ainda comenta que é possível retirar a sazonalidade de uma série, criando uma nova série temporal utilizando o operador de diferença de retardo, tornando-a estacionária.

Figura 16 – Correlograma de um processo não estacionário sazonal



Fonte: Adaptado de Fischer (1982)

Figura 17 – Correlograma de um processo estacionário sazonal



Fonte: Adaptado de Fischer (1982)

2.3.3 Etapas de modelagem

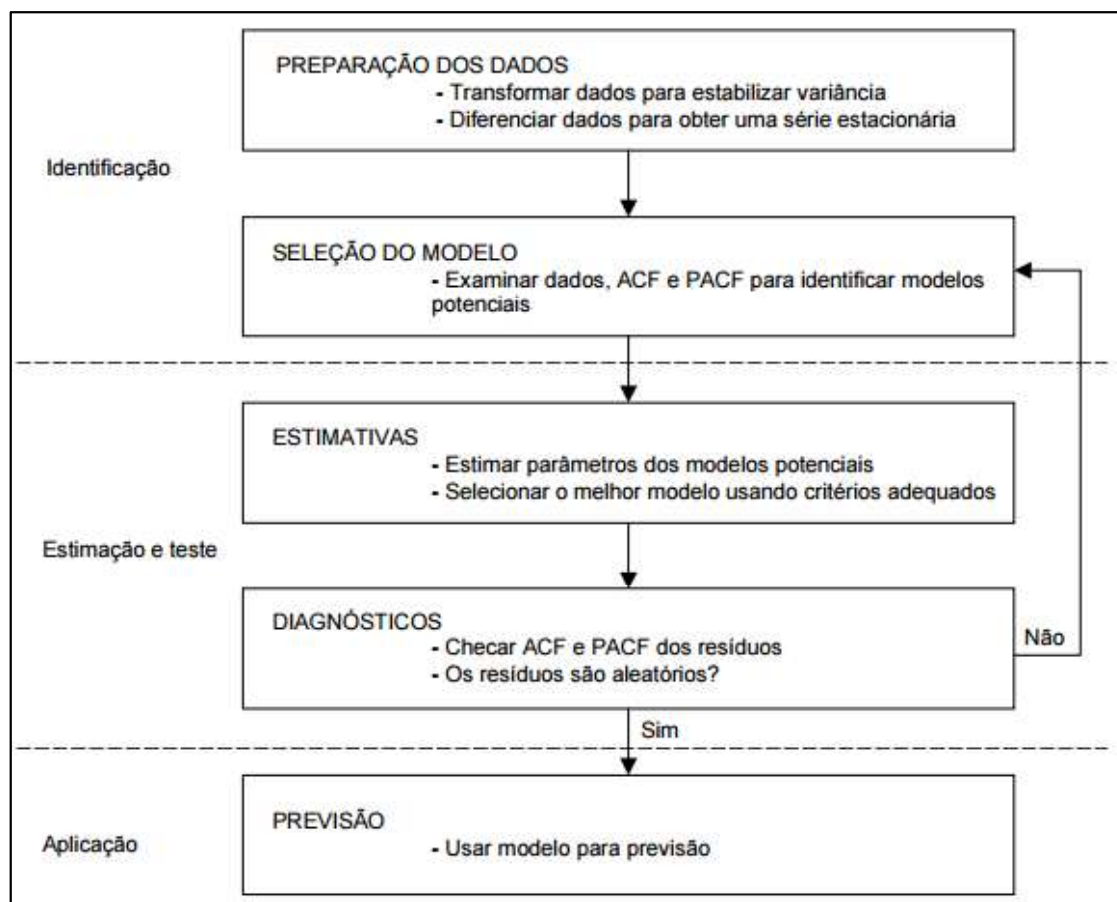
Bressan (2001), Werner (2004), Ramos, Santos e Rebelo (2015), Putro, Koshio e Oktaferdian (2016) e Sen, Roy e Pal (2016) explicam que são necessárias determinadas etapas para a construção do modelo Box-Jenkins, listadas a seguir e ilustradas na Figura 18:

- a) identificação: consiste em analisar qual modelo melhor descreve o comportamento da série, seja sazonal ou não. Pellegrini (2000) sugere que é necessário um mínimo de 50 observações para auxiliar na identificação do modelo. Ainda, o autor explica que as principais ferramentas utilizadas nesta etapa são a função de autocorrelação e a função de autocorrelação parcial, a fim de analisar seus correlogramas e determinar se a série é estacionária ou não. Caso seja uma não estacionária, então ela deve ser

estabilizada com a utilização da diferenciação, a fim de transformar em uma série estacionária, conhecendo, então, o valor de d . Para o próximo passo, Pellegrini (2000), Bressan (2001) e Sharafi et al. (2017) explicam que é necessário determinar os valores de p e q , também analisando os correlogramas. Se a ACF extingue-se rapidamente e a PACF trunca abruptamente após o primeiro *lag*, então $p = 1$, ou se o truncamento ocorre após o segundo *lag*, $p = 2$, conforme apresentado nas Figuras 8 e 9, respectivamente. Pellegrini (2000) e Sharafi et al. (2017) ainda comentam que, se a PACF tende a zero rapidamente e a ACF trunca subitamente após o primeiro *lag*, então $q = 1$, ou se o truncamento ocorre após o segundo *lag*, $q = 2$, conforme apresentado nas Figuras 10 e 11, respectivamente. Quando ACF e PACF extinguem-se rapidamente, é sugerido então a utilização de um modelo misto, e cita que são parâmetros são de difícil identificação;

- b) estimacão: nesta etapa, é necessário estimar os parâmetros γ e/ou Γ do componente auto regressivo (Fischer (1982) sugere a utilização de mínimos quadrados para a obtenção deste valor), e θ e/ou θ da média móvel (por não ser linear, Fischer (1982) propõem o uso do método da máxima verossimilhança), dependendo do modelo utilizado. Com estes parâmetros determinados, é possível calcular os erros aleatórios e_t e sua variância; e
- c) verificacão: avalia se o modelo estimado é apropriado para descrever o comportamento dos dados. Pellegrini (2000), Bressan (2001), Putro, Koshio e Oktaferdian (2016) e Sen, Roy e Pal (2016) ressaltam que, uma maneira de realizar a verificacão do modelo é analisar o comportamento da ACF dos resíduos calculados na fase de estimacão. Fischer (1982) explica que, essa análise é feita comparando a ACF dos resíduos calculados com a ACF dos resíduos de uma série de dados realmente observados; se as duas funções tiverem um comportamento semelhante, significa que o modelo foi bem identificado.

Figura 18 – Etapas de modelagem de Box-Jenkins



Fonte: Adaptado de Lemos (2006)

Caso o modelo não for adequado, este ciclo deve-se repetir, retornando a fase de identificação. Werner (2004) ainda comenta que, quando o modelo estimado for satisfatório, então é realizada uma última fase da metodologia Box-Jenkins: realizar a previsão.

Fischer (1982) e Putro, Koshio e Oktaferdian (2016) conceituam $\hat{Y}_t(h)$ como sendo a estimativa do valor futuro provável. Os autores comentam que, para a proposta de Box e Jenkins, é calculado a previsão do valor para um período h seguinte, ou seja, no tempo $t + h$. Então, posteriormente, é utilizado este resultado no cálculo da previsão subsequente.

2.3.4 Recapitulação da literatura

Para esta subseção, foram reapresentados aspectos importantes para o método Box-Jenkins, utilizando-se tabelas para possibilitar comparações e um melhor entendimento.

2.3.4.1 Agrupamento das equações

As equações de cada modelo Box-Jenkins estão exibidas na Tabela 3, sendo elas já apresentadas nas Equações 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36 e 37, respectivamente.

Tabela 3 – Equações do método Box-Jenkins

Função autocorrelação	População	$\rho_k = \frac{\sum (Y_t - \mu_t)(Y_{t-k} - \mu_{t-k})}{\sqrt{\text{Var}(Y_t) * \text{Var}(Y_{t-k})}} = \frac{\text{Cov}(Y_t, Y_{t-k})}{\sigma(Y_t) * \sigma(Y_{t-k})}$ $\rho_k = \frac{\text{Cov}(Y_t, Y_{t-k})}{\text{Var}(Y_t)} \rightarrow \text{para processo estacionário}$
	Amostral	$r_k = \frac{\sum_{t=k+1}^n (Y_t - \mu)(Y_{t-k} - \mu)}{\sum_{t=1}^n (Y_t - \mu)^2}$
Função autocorrelação parcial		$\phi_{kk} = \frac{r_k - \phi_1 \rho_{k-1} - \phi_2 \rho_{k-2} - \dots - \phi_{k-1} \rho_1}{1 - \phi_1 \rho_1 - \phi_2 \rho_2 - \dots - \phi_{k-1} \rho_{k-1}}$
Modelo AR(p)		$Y_t = c + \gamma_1 Y_{t-1} + \gamma_2 Y_{t-2} + \dots + \gamma_p Y_{t-p} + e_t$ $(1 - \gamma_1 B - \gamma_2 B^2 - \dots - \gamma_p B^p) Y_t = e_t$ $\gamma(B) Y_t = e_t$
Modelo MA(q)		$Y_t = \mu - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} + e_t$ $Y_t = \mu + (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) e_t$ $Y_t = \mu + \theta(B) e_t$
Modelo ARMA(p,q)		$Y_t = c + \gamma_1 Y_{t-1} + \dots + \gamma_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q}$ $\gamma(B) Y_t = c + \theta(B) e_t$
ARIMA(p,d,q)		$\gamma(B) (1 - B)^d Y_t = \theta(B) e_t$
ARIMA(P,Q,D)s		$\Gamma(B^S) (1 - B^S)^D Y_t = \Theta(B^S)$ $\gamma(B) \Gamma(B^S) (1 - B^S)^D Y_t = \theta(B) \Theta(B^S) e_t$

Fonte: Elaborado pela autora (2017)

2.3.4.2 Revisão bibliográfica – Box-Jenkins

Para esta subseção, foi realizada uma síntese das aplicações estudadas nos trabalhos dos autores citados na seção 2.3, apresentados no Apêndice D.

2.4 ANÁLISE ESTATÍSTICA

Para esta seção, foram estudados conceitos de estatística necessários para a análise dos resultados, ao fim deste trabalho.

2.4.1 Erros de previsão

Ballou (2001) comenta que, uma previsão de série temporal sempre implicará algum nível de erro, já que o futuro não é um reflexo exato do passado. Sendo assim, um erro na previsão representa quão próximo da realidade foi o resultado da prognose, ou seja, é a diferença entre o valor real Y_t e o valor previsto $\hat{Y}_t$, conceituado matematicamente na Equação 38.

$$MSE = \frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2 \quad (38)$$

Onde, MSE significa erro quadrático médio, do inglês *mean squared error* e n a quantidade total de dados. De acordo com Pellegrini (2000) e Ramos, Santos e Rebelo (2015), por se tratar do erro do valor estimado em comparação ao real, quanto menor o MSE, melhor.

2.4.2 Coeficiente de determinação

Conforme Freund (2006), Lapponi (2005) e Levine et al. (2012), o coeficiente de determinação, representado por r^2 , avalia a qualidade do ajuste do modelo, ou seja, determina o quanto o modelo foi capaz de explicar os dados coletados, por exemplo, para um r^2 igual a 0,877, significa que 87,7% da variabilidade dos dados é explicado pelo modelo. Os autores ainda apresentam a Equação 39.

$$r^2 = \frac{SQReg}{STQ} = \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (39)$$

Onde, $SQReg$ representa a soma dos quadrados da regressão, e STQ é a soma total dos quadrados. Ainda, $\hat{Y}$ é o valor previsto, $\bar{Y}$ é a média aritmética para o valor de Y , e Y_i é cada um dos valores observados de Y .

Lapponi (2005) também comenta que r^2 é sempre um número positivo, variando de 0 a 1, e que quanto maior for o r^2 , mais forte será a relação linear entre as duas variáveis sendo analisadas, por isso melhor será explicado a relação dos dados.

Montgomery e Runger (2016) explicam que o r^2 deve ser usado com cautela, porque é possível aumentar o seu valor se for adicionado termos suficientes, ou seja, é possível fazer um r^2 ser unitário se tiver n pontos com um polinômio de grau $n - 1$. Então, pode-se aumentar o valor do coeficiente de determinação ao passo que se aumenta as variáveis, mas isso não significa que o novo modelo seja melhor do que o antigo. Os autores também ressaltam que a dispersão da variável independente do problema é outro fator que impacta na magnitude de r^2 , ou seja, quanto maior a dispersão, geralmente maior será o valor de r^2 .

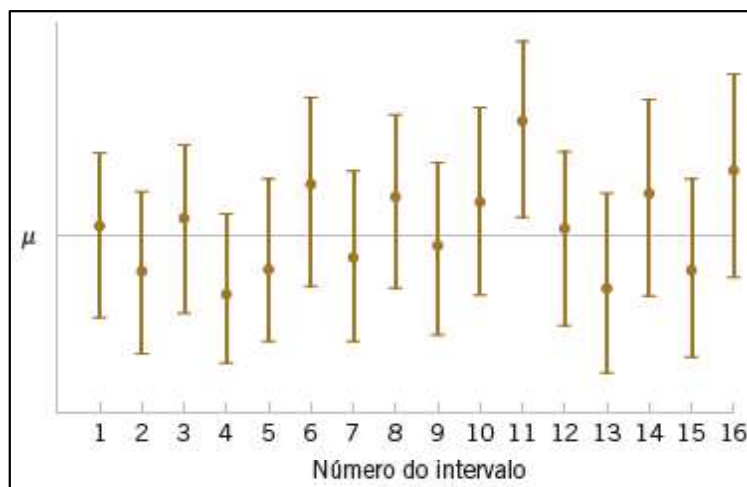
2.4.3 Intervalo de confiança

Lapponi (2005) e Levine et al. (2012) definem intervalo de confiança para a média como sendo uma extensão de valores, que é conhecida a probabilidade de que a média esteja entre este intervalo. Montgomery e Runger (2016) esclarecem essa definição ao comentar que, se for um intervalo de confiança de probabilidade de 95%, significa que 5% dos intervalos não englobam o valor da média μ , explicação ilustrada na Figura 19. Esta probabilidade é determinada pelo pesquisador, e não pelos dados.

Assim, Lapponi (2005), Levine et al. (2012) e Montgomery e Runger (2016) apresentam a Equação 40 para calcular os limites superior e inferior do intervalo de confiança.

$$\bar{X} = \mu + z \frac{\sigma}{\sqrt{n}} \quad (40)$$

Figura 19 – Construção repetida de um intervalo de confiança de μ



Fonte: Montgomery e Runger (2016, p. 217)

Na Equação 40, o $\bar{X}$ representa os limites superior e inferior, μ a média, z a probabilidade, σ o desvio padrão e n é o tamanho da amostra. Para a probabilidade, é necessário analisar a tabela de estatística de distribuição normal padronizada acumulada, geralmente é atribuída uma probabilidade de 95%, então o valor de z na equação será -1,96 para o limite inferior e 1,96 para o superior.

2.4.4 Análise de resíduos

De acordo com Lapponi (2005) e Levine et al. (2012), a análise de resíduo permite avaliar se o modelo utilizado é o mais apropriado. Montgomery e Runger (2016) também comentam que esta análise é usada para verificar a suposição de que erros sejam distribuídos de forma normal. Para isso, é necessário calcular o resíduo ou erro e_t , conforme Equação 41.

$$e_t = Y_t - \hat{Y}_t \quad (41)$$

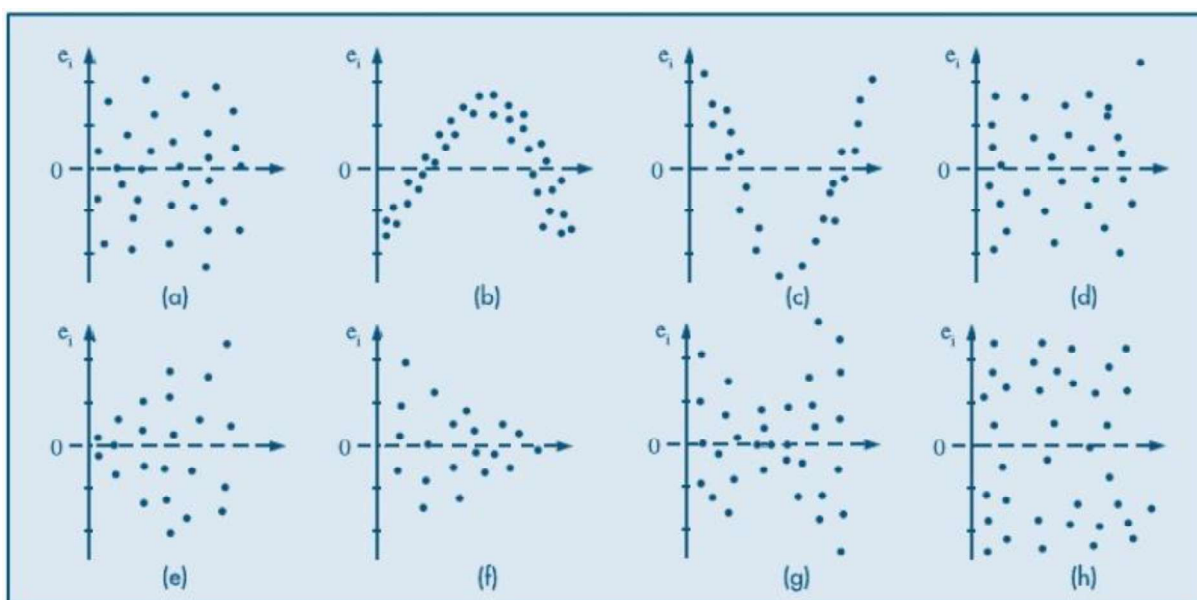
Onde, Y é o valor observado e $\hat{Y}$ é o valor previsto. Ainda, Montgomery e Runger (2016) apresentam a Equação 42, que é para padronizar os resíduos.

$$d_t = e_t / \sqrt{\sigma^2} \quad (42)$$

Onde, d_t é o resíduo padronizado e σ é o desvio padrão. Os autores comentam que os resíduos padronizados variam entre -2 e +2, caso os erros forem distribuídos normalmente. Se houver algum resíduo fora deste intervalo, pode indicar a presença de um *outlier*, ou seja, este dado pode ter sofrido uma influência externa e, depois de uma análise da causa, pode ser descartado.

Com a elaboração de um gráfico com os dados dos erros, é possível analisar visualmente o seu comportamento e obter conclusões. Bussab e Morettin (2013) apresentam, na Figura 20, os tipos usuais de gráficos de resíduos.

Figura 20 – Gráficos de resíduos



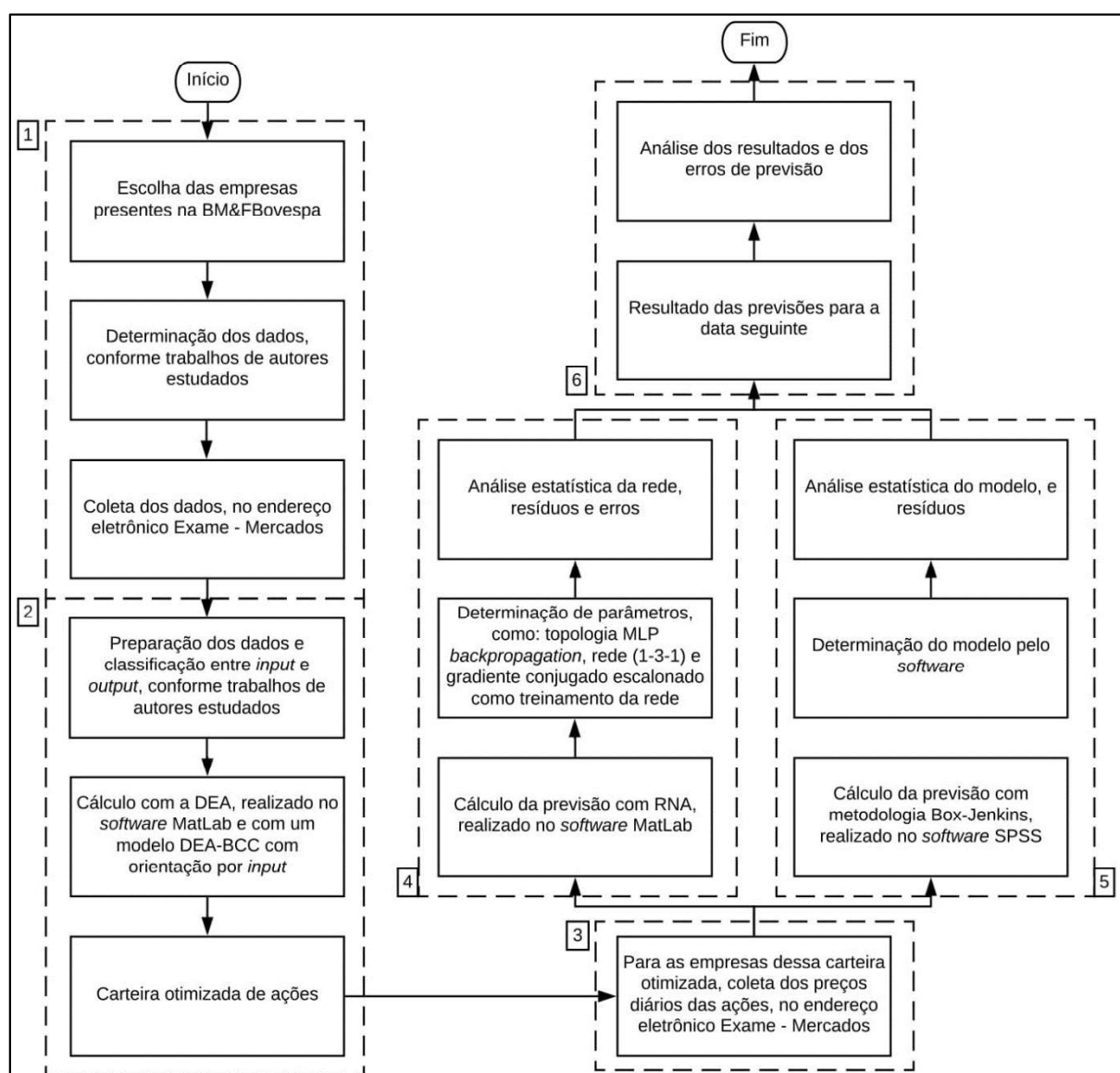
Fonte: Adaptado de Bussab e Morettin (2013)

A Figura 20 (a) representa a situação ideal para os resíduos, distribuído aleatoriamente em torno do zero. Bussab e Morettin (2013) explicam que as curvaturas dos gráficos (b) e (c) representam um modelo não linear, sendo necessário procurar outras funções matemáticas para explicar o modelo. A situação (d) mostra um ponto discrepante e deve ser melhor analisado este desvio. Os autores comentam que pode ser um erro de medida ou é uma discrepância real. Gráficos residuais que parecem com o (e), (f) e (g) indicam desigualdade de variância, e enquanto o (h) possui maior incidência de pontos nos extremos, mostrando haver problema com a normalidade, isto é, os dados podem não ser de uma distribuição normal.

3 MÉTODOS DA DISSERTAÇÃO

Com o intuito de atingir o objetivo geral proposto neste trabalho, o método elaborado apresenta-se ilustrada na Figura 21, sendo que as etapas em tracejado representam cada um dos seis objetivos específicos. Dentro destas etapas em tracejado, há fases que necessitam serem realizadas para que a etapa fique concluída.

Figura 21 – Etapas propostas



Fonte: Elaborado pela autora (2017)

De acordo com a Figura 21, o trabalho iniciou-se com a escolha das empresas que participaram do estudo e a determinação de quais dados foram necessários,

conforme autores com trabalhos similares, para em seguida, realizar a coleta destes dados. A segunda etapa consistiu inicialmente na preparação dos dados para poder calcular com a DEA, assim definindo quais empresas foram consideradas eficientes e, portanto, pertenceram a carteira otimizada de ações. Para estas empresas, então foram coletados os preços diários de suas ações, representando a terceira etapa. A quarta e quinta etapas representam os cálculos de previsão do preço destas ações, e ocorreram simultaneamente. Os cálculos foram realizados com dois métodos de previsão: Redes Neurais Artificiais e metodologia Box-Jenkins, tendo o primeiro ocorrido no *software* MatLab e o segundo, no SPSS. Para finalizar este estudo, realizou-se a sexta etapa, que consiste na análise dos resultados obtidos com os métodos, comparando-os com os preços reais e fazendo um estudo acerca dos erros percentuais.

Cada uma destas seis etapas citadas foi detalhada nas subseções seguintes.

3.1 ESCOLHA DAS EMPRESAS E COLETA DE SEUS DADOS

A Bolsa de Valores de São Paulo (Bovespa) era a bolsa oficial do Brasil, até 2008, quando ocorreu um processo de fusão e ela passou a se chamar Bolsa de Valores, Mercadorias e Futuros de São Paulo (BM&FBovespa). Atualmente, há centenas empresas que negociam suas ações na BM&FBovespa. Este trabalho não tem o objetivo de analisar todas elas, por isso foi utilizado o trabalho de Rotela Junior, Pamplona e Salomon (2014) como base e assim, foram selecionadas as cinquenta empresas com os maiores valores de mercado, sejam do tipo de ações ON (ordinária) ou PN (preferencial), ou seja, respectivamente, ativos que dão direito a voto na empresa e ativos que tem preferência na distribuição de dividendos. Esta listagem de empresas foi obtida do endereço eletrônico da bolsa de valores.

A coleta dos dados foi embasada nos trabalhos de Rotela Junior, Pamplona e Salomon (2014) e de Lopes et al. (2008), e assim, foram necessários os seguintes dados:

- a) preço/lucro (P/L);
- b) volatilidade anualizada (V);
- c) lucro por ação (LPA); e
- d) retorno sobre investimento (R).

Preço/Lucro pode ser definido como sendo o índice que mostra quantas vezes o preço é maior do que o lucro, ou seja, de acordo com Hazzan (1991), indica quantos períodos são necessários para recuperar o preço da ação em forma de lucro. Volatilidade é a taxa de variação de um ativo em um determinado período. Ela é considerada uma medida de risco, pois ela mede a dispersão dos retornos, ou seja, o quanto mais varia o preço de uma ação em um determinado tempo, maior o risco de se ter lucro ou prejuízo com esta ação.

De forma simplificada, lucro por ação (LPA) é a divisão do lucro líquido pelo número de ações em circulação, ambos para o mesmo determinado período de tempo. Ainda, Campos e Scherer (2001) consideram como sendo um quociente que mede a rentabilidade e que indica quão lucrativo foi um empreendimento pela utilização dos recursos disponibilizados pelos acionistas. Já a rentabilidade indica o percentual de retorno sobre o investimento, sendo um importante indicador de desempenho. O seu resultado foi obtido pela divisão do lucro operacional (que é o somatório do lucro líquido com despesas com juros e depreciação) da empresa em um determinado período pelo somatório do patrimônio líquido com o endividamento total.

Todos os dados coletados foram referentes ao último ano fiscal. Os valores de volatilidade de 12, 24 e 36 meses foram calculados de acordo com a Equação 43, conforme apresentada por Malacrida e Yamamoto (2006).

$$VH = \sqrt{252} \sigma \left[\ln \left(\frac{P_t}{P_{t-1}} \right) \right] \quad (43)$$

Onde, VH representa a volatilidade histórica, σ o desvio padrão da amostra, P_t o preço do período t e P_{t-1} como sendo o preço do período $t - 1$.

O intervalo de tempo escolhido para este trabalho foi de 1, 2 e 3 anos, e foi necessário multiplicar o desvio padrão calculado pela raiz quadrada de 252 para a volatilidade ser anualizada, sendo que o valor 252 representa o total de dias úteis no ano. Sendo assim, as informações de preço diário para os 3 anos de análise foram retiradas do endereço eletrônico Exame – Mercados.

Para Preço/Lucro e lucro por ação, os dados foram coletados do endereço eletrônico Exame, na seção de Mercados – Cotações da Bovespa. Já o retorno de investimento de 12, 24 e 36 meses, foi calculado utilizando os dados de lucro

operacional, patrimônio líquido e endividamento total para os respectivos anos. Estas informações também foram extraídas do endereço eletrônico Exame – Mercados.

O resultado desta etapa foi a obtenção da listagem das empresas participantes deste trabalho e também a coleta de seus respectivos dados.

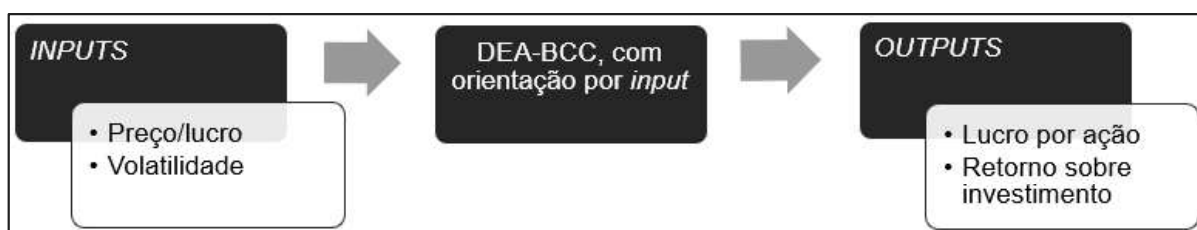
3.2 CÁLCULO DA EFICIÊNCIA COM A UTILIZAÇÃO DA DEA

Na segunda etapa, para o cálculo da Análise Envoltória de Dados, foi utilizado o MatLab, usufruindo de uma programação já existente neste *software*. O propósito deste trabalho foi a análise dos resultados, por isso se optou por realizar os procedimentos matemáticos no *software*, além dos ganhos na agilidade, há uma maior precisão dos resultados.

Este estudo utilizou o modelo DEA-BCC, pois, caso ocorra alguma variação nos dados de entrada, não necessariamente ocorrerá uma variação proporcional nos dados de saída. Ainda, a orientação utilizada foi por *input*, pois se deseja minimizar os insumos (*input*) para uma determinada quantidade de produto (*output*). Para ε , foi considerado um valor igual a 0,001.

Para o uso da DEA, foi necessário classificar as informações como *input* e *output*. Então, os quatro dados explicados anteriormente, preço/lucro, volatilidade, lucro por ação e retorno sobre investimento, ficaram categorizadas de acordo com a Figura 22, ainda seguindo os trabalhos de Rotela Junior, Pamplona e Salomon (2014) e de Lopes et al. (2008) como base.

Figura 22 – Representação do modelo DEA



Fonte: Elaborado pela autora (2017)

Para os dados utilizados na DEA, deve-se ter o cuidado com a presença de valores negativos, pois não são permitidos nos cálculos, conforme explicado no Capítulo 2.1. Essa proibição é por causa das restrições impostas nos modelos matemáticos, como definido por Charnes, Cooper e Rhodes (1978). Em uma

eventualidade de se ter um valor negativo, Martins (2014) sugere excluir a respectiva DMU da análise.

Após a realização dos cálculos, foi obtida uma lista de empresas que tiveram 100% de eficiência, analisando as informações citadas anteriormente. Não se pode afirmar que o resultado foi um ranking com as melhores empresas quando comparado a todas as demais, mas sim, um agrupamento destas de forma a maximizar a rentabilidade. Além disso, dentre as empresas desta lista de empresas com 100% de eficiência, não se pode considerar se há alguma melhor que a outra, ou seja, elas são todas tão eficientes quanto as outras, isto ocorre devido a metodologia do DEA. Esta classificação de eficiência depende do desempenho de transformar insumos em produtos, comparando com as demais unidades, ou seja, as que obtêm uma maior propensão de retorno de lucro. Para este trabalho, uma empresa com 100% de eficiência significa que obteve a melhor transformação de preço/lucro e volatilidade em lucro por ação e retorno sobre investimento.

Para esta seção, o produto final foi a carteira otimizada de ações das empresas consideradas eficientes, calculando com a DEA.

3.3 DETERMINAÇÃO DOS DADOS PARA A ANÁLISE DE PREVISÃO

Para esta terceira etapa, foram necessários outros dados para dar continuidade no estudo. Assim, foi preciso coletar o preço diário de cada ação das empresas consideradas como eficientes pela DEA. Em seguida, foi exigida uma preparação destes dados para utilizar a Rede Neural e a metodologia Box-Jenkins, pois eles devem estar em ordem crescente de suas datas. A organização dos valores foi feita no *software* Excel.

Os dados foram coletados no endereço eletrônico Exame – Mercados, e o intervalo de tempo foi dia a dia comercial, isto é, quando a BM&FBovespa trabalha; e foram eles: preço da ação, o seu respectivo dia e a variação percentual em comparação ao dia anterior.

Puderam ser extraídos valores históricos de até 10 anos, e por isso, foram realizadas dez prognoses para cada empresa, que ocorreu da seguinte forma:

- a) o equivalente a 12 meses de dados entre os anos de 2017 e 2016;
- b) o equivalente a 24 meses de dados entre os anos de 2017, 2016 e 2015;

- c) o equivalente a 36 meses de dados entre os anos de 2017, 2016, 2015 e 2014;
- d) o equivalente a 48 meses de dados entre os anos de 2017 a 2013;
- e) o equivalente a 60 meses de dados entre os anos de 2017 a 2012;
- f) o equivalente a 72 meses de dados entre os anos de 2017 a 2011;
- g) o equivalente a 84 meses de dados entre os anos de 2017 a 2010;
- h) o equivalente a 96 meses de dados entre os anos de 2017 a 2009;
- i) o equivalente a 108 meses de dados entre os anos de 2017 a 2008; e
- j) o equivalente a 120 meses de dados entre os anos de 2017 a 2007.

Assim, com estes dez cálculos para cada empresa da carteira otimizada, foi possível determinar qual a quantidade de dados históricos mais indicadas para cada método de previsão utilizado. Não há muitos trabalhos publicados que tratem desta análise, sendo então um diferencial deste estudo.

Para esta terceira etapa, o resultado foi a coleta dos preços diários das empresas participantes da carteira otimizada.

3.4 UTILIZAÇÃO DAS REDES NEURAIIS ARTIFICIAIS

Esta seção descreve o modelo implementado para a análise de previsão de preço de ações. De acordo com os conceitos abordados no Capítulo 2.2, as RNA possuem parâmetros que foram necessários especificar, conforme detalhado a seguir:

- a) topologia de rede;
- b) processos de aprendizagem;
- c) algoritmo de aprendizagem;
- d) quantidade de camadas escondidas, caso houver;
- e) treinamento da rede; e
- f) quantidade percentual de dados que foram utilizados para treinamento, validação e teste da rede.

Para esta etapa, foi utilizada a RNA para prever o valor da data seguinte dos dados previamente coletados, para após realizar o comparativo dos resultados com os obtidos com a metodologia Box-Jenkins. Esta prognose ocorreu no MatLab, devido à complexidade do cálculo e a programação das redes já existente no *software*. Como

foram feitas análises diárias em períodos de, no mínimo, um ano, a quantidade de dados a ser utilizada foi suficientemente grande para a rede entender o comportamento da série.

Para este trabalho, foi utilizada a topologia *multilayer perceptrons*, por ter uma melhor capacidade de aprendizagem, já que este estudo possui uma grande base de dados, seguindo os trabalhos de diversos autores estudados no Capítulo 2.2 como base. Como não foram fornecidos os dados de saída, o processo de aprendizagem foi não supervisionado, fazendo com que a rede agrupe os valores de maior semelhança.

Para as camadas da rede, foi definida a utilização de uma camada para a de entrada e a de saída, já que foi fornecida uma variável de entrada (preço diário) e se obterá uma variável de resposta (preço da data seguinte). Para determinar a quantidade de camadas ocultas, foi seguido o trabalho de Devadoss e Ligori (2013), em que os autores consideraram uma proporção matemática de $2(\text{quantidade de camada de entrada}) + 1$, como quantificação das camadas ocultas. Assim, foi utilizada uma rede (1-3-1).

Para o algoritmo de aprendizagem, foi escolhido o *backpropagation*, baseando-se nos trabalhos dos autores estudados com aplicação similar. Para o treinamento da rede, foi utilizado o gradiente conjugado escalonado, que é um algoritmo de otimização em que busca a minimização do erro dos pesos w . Ainda, foi escolhido dividir os dados da seguinte maneira: 70% dos valores foram usados para treinamento, 15% para a validação e os restantes 15% para testes da rede, e classificando-os de forma arbitrária.

Após a realização dos cálculos, foi avaliada a performance da rede neural, analisando o erro quadrático médio e o coeficiente de determinação de cada previsão. O propósito foi analisar estatisticamente o resultado do cálculo e se os parâmetros escolhidos foram adequados.

Assim, o produto desta etapa foi os resultados das previsões com diferentes quantidades de dados históricos, análise de resíduos, e informações como MSE e coeficiente de determinação para cada uma das prognoses. Ainda, para este método, foi possível analisar a classificação dos dados em treino, validação e teste, e o histograma dos erros separados nesta classificação.

3.5 APLICAÇÃO DA METODOLOGIA DE BOX-JENKINS

Conforme apresentado no Capítulo 2.3, existem três etapas necessárias para a utilização da metodologia Box-Jenkins: o primeiro passo para determinar qual modelo será utilizado é calcular os coeficientes de autocorrelação e autocorrelação parcial, para analisar seus correlogramas, sendo esta a etapa da identificação. É desta forma que se define se é um processo estacionário ou não estacionário, conseqüentemente, observando a necessidade de diferenciar a série temporal e assim, determinar o valor de d . Ainda com os correlogramas, determinam-se os valores de p e q .

A etapa seguinte é a da estimação, em que são calculados parâmetros ϕ e/ou Φ do componente auto regressivo e θ e/ou Θ da média móvel, além dos erros aleatórios. Para a terceira e última fase, a da verificação, é avaliado se o modelo escolhido foi apropriado, desta vez, observando os correlogramas dos resíduos.

Esta quinta etapa constitui-se na aplicação deste método, para em seguida, comparar com os resultados obtidos com a RNA. Foi utilizado o *software* SPSS para a realização dos cálculos e a definição dos parâmetros, a fim de obter a maior assertividade possível no modelo escolhido. Como foram feitas análises diárias em períodos de, no mínimo, um ano, a quantidade de dados a ser utilizada satisfaz uma exigência deste *software*, que é uma quantidade mínima de 48 dados para poder garantir o cálculo da previsão.

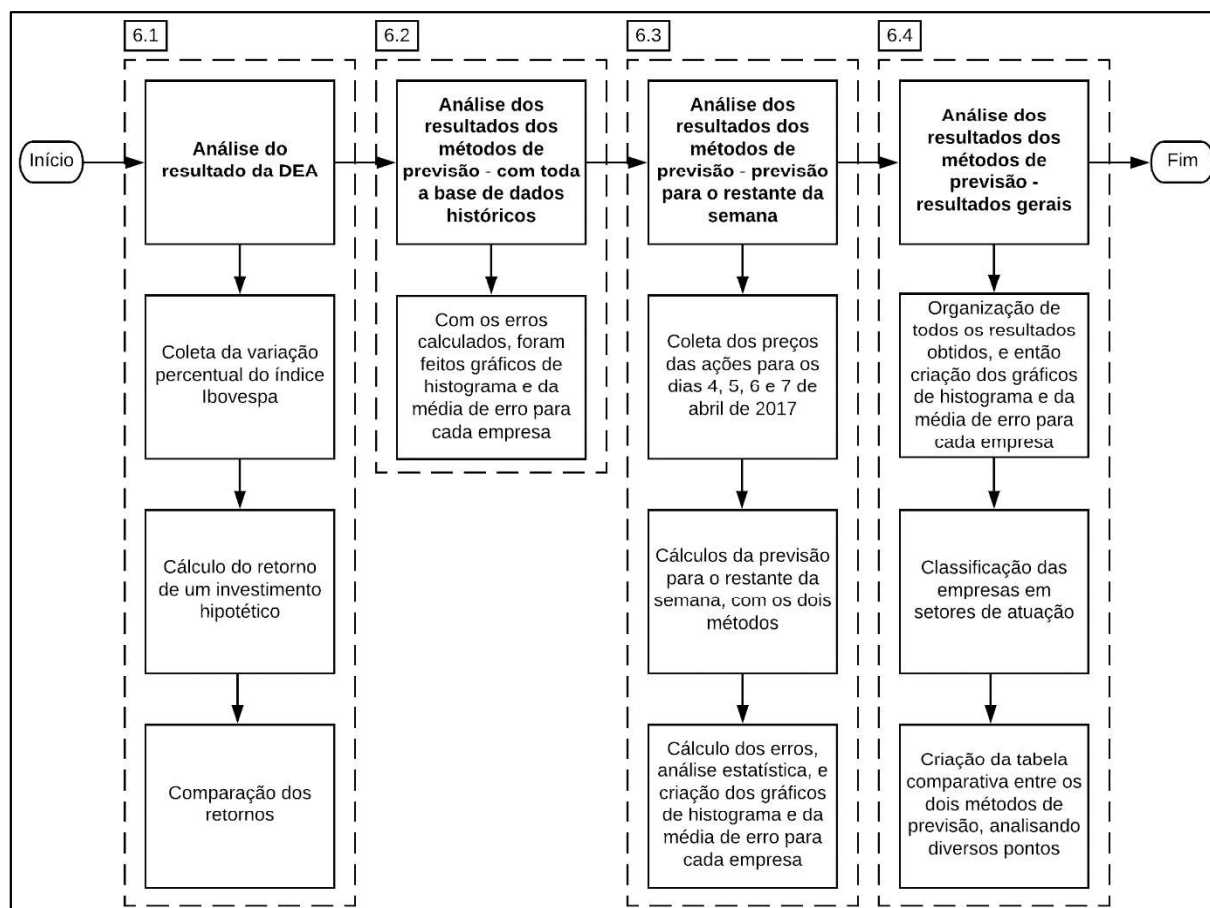
Após a realização dos cálculos, foi avaliada a performance do método, analisando o erro quadrático médio e o coeficiente de determinação de cada previsão. O propósito foi analisar estatisticamente o resultado do cálculo e se os parâmetros determinados foram adequados.

Assim, o produto desta etapa foram os resultados das previsões com diferentes quantidades de dados históricos, e informações como MSE e coeficiente de determinação para cada uma das prognoses. Ainda, para este método, foi possível analisar os correlogramas das funções de autocorrelação e autocorrelação parcial dos dados históricos e também os dos resíduos.

3.6 AVALIAÇÃO DOS RESULTADOS

A avaliação dos resultados obtidos pode ser dividida em quatro etapas, conforme ilustrado na Figura 23.

Figura 23 – Etapas da avaliação dos resultados



Fonte: Elaborado pela autora (2018)

Então, esta subseção foi separada em duas partes: a análise da carteira otimizada com o uso da DEA, e a avaliação das previsões calculadas com RNA e Box-Jenkins. Assim, obtiveram dois resultados para esta última parte do trabalho: o comparativo de retorno da carteira otimizada com o índice Ibovespa para a primeira parte; e o comparativo do preço estimado pelos dois métodos com o real, além da análise dos erros. Para esta análise, foram feitos gráficos de histograma e também da média dos erros de cada empresa, a fim de auxiliar na determinação de qual método de previsão de séries temporais foi o mais adequado. As metodologias destas análises serão explicadas nas subseções a seguir.

3.6.1 Avaliação da DEA

Para avaliar o benefício em utilizar Análise Envoltória de Dados, foi feito um comparativo de retorno da carteira otimizada e do índice Ibovespa. De acordo com o endereço eletrônico da BM&FBovespa, a Ibovespa é uma carteira teórica de empresas, que serve para indicar o desempenho médio das cotações dos ativos de maior negociabilidade e representatividade do mercado de ações brasileiro. Para este comparativo, foi utilizada a variação os preços diários das ações com o equivalente a 12 meses de dados entre 2016 e 2017, e realizando um investimento hipotético de R\$1000,00. Para a carteira otimizada, este valor foi dividido por igual entre todas as ações, e para a Ibovespa, foi aplicado este valor por inteiro.

Esta quantia para este investimento fictício foi escolhida arbitrariamente, para poder comparar o retorno da Ibovespa com o da carteira otimizada. A variação do preço das ações foi coletada na terceira etapa deste estudo. Já para a variação do índice Ibovespa, foi obtida também do endereço eletrônico Exame – Mercados, com a mesma data dos preços das ações, e o intervalo de tempo também foi dia a dia comercial.

3.6.2 Avaliação dos métodos de previsão

Após a realização das duas últimas propostas, foi calculado o erro percentual dos resultados de cada modelo de previsão. Assim, foi possível determinar qual o método mais indicado e qual a melhor quantidade de dados históricos para calcular. Este erro, apresentado na Equação 44, foi determinado com os valores reais e os previstos, para cada uma das dez prognoses realizadas, conforme descrito na terceira proposta.

$$Erro \% = \frac{Valor\ previsto - Valor\ real}{Valor\ real} \quad (44)$$

Para o cálculo desta diferença de valores, foi utilizada a previsão calculada e o preço do dia seguinte da base de dados, que também foi coletado no endereço eletrônico Exame – Mercados. Os cálculos foram realizados no *software* Excel, assim

como a elaboração dos gráficos, a fim de facilitar a visualização dos dados observados na prática e os estimados.

Com o propósito de a reprodutibilidade dos resultados obtidos e garantir que estes métodos possuem uma precisão satisfatória, foi calculada a previsão para o restante da semana em análise. Para estes cálculos, foi utilizada uma base de dados históricos com o equivalente a 12 meses, e os parâmetros dos dois métodos permaneceram os mesmos. Desta análise complementar, também se obteve o histograma dos erros e o gráfico da média de erros para cada empresa.

Para finalizar o estudo com os métodos de previsão, foi observado o comportamento de todos os 194 resultados calculados de cada método, ou seja, as 130 prognoses com diferentes quantidades de dados históricos e mais as 64 previsões do restante da semana, que são referentes as 16 empresas para 4 dias; totalizando 388 cálculos de previsão. Com isto, foi possível fazer um comparativo dos dois métodos, utilizando os seguintes dados: erro quadrático médio, coeficiente de determinação e diversas análises de erro; a fim de determinar qual o método de previsão mais adequado.

4 APLICAÇÃO DOS MÉTODOS

Neste capítulo, encontram-se todos os resultados do método apresentado no Capítulo 3, utilizando a fundamentação teórica estudada no Capítulo 2. Nesta etapa, foram demonstrados os cálculos realizados, assim como as considerações e as conclusões observadas.

4.1 ESCOLHA DAS EMPRESAS E COLETA DE SEUS DADOS

Conforme apresentado no capítulo anterior, foram listadas as 50 empresas com os maiores valores de mercado, de acordo com o endereço eletrônico da BM&FBovespa. Esta coleta foi realizada no dia 14 de junho de 2016, e as informações são com base no histórico mensal. A listagem das empresas com seus respectivos códigos foi apresentada na Tabela 4.

Tabela 4 – Lista das empresas escolhidas

(continua)

	Empresa	Código		Empresa	Código
1	AMBEV	ABEV3	26	SABESP	SBSP3
2	ITAU UNIBANCO	ITUB4	27	TIM	TIMP3
3	BANCO BRADESCO	BBDC4	28	SUZANO PAPEL E CELULOSE	SUZB5
4	PETROBRAS	PETR4	29	BRASKEM	BRKM5
5	CIELO	CIEL3	30	EMBRAER	EMBR3
6	VALE	VALE5	31	LOJAS RENNER	LREN3
7	TELEFÔNICA BRASIL	VIVT4	32	COSAN	CSAN3
8	BANCO SANTANDER	SANB4	33	ELETROBRAS	ELET6
9	BB SEGURIDADE	BBSE3	34	GPA	PCAR4
10	ITAUSA	ITSA4	35	CETIP	CTIP3
11	BANCO DO BRASIL	BBAS3	36	M. DIAS BRANCO	MDIA3
12	ULTRAPAR	UGPA3	37	MULTIPLAN	MULT3
13	BRF	BRFS3	38	NATURA	NATU3
14	BM&FBOVESPA	BVMF3	39	TRANSMISSÃO PAULISTA	TRPL4
15	JBS	JBSS3	40	CIA SIDERÚRGICA NACIONAL	CSNA3
16	GRUPO CCR	CCRO3	41	EQUATORIAL ENERGIA	EQTL3
17	WEG	WEGE3	42	GERDAU	GGBR4

(conclusão)					
	Empresa	Código		Empresa	Código
18	ENGIE BRASIL ENERGIA	EGIE3	43	PORTO SEGURO	PSSA3
19	KLABIN	KLBN4	44	EDP ENERGIAS DO BRASIL	ENBR3
20	LOJAS AMERICANAS	LAME4	45	CIA ENERGÉTICA DE MINAS GERAIS	CMIG4
21	RAIADROGASIL	RADL3	46	LOCALIZA	RENT3
22	CPFL ENERGIA	CPFE3	47	BRMALLS	BRML3
23	FIBRIA	FIBR3	48	TAESA	TAAE11
24	HYPERMARCAS	HYPE3	49	RUMO LOG	RUMO3
25	KROTON EDUCACIONAL	KROT3	50	SÃO MARTINHO	SMT03

Fonte: Elaborado pela autora (2016)

Para o cálculo da DEA, foram necessários dados de *input* e de *output*. Assim, foram utilizadas as seguintes informações de *input*: preço/lucro (P/L) e volatilidade (V) de 1, 2 e 3 anos, disponibilizados na Tabela 5. Já os dados de *output*, que foram lucro por ação (LPA) e retorno sobre investimentos (R) de 1, 2 e 3 anos, foram indicados na Tabela 6. Todas estas informações foram coletadas durante a noite de 15 de junho de 2016, momento em que a bolsa de valores não estava mais trabalhando, garantindo assim a obtenção do valor de fechamento do dia. Esta data para reunir os dados das amostras foi escolhida arbitrariamente.

Tabela 5 – Dados de *input* das empresas

(continua)					
Empresa	Código	P/L	V1	V2	V3
AMBEV	ABEV3	R\$ 24,06	20,42%	24,87%	23,22%
ITAU UNIBANCO	ITUB4	R\$ 7,16	32,56%	35,12%	27,28%
BANCO BRADESCO	BBDC4	R\$ 7,36	34,89%	41,27%	25,81%
PETROBRAS	PETR4	-R\$ 2,92	58,91%	55,07%	36,09%
CIELO	CIEL3	R\$ 20,37	35,15%	32,91%	73,22%
VALE	VALE5	-R\$ 1,92	50,21%	31,08%	31,01%
TELEFÔNICA BRASIL	VIVT4	R\$ 18,12	30,00%	27,11%	22,11%
BANCO SANTANDER	SANB4	R\$ 6,64	47,73%	67,01%	402,97%
BB SEGURIDADE	BBSE3	R\$ 12,89	37,71%	26,87%	24,09%
ITAUSA	ITSA4	R\$ 6,11	32,08%	34,82%	25,64%
BANCO DO BRASIL	BBAS3	R\$ 3,05	48,09%	49,20%	27,21%

(continuação)

Empresa	Código	P/L	V1	V2	V3
ULTRAPAR	UGPA3	R\$ 25,27	22,67%	26,90%	20,85%
BRF	BRFS3	R\$ 12,78	26,85%	24,93%	22,86%
BM&FBOVESPA	BVMF3	R\$ 13,12	33,75%	39,58%	30,72%
JBS	JBSS3	R\$ 6,38	38,41%	43,45%	41,30%
GRUPO CCR	CCRO3	R\$ 28,89	39,22%	35,14%	24,36%
WEG	WEGE3	R\$ 20,79	27,46%	71,08%	35,14%
ENGIE BRASIL ENERGIA	EGIE3	R\$ 15,61	28,59%	28,27%	21,86%
KLABIN	KLBN4	R\$ 4,22	33,25%	26,58%	164,01%
LOJAS AMERICANAS	LAME4	R\$ 87,06	39,33%	24,90%	33,55%
RAIADROGASIL	RADL3	R\$ 56,77	32,55%	30,16%	27,81%
CPFL ENERGIA	CPFE3	R\$ 18,19	36,86%	35,30%	23,14%
FIBRIA	FIBR3	R\$ 133,65	34,38%	31,87%	32,86%
HYPERMARCAS	HYPE3	R\$ 48,65	34,71%	25,18%	23,78%
KROTON EDUCACIONAL	KROT3	R\$ 13,00	60,20%	32,56%	146,89%
SABESP	SBSP3	R\$ 32,17	41,03%	35,15%	34,42%
TIM	TIMP3	R\$ 8,19	36,96%	39,37%	30,19%
SUZANO PAPEL E CELULOSE	SUZB5	R\$ 12,24	35,80%	30,85%	31,29%
BRASKEM	BRKM5	R\$ 5,39	56,89%	26,73%	31,50%
EMBRAER	EMBR3	R\$ 56,93	29,18%	25,57%	32,58%
LOJAS RENNER	LREN3	R\$ 17,36	32,20%	167,62%	27,75%
COSAN	CSAN3	R\$ 25,08	38,34%	34,36%	26,82%
ELETROBRAS	ELET6	-R\$ 1,38	44,87%	43,59%	44,75%
GPA	PCAR4	R\$ 17,85	34,16%	22,53%	23,47%
CETIP	CTIP3	R\$ 22,18	25,26%	19,61%	22,92%
M. DIAS BRANCO	MDIA3	R\$ 17,15	31,62%	24,25%	25,23%
MULTIPLAN	MULT3	R\$ 28,22	33,14%	28,07%	20,84%
NATURA	NATU3	R\$ 20,98	40,82%	30,87%	27,90%
TRANSMISSÃO PAULISTA	TRPL4	R\$ 19,61	30,02%	26,50%	28,49%
CIA SIDERÚRGICA NACIONAL	CSNA3	R\$ 4,91	74,57%	47,75%	40,67%
EQUATORIAL ENERGIA	EQTL3	R\$ 14,91	27,69%	23,29%	22,16%
GERDAU	GGBR4	-R\$ 2,19	51,07%	32,90%	31,39%
PORTO SEGURO	PSSA3	R\$ 8,34	34,97%	27,62%	27,88%
EDP ENERGIAS DO BRASIL	ENBR3	R\$ 4,69	36,52%	36,47%	30,65%
CIA ENERGÉTICA DE MINAS GERAIS	CMIG4	R\$ 2,80	43,77%	39,30%	42,10%
LOCALIZA	RENT3	R\$ 17,24	35,90%	28,83%	22,96%

(conclusão)					
Empresa	Código	P/L	V1	V2	V3
BRMALLS	BRML3	R\$ 1.433,00	46,46%	39,01%	29,43%
TAESA	TAE11	R\$ 7,15	24,95%	27,79%	22,68%
RUMO LOG	RUMO3	-R\$ 21,05	67,44%	46,41%	29,81%
SÃO MARTINHO	SMT03	R\$ 27,71	34,81%	30,74%	30,65%

Fonte: Elaborada pela autora (2016)

Sendo V1, V2 e V3 representando a volatilidade de 1, 2 e 3 anos, respectivamente.

Tabela 6 – Dados de *output* das empresas

(continua)					
Empresa	Código	LPA	R1	R2	R3
AMBEV	ABEV3	R\$ 0,80	36,17%	35,33%	33,64%
ITAU UNIBANCO	ITUB4	R\$ 4,12	2,41%	4,45%	3,54%
BANCO BRADESCO	BBDC4	R\$ 3,15	1,72%	4,24%	3,51%
PETROBRAS	PETR4	-R\$ 0,02	-1,78%	-3,32%	5,61%
CIELO	CIEL3	R\$ 1,57	26,21%	50,07%	57,52%
VALE	VALE5	-R\$ 4,00	-11,06%	7,88%	15,39%
TELEFÔNICA BRASIL	VIVT4	R\$ 2,30	6,65%	9,70%	9,44%
BANCO SANTANDER	SANB4	R\$ 2,66	-1,05%	2,48%	1,76%
BB SEGURIDADE	BBSE3	R\$ 2,10	66,35%	51,53%	42,26%
ITAUSA	ITSA4	R\$ 1,19	18,91%	19,70%	17,64%
BANCO DO BRASIL	BBAS3	R\$ 5,43	1,95%	2,86%	5,30%
ULTRAPAR	UGPA3	R\$ 2,71	17,52%	14,12%	15,86%
BRF	BRFS3	R\$ 3,63	16,22%	14,23%	8,56%
BM&FBOVESPA	BVMF3	R\$ 1,22	11,05%	6,97%	7,26%
JBS	JBSS3	R\$ 1,62	9,81%	12,24%	7,59%
GRUPO CCR	CCRO3	R\$ 0,54	15,23%	20,57%	23,29%
WEG	WEGE3	R\$ 0,72	10,36%	11,98%	13,03%
ENGIE BRASIL ENERGIA	EGIE3	R\$ 2,30	23,00%	23,73%	26,95%
KLABIN	KLBN4	-R\$ 1,70	6,38%	9,42%	9,06%
LOJAS AMERICANAS	LAME4	R\$ 0,17	15,47%	17,34%	16,36%
RAIADROGASIL	RADL3	R\$ 1,05	16,84%	11,15%	5,78%
CPFL ENERGIA	CPFE3	R\$ 1,02	8,46%	10,07%	10,43%
FIBRIA	FIBR3	R\$ 0,25	13,87%	7,24%	7,06%
HYPERMARCAS	HYPE3	R\$ 0,58	6,81%	6,62%	8,04%
KROTON EDUCACIONAL	KROT3	R\$ 0,86	11,31%	8,40%	17,64%

(conclusão)					
Empresa	Código	LPA	R1	R2	R3
SABESP	SBSP3	R\$ 0,81	11,34%	7,93%	14,03%
TIM	TIMP3	R\$ 0,86	12,42%	11,10%	12,40%
SUZANO PAPEL E CELULOSE	SUZB5	-R\$ 0,25	12,84%	5,11%	4,14%
BRASKEM	BRKM5	R\$ 3,94	16,33%	10,27%	8,77%
EMBRAER	EMBR3	R\$ 0,33	3,40%	7,83%	11,97%
LOJAS RENNER	LREN3	R\$ 1,22	22,15%	21,07%	19,86%
COSAN	CSAN3	R\$ 1,33	10,96%	9,14%	5,56%
ELETRONBRAS	ELET6	-R\$ 10,68	-13,89%	-1,99%	-5,45%
GPA	PCAR4	R\$ 2,29	8,55%	19,72%	15,54%
CETIP	CTIP3	R\$ 1,90	22,64%	24,33%	23,83%
M. DIAS BRANCO	MDIA3	R\$ 5,34	12,98%	17,04%	17,06%
MULTIPLAN	MULT3	R\$ 1,93	9,88%	10,18%	8,23%
NATURA	NATU3	R\$ 1,14	19,15%	26,74%	35,05%
TRANSMISSÃO PAULISTA	TRPL4	R\$ 3,07	9,58%	7,05%	-3,41%
CIA SIDERÚRGICA NACIONAL	CSNA3	R\$ 1,30	12,34%	7,92%	8,70%
EQUATORIAL ENERGIA	EQTL3	R\$ 3,06	9,35%	11,74%	5,17%
GERDAU	GGBR4	-R\$ 2,69	-5,45%	5,60%	5,85%
PORTO SEGURO	PSSA3	R\$ 3,10	3,34%	4,51%	14,09%
EDP ENERGIAS DO BRASIL	ENBR3	R\$ 2,62	20,26%	18,22%	14,21%
CIA ENERGÉTICA DE MINAS GERAIS	CMIG4	R\$ 1,98	14,63%	22,51%	19,74%
LOCALIZA	RENT3	R\$ 1,93	14,83%	16,64%	17,70%
BRMALLS	BRML3	R\$ 0,01	9,80%	12,02%	13,41%
TAESA	TAEE11	R\$ 2,64	19,11%	18,79%	15,23%
RUMO LOG	RUMO3	-R\$ 0,58	6,51%	9,94%	11,14%
SÃO MARTINHO	SMT03	R\$ 1,89	7,78%	6,65%	5,09%

Fonte: Elaborada pela autora (2016)

Sendo R1, R2 e R3 representando o retorno sobre investimento de 1, 2 e 3 anos, respectivamente.

4.2 CÁLCULO DA EFICIÊNCIA COM A UTILIZAÇÃO DA DEA

Para esta segunda etapa, foi calculada a eficiência utilizando as técnicas de DEA e todas informações coletadas na fase anterior. Nas Tabelas 5 e 6 apresentadas

anteriormente, há valores negativos e não é permitido nos cálculos de DEA, como definido na metodologia apresentada no Capítulo 2.1. Por isso, conforme sugerido por Martins (2014), foram desconsideradas estas empresas, reduzindo a análise para 41 empresas.

Assim, após os cálculos no *software* MatLab, os dados de eficiência para cada empresa são apresentados na Tabela 7. As empresas que obtiveram um resultado igual a 1 foram consideradas como eficiente e fazem parte da carteira otimizada.

Tabela 7 – Dados de eficiência

(continua)		
Empresa	Código	Eficiência (%)
AMBEV	ABEV3	100%
ITAU UNIBANCO	ITUB4	100%
BANCO BRADESCO	BBDC4	91%
CIELO	CIEL3	100%
TELEFÔNICA BRASIL	VIVT4	98%
BB SEGURIDADE	BBSE3	100%
ITAUSA	ITSA4	100%
BANCO DO BRASIL	BBAS3	100%
ULTRAPAR	UGPA3	100%
BRF	BRFS3	100%
BM&FBOVESPA	BVMF3	73%
JBS	JBSS3	85%
GRUPO CCR	CCRO3	88%
WEG	WEGE3	81%
ENGIE BRASIL ENERGIA	EGIE3	100%
LOJAS AMERICANAS	LAME4	79%
RAIADROGASIL	RADL3	78%
CPFL ENERGIA	CPFE3	94%
FIBRIA	FIBR3	69%
HYPERMARCAS	HYPE3	92%
KROTON EDUCACIONAL	KROT3	78%
SABESP	SBSP3	64%
TIM	TIMP3	79%
BRASKEM	BRKM5	100%
EMBRAER	EMBR3	82%
LOJAS RENNER	LREN3	80%
COSAN	CSAN3	80%

(conclusão)		
Empresa	Código	Eficiência (%)
GPA	PCAR4	98%
CETIP	CTIP3	100%
M. DIAS BRANCO	MDIA3	100%
MULTIPLAN	MULT3	100%
NATURA	NATU3	84%
CIA SIDERÚRGICA NACIONAL	CSNA3	80%
EQUATORIAL ENERGIA	EQTL3	100%
PORTO SEGURO	PSSA3	97%
EDP ENERGIAS DO BRASIL	ENBR3	100%
CIA ENERGÉTICA DE MINAS GERAIS	CMIG4	100%
LOCALIZA	RENT3	95%
BRMALLS	BRML3	71%
TAESA	TAE11	100%
SÃO MARTINHO	SMT03	73%

Fonte: Elaborada pela autora (2016)

Analisando a Tabela 7, pode-se perceber que, das 41 empresas escolhidas para este estudo, 17 obtiveram uma eficiência de 100%, sendo elas: AMBEV, Itau Unibanco, Cielo, BB Seguridade, Itausa, Banco do Brasil, Ultrapar, BRF, Engie Brasil Energia, Braskem, Cetip, M. Dias Branco, Multiplan, Equatorial Energia, EDP Energias do Brasil, Companhia Energética de Minas Gerais e Taesa. Porém, desde o dia 30 de março de 2017, as ações da empresa Cetip não existem mais na BM&FBovespa, então, ela será desconsiderada para as etapas seguintes.

4.3 DETERMINAÇÃO DOS DADOS PARA A ANÁLISE DE PREVISÃO

Com a determinação de quais empresas fazem parte da carteira otimizada, foi possível coletar o preço diário das ações de cada uma delas, utilizando o endereço eletrônico Exame – Mercados. Conforme apresentado no Capítulo 2, não há qualquer restrição de valor para Rede Neural Artificial e nem para Box-Jenkins, então não foi necessário excluir alguma empresa do estudo. Os dados são referentes a 2 de abril de 2007, a 31 de março de 2017, sendo a escolha da data inicial e final para a amostra foi arbitrada.

As análises de previsão ocorreram conforme determinado a seguir:

- a) o equivalente a 12 meses de dados entre os anos de 2017 e 2016 → de 1º de abril de 2016 a 31 de março de 2017;
- b) o equivalente a 24 meses de dados entre os anos de 2017, 2016 e 2015 → de 1º de abril de 2015 a 31 de março de 2017;
- c) o equivalente a 36 meses de dados entre os anos de 2017, 2016, 2015 e 2014 → de 1º de abril de 2014 a 31 de março de 2017;
- d) o equivalente a 48 meses de dados entre os anos de 2017 a 2013 → de 1º de abril de 2013 a 31 de março de 2017;
- e) o equivalente a 60 meses de dados entre os anos de 2017 a 2012 → de 2 de abril de 2012 a 31 de março de 2017;
- f) o equivalente a 72 meses de dados entre os anos de 2017 a 2011 → de 1º de abril de 2011 a 31 de março de 2017;
- g) o equivalente a 84 meses de dados entre os anos de 2017 a 2010 → de 1º de abril de 2010 a 31 de março de 2017;
- h) o equivalente a 96 meses de dados entre os anos de 2017 a 2009 → de 1º de abril de 2009 a 31 de março de 2017;
- i) o equivalente a 108 meses de dados entre os anos de 2017 a 2008 → de 1º de abril de 2008 a 31 de março de 2017; e
- j) o equivalente a 120 meses de dados entre os anos de 2017 a 2007 → de 2 de abril de 2007 a 31 de março de 2017.

Porém, nem todas as empresas possuem dados até o ano de 2007. Então, para estes casos, foi realizada a prognose até a data final disponível. A Figura 24 exibe quantas previsões puderam ser feitas com os dados disponibilizados, totalizando 130 cálculos.

O Apêndice E apresenta todos estes dados respectivamente para cada uma das empresas consideradas eficientes. As tabelas nestes apêndices exibem somente apenas algum destes valores, devido ao grande volume de dados ao expressar, no máximo, todos os 2475 preços de cada empresa.

Figura 24 – Possíveis previsões para cada empresa

Empresa	Código	Possíveis previsões (em meses)									
		12	24	36	48	60	72	84	96	108	120
AMBEV	ABEV3	✓	✓	✓							
ITAU UNIBANCO	ITUB4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
CIELO	CIEL3	✓	✓	✓	✓	✓	✓	✓			
BB SEGURIDADE	BBSE3	✓	✓	✓							
ITAUSA	ITSA4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
BANCO DO BRASIL	BBAS3	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
ULTRAPAR	UGPA3	✓	✓	✓	✓	✓					
BRF	BRFS3	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
ENGIE BRASIL ENERGIA	EGIE3	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
BRASKEM	BRKM5	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
M. DIAS BRANCO	MDIA3	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
MULTIPLAN	MULT3	✓	✓	✓	✓	✓	✓	✓	✓	✓	
EQUATORIAL ENERGIA	EQTL3	✓	✓	✓	✓	✓	✓	✓	✓		
EDP ENERGIAS DO BRASIL	ENBR3	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
CIA ENERGÉTICA DE MINAS GERAIS	CMIG4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
TAESA	TAE11	✓	✓	✓	✓	✓					

Fonte: Elaborado pela autora (2017)

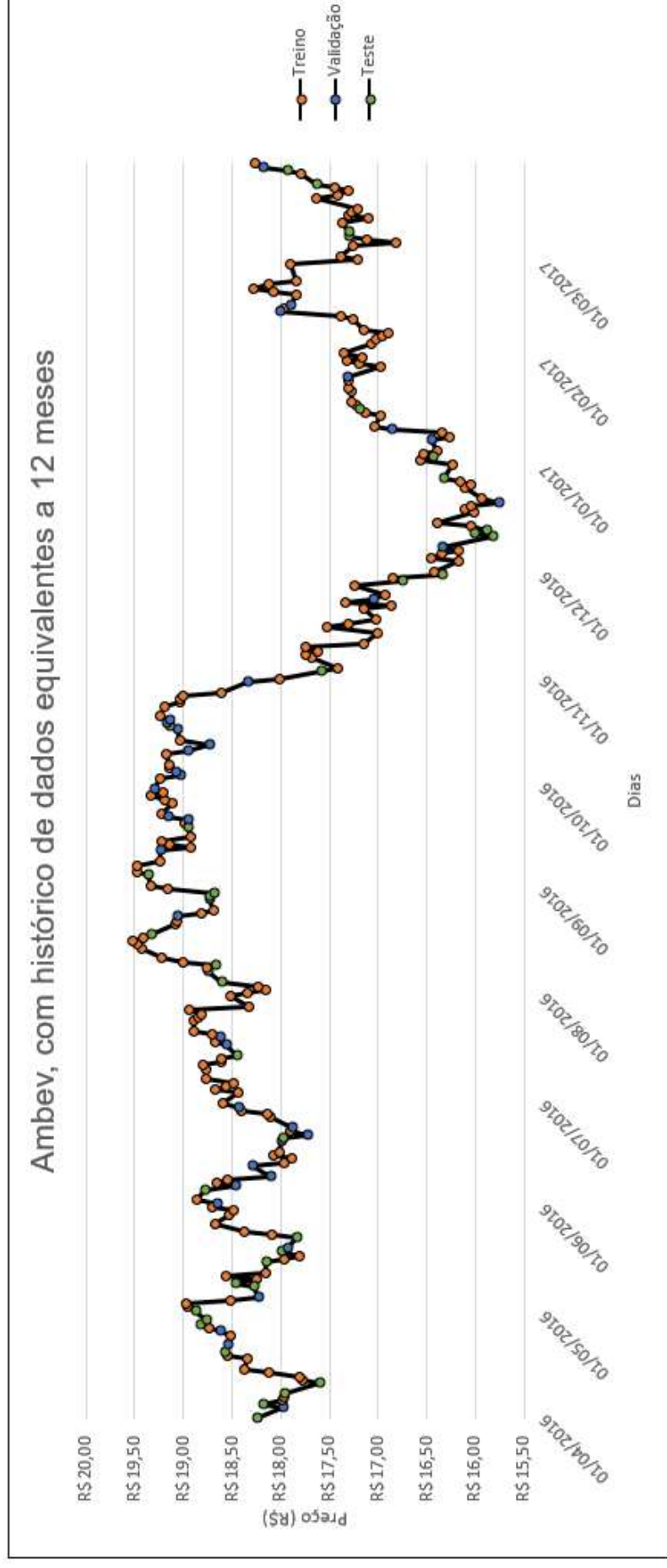
4.4 UTILIZAÇÃO DA RNA

Nesta seção, foi calculada a previsão dos preços das ações, utilizando Redes Neurais Artificiais. Com os parâmetros definidos no Capítulo 3, os cálculos foram realizados no *software* MatLab, com o auxílio do *toolbox* das RNA. Foi criada uma rede diferente para cada análise, para se ter um melhor entendimento do comportamento do conjunto de dados.

4.4.1 Previsão com RNA

Os preços históricos foram divididos aleatoriamente em dados de treino, validação e teste da rede. Utilizando a empresa Ambev com o equivalente a 12 meses de dados como exemplo, os seus 250 preços ficaram classificados como: 174 valores para treino, 38 para validação e 38 para teste, representando respectivamente 70%, 15% e 15% dos dados, conforme determinado no capítulo anterior. O gráfico que apresenta como ficou esta divisão foi exibido na Figura 25.

Figura 25 – Classificação dos dados para RNA



Fonte: Elaborado pela autora (2017)

Então, foi calculada a previsão para as 130 análises previamente detalhadas. Os resultados da prognose de 12, 24, 36, 48 e 60 meses foram apresentados na Tabela 8, enquanto os de 72, 84, 96, 108 e 120 meses estão na Tabela 9. Os dados históricos foram até o dia 31 de março de 2017, e a previsão calculada foi para a data seguinte. Então, este valor calculado representa a prognose do preço do dia 3 de abril de 2017, já que o período de tempo estudado foi dia a dia comercial.

Tabela 8 – Previsão para 12, 24, 36, 48 e 60 meses

Código	Resultados das previsões com RNA				
	12	24	36	48	60
ABEV3	R\$ 18,32	R\$ 18,24	R\$ 18,31	-	-
ITUB4	R\$ 38,04	R\$ 37,77	R\$ 38,18	R\$ 38,23	R\$ 38,39
CIEL3	R\$ 28,13	R\$ 28,22	R\$ 27,83	R\$ 28,27	R\$ 27,76
BBSE3	R\$ 29,44	R\$ 29,16	R\$ 29,30	-	-
ITSA4	R\$ 9,45	R\$ 9,47	R\$ 9,56	R\$ 9,33	R\$ 9,50
BBAS3	R\$ 34,24	R\$ 33,84	R\$ 33,59	R\$ 33,56	R\$ 33,93
UGPA3	R\$ 71,57	R\$ 71,19	R\$ 71,36	R\$ 71,36	R\$ 71,62
BRFS3	R\$ 38,63	R\$ 38,13	R\$ 38,17	R\$ 39,03	R\$ 39,03
EGIE3	R\$ 35,44	R\$ 35,40	R\$ 35,37	R\$ 35,43	R\$ 35,35
BRKM5	R\$ 31,66	R\$ 31,70	R\$ 32,28	R\$ 31,82	R\$ 31,81
MDIA3	R\$ 127,28	R\$ 127,90	R\$ 126,01	R\$ 127,41	R\$ 128,65
MULT3	R\$ 65,60	R\$ 65,84	R\$ 65,82	R\$ 65,56	R\$ 65,70
EQTL3	R\$ 58,72	R\$ 61,08	R\$ 58,86	R\$ 58,60	R\$ 58,37
ENBR3	R\$ 13,88	R\$ 13,77	R\$ 13,86	R\$ 13,94	R\$ 13,95
CMIG4	R\$ 9,98	R\$ 10,39	R\$ 10,15	R\$ 10,19	R\$ 10,31
TAEE11	R\$ 22,46	R\$ 22,35	R\$ 22,81	R\$ 22,44	R\$ 23,10

Fonte: Elaborado pela autora (2017)

Tabela 9 – Previsão para 72, 84, 96, 108 e 120 meses

(continua)

Código	Resultados das previsões com RNA				
	72	84	96	108	120
ABEV3	-	-	-	-	-
ITUB4	R\$ 38,23	R\$ 37,59	R\$ 37,82	R\$ 37,49	R\$ 37,78
CIEL3	R\$ 28,56	R\$ 27,86	-	-	-
BBSE3	-	-	-	-	-
ITSA4	R\$ 9,47	R\$ 9,54	R\$ 9,54	R\$ 9,55	R\$ 9,60
BBAS3	R\$ 33,41	R\$ 34,12	R\$ 33,79	R\$ 33,93	R\$ 34,02
UGPA3	-	-	-	-	-
BRFS3	R\$ 38,55	R\$ 38,39	R\$ 38,14	R\$ 38,60	R\$ 38,28
EGIE3	R\$ 35,25	R\$ 35,25	R\$ 35,09	R\$ 35,34	R\$ 35,34

(conclusão)

Código	Resultados das previsões com RNA				
	72	84	96	108	120
BRKM5	R\$ 31,53	R\$ 31,26	R\$ 31,95	R\$ 31,27	R\$ 31,76
MDIA3	R\$ 128,87	R\$ 129,56	R\$ 128,81	R\$ 129,93	R\$ 127,78
MULT3	R\$ 65,85	R\$ 65,70	R\$ 66,74	R\$ 65,62	-
EQTL3	R\$ 58,63	R\$ 57,84	R\$ 57,81	-	-
ENBR3	R\$ 13,83	R\$ 13,82	R\$ 13,59	R\$ 13,75	R\$ 13,72
CMIG4	R\$ 10,23	R\$ 10,26	R\$ 10,34	R\$ 10,46	R\$ 10,16
TAEE11	-	-	-	-	-

Fonte: Elaborado pela autora (2017)

Com a previsão calculada, foi possível dar continuidade ao estudo com a análise de performance da rede.

4.4.2 Análise estatística

Como medida de performance da rede, foi estudado o MSE e o r^2 de cada previsão realizada, assim como Moghaddam, Moghaddam e Esfandyari (2016) fizeram em seu trabalho. Então, foram apresentados na Tabela 10 os MSE de cada cálculo.

Tabela 10 – Erro quadrático médio de cada prognose

(continua)

Código	MSE para cada previsão realizada									
	12	24	36	48	60	72	84	96	108	120
ABEV3	0,045	0,055	0,097	-	-	-	-	-	-	-
ITUB4	0,345	0,456	0,447	0,443	0,357	0,374	0,384	0,363	0,858	0,562
CIEL3	0,315	0,730	1,058	0,909	1,118	0,940	0,924	-	-	-
BBSE3	0,379	0,417	0,392	-	-	-	-	-	-	-
ITSA4	0,022	0,024	0,028	0,028	0,026	0,034	0,030	0,031	0,041	0,037
BBAS3	0,776	0,466	0,517	0,427	0,371	0,345	0,393	0,332	0,336	0,350
UGPA3	0,953	1,056	0,848	0,755	0,752	-	-	-	-	-
BRFS3	0,822	1,033	0,953	0,868	0,798	0,708	0,638	0,640	0,525	0,606
EGIE3	0,247	0,263	0,290	0,264	0,253	0,232	0,209	0,229	0,186	0,193
BRKM5	0,613	0,585	0,306	0,319	0,251	0,230	0,217	0,194	0,197	0,172
MDIA3	6,149	4,391	4,790	2,973	3,195	2,469	2,656	2,069	3,235	1,642
MULT3	1,271	1,007	0,970	0,867	0,857	0,710	0,705	0,649	0,615	-
EQTL3	0,593	1,297	0,374	0,321	0,253	0,223	0,288	0,246	-	-
ENBR3	0,055	0,073	0,053	0,049	0,051	0,526	0,478	0,491	0,460	0,456

(conclusão)

Código	MSE para cada previsão realizada									
	12	24	36	48	60	72	84	96	108	120
CMIG4	0,131	0,078	0,103	0,094	0,172	0,236	0,313	0,182	0,202	0,202
TAAE11	0,168	0,244	0,118	0,111	0,131	-	-	-	-	-

Fonte: Elaborado pela autora (2017)

Conforme ressalta Pellegrini (2000), quanto mais próximo do zero for o valor do MSE, melhor, por se tratar de um cálculo de erro. Ainda, foi analisado o coeficiente de determinação de cada previsão realizada, cujos valores foram apresentados na Tabela 11.

Tabela 11 – Coeficiente de determinação de cada prognose

Código	r^2 para cada previsão realizada									
	12	24	36	48	60	72	84	96	108	120
ABEV3	0,996	0,996	0,976	-	-	-	-	-	-	-
ITUB4	0,998	0,997	0,998	0,994	0,999	0,998	0,998	0,999	0,979	0,994
CIEL3	0,997	0,977	0,948	0,967	0,995	0,999	0,997	-	-	-
BBSE3	0,968	0,997	0,998	-	-	-	-	-	-	-
ITSA4	0,997	0,999	0,997	0,995	0,997	0,995	0,999	0,999	0,998	0,999
BBAS3	0,989	0,998	0,998	0,999	1,000	1,000	0,996	0,999	0,999	0,998
UGPA3	0,968	0,991	1,000	0,999	0,999	-	-	-	-	-
BRFS3	0,997	0,999	0,999	0,999	0,999	1,000	1,000	0,999	1,000	1,000
EGIE3	0,989	0,997	0,998	0,999	0,998	0,999	1,000	0,999	1,000	1,000
BRKM5	0,995	0,996	0,999	0,998	0,999	0,999	0,999	0,999	0,999	1,000
MDIA3	0,997	0,998	0,996	1,000	0,998	1,000	0,999	1,000	0,998	1,000
MULT3	0,968	0,998	0,999	0,998	0,998	1,000	1,000	1,000	0,998	-
EQTL3	0,998	0,991	1,000	1,000	1,000	1,000	0,999	1,000	-	-
ENBR3	0,972	0,980	0,997	1,000	0,999	0,999	1,000	0,999	1,000	1,000
CMIG4	0,980	0,998	0,998	0,999	0,999	0,999	0,997	1,000	1,000	1,000
TAAE11	0,990	0,979	0,997	0,995	0,997	-	-	-	-	-

Fonte: Elaborado pela autora (2017)

Conforme Lapponi (2005), quanto mais próximo de 1 for o r^2 , mais forte é a relação linear entre as variáveis, e maior a porcentagem de dados explicados pelo modelo.

4.4.3 Análise dos resíduos e erro

Por ser uma rede com *backpropagation* como algoritmo de aprendizagem, foi possível coletar o valor calculado respectivo de cada dado de entrada, conforme explicado por Guimarães (2006). Então, foi possível compará-los com os valores reais, que está exibido na Figura 26. Também foi utilizada a empresa Ambev com 12 meses de dados para exemplificar esta comparação.

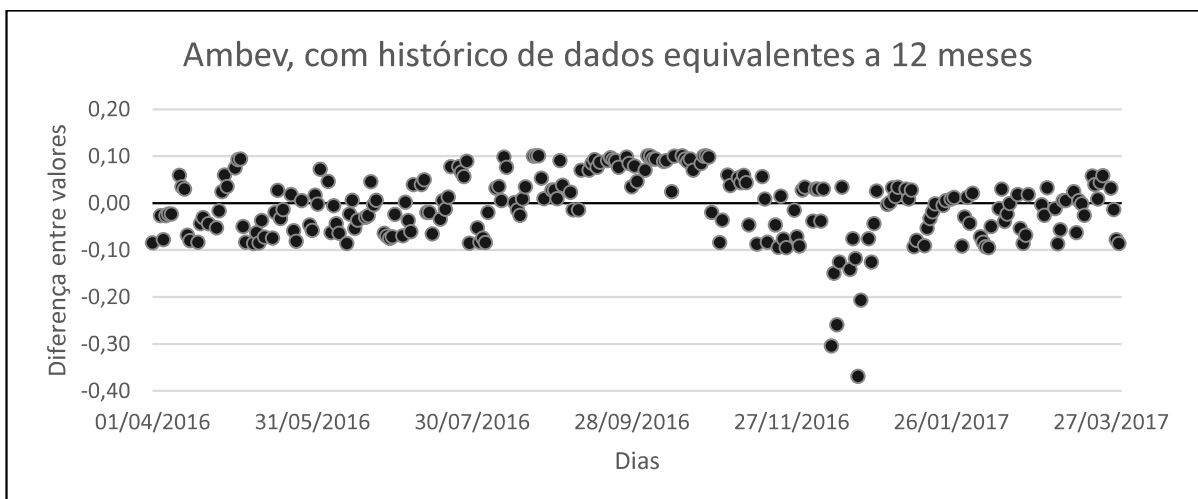
Figura 26 – Comparativo dos dados reais com os calculados com RNA



Fonte: Elaborado pela autora (2017)

A diferença entre cada um dos valores reais e os calculados representa o erro da rede. Esta análise dos resíduos deste exemplo foi apresentada na Figura 27.

Figura 27 – Análise dos resíduos

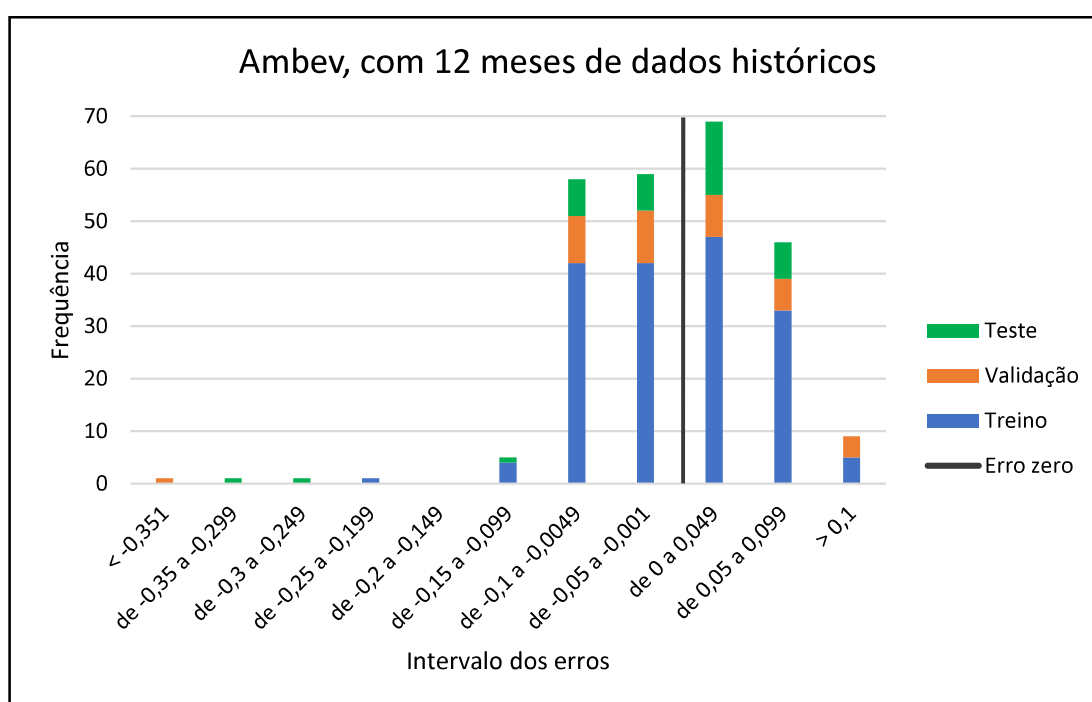


Fonte: Elaborado pela autora (2017)

Por se tratar de erro, então quanto mais próximo do zero estes valores, melhor, pois representa uma rede com uma boa acuracidade.

Por ter a classificação dos dados entre treino, validação e teste da rede, apresentado na Figura 25, foi possível analisar o seu comportamento para estas categorias. Com o propósito de poder fazer uma observação geral destes erros, eles foram separados em intervalos de variação de 0,05, para auxiliar na análise de frequência que ocorrem. Ainda utilizando a empresa Ambev com o equivalente a 12 meses de dados como exemplo, o histograma destes erros foi exibido na Figura 28.

Figura 28 – Histograma dos erros



Fonte: Elaborado pela autora (2017)

No gráfico, a linha vertical representa o erro igual a zero, ou seja, quando a rede calcula o preço para um determinado dia e foi exatamente igual ao observado neste mesmo dia. Então, a maioria dos erros deveria estar próximo a esta linha, para significar que a rede obteve bons resultados, o que ocorre neste exemplo. Caso a distribuição dos erros tivesse ficado afastada dessa linha, seria necessário alterar os parâmetros escolhidos para a rede. Todos outros 129 histogramas apresentaram comportamento similar a este.

4.5 APLICAÇÃO DA METODOLOGIA DE BOX-JENKINS

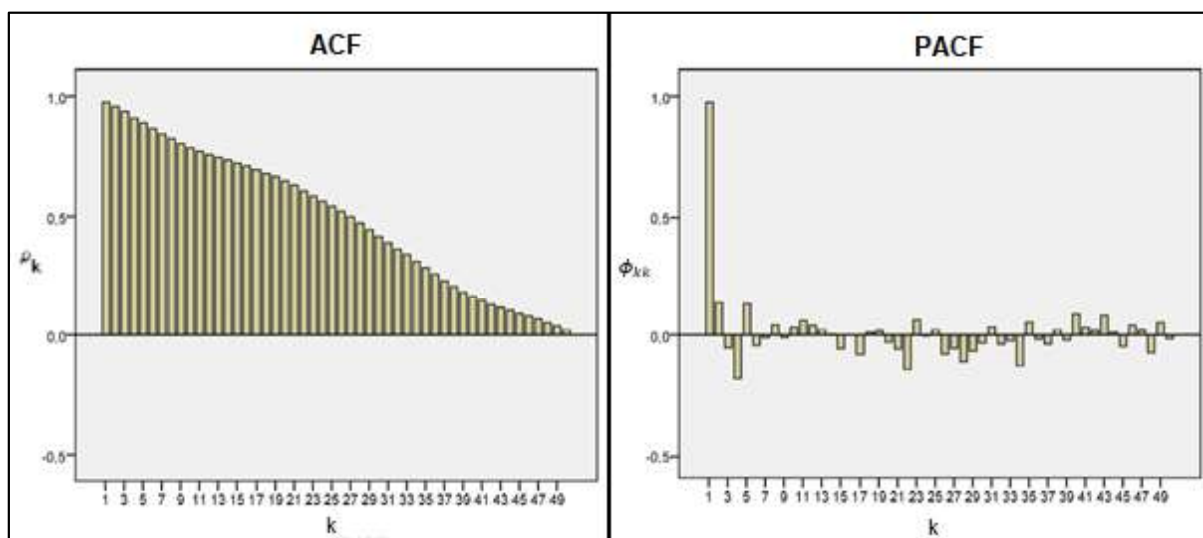
Nesta etapa, foi calculada a previsão com a metodologia Box-Jenkins, e para isso, foi utilizado o *software* SPSS. A análise com menor quantidade de dados foi a equivalente a 12 meses, contendo 252 dados, satisfazendo a quantidade mínima sugerida por Pellegrini (2000).

Conforme explicado no Capítulo 2.3, deve-se realizar algumas etapas para esta metodologia: observação dos gráficos das funções de autocorrelação e autocorrelação parcial; escolha do modelo a ser utilizado, e a previsão dos valores. Foi, então, finalizado com o estudo dos correlogramas dos resíduos. Todas estas informações foram obtidas no *software* SPSS.

4.5.1 Previsão com Box-Jenkins

Iniciando o estudo com Box-Jenkins, foi apresentado, na Figura 29, os gráficos das funções de autocorrelação e autocorrelação parcial, para a empresa Ambev e com o equivalente a um ano de dados. Por serem 130 análises de previsões realizadas, obteve-se 260 correlogramas, sendo um grande volume de dados, por isso foi utilizado este caso para demonstrar os gráficos. Este exemplo foi escolhido para retratar todas as informações por ser o primeiro da lista, sem nenhum outro motivo.

Figura 29 – Correlogramas



Fonte: Elaborado pela autora (2017)

Analisando o gráfico da ACF, percebe-se um comportamento de série temporal não estacionária, pelo coeficiente ρ_k não se aproximar de zero rapidamente, o que indica a necessidade de diferenciar a série. Todos outros 129 correlogramas apresentaram comportamento similar.

Após a determinação dos modelos realizada pelo *software*, foi calculada a previsão da data seguinte para cada caso. Para a primeira empresa da lista, Ambev, são apresentados, na Tabela 12, os modelos e a prognose de cada análise possível. As demais são exibidas no Apêndice F.

Tabela 12 – Resultados do cálculo de previsão

AMBEV - ABEV3		
12 meses	Modelo	ARIMA(1,1,0)
	Valor calculado	R\$ 18,24
24 meses	Modelo	ARIMA(2,1,2)
	Valor calculado	R\$ 18,29
36 meses	Modelo	ARIMA(0,1,0)
	Valor calculado	R\$ 18,24

Fonte: Elaborado pela autora (2017)

Conforme apresentado anteriormente, os dados históricos foram até o dia 31 de março de 2017, e a previsão calculada foi para a data seguinte. Então, este valor calculado representa a prognose do preço do dia 3 de abril de 2017, já que o período de tempo estudado foi dia a dia comercial. Todos os valores foram calculados com um intervalo de confiança de 95%, assim como Sharafi et al. (2017) calcularam em seu trabalho.

4.5.2 Análise estatística

Como análise de performance da previsão, foram analisados o MSE e o coeficiente de determinação de cada cálculo, informações fornecidas pelo *software*, para comprovar se o modelo escolhido foi o mais adequado. Estes dados foram apresentados a seguir, sendo a Tabela 13 para o erro quadrático médio de cada análise.

Tabela 13 – Erro quadrático médio de cada prognose

Código	MSE para cada previsão realizada									
	12	24	36	48	60	72	84	96	108	120
ABEV3	0,044	0,055	0,056	-	-	-	-	-	-	-
ITUB4	0,340	0,370	0,412	0,367	0,352	0,350	0,353	0,355	0,416	0,419
CIEL3	0,306	0,462	0,484	0,972	0,992	0,920	0,867	-	-	-
BBSE3	0,285	0,403	0,373	-	-	-	-	-	-	-
ITSA4	0,020	0,023	0,026	0,024	0,024	0,027	0,028	0,029	0,033	0,033
BBAS3	0,415	0,425	0,487	0,407	0,362	0,336	0,320	0,307	0,324	0,321
UGPA3	0,803	0,870	0,843	0,746	0,696	-	-	-	-	-
BRFS3	0,760	0,982	0,914	0,808	0,717	0,656	0,585	0,536	0,516	0,491
EGIE3	0,211	0,240	0,285	0,259	0,247	0,221	0,199	0,180	0,176	0,173
BRKM5	0,403	0,345	0,265	0,226	0,208	0,209	0,191	0,176	0,169	0,165
MDIA3	4,973	3,549	3,031	2,809	2,446	2,120	1,893	1,703	1,550	1,421
MULT3	0,978	0,882	0,904	0,799	0,766	0,704	0,659	0,616	0,579	-
EQTL3	0,569	0,412	0,327	0,267	0,234	0,200	0,176	0,156	-	-
ENBR3	0,036	0,047	0,048	0,047	0,048	0,510	0,471	0,438	0,432	0,426
CMIG4	0,066	0,070	0,088	0,081	0,149	0,185	0,180	0,173	0,188	0,187
TAAE11	0,158	0,123	0,112	0,102	0,118	-	-	-	-	-

Fonte: Elaborado pela autora (2017)

De acordo com Pellegrini (2000), quanto mais baixo for este valor, melhor, já que se trata de uma análise de erro de previsão. Ainda, são apresentados os valores dos coeficientes de determinação, que estão na Tabela 14.

Tabela 14 – Coeficiente de determinação de cada prognose

(continua)

Código	r^2 para cada previsão realizada									
	12	24	36	48	60	72	84	96	108	120
ABEV3	0,952	0,930	0,969	-	-	-	-	-	-	-
ITUB4	0,983	0,985	0,977	0,976	0,974	0,976	0,981	0,981	0,982	0,981
CIEL3	0,966	0,957	0,954	0,918	0,978	0,979	0,977	-	-	-
BBSE3	0,896	0,956	0,956	-	-	-	-	-	-	-
ITSA4	0,979	0,983	0,976	0,972	0,970	0,983	0,991	0,990	0,989	0,988
BBAS3	0,986	0,987	0,982	0,981	0,980	0,980	0,982	0,983	0,986	0,985
UGPA3	0,891	0,951	0,985	0,987	0,992	-	-	-	-	-
BRFS3	0,969	0,987	0,985	0,987	0,993	0,995	0,997	0,998	0,998	0,998
EGIE3	0,949	0,965	0,963	0,963	0,973	0,991	0,995	0,997	0,998	0,998
BRKM5	0,989	0,993	0,994	0,993	0,993	0,993	0,993	0,994	0,995	0,994
MDIA3	0,986	0,995	0,993	0,992	0,994	0,997	0,998	0,998	0,998	0,999
MULT3	0,933	0,989	0,984	0,983	0,983	0,991	0,994	0,996	0,997	-
EQTL3	0,979	0,995	0,998	0,998	0,999	0,999	0,999	0,999	-	-

(conclusão)

Código	r^2 para cada previsão realizada									
	12	24	36	48	60	72	84	96	108	120
ENBR3	0,960	0,975	0,985	0,983	0,980	0,981	0,994	0,995	0,994	0,994
CMIG4	0,973	0,988	0,993	0,993	0,996	0,996	0,997	0,996	0,996	0,995
TAE11	0,955	0,980	0,975	0,976	0,975	-	-	-	-	-

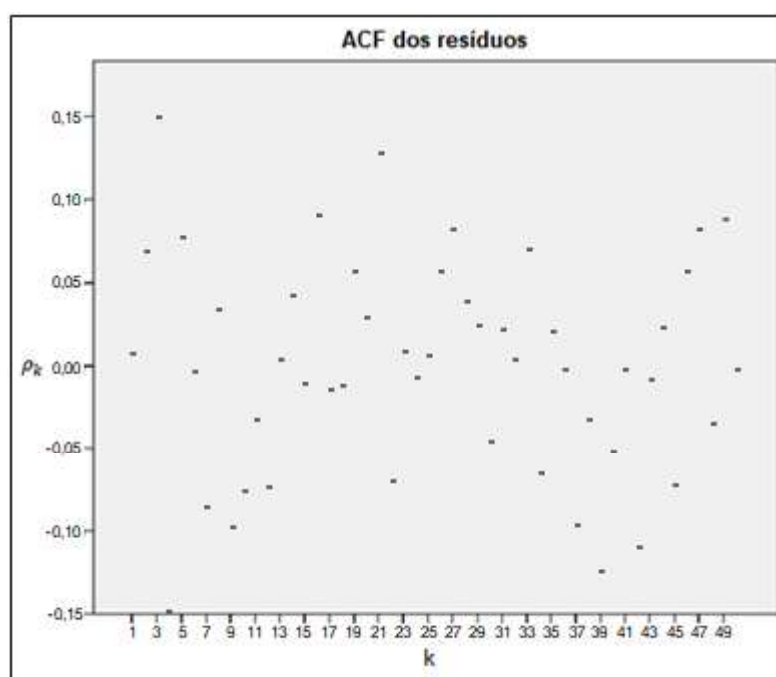
Fonte: Elaborado pela autora (2017)

Conforme Lapponi (2005), quanto mais próximo de 1 for o r^2 , mais forte é a relação linear entre as variáveis, e maior a porcentagem de dados explicados pelo modelo.

4.5.3 Análise dos resíduos

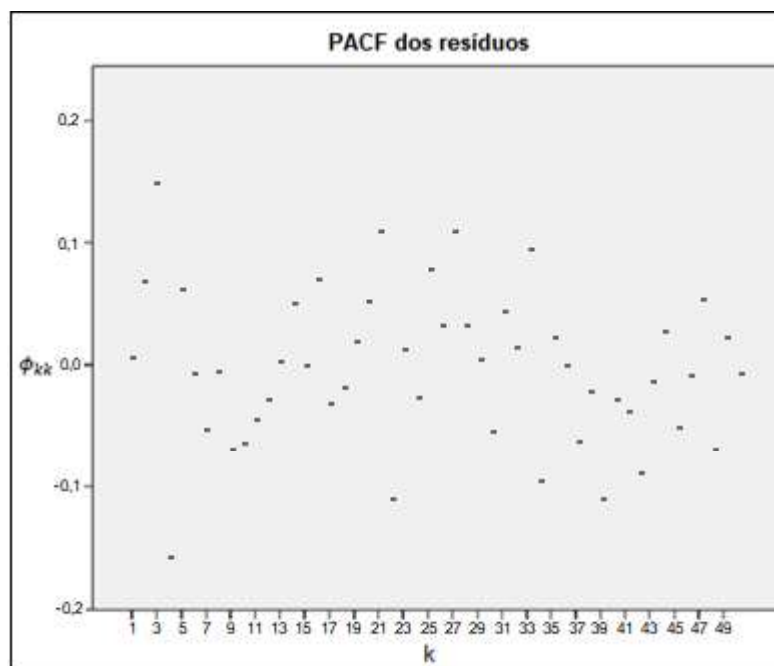
Para comprovar que o modelo escolhido foi mais apropriado, deve-se analisar também os gráficos das funções de autocorrelação e autocorrelação parcial dos resíduos. Pelo grande volume de dados gerados também por todos 130 correlogramas dos resíduos, ainda foi utilizada a empresa Ambev com o equivalente a um ano de dados como exemplo. As Figuras 30 e 31 retratam estes correlogramas.

Figura 30 – Correlograma da ACF dos resíduos



Fonte: Elaborado pela autora (2017)

Figura 31 – Correlograma da PACF dos resíduos



Fonte: Elaborado pela autora (2017)

Analizando estes dois gráficos, foi possível observar que representam a situação ideal para os resíduos, por estarem distribuídos aleatoriamente em torno do zero, conforme estudado no Capítulo 2.3. Portanto, foi comprovado que o modelo ARIMA(1,1,0) é o mais adequado para calcular a previsão do preço da ação da empresa Ambev, para o equivalente a um ano de dados. Todos outros 129 correlogramas apresentaram comportamento similar.

4.6 AVALIAÇÃO DOS RESULTADOS

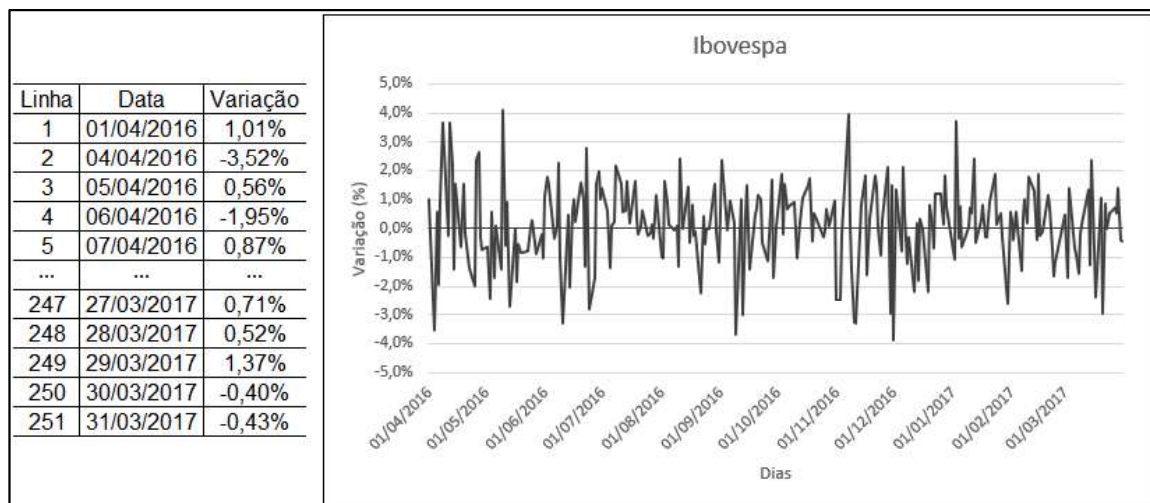
Nesta última etapa do trabalho, foram avaliados todos os resultados obtidos: a carteira otimizada com Análise Envoltória de Dados e os valores previstos com Redes Neurais Artificiais e metodologia Box-Jenkins.

4.6.1 Resultado da DEA

Para avaliar a vantagem de usar Análise Envoltória de Dados, foi comparado o retorno de um investimento hipotético, conforme explicado no Capítulo 3.6. Este retorno de capital foi calculado com o preço das ações presentes na carteira otimizada. Para a comparação, foi necessário coletar a variação percentual do índice Ibovespa,

que foi retirado do endereço eletrônico Exame – Mercados e corresponde a 1º de abril de 2016 até 31 de março de 2017. Devido ao grande volume de dados, a Figura 32 apresenta alguns destes valores.

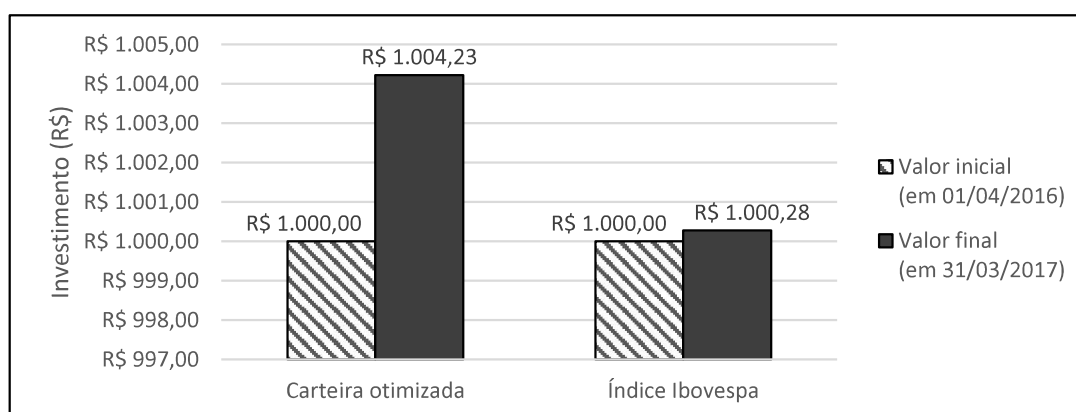
Figura 32 – Variação percentual do índice Ibovespa



Fonte: Elaborado pela autora (2017)

Para a carteira otimizada, foram utilizados os preços deste mesmo período de tempo, ou seja, dados de 1º de abril de 2016 até 31 de março de 2017. Foi arbitrado uma quantia de R\$1000,00 para este investimento hipotético, valor em que foi dividido por igual entre as 16 empresas consideradas eficientes pela DEA, enquanto foi utilizado o valor integral para o índice Ibovespa. Então, foi calculado o quanto esta quantia alterava a cada dia, até a última data. Para a carteira otimizada, foi somado, de todas empresas, o valor correspondente ao dia 31 de março de 2017, para obter a quantia final do investimento. O gráfico da Figura 33 apresenta estes valores.

Figura 33 – Gráfico comparativo de investimentos



Fonte: Elaborado pela autora (2017)

Observando os valores presentes no gráfico, foi possível calcular o retorno de cada investimento. A carteira otimizada com DEA obteve um retorno de 0,423%, enquanto o da Ibovespa foi de 0,028%.

4.6.2 Resultados dos métodos de previsão

Para esta última etapa, foi calculado o erro das previsões calculadas previamente, conforme explicado na última seção do Capítulo 3. Como se calcula a previsão para a data seguinte do conjunto de dados e os dados históricos foram até o dia 31 de março de 2017, então o valor considerado como preço real foi o correspondente ao próximo dia útil, que foi em 3 de abril. O valor para este dia também foi coletado no endereço eletrônico Exame – Mercados.

O erro foi calculado para todas 130 previsões das RNA e também da metodologia Box-Jenkins. A Tabela 15 apresenta o valor real do dia 3 de abril de 2017 e os erros calculados, referentes a empresa Ambev. Os demais erros estão exibidos no Apêndice G.

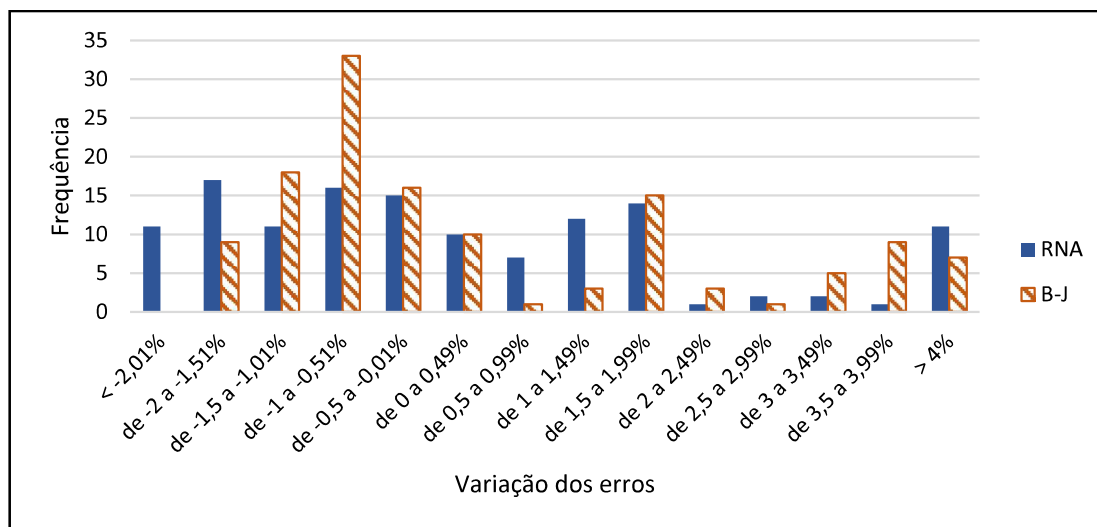
Tabela 15 – Erro de previsão da empresa Ambev

AMBEV - ABEV3			
Preço do dia 03/04/2017:		R\$ 18,00	
Método utilizado:		RNA	B-J
12 meses	Valor calculado	R\$ 18,32	R\$ 18,24
	Erro	1,78%	1,33%
24 meses	Valor calculado	R\$ 18,24	R\$ 18,29
	Erro	1,33%	1,61%
36 meses	Valor calculado	R\$ 18,31	R\$ 18,24
	Erro	1,72%	1,33%

Fonte: Elaborado pela autora (2017)

Com o propósito de poder fazer uma observação geral destes erros, eles foram separados em intervalos de variação de 0,5%, para auxiliar na análise de frequência que ocorrem. Este histograma é apresentado na Figura 34.

Figura 34 – Histograma dos erros de previsão

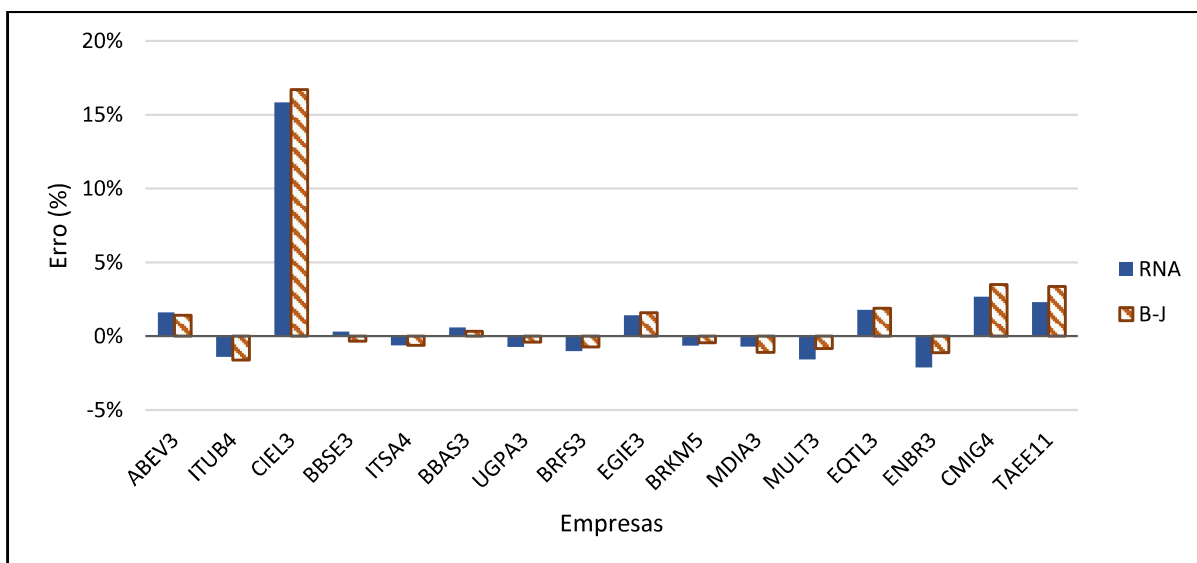


Fonte: Elaborado pela autora (2017)

Observando o gráfico, verifica-se que a maior frequência de erro ocorreu no intervalo de -2 a -1,51%, para as Redes Neurais Artificiais, e ainda se percebe que, para este estudo, este método tende a prever valores menores do que o observado na prática, ou seja, possui mais erros negativos do que positivos. Enquanto para a metodologia de Box-Jenkins, a maior ocorrência foi no intervalo de -1 a -0,51%, e repete o comportamento de se ter mais erros negativos do que positivos.

Para uma análise mais detalhada dos erros, foi calculado a média dos valores, possibilitando a comparação de qual método de previsão foi o mais adequado para cada empresa. O gráfico das médias dos erros é apresentado na Figura 35.

Figura 35 – Média dos erros de cada empresa



Fonte: Elaborado pela autora (2017)

Com o gráfico da Figura 35, foi possível observar que as médias dos erros dos dois métodos de previsão ficaram parecidos para todas empresas, ou seja, os métodos calcularam valores de previsão similares.

Com o objetivo de examinar a reprodutibilidade dos resultados obtidos e garantir que estes métodos possuem uma boa precisão, não sendo somente um acaso exclusivo ao dia 3 de abril de 2017, foi calculado a previsão para o restante dessa semana, ou seja, os dias 4, 5, 6 e 7 de abril de 2017. Tanto para RNA quanto para metodologia B-J, para esta prognose, foi utilizada dados históricos equivalentes a 12 meses e acrescentado o último valor real para prever o próximo, ou seja, preços de até o dia 3 de abril para prever dia 4, com a análise seguinte tendo preços de até o dia 4 para prever o dia 5, e assim sucessivamente.

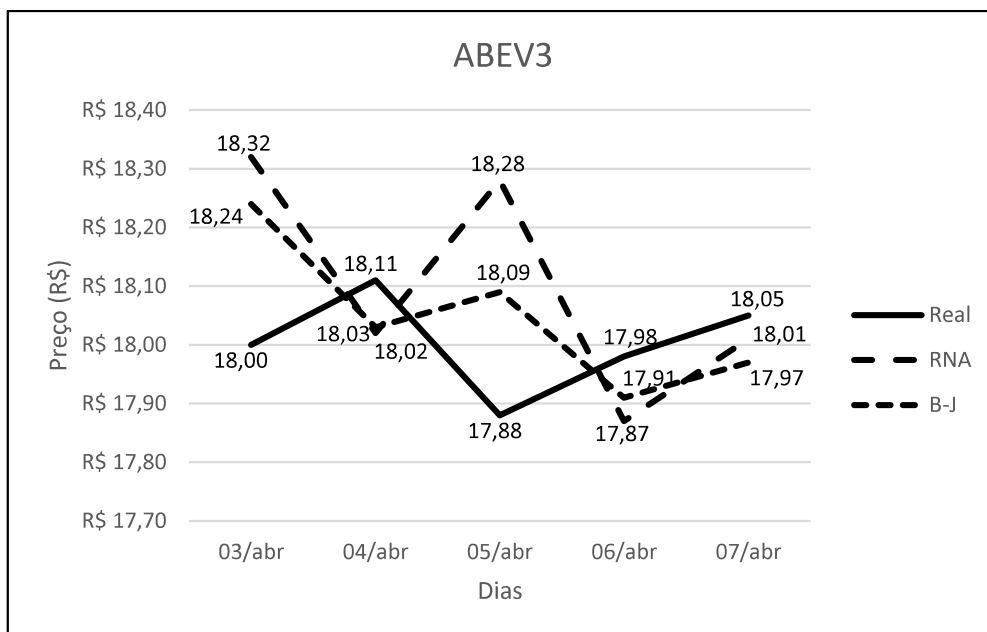
Estas redes neurais também foram criadas no *software* MatLab e com os mesmos parâmetros determinados, sendo eles: topologia *multilayer perceptrons*, uma camada de entrada com três ocultas e mais uma de saída, *backpropagation* como algoritmo de aprendizagem, para o treinamento de rede foi escolhido o gradiente conjugado escalonado, e 70% dos dados foram para treino, 15% para validação e 15% para teste da rede, sendo divididos aleatoriamente.

Para finalizar esta análise complementar, foi calculada a previsão do restante da semana utilizando a metodologia Box-Jenkins. O cálculo também foi realizado no *software* SPSS, e com um intervalo de confiança com probabilidade de 95%. Os correlogramas, tanto das funções de autocorrelação quanto das de autocorrelação parciais, tiveram comportamento similar aos apresentados na subseção anterior deste mesmo capítulo.

A Figura 36 mostra os valores previstos para a empresa Ambev, tanto para as RNA quanto para Box-Jenkins. Os demais resultados foram exibidos no Apêndice H.

As informações das redes, como o MSE, e o r^2 de cada análise, estão no Apêndice I. Para a metodologia B-J, os modelos determinados pelo *software*, o erro quadrático médio e o coeficiente de determinação de cada prognose foram exibidos no Apêndice J.

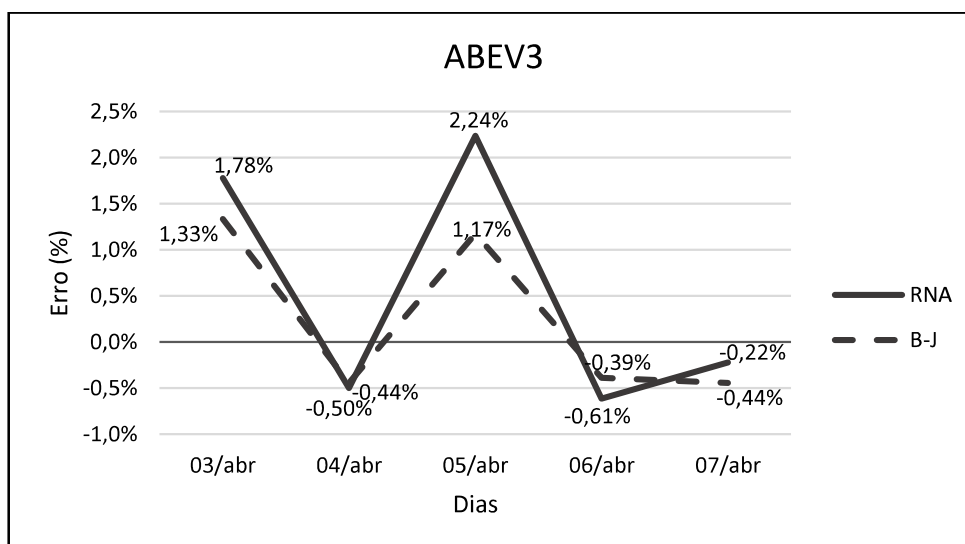
Figura 36 – Resultados das previsões para a empresa Ambev



Fonte: Elaborado pela autora (2017)

Após as previsões complementar de ambos os métodos, foi calculado o erro de cada análise, conforme explicado na seção 3.6 do Capítulo 3. O comparativo dos erros da empresa Ambev foi apresentado na Figura 37, e os demais estão no Apêndice L.

Figura 37 – Erros das previsões do restante da semana da empresa Ambev

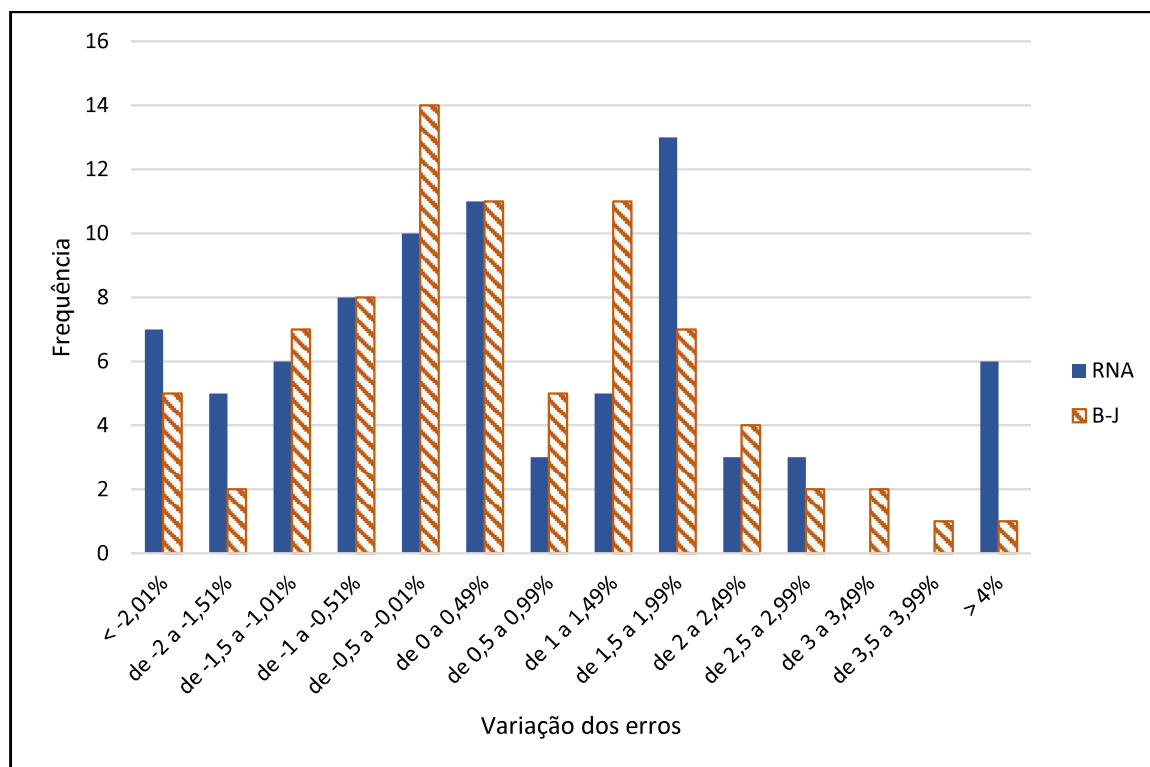


Fonte: Elaborado pela autora (2017)

Observa-se que no gráfico da Figura 37, os erros percentuais estão uniformemente distribuídos e próximos do valor zero. A fim de fazer uma observação

geral destes erros, eles foram separados em intervalos de variação de 0,5%, para auxiliar na análise de frequência que ocorrem. Este histograma foi apresentado na Figura 38.

Figura 38 – Histograma dos erros de previsão do restante da semana

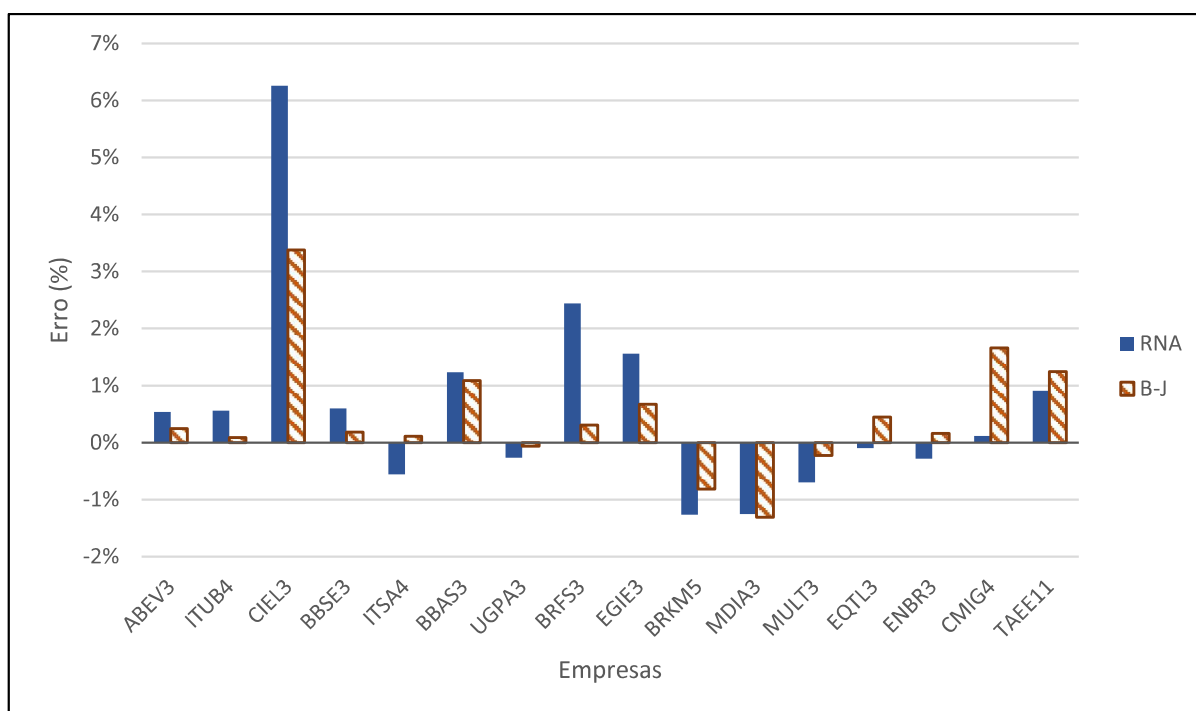


Fonte: Elaborado pela autora (2017)

Observando o gráfico, verifica-se que a maior frequência de erro ocorreu no intervalo de 1,5 a 1,99%, para as Redes Neurais Artificiais, e ainda se percebe que este método tende a prever valores maiores do que o observado na prática, ou seja, levemente possui mais erros positivos do que negativos. Enquanto para a metodologia de Box-Jenkins, a maior ocorrência foi no intervalo de -0,5 a -0,01%, e repete o comportamento de se ter mais erros positivos do que negativos.

Para uma análise mais detalhada dos erros, foi calculada a média dos valores, possibilitando a comparação de qual método de previsão foi o mais adequado para cada empresa. O gráfico das médias dos erros é exibido na Figura 39.

Figura 39 – Média de erros por empresa para previsão do restante da semana



Fonte: Elaborado pela autora (2017)

Com o gráfico da Figura 39, foi possível observar que as médias dos erros dos dois métodos de previsão ficaram parecidos para todas empresas, ou seja, os métodos calcularam valores de previsão similares.

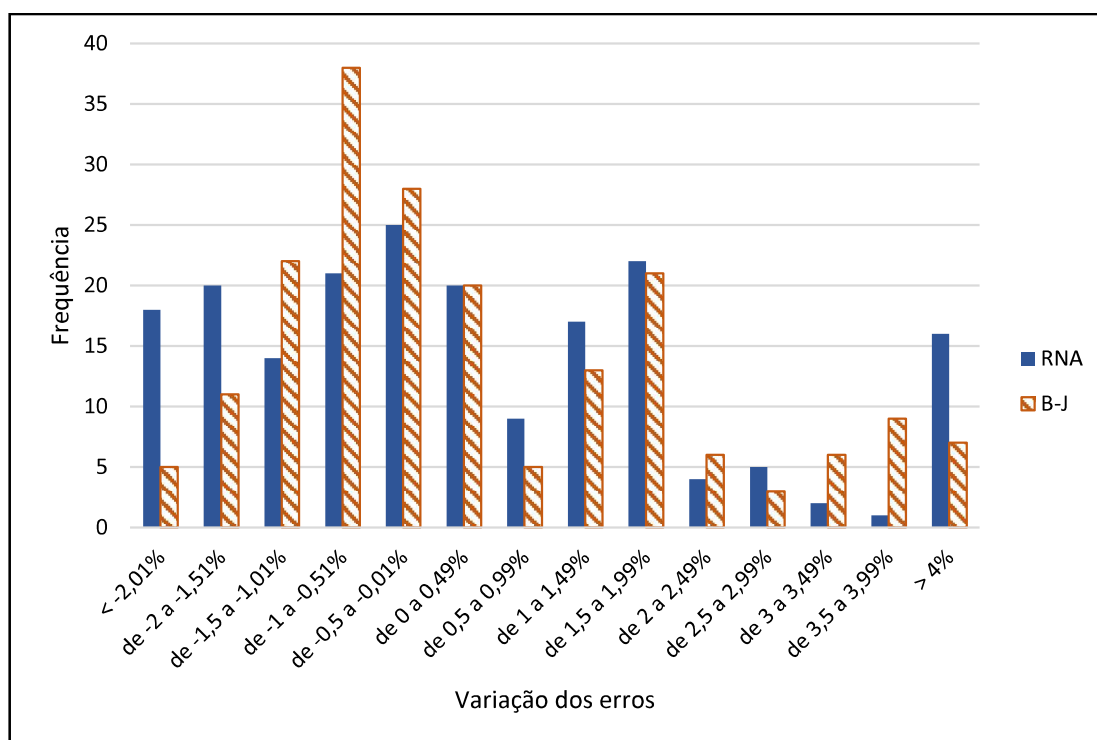
Os gráficos do histograma e da média de erros para cada empresa ficaram similares entre a análise com toda a base de dados históricos e a da previsão para o restante da semana, comprovando, assim, a reprodutibilidade dos dois métodos.

4.6.2.1 Considerações sobre os métodos de previsão

Neste trabalho, foram realizados 194 cálculos de previsão para cada método, em que consistem nas 130 prognoses com diferentes quantidades de dados históricos e mais as 64 prognoses do restante da semana, que são referentes as 16 empresas para 4 dias; totalizando 388 cálculos de previsão. Assim, a fim de poder analisar de maneira geral todos estes erros apresentados anteriormente, foi feito um histograma geral, conforme Figura 40.

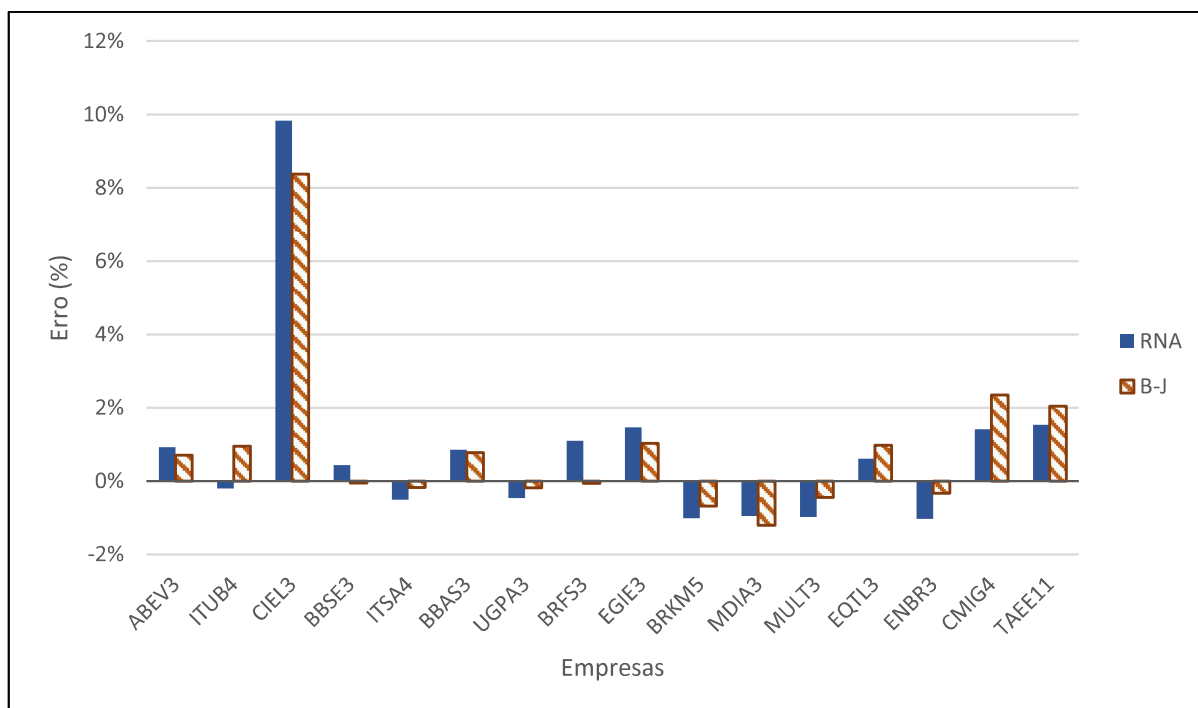
Para analisar todos os erros em relação as empresas, foi feito o gráfico das médias dos erros, exibido na Figura 41.

Figura 40 – Histograma de todos os erros



Fonte: Elaborado pela autora (2017)

Figura 41 – Média de todos os erros para cada empresa



Fonte: Elaborado pela autora (2017)

Ainda, as empresas presentes na carteira otimizada foram classificadas de acordo com seus setores de atuação, para examinar o comportamento dos erros em relação a estes segmentos. Esta classificação foi apresentada na Tabela 16.

Tabela 16 – Classificação das empresas

Setor	Empresas
Consumo e varejo	AMBEV BRF M. DIAS BRANCO
Energia e saneamento	ENGIE BRASIL ENERGIA EQUATORIAL ENERGIA EDP ENERGIAS DO BRASIL CIA ENERGÉTICA DE MINAS GERAIS TAESA
Financeiro	ITAU UNIBANCO CIELO BANCO DO BRASIL
Holding	ITAUSA
Imobiliário e construção	MULTIPLAN
Petroquímico	ULTRAPAR BRASKEM
Seguros	BB SEGURIDADE

Fonte: Elaborado pela autora (2017)

Estas informações das classificações dos respectivos setores foram retiradas do endereço eletrônico InfoMoney.

Então, utilizando os dados estudados neste trabalho, foi possível desenvolver uma tabela para comparar os dois métodos de previsão. Esta comparação foi apresentada na Tabela 17, e nela possui uma marcação do melhor resultado para cada uma das análises realizadas.

Tabela 17 – Comparativo dos métodos de previsão

(continua)		
	RNA	Box-Jenkins
Média geral do MSE	0,710	0,589 ✓
Média geral do r^2	0,991 ✓	0,970
Frequência de 0,00% de erro	3 ✓	1
% de previsões com erro dentro do intervalo de -0,5 a -0,01%	12,9%	14,4% ✓

	(conclusão)	
	RNA	Box-Jenkins
% de previsões com erro dentro do intervalo de 0,01 a 0,5%	10,3% ✓	10,3% ✓
Intervalo com maior frequência de erros	-0,5 a 0% ✓	-1 a -0,5%
Média geral de erros	0,81%	0,78% ✓
Média dos erros do setor Consumo e varejo	0,35%	-0,19% ✓
Média dos erros do setor Energia e saneamento	0,80% ✓	1,21%
Média dos erros do setor Financeiro	3,49%	3,36% ✓
Média dos erros do setor Holding	-0,51%	-0,17% ✓
Média dos erros do setor Imobiliário e construção	-0,98%	-0,45% ✓
Média dos erros do setor Petroquímico	-0,74%	-0,44% ✓
Média dos erros do setor Seguros	0,43%	-0,06% ✓

Fonte: Elaborado pela autora (2017)

Para o erro quadrático médio geral e o coeficiente de determinação geral, foi calculada a média de todos 194 cálculos de previsão realizados para cada método, assim como para o percentual médio de erros. De todos os resultados, foi contabilizado a frequência que ocorreu 0,00% de erro, quando comparado os valores das ações até a terceira casa decimal. As análises dos intervalos de erros basearam-se no histograma geral apresentado anteriormente.

Com a classificação das empresas, foi possível analisar a média dos erros de cada setor. Para esta análise, foi inicialmente calculada a média de erros para cada empresa, para todos 194 cálculos de cada um dos métodos. Em seguida, foi calculada a média para os setores, de acordo com a classificação apresentada na Tabela 16.

O MSE trata do erro, por isso, quanto menor o seu valor, melhor, por esta razão a metodologia Box-Jenkins obteve um melhor desempenho. Já para o coeficiente de determinação, é melhor que este valor esteja mais próximo do 1, assim as RNA possuíram uma melhor performance. Para análise dos erros gerais, as RNA possuem melhor resultado para a frequência de 0,00% de erro, a metodologia B-J apresentou um melhor resultado no comparativo para o percentual de previsões com erro dentro do intervalo de -0,5 a -0,01% por ter mais erros próximos do zero, enquanto houve um empate dos dois métodos para o percentual de previsões com erro dentro do intervalo de 0,01 a 0,5%.

Para o intervalo com a maior frequência de erros, as RNA obtiveram melhor desempenho, devido ao seu intervalo estar mais próximo de 0% de erro, em contrapartida, a média geral dos erros da metodologia B-J foi menor. Para a análise de média de erro por setor, o método que obteve melhor performance foi o que mais se aproximou de 0 % de erro.

5 CONCLUSÃO

Este capítulo final foi dividido em duas partes: as considerações finais e sugestões de trabalhos futuros, conforme apresentados a seguir.

5.1 CONSIDERAÇÕES FINAIS

Este trabalho foi composto por dois estudos: criação de um portfólio otimizado e análise de previsão de séries temporais. O primeiro foi com a utilização de Análise Envoltória de Dados e empresas presentes na bolsa de valores brasileira que participaram da pesquisa. Já para a análise de previsão, foi realizado um comparativo entre Redes Neurais Artificiais e metodologia Box-Jenkins, calculando a prognose de preço de ações das empresas participantes da carteira otimizada.

Com o uso destes três conceitos, foi possível perceber algumas vantagens e desvantagens. Para DEA, podem-se citar como algumas vantagens: consegue calcular com múltiplas entradas e saídas, podendo utilizar qualquer medida para estes dados; as DMUs ineficientes podem ser analisadas e quantificadas para cada unidade avaliada, realizando um *benchmark* destas a fim de melhorá-las e torná-las eficientes. Já algumas desvantagens foram: os resultados são sensíveis a entradas e/ou saída de dados e de DMUs; não é possível calcular com números negativos devido a restrições matemáticas.

Então, com o comparativo do retorno entre a carteira otimizada e o índice Ibovespa, apresentado na Figura 33, foi possível perceber o ganho em utilizar DEA para seleção de portfólio. Enquanto a Ibovespa obteve um retorno de 0,028%, a carteira otimizada alcançou 0,423%, representando um ganho 15 vezes maior. Assim, a aplicação da DEA foi satisfatória para o objetivo deste trabalho.

Já para a utilização das RNA, algumas vantagens foram: têm a capacidade de aprender e modelar relações entre entradas e saídas não-lineares e complexos, como ocorre em exemplos reais; pode aproximar graficamente qualquer função, independentemente da sua linearidade; não há restrições às variáveis de entrada, como por exemplo a forma como elas devem ser distribuídas; é possível utilizar com grandes quantidades de dados de entrada. Em contrapartida, foram observadas algumas desvantagens também: pode haver uma grande quantidade de treinamento, necessitando elevado custo computacional para processamento, havendo uma

dependência de *hardware*; pode ocorrer *overfitting*; não existe uma regra específica para determinar a estrutura das RNA, uma arquitetura apropriada é conseguida através da experiência e do teste e erro.

Para a metodologia Box-Jenkins, podem-se citar algumas vantagens: um *software* pode estimar os parâmetros; pode ser usado para modelar séries temporárias estacionárias e não estacionárias com e sem componente sazonal; não são necessárias muitas observações passadas. Quanto às desvantagens, podem-se mencionar: dificuldade de estimar os parâmetros, analisando somente os correlogramas; não há recurso de atualização automática, ou seja, à medida que novos dados se tornam disponíveis, todo o procedimento de modelagem deve ser repetido.

Por meio da análise dos resultados, os dois métodos de previsão se mostraram satisfatórios para a aplicabilidade escolhida neste trabalho, mas, após analisar diversos pontos, conforme exibido na Tabela 17, foi possível notar que a metodologia Box-Jenkins foi a mais indicada para prever preço de ações, quando comparada com RNA. Este método obteve menores médias de erros, apesar de que as redes neurais tiveram uma maior frequência de 0,00% de erro, isso é devido ao fato que, no histograma geral, os erros da RNA ficaram mais distribuídos no gráfico, enquanto os da metodologia B-J ficaram mais próximos de 0%, diminuindo assim o valor da média. Para cálculos de prognose para empresas do setor Energia e saneamento, foi recomendado o uso de RNA, por ter uma menor média de erro; enquanto Box-Jenkins foi aconselhável para os demais setores analisados neste trabalho. Ainda, foi possível mostrar por meio de análises quantitativas que não é necessária uma grande quantidade de dados históricos para calcular previsão, por isso as prognoses do restante da semana foram realizadas com o equivalente a 12 meses de dados históricos.

Os objetivos previamente estipulados foram satisfatoriamente alcançados, com o estudo teórico dos conceitos apresentados no Capítulo 2, a descrição de como ocorreram as etapas no Capítulo 3 e as suas aplicações relatadas no Capítulo 4. Com este trabalho, foi evidenciada a importância do uso de técnicas corretas para o apoio a tomada de decisão em relação a investimentos na bolsa de valores, além também do auxílio que os *softwares* matemáticos proporcionaram para estes dois exemplos de análises matemática, ou seja, seleção de portfólio otimizado e métodos de previsão.

5.2 SUGESTÕES DE TRABALHOS FUTUROS

As metodologias utilizadas apresentaram resultados satisfatórios, porém podem ser complementadas. Algumas sugestões podem ser citadas para trabalhos futuros, como por exemplo:

- a) estudar o comportamento dos preços das ações, como são influenciados pelo mercado, eventos históricos, questões políticas e econômicas, entre outros fatores;
- b) estabelecer uma relação correlacional entre dólar e o índice Ibovespa;
- c) para a análise das RNA, calcular a previsão com diferentes parâmetros determinados neste trabalho, a fim de determinar qual a quantidade de camadas mais adequada, por exemplo; e
- d) comparar os resultados obtidos das prognoses com outros métodos de previsão de séries temporais.

REFERÊNCIA BIBLIOGRÁFICA

ANDRADE, Eduardo Leopoldino de. **Introdução à Pesquisa Operacional**. 2. ed. Rio de Janeiro: LTC, 1998.

ANDRADE, Livia Naiara de. **Redes neurais artificiais aplicadas na identificação automática de áreas cafeeiras em imagens de satélite**. 2011, 92 f. Dissertação (Mestrado em Ciência da Computação) – Universidade Federal de Minas Gerais, Belo Horizonte, 2011. Disponível em:

<<http://www.bibliotecadigital.ufmg.br/dspace/bitstream/handle/1843/SLSS-8GPR28/livianaiaaraandade.pdf?sequence=1>>. Acesso em: 2 dez. 2017.

BALLOU, Ronald. **Gerenciamento da cadeia de suprimentos: planejamento, organização e logística empresarial**. 4. ed. Porto Alegre: Bookman, 2001.

BANKER, Rajiv D.; CHARNES, Abraham; COOPER, William Wager. Some models for estimating technical and scale inefficiencies in data envelopment analysis.

Management Science, v. 30, n. 9, p. 1078-1092, 1984. Disponível em:

<[http://psyso2012.khuisf.ac.ir/Prof/images/Uploaded_files/BCC%201984\[4319189\].PDF](http://psyso2012.khuisf.ac.ir/Prof/images/Uploaded_files/BCC%201984[4319189].PDF)>. Acesso em: 15 jun. 2016.

BATISTA, Fabiana Dias. **Metodologia para o uso da Análise Envoltória de Dados no auxílio à decisão**. 2009, 107 f. Dissertação (Mestrado em Ciências em Engenharia de Produção) – Universidade Federal de Itajubá, Itajubá, 2009.

Disponível em: <<http://saturno.unifei.edu.br/bim/0034807.pdf>>. Acesso em: 20 nov. 2016.

BM&FBovespa. Bolsa de Valores, Mercadorias e Futuros de São Paulo. **Site institucional**. Disponível em: <http://www.bmfbovespa.com.br/pt_br/>. Acesso em: 20 jun. 2016.

BRESSAN, Aureliano Angel. **Modelos de previsão de preços aplicados aos contratos futuros agropecuários**. 2001, 152 f. Tese (Doutorado em Economia Rural) – Universidade Federal de Viçosa, Viçosa, 2001. Disponível em:

<<http://www.sbicafe.ufv.br/handle/123456789/375>>. Acesso em: 15 mar. 2017.

BUSSAB, Wilton; MORETTIN, Pedro. **Estatística básica**. 8ª. ed. São Paulo: Saraiva, 2013.

CAMPOS, Gabriel Moreira; SCHERER, Luciano Márcio. Lucro por ação. **Revista Contabilidade & Finanças**, v. 12, n. 26, p. 81-94, 2001. Disponível em:

<<http://www.scielo.br/pdf/rcf/v12n26/v12n26a05.pdf>>. Acesso em: 20 jun. 2016.

CHARNES, Abraham; COOPER, William W.; RHODES, Edwardo. Measuring the efficiency of decision making units. **European journal of operational research**, v. 2, n. 6, p. 429-444, 1978. Disponível em: <http://ac.els-cdn.com/0377221778901388/1-s2.0-0377221778901388-main.pdf?_tid=0af190ee-2dc4-11e6-a06e-00000aab0f6b&acdnat=1465423250_a200c4f7a95a7612aa519476e3f54e0c>.

Acesso em: 15 jun. 2016.

COLIN, Emerson Carlos. **Pesquisa Operacional**: 170 aplicações em estratégia, finanças, logística, produção, marketing e vendas. Rio de Janeiro: LTC, 2007.

DARTORA, Gery Antônio. **Redes neurais artificiais**: aplicação na tipificação de carcaça de suínos. 2003, 105 f. Dissertação (Mestrado em Ciência da Computação) – Universidade Federal de Santa Catarina, Florianópolis, 2003. Disponível em: <<https://repositorio.ufsc.br/xmlui/bitstream/handle/123456789/84537/195587.pdf?sequence=1&isAllowed=y>>. Acesso em: 16 out. 2016.

DEVADOSS, A. Victor; LIGORI, T. Antony Alphonse. Stock prediction using artificial neural networks. **International Journal of Data Mining Techniques and Applications**, v. 2, p. 283-291, 2013. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=18E1B398DB224DF4E15BEAE7EBD17DD7?doi=10.1.1.678.6606&rep=rep1&type=pdf>>. Acesso em: 15 jun. 2017.

DUMITRU, Cristian-Dragos; GLIGOR, Adrian. Daily Average Wind Energy Forecasting Using Artificial Neural Networks. **Procedia Engineering**, v. 181, p. 829-836, 2017. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1877705817310597>>. Acesso em: 19 set. 2017.

Exame – Mercados. **Site institucional**. Disponível em: <<http://exame.abril.com.br/mercados/cotacoes-bovespa>>. Acesso em: 20 jun. 2016.

FISCHER, Sergio. **Séries univariantes de tempo** – metodologia Box-Jenkins. Porto Alegre: FEE, 1982.

FREITAS, Sander Oliveira de. **Utilização de um modelo baseado em redes neurais para a precificação de opções**. 2001, 103 f. Dissertação (Mestrado em Administração) – Universidade Federal de Minas Gerais, Belo Horizonte, 2001. Disponível em: <http://www.bibliotecadigital.ufmg.br/dspace/bitstream/handle/1843/BUBD-9BHGGH/disserta__o_sander_oliveira_de_freitas.pdf?sequence=1>. Acesso em: 2 fev. 2017.

FREUND, John E. **Estatística aplicada**: economia, administração e contabilidade. 11. ed. Porto Alegre: Bookman, 2006.

GAIO, Luiz Eduardo; CASTRO JÚNIOR, Luiz Gonzaga de; OLIVEIRA, André Ribeiro de. Previsão de preço futuro do boi gordo na BM&F: uma comparação entre modelos de séries temporais e redes neurais. **Organizações Rurais & Agroindustriais**, v. 9, n. 2, 2007. Disponível em: <<http://revista.dae.ufla.br/index.php/ora/article/view/117>>. Acesso em: 7 jan. 2017.

GALVÃO, Paulo José Lopes Normande. **Análise Envoltória de Dados aplicada ao setor brasileiro de distribuição de energia elétrica**. 2008, 109 f. Dissertação (Mestrado profissionalizante em Administração) – Faculdade de Economia e Finanças IBMEC, Rio de Janeiro, 2008. Disponível em:

<http://www2.aneel.gov.br/biblioteca/trabalhos/trabalhos/Dissertacao_Paulo_Galvao.pdf>. Acesso em: 20 nov. 2016.

GIACOMEL, Felipe dos Santos. **Um método algorítmico para operações na bolsa de valores em ensembles de Redes Neurais para modelar e prever os movimentos dos mercados de ações**. 2016, 92 f. Dissertação (Mestrado em Ciência da Computação) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2016. Disponível em: <<http://www.lume.ufrgs.br/bitstream/handle/10183/134586/000988403.pdf?sequence=1>>. Acesso em: 7 dez. 2016.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. 5. ed. São Paulo: Atlas, 2010.

GÖKŞEN, Yılmaz; DOĞAN, Onur; ÖZKARABACAK, Bilge. A Data Envelopment Analysis Application for Measuring Efficiency of University Departments. **Procedia Economics and Finance**, v. 19, p. 226-237, 2015. Disponível em: <http://ac.els-cdn.com/S2212567115000246/1-s2.0-S2212567115000246-main.pdf?_tid=07b82e3a-4514-11e6-8318-00000aab0f01&acdnat=1467986481_0188338e9e42c0e59f58960dcc66febf>. Acesso em: 8 jul. 2016.

GUIMARÃES, Lourenço da Rocha. **Previsão de inadimplência e redes neurais artificiais**. 2006, 85 f. Dissertação (Mestrado em Modelagem Computacional) – Universidade do Estado do Rio de Janeiro, Nova Friburgo, 2006. Disponível em: <http://www.bdtd.uerj.br/tde_busca/arquivo.php?codArquivo=383>. Acesso em: 19 set. 2017.

GUO, Xiaoying; LU, Ching-Cheng; LEE, Jen-Hui; CHIU, Yung-Ho. Applying the dynamic DEA model to evaluate the energy efficiency of OECD countries and China. **Energy**, v. 134 p. 392-399. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0360544217310344>>. Acesso em: 2 dez. 2017.

HAYKIN, Simon. **Redes Neurais: princípios e prática**. 2. ed. Porto Alegre: Bookman, 2001.

HAZZAN, Samuel. **Desempenho de ações na Bolsa de Valores de São Paulo e sua relação com o índice preço-lucro**. 1991, 270 f. Tese (Doutorado em Administração) – Escola de Administração de Empresas de São Paulo da Fundação Getúlio Vargas, São Paulo, 1991. Disponível em: <http://bibliotecadigital.fgv.br/dspace/bitstream/handle/10438/4558/HAZZAN_Samuel%28tese%29_OCR.pdf?sequence=1&isAllowed=y>. Acesso em: 20 jun. 2016.

InfoMoney. **Site institucional**. Disponível em: <<http://www.infomoney.com.br/>>. Acesso em: 7 dez. 2017.

JEBALI, Eya; ESSID, Hédi; KHRAIEF, Naceur. The analysis of energy efficiency of the Mediterranean countries: a two-stage double bootstrap DEA approach. **Energy**, v. 134, p. 991-1000, 2017. Disponível em:

<<http://www.sciencedirect.com/science/article/pii/S0360544217310575>>. Acesso em: 3 jul. 2017.

KIM, Sungki; KO, Wonil; NAM, Hyoon; KIM, Chulmin; CHUNG, Yangdon; BANG, Sungsig. Statistical model for forecasting uranium prices to estimate the nuclear fuel cycle cost. **Nuclear Engineering and Technology**, v. 49, p. 1063-1070, 2017.

Disponível em:

<<https://www.sciencedirect.com/science/article/pii/S1738573317300992>>. Acesso em: 21 out. 2017.

KIRCHNER, Rosane Maria. **Modelagem e previsão por meio de metodologia Box-Jenkins: uma ferramenta de gestão**. 2006, 64 f. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal de Santa Maria, Santa Maria, 2006. Disponível em:

<http://cascavel.ufsm.br/tede//tde_busca/arquivo.php?codArquivo=922>. Acesso em: 21 fev. 2017.

KRISHNAN, N.M. Anoop; MANGALATHU, Sujith; SMEDSKJAER, Morten M.; TANDIA, Adama; BURTON, Henry; BAUCHY, Mathieu. Predicting the dissolution kinetics of silicate glasses using machine learning. **Journal of Non-Crystalline Solids**, v. 487, p. 37-45, 2018. Disponível em:

<<https://www.sciencedirect.com/science/article/pii/S0022309318300905>>. Acesso em: 10 mar. 2018.

LAPPONI, Juan Carlos. **Estatística usando Excel**. Rio de Janeiro: Elsevier, 2005.

LEMOES, Fernando de Oliveira. **Metodologia para seleção de métodos de previsão de demanda**. 2006, 183 f. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2006.

Disponível em:

<<http://www.producao.ufrgs.br/arquivos/publicacoes/fernandooliveiralemos.pdf>>. Acesso em: 15 fev. 2017.

LEVINE, David; STEPHAN, David; KREHBIEL, Timothy; BERENSON, Mark.

Estatística – Teoria e aplicações usando o Microsoft® Excel em português. Rio de Janeiro: LTC, 2012.

LIM, Sungmook; OH, Kwang Wuk; ZHU, Joe. Use of DEA cross-efficiency evaluation in portfolio selection: An application to Korean stock market. **European Journal of Operational Research**, v. 236, n. 1, p. 361-368, 2014. Disponível em:

<<http://www.sciencedirect.com/science/article/pii/S0377221713009776>>. Acesso em: 3 jul. 2017.

LOPES, Ana Lúcia Miranda; LIMA, Marcus Vinicius Andrade; DUTRA, Ademar; SAURIN, Valter. Data Envelopment Analysis – DEA como Estratégia para Seleção de Carteiras de Investimento: uma aplicação a 13 anos do mercado de ações brasileiro. **Encontro da Associação Nacional de Pós-Graduação e Pesquisa em Administração**, v. 32, p. 1-16, 2008. Disponível em:

<<http://pesquisas.face.ufmg.br/nipe-log/wp-content/uploads/sites/4/2015/11/Artigo28.pdf>>. Acesso em: 1 jul. 2016.

MALACRIDA, Mara Jane Contrera; YAMAMOTO, Marina Mitiyo. Governança corporativa: nível de evidenciação das informações e sua relação com a volatilidade das ações do Ibovespa. **Revista contabilidade e finanças**, v. 17, p. 65-79, 2006. Disponível em: <<http://www.scielo.br/pdf/rcf/v17nspe/v17nspea06>>. Acesso em: 11 jul. 2016.

MARDANI, Abbas; ZAVADSKAS, Edmundas Kazimieras; STREIMIKIENE, Dalia; JUSOH, Ahmad; KHOSHNOUDI, Masoumeh. A comprehensive review of data envelopment analysis (DEA) approach in energy efficiency. **Renewable and Sustainable Energy Reviews**, v. 70, p. 1298-1322, 2017. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1364032116310875>>. Acesso em: 19 set. 2017.

MARIANO, Enzo B.; ALMEIDA, Mariana R.; REBELATTO, Daisy AN. Peculiaridades da análise por envoltória de dados. **Anais do XII SIMPEP**, 2006a. Disponível em: <https://www.researchgate.net/profile/Enzo_Mariano/publication/240620915_Peculiaridades_da_Analise_por_Envoltoria_de_Dados/links/0deec5252fadb1249b000000.pdf>. Acesso em: 21 jul. 2016.

_____. Princípios Básicos para uma proposta de ensino sobre análise por envoltória de dados. In: COBENGE, 34 (Ed.). **XXXIV CONGRESSO BRASILEIRO DE ENSINO DE ENGENHARIA (COBENGE 2006), Universidade de Passo Fundo–UPF**. 2006b. Disponível em: <https://www.researchgate.net/profile/Enzo_Mariano/publication/257409786_Principios_basicos_para_uma_proposta_de_ensino_sobre_analise_por_envoltoria_de_dados/links/00463525379e8e2fbc000000.pdf>. Acesso em: 21 jul. 2016.

MARTINS, Vanessa de Quadros. **Eficiência econômica em empresas distribuidoras de energia elétrica: um estudo com base na análise envoltória de dados**. 2014, 86 f. Dissertação (Mestrado em Ciências Contábeis) – Universidade do Vale do Rio dos Sinos, UNISINOS, São Leopoldo, 2014. Disponível em: <<http://www.repositorio.jesuita.org.br/handle/UNISINOS/3770>>. Acesso em: 24 ago. 2016.

MELLO, João Carlos C. B. S. de; MEZA, Lidia Angulo; GOMES, Eliane Gonçalves; NETO, Luiz Biondi. Curso de Análise Envoltória de Dados. **Simpósio Brasileiro de Pesquisa Operacional**, v. 37, p. 2521-2547, 2005. Disponível em: <http://www.uff.br/decisao/sbpo2005_curso.pdf>. Acesso em: 20 jun. 2016.

MOGHADDAM, Amin Hedayati; MOGHADDAM, Moein Hedayati; ESFANDYARI, Morteza. Stock market index prediction using artificial neural network. **Journal of Economics, Finance and Administrative Science**, v. 21, n. 41, p. 89-93, 2016. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2077188616300245>>. Acesso em: 19 set. 2017.

MONTGOMERY, Douglas; RUNGER, George. **Estatística aplicada e probabilidade para engenheiros**. 6ª. ed. Rio de Janeiro: LTC, 2016.

MULERO, Ángel; PIERANTOZZI, Mariano; CACHADIÑA, Isidro; DI NICOLA, Giovanni. An Artificial Neural Network for the surface tension of alcohols. **Fluid Phase Equilibria**, v. 449, p. 28-40, 2017. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0378381217302248>>. Acesso em: 2 dez. 2017.

PELLEGRINI, Fernando Rezende. **Metodologia para implementação de sistemas de previsão de demanda**. 2000, 146 f. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2000. Disponível em: <<http://www.producao.ufrgs.br/arquivos/publicacoes/Fernando%20R%20Pellegrini.pdf>>. Acesso em: 15 fev. 2017.

PUTRO, Sapto Purnomo; KOSHIO, Sunshuke; OKTAFERDIAN, Vivin. Implementation of ARIMA Model to Asses Seasonal Variability Macrobenthic Assemblages. **Aquatic Procedia**, vol. 7, p. 277-284, 2016. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2214241X16300608>>. Acesso em: 21 out. 2017.

QUEIRÓS, Emerson Oliveira de. **Modelo de previsão para receita tributária estadual: aplicação do modelo de Box-Jenkins**. 2012, 101 f. Dissertação (Mestrado em Economia Política) – Pontifícia Universidade Católica de São Paulo, São Paulo, 2012. Disponível em: <<https://tede2.pucsp.br/handle/handle/9196>>. Acesso em: 15 fev. 2017.

RAFAELI, Leonardo. **A análise envoltória de dados como ferramenta para avaliação de desempenho relativo**. 2009, 166 f. Dissertação (Mestrado Acadêmico em Engenharia de Produção) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2009. Disponível em: <<http://www.lume.ufrgs.br/bitstream/handle/10183/16477/000701359.pdf?sequence=1>>. Acesso em: 24 ago. 2016.

RAIA JUNIOR, Archimedes Azevedo. **Acessibilidade e mobilidade na estimativa de um índice de potencial de viagens utilizando redes neurais artificiais e sistemas de informações geográficas**. 2000, 212 f. Tese (Doutorado em Engenharia Civil) – Escola de Engenharia de São Carlos, São Carlos, 2000. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/18/18137/tde-10112001-160812/en.php>>. Acesso em: 6 dez. 2016.

RAMOS, Patrícia; SANTOS, Nicolau; REBELO, Rui. Performance of state space and ARIMA models for consumer retail sales forecasting. **Robotics and computer-integrated manufacturing**, v. 34, p. 151-163, 2015. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0736584515000137>>. Acesso em: 26 jul. 2017.

RICH, Elaine; KNIGHT, Kevin; NAIR, Shivashankar. **Artificial intelligence**. 3. ed. Nova Deli: Tata McGraw-Hill Publishing Company Limited, 2009.

ROSANO-PEÑA, Carlos. Eficiência e impacto do contexto na gestão através do DEA: o caso da UEG. **Produção**, São Paulo, v. 22, n. 4, p. 778-787, 2012.

Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-65132012000400010>. Acesso em: 26 jul. 2017.

ROSSI, Daniela Jardim. **Previsão da velocidade dos ventos por Redes Neurais Artificiais e ARIMA de Box-Jenkins**. 2013, 76 f. Dissertação (Mestrado em Engenharia Elétrica) – Universidade Estadual Paulista Júlio de Mesquita Filho, Ilha Solteira, 2013. Disponível em: <<http://repositorio.unesp.br/handle/11449/111121>>. Acesso em: 28 fev. 2017.

ROTELA JUNIOR, Paulo; PAMPLONA, Edson de Oliveira; SALOMON, Fernando Luiz Riêra. Otimização de Portfólios: Análise de Eficiência. **Revista de Administração de Empresas**, v. 54, n. 4, p. 405-413, 2014. Disponível em: <<http://www.scielo.br/pdf/rae/v54n4/0034-7590-rae-54-04-00405.pdf>>. Acesso em: 14 maio 2016.

SEN, Parag; ROY, Mousumi; PAL, Parimal. Application of ARIMA for forecasting energy consumption and GHG emission: A case study of an Indian pig iron manufacturing organization. **Energy**, v. 116, p. 1031-1038, 2016. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0360544216315043>>. Acesso em: 26 jul. 2017.

SHARAFI, Mehdi; GHAEM, Haleh; TABATABAEE, Hamid Reza; FARAMARZI, Hossein. Forecasting the number of zoonotic cutaneous leishmaniasis cases in south of Fars province, Iran using seasonal ARIMA time series method. **Asian Pacific Journal of Tropical Medicine**, v. 10, p. 79-86, 2017. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1995764516307374>>. Acesso em: 21 out. 2017.

SOUZA, Rodrigo Clemente Thom de. **Previsão de séries temporais utilizando Rede Neural treinada por filtro de Kalman e evolução diferencial**. 2008, 98f. Dissertação (Mestrado em Engenharia de Produção e Sistemas) – Pontifícia Universidade Católica do Paraná, Curitiba, 2008. Disponível em: <<https://docs.ufpr.br/~thom/papers/dissertation.pdf>>. Acesso em: 7 jan. 2017.

TANEJA, Kanika; AHMAD, Shamshad; AHMAD Kafeel; ATTRI, S.D. Time series analysis of aerosol optical depth over New Delhi using Box–Jenkins ARIMA modeling approach. **Atmospheric Pollution Research**, v. 7, n. 4, p. 585-596, 2016. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1309104215301082>>. Acesso em: 26 jul. 2017.

THAKUR, Gour Sundar Mitra; BHATTACHARYYA, Rupak; MONDAL, Seema Sarkar. Artificial Neural Network Based Model for Forecasting of Inflation in India. **Fuzzy Information and Engineering**, v. 8, n. 1, p. 87-100, 2016. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1616865816300061>>. Acesso em: 19 set. 2017.

Universidade de Caxias do Sul. **Site institucional**. Disponível em: <<http://www.ucs.br/site>>. Acesso em: 20 jun. 2016.

WEI, William Wu-Shyong. **Time series analysis**. 2. ed. Nova Iorque: Pearson, 2005.

WERNER, Liane. **Um modelo composto para realização de previsão de demanda através da integração da combinação de previsões e do ajuste baseado na opinião**. 2004, 166 f. Tese (Doutorado em Engenharia de Produção) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2004. Disponível em: <<http://hdl.handle.net/10183/4189>>. Acesso em: 15 fev. 2017.

YUAN, Chaoqing; LIU, Sifeng; FANG, Zhigeng. Comparison of China's primary energy consumption forecasting by using ARIMA (the autoregressive integrated moving average) model and GM(1,1) model. **Energy**, v. 100, p. 384-390, 2016. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0360544216300421>>. Acesso em: 21 out. 2017.

ZHOU, Peng; POH, Kim Leng; ANG, Beng Wah. A non-radial DEA approach to measuring environmental performance. **European Journal of Operational Research**, v. 178, n. 1, p. 1-9, 2007. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0377221706003407>>. Acesso em: 26 jul. 2017.

APÊNDICE A – SÍNTESE DOS TRABALHOS ESTUDADOS REFERENTES A ANÁLISE ENVOLTÓRIA DE DADOS

(continua)

Autor (ano)	Título do trabalho	Aplicação	Tipo do modelo utilizado	Input	Output	Principais conclusões sobre a aplicação
Batista (2009)	Metodologia para uso da Análise Envoltória de Dados no auxílio à decisão	Estudo feito em uma empresa multinacional de grande porte do setor automotivo, formada por 14 unidades localizadas em diferentes países.	Modelo BCC, com orientação por <i>input</i>	- Quantidade de peças vendidas - Estoque (em dólares) - Estoque (em dias) - Investimentos - Capital empregado - Funcionários - Custos com pessoal nas vendas	- Vendas - Resultado de vendas - Rentabilidade de vendas - Retorno do capital empregado - Vendas por funcionário	Apesar do potencial de utilização da DEA ser muito grande, o gestor que a utilizará deve estar atento a algumas particularidades para que haja credibilidade na análise.
Colin (2007)	Pesquisa Operacional: 170 aplicações em estratégia, finanças, logística, produção, marketing e vendas	Utilizou um exemplo teórico com três hospitais e aplicou DEA para saber qual a eficiência de cada um deles.	Modelo CCR, com orientação por <i>input</i>	- Capital investido - Mão-de-obra	- Pacientes jovens - Pacientes adultos - Pacientes idosos	O hospital 2 foi o único ineficiente, e calculou o seu <i>benchmark</i> para torná-lo eficiente.

(continuação)

Autor (ano)	Título do trabalho	Aplicação	Tipo do modelo utilizado	Input	Output	Principais conclusões sobre a aplicação
Galvão (2008)	Análise Envoltória de Dados aplicada ao setor brasileiro de distribuição de energia elétrica	Classificar as distribuidoras do setor elétrico brasileiro de forma eficiente.	Modelos CCR e BCC, ambos com orientação por <i>input</i>	- Número de postes instalados - Número de transformadores instalados - Extensão, em quilômetros, de todas linhas de transmissão - Total de despesas operacionais - Número de funcionários	- Área total, em quilômetros quadrados, coberta pela distribuidora - Número total de consumidores atendidos - Total da energia fornecida, em gigawatts por ano	Deve-se analisar os resultados obtidos com certa cautela, pois, às vezes, são atribuídos valores de pesos que não estejam de acordo com a realidade.

(continuação)

Autor (ano)	Título do trabalho	Aplicação	Tipo do modelo utilizado	Input	Output	Principais conclusões sobre a aplicação
Guo, Lu, Lee e Chiu (2017)	Aplicando o modelo dinâmico DEA para avaliar a eficiência energética dos países da OCDE e da China	Este estudo avalia a eficiência intertemporal para a eficiência executiva com base em emissões de CO2 de combustíveis fósseis nos países da OCDE e na China.	Modelo SBM dinâmico com orientação por <i>output</i>	- Área da terra - População - Uso de energia	- Emissões de CO2 - PIB	O índice de ajuste no estoque de energia mostra que a maioria dos países apresenta uma melhoria da eficiência. Esses 27 países devem, assim, aumentar o número de estoque de energia para melhorar sua eficiência.
Jebali, Essid e Khraief (2017)	A análise da eficiência energética dos países mediterrânicos: uma abordagem DEA de duas etapas de inicialização dupla	O objetivo principal deste trabalho é examinar os determinantes da eficiência energética nos países do Mediterrâneo durante os anos de 2009 a 2012.	Modelo BCC, ambos com orientação por <i>input</i>	- Força de trabalho - Formação bruta de capital fixo	- Produto Interno Bruto	Resultados empíricos mostram que os níveis de eficiência energética nos países mediterrânicos são altos e em declínio ao longo do tempo.

(continuação)

Autor (ano)	Título do trabalho	Aplicação	Tipo do modelo utilizado	Input	Output	Principais conclusões sobre a aplicação
Lim, Oh e Zhu (2014)	Uso da avaliação de eficiência cruzada da DEA em seleção de portfólio: uma aplicação na bolsa de valores coreana	É proposto uma maneira de usar a avaliação de eficiência cruzada DEA na seleção de portfólio. Embora a eficiência cruzada seja uma abordagem desenvolvida para a avaliação de pares, foi melhorado seu uso na seleção de portfólio.	Modelo BCC aditivo, com uma medida de ineficiência ajustada em escala	Diversos dados relacionados a: - Utilização de ativos - Liquidez - Influência econômica	Diversos dados relacionados a: - Rentabilidade - Crescimento	Foi demonstrado que a abordagem proposta pode ser uma ferramenta promissora para a seleção de portfólio de ações presentes na bolsa de valores coreana, mostrando que a carteira selecionada produz maiores rendimentos ajustados ao risco do que outras carteiras de referência para um período de amostra de 9 anos, de 2002 a 2011.
Lopes, Lima, Dutra e Saurin (2008)	DEA como Estratégia para Seleção de Carteiras de Investimento: uma aplicação a 13 anos do mercado de ações brasileiro	Seleção de carteira de investimentos de empresas presentes na Bolsa de Valores de São Paulo.	Modelos CCR e BCC, ambos com orientação por <i>input</i>	- Preço/lucro - Beta - Volatilidade	- Lucro por ação - Retorno de 12, 36 e 60 meses	Carteiras DEA obtiveram um desempenho superior quando comparadas com a Ibovespa, e que a mais indicada foi a carteira anual do modelo BCC, pois é a que teve o maior retorno.

(continuação)

Autor (ano)	Título do trabalho	Aplicação	Tipo do modelo utilizado	Input	Output	Principais conclusões sobre a aplicação
Martins (2014)	Eficiência econômica em empresas distribuidoras de energia elétrica: um estudo com base na Análise Envoltória de Dados	Busca a eficiência econômica das empresas distribuidoras de energia elétrica.	Modelo CCR, com orientação por <i>output</i>	- Ativo total - Ativo imobilizado - Total de empregados	- Lucro antes dos juros, impostos, depreciação e amortização - Resultado do período - Receita líquida	Das 18 empresas estudadas, apenas 7 foram consideradas eficientes, representando 38,8% da amostra; e fatores como localização da empresa, densidade populacional, tamanho da área de concessão e da extensão da rede não são relevantes ou determinantes para explicar a eficiência das empresas.
Rafaeli (2009)	A análise envoltória de dados como ferramenta para avaliação de desempenho relativo	Sugerir uma nova forma de avaliar o desempenho dos países do globo, considerando seu grau de desenvolvimento.	Modelo CCR, com orientação por <i>input</i>	- Despesas públicas com educação - Consumo de energia elétrica - Mais outros 11 <i>inputs</i>	- Taxa de desemprego - Expectativa de vida ao nascer - Mais outros 12 <i>outputs</i>	Demonstrar outros exemplos de aplicações de DEA além dos clássicos de comparação entre empresas ou de unidades internas de uma organização.

(continuação)

Autor (ano)	Título do trabalho	Aplicação	Tipo do modelo utilizado	Input	Output	Principais conclusões sobre a aplicação
Rosano- Peña (2012)	Eficiência e impacto do contexto na gestão através do DEA: o caso da UEG	Este trabalho mostra a contribuição do método DEA na avaliação da eficiência técnica de organizações que operam em contextos diferenciados, aplicado na Universidade Estadual de Goiás.	Modelo BCC, com orientação por <i>input</i>	- Número de funcionários - Número de professores - Área construída em m²	- Projetos de pesquisa registrados - Cursos de graduação - Cursos superiores de qualificação - Cursos de pós-graduação - Número de alunos regulares	A UEG poderia incrementar sua produção em 7%. A pesquisa também mostrou que das 20 unidades pesquisadas 11 estão subdimensionadas e 9 superdimensionadas. Essa informação poderia ser utilizada para reduzir custos e otimizar o porte dessas unidades. Após a otimização, foi utilizado também a abordagem de Sharpe e na comparação das carteiras, observa-se que a resultante da combinação de ambos os modelos apresentou melhor desempenho.
Rotela Junior, Pomplona e Salomon (2014)	Otimização de Portfólios: Análise de Eficiência	Construir uma carteira otimizada de ações de empresas de capital aberto na Bolsa de Valores de São Paulo.	Modelo BCC, com orientação por <i>input</i>	- Retorno de um, dois e três anos - Lucro por ação	- Beta - Preço/lucro - Volatilidade de um, dois e três anos	Após a otimização, foi utilizado também a abordagem de Sharpe e na comparação das carteiras, observa-se que a resultante da combinação de ambos os modelos apresentou melhor desempenho.

(conclusão)

Autor (ano)	Título do trabalho	Aplicação	Tipo do modelo utilizado	Input	Output	Principais conclusões sobre a aplicação
Zhou, Poh e Ang (2007)	Uma abordagem não- radial de DEA para medir desempenho ambiental	É apresentado uma abordagem não-radial de DEA para medir um índice de desempenho ambiental que modela a mudança de desempenho ambiental ao longo do tempo, aplicado a 26 países da OCDE.	Modelo baseado em DEA para medir o desempenho ambiental, com orientação por <i>output</i>	- Força de trabalho - Consumo de energia	- Produto Interno Bruto	Verificou-se que o desempenho ambiental dos países da OCDE como um todo foi melhorado de 1995 a 1997.

APÊNDICE B – SÍNTESE DOS TRABALHOS ESTUDADOS REFERENTES A RNA: PARÂMETROS ESCOLHIDOS

(continua)

Autor (ano)	Título do trabalho	Aplicação	Topologia	Algoritmo de aprendizagem
Andrade (2011)	Redes Neurais Artificiais aplicadas na identificação automática de áreas cafeeiras em imagens de satélite	Aplicação de Redes Neurais Artificiais para a classificação automática de dados de sensoriamento remoto, que fosse adequada para a identificação e mapeamento do uso e ocupação das terras, com ênfase na identificação de 3 áreas cafeeiras de Minas Gerais.	<i>Multilayer Perceptron</i> (MLP): - 1 neurônio na camada de entrada, 14 na intermediária e 1 na de saída, - 1 neurônio na camada de entrada, 18 na intermediária e 5 na de saída, e - 1 neurônio na camada de entrada, 18 na intermediária e 1 na de saída	<i>Backpropagation</i>
Bressan (2001)	Modelos de previsão de preços aplicados aos contratos futuros agropecuários	Aplicação de modelos de previsão como ferramenta de decisão de compra e venda de ações da BM&F de <i>commodities</i> boi gordo, café e soja.	Não foi especificado.	Não foi especificado.

(continuação)

Autor (ano)	Título do trabalho	Aplicação	Topologia	Algoritmo de aprendizagem
Dartora (2003)	Redes neurais artificiais: aplicação na tipificação de carcaça de suínos	Aplicar RNA na tipificação de carcaça de suínos, e comparar com o resultado de outras técnicas para analisar qual teve o melhor desempenho.	<i>Multilayer Perceptron</i> (MLP) totalmente conectada, com 1 camada escondida, sendo 3 neurônios na camada de entrada, 2 na intermediária e 1 na de saída	Método de retropropagação do erro com momento
Devadoss e Ligor (2013)	Predições de ações com o uso de Redes Neurais Artificiais	Aplicar RNA na previsão de preço de ações de quatro empresas selecionadas que operam na Bolsa de Valores de Bombaim, localizada na Índia.	<i>Multilayer Perceptron</i> (MLP) com 5 neurônios na camada de entrada, 11 na intermediária e 1 na de saída	<i>Backpropagation</i>
Dumitru e Gligor (2017)	Previsão média diária de energia eólica usando redes de neurônios artificiais	Este artigo apresenta uma proposta relacionada à previsão do poder eólico usando redes neurais artificiais. A solução proposta é investigada para estabelecer a configuração ideal e os desempenhos para o Sudeste da Europa.	<i>Multilayer Perceptron</i> (MLP) com 1 neurônio na camada de entrada, 15 na intermediária e 1 na de saída	<i>Backpropagation</i>

(continuação)

Autor (ano)	Título do trabalho	Aplicação	Topologia	Algoritmo de aprendizagem
Freitas (2001)	Utilização de um modelo baseado em redes neurais para a precificação de opções	Verificar se a técnica de redes neurais seria capaz de superar o modelo utilizado para precificação de opções na bolsa de valores.	<i>Multilayer Perceptron</i> (MLP) totalmente conectada, com 4 neurônios na intermediária e 1 na de saída	<i>Backpropagation</i>
Gaio, Castro Junior e Oliveira (2007)	Previsão de preço futuro do boi gordo na BM&F: uma comparação entre modelos de séries temporais e redes neurais	O objetivo do trabalho é comparar modelos de previsão de preço, para o mercado de boi gordo na bolsa de valores, utilizando redes neurais e ferramentas estatística de modelagem de séries temporais.	<i>Multilayer Perceptron</i> (MLP) com 2 neurônios na intermediária e 1 na de saída	<i>Backpropagation</i>
Giacomet (2016)	Um método algorítmico para operações na bolsa de valores em <i>ensembles</i> de Redes Neurais para modelar e prever os movimentos dos mercados de ações	Prever a variação dos preços de ações nas bolsas de valores, e não o preço em si.	<i>Multilayer Perceptron</i> (MLP) com 20 neurônios na camada de entrada, 6 na intermediária e 2 na de saída	<i>Backpropagation</i>
Guimarães (2006)	Previsão de inadimplência e redes neurais artificiais	Prever inadimplência com características de usuários de crédito, aplicando a uma RNA	<i>Multilayer Perceptron</i> (MLP) com 3 neurônios na camada de entrada, 5 na intermediária e 1 na de saída	<i>Backpropagation</i>

(continuação)

Autor (ano)	Título do trabalho	Aplicação	Topologia	Algoritmo de aprendizagem
Moghaddam, Moghaddam e Esfandyari (2016)	Previsão do índice de mercado de ações usando rede neural artificial	Neste estudo, foi investigada a capacidade da rede neural artificial (RNA) na previsão da taxa de câmbio diária NASDAQ.	Foram testadas diversas configurações de redes, mas todas com <i>Multilayer Perceptron</i> (MLP)	<i>Backpropagation</i>
Mulero et al. (2017)	Uma Rede Neural Artificial para a tensão superficial de álcoois	Um modelo de rede neural artificial é proposto para o cálculo e previsão da tensão superficial dos álcoois.	<i>Multilayer Perceptron</i> (MLP) com uma camada de entrada, 21 neurônios em 2 camadas intermediárias e 1 na de saída	<i>Backpropagation</i>
Raia Junior (2000)	Acessibilidade e mobilidade na estimativa de um índice de potencial de viagens utilizando redes neurais artificiais e sistemas de informações geográficas	Propor um processo de modelagem destinado a estimar potenciais de viagens integrando aspectos como acessibilidade e mobilidade.	<i>Multilayer Perceptron</i> (MLP) com até 3 camadas escondidas	<i>Backpropagation</i>
Souza (2008)	Previsão de séries temporais utilizando Rede Neural treinada por filtro de Kalman e evolução diferencial	Propõe o uso do Filtro de Kalman Estendido para minimização do erro de treinamento da rede na previsão de séries temporais financeiras de <i>commodities</i> do agronegócio.	<i>Multilayer Perceptron</i> (MLP)	Filtro de Kalman

(conclusão)

Autor (ano)	Título do trabalho	Aplicação	Topologia	Algoritmo de aprendizagem
Thakur, Bhattacharyya e Mondal (2016)	Modelo baseado em Rede Neural Artificial para previsão de inflação na Índia	Propor uma RNA para prever a inflação mensal da Índia com base em diversos dados históricos.	<i>Multilayer Perceptron</i> (MLP) com 10 camadas de entrada, 25 neurônios em 4 camadas intermediárias e 1 na de saída	<i>Backpropagation</i>

APÊNDICE C – SÍNTESE DOS TRABALHOS ESTUDADOS REFERENTES A RNA: DADOS UTILIZADOS

(continua)

Autor (ano)	Título do trabalho	Input	Output	Principais conclusões sobre a aplicação
Andrade (2011)	Redes Neurais Artificiais aplicadas na identificação automática de áreas cafeeiras em imagens de satélite	- Imagens do terreno com amostras de café, mata, água, outros usos e área urbana	- Classificação das áreas	A metodologia de RNA apresentou um bom resultado, porém, sugere-se que utilize outros dados de entrada para a RNA, uma vez que somente as bandas espectrais não são suficientes para uma classificação otimizada.
Bressan (2001)	Modelos de previsão de preços aplicados aos contratos futuros agropecuários	- Preço de toda quarta-feira útil de boi gordo, café e soja, de 9/11/1996 a 31/12/1999	- Preços das ações para data seguinte	Modelos de previsão captam, de modo coerente, o comportamento dos preços, porém o modelo ARIMA obteve melhor desempenho nos três mercados, enquanto as RNA indicaram um desempenho insatisfatório em termos do erro percentual.

(continuação)

Autor (ano)	Título do trabalho	Input	Output	Principais conclusões sobre a aplicação
Dartora (2003)	Redes neurais artificiais: aplicação na tipificação de carcaça de suínos	- Peso da carcaça quente - Espessura de toucinho avaliada entre a última e a penúltima costela - Profundidade de músculo avaliada entre a última e a penúltima costela	- Porcentagem de carne na carcaça resfriada	RNA apresentou superior desempenho, quando comparado com o modelo linear de regressão múltipla. A RNA pode explicar 85% da variação no percentual de carne na carcaça resfriada, enquanto o método utilizado explica 76,91%.
Devadoss e Ligori (2013)	Predições de ações com o uso de Redes Neurais Artificiais	- Preço diário das ações de quatro empresas indianas, no mês de abril de 2013	- Preço da ação para os 5 dias seguintes	A RNA conseguiu prever os preços das ações com maior acurácia se for aumentado a quantidade de dados de entrada.
Dumitru e Gligor (2017)	Previsão média diária de energia eólica usando redes de neurônios artificiais	- Dados históricos de até 365 dias da produção média diária de energia	- Previsão da produção média diária de energia	O RNA é uma solução viável para identificar padrões e pode ser usado na estimativa de tendência ou produção de energia elétrica.
Freitas (2001)	Utilização de um modelo baseado em redes neurais para a precificação de opções	- Preço diário da ação da empresa Telebrás, de 01/01/1995 a 20/12/1999	- Preço da ação para data seguinte	O desempenho da rede neural foi superior em todas as maneiras analisadas.

(continuação)

Autor (ano)	Título do trabalho	Input	Output	Principais conclusões sobre a aplicação
Gaio, Castro Junior e Oliveira (2007)	Previsão de preço futuro do boi gordo na BM&F: uma comparação entre modelos de séries temporais e redes neurais	- Preço diário do boi gordo, de 01/08/1997 a 27/05/2005	- Preço do boi gordo para data seguinte	A Rede Neural obteve um desempenho superior quando comparado com a previsão dos modelos de séries temporais, tendo um menor valor de erro.
Giacomel (2016)	Um método algorítmico para operações na bolsa de valores em <i>ensembles</i> de Redes Neurais para modelar e prever os movimentos dos mercados de ações	Preço das nove ações mais negociadas nos mercados norte-americano e brasileiro, sendo os períodos: - Dados diários das ações norte-americanas de 19/05/2008 a 14/01/2015, totalizando 1677 períodos - Dados de 15 em 15 minutos das ações brasileiras de 01/07/2014 a 18/09/2014, totalizando 1914 períodos	- Variação do preço da ação	Operações reais foram simuladas e foi possível lucrar em 89% dos casos avaliados, utilizando a RNA.

(continuação)

Autor (ano)	Título do trabalho	Input	Output	Principais conclusões sobre a aplicação
Guimarães (2006)	Previsão de inadimplência e redes neurais artificiais	- Características de 130 clientes de crédito, metade sendo adimplentes e metade inadimplentes, sendo elas: renda, estado civil, tipo de emprego, gênero, tipo e tempo de residência	- Score que representa a previsão de inadimplência ou inadimplência dos clientes	A rede mostrou-se um instrumento eficaz para previsões de inadimplência, classificando corretamente 86,99% dos candidatos.
Moghaddam, Moghaddam e Esfandyari (2016)	Previsão do índice de mercado de ações usando rede neural artificial	- Taxas de câmbio diárias do NASDAQ de 28 de janeiro de 2015 a 18 de junho de 2015	- Previsão da taxa de câmbio do NASDAQ	Os resultados do modelo mostram que não há diferença distinta entre a capacidade de previsão de quatro e nove dias úteis anteriores como parâmetros de entrada.
Mulero et al. (2017)	Uma Rede Neural Artificial para a tensão superficial de álcoois	- Temperatura, temperatura do ponto crítico, densidade crítica, e raio de rotação de 147 álcoois	- Tensão superficial dos álcoois	Foi encontrado um desvio total de percentagem absoluta de 1,04%, RMSE de 0,004 e r^2 de 0,9985. Então, RNA pode ser considerada como uma ferramenta muito precisa para reproduzir os dados de tensão superficial atualmente disponíveis para álcoois e até mesmo para fins de predição.

(conclusão)

Autor (ano)	Título do trabalho	Input	Output	Principais conclusões sobre a aplicação
Raia Junior (2000)	Acessibilidade e mobilidade na estimativa de um índice de potencial de viagens utilizando redes neurais artificiais e sistemas de informações geográficas	- Renda familiar na família - Número de pessoas na família - Proporção de pessoas na família com instrução até: nível primário, nível ginásial, nível colegial, com superior incompleto, e superior completo - Porcentagem de pessoas que possuem carteira de habilitação	- Número de viagens realizadas - Distância percorrida nas viagens	A metodologia empregada é uma ferramenta útil para o planejamento urbano e de transportes. Porém, deixa evidente que não basta apenas prover acessibilidade física, é preciso propiciar melhores índices de mobilidade.
Souza (2008)	Previsão de séries temporais utilizando Rede Neural treinada por filtro de Kalman e evolução diferencial	- Preço diário da soja, de 29/07/1997 a 26/05/2008	- Preço da soja para data seguinte	A rede que foi otimizada com a evolução diferencial obteve um resultado amplamente superior, em termos do menor MSE, comparando com o do algoritmo genético.
Thakur, Bhattacharyya e Mondal (2016)	Modelo baseado em Rede Neural Artificial para previsão de inflação na Índia	- Histórico mensal de jan/2000 a dez/2012, das seguintes informações: PIB, taxa de câmbio, exportação, oferta de moeda, reservas estrangeiras, importação, preços do ouro e do petróleo, balança comercial	- Previsão da inflação mensal para os meses de 2014	A exatidão do modelo se revela satisfatória quando comparada com a previsão de algumas agências bem conhecidas.

APÊNDICE D – SÍNTESE DOS TRABALHOS ESTUDADOS REFERENTES A METODOLOGIA BOX-JENKINS

(continua)

Autor (ano)	Título do trabalho	Aplicação	Dados de entrada	Modelo utilizado	Principais conclusões sobre a previsão
Bressan (2001)	Modelos de previsão de preços aplicados aos contratos futuros agropecuários	Aplicação de modelos de previsão como ferramenta de decisão de compra e venda de ações da BM&F de <i>commodities</i> boi gordo, café e soja.	Preço de toda quarta-feira útil de boi gordo, café e soja, de 9/11/1996 a 31/12/1999	Não foi especificado.	Modelos de previsão captam, de modo coerente, o comportamento dos preços e o modelo ARIMA obteve melhor desempenho nos três mercados, em função de sua rápida adaptabilidade.
Kim, Ko, Nam, Kim, Chung e Bang (2017)	Modelo estatístico para previsão dos preços do urânio para estimar o custo do ciclo do combustível nuclear	Este artigo apresenta um método para prever os futuros preços do urânio que são utilizados como dados de entrada para calcular o custo do urânio, que é um condutor do custo chave racional do custo do ciclo do combustível nuclear.	Preço trimestral do urânio de 2000 a 2014	ARIMA(1,0,0), ARIMA(0,1,1), ARIMA(2,1,1), entre outros	Quando o preço do urânio estava previsto em 2015, a margem de erro da previsão do modelo ARIMA foi calculada 5,4%, enquanto que o modelo de taxa de escalonamento apresentou margem de erro de 7,32%. Assim, verificou-se que o modelo ARIMA é mais adequado na diminuição da incerteza.

(continuação)

Autor (ano)	Título do trabalho	Aplicação	Dados de entrada	Modelo utilizado	Principais conclusões sobre a previsão
Kirchner (2006)	Modelagem e previsão por meio de metodologia Box-Jenkins: uma ferramenta de gestão	Conhecer o comportamento do faturamento de uma empresa do ramo agrícola e duas filiais, para modelar e fazer previsões.	Série histórica do faturamento mensal, do período de janeiro de 2001 a maio de 2006	Na matriz: ARMA(1,1); em uma filial: ARIMA(1,0,0)* (2,0,0)s; na outra filial: ARIMA(1,0,0)* (1,0,0)s	Todos os valores encontraram-se no intervalo de confiança de 95%; e concluiu-se que a metodologia pode ser uma importante ferramenta a ser utilizada pelos gestores.
Lemos (2006)	Metodologia para seleção de métodos de previsão de demanda	Mercado de eletrônicos, que possui histórico de demanda.	Demanda mensal de três produtos, com histórico de 50 meses	Produto 1: ARIMA(0,1,0)* (0,1,0)s Produto 2: ARIMA(1,0,0)* (0,1,0)s Produto 3: ARIMA(1,0,0)* (0,1,0)s	Para os produtos 1 e 3, o modelo de Box-Jenkins teve a menor média dos erros percentuais absolutos, sendo o método de previsão mais adequado.
Pellegrini (2000)	Metodologia para implementação de sistemas de previsão de demanda	Previsão de demanda para 8 semanas, em uma empresa do ramo alimentício, localizada na cidade de Santa Cruz do Sul, RS.	Demanda semanal de 15 famílias de produtos, do período da primeira semana de 1998 a vigésima semana de 2000, totalizando 124 semanas	ARIMA(0,1,1), ARIMA(0,1,1)* (1,0,0)s, ARIMA(0,1,0)*(0,1,0)s, entre outros	Os modelos de Box-Jenkins modelam de forma satisfatória a grande maioria das séries temporais encontradas, porém este trabalho ainda expôs as dificuldades de modelagem de dados reais.

(continuação)

Autor (ano)	Título do trabalho	Aplicação	Dados de entrada	Modelo utilizado	Principais conclusões sobre a previsão
Putro, Koshio e Oktaferdian (2016)	Implementação do modelo ARIMA para ensaios da variação sazonal Conjuntos macrobênicos	Este estudo discutiu a aplicação do método da Média Mover Integrada Autoregressiva (ARIMA) para avaliar a variabilidade sazonal das montagens macrobentônicas.	A abundância de macrobentos nas águas	ARIMA(0, 1, 1)	Embora as assembleias macrobentônicas possam ser influenciadas pelo tempo como resposta às mudanças sazonais da qualidade da água e propriedades do sedimento, sua estrutura da comunidade sazonal pode ser predita usando o método ARIMA.
Queirós (2012)	Modelo de previsão para receita tributária estadual: aplicação da metodologia Box-Jenkins	Construir um modelo de previsão de curto prazo para o imposto estadual (ICMS).	Valor arrecadado de 5 estados brasileiros, representando cada região, no período de janeiro de 1997 a dezembro de 2011	Amazonas: ARIMA (0, 1, 1)(0, 1, 1)s Bahia: ARIMA (2, 1, 0)(0, 1, 1)s Goiás: ARIMA(3, 1, 0) São Paulo: ARIMA (0, 1, 2)(0, 1, 1)s Rio Grande do Sul: ARIMA(0, 1, 2)* (0, 1, 1)s	O resultado da previsão mostrou-se satisfatório para os cinco estados selecionados, e uma previsão deste imposto pode auxiliar na elaboração do orçamento anual dos estados brasileiros para o ano seguinte.

(continuação)

Autor (ano)	Título do trabalho	Aplicação	Dados de entrada	Modelo utilizado	Principais conclusões sobre a previsão
Ramos, Santos e Rebelo (2015)	Desempenho do espaço de estados e modelos ARIMA para previsão de vendas no varejo ao consumidor	Comparar os dois modelos para previsão de vendas mensais de cinco categorias de sapatos femininos de uma marca de sapatos.	Média mensal de cinco categorias de sapatos femininos, de janeiro de 2007 a abril de 2012	ARIMA(0,0,3)* (0,1,0)s, ARIMA (1,0,0)*(0,1,2)s, entre outros	Em algumas análises, espaço de estados obteve um melhor desempenho, em outras o ARIMA se saiu melhor, mas pode-se considerar que ambos métodos podem prever bem o movimento da tendência e as flutuações sazonais.
Rossi (2013)	Previsão de velocidade dos ventos por Redes Neurais Artificiais e ARIMA de Box-Jenkins	Previsão das velocidades dos ventos, com dados agrometeorológicos reais, observados na região da Ilha Solteira, em SP.	Dados diários de velocidades dos ventos, coletados no ano de 2011	ARIMA(2,1,1)* (1,0,0)s	A previsão do modelo ARIMA teve o maior percentual de erro, quando comparado com a prognose da Rede Neural e do modelo dos dois métodos combinados.

(continuação)

Autor (ano)	Título do trabalho	Aplicação	Dados de entrada	Modelo utilizado	Principais conclusões sobre a previsão
Sen, Roy e Pal (2016)	Aplicação do ARIMA para previsão de consumo de energia e emissão de gases do efeito estufa: estudo de caso de uma organização indiana de fabricação de ferro gusa	Foi realizado esta aplicação, para que os gerentes pudessem conhecer as tendências atuais e futuras desses indicadores, para criar uma melhor política ambiental.	Dados mensais de consumo de energia e emissão de gases do efeito estufa dos anos de 2002 a 2013	Para consumo de energia: ARIMA(1,0,0)* (0,1,1)s Para emissão de GEE: ARIMA (0,1,4)*(0,1,1)s	Em ambos os casos, as previsões se assemelham às do modelo de tendência aleatória sazonal, no entanto, elas parecem mais suaves. A seleção correta dos modelos ARIMA para esses indicadores ajudará na previsão precisa para conseguir uma melhor prática de gestão ambiental.
Sharafi, Ghaem, Tabatabaee e Faramarzi (2017)	Previsão do número de casos de leishmaniose cutânea zoonótica no sul da província de Fars, no Irã, usando o método sazonal de séries temporais ARIMA	O objetivo é prever a tendência da leishmaniose cutânea e avaliar a relação entre a tendência da doença e as variáveis climáticas no sul da província de Fars, usando ARIMA sazonal.	Informações sobre a doença e as condições climáticas foram coletadas mensalmente de janeiro de 2010 a março de 2016	ARIMA(4,1,4)* (0,1,0) ₍₁₂₎ e ARIMA(4,1,4)* (0,1,1) ₍₁₂₎	Com base nos resultados, a leishmaniose segue uma tendência descendente na área de estudo, caso a condição de seca continue, os modelos SARIMA podem adequadamente medir a tendência da doença, e a doença segue uma tendência sazonal.

(continuação)

Autor (ano)	Título do trabalho	Aplicação	Dados de entrada	Modelo utilizado	Principais conclusões sobre a previsão
Taneja, Ahmad, e Attri (2016)	Análise de séries temporais da profundidade óptica de aerossóis em Nova Deli usando abordagem de modelagem ARIMA de Box-Jenkins	Aplicar técnica de modelagem estocástica na análise das tendências futuras das propriedades ópticas do aerossol, em Nova Deli, na Índia.	Média mensal da profundidade óptica do aerossol, dos anos de 2004 a 2014	ARIMA(1,0,0)* (0,1,2)s e ARIMA(1,0,0)* (2,1,1)s	Os dois modelos foram considerados os mais adequados, e foram utilizados para calcular a previsão. Resultados satisfatórios foram obtidos e provou-se viável o uso de uma técnica de modelagem simplista.
Werner (2004)	Um modelo composto para realizar previsão de demanda através da integração da combinação de previsões e do ajuste baseado na opinião	Previsão de demanda na manutenção de equipamentos de informática.	Demanda mensal de atendimento de clientes em um período de 5 anos	ARIMA(1,1,0)	Com os dados disponíveis da empresa, foi possível comprovar a exequibilidade da proposta do trabalho, ressaltando a dificuldade na aplicação, por ser necessário ter grande conhecimento das técnicas de estatísticas, para escolher o modelo mais apropriado para o problema.

(conclusão)

Autor (ano)	Título do trabalho	Aplicação	Dados de entrada	Modelo utilizado	Principais conclusões sobre a previsão
Yuan, Liu e Fang (2016)	Comparação da previsão de consumo de energia primária da China usando o modelo ARIMA (média móvel integrada autorregressiva) e o modelo GM(1,1)	O consumo primário de energia da China aumenta rapidamente, o que está altamente relacionado com o desenvolvimento sustentável da China e tem um grande impacto no mercado global de energia. Dois modelos univariados, modelo ARIMA e modelo GM (1,1), são usados para prever o consumo de energia primária da China.	Consumo anual de energia primária da China de 1965 a 2014	ARIMA(2,1,1)	Os resultados dos dois modelos estão em conformidade com os requisitos. Através da comparação, verifica-se que os valores ajustados do modelo ARIMA respondem menos às flutuações porque estão limitadas pela sua tendência de longo prazo, enquanto as do modelo GM (1,1) respondem mais devido ao uso dos últimos quatro dados. Assim, um modelo híbrido é construído com esses dois modelos, e seu MAPE é menor que modelo ARIMA e modelo GM(1,1).

APÊNDICE E – PREÇOS DIÁRIO DAS EMPRESAS QUE FORAM CONSIDERADAS EFICIENTES

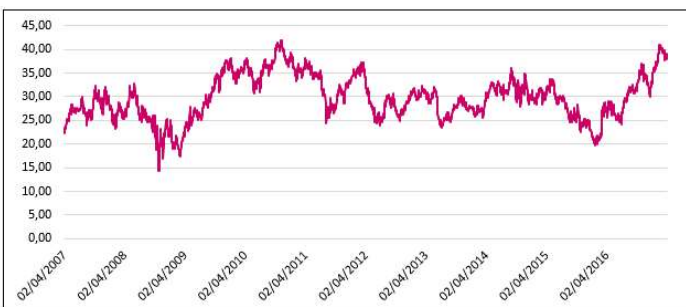
1. AMBEV – ABEV3

Linha	Data	Preço (R\$)
1	01/04/2014	16,26
2	02/04/2014	16,45
3	03/04/2014	16,46
4	04/04/2014	16,19
5	07/04/2014	16,19
...	...	...
740	27/03/2017	17,79
741	28/03/2017	17,94
742	29/03/2017	18,18
743	30/03/2017	18,27
744	31/03/2017	18,24



2. ITAU UNIBANCO – ITUB4

Linha	Data	Preço (R\$)
1	02/04/2007	22,38
2	03/04/2007	23,08
3	04/04/2007	23,34
4	05/04/2007	23,52
5	09/04/2007	23,69
...	...	...
2471	27/03/2017	38,56
2472	28/03/2017	38,43
2473	29/03/2017	39,00
2474	30/03/2017	38,54
2475	31/03/2017	37,90



3. CIELO – CIEL3

Linha	Data	Preço (R\$)
1	01/04/2010	38,12
2	05/04/2010	37,56
3	06/04/2010	36,44
4	07/04/2010	35,20
5	08/04/2010	36,21
...	...	...
1732	27/03/2017	28,02
1733	28/03/2017	28,14
1734	29/03/2017	28,11
1735	30/03/2017	28,05
1736	31/03/2017	28,31



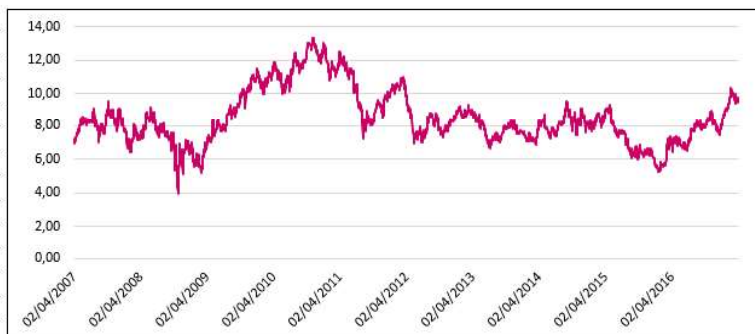
4. BB SEGURIDADE – BBSE3

Linha	Data	Preço (R\$)
1	01/04/2014	24,22
2	02/04/2014	24,40
3	03/04/2014	24,21
4	04/04/2014	24,08
5	07/04/2014	24,66
...	...	...
740	27/03/2017	28,37
741	28/03/2017	28,65
742	29/03/2017	29,39
743	30/03/2017	29,45
744	31/03/2017	29,20



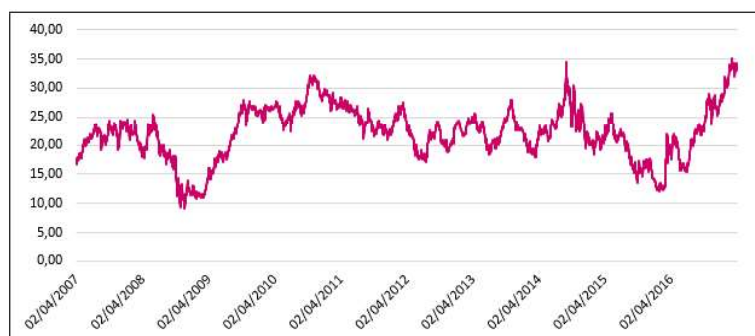
5. ITAUSA – ITSA4

Linha	Data	Preço (R\$)
1	02/04/2007	6,96
2	03/04/2007	7,14
3	04/04/2007	7,20
4	05/04/2007	7,20
5	09/04/2007	7,11
...	...	...
2471	27/03/2017	9,75
2472	28/03/2017	9,68
2473	29/03/2017	9,77
2474	30/03/2017	9,65
2475	31/03/2017	9,50



6. BANCO DO BRASIL – BBAS3

Linha	Data	Preço (R\$)
1	02/04/2007	16,57
2	03/04/2007	17,42
3	04/04/2007	17,33
4	05/04/2007	17,68
5	09/04/2007	17,83
...	...	...
2471	27/03/2017	33,31
2472	28/03/2017	33,10
2473	29/03/2017	34,30
2474	30/03/2017	34,07
2475	31/03/2017	33,77



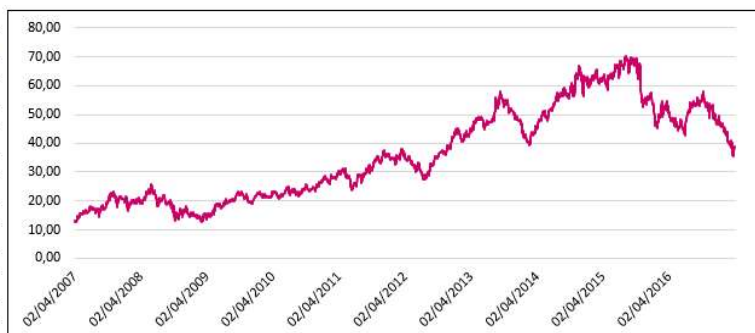
7. ULTRAPAR – UGPA3

Linha	Data	Preço (R\$)
1	02/04/2012	39,35
2	03/04/2012	40,03
3	04/04/2012	39,60
4	05/04/2012	39,54
5	09/04/2012	39,06
...	...	...
1234	27/03/2017	71,33
1235	28/03/2017	71,48
1236	29/03/2017	72,50
1237	30/03/2017	72,24
1238	31/03/2017	71,62



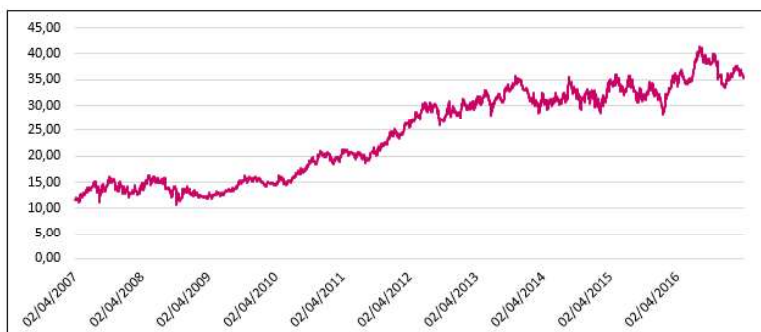
8. BRF – BRFS3

Linha	Data	Preço (R\$)
1	02/04/2007	12,62
2	03/04/2007	12,98
3	04/04/2007	13,05
4	05/04/2007	13,31
5	09/04/2007	13,19
...	...	...
2469	27/03/2017	37,09
2470	28/03/2017	38,65
2471	29/03/2017	38,87
2472	30/03/2017	38,54
2473	31/03/2017	38,61



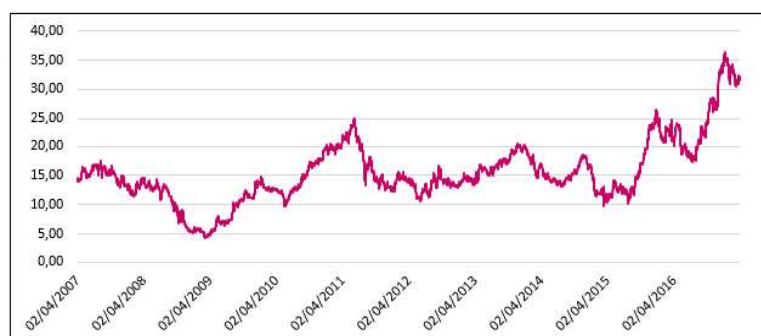
9. ENGIE BRASIL ENERGIA – EGIE3

Linha	Data	Preço (R\$)
1	02/04/2007	11,52
2	03/04/2007	11,67
3	04/04/2007	11,61
4	05/04/2007	11,62
5	09/04/2007	11,90
...	...	...
2470	27/03/2017	35,93
2471	28/03/2017	35,53
2472	29/03/2017	35,79
2473	30/03/2017	35,40
2474	31/03/2017	35,37



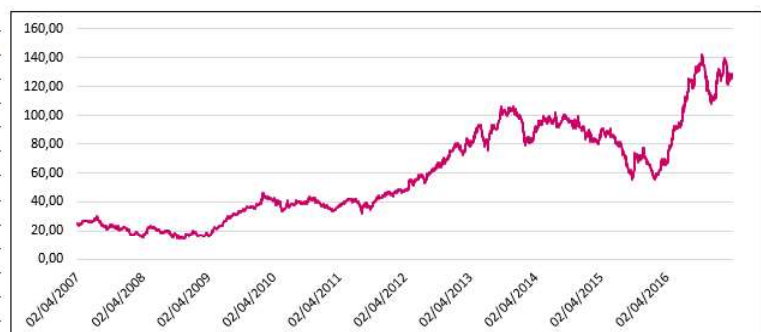
10. BRASKEM – BRKM5

Linha	Data	Preço (R\$)
1	02/04/2007	13,90
2	03/04/2007	14,44
3	04/04/2007	14,48
4	05/04/2007	14,48
5	09/04/2007	14,36
...	...	...
2471	27/03/2017	32,28
2472	28/03/2017	31,75
2473	29/03/2017	32,19
2474	30/03/2017	31,54
2475	31/03/2017	31,75



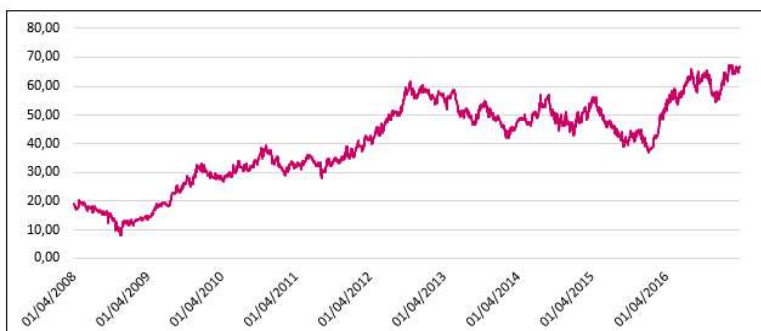
11. M. DIAS BRANCO – MDIA3

Linha	Data	Preço (R\$)
1	02/04/2007	25,23
2	03/04/2007	24,83
3	04/04/2007	24,43
4	05/04/2007	24,54
5	09/04/2007	24,40
...	...	...
2463	27/03/2017	126,11
2464	28/03/2017	128,50
2465	29/03/2017	126,52
2466	30/03/2017	125,67
2467	31/03/2017	127,80



12. MULTIPLAN – MULT3

Linha	Data	Preço (R\$)
1	01/04/2008	19,34
2	02/04/2008	19,14
3	03/04/2008	18,95
4	04/04/2008	18,77
5	07/04/2008	18,13
...	...	...
2227	27/03/2017	65,56
2228	28/03/2017	66,25
2229	29/03/2017	66,54
2230	30/03/2017	66,67
2231	31/03/2017	66,30



13. EQUATORIAL ENERGIA – EQTL3

Linha	Data	Preço (R\$)
1	01/04/2009	5,61
2	02/04/2009	5,89
3	03/04/2009	5,94
4	06/04/2009	5,99
5	07/04/2009	6,08
...	...	...
1977	27/03/2017	56,90
1978	28/03/2017	57,00
1979	29/03/2017	57,49
1980	30/03/2017	58,03
1981	31/03/2017	58,79



14. EDP ENERGIAS DO BRASIL – ENBR3

Linha	Data	Preço (R\$)
1	02/04/2007	22,55
2	03/04/2007	23,12
3	04/04/2007	23,23
4	05/04/2007	23,27
5	09/04/2007	23,88
...	...	...
2472	27/03/2017	13,60
2473	28/03/2017	13,69
2474	29/03/2017	13,70
2475	30/03/2017	13,99
2476	31/03/2017	13,95



15. CIA ENERGÉTICA DE MINAS GERAIS – CMIG4

Linha	Data	Preço (R\$)
1	02/04/2007	17,50
2	03/04/2007	17,75
3	04/04/2007	17,80
4	05/04/2007	17,82
5	09/04/2007	17,94
...	...	...
2471	27/03/2017	10,54
2472	28/03/2017	10,37
2473	29/03/2017	10,35
2474	30/03/2017	10,26
2475	31/03/2017	10,32



16. TAESA – TAE11

Linha	Data	Preço (R\$)
1	02/04/2012	18,13
2	03/04/2012	17,99
3	04/04/2012	18,04
4	05/04/2012	18,35
5	09/04/2012	18,13
...	...	...
1234	27/03/2017	22,05
1235	28/03/2017	22,43
1236	29/03/2017	22,75
1237	30/03/2017	22,80
1238	31/03/2017	22,95



APÊNDICE F – RESULTADOS OBTIDOS UTILIZANDO BOX-JENKINS

(continua)

2. ITAU UNIBANCO – ITUB4			3. CIELO – CIEL3			4. BB SEGURIDADE – BBSE3		
12 meses	Modelo	ARIMA(0,1,0)	12 meses	Modelo	ARIMA(0,1,0)	12 meses	Modelo	ARIMA(0,1,0)
	Valor calculado	R\$ 37,94		Valor calculado	R\$ 28,30		Valor calculado	R\$ 29,21
24 meses	Modelo	ARIMA(0,1,0)	24 meses	Modelo	ARIMA(0,1,0)	24 meses	Modelo	ARIMA(0,1,8)
	Valor calculado	R\$ 37,91		Valor calculado	R\$ 28,29		Valor calculado	R\$ 29,02
36 meses	Modelo	ARIMA(0,1,3)	36 meses	Modelo	ARIMA(0,1,0)	36 meses	Modelo	ARIMA(0,1,8)
	Valor calculado	R\$ 37,85		Valor calculado	R\$ 28,30		Valor calculado	R\$ 29,10
48 meses	Modelo	ARIMA(0,1,3)	48 meses	Modelo	ARIMA(0,1,0)			
	Valor calculado	R\$ 37,85		Valor calculado	R\$ 28,29			
60 meses	Modelo	ARIMA(0,1,3)	60 meses	Modelo	ARIMA(2,1,2)			
	Valor calculado	R\$ 37,85		Valor calculado	R\$ 28,29			
72 meses	Modelo	ARIMA(0,1,3)	72 meses	Modelo	ARIMA(2,1,2)			
	Valor calculado	R\$ 37,86		Valor calculado	R\$ 28,31			
84 meses	Modelo	ARIMA(0,1,3)	84 meses	Modelo	ARIMA(2,1,2)			
	Valor calculado	R\$ 37,87		Valor calculado	R\$ 28,31			
96 meses	Modelo	ARIMA(2,1,2)						
	Valor calculado	R\$ 37,86						
108 meses	Modelo	ARIMA(0,1,3)						
	Valor calculado	R\$ 37,85						
120 meses	Modelo	ARIMA(0,1,3)						
	Valor calculado	R\$ 37,85						

(continuação)

5. ITAUSA – ITSA4			6. BANCO DO BRASIL – BBAS3			7. ULTRAPAR – UGPA3		
12 meses	Modelo	ARIMA(0,1,0)	12 meses	Modelo	ARIMA(0,1,0)	12 meses	Modelo	ARIMA(0,1,0)
	Valor calculado	R\$ 9,51		Valor calculado	R\$ 33,83		Valor calculado	R\$ 71,63
24 meses	Modelo	ARIMA(0,1,0)	24 meses	Modelo	ARIMA(0,1,0)	24 meses	Modelo	ARIMA(0,1,0)
	Valor calculado	R\$ 9,50		Valor calculado	R\$ 33,79		Valor calculado	R\$ 71,63
36 meses	Modelo	ARIMA(0,1,0)	36 meses	Modelo	ARIMA(0,1,0)	36 meses	Modelo	ARIMA(0,1,0)
	Valor calculado	R\$ 9,50		Valor calculado	R\$ 33,79		Valor calculado	R\$ 71,64
48 meses	Modelo	ARIMA(0,1,0)	48 meses	Modelo	ARIMA(2,1,2)	48 meses	Modelo	ARIMA(0,1,1)
	Valor calculado	R\$ 9,50		Valor calculado	R\$ 33,69		Valor calculado	R\$ 71,66
60 meses	Modelo	ARIMA(1,1,1)	60 meses	Modelo	ARIMA(0,1,8)	60 meses	Modelo	ARIMA(0,1,1)
	Valor calculado	R\$ 9,50		Valor calculado	R\$ 33,68		Valor calculado	R\$ 71,66
72 meses	Modelo	ARIMA(0,1,0)	72 meses	Modelo	ARIMA(2,1,2)			
	Valor calculado	R\$ 9,50		Valor calculado	R\$ 33,71			
84 meses	Modelo	ARIMA(0,1,0)	84 meses	Modelo	ARIMA(1,1,10)			
	Valor calculado	R\$ 9,50		Valor calculado	R\$ 33,77			
96 meses	Modelo	ARIMA(1,1,1)	96 meses	Modelo	ARIMA(2,1,2)			
	Valor calculado	R\$ 9,50		Valor calculado	R\$ 33,70			
108 meses	Modelo	ARIMA(3,1,3)	108 meses	Modelo	ARIMA(0,1,0)			
	Valor calculado	R\$ 9,50		Valor calculado	R\$ 33,78			
120 meses	Modelo	ARIMA(0,1,7)	120 meses	Modelo	ARIMA(2,1,2)			
	Valor calculado	R\$ 9,50		Valor calculado	R\$ 33,78			

(continuação)

8. BRF – BRFS3			9. ENGIE BRASIL ENERGIA – EGIE3			10. BRASKEM – BRKM5		
12 meses	Modelo	ARIMA(0,1,0)	12 meses	Modelo	ARIMA(1,1,0)	12 meses	Modelo	ARIMA(0,1,0)
	Valor calculado	R\$ 38,56		Valor calculado	R\$ 35,37		Valor calculado	R\$ 31,79
24 meses	Modelo	ARIMA(0,1,0)	24 meses	Modelo	ARIMA(0,1,6)	24 meses	Modelo	ARIMA(0,1,0)
	Valor calculado	R\$ 38,56		Valor calculado	R\$ 35,40		Valor calculado	R\$ 31,79
36 meses	Modelo	ARIMA(1,1,1)	36 meses	Modelo	ARIMA(0,1,3)	36 meses	Modelo	ARIMA(0,1,0)
	Valor calculado	R\$ 38,61		Valor calculado	R\$ 35,34		Valor calculado	R\$ 31,78
48 meses	Modelo	ARIMA(0,1,0)	48 meses	Modelo	ARIMA(0,1,13)	48 meses	Modelo	ARIMA(0,1,0)
	Valor calculado	R\$ 38,60		Valor calculado	R\$ 35,37		Valor calculado	R\$ 31,77
60 meses	Modelo	ARIMA(0,1,0)	60 meses	Modelo	ARIMA(0,1,3)	60 meses	Modelo	ARIMA(0,1,0)
	Valor calculado	R\$ 38,61		Valor calculado	R\$ 35,36		Valor calculado	R\$ 31,77
72 meses	Modelo	ARIMA(0,1,0)	72 meses	Modelo	ARIMA(0,1,3)	72 meses	Modelo	ARIMA(0,1,0)
	Valor calculado	R\$ 38,62		Valor calculado	R\$ 35,36		Valor calculado	R\$ 31,76
84 meses	Modelo	ARIMA(0,1,0)	84 meses	Modelo	ARIMA(1,1,1)	84 meses	Modelo	ARIMA(0,1,0)
	Valor calculado	R\$ 38,63		Valor calculado	R\$ 35,41		Valor calculado	R\$ 31,76
96 meses	Modelo	ARIMA(0,1,0)	96 meses	Modelo	ARIMA(1,1,1)	96 meses	Modelo	ARIMA(0,1,0)
	Valor calculado	R\$ 38,64		Valor calculado	R\$ 35,44		Valor calculado	R\$ 31,77
108 meses	Modelo	ARIMA(0,1,0)	108 meses	Modelo	ARIMA(2,1,1)	108 meses	Modelo	ARIMA(0,1,0)
	Valor calculado	R\$ 38,62		Valor calculado	R\$ 35,41		Valor calculado	R\$ 31,76
120 meses	Modelo	ARIMA(0,1,0)	120 meses	Modelo	ARIMA(2,1,6)	120 meses	Modelo	ARIMA(0,1,0)
	Valor calculado	R\$ 38,63		Valor calculado	R\$ 35,39		Valor calculado	R\$ 31,76

(continuação)

11. M. DIAS BRANCO – MDIA3			12. MULTIPLAN – MULT3			13. EQUATORIAL ENERGIA – EQTL3		
12 meses	Modelo	ARIMA(0,1,10)	12 meses	Modelo	ARIMA(1,1,1)	12 meses	Modelo	ARIMA(0,1,0)
	Valor calculado	R\$ 127,53		Valor calculado	R\$ 66,28		Valor calculado	R\$ 58,87
24 meses	Modelo	ARIMA(1,1,1)	24 meses	Modelo	ARIMA(1,1,1)	24 meses	Modelo	ARIMA(0,1,0)
	Valor calculado	R\$ 127,87		Valor calculado	R\$ 66,29		Valor calculado	R\$ 58,87
36 meses	Modelo	ARIMA(1,1,10)	36 meses	Modelo	ARIMA(0,1,0)	36 meses	Modelo	ARIMA(0,1,0)
	Valor calculado	R\$ 127,44		Valor calculado	R\$ 66,33		Valor calculado	R\$ 58,88
48 meses	Modelo	ARIMA(1,1,12)	48 meses	Modelo	ARIMA(0,1,0)	48 meses	Modelo	ARIMA(1,1,1)
	Valor calculado	R\$ 127,37		Valor calculado	R\$ 66,31		Valor calculado	R\$ 58,80
60 meses	Modelo	ARIMA(1,1,8)	60 meses	Modelo	ARIMA(0,1,1)	60 meses	Modelo	ARIMA(1,1,1)
	Valor calculado	R\$ 127,76		Valor calculado	R\$ 66,33		Valor calculado	R\$ 58,75
72 meses	Modelo	ARIMA(1,1,8)	72 meses	Modelo	ARIMA(0,1,1)	72 meses	Modelo	ARIMA(1,1,1)
	Valor calculado	R\$ 127,71		Valor calculado	R\$ 66,32		Valor calculado	R\$ 58,74
84 meses	Modelo	ARIMA(0,1,1)	84 meses	Modelo	ARIMA(0,1,1)	84 meses	Modelo	ARIMA(0,1,3)
	Valor calculado	R\$ 127,66		Valor calculado	R\$ 66,32		Valor calculado	R\$ 58,74
96 meses	Modelo	ARIMA(1,1,7)	96 meses	Modelo	ARIMA(0,1,1)	96 meses	Modelo	ARIMA(0,1,3)
	Valor calculado	R\$ 127,79		Valor calculado	R\$ 66,32		Valor calculado	R\$ 58,76
108 meses	Modelo	ARIMA(0,1,9)	108 meses	Modelo	ARIMA(0,1,1)			
	Valor calculado	R\$ 128,05		Valor calculado	R\$ 66,32			
120 meses	Modelo	ARIMA(0,1,0)						
	Valor calculado	R\$ 127,91						

(conclusão)

14. EDP ENERGIAS DO BRASIL – ENBR3			15. CIA ENERGÉTICA DE MINAS GERAIS – CMIG4			16. TAESA – TAEE11		
12 meses	Modelo	ARIMA(0,1,0)	12 meses	Modelo	ARIMA(0,1,0)	12 meses	Modelo	ARIMA(0,1,4)
	Valor calculado	R\$ 13,96		Valor calculado	R\$ 10,33		Valor calculado	R\$ 22,87
24 meses	Modelo	ARIMA(0,1,1)	24 meses	Modelo	ARIMA(1,1,1)	24 meses	Modelo	ARIMA(0,1,6)
	Valor calculado	R\$ 13,95		Valor calculado	R\$ 10,32		Valor calculado	R\$ 22,92
36 meses	Modelo	ARIMA(0,1,1)	36 meses	Modelo	ARIMA(0,1,0)	36 meses	Modelo	ARIMA(2,1,14)
	Valor calculado	R\$ 13,95		Valor calculado	R\$ 10,32		Valor calculado	R\$ 22,78
48 meses	Modelo	ARIMA(1,1,0)	48 meses	Modelo	ARIMA(0,1,0)	48 meses	Modelo	ARIMA(0,1,14)
	Valor calculado	R\$ 13,95		Valor calculado	R\$ 10,32		Valor calculado	R\$ 22,86
60 meses	Modelo	ARIMA(0,1,1)	60 meses	Modelo	ARIMA(0,1,4)	60 meses	Modelo	ARIMA(0,1,4)
	Valor calculado	R\$ 13,95		Valor calculado	R\$ 10,34		Valor calculado	R\$ 22,90
72 meses	Modelo	ARIMA(0,1,0)	72 meses	Modelo	ARIMA(2,1,2)			
	Valor calculado	R\$ 13,95		Valor calculado	R\$ 10,34			
84 meses	Modelo	ARIMA(0,1,0)	84 meses	Modelo	ARIMA(0,1,4)			
	Valor calculado	R\$ 13,95		Valor calculado	R\$ 10,33			
96 meses	Modelo	ARIMA(0,1,0)	96 meses	Modelo	ARIMA(0,1,4)			
	Valor calculado	R\$ 13,95		Valor calculado	R\$ 10,33			
108 meses	Modelo	ARIMA(0,1,0)	108 meses	Modelo	ARIMA(0,1,4)			
	Valor calculado	R\$ 13,95		Valor calculado	R\$ 10,33			
120 meses	Modelo	ARIMA(0,1,0)	120 meses	Modelo	ARIMA(0,1,4)			
	Valor calculado	R\$ 13,95		Valor calculado	R\$ 10,33			

APÊNDICE G – ERRO DOS RESULTADOS

2. ITAU UNIBANCO – ITUB4

Preço do dia 03/04/2017:		R\$ 38,49	
Método utilizado:		RNA	B-J
12 meses	Valor calculado	R\$ 38,04	R\$ 37,94
	Erro	-1,17%	-1,43%
24 meses	Valor calculado	R\$ 37,77	R\$ 37,91
	Erro	-1,87%	-1,51%
36 meses	Valor calculado	R\$ 38,18	R\$ 37,85
	Erro	-0,81%	-1,66%
48 meses	Valor calculado	R\$ 38,23	R\$ 37,85
	Erro	-0,68%	-1,66%
60 meses	Valor calculado	R\$ 38,39	R\$ 37,85
	Erro	-0,26%	-1,66%
72 meses	Valor calculado	R\$ 38,23	R\$ 37,86
	Erro	-0,68%	-1,64%
84 meses	Valor calculado	R\$ 37,59	R\$ 37,87
	Erro	-2,34%	-1,61%
96 meses	Valor calculado	R\$ 37,82	R\$ 37,86
	Erro	-1,74%	-1,64%
108 meses	Valor calculado	R\$ 37,49	R\$ 37,85
	Erro	-2,60%	-1,66%
120 meses	Valor calculado	R\$ 37,78	R\$ 37,85
	Erro	-1,84%	-1,66%

3. CIELO – CIEL3

Preço do dia 03/04/2017:		R\$ 24,25	
Método utilizado:		RNA	B-J
12 meses	Valor calculado	R\$ 28,13	R\$ 28,30
	Erro	16,00%	16,70%
24 meses	Valor calculado	R\$ 28,22	R\$ 28,29
	Erro	16,37%	16,66%
36 meses	Valor calculado	R\$ 27,83	R\$ 28,30
	Erro	14,76%	16,70%
48 meses	Valor calculado	R\$ 28,27	R\$ 28,29
	Erro	16,58%	16,66%
60 meses	Valor calculado	R\$ 27,76	R\$ 28,29
	Erro	14,47%	16,66%
72 meses	Valor calculado	R\$ 28,56	R\$ 28,31
	Erro	17,77%	16,74%
84 meses	Valor calculado	R\$ 27,86	R\$ 28,31
	Erro	14,89%	16,74%

4. BB SEGURIDADE – BBSE3

Preço do dia 03/04/2017:		R\$ 29,21	
Método utilizado:		RNA	B-J
12 meses	Valor calculado	R\$ 29,44	R\$ 29,21
	Erro	0,79%	0,00%
24 meses	Valor calculado	R\$ 29,16	R\$ 29,02
	Erro	-0,17%	-0,65%
36 meses	Valor calculado	R\$ 29,30	R\$ 29,10
	Erro	0,31%	-0,38%

5. ITAUSA – ITSA4

Preço do dia 03/04/2017:		R\$ 9,56	
Método utilizado:		RNA	B-J
12 meses	Valor calculado	R\$ 9,45	R\$ 9,51
	Erro	-1,15%	-0,52%
24 meses	Valor calculado	R\$ 9,47	R\$ 9,50
	Erro	-0,94%	-0,63%
36 meses	Valor calculado	R\$ 9,56	R\$ 9,50
	Erro	0,00%	-0,63%
48 meses	Valor calculado	R\$ 9,33	R\$ 9,50
	Erro	-2,41%	-0,63%
60 meses	Valor calculado	R\$ 9,50	R\$ 9,50
	Erro	-0,63%	-0,63%
72 meses	Valor calculado	R\$ 9,47	R\$ 9,50
	Erro	-0,94%	-0,63%
84 meses	Valor calculado	R\$ 9,54	R\$ 9,50
	Erro	-0,21%	-0,63%
96 meses	Valor calculado	R\$ 9,54	R\$ 9,50
	Erro	-0,21%	-0,63%
108 meses	Valor calculado	R\$ 9,55	R\$ 9,50
	Erro	-0,10%	-0,63%
120 meses	Valor calculado	R\$ 9,60	R\$ 9,50
	Erro	0,42%	-0,63%

6. BANCO DO BRASIL – BBAS3

(continua)

Preço do dia 03/04/2017:		R\$ 33,64	
Método utilizado:		RNA	B-J
12 meses	Valor calculado	R\$ 34,24	R\$ 33,83
	Erro	1,78%	0,56%
24 meses	Valor calculado	R\$ 33,84	R\$ 33,79
	Erro	0,59%	0,45%

(conclusão)

Método utilizado:		RNA	B-J
36 meses	Valor calculado	R\$ 33,59	R\$ 33,79
	Erro	-0,15%	0,45%
48 meses	Valor calculado	R\$ 33,56	R\$ 33,69
	Erro	-0,24%	0,15%
60 meses	Valor calculado	R\$ 33,93	R\$ 33,68
	Erro	0,86%	0,12%
72 meses	Valor calculado	R\$ 33,41	R\$ 33,71
	Erro	-0,68%	0,21%
84 meses	Valor calculado	R\$ 34,12	R\$ 33,77
	Erro	1,43%	0,39%
96 meses	Valor calculado	R\$ 33,79	R\$ 33,70
	Erro	0,45%	0,18%
108 meses	Valor calculado	R\$ 33,93	R\$ 33,78
	Erro	0,86%	0,42%
120 meses	Valor calculado	R\$ 34,02	R\$ 33,78
	Erro	1,13%	0,42%

7. ULTRAPAR – UGPA3

Preço do dia 03/04/2017:		R\$ 71,94	
Método utilizado:		RNA	B-J
12 meses	Valor calculado	R\$ 71,57	R\$ 71,63
	Erro	-0,51%	-0,43%
24 meses	Valor calculado	R\$ 71,19	R\$ 71,63
	Erro	-1,04%	-0,43%
36 meses	Valor calculado	R\$ 71,36	R\$ 71,64
	Erro	-0,81%	-0,42%
48 meses	Valor calculado	R\$ 71,36	R\$ 71,66
	Erro	-0,81%	-0,39%
60 meses	Valor calculado	R\$ 71,62	R\$ 71,66
	Erro	-0,44%	-0,39%

8. BRF – BRFS3

(continua)

Preço do dia 03/04/2017:		R\$ 38,89	
Método utilizado:		RNA	B-J
12 meses	Valor calculado	R\$ 38,63	R\$ 38,56
	Erro	-0,67%	-0,85%
24 meses	Valor calculado	R\$ 38,13	R\$ 38,56
	Erro	-1,95%	-0,85%
36 meses	Valor calculado	R\$ 38,17	R\$ 38,61
	Erro	-1,85%	-0,72%

(conclusão)			
Método utilizado:		RNA	B-J
48 meses	Valor calculado	R\$ 39,03	R\$ 38,60
	Erro	0,36%	-0,75%
60 meses	Valor calculado	R\$ 39,03	R\$ 38,61
	Erro	0,36%	-0,72%
72 meses	Valor calculado	R\$ 38,55	R\$ 38,62
	Erro	-0,87%	-0,69%
84 meses	Valor calculado	R\$ 38,39	R\$ 38,63
	Erro	-1,29%	-0,67%
96 meses	Valor calculado	R\$ 38,14	R\$ 38,64
	Erro	-1,93%	-0,64%
108 meses	Valor calculado	R\$ 38,60	R\$ 38,62
	Erro	-0,75%	-0,69%
120 meses	Valor calculado	R\$ 38,28	R\$ 38,63
	Erro	-1,57%	-0,67%

9. ENGIE BRASIL ENERGIA – EGIE3

Preço do dia 03/04/2017:		R\$ 34,83	
Método utilizado:		RNA	B-J
12 meses	Valor calculado	R\$ 35,44	R\$ 35,37
	Erro	1,75%	1,55%
24 meses	Valor calculado	R\$ 35,40	R\$ 35,40
	Erro	1,64%	1,64%
36 meses	Valor calculado	R\$ 35,37	R\$ 35,34
	Erro	1,55%	1,46%
48 meses	Valor calculado	R\$ 35,43	R\$ 35,37
	Erro	1,72%	1,55%
60 meses	Valor calculado	R\$ 35,35	R\$ 35,36
	Erro	1,49%	1,52%
72 meses	Valor calculado	R\$ 35,25	R\$ 35,36
	Erro	1,21%	1,52%
84 meses	Valor calculado	R\$ 35,25	R\$ 35,41
	Erro	1,21%	1,67%
96 meses	Valor calculado	R\$ 35,09	R\$ 35,44
	Erro	0,75%	1,75%
108 meses	Valor calculado	R\$ 35,34	R\$ 35,41
	Erro	1,46%	1,67%
120 meses	Valor calculado	R\$ 35,34	R\$ 35,39
	Erro	1,46%	1,61%

10. BRASKEM – BRKM5

Preço do dia 03/04/2017:		R\$ 31,91	
Método utilizado:		RNA	B-J
12 meses	Valor calculado	R\$ 31,66	R\$ 31,79
	Erro	-0,78%	-0,38%
24 meses	Valor calculado	R\$ 31,70	R\$ 31,79
	Erro	-0,66%	-0,38%
36 meses	Valor calculado	R\$ 32,28	R\$ 31,78
	Erro	1,16%	-0,41%
48 meses	Valor calculado	R\$ 31,82	R\$ 31,77
	Erro	-0,28%	-0,44%
60 meses	Valor calculado	R\$ 31,81	R\$ 31,77
	Erro	-0,31%	-0,44%
72 meses	Valor calculado	R\$ 31,53	R\$ 31,76
	Erro	-1,19%	-0,47%
84 meses	Valor calculado	R\$ 31,26	R\$ 31,76
	Erro	-2,04%	-0,47%
96 meses	Valor calculado	R\$ 31,95	R\$ 31,77
	Erro	0,13%	-0,44%
108 meses	Valor calculado	R\$ 31,27	R\$ 31,76
	Erro	-2,01%	-0,47%
120 meses	Valor calculado	R\$ 31,76	R\$ 31,76
	Erro	-0,47%	-0,47%

11. M. DIAS BRANCO – MDIA3

(continua)

Preço do dia 03/04/2017:		R\$ 129,13	
Método utilizado:		RNA	B-J
12 meses	Valor calculado	R\$ 127,28	R\$ 127,53
	Erro	-1,43%	-1,24%
24 meses	Valor calculado	R\$ 127,90	R\$ 127,87
	Erro	-0,95%	-0,98%
36 meses	Valor calculado	R\$ 126,01	R\$ 127,44
	Erro	-2,42%	-1,31%
48 meses	Valor calculado	R\$ 127,41	R\$ 127,37
	Erro	-1,33%	-1,36%
60 meses	Valor calculado	R\$ 128,65	R\$ 127,76
	Erro	-0,37%	-1,06%
72 meses	Valor calculado	R\$ 128,87	R\$ 127,71
	Erro	-0,20%	-1,10%
84 meses	Valor calculado	R\$ 129,56	R\$ 127,66
	Erro	0,33%	-1,14%
96 meses	Valor calculado	R\$ 128,81	R\$ 127,79
	Erro	-0,25%	-1,04%

(conclusão)			
Método utilizado:		RNA	B-J
108 meses	Valor calculado	R\$ 129,93	R\$ 128,05
	Erro	0,62%	-0,84%
120 meses	Valor calculado	R\$ 127,78	R\$ 127,91
	Erro	-1,05%	-0,94%

12. MULTIPLAN – MULT3

Preço do dia 03/04/2017:		R\$ 66,87	
Método utilizado:		RNA	B-J
12 meses	Valor calculado	R\$ 65,60	R\$ 66,28
	Erro	-1,90%	-0,88%
24 meses	Valor calculado	R\$ 65,84	R\$ 66,29
	Erro	-1,54%	-0,87%
36 meses	Valor calculado	R\$ 65,82	R\$ 66,33
	Erro	-1,57%	-0,81%
48 meses	Valor calculado	R\$ 65,56	R\$ 66,31
	Erro	-1,96%	-0,84%
60 meses	Valor calculado	R\$ 65,70	R\$ 66,33
	Erro	-1,75%	-0,81%
72 meses	Valor calculado	R\$ 65,85	R\$ 66,32
	Erro	-1,53%	-0,82%
84 meses	Valor calculado	R\$ 65,70	R\$ 66,32
	Erro	-1,75%	-0,82%
96 meses	Valor calculado	R\$ 66,74	R\$ 66,32
	Erro	-0,19%	-0,82%
108 meses	Valor calculado	R\$ 65,62	R\$ 66,32
	Erro	-1,87%	-0,82%

13. EQUATORIAL ENERGIA – EQTL3

(continua)			
Preço do dia 03/04/2017:		R\$ 57,71	
Método utilizado:		RNA	B-J
12 meses	Valor calculado	R\$ 58,72	R\$ 58,87
	Erro	1,75%	2,01%
24 meses	Valor calculado	R\$ 61,08	R\$ 58,87
	Erro	5,84%	2,01%
36 meses	Valor calculado	R\$ 58,86	R\$ 58,88
	Erro	1,99%	2,03%
48 meses	Valor calculado	R\$ 58,60	R\$ 58,80
	Erro	1,54%	1,89%
60 meses	Valor calculado	R\$ 58,37	R\$ 58,75
	Erro	1,14%	1,80%

(conclusão)			
Método utilizado:		RNA	B-J
72 meses	Valor calculado	R\$ 58,63	R\$ 58,74
	Erro	1,59%	1,78%
84 meses	Valor calculado	R\$ 57,84	R\$ 58,74
	Erro	0,23%	1,78%
96 meses	Valor calculado	R\$ 57,81	R\$ 58,76
	Erro	0,17%	1,82%

14. EDP ENERGIAS DO BRASIL – ENBR3

Preço do dia 03/04/2017:		R\$ 14,11	
Método utilizado:		RNA	B-J
12 meses	Valor calculado	R\$ 13,88	R\$ 13,96
	Erro	-1,63%	-1,06%
24 meses	Valor calculado	R\$ 13,77	R\$ 13,95
	Erro	-2,41%	-1,13%
36 meses	Valor calculado	R\$ 13,86	R\$ 13,95
	Erro	-1,77%	-1,13%
48 meses	Valor calculado	R\$ 13,94	R\$ 13,95
	Erro	-1,20%	-1,13%
60 meses	Valor calculado	R\$ 13,95	R\$ 13,95
	Erro	-1,13%	-1,13%
72 meses	Valor calculado	R\$ 13,83	R\$ 13,95
	Erro	-1,98%	-1,13%
84 meses	Valor calculado	R\$ 13,82	R\$ 13,95
	Erro	-2,06%	-1,13%
96 meses	Valor calculado	R\$ 13,59	R\$ 13,95
	Erro	-3,69%	-1,13%
108 meses	Valor calculado	R\$ 13,75	R\$ 13,95
	Erro	-2,55%	-1,13%
120 meses	Valor calculado	R\$ 13,72	R\$ 13,95
	Erro	-2,76%	-1,13%

15. CIA ENERGÉTICA DE MINAS GERAIS – CMIG4

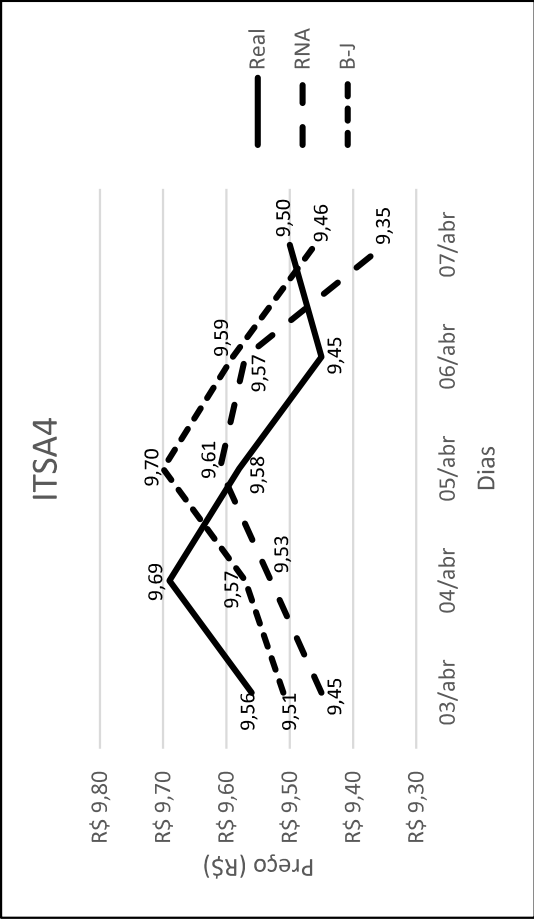
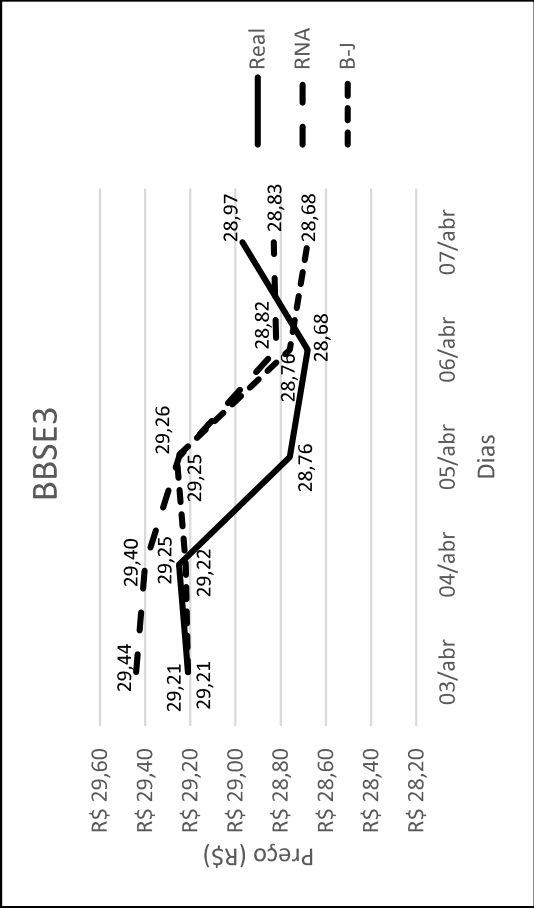
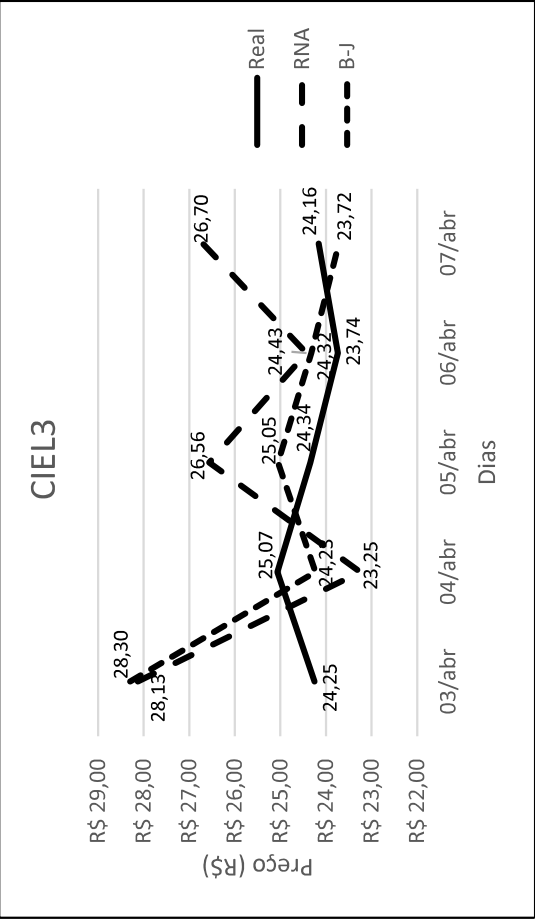
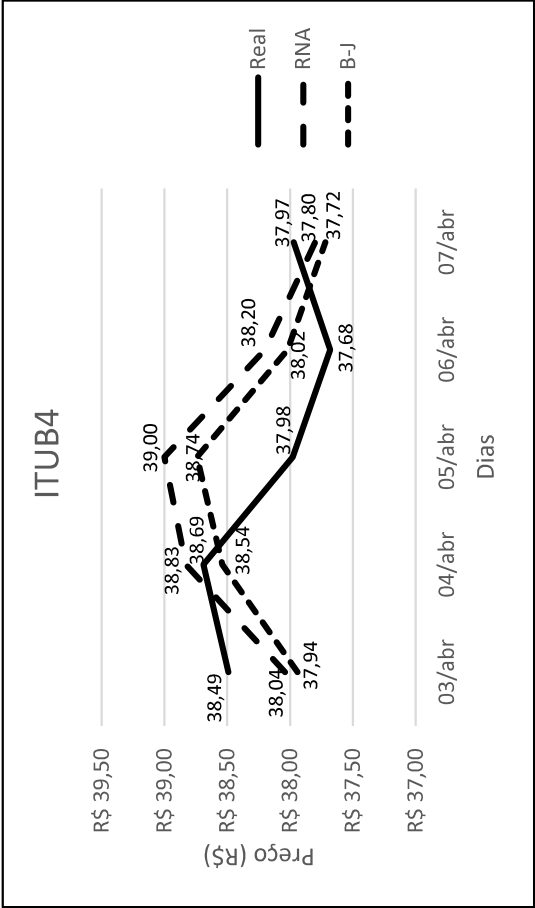
(continua)			
Preço do dia 03/04/2017:		R\$ 9,98	
Método utilizado:		RNA	B-J
12 meses	Valor calculado	R\$ 9,98	R\$ 10,33
	Erro	0,00%	3,51%
24 meses	Valor calculado	R\$ 10,39	R\$ 10,32
	Erro	4,11%	3,41%
36 meses	Valor calculado	R\$ 10,15	R\$ 10,32
	Erro	1,70%	3,41%

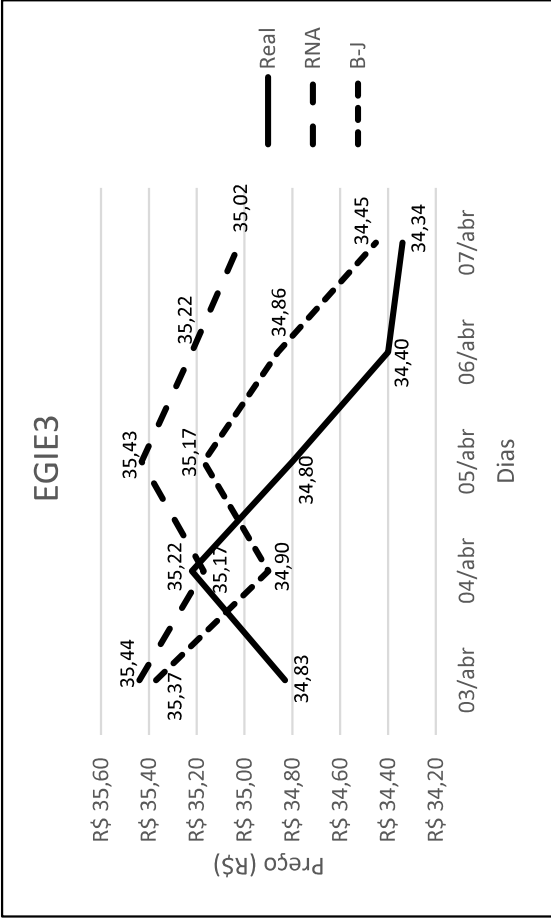
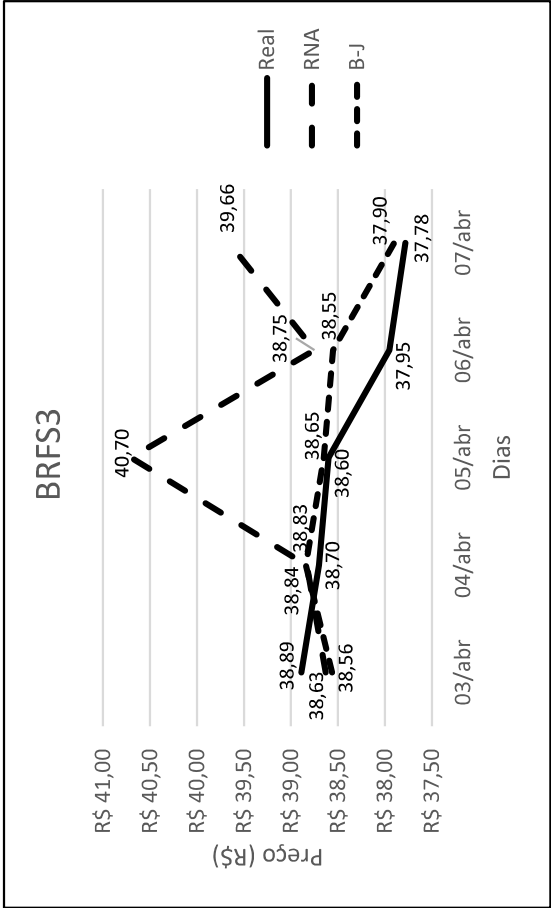
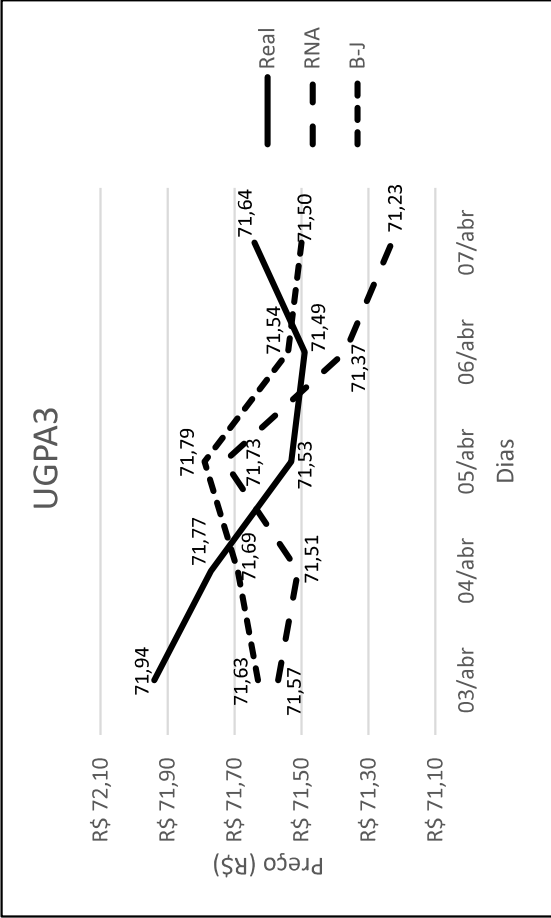
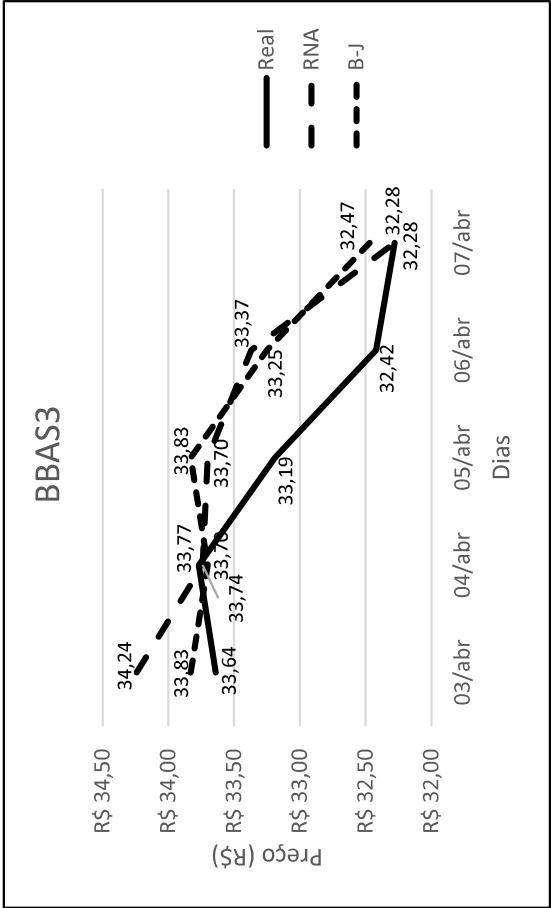
(conclusão)			
Método utilizado:		RNA	B-J
48 meses	Valor calculado	R\$ 10,19	R\$ 10,32
	Erro	2,10%	3,41%
60 meses	Valor calculado	R\$ 10,31	R\$ 10,34
	Erro	3,31%	3,61%
72 meses	Valor calculado	R\$ 10,23	R\$ 10,34
	Erro	2,51%	3,61%
84 meses	Valor calculado	R\$ 10,26	R\$ 10,33
	Erro	2,81%	3,51%
96 meses	Valor calculado	R\$ 10,34	R\$ 10,33
	Erro	3,61%	3,51%
108 meses	Valor calculado	R\$ 10,46	R\$ 10,33
	Erro	4,81%	3,51%
120 meses	Valor calculado	R\$ 10,16	R\$ 10,33
	Erro	1,80%	3,51%

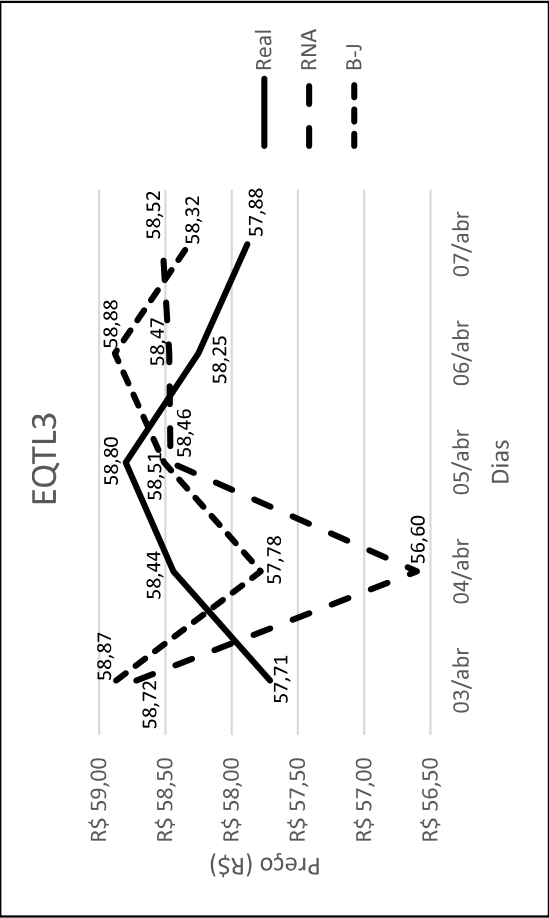
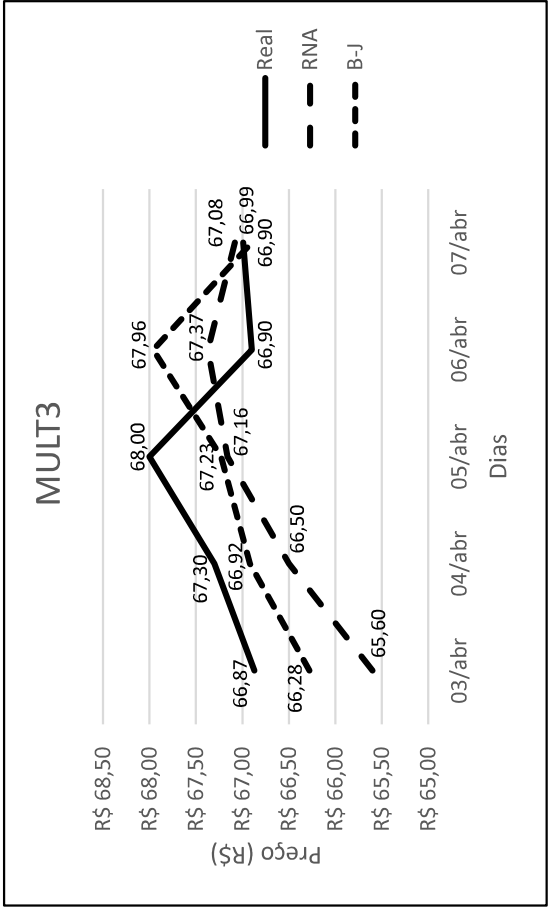
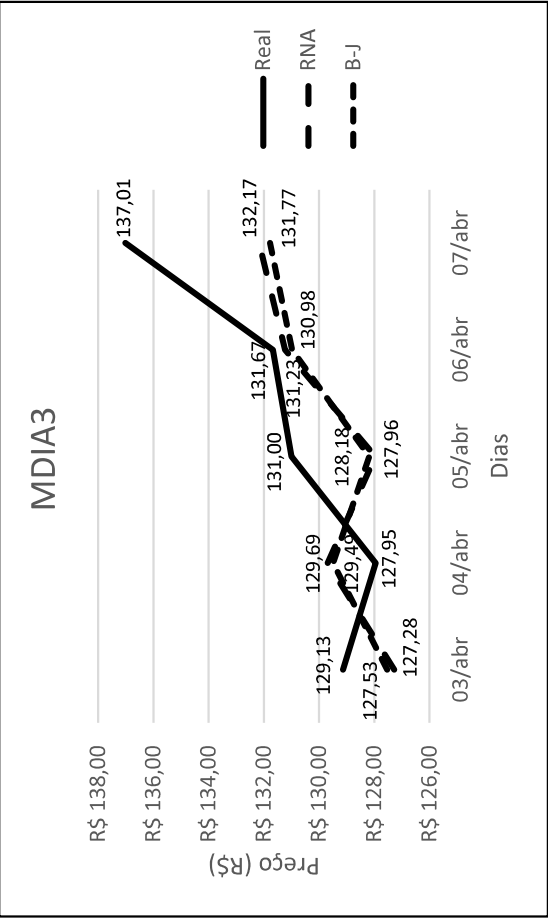
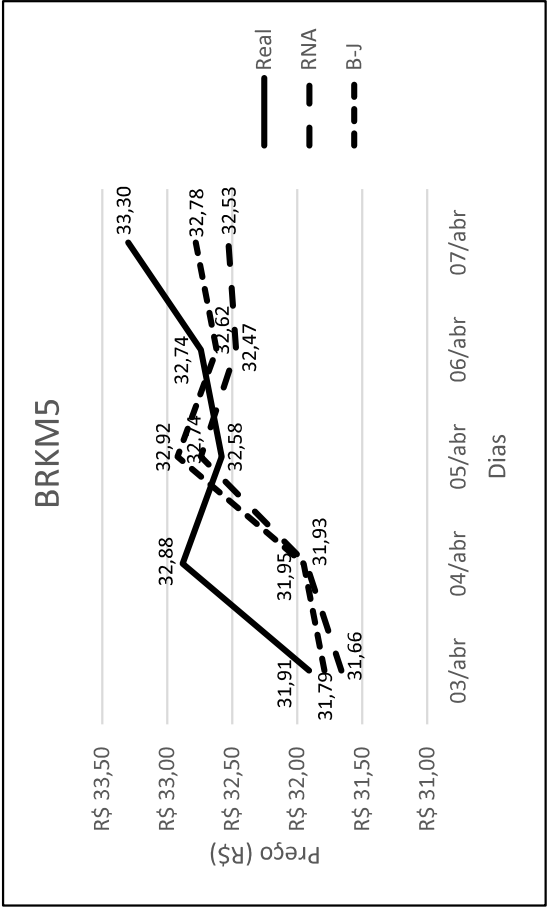
16. TAESA – TAEE11

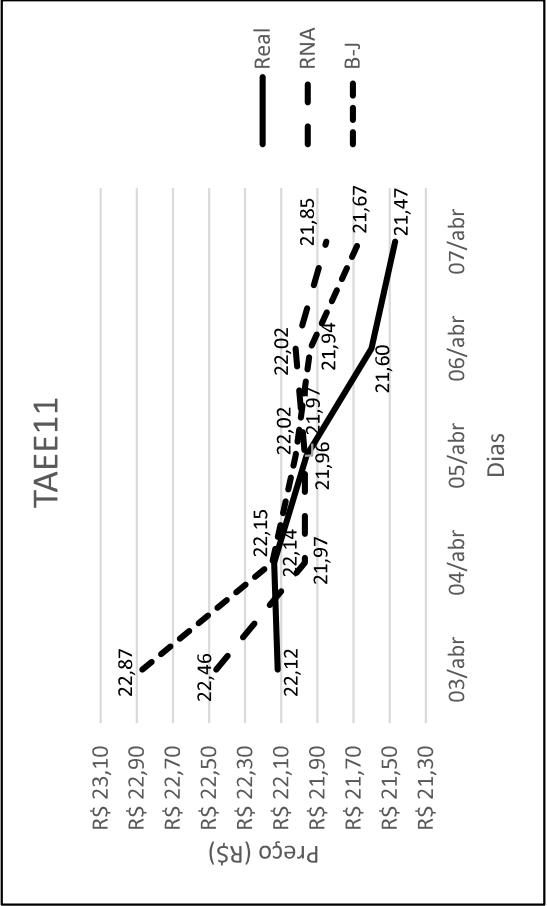
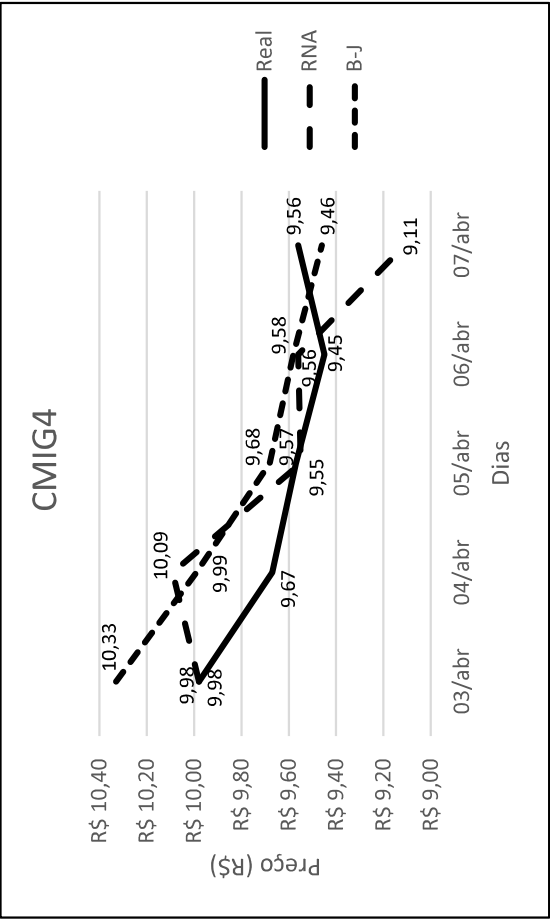
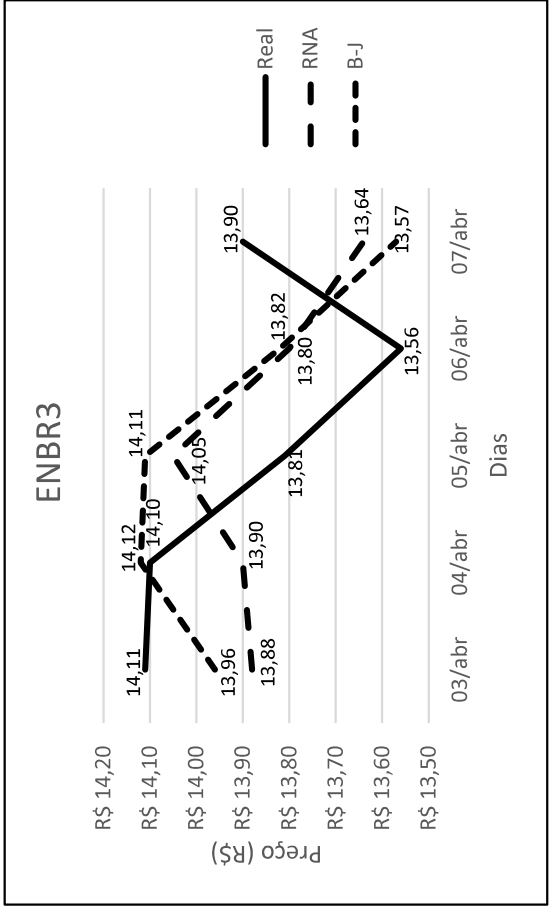
Preço do dia 03/04/2017:		R\$ 22,12	
Método utilizado:		RNA	B-J
12 meses	Valor calculado	R\$ 22,46	R\$ 22,87
	Erro	1,54%	3,39%
24 meses	Valor calculado	R\$ 22,35	R\$ 22,92
	Erro	1,04%	3,62%
36 meses	Valor calculado	R\$ 22,81	R\$ 22,78
	Erro	3,12%	2,98%
48 meses	Valor calculado	R\$ 22,44	R\$ 22,86
	Erro	1,45%	3,35%
60 meses	Valor calculado	R\$ 23,10	R\$ 22,90
	Erro	4,43%	3,53%

APÊNDICE H – PREVISÕES PARA O RESTANTE DA SEMANA









APÊNDICE I – INFORMAÇÕES DAS RNA PARA PREVISÕES PARA O REstante DA SEMANA

Erro médio quadrático e coeficiente de determinação de cada prognose

	04/abr		05/abr		06/abr		07/abr	
	MSE	r^2	MSE	r^2	MSE	r^2	MSE	r^2
ABEV3	0,044	0,997	0,066	0,978	0,044	0,997	0,048	0,995
ITUB4	0,357	0,998	0,366	0,997	0,349	0,999	0,376	0,997
CIEL3	0,506	0,979	0,607	0,963	0,402	0,994	0,683	0,952
BBSE3	0,288	0,990	0,264	0,996	0,278	0,991	0,295	0,987
ITSA4	0,030	0,991	0,022	0,997	0,021	0,997	0,025	0,996
BBAS3	0,429	0,998	0,419	0,998	0,461	0,997	0,554	0,995
UGPA3	0,788	0,984	0,954	0,969	0,743	0,993	0,754	0,988
BRFS3	0,773	0,999	1,727	0,964	0,817	0,998	1,817	0,962
EGIE3	0,249	0,987	0,222	0,995	0,314	0,964	0,410	0,937
BRKM5	0,486	0,998	0,426	0,999	0,417	0,999	0,474	0,998
MDIA3	5,305	0,998	6,810	0,995	5,105	0,999	5,331	0,998
MULT3	0,970	0,996	1,146	0,978	1,028	0,988	1,076	0,986
EQTL3	0,897	0,991	0,620	0,997	0,725	0,995	0,692	0,996
ENBR3	0,057	0,968	0,035	0,998	0,037	0,994	0,038	0,993
CMIG4	0,070	0,992	0,064	0,997	0,162	0,967	0,125	0,978
TAEE11	0,224	0,971	0,328	0,927	0,171	0,987	0,176	0,989

APÊNDICE J – INFORMAÇÕES DO MÉTODO DE BOX-JENKINS PARA PREVISÕES PARA O RESTANTE DA SEMANA

Modelos determinados para cada prognose

	04/abr	05/abr	06/abr	07/abr
ABEV3	ARIMA(1,1,0)	ARIMA(1,1,0)	ARIMA(1,1,0)	ARIMA(1,1,0)
ITUB4	ARIMA(0,1,0)	ARIMA(0,1,0)	ARIMA(0,1,0)	ARIMA(0,1,0)
CIEL3	ARIMA(0,1,0)	ARIMA(0,1,0)	ARIMA(0,1,0)	ARIMA(0,1,0)
BBSE3	ARIMA(0,1,0)	ARIMA(0,1,0)	ARIMA(0,1,0)	ARIMA(0,1,0)
ITSA4	ARIMA(0,1,0)	ARIMA(0,1,0)	ARIMA(0,1,0)	ARIMA(0,1,0)
BBAS3	ARIMA(0,1,0)	ARIMA(0,1,0)	ARIMA(0,1,0)	ARIMA(0,1,0)
UGPA3	ARIMA(0,1,0)	ARIMA(0,1,0)	ARIMA(0,1,0)	ARIMA(0,1,0)
BRFS3	ARIMA(0,1,0)	ARIMA(0,1,0)	ARIMA(0,1,0)	ARIMA(0,1,0)
EGIE3	ARIMA(1,1,0)	ARIMA(1,1,0)	ARIMA(1,1,0)	ARIMA(1,1,0)
BRKM5	ARIMA(0,1,0)	ARIMA(0,1,0)	ARIMA(0,1,0)	ARIMA(0,1,0)
MDIA3	ARIMA(0,1,0)	ARIMA(0,1,10)	ARIMA(0,1,10)	ARIMA(0,1,10)
MULT3	ARIMA(0,1,0)	ARIMA(1,1,1)	ARIMA(1,1,1)	ARIMA(1,1,1)
EQTL3	ARIMA(0,1,0)	ARIMA(0,1,0)	ARIMA(0,1,0)	ARIMA(0,1,0)
ENBR3	ARIMA(0,1,0)	ARIMA(0,1,0)	ARIMA(0,1,0)	ARIMA(0,1,0)
CMIG4	ARIMA(0,1,0)	ARIMA(0,1,0)	ARIMA(0,1,0)	ARIMA(0,1,0)
TAE11	ARIMA(2,1,2)	ARIMA(2,1,2)	ARIMA(2,1,2)	ARIMA(2,1,2)

Erro médio quadrático e coeficiente de determinação de cada prognose

	04/abr		05/abr		06/abr		07/abr	
	MSE	r^2	MSE	r^2	MSE	r^2	MSE	r^2
ABEV3	0,044	0,952	0,044	0,952	0,044	0,951	0,044	0,951
ITUB4	0,339	0,983	0,338	0,983	0,339	0,983	0,338	0,983
CIEL3	0,371	0,959	0,372	0,959	0,372	0,960	0,372	0,961
BBSE3	0,284	0,896	0,283	0,896	0,283	0,896	0,282	0,896
ITSA4	0,020	0,979	0,020	0,979	0,020	0,979	0,020	0,979
BBAS3	0,413	0,987	0,412	0,987	0,412	0,987	0,413	0,987
UGPA3	0,799	0,891	0,796	0,892	0,794	0,892	0,790	0,892
BRFS3	0,757	0,970	0,755	0,970	0,752	0,971	0,750	0,971
EGIE3	0,212	0,949	0,211	0,949	0,211	0,949	0,211	0,949
BRKM5	0,402	0,989	0,404	0,989	0,403	0,989	0,402	0,989
MDIA3	5,063	0,985	4,951	0,986	4,964	0,986	4,946	0,986
MULT3	0,978	0,933	0,972	0,934	0,964	0,936	0,962	0,936
EQTL3	0,572	0,979	0,572	0,979	0,569	0,979	0,569	0,979
ENBR3	0,036	0,960	0,036	0,960	0,036	0,960	0,036	0,960
CMIG4	0,066	0,973	0,066	0,973	0,066	0,973	0,066	0,973
TAE11	0,162	0,954	0,162	0,954	0,161	0,954	0,161	0,954

APÊNDICE L – ERRO DOS RESULTADOS DAS PREVISÕES PARA O RESTANTE DA SEMANA

