

PONTÍFICA UNIVERSIDADE CATÓLICA DE GOIÁS  
PROGRAMA DE PÓS-GRADUAÇÃO *STRICTO SENSU*  
EM ENGENHARIA DE PRODUÇÃO E SISTEMAS

**PREVISÃO DE SÉRIES TEMPORAIS EM UMA INDÚSTRIA METAL  
MECÂNICA UTILIZANDO MÉTODO CLÁSSICO DE BOX-JENKINS E  
REDES NEURAS ARTIFICIAIS MLP**

RAFAEL GOMES LOIOLA

MARÇO  
2016

**PREVISÃO DE SÉRIES TEMPORAIS EM UMA INDÚSTRIA METAL  
MECÂNICA UTILIZANDO MÉTODO CLÁSSICO DE BOX-JENKINS E  
REDES NEURAS ARTIFICIAIS MLP**

**RAFAEL GOMES LOIOLA**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia de Produção e Sistemas da Pontifícia Universidade Católica de Goiás, como parte dos requisitos para obtenção do título de Mestre em Engenharia de Produção e Sistemas.

Orientadora: Prof.<sup>a</sup> Maria José Pereira Dantas, Dr.<sup>a</sup>.

MARÇO  
2016

Dados Internacionais de Catalogação da Publicação (CIP)  
(Sistema de Bibliotecas PUC Goiás)

L834p      Loiola, Rafael Gomes.  
Previsão de séries temporais em uma indústria metal  
mecânica utilizando método clássico de Box-Jenkins e redes  
neurais artificiais MLP [manuscrito] / Rafael Gomes Loiola –  
Goiânia, 2016.  
120 f. ; 30 cm.

Dissertação (mestrado) – Pontifícia Universidade Católica  
de Goiás, Programa de Pós-Graduação *Stricto Sensu* em  
Engenharia de Produção e Sistemas, 2016.

“Orientadora: Profa. Dra. Maria José Pereira Dantas”.

Bibliografia.

1. Previsão de séries temporais. 2. Métodos estatísticos.  
3. Redes neurais artificiais. I. Título.

CDU 004.032.26(043)

**PREVISÃO DE SÉRIES TEMPORAIS EM UMA INDÚSTRIA METAL  
MECÂNICA UTILIZANDO MÉTODO CLÁSSICO DE BOX-JENKINS E  
REDES NEURAIS ARTIFICIAIS MLP**

**RAFAEL GOMES LOIOLA**

Esta Dissertação foi julgada para obtenção do título de Mestre em Engenharia de Produção e Sistemas, pelo Programa de Pós-Graduação em Engenharia de Produção e Sistemas da Pontifícia Universidade Católica de Goiás em março de 2016.

---

Prof. Ricardo Luiz Machado, Dr.

Coordenador do Programa de Pós-Graduação em  
Engenharia de Produção e Sistemas

**Banca Examinadora:**

---

Prof.<sup>a</sup> Maria José Pereira Dantas, Dr.<sup>a</sup>

Orientadora

---

Prof. Iran Martins do Carmo, Dr.

Membro externo

---

Prof. Sibélius Lellis Vieira, Dr.

Membro interno

GOIÂNIA – GOIÁS

MARÇO DE 2016

# DEDICATÓRIA

À minha família.

## **AGRADECIMENTOS**

Agradeço primeiramente a Deus, por ter me ajudado e permitido chegar até aqui.

Aos meus pais, por terem me proporcionado todos esses anos de aprendizado e pela confiança em mim depositada de que concluiria mais esta etapa.

À minha irmã, que sempre torceu por mim mesmo estando longe, sempre acreditou no meu potencial e me ajudou quando necessário.

À minha namorada, Ana Flávia, por ter me apoiado nos momentos de maior dificuldade, sempre buscando me trazer motivação para continuar, me auxiliando com muita paciência, compreensão, carinho e ficando ao meu lado durante todo este tempo.

À minha orientadora, professora Dr<sup>a</sup> Maria José Pereira Dantas, pela confiança que depositou em mim, apoio e compartilhamento de ensinamentos durante todo este tempo.

Ao meu cunhado, Diogo, por sempre estar disposto a me auxiliar e incentivar durante toda esta caminhada mesmo estando longe.

Aos meus amigos Paulo Henrique Borba, Agenor Sousa e Marina Meireles pelo compartilhamento de sugestões pertinentes e valiosas para finalização deste trabalho.

À empresa pela contribuição dos dados para o estudo e por todo o apoio e incentivo que permitiram a conclusão deste trabalho.

A todos os que não foram citados aqui, porém, de alguma forma, contribuíram para que esta conquista se realizasse.

## RESUMO

A previsão de demanda é de essencial importância em ambientes organizacionais, de forma a servir como ferramenta de apoio a tomada de decisão durante o desenvolvimento do planejamento estratégico das empresas. Este trabalho teve como principal objetivo comparar modelos estatísticos e de inteligência artificial para problemas de previsão de demanda utilizando séries temporais por meio dos métodos de Box-Jenkins e rede neural artificial Multilayer Perceptron (MLP). Realizou-se o estudo para identificação e definição dos principais métodos de previsão de demanda. Posteriormente, aplicaram-se os métodos de previsão selecionados para a análise dos três produtos mais relevantes de uma indústria metal mecânica, no período de 2012 até 2014. Os quatro últimos períodos da série foram utilizados apenas para validação de desempenho de ambos os métodos propostos através das análises dos erros de previsão. Os *softwares* R, Matlab e SPSS apoiaram a aplicação, modelagem e análise dos dados. A partir dos modelos, realizou-se a previsão um passo a frente das vendas de uma indústria metal mecânica e posteriormente fez-se o comparativo de seus resultados através das medidas de erros referentes à raiz quadrada do erro quadrático médio, RMSE, e o erro percentual absoluto médio, MAPE, para identificar o modelo mais satisfatório e adequado para a predição. Os resultados indicaram que o desempenho das previsões utilizando o método estatístico de Box-Jenkins nos Produtos 1 e 3 foram superiores à aplicação dos modelos de rede neural MLP. Enquanto que para o Produto 2, o método de redes neurais alcançou melhores resultados. Nas análises estatísticas verificou-se que as séries apresentam padrões de comportamento referente à sazonalidade e oscilações, sendo possível observar que ambos os métodos apresentam resultados satisfatórios para cada característica de dados das séries temporais estudadas.

**Palavras-chave:** Séries temporais, Box-Jenkins, Redes Neurais MLP, Indústria, Previsão de demanda.

## ABSTRACT

The demand forecasting is of essential importance for business environments, in a way to serve as a decision making supporting tool during the development of companies strategic planning. This work strived to compare statistics with artificial intelligence methods applied to provisioning on demand issues using temporal series through Box-Jenkins and Artificial Neural Networks Multilayer Perceptron (MLP) methods. Studies were performed to identify and define the main demand forecasting methods. Subsequently, the selected prediction methods for the analysis of the three most relevant products of a metalworking industry were applied in the period 2012 to 2014. The four last periods were used only for performance validation of both methods, through the analysis of forecast errors. Softwares R, Matlab and SPSS supported the data deployment, modeling and analysis. From those models, a step ahead provisioning of sales of a metal mechanic industry was performed, followed by the comparison of the errors of each method based on root mean squared error, RMSE, and mean absolute percentage error, MAPE, to identify the most satisfactory and adequate provisioning method. The results indicated that the performance of the forecasts using the statistical method of Box-Jenkins in Products 1 and 3 were higher than the application of the MLP neural network models. While, for Product 2 the method of neural networks achieved better results. In the statistics analysis, one could verify that the series present some behavior patterns associated to seasonality and oscillations, being possible to observe that both methods show satisfactory results for each data characteristics of the temporal series.

**Keywords:** *Temporal series, Box-Jenkins, Artificial Neural Networks MLP, Industry, Demand forecasting.*



## LISTA DE FIGURAS

Figura 2.1 – Métodos de previsão .....	26
Figura 2.2 – Modelo estatístico usado nas previsões .....	30
Figura 2.3 – Período dos dados utilizados na previsão .....	30
Figura 2.4 – Técnicas mais utilizadas na previsão .....	30
Figura 2.5 – Fluxograma do ciclo iterativo de Box & Jenkins .....	32
Figura 2.6 – Modelo de rede neural biológica .....	40
Figura 2.7 – Modelo computacional de rede neural .....	42
Figura 2.8 – Modelo de rede <i>feedforward</i> simples. ....	45
Figura 2.9 – Modelo de rede <i>feedforward</i> de camadas múltiplas. ....	45
Figura 2.10 – Arquitetura rede neural de múltiplas camadas .....	47
Figura 3.1 – Classificação da pesquisa científica em Engenharia de Produção .....	53
Figura 3.2 – Processo de modelagem .....	54
Figura 4.1 – Série temporal real do Produto 1 .....	61
Figura 4.2 – BoxPlot anual do Produto 1 .....	62
Figura 4.3 – BoxPlot anual após tratamento de <i>outliers</i> do Produto 1 .....	63
Figura 4.4 – Série temporal após tratamento de <i>outliers</i> do Produto 1 .....	63
Figura 4.5 – Função de autocorrelação do Produto 1 .....	65
Figura 4.6 – Função de autocorrelação parcial do Produto 1 .....	65
Figura 4.7 – Histograma do Produto 1 .....	66
Figura 4.8 – Correlogramas de Padronização e ACF dos resíduos do Produto 1 .....	67
Figura 4.9 – <i>P-values</i> para teste Ljung-Box .....	68
Figura 4.10 – QQ-Plot dos resíduos para o Produto 1 .....	69
Figura 4.11 – Comparativo entre a demanda real e predita do Produto 1 .....	69
Figura 4.12 – Previsão método Box-Jenkins do Produto 1 .....	70

Figura 4.13 – Série temporal real do Produto 2 .....	71
Figura 4.14 – BoxPlot anual do Produto 2 .....	71
Figura 4.15 – Função de autocorrelação do Produto 2 .....	72
Figura 4.16 – Função de autocorrelação parcial do Produto 2 .....	73
Figura 4.17 – Histograma do Produto 2 .....	73
Figura 4.18 – Correlogramas de Padronização e ACF dos resíduos do Produto 2 .....	74
Figura 4.19 – <i>P-values</i> para teste Ljung-Box .....	75
Figura 4.20 – QQ-Plot dos resíduos para o Produto 2 .....	76
Figura 4.21 – Comparativo entre a demanda real e predita do Produto 2 .....	76
Figura 4.22 – Previsão método Box-Jenkins do Produto 2 .....	77
Figura 4.23 – Série temporal real do Produto 3 .....	77
Figura 4.24 – Série temporal após tratamento de <i>outliers</i> do Produto 3 .....	78
Figura 4.25 – BoxPlot anual sem tratamento de <i>outliers</i> do Produto 3 .....	78
Figura 4.26 – BoxPlot anual após 1º tratamento de <i>outliers</i> do Produto 3 .....	79
Figura 4.27 – BoxPlot anual após último tratamento de <i>outliers</i> do Produto 3 .....	79
Figura 4.28 – Função de autocorrelação do Produto 3 .....	80
Figura 4.29 – Função de autocorrelação parcial do Produto 3 .....	81
Figura 4.30 – Histograma do Produto 3 .....	81
Figura 4.31 – Correlogramas de Padronização e ACF dos resíduos do Produto 3 .....	82
Figura 4.32 – <i>P-values</i> para teste Ljung-Box .....	83
Figura 4.33 – QQ-Plot dos resíduos para o Produto 3 .....	84
Figura 4.34 – Comparativo entre a demanda real e predita do Produto 3 .....	84
Figura 4.35 – Previsão método Box-Jenkins do Produto 3 .....	85
Figura 4.36 – Treinamento, teste e previsão da Rede Neural MLP - Produto 1 .....	87
Figura 4.37 – Treinamento, teste e previsão da Rede Neural MLP - Produto 2 .....	88
Figura 4.38 – Treinamento, teste e previsão da Rede Neural MLP - Produto 3 .....	89

## LISTA DE TABELAS

Tabela 2.1: Exemplos de técnicas para previsão de demanda. ....	27
Tabela 4.1 – Medidas descritivas da série de dados do Produto 1 .....	64
Tabela 4.2 – Descrição do modelo ARIMA do Produto 1 .....	67
Tabela 4.3 – Teste de autocorrelação Ljung-Box dos resíduos do Produto 1 .....	68
Tabela 4.4 – Teste de normalidade Shapiro-Wilk dos resíduos do Produto 1 .....	68
Tabela 4.5 – Medidas descritivas da série de dados do Produto 2 .....	72
Tabela 4.6 – Descrição do modelo ARIMA do Produto 2 .....	74
Tabela 4.7 – Teste de autocorrelação Ljung-Box dos resíduos do Produto 2 .....	75
Tabela 4.8 – Teste de normalidade Shapiro-Wilk dos resíduos do Produto 2 .....	75
Tabela 4.9 – Medidas descritivas da série de dados do Produto 3 .....	80
Tabela 4.10 – Descrição do modelo ARIMA do Produto 3 .....	82
Tabela 4.11 – Teste de autocorrelação Ljung-Box dos resíduos do Produto 3 .....	83
Tabela 4.12 – Teste de normalidade Shapiro-Wilk dos resíduos do Produto 3 .....	83
Tabela 4.13 – Resumo comparativo de erros dos modelos estudados .....	90

## LISTA DE SIGLAS

ACF – *Autocorrelation Function*

AR – *AutoRegressive*

ARIMA – *AutoRegressive Integrated Moving Average*

ARMA – *AutoRegressive Moving Average*

MA – *Moving Average*

MAPE – *Mean Average Percentage Error*

MATLAB – *Matrix Laboratory*

MLP – *MultiLayer Perceptron*

MSE – *Mean Square Error*

PACF – *Partial Autocorrelation Function*

PCP – *Planejamento e Controle de Produção*

RBF – *Radial Base Function*

RMSE – *Root Mean Square Error*

RNA – *Redes Neurais Artificiais*

SARIMA – *Seasonal Auto Regressive Integrated Moving Average*

SKU – *Stock Keeping Unit*

SLA – *Service Level Agreement*

SPSS – *Statistical Package for the Social Sciences*

TDNN – *Time Delay Neural Network*

UFLA – *Universidade Federal de Lavras*

VLSI – *Very Large Scale Integration*

# SUMÁRIO

<b>1. INTRODUÇÃO .....</b>	<b>14</b>
1.1 PROBLEMÁTICA DA PESQUISA .....	16
1.2 OBJETIVO GERAL .....	17
1.3 OBJETIVOS ESPECÍFICOS .....	17
1.4 JUSTIFICATIVA .....	18
1.5 ESTRUTURA DO TRABALHO .....	21
<b>2. REVISÃO BIBLIOGRÁFICA.....</b>	<b>22</b>
2.1 SÉRIES TEMPORAIS.....	22
2.2 PREVISÃO DE DEMANDA .....	24
2.3 MÉTODOS DE PREVISÃO DE DEMANDA .....	25
2.4 IMPLANTAÇÃO DE SISTEMAS DE PREVISÃO.....	28
2.5 MODELOS DE BOX-JENKINS .....	31
2.5.1 Função de autocorrelação (ACF) e autocorrelação parcial (PACF) .....	35
2.5.2 Componente autorregressivo (AR) .....	35
2.5.3 Componente de médias móveis (MA) .....	36
2.5.4 Modelo autorregressivo de médias móveis (ARMA) .....	37
2.5.5 Modelo autorregressivo integrado de médias móveis (ARIMA) .....	37
2.5.6 Modelo autorregressivo integrado de médias móveis e componentes sazonais (SARIMA).....	38
2.6 REDES NEURAS .....	38
2.6.1 Principais arquiteturas de redes neurais artificiais .....	44
2.6.2 Processo de aprendizagem .....	46
2.7 REDE NEURAL MULTI-LAYER PERCEPTRON (MLP) .....	47
2.8 TRANSFORMAÇÃO DE YEO-JOHNSON.....	51
2.9 ERROS DE PREVISÃO .....	51
<b>3. METODOLOGIA .....</b>	<b>53</b>
3.1 CARACTERIZAÇÃO DA PESQUISA.....	53
3.2 METODOLOGIA DA PESQUISA .....	55
3.3 DADOS.....	56
3.4 PROGRAMA SPSS .....	56
3.5 PROGRAMA R .....	56
3.6 PROGRAMA MATLAB.....	57
3.7 MÉTODO DE SELEÇÃO DOS DADOS .....	57
3.8 MODELOS DE PREVISÃO .....	58
3.8.1 Modelo ARIMA.....	58
3.8.2 Modelo de Redes Neurais .....	59
<b>4. RESULTADOS E DISCUSSÕES .....</b>	<b>61</b>

4.1 ANÁLISE ESTATÍSTICA DO PRODUTO 1 .....	61
4.2 MÉTODO ARIMA PARA O PRODUTO 1 .....	66
4.3 ANÁLISE ESTATÍSTICA DO PRODUTO 2 .....	70
4.4 MÉTODO ARIMA PARA O PRODUTO 2 .....	73
4.5 ANÁLISE ESTATÍSTICA DO PRODUTO 3 .....	77
4.6 MÉTODO ARIMA PARA O PRODUTO 3 .....	82
4.7 REDES NEURAIS .....	85
4.8 ANÁLISE COMPARATIVA DOS MODELOS .....	90
<b>5. CONCLUSÃO .....</b>	<b>92</b>
<b>REFERÊNCIAS .....</b>	<b>95</b>
<b>APÊNDICE A – DEMANDA DE PRODUTOS .....</b>	<b>101</b>
<b>APÊNDICE B – CÓDIGOS NO MATLAB .....</b>	<b>102</b>
<b>APÊNDICE C – CURVA AB DOS PRODUTOS.....</b>	<b>112</b>
<b>APÊNDICE C – CURVA AB DOS PRODUTOS (CONTINUAÇÃO) .....</b>	<b>113</b>
<b>APÊNDICE D – SÉRIE TEMPORAL REAL DOS PRODUTOS X SÉRIE PROCESSADA PELA RNA MLP.....</b>	<b>114</b>
<b>APÊNDICE D – SÉRIE TEMPORAL REAL DOS PRODUTOS X SÉRIE PROCESSADA PELA RNA MLP (CONTINUAÇÃO) .....</b>	<b>115</b>
<b>APÊNDICE D – SÉRIE TEMPORAL REAL DOS PRODUTOS X SÉRIE PROCESSADA PELA RNA MLP (CONTINUAÇÃO) .....</b>	<b>116</b>
<b>APÊNDICE E – SAÍDAS GRÁFICAS DOS TRÊS PRODUTOS GERADAS PELO SOFTWARE MATLAB .....</b>	<b>117</b>
<b>APÊNDICE E – SAÍDAS GRÁFICAS DOS TRÊS PRODUTOS GERADAS PELO SOFTWARE MATLAB (CONTINUAÇÃO) .....</b>	<b>118</b>
<b>APÊNDICE F – CÓDIGOS NO R.....</b>	<b>119</b>

## 1. INTRODUÇÃO

Atualmente, o mercado vem exigindo das empresas propriedades imprescindíveis para sua sobrevivência diante de um ambiente no qual prevalecem características como alta competitividade, incertezas e instabilidades. Desta forma, buscam-se situações que garantam maiores graus de precisão e certeza nas tomadas de decisão desempenhadas pela área gerencial.

Estas áreas precisam de orientações quanto aos direcionamentos futuros a serem tomados, os quais influenciarão diretamente no planejamento de cenários, analisando horizontes de tempo de curto, médio e longo prazo. Para Georgoff e Murdick (1986), uma orientação apropriada aumenta as chances de sucesso das organizações, evita perigos futuros, auxilia na manutenção das vantagens competitivas, influencia as organizações em escolhas e investimentos corretos, e favorece o aproveitamento de oportunidades.

O gerenciamento da demanda tem uma grande relevância na sincronização da cadeia de abastecimento e várias empresas já estão adotando sistemas baseados em *forecasting* colaborativo com o intuito de envolver todos os elos da cadeia de abastecimento na coleta de dados e administração das informações. O resultado almejado é uma rede competitiva de alto desempenho, flexível, ágil e independente da variabilidade da demanda, quantidade de SKU's (*stock keeping units*), SLA (*Service Level Agreement*) ou complexidade. (BARBOSA e CHAVES, 2012).

Para Castro (2001), a predição de séries temporais é um estudo de extrema relevância, visto que são encontrados vários exemplos de sua utilização, na natureza (em campos como geofísica, astrofísica e meteorologia), em ciências sociais (em campos como a demografia), nas ciências médicas (em estudos de processos fisiológicos involuntários), na área de ciências econômicas (voltadas para acompanhamento das taxas de câmbio de moedas e mercado de ações) e nas diversas engenharias (a qual será o foco da pesquisa).

As empresas direcionam suas estratégias de acordo com o que elas acham que o seu negócio irá trazer de resultados. Estas estratégias são definidas através de previsões, sendo a previsão de demanda a principal delas. A previsão de demanda é a base para o planejamento estratégico das áreas de produção, vendas e finanças de toda e qualquer empresa. A partir deste princípio, desenvolvem-se nas empresas os planos de capacidade, de vendas, fluxo de caixa, produção, estoques, mão de obra, compras, entre outros. A grande importância das previsões ocorre a partir da antecipação de cenários e planejamento adequado de ações.

A previsão de demanda é de fundamental importância em ambientes organizacionais, de forma a servir como ferramenta de apoio a tomada de decisão durante o desenvolvimento do planejamento estratégico das empresas, tal como, por exemplo, na definição de quantidades a serem produzidas e vendidas, além de estratégias de definição dos preços a serem cobrados e redução dos possíveis gastos desnecessários.

De acordo com Sharda (1994), o funcionamento dos modelos de previsão de demanda utilizando redes neurais é considerado mais flexível quando comparados aos métodos estatísticos, visto que o primeiro assume uma relação subjacente entre as entradas, as quais são representadas por séries históricas passadas e em alguns casos outras variáveis relevantes e as saídas que serão os valores futuros encontrados.

Segundo Slack, Chambers e Johnston (2009), para realizar o prévio conhecimento do comportamento do consumo do mercado é imprescindível que se faça um dimensionamento adequado das necessidades da empresa. Tal conhecimento pode ser obtido por meio de previsões. As técnicas de previsão de demanda são frequentemente utilizadas para auxiliar no processo de tomada de decisão.

Para Tubino (2009), normalmente a responsabilidade pela preparação da previsão de demanda é dos setores de Marketing ou Vendas. No entanto, existem dois bons motivos para que a área de Planejamento e Controle de Produção (PCP) entenda como esta atividade é realizada. Primeiramente, o PCP utiliza como principal informação na elaboração de suas atividades, a previsão de demanda. Além disso, a mesma afeta diretamente todo o desempenho esperado de suas funções de planejamento e controle do sistema produtivo. Desta maneira, se torna essencial o conhecimento efetivo de como obteve-se os dados, em quais bases foram empregadas as técnicas de previsão e quais as limitações, para que assim, venha facilitar a comunicação entre as áreas. Além disso, em empresas de pequeno e médio porte não há uma especialização muito avançada das atividades, cabendo à área de PCP elaborar estas previsões.

De acordo com Pellegrini e Fogliatto (2001), os métodos de previsão podem ser classificados em grupos distintos que levam em consideração o tipo de abordagem utilizada: (i) métodos quantitativos de previsão e (ii) métodos qualitativos de previsão. Os métodos quantitativos, também denominados como métodos de *forecasting*, baseiam-se na análise de séries temporais, ou seja, dados que apresentam uma variação ao longo do tempo. A previsão de demanda utilizando métodos quantitativos pode ser feita através de vários modelos matemáticos.



Em contrapartida, têm-se os métodos qualitativos, que baseiam-se em opiniões de especialistas, se tornando vulneráveis a apresentar tendências que venham comprometer a confiabilidade dos seus resultados. Os métodos qualitativos envolvem estimativas subjetivas por meio da opinião de especialistas ou consumidores. Apesar de alguns métodos qualitativos apresentarem técnicas estruturadas, tais como, por exemplo, pesquisas de intenções e Delphi, o processo para obtenção da previsão é subjetivo (MONTGOMERY; JOHNSON; GARDINER, 1990).

O planejamento é necessário a longo, médio e curto prazo, sendo interessante trabalhar com métodos estatísticos e computacionais de previsão que permitam garantir de maneira assertiva e precisa o planejamento do quanto a empresa busca vender seus produtos ou serviços durante um determinado período futuro, pois este será o início de todas as tomadas de decisão da organização a fim de se garantir o atendimento das metas propostas.

Dessa forma, a partir de modelos matemáticos voltados para projeções de séries temporais, busca-se antecipar cenários futuros e garantir maior assertividade no desenvolvimento do planejamento e definição de estratégias.

### **1.1 Problema da pesquisa**

Para Makridakis, Wheelwright e Hyndman (1998), a previsão é uma parte integrante da tomada de atividades de gestão de decisões. Uma organização estabelece metas e objetivos, e então seleciona ações que esperam resultar na execução desses objetivos e metas. A necessidade de previsão está aumentando enquanto as áreas gerenciais tentam reduzir sua dependência do acaso e se tornam mais científicas em lidar com o seu ambiente, uma vez que cada área de uma organização está relacionada com todas as outras, uma previsão boa ou ruim pode afetar toda a organização. Algumas das áreas em que a previsão atualmente tem um papel importante são:

- Programação;
- Aquisição de recursos;
- Determinação dos requisitos de recursos.

A partir do desenvolvimento de um sistema eficaz e aderente à realidade da empresa para a projeção de cenários futuros da demanda de vendas, é possível obter uma evolução e rapidez nas tomadas de decisões. Além disto, a criação de um modelo que forneça uma visão

geral dos itens demandados, considerando, por exemplo, variáveis históricas referentes a volume e faturamento durante um determinado período será de extrema importância para a área de Planejamento e Controle da Produção (PCP), que passaria a ter dados essenciais para a otimização da programação, planejamento e controle de toda a produção, garantindo assim, uma capacidade de maximização da produção. Além disso, a garantia da previsibilidade da demanda afetaria diretamente a própria satisfação do cliente, já que os processos logísticos serão executados da maneira com que foram planejados, resultando em atendimentos eficientes.

Dessa forma, um dos pontos a serem estudados neste trabalho é garantir tomadas de decisões mais otimizadas e o tema central do estudo é a identificação da melhor ferramenta de previsão de demanda por meio da análise de dados históricos disponíveis na empresa.

Assim, a questão principal para a elaboração desta pesquisa é:

Considerando as características dos dados estudados na Empresa avaliada e os métodos de previsão de demanda, Box-Jenkins e Redes Neurais, qual a capacidade de previsão de cada um deles e qual apresenta resultados mais satisfatórios perante a empresa?

## **1.2 Objetivo geral**

O objetivo deste trabalho é desenvolver um estudo utilizando séries temporais, observadas durante o período de 03 anos, em uma empresa localizada na região de Aparecida de Goiânia – GO, a qual será identificada como Empresa X, por motivos de sigilo industrial. Aplicar a metodologia estatística de Box-Jenkins e modelagem de uma Rede Neural, especificamente do tipo *MultiLayer Perceptron* (MLP), com o intuito de realizar a previsão de demanda de vendas de uma empresa do ramo metal mecânico; para que se possa, a partir dos resultados obtidos, apresentar uma proposta para dar suporte principalmente às tomadas de decisão da diretoria durante o planejamento estratégico e o responsável pela área de PCP durante o planejamento mestre e a programação da produção.

## **1.3 Objetivos específicos**

- Realizar a previsão da série temporal utilizando método estatístico clássico de Box-Jenkins;
- Realizar a previsão da série temporal utilizando a técnica de Redes Neurais,

especificamente Redes Neurais *MultiLayer Perceptron* (MLP);

- Analisar os resultados de ambos os métodos;
- Comparar qual dos métodos é mais satisfatório perante aos dados reais.

#### **1.4 Justificativa**

É claramente observado por parte dos executivos de companhias o quanto o planejamento é extremamente importante e comum a qualquer tipo de empresa e em todos os níveis e áreas da organização, independentemente do seu porte ou ramo. De acordo com Pellegrini e Fogliatto (2001), as previsões de demanda desempenham um importante papel em diversas áreas na gestão de organizações, tais como, por exemplo, na área financeira, focando no planejamento da necessidade de recursos, na área de recursos humanos, na área de vendas (no agendamento de promoções) e na área de produção, voltado na operacionalização de diversos aspectos do gerenciamento da produção, gestão de estoques e desenvolvimento de planos agregados de produção.

De maneira geral, quase todas as empresas devem prever sua produção sob pena de ter, seja na compra de matéria-prima, na alocação de funcionários, no tempo de trabalho das máquinas e até mesmo no estoque, seus custos aumentados.

A previsão de demanda é realizada a partir de um conjunto de observações sequenciais ordenadas durante um determinado período do tempo, no qual é denominado como série temporal e pode ser classificada como sendo estocástica ou determinística, de acordo com McCabe e Tremayne (1995).

Dessa forma, realizar a predição de séries temporais utilizando como alternativa a comparação entre modelos matemáticos clássicos com técnicas baseadas em Redes Neurais Artificiais (RNAs) pode trazer contribuições relevantes para o estudo de predição.

De acordo com Soares (2008), a proposta de Connor, Martin e Atlas (1994), foi a primeira na qual buscou-se realizar a combinação de RNAs com modelos clássicos. Neste, os autores sugerem propor um algoritmo de aprendizado para aplicar as redes neurais.

Para Sharda (1994), uma área de grande aplicação para as redes neurais é a previsão de demanda, visto que, ao contrário dos métodos tradicionais que são baseados em métodos estatísticos, as redes neurais são dirigidas para os dados e métodos adaptativos que condicionam as redes a captar relações funcionais entre os dados. Dessa forma, são mais

apropriadas para problemas em que as soluções requerem um conhecimento de difícil especificação, desde que existam dados para que a rede tenha um processo de aprendizado e possa inferir corretamente sobre parte da população, mesmo se os dados apresentarem informações ruidosas.

De acordo com Werbos (1988), as redes treinadas com *backpropagation* superaram os métodos estatísticos tradicionais, como as abordagens de regressão e Box-Jenkins.

Na literatura é possível encontrar inúmeros trabalhos voltados para previsão de demanda e o desempenho dos resultados obtidos varia bastante de acordo com os métodos utilizados.

No trabalho desenvolvido por Figueredo (2008), realiza-se um estudo de caso utilizando os dados históricos de uma empresa em Santa Catarina, no qual focou-se no Planejamento e Controle da Produção (PCP) de produtos. Utilizou-se como período de pesquisa oito anos, contabilizando os valores de venda mensais dos dez produtos de maior expressividade da empresa, sendo que para cada um dos 10 produtos foram analisadas 96 valores. O foco de estudo da pesquisa se deu a partir da combinação da metodologia de Box-Jenkins para a previsão de séries temporais e a técnica conhecida como Redes Neurais de Funções de Bases Radiais (RBF). Neste trabalho concluiu-se que a metodologia de Redes Neurais apresenta maior precisão para previsão de series temporais, já que apresentou um valor menor do RMSE.

Zoucas (2009) estudou a importância da aplicação de previsão de séries temporais no segmento industrial e utilizou como ferramenta de cálculo a aplicação de Redes Neurais, a fim de analisar o comportamento e a precisão de demanda de onze segmentos industriais distintos. Utilizou-se como conteúdo bibliográfico a importância das técnicas de previsão em séries temporais, classificação das técnicas, etapas para implantação de previsão, entre outros tópicos referentes ao estudo das técnicas mais estudadas atualmente. A partir da análise prévia dos dados coletados, definiu-se a utilização de uma Rede *Perceptron* Multi Camadas com atraso no tempo, conhecida também como TDNN.

Villamagna (2013) analisou séries mensais de consumo e demanda de energia elétrica da Universidade Federal de Lavras (UFLA), durante o período de janeiro de 1995 a dezembro de 2011, com o intuito de prever os valores futuros. Nesta pesquisa, a autora utilizou como ferramentas os métodos de Box-Jenkins e Redes Neurais Artificiais. Apesar da metodologia de Box-Jenkins apresentar-se eficiente para representar a série analisada, os resultados

mostrados pelas Redes Neurais Artificiais obtiveram melhores resultados, apresentando como um método satisfatório para a previsão de demanda.

Outro estudo que buscou a comparação do desempenho entre técnicas de previsão é apresentado por Viglioni (2006), no qual realizou-se a comparação das técnicas clássicas de previsão de curto prazo (Média Móvel, Suavização Exponencial, Suavização Exponencial com Séries de Tendências (método de Holt) e Regressão Linear) com Redes Neurais Artificiais (RNA), utilizando o conceito de rede neural do tipo *feedforward* com algoritmo *back-propagation* para analisar a solicitação diária de transporte de uma ferrovia brasileira. Realizou-se um estudo de caso, com base na série histórica de solicitações de transporte ferroviário no Brasil, comparou-se criticamente o desempenho de ambas as técnicas e chegou-se a conclusão que as redes neurais apresentaram melhores resultados para as previsões de demanda solicitadas pelos clientes.

Pasquotto (2010) explorou a aplicação de redes neurais de Elman na previsão de três séries temporais mensais do varejo de bens e serviços no Brasil nas quais as variáveis desta série estão relacionadas com a demanda de produtos farmacêuticos, adubos e tráfego aéreo. As previsões utilizando redes neurais de Elman posteriormente foram comparadas com os modelos utilizando a metodologia clássica Box-Jenkins. Para se verificar a qualidade de cada modelo utilizou-se o cálculo do erro médio quadrático juntamente do coeficiente de Theil. Nesta pesquisa concluiu-se que a previsão utilizando o método de Box-Jenkins foi superior a rede neural de Elman para todas as três variáveis analisadas e que houve um impacto negativo na qualidade da previsão das redes neurais, devido a presença de tendência e sazonalidade das séries.

Para Mohammadi, Fatemi e Zeinali (2014), a otimização das precisões nas previsões de series temporais tem recebido uma atenção considerável nos últimos anos. Dessa forma, os autores apresentaram um novo algoritmo evolutivo híbrido para determinar a arquitetura (variáveis de entrada e neurônios da camada escondida) e os parâmetros de Redes Neurais de Funções de Bases Radiais simultaneamente. Neste estudo, verificou-se que modelo híbrido desenvolvido foi bastante satisfatório para a resolução de problemas de previsão de demanda e uma ótima alternativa de otimização nas predições.

Um dos pontos necessários para a melhoria dos processos produtivos é desenvolver e aplicar modelos quantitativos de previsão, que melhor se adequam aos valores de produção ou de demanda. Dessa forma, usando como referência os dados históricos de vendas, buscou-se prever a quantidade de produtos de uma empresa do ramo metal mecânico. A partir dos

valores obtidos nas previsões, utilizar uma das metodologias aplicadas, com o intuito de elaborar uma política de planejamento mais adequada e eficiente às necessidades do mercado na qual a empresa está inserida.

## 1.5 Estrutura do trabalho

A dissertação está estruturada em cinco capítulos, sendo que o conteúdo de cada um deles, a partir do capítulo de revisão bibliográfica, está descrito na sequência.

O capítulo 2, referente a revisão bibliográfica, apresenta toda a fundamentação teórica pertinente com a temática principal da pesquisa, sendo esta voltada principalmente para o modelo estatístico de previsão de séries temporais Box-Jenkins, além da revisão voltada para redes neurais, com ênfase nas redes neurais MLP.

O capítulo 3 apresenta a metodologia da pesquisa, envolvendo toda sua caracterização, o tipo de série temporal escolhida, os métodos de aplicação dos modelos utilizados para a previsão de demanda e os resultados de desempenho de cada um deles.

Durante o capítulo 4 são exibidos os resultados alcançados com a aplicação dos métodos juntamente com as discussões e comentários críticos de análise de desempenho.

O capítulo 5, e último, apresenta todas as conclusões obtidas durante o desenvolvimento da pesquisa e possíveis sugestões de melhoria para trabalhos futuros.

Após o capítulo 5, são apresentadas as referências, seguido de seis apêndices utilizados para o desenvolvimento da pesquisa, nos quais apresentam demanda de produtos, códigos utilizados, curva ABC e os comparativos entre as saídas reais e obtidas para cada um dos modelos.

Os cenários de previsão de demanda são aplicados para prever a venda de produtos em uma indústria do ramo metal mecânico, que produz peças mecânicas em alumínio, tais como: engates rápidos, polias, rodas, conexões hidráulicas, molas pneumáticas (*airspring*), dentre outras. A empresa tem bastante interesse no estudo em questão, e por isso forneceu todos os dados necessários para a pesquisa.

Desta forma, busca-se propor o desenvolvimento de modelos de previsão de demanda das vendas a fim de auxiliar a tomada de decisão dos executivos da empresa com o intuito de estimar toda a produção em longo prazo, além de sua capacidade fabril.

## 2. REVISÃO BIBLIOGRÁFICA

A revisão bibliográfica tem como objetivo embasar esse trabalho com conceitos encontrados na literatura sobre temas relevantes para sua realização.

Os temas importantes que serão abordados neste trabalho são o comparativo entre métodos estatísticos de previsão de demanda, Box-Jenkins, com métodos utilizando redes neurais artificiais, especificamente rede neural MLP.

### 2.1 Séries temporais

Para Morettin e Toloí (1986), denomina-se série temporal um conjunto de observações ordenadas durante um determinado período de tempo. Alguns exemplos de séries temporais são:

- Valores diários do preço das ações de uma empresa;
- Valores mensais de temperatura de uma cidade;
- Valores mensais do índice de produtos no país.

Ainda conforme Morettin e Toloí (1986), uma série temporal pode ser definida como qualquer conjunto de observações ordenadas no tempo e compostas por quatro elementos:

- Tendência: sentido de deslocamento da série ao longo de vários anos;
- Ciclo: movimento que tende a ser periódico ao longo de vários anos;
- Sazonalidade: movimento ondulatório de curta duração, em geral, inferior a um ano;
- Erro ou ruído aleatório: variabilidade intrínseca aos dados que não pode ser modelado.

Dentre as séries temporais, têm-se as séries discretas e contínuas. Quando se fala em séries temporais discretas, nota-se que as mesmas partem de observações realizadas em instantes de tempo igualmente espaçadas, porém, sendo possível que suas observações se apresentem de maneira irregular. Ainda de acordo com Morettin e Toloí (1986), é importante ressaltar que mesmo que uma série seja obtida continuamente durante um intervalo de tempo de amplitude  $T$ , ainda assim, será necessário transformá-la em uma série discreta, por meio da amostragem em intervalos de tempo equiespaçados.

De acordo com Souza (2006) pode-se classificar a formação das séries temporais como determinística, ou seja, quando os valores futuros constituídos por meio da função matemática do tipo  $y = f(\text{tempo})$  ou estocástica, quando esses valores são expostos por termos probabilísticos, uma vez que a série está descrita por meio de uma função na qual não envolve apenas o tempo, mas também uma variável aleatória do tipo  $y = f(\text{tempo}, y_t)$ , onde  $y_t$  se refere ao termo aleatório residual, utilizado no momento que em não se consegue explicar completamente algum movimento irregular da série apenas por meio da relação matemática.

Para Morettin e Tolo (1986), série temporal é parte de uma trajetória dentre muitas que poderiam ter sido observadas, dentro de um conjunto de observações disponíveis para análise, enquanto processo estocástico pode ser pensado de duas maneiras: como um conjunto de possíveis trajetórias que podem ser observadas e como um conjunto de variáveis aleatórias, sendo uma para cada tempo  $t$ .

Souza (2006) define o processo como estocástico quando suas médias e variâncias forem constantes ao longo do tempo e o valor da covariância entre dois períodos de tempo depender apenas da distância ou defasagem entre dois períodos, e não do período de tempo efetivamente em que a covariância é calculada. Caso haja não-estacionariedade, é possível que se obtenha resultados espúrios como consequência. Uma família de variáveis aleatórias definidas no mesmo espaço de probabilidades pode ser considerada como um processo estocástico.

De acordo com Mantovani (2011), a obtenção de previsões por meio de séries temporais é realizada a partir da análise dos movimentos da variável de estudo durante determinado período de tempo, necessitando a identificação de componentes básicos existentes na série histórica, tais como, por exemplo, tendência de aumento ou queda, aleatoriedade, sazonalidade, a fim de estabelecer relações entre as observações e determinação de um modelo matemático a fim de prever o comportamento futuro.

A definição de série temporal para Nascimento *et al*, 1996; Levine, Berenson e Stephan, 2000 *apud* Nardi *et al*, 2006, considera como sendo um conjunto de dados numéricos obtidos durante períodos regulares ao longo do tempo e que a sua análise tem como objetivo formular modelos capazes de descrever a dependência do dado com relação ao tempo e desta forma, prever valores futuros. Uma série temporal é composta por: tendência, sazonalidade, ciclos e aleatoriedade.



As técnicas de utilização de séries temporais são baseadas em modelos quantitativos que executam uma lógica utilizando operações matemáticas a fim de gerar uma previsão estatística. Esta lógica permite avaliar os dados históricos para conhecer o comportamento do sistema, a partir do pressuposto de que o mesmo é estável, extrapolando o conhecimento adquirido para o futuro (STAUDT, 2011).

Já para Santos (2006), uma série temporal nada mais é que uma sequência ordenada no tempo de observações de uma determinada variável. O autor descreve que estas podem ser classificadas em contínuas, quando são medidas sem interrupções na variável tempo, ou discretas quando mensuradas em intervalos sucessivos de tempo, sendo em geral, durante períodos igualmente espaçados.

## **2.2 Previsão de demanda**

O extremo dinamismo e complexidade do mercado faz com que haja a necessidade de tomadas de decisões cada vez mais rápidas e eficientes. Dessa forma, torna-se importante a utilização de ferramentas que possibilitem garantir uma maior proatividade, identificação e assertividade quanto às necessidades do mercado. De maneira geral, em qualquer organização em que há uma estratégia de planejamento voltada para atender seus consumidores de maneira prontamente eficaz, prever a demanda por meio de técnicas de previsão impactam diretamente na melhoria das tomadas de decisão em todos os níveis de planejamento e, por consequência, garantem também a sobrevivência no mercado que se mostra cada vez mais competitivo.

De maneira geral, os métodos mais utilizados para a previsão de demanda são aqueles baseados em dados históricos, os quais necessitam de revisões contínuas com o intuito de garantir que as variáveis originais levadas em consideração no momento da previsão ainda continuam válidas se comparadas ao seu tempo de atuação.

Para Ballou (2006), a previsão de demanda é de suma importância para toda a empresa, à medida que proporciona a entrada básica para o planejamento e controle de todas as diversas áreas funcionais das organizações. Os volumes de demanda e os momentos em que ocorrerão afetam fundamentalmente os índices de capacidade, as necessidades financeiras e a estrutura geral de qualquer negócio.

Morettin e Toli (1986) classificam os procedimentos de previsão de séries temporais em duas categorias:

- Automáticos: aqueles aplicados diretamente, com a utilização de programas simples de computador;
- Não automáticos: aqueles que exigem intervenção de pessoas especializadas para serem aplicadas.

Alguns procedimentos automáticos que podemos citar são os modelos de alisamento exponencial (alisamento exponencial simples, linear de Brown, biparamétrico de Holt, sazonal de Holt-Winters, entre outros), autorregressão *stepwise*, entre outros. Já os modelos de previsão não automáticos incluem os métodos de Box-Jenkins e Bayesiano.

### **2.3 Métodos de previsão de demanda**

De acordo com Bacci (2007), para se determinar o modelo, ou os modelos de previsão de demanda que uma empresa deve adotar, utiliza-se de uma série de fatores, tais como o horizonte de previsão, a disponibilidade dos dados, a precisão desejável, o custo para formulação da previsão, a disponibilidade de pessoas qualificadas, a natureza dos produtos e serviços.

Para Laboissiere, Fernandes e Lage (2015), os modelos de previsão têm como objetivo determinar valores e tendências de uma série temporal, que normalmente são tratados por meio de modelos de regressão futura.

Ainda segundo com Bacci (2007), a etapa mais relevante e complexa para a construção do modelo de previsão é a definição da técnica apropriada para utilização, visto que, balancear as vantagens e desvantagens das técnicas com respeito às limitações e situações requeridas é considerada uma tarefa bastante difícil.

Para Lustosa *et al.* (2008), podem-se classificar os vários métodos de previsão em duas categorias macros, sendo estes qualitativos e quantitativos, conforme mostrado na Figura 2.1.

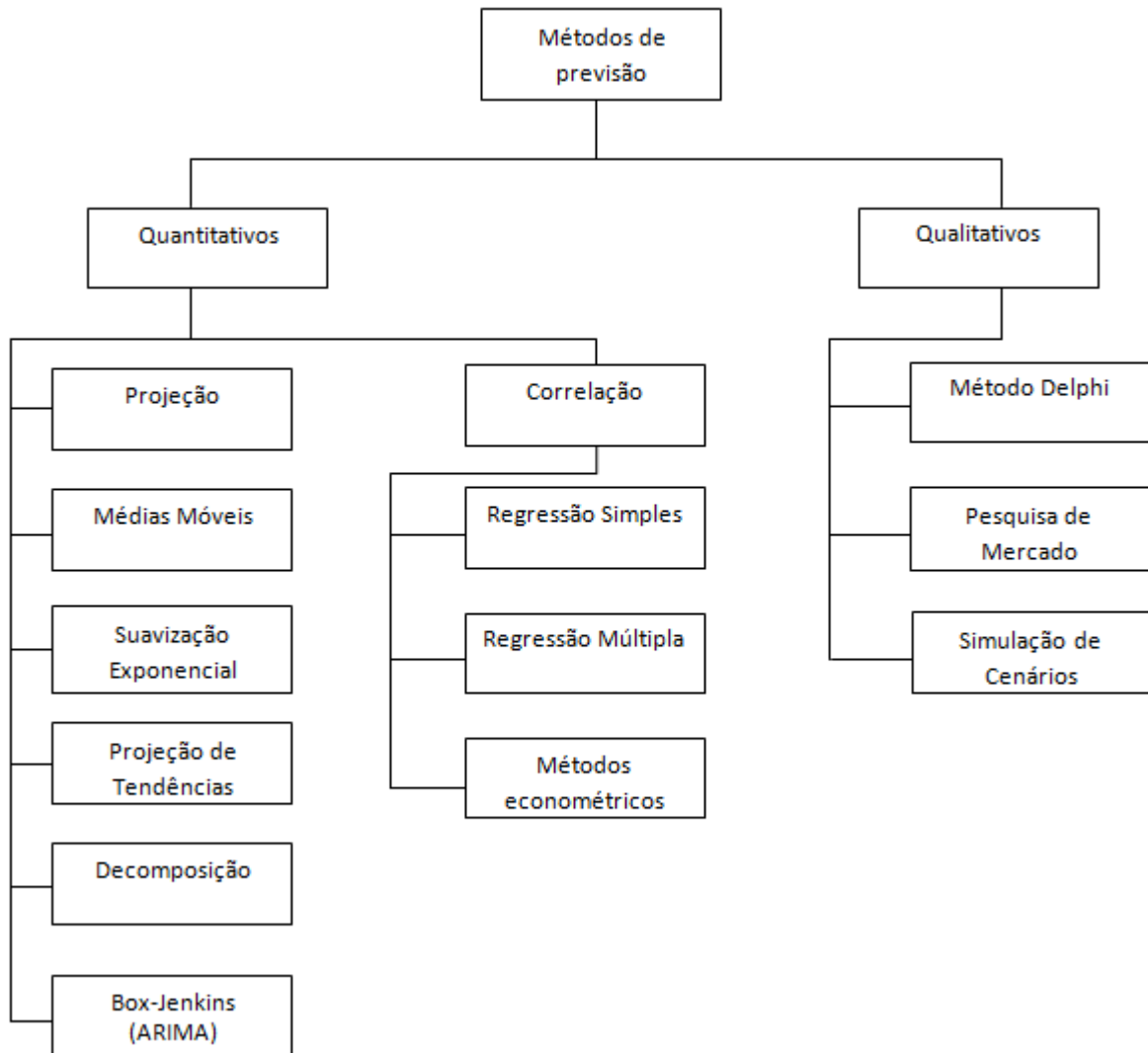


Figura 2.1 – Métodos de previsão

Adaptado de Lustosa *et al.* (2008)

Os métodos qualitativos se baseiam em previsões subjetivas ou optativas por natureza. Além disso, são baseados em estimativas e opiniões de pessoas experientes na área ou especialistas capacitados. Segundo Davis (2001), dentre essas técnicas destacam-se a Pesquisa de Mercado, Analogia Histórica e o Método Delphi. Lustosa *et al.* (2008) acrescenta como técnica qualitativa a simulação de cenários.

Quanto aos métodos quantitativos, estes buscam analisar os dados históricos de um ou mais produtos por meio de modelos matemáticos a fim de projetar uma demanda futura. Para Davis (2001), os métodos quantitativos podem ser divididos em dois grupos principais: técnicas baseadas na análise de séries temporais e causal, nos quais tentam compreender o sistema que envolve o item a ser previsto.

A utilização e análise das séries temporais baseiam-se na ideia de que os dados referentes à demanda do passado estão diretamente relacionados com a previsão da demanda futura, ou seja, existe uma tendência gerada pelos dados históricos que possivelmente continuará gerando a demanda no futuro. De forma técnica, equivale dizer que o processo estocástico gerador das observações passadas continuará gerando observações no futuro.

Quanto às técnicas utilizadas nos métodos quantitativos, destacam-se as médias móveis simples, a média ponderada exponencial, a Regressão Linear e Projeção de Tendência, entre outras, além da Metodologia Box-Jenkins, que é considerada a mais importante.

Já os métodos causais, segundo Davis (2001), procuram compensar as falhas da análise de séries temporais, buscando identificar os fatores que influenciam o comportamento da demanda. Alguns exemplos de métodos mais citados são a análise de regressão e os métodos econométricos. Na Tabela 2.1 é apresentado um comparativo entre os métodos de previsão de demanda.

Tabela 2.1: Exemplos de técnicas para previsão de demanda.

<b>Técnica</b>	<b>Horizonte de tempo</b>	<b>Complexidade do modelo</b>	<b>Precisão do modelo</b>	<b>Dados necessários</b>
<b>Qualitativo</b>				
Método Delphi	Longo	Alta	Variável	Muitos
<b>Quantitativos Séries temporais</b>				
Média Móvel	Curto	Muito Baixa	Média	Poucos
Média Ponderada Exponencial	Curto	Baixa	Adequada	Muito Pouco
Regressão Linear	Longo	Média Baixa	Média Alta	Muitos
<b>Causal</b>				
Análise de Regressão	Longo	Adequada	Alta	Muitos

Fonte: Makridakis, Wheelwright e Hyndman (1998)

Para Makridakis, Wheelwright e Hyndman (1998), a previsão de métodos quantitativos pode ser aplicada quando existem basicamente três condições:

- Informações do passado disponíveis;
- Estas informações possam ser quantificadas por meio de dados numéricos;

- Pode-se supor que alguns aspectos do padrão passado poderão continuar no futuro, colocação definida também como pressuposto da continuidade.

## **2.4 Implantação de sistemas de previsão**

Para Pellegrini e Fogliatto (2001), os problemas relacionados à previsão devem estar integrados a diversos fatores que necessitam ser analisados na etapa da implantação de um sistema de previsão, tais como:

- Como e onde a previsão será utilizada;
- Disponibilidade de dados;
- Acurácia;
- Custo da análise e preferências gerenciais;
- Período ou tempo requerido, horizonte (números de períodos futuros cobertos pela previsão), e intervalo da previsão (frequência com a qual novas previsões são preparadas).

Ainda conforme Pellegrini e Fogliatto (2001), o período consiste no fator que mais influenciará na escolha do modelo estatístico. A determinação do horizonte da previsão está diretamente associada à agilidade da organização, visto que, quanto menos ágil for ela, maior será o horizonte demandado e o intervalo dependerá da estabilidade do processo e custos de replanejamento.

Uma etapa fundamental para a implantação de sistemas de previsão é a seleção de dados, visto que são classificados como as entradas dos modelos de previsão e uma vez que estas sejam consideradas ruins, a probabilidade dos seus resultados serem ruins é diretamente proporcional.

Dentro da metodologia de previsão utilizando séries temporais, nem sempre há a necessidade de análise individual de todos os produtos no banco de dados ou produzidos numa determinada organização, visto que os mesmos podem apresentar inúmeros outros produtos que podem ser analisados de forma conjunta ou simplesmente agregada, tais como, por exemplo, aqueles que fazem parte de uma mesma família.

Uma das metodologias mais aplicadas para a agregação de produtos, segundo Pellegrini e Fogliatto (2001) é a classificação ABC dos produtos, na qual relaciona-se a

importância do produto de acordo com a demanda e faturamento. De forma simplificada, os produtos de classe A representam 80% do faturamento e 20% dos produtos vendidos, enquanto os da classe B representam 15% do faturamento e aproximadamente 30% dos produtos vendidos pela empresa e por último os itens referentes a classe C, os quais estão inseridos nos 5% de faturamento e cerca de 50% dos produtos vendidos e neste último caso, é mais indicada uma previsão agregada de demanda em função do baixo percentual de resultado financeiro da empresa.

Outra etapa importante para implantação do sistema de previsão é a análise dos dados. Nesta etapa, é importante que se faça a limpeza e ajuste nos dados trabalhados, a fim de se obter melhores resultados nas previsões.

Durante essa fase é interessante observar se os dados estão completos, se existem pontos sazonais, tendência e, além disso, definir qual o número de períodos dos dados serão utilizados para a preparação da previsão.

A definição de períodos de dados a serem analisados pode ser justificada pela pesquisa desenvolvida por Jain (2007 *apud* ZOUCCAS, 2009). Na mesma, o autor busca analisar 10 empresas que empregam técnicas e modelos estatísticos para previsão de demanda e a partir disto, realiza observações interessantes para o tema em questão, conforme pode-se verificar nas Figuras 2.2, Figura 2.3 e Figura 2.4.

A partir da análise das Figuras 2.2, 2.3 e 2.4, observou-se respectivamente, que o modelo estatístico mais empregado nas previsões é realizado a partir de séries de tempo, bem como o período de 31 a 36 meses seja considerado como o mais utilizado nas previsões, por meio de técnicas de média e tendência simples. O método de Box-Jenkins (ARIMA), um dos métodos de estudo desta pesquisa, representou 7% entre os métodos pesquisados por Jain (2007 *apud* ZOUCCAS, 2009), além de que o período destacado como mais utilizado por este autor, também é equivalente ao usado para o desenvolvimento desta pesquisa.

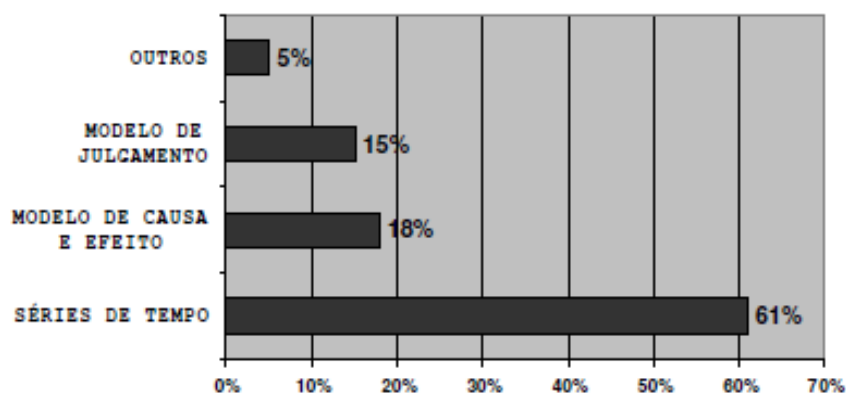


Figura 2.2 – Modelo estatístico usado nas previsões

Fonte: Jain (2007)

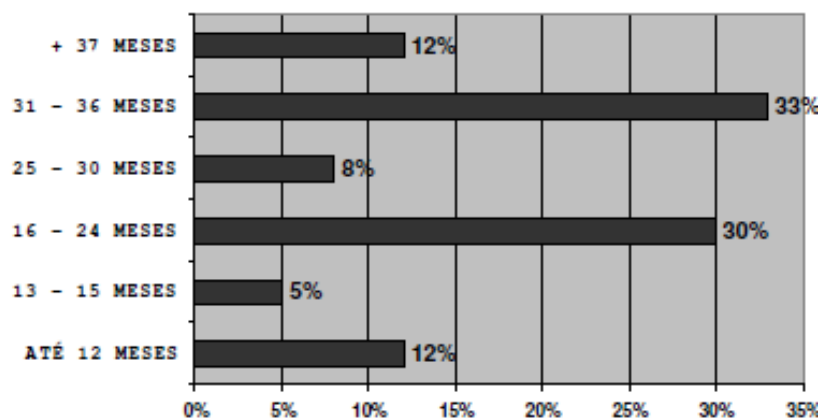


Figura 2.3 – Período dos dados utilizados na previsão

Fonte: Jain (2007)

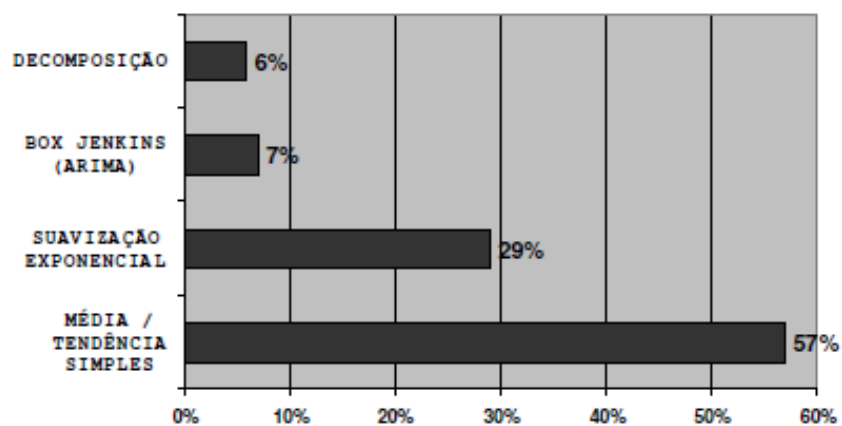


Figura 2.4 – Técnicas mais utilizadas na previsão

Fonte: Jain (2007)

A pesquisa desenvolvida por Jain (2007 *apud* ZOUCAS, 2009), serviu como referência para as etapas de implantação dos métodos de previsão de demanda, dando embasamento teórico para as etapas de definição do período de coleta de dados, modelos e técnicas interessantes a serem utilizadas.

## 2.5 Modelos de Box-Jenkins

De acordo com Pellegrini e Fogliatto (2001), os modelos de Box-Jenkins são conhecidos também como ARIMA (*Autoregressive Integrated Moving Average*) ou Modelos Autoregressivos Integrados a Média Móvel e foram propostos por George Box e Gwilym Jenkins no início dos anos 70.

O método permite obter resultados satisfatórios de previsão por meio da modelagem matemática do comportamento referente à correlação seriada ou autocorrelação entre os valores da série temporal.

Ainda de acordo com Pellegrini e Fogliatto (2001), os métodos de Box-Jenkins abrangem como idéia principal o fato que os valores de uma série temporal são bastante dependentes, ou seja, existe uma explicação para cada um dos valores prévios da série. Além disso, os modelos ARIMA estão entre a classe mais geral de modelos para a análise de séries temporais.

Cavalcanti-Netto (2007) afirma que os modelos ARIMA apresentam os melhores resultados quando se trata de previsões de curto prazo com informações mensais referentes a 05 ou 10 anos e que não apresentam sazonalidade. Além disso, é importante observar se o processo apresenta estacionariedade ou não, para que seja possível fixar os parâmetros dos modelos e utilizá-los nas previsões do futuro a partir do passado.

Flores (2009) afirma que a metodologia Box-Jenkins ainda se destaca quanto à escolha de outros modelos devido sua facilidade de uso, especialmente em relação a modelos não paramétricos, como por exemplo, os modelos de Redes Neurais Artificiais.

Para Souza (2006) os métodos de Box-Jenkins estão relacionados aos métodos de identificação, ajuste, checagem e uso de modelos autorregressivos integrados a média móvel, simplesmente conhecido como modelo ARIMA. São modelos matemáticos que buscam aprender o comportamento da correlação seriada ou autocorrelação entre os valores da série temporal, e, a partir desse comportamento, realizar as futuras previsões.



De acordo com Morretin & Toli (1986), os modelos de Box-Jenkins são construídos em um meio iterativo, conforme mostrado na Figura 2.5, em que a escolha do modelo é realizada com relação à base de dados apresentada. As etapas para construção do modelo, conforme Box & Jenkins (1976) são:

- Identificação: essa etapa consiste em encontrar qual o comportamento da série, dentre as várias versões dos modelos de Box-Jenkins, sejam eles sazonais ou não. A identificação do modelo é embasada por meio do comportamento das funções de autocorrelações (ACF) e autocorrelações parciais (PACF).
- Estimação: consiste em estimar os parâmetros  $\phi$  e  $\Phi$  do componente autorregressivo,  $\theta$  e  $\Theta$  do componente de médias móveis e variância de  $\varepsilon_T$  (ruído branco).
- Verificação: consiste em avaliar se o modelo estimado é apropriado para descrever o comportamento dos dados. Caso não seja, repete-se o ciclo, iniciando na fase de identificação. É bastante comum identificar mais de um dos modelos, os quais serão estimados e verificados. Quando se obtém um modelo satisfatório, passa-se para a última etapa da metodologia de Box-Jenkins, que é o objetivo principal: realizar previsões.

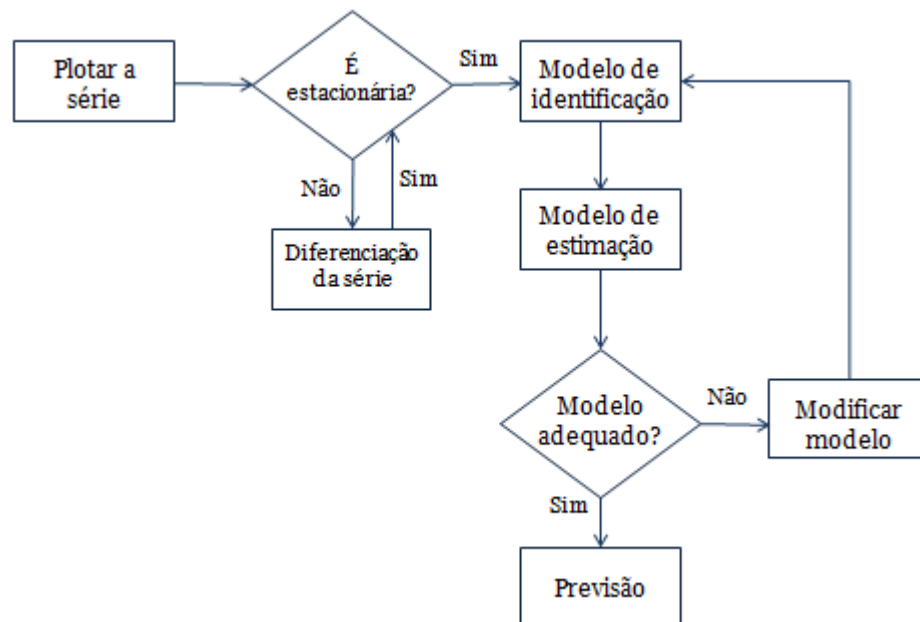


Figura 2.5 – Fluxograma do ciclo iterativo de Box & Jenkins

Fonte: Adaptado por Morretin & Toli (1986)

Durante a etapa anterior à realização das previsões, a etapa de verificação, avalia-se a significância estatística dos parâmetros e faz-se a análise dos resíduos do modelo. Nesta etapa, aplica-se o teste  $Q$  de Ljung-Box, o qual conforme Sartoris (2003) testa a hipótese nula da ausência de autocorrelação em uma série, e é aplicado aos resíduos de modelos selecionados inicialmente. É importante ressaltar que os resíduos de um modelo adequado não devem apresentar autocorrelação, apresentando um aspecto de ruído branco.

A vantagem de utilização do teste Ljung-Box é seu desempenho perante a utilização de amostras de dados pequenas ou moderadas. Dessa forma, utiliza-se o teste Ljung-Box, no qual sua estatística é descrita pela Equação 1:

$$Q(K) = n(n + 2) \sum_{k=1}^m \frac{r_k^2}{n-k} \quad (1)$$

Tal que:

$n$  refere-se ao número total de observações e as hipóteses são:

$$H_0 = r_1 = \dots = r_k = 0$$

$$H_1 = r_k \neq 0$$

$r_k$  = autocorrelações residuais,  $k = 1, \dots, m$ .

Ainda de acordo com Sartoris (2003), pode-se assumir independência estatística dos dados apenas quando  $Q$  é maior que o valor da distribuição  $\chi^2$ , ou seja, um  $p$ -value  $< \alpha$  indica rejeição da hipótese nula, assumindo na prática, a presença de autocorrelação nos dados. Ou seja, caso o  $Q$  calculado seja superior ao  $Q$  tabelado para determinado nível de significância, não se aceita a hipótese nula  $H_0$  de que os resíduos não são autocorrelacionados. No entanto, isto implica em aceitar a hipótese alternativa  $H_1$  de que pelo menos uma autocorrelação é estatisticamente diferente de zero, ou seja, não é ruído branco.

Durante a etapa de verificação, utilizou-se também o teste de Shapiro-Wilk, o qual segundo Shapiro e Wilk (1965), tem como finalidade verificar a aderência ao modelo Normal. Na pesquisa, o teste de Shapiro-Wilk foi utilizado para testar a normalidade dos resíduos.

A partir dos modelos autorregressivos (AR) e os de médias móveis (MA), vários outros foram criados. Entre eles, estão os modelos autorregressivos integrados de médias móveis (ARIMA), onde é necessário tomar diferenças de uma série não estacionária a fim de transformá-la em uma série estacionária. Além disso, tem-se os modelos autorregressivos integrados de médias móveis e componentes sazonais (SARIMA).

Para Cavalcanti-Netto (2007), a sistemática da metodologia de Box-Jenkins permite que o modelo ARIMA descreva tanto o comportamento estacionário como o não estacionário, trazendo uma característica de flexibilidade para a metodologia. Suas previsões são embasadas em valores correntes e passados da série. Mueller (1996) destaca a flexibilidade da metodologia de Box-Jenkins, citando que o modelo ARIMA apresenta capacidade de manipular séries temporais de qualquer natureza a partir da combinação dos valores reais passados ou dos erros ocorridos, características primordiais para a escolha de aplicação do estudo em questão.

Caldwell (2006) afirma que os modelos de Box-Jenkins podem ser usados para representar processos que são estacionários ou não-estacionários. Além disso, o autor define os processos estacionários como sendo aqueles em que as propriedades estatísticas são as mesmas ao longo do tempo, oscilando apenas em torno de um valor médio fixo. Já séries temporais não-estacionárias incluem características como mudanças no nível, tendências ou mudanças no comportamento sazonal. Werner e Ribeiro (2003) definem modelos estacionários como aqueles que assumem que o processo está em “equilíbrio”.

Pode-se afirmar, intuitivamente, que os modelos ARIMA representam as séries temporais como uma ponderação dos próprios valores e erros passados da série.

Conforme Fava (2000), a definição do modelo ARIMA resulta da combinação de três componentes, os quais são denominados como “filtros”:

- Componente autorregressivo (AR);
- Filtro de integração (I);
- Componente de médias móveis (MA).

Os modelos ARIMA ( $p, d, q$ ) oferecem parâmetros desconhecidos, os quais devem ser identificados e estimados a partir das características da série de dados. Para identificação de tais parâmetros de um modelo ARIMA é necessário que se determine:

- Ordem máxima dos termos autorregressivos, conhecido como ordem  $p$ ;
- Nível de diferenciação a partir do qual a série se torna estacionária, conhecido como ordem  $d$ ;
- Ordem máxima dos termos de médias móveis, conhecido como ordem  $q$ .

### 2.5.1 Função de autocorrelação (ACF) e autocorrelação parcial (PACF)

De acordo com Gujarati (2006), a função de autocorrelação parcial (PACF) e a função de autocorrelação (ACF) são importantes para a etapa de identificação da ordem  $q$  do processo MA e ordem  $p$  do processo AR, respectivamente. Ou seja, para que os modelos (AR, MA e ARIMA) sejam identificados e selecionados, é necessário definir duas funções primordiais, que são a função de autocorrelação (ACF) e a função de autocorrelação parcial (PACF).

A ACF é gerada a partir da correlação linear de cada valor  $y_T$ , da série de dados trabalhada com os demais valores em *lags* distintos, como por exemplo  $y_{T-1}$ ,  $y_{T-2}$ , e assim por diante. Já a autocorrelação parcial (PACF), apesar de apresentar conceito semelhante, não leva em consideração todos os valores em *lags* da série de dados, ou seja, faz-se uma correlação entre  $y_T$  e  $y_{T-3}$ , porém sem levar em consideração os efeitos causados por  $y_{T-1}$ ,  $y_{T-2}$ .

### 2.5.2 Componente autorregressivo (AR)

O modelo autorregressivo  $AR(p)$  apresenta uma série de dados formada por valores regredidos no tempo e pelo ruído branco ou aleatório, representado por  $\varepsilon_T$ . Dessa maneira, o modelo  $AR(p)$  utilizado em processos estacionários é expresso pela Equação 2:

$$y_T = \phi_1 y_{T-1} + \phi_2 y_{T-2} + \dots + \phi_p y_{T-p} + \varepsilon_T \quad (2)$$

Tal que:

$y_T$  refere-se ao valor correspondente à observação da série temporal no tempo  $T$ ;

$\phi_p$  refere-se ao parâmetro do modelo AR de ordem  $P$ ;

$\varepsilon_T$  correspondente ao valores que não podem ser explicados pelo modelo, ou seja, o resíduo.

Os modelos autorregressivos foram desenvolvidos com o intuito de mostrar que a presente observação da série  $y_T$  pode ser explicada como uma função das  $P$  observações anteriores,  $y_{T-1}$ ,  $y_{T-2}$ , ...,  $y_{T-P}$ , onde  $P$  determina o número de passos entre as observações anteriores e a previsão da próxima observação.

Também pode-se escrever o processo  $AR(p)$  usando o operador de defasagem  $B$ , também conhecido como operador *backshift*  $B$ , onde:

$$y_T = \phi_1 B y_T + \phi_2 B^2 y_T + \dots + \phi_p B^p y_T + \varepsilon_T \quad (3)$$

$$y_T - \phi_1 B y_T - \phi_2 B^2 y_T - \dots - \phi_p B^p y_T = \varepsilon_T \quad (4)$$

Apesar do operador de defasagem  $B$  não ser considerado uma variável, o mesmo pode ser tratado de maneira algébrica, colocando  $y_T$  em evidência, tem-se:

$$\underbrace{(1 - \phi_1 B^1 - \phi_2 B^2 - \dots - \phi_p B^p)}_{\phi_p(B)} y_T = \varepsilon_T \quad (5)$$

$$\phi_p(B) y_T = \varepsilon_T$$

### 2.5.3 Componente de médias móveis (MA)

Na componente de médias móveis (MA), os resultados da série  $y_T$  são definidos pela combinação linear dos resíduos ou ruídos brancos ( $\varepsilon$ ) do período corrente com os ocorridos nos períodos passados. A partir disso, tem-se a seguinte formulação:

$$y_T = \varepsilon_T - \Phi_1 \varepsilon_{T-1} - \Phi_2 \varepsilon_{T-2} - \dots - \Phi_q \varepsilon_{T-q} \quad (6)$$

Onde:

$y_T$  refere-se ao valor correspondente à observação da série temporal no tempo  $T$ ;

$\Phi_q$  refere-se ao parâmetro do modelo MA de ordem  $q$ ;

$\varepsilon_T$  representa os resíduos ou ruídos brancos, isto é, as diferenças entre os valores reais e os valores previstos da série em cada instante  $t$ ;

$q$  representa o número de defasagens da média móvel.

Da mesma maneira que fez-se para o modelo autorregressivo  $AR(p)$ , também pode-se tratar o modelo  $MA(q)$  utilizando o operador de defasagem  $B$  da seguinte forma:

$$y_T = \underbrace{(1 - \Phi_1 B^1 - \Phi_2 B^2 - \dots - \Phi_q B^q)}_{\Phi_q(B)} \varepsilon_T \quad (7)$$

$$y_T = \Phi_q(B) \varepsilon_T$$

### 2.5.4 Modelo autorregressivo de médias móveis (ARMA)

Este modelo é resultado da combinação dos processos  $AR(p)$  e  $MA(q)$ , ou seja, é descrito pelos seus  $p$  valores passados e  $q$  resíduos ou ruídos brancos correntes e passados das séries temporais estacionárias. Tal que:

$$y_T = \underbrace{\phi_1 y_{T-1} + \phi_2 y_{T-2} + \dots + \phi_p y_{T-p}}_{\text{Modelo autorregressivo}} + \underbrace{\varepsilon_T - \Phi_1 \varepsilon_{T-1} - \Phi_2 \varepsilon_{T-2} - \dots - \Phi_q \varepsilon_{T-q}}_{\text{Modelo de média móvel}} \quad (8)$$

Onde:

$\phi$  representa o parâmetro autorregressivo,  $i = 1, \dots, p$ ;

$\Phi$  representa o parâmetro de média móvel,  $i = 1, \dots, q$ ;

$\varepsilon_T$  ruído branco.

Utilizando o operador de defasagem  $B$ , tem-se

$$\phi_p(B) y_T = \Phi_q(B) \varepsilon_T \quad (9)$$

### 2.5.5 Modelo autorregressivo integrado de médias móveis (ARIMA)

Um dos modelos mais importantes e amplamente utilizados em séries temporais é o modelo autorregressivo integrado de médias móveis (ARIMA). A sua popularidade é devido às suas propriedades estatísticas, tais como, por exemplo, a metodologia de Box-Jenkins, utilizada no processo de criação de modelos (ZHANG, 2001).

Os modelos ARIMA resultam da combinação dos componentes autorregressivos (AR), integração (I) e médias móveis (MA). Esses modelos são utilizados para analisar séries temporais não estacionárias, nos quais, um modelo ARIMA  $(p, d, q)$  nada mais é que o resultado de uma transformação de um modelo ARMA  $(p, q)$ , onde substituiu-se  $\phi_p(B) y_T$  pela componente  $\phi_p(B)(1-B)^d y_T$  de acordo com Pacheco, (2001).

A representação de um modelo ARIMA  $(p, d, q)$  considera  $p$  como sendo o número de defasagens da série, enquanto  $d$  é a ordem de integração e  $q$  o número de defasagens dos erros aleatórios. As ordens de  $p$  e  $q$  são determinadas a partir das funções de autocorrelação e autocorrelação parcial, visto que tratam da combinação dos processos AR e MA.

No caso de  $y_T$  se tornar uma série estacionária, após passar por  $d$  diferenças, resultadas da série  $z_T$  do tipo ARMA ( $p, q$ ), diz-se que  $y_T$  é descrito por um modelo ARIMA ( $p, d, q$ ), tal que, segundo a Equação 10:

$$z_T = \phi_1 z_{T-1} + \phi_2 z_{T-2} + \dots + \phi_p z_{T-p} + \varepsilon_T - v_1 \varepsilon_{T-1} - v_2 \varepsilon_{T-2} - \dots - v_q \varepsilon_{T-q} \quad (10)$$

Um modelo ARIMA ( $p, d, q$ ) pode ser representado de forma mais resumida como apresentada na Equação 11.

$$\phi_p(B)(1-B)^d y_T = \Phi_q(B) \varepsilon_T \quad (11)$$

Dessa forma, é possível verificar que um modelo ARIMA ( $p, d, q$ ) nada mais é que um modelo ARMA ( $p, q$ ) que sofreu uma transformação, substituindo  $\phi_p(B) y_T$  pelo componente  $\phi_p(B)(1-B)^d y_T$ . Portanto, ambos apresentam as mesmas características, segundo Pacheco (2001).

### 2.5.6 Modelo autorregressivo integrado de médias móveis e componentes sazonais (SARIMA)

Algumas séries temporais apresentam uma componente sazonal importante. A componente sazonal de uma série  $l$  irá depender de valores passados, ou seja,  $y_{t-l}, y_{t-2l}, \dots$ . Este efeito sazonal pode ser interpretado como um fenômeno estocástico.

Dessa forma, denomina-se como SARIMA ( $p, d, q$ ) x ( $P, D, Q$ ), a combinação de componentes estocásticas regulares com sazonais e a mesma pode ser apresentada matematicamente como ilustrado na Equação 12.

$$\phi_p(B) (1-B)^d \theta_p(B_S)(1-B_S)^D y_T = \Phi_q(B) \Theta_Q(B_S) \varepsilon_t \quad (12)$$

A determinação de modelos do tipo SARIMA apresenta os mesmos procedimentos referentes à metodologia de Box-Jenkins utilizados para encontrar o modelo ARIMA não sazonal, ou seja, ambos levam em consideração as observações referentes ao comportamento da ACF e PACF. No entanto, neste caso, observa-se as defasagens sazonais.

## 2.6 Redes neurais

A utilização de técnicas voltadas para a previsão do futuro pode ser de fundamental interesse para diversas áreas, não apenas no nível estratégico, mas também no nível

operacional e de tomadas de decisão. A principal finalidade da previsão se baseia em gerar novos dados a partir de modelos e informações pré-existentes, sendo que estas podem ser classificadas e definidas de inúmeras maneiras, incluindo erros, falta de valores, ruídos e várias outras características que dificultam a criação de valores de previsão.

Segundo Braga, Carvalho e Ludermir (2000), os pioneiros em desenvolver um trabalho destinado à criação do modelo artificial de um neurônio biológico foram Warren McCulloch e Walter Pitts em 1943. Neste trabalho, os autores concentraram muito mais na descrição de um modelo artificial de um neurônio e na apresentação de suas capacidades computacionais do que em apresentar técnicas de aprendizado.

As redes neurais artificiais são modelos computacionais inspirados no funcionamento do sistema nervoso dos seres humanos. São capazes de adquirir e realizar a manutenção do conhecimento por meio das informações inseridas. Além disso, também permitem ser definidas por um conjunto de unidades de processamento, denominadas neurônios artificiais. Estes neurônios são interligados por interconexões, as quais são representadas por vetores/matrizes de pesos (SILVA; SPATTI; FLAUZINHO, 2010).

As redes neurais artificiais (RNA) são, como seu nome indica, redes computacionais que simulam, a grosso modo, a rede de células nervosas (neurônios) de um sistema nervoso central biológico, seja humano ou animal. Esta simulação é basicamente uma simulação célula por célula, isto é, neurônio por neurônio (GRAUPE, 2007).

De acordo com Haykin (2001), o cérebro humano é um computador (sistema de processamento de informação) altamente poderoso, complexo, não linear e paralelo. Ele é capaz de processar e organizar uma imensa quantidade de informações em um tempo ínfimo. Esta capacidade de organização acontece devido à presença de constituintes estruturais conhecidas como neurônios.

O funcionamento do cérebro humano apresenta características de desenvolver suas regras por meio de “experiências” adquiridas nas mais diversas situações do cotidiano e estas acabaram por influenciar os pesquisadores a buscarem simular o funcionamento do cérebro a projetar modelos a partir do seu ambiente por meio do processo de aprendizagem e, além disso, armazenar o conhecimento adquirido.

Conforme Silva *et al.* (2010), os neurônios biológicos são regidos por elementos processadores biológicos que operam em paralelo e tem como objetivo principal conduzir impulsos sob determinadas condições de operação. Os neurônios estão conectados juntamente



uns aos outros por meio de sinapses, formando uma enorme rede. Os prolongamentos que conduzem informações em uma célula são denominados dendritos, e o axônio é constituído por um único prolongamento, cuja missão é conduzir impulsos elétricos para fora da célula. A função das sinapses é transmitir estímulos por todo o corpo humano, conforme mostra a Figura 2.6.

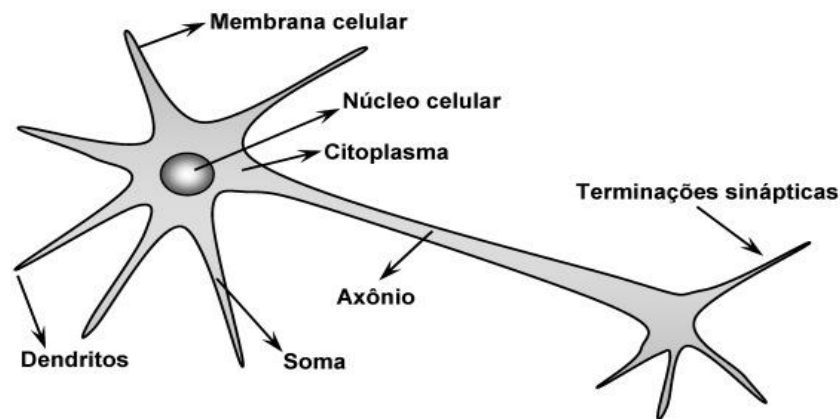


Figura 2.6 – Modelo de rede neural biológica

Fonte: Silva *et al.* (2010)

As Redes Neurais, ou redes neurais artificiais, apresentam uma tecnologia embasada em muitas disciplinas, tais como, por exemplo, a neurociência, a matemática, a estatística, a física, a ciência da computação e a engenharia. Além disso, as redes neurais encontram diversas áreas de aplicação, tais como modelagem, análise de séries temporais, reconhecimento de padrões, processamento de sinais e controle, em virtude de uma propriedade importante, a habilidade de aprender a partir de dados de entrada com ou sem supervisão (Haykin, 2001).

Algumas técnicas de programação são utilizadas para resolução de problemas e trazem soluções com resultados bastante satisfatórios, sejam estes no produto final ou no desenvolvimento da tarefa. Dentre estas técnicas, as RNAs se destacam devido às características como capacidade de aprendizado e generalização. Além disso, são utilizadas para o reconhecimento de padrões, processamento de sinais, digitais e sinais, otimização, visão computacional, entre outras, e previsão de séries temporais, foco dessa pesquisa.

De acordo com Silva *et al.* (2010), as redes neurais apresentam aplicações bastante relevantes, como por exemplo: facilidade para adaptação por experiência; por meio do ajuste dos parâmetros apresentados sucessivamente, capacidade de aprendizado; através do relacionamento de todas as variáveis compostas na aplicação, habilidade de generalização;

por meio da estimação de soluções que até o momento eram desconhecidas, organização dos dados; pelo agrupamento dos dados internos seguindo um padrão naqueles que apresentam particularidades em comum; tolerância a falhas, entre outras características.

Carbonneau, Laframboise e Vahidov (2008) concluíram em seu estudo que as redes neurais apresentam resultados melhores que um modelo de regressão linear, mas sua aplicação deve ser avaliada em função da complexidade computacional a qual a RNA apresenta. Além disso, ele cita que as RNAs são incluídas em estudos de previsão de demanda em função dos fabricantes considerarem a demanda como uma série caótica.

Para Russell e Norvig (2004) uma Rede Neural Artificial (RNA) é um modelo computacional que compartilha propriedades semelhantes as que são encontradas no cérebro humano, o qual consiste de várias unidades simples trabalhando em paralelo sem um controle central e as conexões entre unidades possuem pesos numéricos que podem ser modificados pelo elemento de aprendizado.

Para Haykin (2001), as redes neurais apresentam inúmeras arquiteturas e possuem diversas propriedades e capacidades úteis, tais como, por exemplo:

- Não-linearidade: propriedade muito importante das redes neurais, no qual a mesma apresenta uma conexão de neurônios não-lineares por toda a rede;
- Mapeamento de entrada-saída: neste caso, o aprendizado é supervisionado e os pesos sinápticos da rede são modificados pela aplicação de exemplos de tarefa. Essas modificações tem o intuito de minimizar a diferença entre a resposta desejada e a resposta real da rede, produzida pelo sinal de entrada, conforme critério estatístico apropriado. Dessa forma, o treinamento da rede é repetido por várias vezes até que a mesma alcance um estado estável onde não existam mais modificações significativas nos pesos sinápticos. Assim, a rede aprende pelo exemplos ao construir um mapeamento de entrada-saída para o problema em questão.
- Adaptabilidade: as redes neurais tem a capacidade de adaptação dos seus pesos sinápticos conforme haja modificações no meio ambiente, ou seja, uma rede treinada em um ambiente específico pode ser facilmente “retreinada” para lidar com pequenas mudanças nas condições operativas do ambiente, modificando em tempo real seus pesos sinápticos.

- Resposta a evidências: uma rede neural pode ser desenvolvida com o objetivo de fornecer informações não somente para quais padrões deverão ser selecionados, mas também sobre a confiança ou crença na tomada de decisão.
- Informação contextual: cada um dos neurônios da rede são afetados pelas atividades de todos os outros. Dessa maneira, a informação contextual é tratada naturalmente pela rede neural.
- Tolerância às falhas: no sentido de que o desempenho da rede se degrada suavemente sob condições de operação adversas, uma rede neural implementada na forma física (em *hardware*), tem o potencial de ser inerentemente tolerante às falhas ou capaz de realizar computação robusta.
- Implementação em *Very-Large-Scale-Integration* (VLSI): a característica de a rede neural apresentar a natureza maciçamente paralela faz com que seja potencialmente rápida na execução de certas tarefas da computação e isso possibilita que a rede se torne adequada para a implementação da tecnologia de integração em escala muito ampla, a qual fornece meio de capturar comportamentos realmente complexos de uma forma altamente hierárquica.
- Uniformidade de Análise e Projeto: a propriedade da rede neural artificial que permite construir redes modulares através de uma integração homogênea de módulos, isto quer dizer que é possível criar redes específicas e conectá-las, criando assim uma rede complexa.
- Analogia neurobiológica: as redes neurais são inspiradas no cérebro humano.

O modelo mais simplificado de um elemento computacional ou unidade processadora, denominada neurônio artificial, proposto por Warren McCulloch e Walter Pitts em 1943, pode ser visto conforme a Figura 2.7.

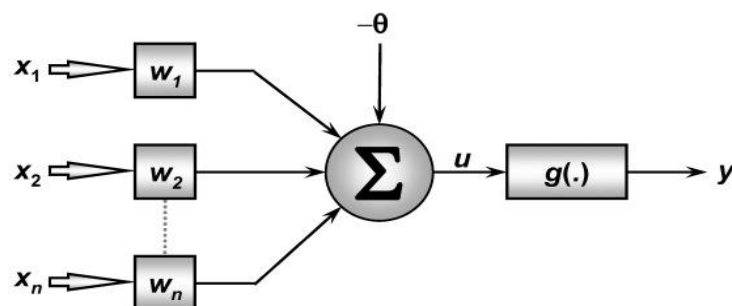


Figura 2.7 – Modelo computacional de rede neural

Fonte: Silva *et al.* (2010)

Realizando uma analogia ao modelo biológico, cada neurônio pode ser representado conforme a Figura 2.7 e as inúmeras entradas advindas da aplicação são diretamente relacionadas pelo conjunto  $\{x_1, x_2, \dots, x_n\}$ , os quais são equivalentes aos impulsos elétricos externos captados pelos dendritos no neurônio biológico, conforme Silva *et al.* (2010).

O terminal de saída  $y$  representa o axônio no neurônio biológico e para representar o comportamento das sinapses, têm-se os terminais de entrada com pesos acoplados  $\{w_1, w_2, \dots, w_n\}$  cujos valores podem ser positivos ou negativos dependendo das correspondentes sinapses (BRAGA; CARVALHO; LUDERMIR, 2000).

Para Silva *et al.* (2010), observando a Figura 2.7, verifica-se que o neurônio artificial proposto por McCulloch e Pitts é constituído por sete elementos básicos:

- Sinais de entrada  $\{x_1, x_2, \dots, x_n\}$ : são os sinais advindos do meio externo e que representam os valores assumidos pelas variáveis de uma aplicação específica. Geralmente estes sinais são normalizados com o intuito de incrementar a eficiência computacional dos algoritmos de aprendizagem;
- Pesos sinápticos  $\{w_1, w_2, \dots, w_n\}$ : definidos como os valores que irão ponderar cada uma das variáveis de entrada da rede, permitindo assim, quantificar a sua relevância em relação a funcionalidade do respectivo neurônio;
- Combinador linear  $\{\Sigma\}$ : agrega todos os sinais de entrada que foram ponderados pelos respectivos pesos sinápticos, com o objetivo de produzir um valor de potencial de ativação;
- Limiar de ativação  $\{\theta\}$ : variável que especifica qual o patamar será apropriado para que o resultado produzido pelo combinador linear possa gerar um valor de disparo em direção à saída do neurônio;
- Potencial de ativação  $\{u\}$ : resultado produzido pela diferença do valor produzido entre o combinador e o limiar da ativação. Caso o valor seja positivo, ou seja,  $u \geq \theta$ , o neurônio produzirá um potencial excitatório, caso contrário, será inibitório;
- Função de ativação  $\{g\}$ : tem como objetivo limitar a saída do neurônio dentro de um intervalo de valores razoáveis a serem assumidos pela sua própria imagem funcional;

- Sinal de saída  $\{y\}$ : valor produzido pelo neurônio em relação a um determinado conjunto de sinais de entrada, podendo ser também utilizado por outros neurônios que estão sequencialmente interligados.

Ainda de acordo com Silva *et al.* (2010), as Equações 13 e 14 seguintes sintetizam todo o resultado desenvolvido pelo neurônio artificial proposto por McCulloch e Pitts:

$$u = \sum_{i=1}^n w_i \cdot x_i - \theta \quad (13)$$

$$y = g(u) \quad (14)$$

A habilidade de um ser humano em realizar funções complexas e principalmente a sua capacidade de aprender advêm do processamento paralelo e distribuído da rede de neurônios do cérebro. (Haykin, 2001).

### 2.6.1 Principais arquiteturas de redes neurais artificiais

Para Silva *et al.* (2010), uma rede neural, basicamente é composta por três partes denominadas de camadas. Estas camadas são:

- Camada de entrada: responsável pelo recebimento de dados advindos do meio externo;
- Camadas escondidas, intermediárias ou ocultas: camadas situadas entre a camada de entrada e a camada de saída da rede neural, compostas por neurônios responsáveis por extrair as características associadas ao processo a ser inferido. Nestas camadas quase todo o processo interno da rede é realizado;
- Camada de saída: camada responsável pela produção e apresentação dos resultados finais da rede, advindos dos processamentos realizados pelos neurônios nas camadas anteriores.

A disposição das redes neurais, de acordo com Beale e Jackson (1990), pode ser por meio de uma única camada, sendo esta uma configuração mais simples de uma rede neural, ou então por meio de múltiplas camadas, na qual a rede é especificada principalmente pela sua topologia, pelas características dos nós e pelas regras de treinamento.

Silva *et al.* (2010) descreve as principais arquiteturas das redes neurais artificiais, levando em consideração a disposição dos neurônios, além das suas formas de interligação entre eles e a constituição de suas camadas. As redes podem ser divididas em redes

*feedforward* (alimentação à frente) de camada simples, redes *feedforward* de camadas múltiplas, redes recorrentes e redes reticuladas.

Tafner (1996) conclui que as conexões entre as camadas podem gerar  $n$  números de estruturas diferentes. Quando uma rede possui todas as saídas dos neurônios de uma camada conectadas com todos os neurônios da próxima camada denomina-se como *fullyconnected* (amplamente conectada), conforme mostrado na Figura 2.8.

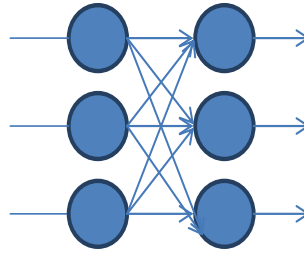


Figura 2.8 – Modelo de rede *feedforward* simples.

Fonte: Adaptado Tafner, 1996

Para Silva *et al.* (2010), a arquitetura da rede neural representada na Figura 2.8, é denominada por rede *feedforward* de camada única, já que o fluxo de informações desta rede segue sempre em uma mesma direção, sendo da camada de entrada em direção à camada de saída. Nesta arquitetura, a quantidade de saídas da rede sempre coincidirá com o número de neurônios. Os principais tipos de redes que contém a arquitetura *feedforward* de camada simples são: *Perceptron* e *Adaline*.

Para Haykin (2001), a segunda classe de rede neural, classificada por redes alimentadas diretamente com múltiplas camadas ou rede *feedforward* de camadas múltiplas se distingue por apresentar uma ou mais camadas ocultas, cujos nós computacionais são chamados de neurônios ocultos ou unidades ocultas. Sua única diferença é não ter contato com o mundo externo, intervindo entre a camada externa e a saída da rede. A Figura 2.9 é um exemplo de rede *feedforward* de múltiplas camadas.

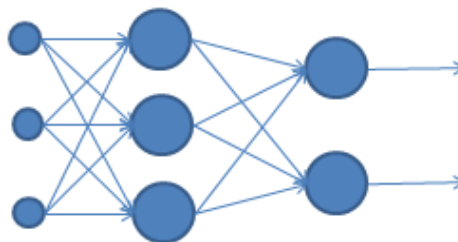


Figura 2.9 – Modelo de rede *feedforward* de camadas múltiplas.

Fonte: Adaptado Tafner, 1996

Para Figueredo (2008), a escolha dos parâmetros de treinamento de uma rede neural é uma tarefa crítica, onde características tais como, a escolha do tamanho da rede neural artificial, a quantidade de camadas escondidas e de neurônios e taxa de aprendizado exercem enorme influência na definição das características que se deseja obter, como, por exemplo, grau de precisão das respostas e tempo de convergência do treinamento. Dessa forma, justifica-se a utilização, em grande parte dos trabalhos, de métodos empíricos para definição destes parâmetros.

### **2.6.2 Processo de aprendizagem**

Haykin (2001) considera uma rede neural como um processador paralelamente distribuído constituído de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para uso. A rede neural se assemelha ao cérebro humano em dois aspectos:

- O conhecimento é adquirido pela rede a partir de seu ambiente através de um processo de aprendizagem;
- As forças de conexão entre neurônios, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido.

Ainda de acordo com Haykin (2001), o procedimento utilizado para realizar o processo de aprendizagem é chamado de algoritmo de aprendizagem, cuja função é modificar os pesos sinápticos da rede de uma forma ordenada para alcançar um objetivo de projeto desejado.

De acordo com Figueredo (2008), as redes neurais artificiais podem alterar seu comportamento, respondendo a estímulos produzidos pelo ambiente, por meio da regulação da força de conexão entre unidades de processamento adjacentes pela adaptação dos pesos sinápticos, reconhecendo as informações apresentadas às suas unidades visíveis.

Conforme Braga, Carvalho e Ludermir (2000), o aprendizado de uma rede neural pode ser classificado em dois tipos: aprendizado supervisionado e aprendizado não supervisionado.

Para que o aprendizado de uma rede neural seja considerado como aprendizado supervisionado, é necessário que suas entradas e saídas desejadas sejam fornecidas por uma espécie de supervisor externo, para que, dessa maneira, os parâmetros da rede sejam ajustados e haja uma ligação entre os pares de entrada e saída fornecidos.

Dessa forma, na ausência de um supervisor, a rede apresentará como desvantagem não conseguir aprender novas situações que não encaixam nas estratégias cobertas pelos exemplos do treinamento da rede.

Conforme Sousa (2012), a definição da taxa de aprendizado de uma rede neural tem enorme influência durante o processo de treinamento da rede.

## 2.7 Rede Neural Multi-Layer Perceptron (MLP)

As redes neurais *MultiLayer Perceptron* (MLP) ou simplesmente Redes Neurais de Múltiplas Camadas são exemplos de redes que utilizam camadas não lineares. Quanto à sua estrutura, as redes MLP são compostas por camadas sucessivas, com a primeira delas denominada camada de entrada ou *input layer*, uma ou mais camadas intermediárias ou *hiddenlayers* e a camada de saída, também denominada como *output layer* (HAYKIN, 2001).

A Figura 2.10 mostra a arquitetura de uma rede neural de múltiplas camadas com duas camadas ocultas e uma camada de saída. Conforme é possível observar na Figura 2.10, o exemplo de rede apresentada está totalmente conectada, ou seja, um neurônio em qualquer local da rede está conectado com todos os demais nós (neurônios) da camada anterior e o fluxo de sinal da rede segue para frente, da esquerda para a direita e de camada em camada, conforme Haykin (2001).

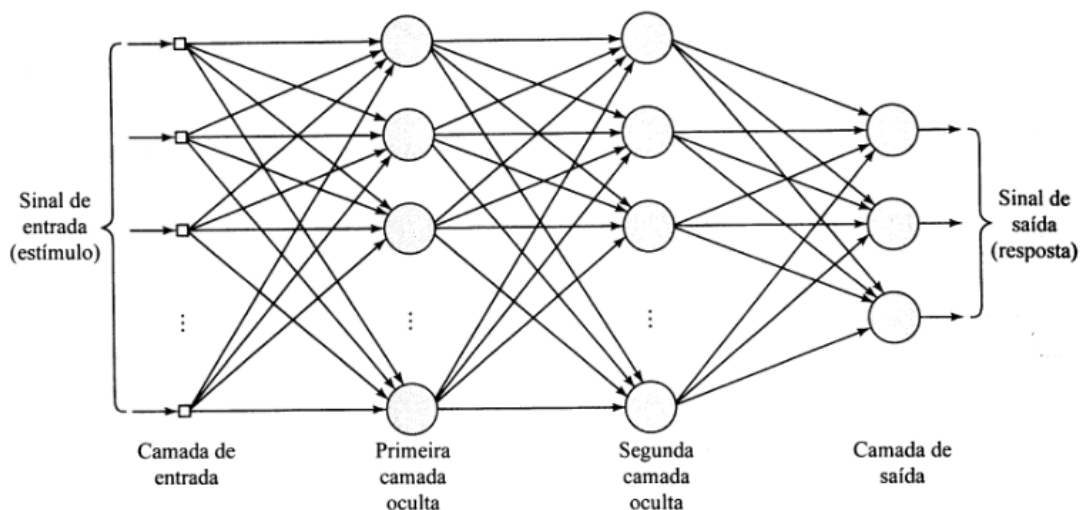


Figura 2.10 – Arquitetura rede neural de múltiplas camadas

Fonte: Haykin (2001)

De acordo com Szoplik (2015), as Redes *MultiLayer Perceptrons* (MLP) são as mais populares e amplamente utilizadas em aplicações práticas. A rede MLP consiste em uma



camada de entrada e uma camada de saída, porém, diferentemente de outras arquiteturas de rede, o número de camadas escondidas pode variar, conforme pode ser visto na Figura 2.10.

O modelo de RNA utilizado nessa pesquisa é o modelo de *perceptrons* de múltiplas camadas, MLP, definido como uma evolução do *perceptron* desenvolvido por Rosenblatt. O *perceptron* é construído em torno de um neurônio não-linear e tem sido aplicado com bastante sucesso a fim de solucionar problemas com grau de dificuldade elevado, por meio do seu treinamento supervisionado, utilizando o algoritmo de retropropagação do erro. Este algoritmo tem por finalidade realizar a aprendizagem por retropropagação do erro.

Para Haykin (2001), a aprendizagem por retropropagação do erro se baseia em dois passos por meio das diferentes camadas da rede neural:

- Propagação ou um passo a frente: durante esta etapa, os pesos sinápticos da rede são todos fixos;
- Retropropagação ou um passo para trás: durante esta etapa, os pesos sinápticos são todos ajustados de acordo com uma regra de correção do erro.

A resposta real da rede é subtraída de uma resposta desejada (alvo) para produzir o sinal do erro, que será propagado para trás através da rede, ou seja, contra as direções sinápticas, o que caracteriza o nome *error back-propagation*.

Para Bocanegra (2002), a solução de diversos problemas de engenharia por meio das RNA's é bastante interessante, tanto pela maneira pela qual a mesma é representada internamente, quanto pelos resultados que podem ser alcançados durante sua utilização. As RNA's podem chegar a desempenhos superiores aos modelos tradicionais, tais como, por exemplo, as análises de regressão.

Ainda conforme Haykin (2001), umas das características de um *perceptron* de múltiplas camadas é a função de ativação não-linear, na qual normalmente utiliza-se a não-linearidade sigmóide, que é definida pela função logística apresentada na Equação 15.

$$y_j = \frac{1}{(1 + \exp(-v_j))} \quad (15)$$

Além disso, apresenta uma ou mais camadas de neurônios ocultos, que permitem que a rede possa realizar tarefas complexas por meio da extração progressiva das características padrões dos vetores de entrada. Outra característica importante de um *perceptron* de múltiplas camadas é o alto grau de conectividade, definida pelas conexões sinápticas da rede.

De acordo com Silva e Oliveira (2002), não se recomenda a utilização de um grande número de camadas escondidas, visto que testes empíricos com a rede neural MLP *backpropagation* não apresentam vantagem significativa no uso de duas camadas escondidas ao invés de uma para solucionar problemas menores. Dessa forma, para a grande maioria dos problemas, utiliza-se apenas uma camada escondida.

Normalmente, o número de neurônios da camada intermediária é definido de forma empírica, dependendo da distribuição dos padrões de treinamento e validação da rede. A utilização de vários neurônios poderá acarretar na memorização do conjunto de treinamento, ao invés de extrair todas as suas características. Este processo é denominado como *overfitting*. Em contrapartida, quando se utiliza poucos neurônios, além da rede demorar mais tempo durante o processo de treinamento, também afeta os resultados ótimos do problema em questão.

Para Branco (2008), as redes neurais caracterizadas como *perceptron* de múltiplas camadas (MLP) têm sido aplicadas com sucesso na resolução de inúmeros problemas, inclusive quanto à previsão de demanda, por meio do algoritmo *backpropagation*.

Para a aplicação do algoritmo de retropropagação, um dos mais utilizados para o treinamento de redes neurais artificiais, proposto por Haykin (2001), deve ser considerada a etapa de inicialização dos dados, onde assume-se que nenhuma informação prévia esteja disponível. Nesta etapa, retiram-se os pesos sinápticos e limiares de uma distribuição uniforme cuja média é zero e cuja variância é escolhida tal que o desvio padrão dos campos locais induzidos dos neurônios se encontre na transição entre a parte linear e saturada da função de ativação sigmóide.

Além disso, é importante realizar a apresentação dos exemplos de treinamento, etapa em que apresenta-se na programação uma época de exemplos de treinamento à rede e para cada exemplo do conjunto, ordena-se de forma a realizar a sequência de computações para frente e para trás.

A próxima etapa consiste na computação para frente (propagação), onde realiza-se o cálculo dos campos locais induzidos e os sinais funcionais da rede, prosseguindo para frente através da rede, camada por camada. O campo local induzido  $v_j^l(n)$  para o neurônio  $j$  na camada  $l$  é definido conforme Equação 16.

$$v_j^l(n) = \sum_{i=0}^{m_o} w_{ji}^{(l)}(n) y_i^{(l-1)}(n) \quad (16)$$

Onde  $y_i^{(l-1)}(n)$  é o sinal (função) de saída do neurônio  $i$  na camada anterior  $l - 1$ , na iteração  $n$ , e  $w_{ji}^{(l)}(n)$  é o peso sináptico do neurônio  $j$  da camada  $l$ , que é alimentado pelo neurônio  $i$  da camada  $l - 1$ . Para  $i = 0$ , tem-se  $y_0^{(l-1)}(n) = +1$  e  $w_{j0}^{(l)}(n)$  é o *bias* aplicado ao neurônio  $j$  na camada  $l$ . Assumindo-se o uso de uma função sigmóide, o sinal de saída do neurônio  $j$  na camada  $l$  é dado pela Equação 17.

$$y_j^l = \varphi_j(v_j(n)) \quad (17)$$

Caso o neurônio  $j$  esteja na primeira camada oculta (i.e.,  $l = 1$ ),  $y_j$  é dado pela Equação 18.

$$y_j^{(0)}(n) = x_j(n) \quad (18)$$

Onde  $x_j(n)$  é o  $j$ -ésimo elemento do vetor de entrada  $x(n)$ . Se o neurônio  $j$  está na camada de saída (i.e.,  $l = L$  é denominado profundidade da rede),  $y_j$  é dado conforme Equação 19.

$$y_j^{(L)} = o_j(n) \quad (19)$$

Calcule-se o sinal de erro, de acordo com a Equação 20.

$$e_j(n) = d_j(n) - o_j(n) \quad (20)$$

Onde  $d_j(n)$  é  $j$ -ésimo elemento do vetor resposta desejada  $d(n)$ .

Aplica-se também a etapa de computação para trás (retropropagação), onde realiza-se o cálculo dos  $\delta$ s (i.e., gradientes locais) da rede, definidos conforme a Equação 21.

$$\delta_j^{(l)}(n) = \begin{cases} e_j^{(L)}(n) \varphi_j'(v_j^{(L)}(n)) & \text{para o neurônio } j \text{ da camada de saída } L \\ \varphi_j'(v_j^{(l)}(n))' \sum_k \delta_k^{(l+1)}(n) w_{kj}^{(l+1)}(n) & \text{para o neurônio } j \text{ na camada oculta } l \end{cases} \quad (21)$$

Onde o apóstrofe em  $\varphi_j'(\cdot)$  representa a diferenciação em relação ao argumento. O ajuste dos pesos sinápticos da rede na camada  $l$  é feito de acordo com a regra delta generalizada mostrado na Equação 22.

$$w_{ji}^{(l)}(n+1) = w_{ji}^{(l)}(n) + \alpha [w_{ji}^{(l)}(n-1)] + \eta \delta_j^{(l)}(n) y_i^{(l-1)}(n) \quad (22)$$

Onde  $\eta$  é o parâmetro da taxa de aprendizagem e  $\alpha$  é a constante de momento.

Na última etapa, iteração, busca-se realizar a iteração das computações para frente e para trás conforme citado nas etapas propagação e retropropagação citadas anteriormente, apresentando novas épocas de exemplos de treinamento para a rede, de maneira com que apresente um resultado satisfatório mediante ao critério de parada estabelecido.

## 2.8 Transformação de Yeo-Johnson

Uma estratégia geralmente utilizada para normalização de dados considerados não normais é buscar realizar a transformação da série em estudo. No entanto, determinar a transformação ideal para as características dos dados não é uma tarefa simples e que produz resultados conforme o esperado.

De acordo com Yeo & Johnson (2000 *apud* MORAES, FERREIRA e BALESTRASSI, 2006), apesar da transformação de Box-Cox representar um enorme avanço no quesito de determinar objetivamente a transformação de um conjunto de dados, a mesma somente se faz válida para valores positivos de  $x$ . Dessa forma, Yeo & Johnson sugeriram uma melhoria em que buscou-se a criação de uma nova família de transformação de dados, válida tanto para valores positivos quanto para valores negativos da variável  $x$ , conforme mostrado na Equação 23.

$$\psi^{(\lambda, x)} = \begin{cases} \{(x+1)^\lambda - 1\}/\lambda & (x \geq 0, \lambda \neq 0) \\ \log(x+1) & (x \geq 0, \lambda = 0) \\ -\{(-x+1)^{2-\lambda} - 1\}/2 - \lambda & (x < 0, \lambda \neq 0) \\ -\log(-x+1) & (x < 0, \lambda = 0) \end{cases} \quad (23)$$

## 2.9 Erros de previsão

Para Lopes (2002), a precisão de um modelo de previsão pode ser medida por meio de cálculos de seus erros. Essas medições são necessárias para verificação da acurácia dos valores preditos, identificação, isolamento, possíveis correções de variações anormais e até mesmo para se determinar técnicas ou parâmetros mais eficientes.

De acordo com Mine (2010), define-se um erro de previsão a partir da diferença entre o valor real e o valor previsto durante um mesmo período, ou seja, as medidas de erros são

relacionadas com a diferença entre os valores observados e valores previstos, conforme observado na Equação 24:

$$e_t = Y_t - F_t \quad (24)$$

Onde:

$e_t$ : representa o erro da previsão;

$Y_t$ : refere-se à observação atual para o período  $t$ ;

$F_t$ : correspondente a previsão para o período  $t$ .

Para a pesquisa em questão utilizou-se como método comparativo os valores do MAPE (Média Absoluta Percentual dos Erros) e RMSE (Raiz do Erro Quadrático Médio). A primeira medida é dada pela Equação 25:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{e_t}{y_t} \times 100 \right| \quad (25)$$

A segunda é definida pela raiz quadrada do MSE, o qual é estimado a partir da média do quadrado do desvio entre a previsão processada e a demanda real. Dessa forma, tem-se a medida RMSE conforme mostrado na Equação 26:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n e_t^2} \quad (26)$$

### 3. METODOLOGIA

#### 3.1 Caracterização da pesquisa

A pesquisa apresenta um conjunto de experimentos metodológicos, com embasamento científico que tem por objetivo obter resultados ao problema proposto.

De acordo com Turrioni e Mello (2012), os processos de classificação de pesquisas científicas podem ser considerados bastante controversos, pelo fato de depender bastante da maneira com que o autor irá focar a sua pesquisa. No entanto, uma forma clássica de se classificar as pesquisas científicas pode ser vista conforme a Figura 3.1.

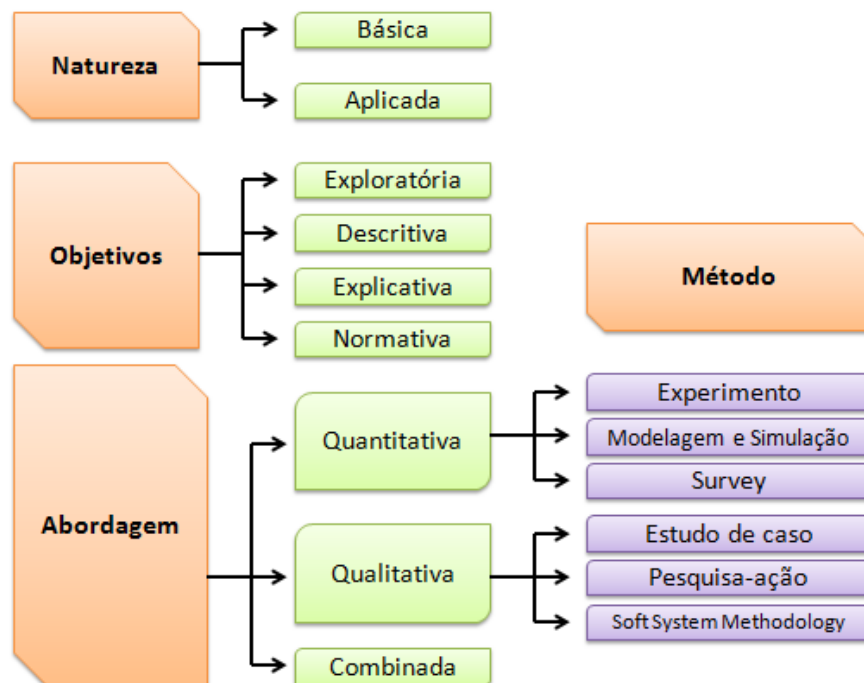


Figura 3.1 – Classificação da pesquisa científica em Engenharia de Produção

Fonte: Turrioni e Mello (2012)

Partindo do pressuposto de que a aplicação dos métodos de previsão de demanda de vendas na empresa resultará em tomadas de decisões assertivas para a organização, é importante que observe alguns passos para esta elaboração. Para Pellegrini e Fogliatto (2001), a elaboração de um sistema de previsão requer conhecimento e habilidade de uma organização em quatro áreas básicas:

- Identificação e definição dos problemas a serem tratados na previsão;
- Aplicação dos métodos de previsão;

- Procedimentos para seleção dos métodos apropriados a situações específicas;
- Suporte organizacional para adaptar e usar os métodos de previsão requeridos

O método científico pode ser definido como a linha de raciocínio seguida durante o processo de pesquisa, e este trabalho pode ser classificado como do tipo modelagem e simulação, com ênfase especial na metodologia em pesquisa operacional no contexto de engenharia de produção. Esta é a abordagem científica utilizada para auxiliar no processo de tomada de decisões, visto que se deseja experimentar, por meio de um modelo, um sistema real, determinando-se como este sistema responderá a modificações que lhe são propostas.

A Figura 3.2, extraída de Arenales *et al.* (2007 *apud* Neto, Pureza *et al.*, 2012) ilustra o processo de modelagem simplificado de modelagem matemática utilizado na pesquisa.

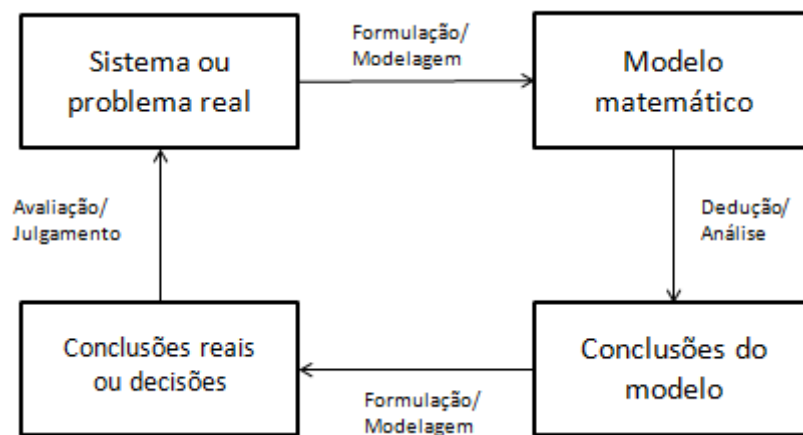


Figura 3.2 – Processo de modelagem

Fonte: Arenales *et al.* (2007 *apud* Neto, Pureza *et al.*, 2012)

Tendo em vista a natureza da pesquisa, considera-se uma pesquisa aplicada, visto que busca aplicação da teoria para solução de problemas que ocorrem na realidade, ou seja, apresenta um interesse prático. De acordo com a Appolinário (2006), a pesquisa aplicada estaria ligada a objetivos comerciais por meio do desenvolvimento de processos ou produtos orientados para as necessidades do mercado, ao contrário da pesquisa básica, a qual se envolve no desenvolvimento do conhecimento científico.

Quanto aos objetivos, a pesquisa é caracterizada como pesquisa exploratória já que busca proporcionar maior familiaridade ao problema de forma a torná-lo explícito e permitindo a criação de hipóteses. Além disso, envolve o levantamento bibliográfico, desenvolvimento e esclarecimento de conceitos e ideias.

De acordo com Neto, Pureza *et al* (2012), os modelos quantitativos podem ser definidos como modelos abstratos, descritos em linguagem matemática e computacional, que usam técnicas experimentais, voltadas para simulação e analíticas, voltadas para matemática e estatística. Dessa forma, tendo em vista a abordagem do problema, tem-se uma pesquisa quantitativa, já que utiliza de técnicas estatísticas e informações embasadas em números para a geração de informações, classificação e análise do problema.

Ainda conforme Neto, Pureza *et al* (2012), a abordagem quantitativa utilizada na pesquisa é caracterizada pelo método do tipo pesquisa operacional, visto que aborda a utilização de técnicas de previsão (*forecasting*) e conseqüentemente uma abordagem científica voltada para o auxílio do processo de tomada de decisões, análises estatísticas de dados históricos e projeção no futuro por meio de modelos matemáticos e estatísticos.

Partindo do princípio de que a premissa fundamental das técnicas de previsão é o entendimento e aplicação do comportamento passado com o objetivo de prever eventos futuros baseados na tendência passada (históricos), a pesquisa pode ser considerada também como pesquisa histórica, já que supõe que pelos menos grande parte do futuro seja réplica do passado.

### **3.2 Metodologia da pesquisa**

- Estudar e identificar alguns dos principais métodos de previsão de demanda e aplicações;
- Propor aplicação da metodologia de previsão de demanda baseada em Redes Neurais e Box-Jenkins;
- Coletar dados referentes ao histórico de vendas dos diferentes tipos de produtos;
- Modelar matematicamente as séries temporais;
- Confrontar os resultados obtidos mediante a referente literatura e ambos os métodos de previsão estudados;
- A partir dos resultados de demanda alcançados, dar suporte para as tomadas de decisão em todos os níveis de planejamento.



### 3.3 Dados

As demandas de produtos analisadas durante a dissertação foram coletadas em uma empresa que atua no mercado industrial do ramo metal mecânico, a qual, por questões de sigilo foi denominada como Empresa X.

Situada no Polo Empresarial da cidade de Aparecida de Goiânia – Goiás, a mesma produz peças mecânicas em alumínio, tais como: polias, engates rápidos, rodas, conexões hidráulicas, molas pneumáticas (*airspring*), dentre outras.

Os dados coletados e utilizados compõe o período de três anos, sendo estes, referentes ao período de janeiro de 2012 até dezembro até 2014. Após a coleta de dados, optou-se pela realização da curva ABC, na qual levou-se em consideração os aspectos de itens mais representativos mediante aos critérios de volume de vendas e faturamento da empresa.

### 3.4 Programa SPSS

A escolha de softwares para o desenvolvimento e interpretação de resultados é de fundamental importância para a pesquisa. Porém, existem algumas restrições que devem ser analisadas, como por exemplo, custo elevado de aquisição e dificuldade de interpretação e aprendizado com relação às suas funcionalidades.

O programa SPSS é uma importante ferramenta que apresenta um pacote de técnicas estatísticas que possibilitam a análise e interpretação de dados referentes a diversos segmentos. O objetivo principal da estatística é compreender o significado das informações contidas em diferentes conjuntos de dados. A utilização de softwares para interpretação dessas informações e resultados é de fundamental importância.

Bruni (2012) cita que o SPSS é um dos softwares considerados mais utilizados para análises estatísticas, além de apresentar uma interface bastante amigável com as versões mais recentes em ambiente Windows e ser considerado um recurso referencial para análise de dados.

### 3.5 Programa R

O programa R é uma ferramenta alternativa, de domínio público, utilizada para o ambiente da computação estatística, o qual fornece uma ampla variedade de técnicas e métodos estatísticos através dos vários pacotes com conjunto de funções disponíveis

(modelagem linear e não linear, testes estatísticos clássicos, análises de séries temporais, clusterização, classificação, entre outros). Além disso, o mesmo oferece ajuda para aplicação de técnicas gráficas com ótima qualidade para publicação, incluindo símbolos matemáticos e fórmulas quando necessário. Permite a instalação em muitas plataformas incluindo Unix, Linux, Macintosh e Windows. Os pacotes mais importantes utilizados pelo R na pesquisa foram o pacote *Forecast* e *tseries*, o qual apresenta os métodos e ferramentas utilizadas no estudos de previsões e séries temporais, respectivamente.

### 3.6 Programa Matlab

O software Matlab é uma ferramenta bastante utilizada em otimização, processamento de sinais e inclusive, soluções de aprendizagem de máquinas. Apresenta enorme poder quando se trata de cálculos numéricos, já que contempla uma extensa parte matemática e capacidade gráfica, as quais são características fundamentais para treinamento e desenvolvimento de redes neurais.

O Matlab apresenta usabilidade interessante, já que a maioria de suas ferramentas encontra-se dividida em caixas de ferramentas (*toolboxes*), o que permite facilitar a visualização das funções existentes quando existe a necessidade de acessar a função *Help* do programa. Além disso, também apresenta a opção de utilização de linhas de comando.

Apesar da facilidade de utilização da ferramenta de *toolbox*, na qual pode ser utilizada a função clássica *newff* ou outra alternativa denominada como *feedforwardnet*, para redes do tipo Perceptrons de Multicamadas, optou-se neste trabalho pela implementação da rede neural MLP por meio de linhas de comandos, visto que esta mostrou-se mais ideal a nível de pesquisa, já que permitiu-se criar uma rede genérica a qual apresenta uma lógica capaz de realizar testes e simulações dos principais parâmetros e por fim realizar várias comparações.

### 3.7 Método de seleção dos dados

A escolha do material a ser analisado é uma etapa primordial para a pesquisa em questão, visto que apesar da indústria ser de pequeno porte, a mesma apresenta um elevado mix de produtos nas mais diversas famílias. E, além disso, desenvolver os modelos de previsão em todos os itens se mostra praticamente inviável para um trabalho acadêmico.

Dessa forma, a amostra de dados selecionada para o estudo foi definida por meio da aplicação da curva ABC, que se preocupa em considerá-los de acordo com sua

representatividade nos resultados da organização, ou seja, quando um pequeno número de itens demandados frequentemente apresenta uma enorme porcentagem de dominância. Quando tem-se exemplos dessa forma, onde uma pequena parcela de itens corresponde a uma representatividade enorme para a organização, estes são classificados como pertencentes à Curva A e a medida que a representatividade decresce e o número de itens aumenta estes devem ser classificados como itens da Curva B e C, respectivamente.

Conforme o Apêndice C, após a realização da Curva ABC, chegou-se a conclusão de que, durante este período analisado, 201 itens são pertencentes à Curva A, o que ainda assim, é considerado como uma enorme amostra de dados e por isso, optou-se por selecionar os 03 primeiros itens mais representativos da demanda durante o período de janeiro de 2012 a dezembro de 2014.

Os itens selecionados pertencem a três famílias importantes de produtos da empresa, sendo elas: família de elementos diversos de máquinas e equipamentos, e conexões. Os produtos escolhidos foram denominados Produto 1, Produto 2 e Produto 3.

Após a seleção dos produtos mais relevantes a serem estudados, a partir dos resultados gerados na curva ABC, tabulou-se os dados de cada um deles a fim de se obter a série temporal da demanda do produto.

### **3.8 Modelos de previsão**

#### **3.8.1 Modelo ARIMA**

Para que o modelo de Box-Jenkins seja construído é importante que se estabeleça quatro etapas fundamentais de acordo com Gujarati (2006), sendo elas: identificação, estimação, verificação ou checagem e previsão.

Para que se possa identificar os modelos ideais, é importante analisar graficamente a série temporal dos modelos em estudo, de forma que seja possível verificar a presença de tendência ou alterações na variância, que permitem revelar se a série é classificada como estacionária ou não, ou seja, se apresenta um processo em “equilíbrio”. Uma vez que a série se mostra com uma inclinação bastante pequena, a mesma será considerada como estacionária.

Durante a etapa de identificação é que os critérios de definição do comportamento da série são identificados. É nesta etapa que busca-se analisar se a série segue um processo

autorregressivo AR, que define o valor de  $p$ , ou um processo de médias móveis MA, o qual determina-se o valor de  $q$ , ou um processo autorregressivo integrado de médias móveis ARIMA, onde os valores de  $p$ ,  $d$  e  $q$  são definidos.

Para que se possa gerar a ordem dos termos  $p$ ,  $d$  e  $q$  são utilizados os métodos da função de autocorrelação (ACF) e autocorrelação parcial (PACF).

De acordo com Hyndman e Khandakar (2008), para muitas pessoas a etapa de seleção de ordem dos parâmetros  $(p, d, q)$  no ARIMA é considerada como um obstáculo comum, já que é considerada subjetiva e difícil de ser aplicada.

Para a aplicação dos modelos ARIMA primeiramente realizou-se o tratamento dos dados primários, a partir da retirada dos pontos fora da curva (*outliers*) presentes nas séries observadas. Os *outliers* foram encontrados por meio do gráfico de Box-Plot, analisados e tratados no software estatístico SPSS, a fim de se obter a suavização da série temporal em cada um dos produtos observados. Após o ajuste da série, buscou-se determinar qual o modelo ARIMA adequado para os dados. Para tal, utilizou-se o software R, o qual possibilita determinar de maneira automática, a partir da função “auto.arima”, o melhor modelo a ser utilizado para as séries de dados inseridas e conseqüentemente todos os seus parâmetros associados.

Para cada uma das séries de dados estudadas gerou-se o modelo ARIMA, fez-se a análise dos resultados referentes aos critérios estatísticos e realizou-se o comparativo com o modelo de rede neural desenvolvido. Uns dos critérios principais utilizados como comparativo dos modelos foram os resultados encontrados no RMSE juntamente com a análise referente à tendência da previsão gerada em todos os modelos.

### **3.8.2 Modelo de Redes Neurais**

Para a implementação da rede neural, inicialmente organizou-se a base de dados com as informações relevantes, no período observado e quanto aos produtos estudados. Por meio do *software* Matlab R2014b, fez-se a programação da rede neural seguindo as etapas do algoritmo de retropropagação sugerido por Haykin (2001) e citado no Capítulo 2.

Antes de se iniciar a etapa de treinamento da rede neural realizou-se a normalização dos dados reais coletados com o objetivo de aprimorar o desempenho da rede estudada. Durante este processo aplicou-se a normalização sigmoïdal, conforme a Equação 27:

$$y = \frac{1}{1 + e^{-\frac{x - \bar{x}}{\sigma_x}}} \quad (27)$$

Durante a própria programação do algoritmo de retropropagação definiu-se as porcentagens de dados utilizadas durante o treinamento da rede, como 70%. Os 30% restantes foram destinados às etapas de teste. Além disso, a programação contou com a inserção dos dados primários como sendo os valores de entrada da rede MLP desenvolvida.

As demais etapas se resumiram na seleção do grupo de dados trabalhados, bem como as etapas de treinamento, teste e geração da previsão da rede neural programada, tendo sido realizados repetidos testes, até que se encontrassem resultados consideravelmente adequados ao cenário de demanda.

Os parâmetros referentes à taxa de aprendizado, quantidade de neurônios, número de iterações e camadas de entrada foram alterados empiricamente com o intuito de identificar um modelo com resultados satisfatórios.

A determinação do algoritmo de treinamento a ser utilizado e a definição de todos os seus parâmetros estão diretamente ligados ao desempenho desejado e ao tempo de treinamento da rede. Desta maneira, taxas de aprendizado muito baixas, por exemplo, acarretam em um aprendizado lento, enquanto que taxas de aprendizado muito elevadas podem causar oscilações no processo de treinamento, dificultando ou até mesmo impedindo a convergência do algoritmo de treinamento *error backpropagation* (HAYKIN, 2001). Dessa forma, optou-se por variar a taxa de aprendizagem entre 0,3 e 0,7.

O critério de parada da rede neural durante a etapa de treinamento foi a partir do limite definido de iterações, de forma que a rede será testada até o momento em que encontrar o menor erro RMSE, garantindo assim um melhor desempenho do modelo.

É importante ressaltar que a rede neural criada durante a pesquisa é genérica, ou seja, possibilita ajuste dos principais parâmetros, como por exemplo: taxa de aprendizagem, número de entradas, quantidade de neurônios, entre outros.

## 4. RESULTADOS E DISCUSSÕES

Nesta seção, os dois modelos de análise propostos são avaliados para cada um dos produtos estudados e posteriormente, seus resultados são comparados mediante aos critérios dos valores encontrados pelo MAPE e RMSE em cada um dos métodos.

As análises foram realizadas com os três produtos mais relevantes para a empresa X durante 36 períodos dos anos referentes a 2012 até 2014, sendo que os 04 últimos períodos da série, correspondente aos meses de setembro/2014, outubro/2014, novembro/2014 e dezembro/2014, foram utilizados apenas para validação de desempenho de ambos os métodos propostos a partir das análises dos erros de previsão RMSE e MAPE. Os itens foram denominados na sequência do Produto 1 até o Produto 506, representando toda a curva A e B, conforme mostrado no Apêndice C, sendo que destes, os produtos escolhidos para serem estudados foram o Produto 1, Produto 2 e Produto 3.

### 4.1 Análise estatística do Produto 1

O processo de análise dos dados selecionados inicialmente se deu a partir da análise das características estatísticas, realizadas por meio da observação dos gráficos das séries, gráfico de autocorrelação ACF, autocorrelação parcial PACF, gráfico de BoxPlot, Histograma, teste de normalidade e de autocorrelação dos resíduos, conforme apresentado nas Figuras 4.1 até Figura 4.7

A Figura 4.1 representa a série temporal com os dados reais do Produto 1 dos 32 períodos analisados, onde é possível observar por meio de um diagnóstico preliminar que não há existência de padrões sazonais da demanda deste produto. Além disso, percebe-se ainda oscilações mensais ao longo dos 03 anos analisados.

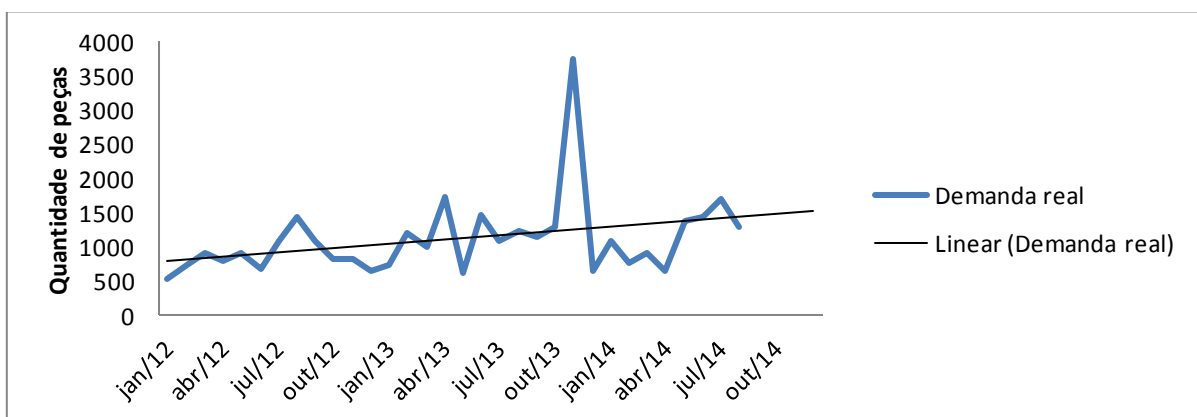


Figura 4.1 – Série temporal real do Produto 1

Fonte: Desenvolvido pelo autor

Para identificar se o modelo estudado é adequado para realização de previsões, inicialmente é necessário analisar o gráfico de tempo da série em estudo, conforme mostrado na Figura 4.1. A partir desta análise, é possível verificar se existe a presença de tendência ou alterações na variância dos dados e, a partir disto, determinar a estacionariedade ou não da série.

Quanto à quantidade de vendas do Produto 1, ilustrado na Figura 4.1, é possível observar que a série não apresenta uma tendência com inclinação abrupta, apenas uma mínima inclinação ascendente, possibilitando considerá-la como uma série estacionária.

Através do BoxPlot, Figura 4.2, é possível observar que os dados foram agrupados anualmente, notando-se que existe um crescimento ascendente da demanda deste produto quando se comparada entre os anos de 2012 e 2013, enquanto entre 2013 e 2014 há uma certa estabilidade. Além disso, nota-se a presença de *outliers* no período 8, referente ao ano de 2012 e no período 23 do ano de 2013.

Os principais motivos apontados pela área comercial da empresa para a presença dos *outliers* encontrados durante estes períodos observados seriam a captação e venda de dois clientes com alto potencial de compra desse item durante os períodos observados, sendo um deles referente a uma importadora que apresentou um consumo acima dos padrões mensais.

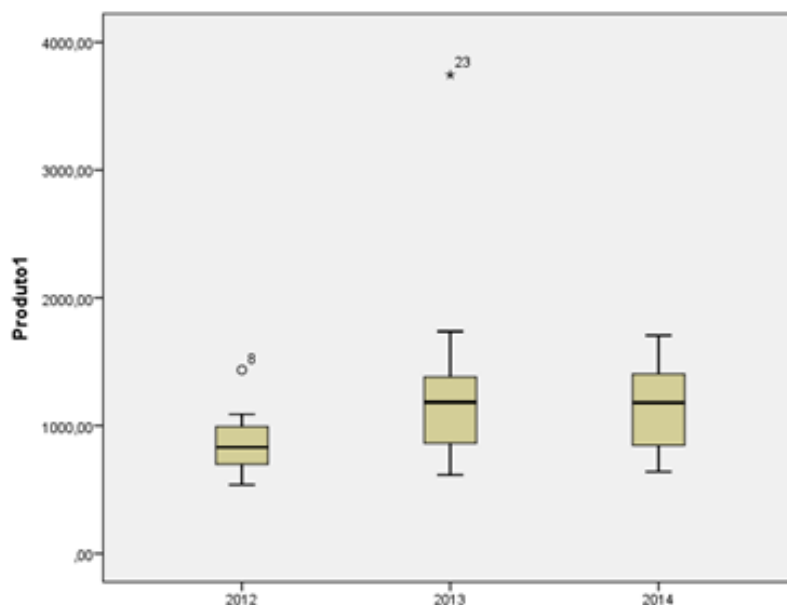


Figura 4.2 – BoxPlot anual do Produto 1

Fonte: Desenvolvido pelo autor

Embora tenha sido justificada a presença de *outliers* na série observada, fez-se o tratamento dos dados a fim de se retirar estes pontos que apresentam um grande afastamento com relação aos demais da série. Este tratamento consistiu na substituição do ponto fora da série pelo valor referente à média do período analisado, onde no caso, é definido pelo período anual da série. Dessa forma, tem-se um resultado de BoxPlot com uma série de dados tratada e conseqüentemente, sem a presença de *outliers*, conforme consta na Figura 4.3.

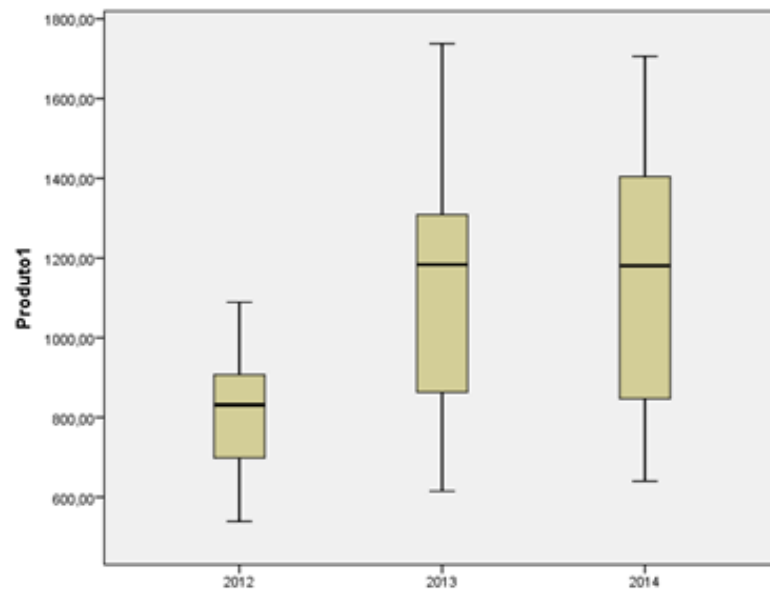


Figura 4.3 – BoxPlot anual após tratamento de *outliers* do Produto 1

Fonte: Desenvolvido pelo autor

Após o tratamento dos *outliers* encontrados, tem-se a nova série de dados representada a partir da Figura 4.4.

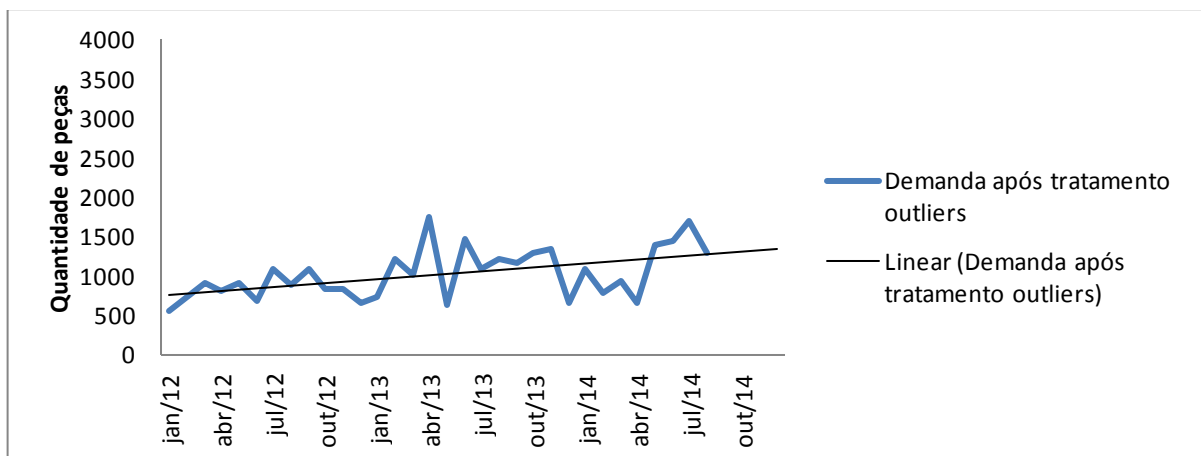


Figura 4.4– Série temporal após tratamento de *outliers* do Produto 1

Fonte: Desenvolvido pelo autor



Na Tabela 4.1 são apresentadas as principais medidas descritivas pela série de demanda de vendas do Produto 1 após o tratamento dos dados para a retirada dos *outliers* presentes no modelo inicial. A partir destas medidas, é possível ter uma visão geral do comportamento dessa variável ao longo do tempo descrito. Nota-se uma variância nos dados ao longo do período analisado, representando 102.811,9. Os valores 763,2, e 1239,0 deixam 25% (1º Quartil) e 75% (3º Quartil) da amostragem abaixo deles, respectivamente, enquanto tem-se o valor da mediana, indicando que 50% das observações estão abaixo desse valor e, conseqüentemente 50% acima dele, representando 957,5.

Tabela 4.1– Medidas descritivas da série de dados do Produto 1

Parâmetros	Média	Mediana	Desvio Padrão	Mínimo	Máximo	1º Quartil	3º Quartil	Variância
<b>Valores</b>	1.019,0	957,5	320,6	539,0	1.738,0	763,2	1.239,0	102.811,9

Fonte: Desenvolvido pelo autor

O próximo passo após as análises de medidas descritivas da série temporal é analisar as funções de autocorrelação (ACF) e autocorrelação parcial (PACF), que buscam auxiliar na verificação de estacionariedade e posteriormente, na definição do modelo mais adequado às características dos dados.

A finalidade da função de autocorrelação (ACF) e autocorrelação parcial (PACF) é identificar qual o melhor modelo ARIMA a ser utilizado a partir da medida da significância das correlações entre as observações presentes e as passadas, e determinar quantas defasagens no tempo estão correlacionadas.

Nas Figuras 4.5 e 4.6, referente às séries temporais da demanda do Produto 1, é possível observar, tanto na ACF quanto na PACF, estacionariedade dos dados, visto que existe a presença de aleatoriedade próxima de 0.

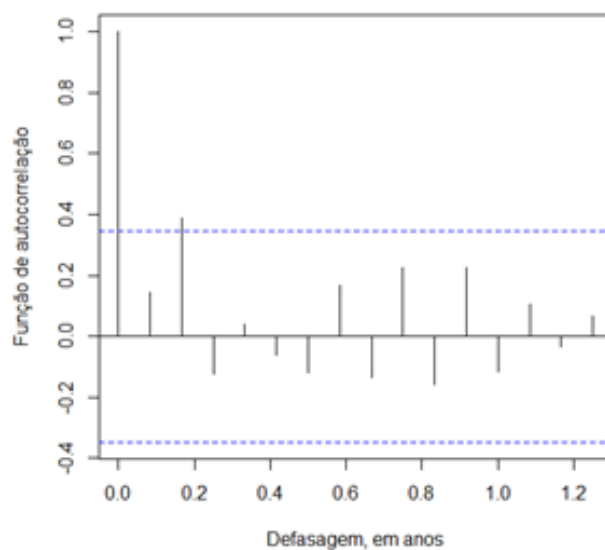


Figura 4.5 – Função de autocorrelação do Produto 1

Fonte: Desenvolvido pelo autor

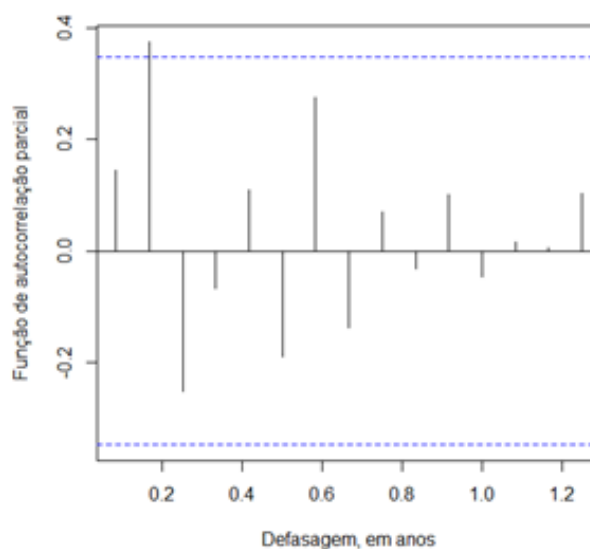


Figura 4.6 – Função de autocorrelação parcial do Produto 1

Fonte: Desenvolvido pelo autor

Pela Figura 4.7, nota-se que o histograma da série de dados apresenta uma mediana inferior à média, sendo a primeira com o valor correspondente à 957,5 enquanto a segunda é de 1019,0. Além disso, observa-se que os períodos finais do histograma terminam abruptamente de um dos lados. Provavelmente, estas observações são causadas após a eliminação dos dados discrepantes (*outliers*) na série temporal devido anomalias nas vendas e este “corte” coincide com os limites de especificação.

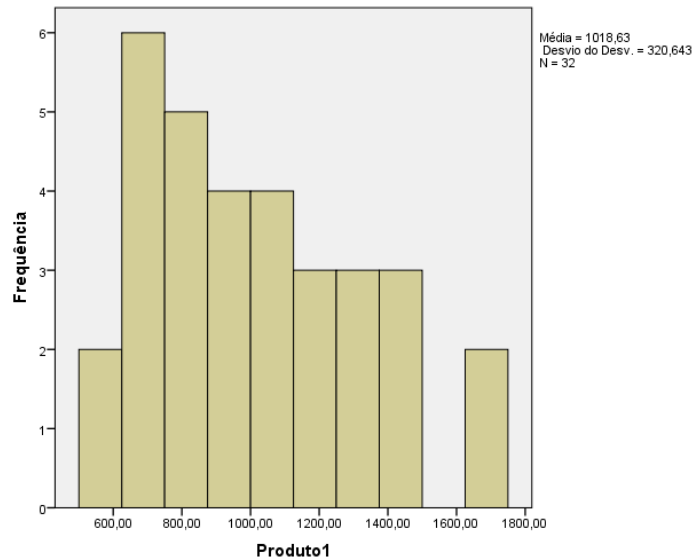


Figura 4.7 – Histograma do Produto 1

Fonte: Desenvolvido pelo autor

## 4.2 Método ARIMA para o Produto 1

Para a análise do ARIMA referente ao Produto 1, realizou-se automaticamente a seleção dos parâmetros  $p$ ,  $d$  e  $q$  por meio do *software* R e obteve-se os resultados apresentados a seguir.

O correlograma dos resíduos indicam que na amostra de vendas do Produto 1 há um padrão sinusoidal sobre 0, ou seja, há uma persistente aleatoriedade positiva e negativa, com valores próximos de zero, padrão típico de ACF. Dessa forma, de acordo com a Figura 4.8, conclui-se que tanto para o correlograma que apresenta os resíduos padronizados quanto para o ACF dos resíduos, há presença de aleatoriedade na série. Logo, a série temporal é considerada estacionária, não sendo necessário diferenciá-la.

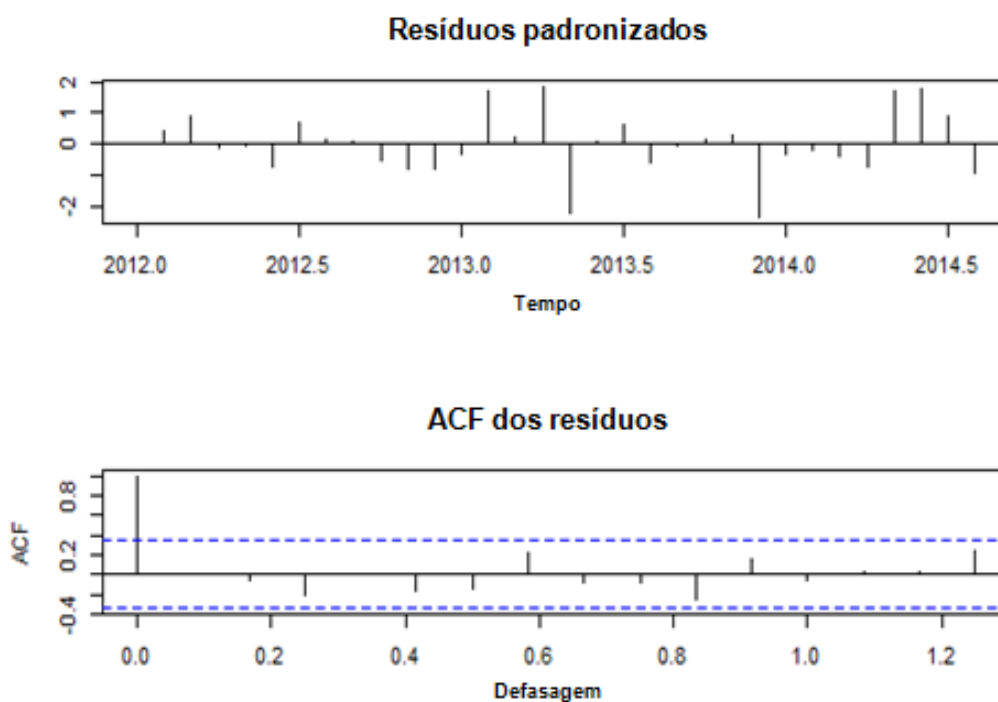


Figura 4.8 – Correlogramas de Padronização e ACF dos resíduos do Produto 1

Fonte: Desenvolvido pelo autor

Além disso, a partir dos correlogramas da Figura 4.8 determina-se o tipo do modelo. Devido aos dados não apresentarem sazonalidade, considerou-se o ARIMA ( $p$ ,  $d$  e  $q$ ). O modelo ótimo estimado pelo *software* R para série de dados trabalhada foi o modelo ARIMA autorregressivo diferenciado de primeira ordem, ou ARIMA (1,1,0)(0,0,0), conforme mostrado na Tabela 4.2.

Tabela 4.2 – Descrição do modelo ARIMA do Produto 1

Descrição do modelo	
<b>Produto 1</b>	<b>Tipo do modelo</b> ARIMA (1,1,0)(0,0,0)

Fonte: Desenvolvido pelo autor

Para etapa de verificação de autocorrelação, aplicou-se o teste de Ljung-Box, também conhecido como estatística de Box-Pierce modificado, aos resíduos da série, conforme mostrado na Tabela 4.3, onde é possível indicar se existe uma especificação correta do modelo. Um  $p$ -value de 0,9866 significa que as autocorrelações residuais do modelo são estatisticamente não significativas, ou seja, os resíduos se comportam como ruído branco,

sendo considerado como corretamente especificado e adequados para a realização de previsões.

Tabela 4.3 – Teste de autocorrelação Ljung-Box dos resíduos do Produto 1

Produto 1	Autocorrelação Ljung-Box	
	Estatísticas ( $X^2$ )	p-value
	0,0003	0,9866

Fonte: Desenvolvido pelo autor

Além disso, observa-se pela Figura 4.9 que todos os valores das autocorrelações apresentados entre os *lags* de 1 a 10 são estatisticamente não significativos e, conseqüentemente, especificados corretamente para a previsão.

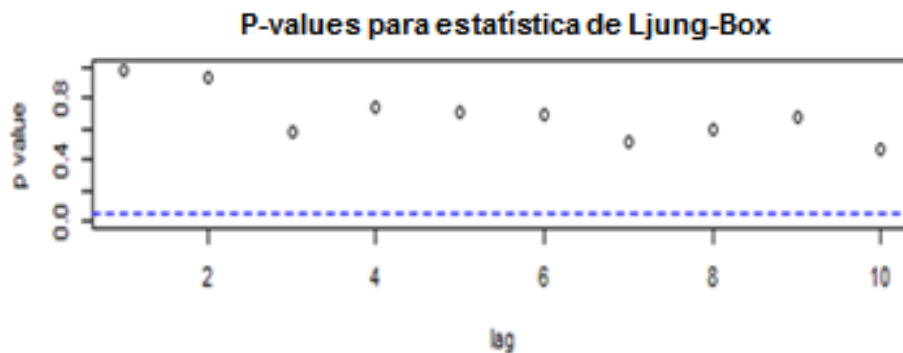


Figura 4.9 – *P-values* para teste Ljung-Box

Fonte: Desenvolvido pelo autor

A Figura 4.10 ilustra o QQ-Plot dos resíduos para o Produto 1, que quando aplicados ao teste de normalidade Shapiro-Wilk, conforme mostrado na Tabela 4.4, apresentaram como resultado  $W = 0,9454$  e  $p\text{-value} = 0,1065$ , ou seja, os resíduos apresentam distribuição normal, com nível de significância de 5%.

Tabela 4.4 – Teste de normalidade Shapiro-Wilk dos resíduos do Produto 1

Produto 1	Teste de normalidade	
	Shapiro-Wilk	
	W	p-value
	0,9454	0,1065

Fonte: Desenvolvido pelo autor

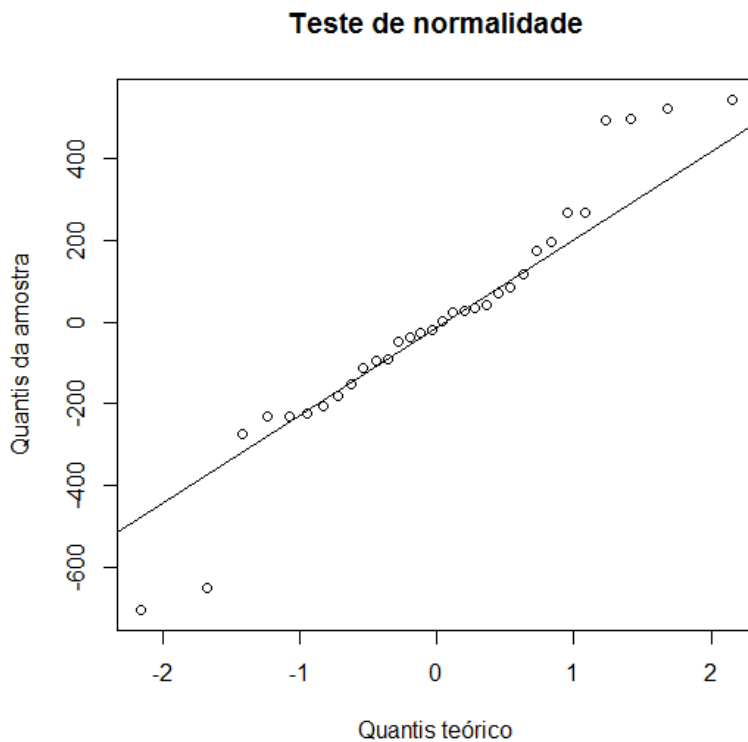


Figura 4.10 – QQ-Plot dos resíduos para o Produto 1

Fonte: Desenvolvido pelo autor

Na Figura 4.11 apresenta-se o gráfico comparativo entre a série temporal real analisada, dados originais, e a série temporal predita, dados processados, pelo *software* R.

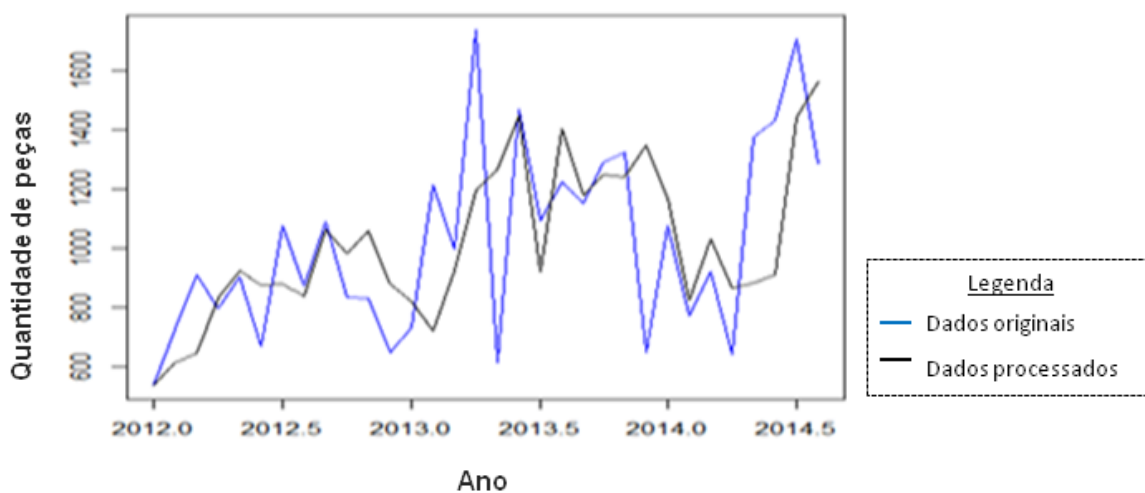


Figura 4.11 – Comparativo entre a demanda real e predita do Produto 1

Fonte: Desenvolvido pelo autor

Para o gráfico do Produto 1 apresentado na Figura 4.12, é possível analisar os valores da série observada, bem como os valores dos limites inferiores e superiores para o período da

série predita. Por fim, tem o resultado final da metodologia de Box-Jenkins que é a previsão do modelo gerado.

O resultado da previsão apresenta pequenas variações nos primeiros meses e posteriormente uma constância dos valores gerados, mostrando pequenas variações na demanda das vendas.

A partir dos resultados preditos pelo ARIMA no *software* R fez-se a validação de desempenho dos erros de previsão RMSE e MAPE comparando os valores reais dos 04 últimos períodos do histórico de dados analisado com os valores gerados no modelo em estudo. A partir disso, a estatística de ajuste do modelo de previsão apresentou um valor referente à Raiz do Erro Médio Quadrático (RMSE) correspondente a 572,16, enquanto o MAPE representou 20%. Ambos os parâmetros foram utilizados para a comparação entre o modelo desenvolvido pela Rede Neural MLP em todos os produtos.

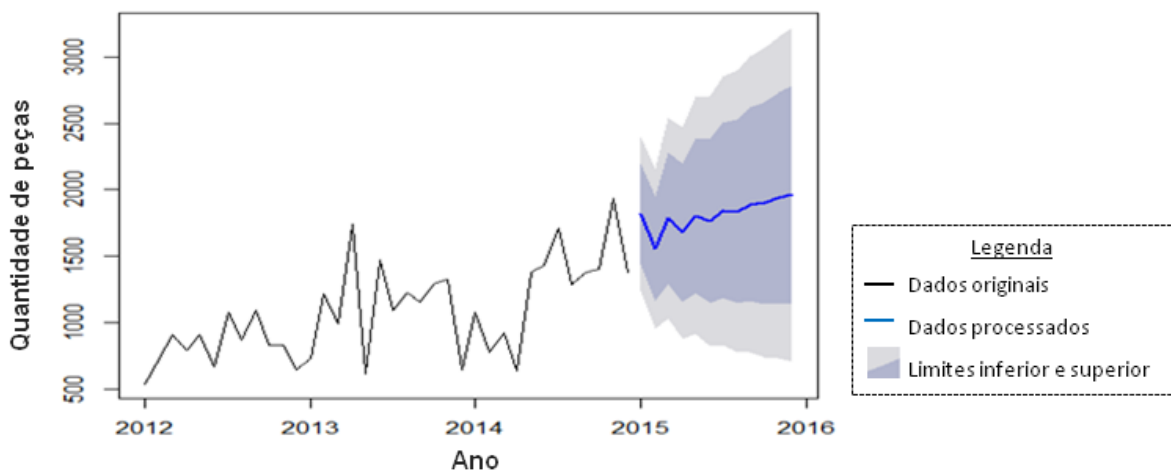


Figura 4.12 – Previsão método Box-Jenkins do Produto 1

Fonte: Desenvolvido pelo autor

### 4.3 Análise estatística do Produto 2

Na Figura 4.13 a série temporal do Produto 2 é apresentada contemplando os dados referentes as vendas da Empresa X com a transformação de Yeo-Johnson, utilizada devido os dados da série deste produto apresentarem valores iguais a 0, característica válida para este tipo de transformação. Diferentemente dos demais produtos analisados, o Produto 2 não apresentou *outliers* na série, e dessa forma, não foi necessário realizar nenhum tratamento na série.

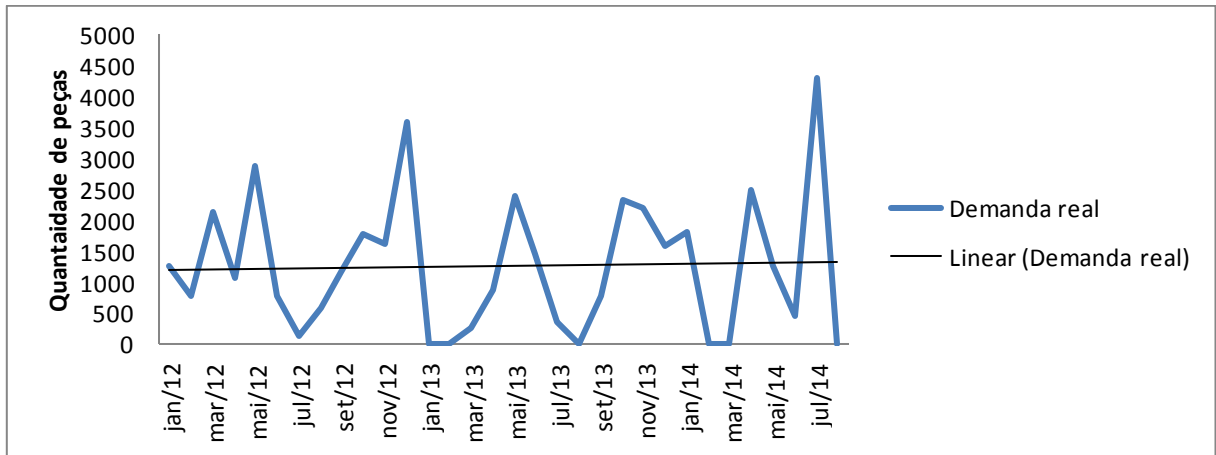


Figura 4.13 – Série temporal real do Produto 2

Fonte: Desenvolvido pelo autor

A Figura 4.14 apresenta o BoxPlot da série temporal do Produto 2, onde observa-se um pequeno declínio das vendas entre os anos de 2012 e 2013. Já entre os anos de 2013 e 2014 observou-se uma constância das vendas.

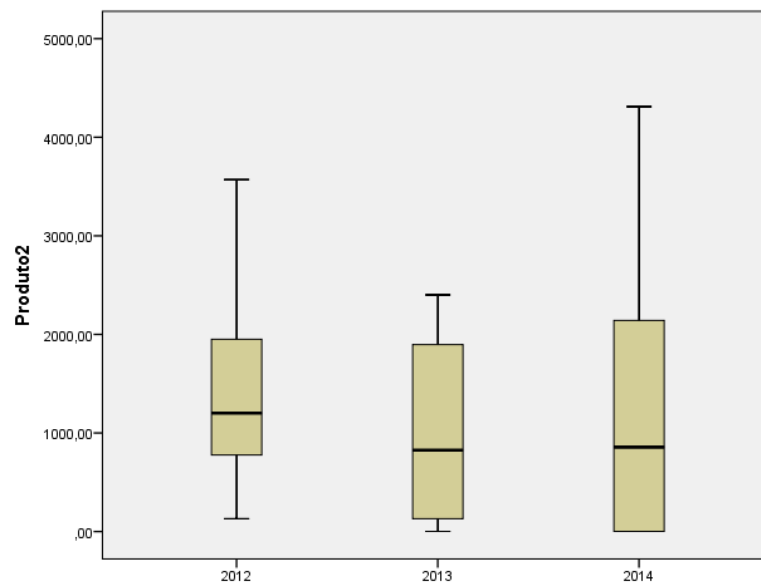


Figura 4.14 – BoxPlot anual do Produto 2

Fonte: Desenvolvido pelo autor

As principais medidas descritivas pela série de demanda de vendas do Produto 2 foram apresentadas na Tabela 4.5 e através destas medidas, é possível ter uma visão geral do comportamento dessa variável ao longo do tempo. Observa-se variância nos dados ao longo do período analisado, representando 1.237.619,0, os valores 336,0 e 1.886,0 representam o 1º Quartil e o 3º Quartil da amostragem, respectivamente. Enquanto o valor da mediana,



indicando 50% das observações, estão abaixo desse valor e, conseqüentemente 50% acima dele, representando 1101.

Tabela 4.5 – Medidas descritivas da série de dados do Produto 2

Parâmetros	Média	Mediana	Desvio Padrão	Mínimo	Máximo	1º Quartil	3º Quartil	Variância
<b>Valores</b>	1.258,0	1.101,0	1.112,5	1,0	4.311,0	336,0	1.886,0	1.237.619,0

Fonte: Desenvolvido pelo autor

Após as análises de medidas descritivas da série temporal é necessário avaliar as funções de autocorrelação (ACF) e autocorrelação parcial (PACF) a fim de auxiliar na verificação de estacionariedade e posteriormente na definição do modelo mais adequado às características dos dados.

Nas Figuras 4.15 e 4.16 referentes à série temporal da demanda do Produto 2 é possível observar, tanto na ACF quanto na PACF, estacionariedade dos dados, visto que existe aleatoriedade próxima de 0.

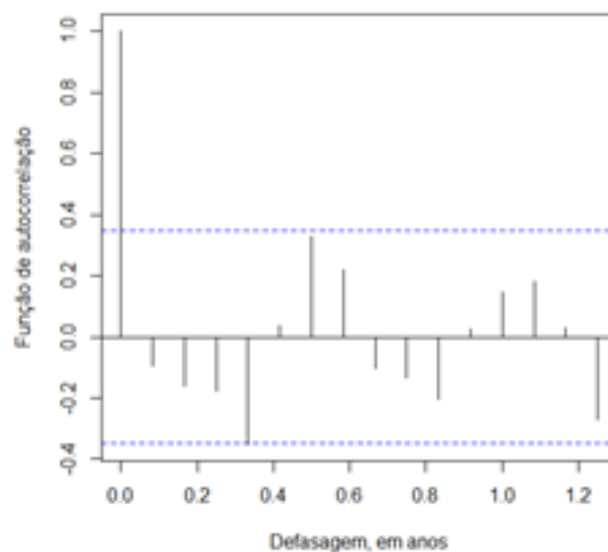


Figura 4.15 – Função de autocorrelação do Produto 2

Fonte: Desenvolvido pelo autor

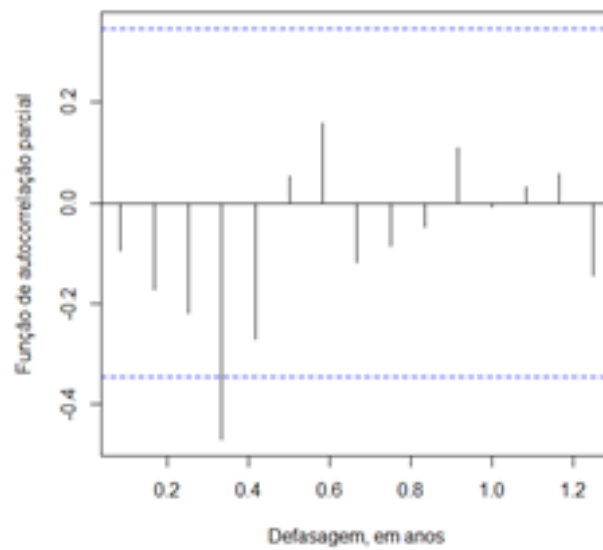


Figura 4.16 – Função de autocorrelação parcial do Produto 2

Fonte: Desenvolvido pelo autor

A Figura 4.17 apresenta o histograma da série de dados exibindo características semelhantes ao histograma do Produto 1, terminando abruptamente de um dos lados, originadas provavelmente pelas características dos dados que apresentam anomalias nas vendas em determinados períodos, conforme pode ser visto no Apêndice A.

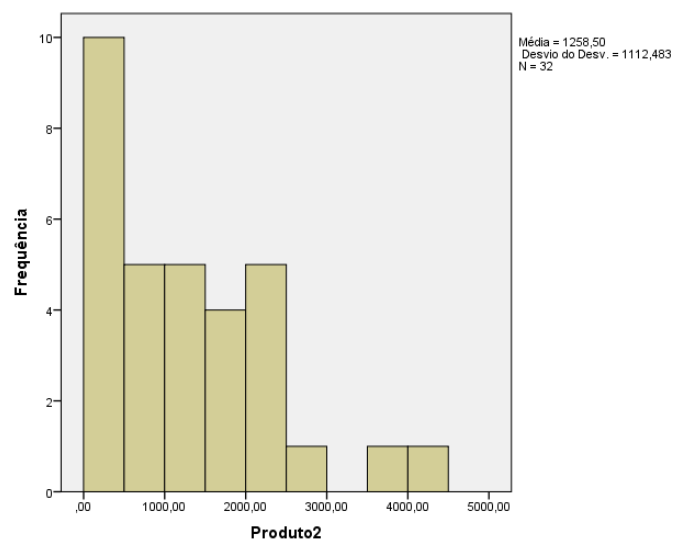


Figura 4.17 – Histograma do Produto 2

Fonte: Desenvolvido pelo autor

#### 4.4 Método ARIMA para o Produto 2

O correlograma dos resíduos indicam que, na amostra de vendas do Produto 2, também existe uma aleatoriedade positiva e negativa, com valores próximos de zero, padrão

típico de ACF. Assim, de acordo com a Figura 4.18, conclui-se que a série temporal é avaliada como estacionária, visto que, tanto para o correlograma que apresenta os resíduos padronizados quanto para o ACF dos resíduos há presença de aleatoriedade na série.

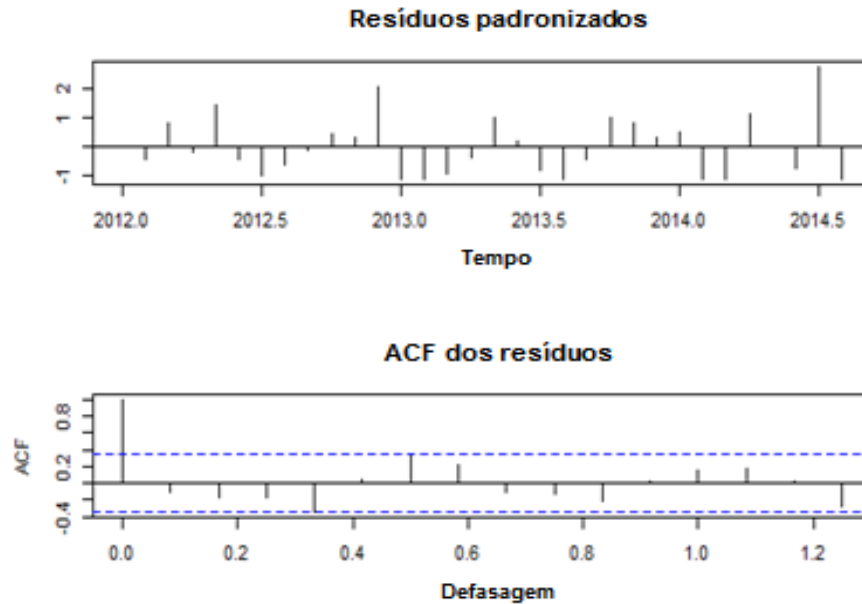


Figura 4.18 – Correlogramas de Padronização e ACF dos resíduos do Produto 2

Fonte: Desenvolvido pelo autor

Os dados do Produto 2 não apresentam sazonalidade, sendo considerados como ARIMA ( $p$ ,  $d$  e  $q$ ). O modelo determinado pelo *software* R para série de dados trabalhada foi o ARIMA de ruído branco, ou ARIMA (0,0,0)(0,0,0), conforme mostrado na Tabela 4.6.

Tabela 4.6 – Descrição do modelo ARIMA do Produto 2

Descrição do modelo	
<b>Produto 2</b>	<b>Tipo do modelo</b> ARIMA (0,0,0)(0,0,0)

Fonte: Desenvolvido pelo autor

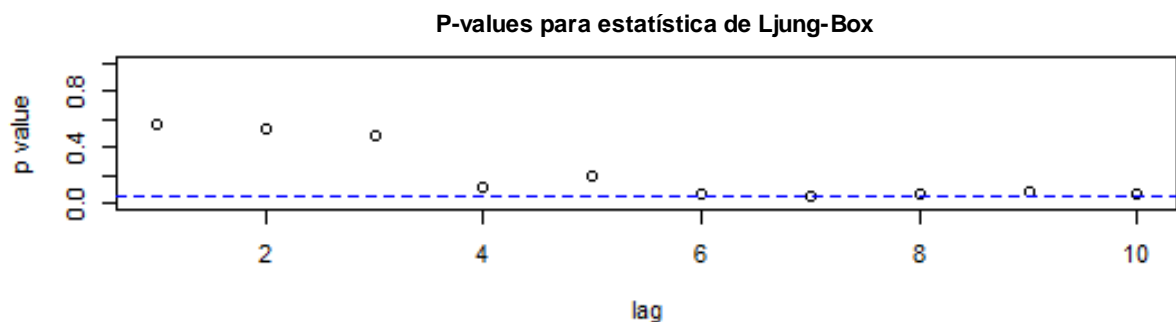
O teste de Ljung-Box, utilizado na etapa de verificação de autocorrelação, mostrou que os resíduos se comportam como ruído branco, devido ao valor de 0,57 apresentado. Dessa maneira, o modelo especificado está devidamente adequado para a determinação da previsão.

Tabela 4.7 – Teste de autocorrelação Ljung-Box dos resíduos do Produto 2

Produto 2	Autocorrelação Ljung-Box	
	Estatísticas ( $X^2$ )	p-value
	0,3227	0,57

Fonte: Desenvolvido pelo autor

De acordo com a Figura 4.19, os valores das autocorrelações apresentados entre os *lags* de 1 a 3 são estatisticamente não significativos e conseqüentemente, especificados corretamente para a previsão. Já a partir do quarto *lag*, os valores tendem a 0.

Figura 4.19 – *P-values* para teste Ljung-Box

Fonte: Desenvolvido pelo autor

Na Figura 4.20 observa o QQ-Plot dos resíduos para o Produto 2, que quando aplicados ao teste de normalidade Shapiro-Wilk, apresentaram como resultado  $W = 0,9185$  e  $p\text{-value} = 0,01889$ , ou seja, mesmo com a transformação de Yeo & Johnson aplicada, os resíduos não apresentaram distribuição normal, com nível de significância de 5%.

Tabela 4.8 – Teste de normalidade Shapiro-Wilk dos resíduos do Produto 2

Produto 2	Teste de normalidade	
	Shapiro-Wilk	
	W	p-value
	0,9185	0,01889

Fonte: Desenvolvido pelo autor

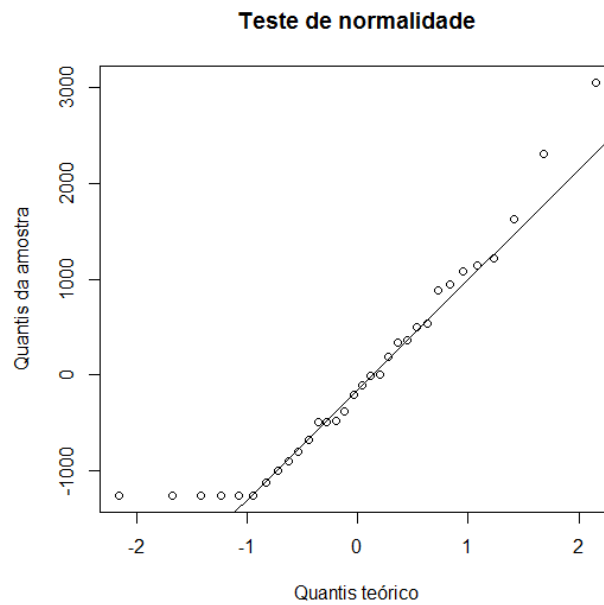


Figura 4.20 – QQ-Plot dos resíduos para o Produto 2

Fonte: Desenvolvido pelo autor

A Figura 4.21 ilustra o gráfico comparativo entre a série temporal real analisada, dados originais, e a série temporal predita, dados processados, pelo *software* R do Produto 2. Uma vez que o ARIMA da série é um ruído branco, os resultados processados pelo modelo são constantes.

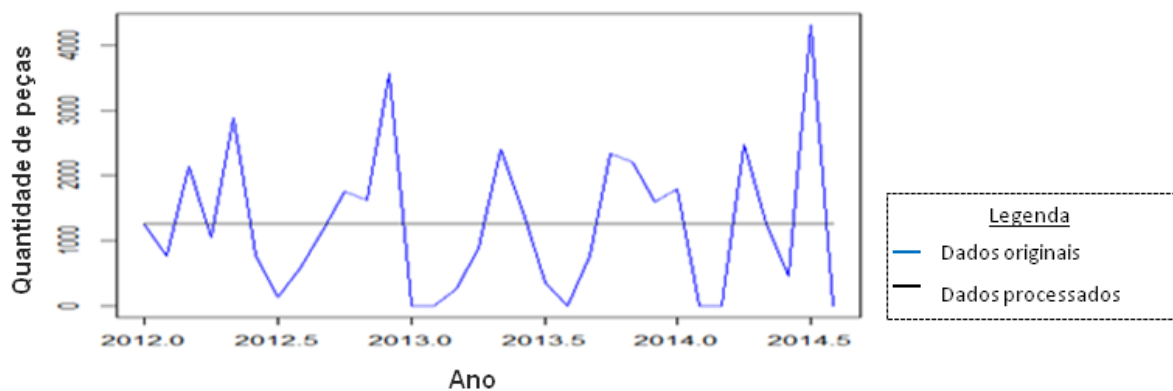


Figura 4.21 – Comparativo entre a demanda real e predita do Produto 2

Fonte: Desenvolvido pelo autor

Conforme a Figura 4.22, verifica-se o resultado da previsão 2, com tendência constante e ARIMA (0,0,0).

Por meio dos resultados previstos pela metodologia de Box-Jenkins no *software* R, realizou a validação de desempenho dos erros de previsão RMSE e MAPE e suas estatísticas

de ajuste do modelo de previsão apresentou um valor referente à Raiz do Erro Médio Quadrático (RMSE) correspondente a 817,994, enquanto o MAPE representou 107%.

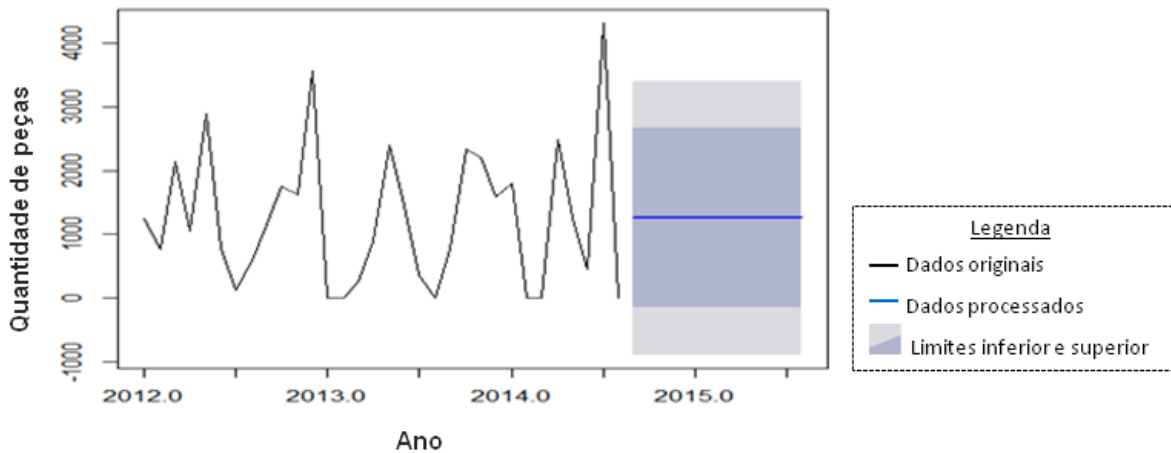


Figura 4.22 – Previsão método Box-Jenkins do Produto 2

Fonte: Desenvolvido pelo autor

#### 4.5 Análise estatística do Produto 3

Na Figura 4.23 a série temporal original do Produto 3 é apresentada, contemplando os dados reais referentes as vendas da Empresa X. Na Figura 4.24, a série após a análise do BoxPlot e tratamentos dos *outliers* é apresentada.

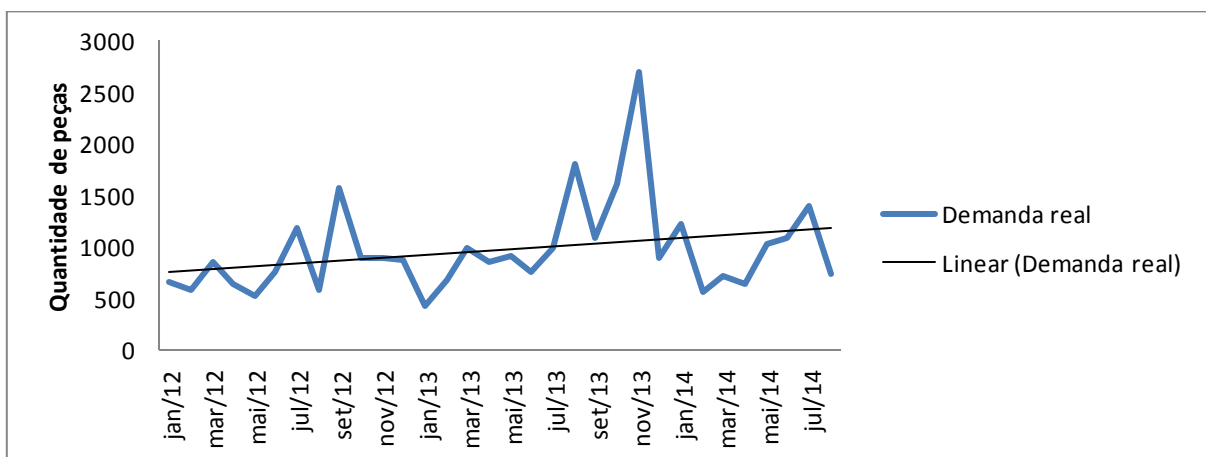


Figura 4.23 – Série temporal real do Produto 3

Fonte: Desenvolvido pelo autor

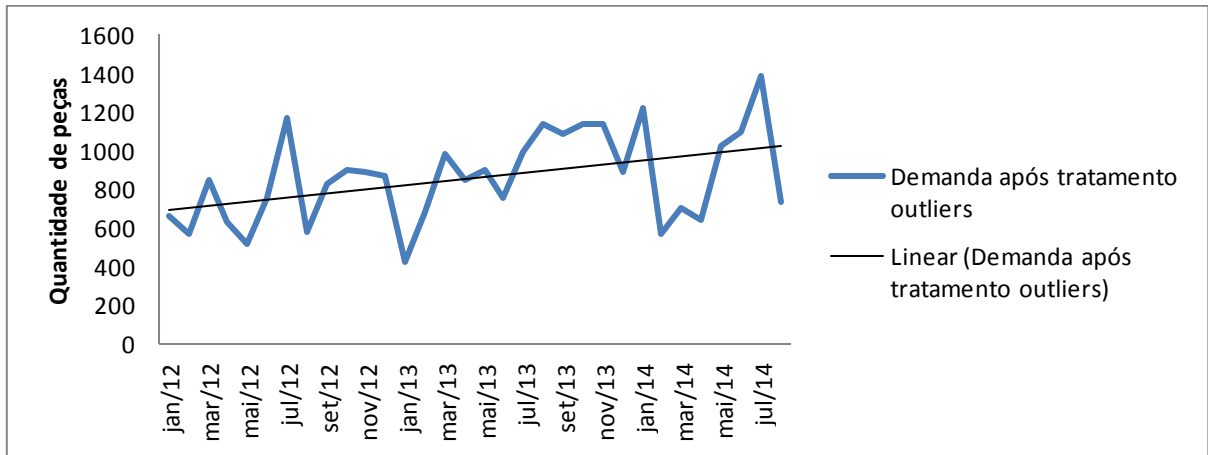


Figura 4.24 – Série temporal após tratamento de *outliers* do Produto 3

Fonte: Desenvolvido pelo autor

O mesmo processo de tratamento de *outliers* utilizado para o Produto 1 foi aplicado para o Produto 3, sendo que para este caso em específico, foram necessárias duas etapas de tratamento dos dados, conforme pode ser visto nos períodos coloridos (verde e roxo) e indicados com (1) e (2) do Apêndice A, e nas Figuras 4.25 até 4.27. No Apêndice A, os períodos que apresentam (1), na coluna de demanda tratada, representam os primeiros períodos em que o gráfico de BoxPlot apresentou pontos fora da curva, enquanto os que contém (2) são aqueles que necessitaram ser tratados posteriormente até a retirada completa dos pontos fora da curva.

A Figura 4.25 mostra os primeiros pontos com *outliers* encontrados no 8º período e 23º período da série de dados.

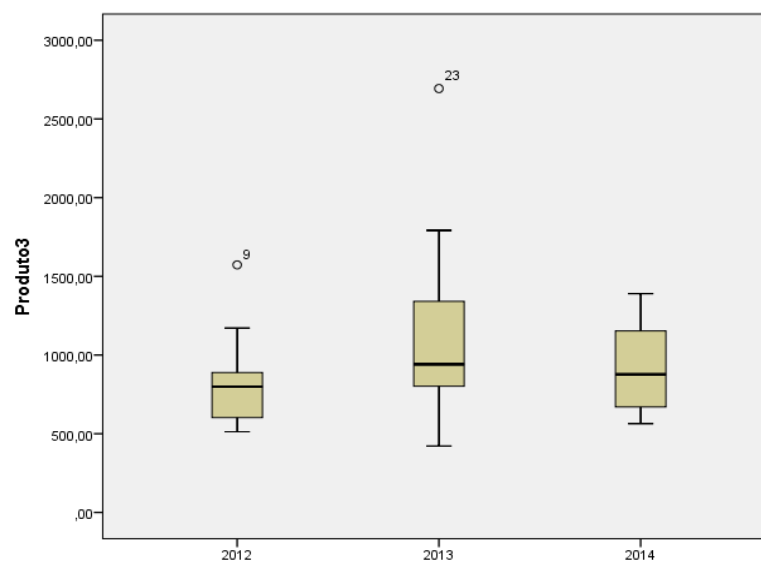


Figura 4.25 – BoxPlot anual sem tratamento de *outliers* do Produto 3

Fonte: Desenvolvido pelo autor

A Figura 4.26 apresenta os segundos pontos com *outliers* encontrados após o primeiro tratamento dos dados, sendo estes no 20º período e 22º período da série.

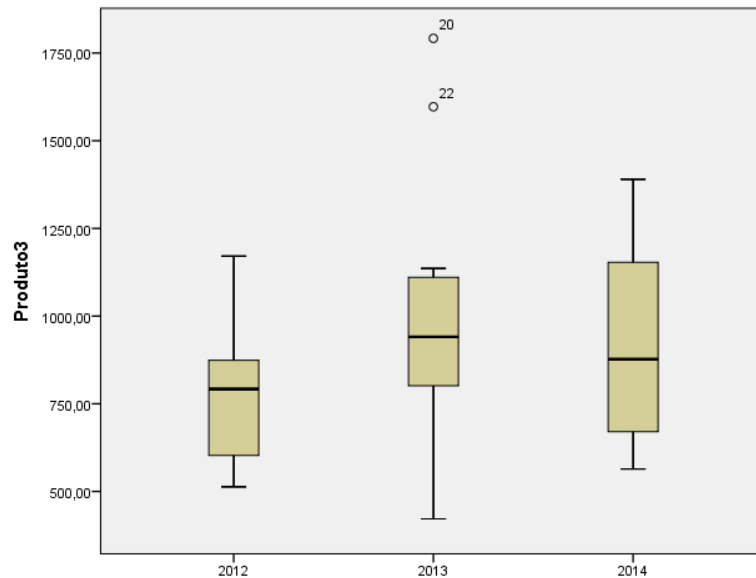


Figura 4.26 – BoxPlot anual após 1º tratamento de *outliers* do Produto 3

Fonte: Desenvolvido pelo autor

Por fim, na Figura 4.27, após o tratamento completo de todos os *outliers* detectados na amostra estudada, é possível observar oscilações das vendas desse produto durante os três anos analisados.

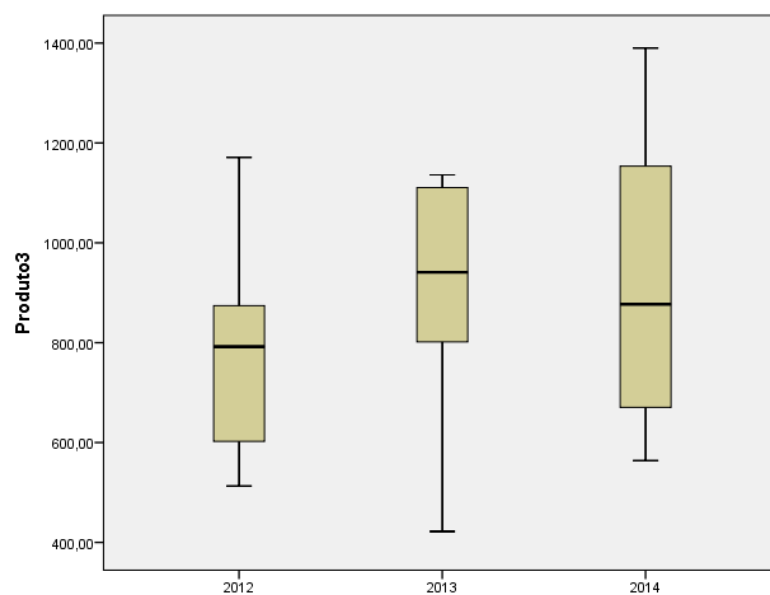


Figura 4.27 – BoxPlot anual após último tratamento de *outliers* do Produto 3

Fonte: Desenvolvido pelo autor



A Tabela 4.9 exibe as medidas descritivas do Produto 3, onde observa-se os parâmetros correspondentes a média, mediana, desvio padrão, valor máximo e mínimo presente na série, 1º quartil, 3º quartil e variância. A partir da Tabela 4.9 é possível verificar que o valor da média é superior ao valor da mediana, representando 859,4 e 857,5 respectivamente.

Tabela 4.9 – Medidas descritivas da série de dados do Produto 3

Parâmetros	Média	Mediana	Desvio Padrão	Mínimo	Máximo	1º Quartil	3º Quartil	Variância
<b>Valores</b>	859,4	857,5	234,5	422,0	1.390,0	671,8	1.035,0	54977,8

Fonte: Desenvolvido pelo autor

Nas Figuras 4.28 e 4.29, referentes à série temporal da demanda do Produto 3, observa-se estacionariedade dos dados, visto que há a presença de aleatoriedade próxima de 0.

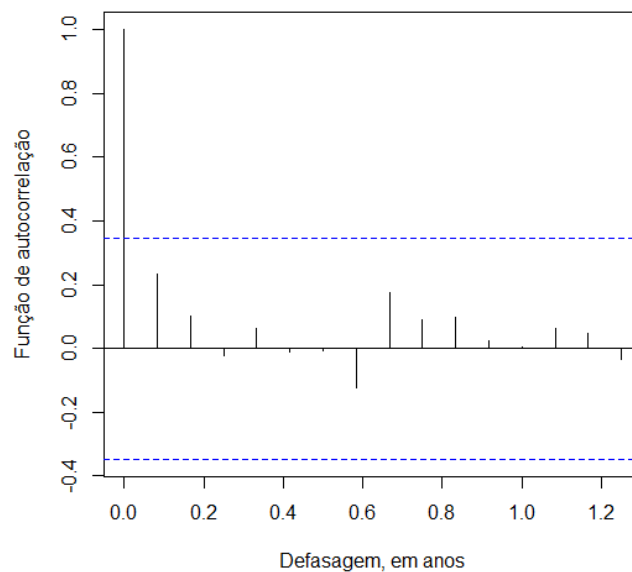


Figura 4.28 – Função de autocorrelação do Produto 3

Fonte: Desenvolvido pelo autor

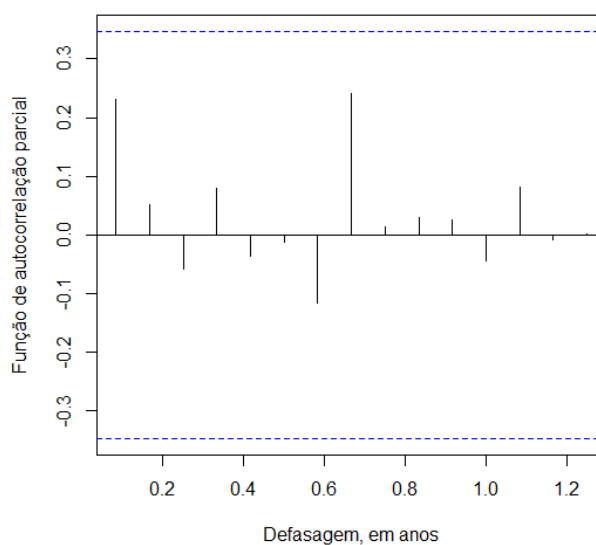


Figura 4.29 – Função de autocorrelação parcial do Produto 3

Fonte: Desenvolvido pelo autor

Na Figura 4.30, o histograma da série de dados apresenta um perfil de histograma multimodal, tendo suas barras frequências altas e baixas, alternativamente. O Produto 3 exibiu esse perfil de histograma devido a quantidade de dados incluídos na classe (barra) variar de classe para classe e apresentar tendência na maneira como os dados foram tratados.

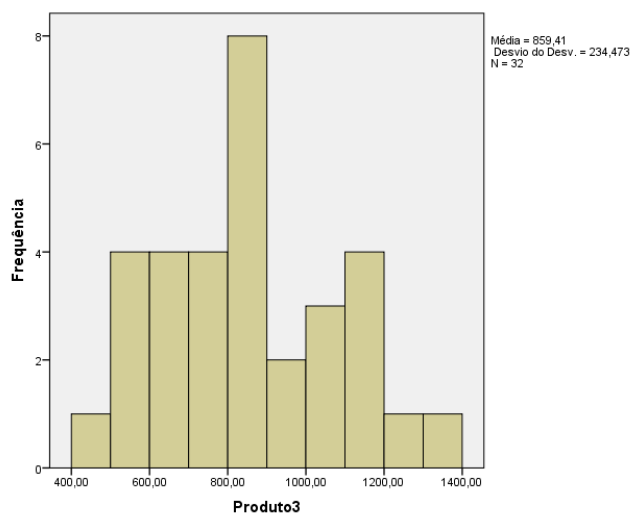


Figura 4.30 – Histograma do Produto 3

Fonte: Desenvolvido pelo autor

#### 4.6 Método ARIMA para o Produto 3

Conforme explicitado na subseção 3.8.1, realizou-se a modelagem de Box-Jenkins conforme as três etapas explicitadas na metodologia. Na etapa de identificação, realizou-se automaticamente a seleção dos parâmetros  $p$ ,  $d$  e  $q$  por meio do *software* R, da mesma forma que foi aplicado a todos os demais produtos.

A partir do correlograma dos resíduos referentes ao Produto 3 mostrado na Figura 4.31, verifica-se estacionariedade da série, devido a presença de aleatoriedade tanto no gráfico de resíduos padronizados quanto para o ACF dos resíduos durante os períodos analisados.

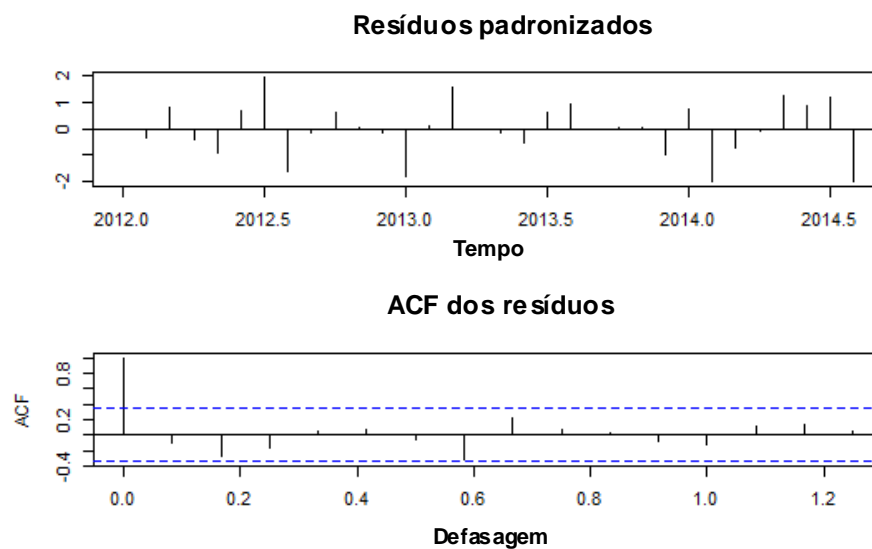


Figura 4.31 – Correlogramas de Padronização e ACF dos resíduos do Produto 3

Fonte: Desenvolvido pelo autor

Os parâmetros do modelo ótimo estimado pelo *software* R para série de dados do Produto 3 foi o modelo ARIMA autorregressivo diferenciado de primeira ordem, ou ARIMA (1,1,0)(0,0,0), conforme mostrado na Tabela 4.10.

Tabela 4.10 – Descrição do modelo ARIMA do Produto 3

Descrição do modelo	
Produto 3	Tipo do modelo ARIMA (1,1,0)(0,0,0)

Fonte: Desenvolvido pelo autor

Aplicou-se o teste de Ljung-Box para a etapa de verificação de autocorrelação dos resíduos da série, conforme mostrado na Tabela 4.11. O *p-value* encontrado no Produto 3 foi

correspondente a 0,5379, ou seja, indica que as autocorrelações residuais do modelo são estatisticamente não significativas. Dessa forma, os resíduos se comportam como ruído branco e estão corretamente especificados e adequados para a geração de previsões.

Tabela 4.11 – Teste de autocorrelação Ljung-Box dos resíduos do Produto 3

<b>Produto 3</b>	<b>Autocorrelação Ljung-Box</b>	
	<b>Estatísticas (<math>X^2</math>)</b>	<b>p-value</b>
	0,3795	0,5379

Fonte: Desenvolvido pelo autor

A Figura 4.32 apresenta os valores das autocorrelações apresentados entre os *lags* 1 até 10, onde é possível verificar que são estatisticamente não significativos e estão corretamente aplicados para a previsão.

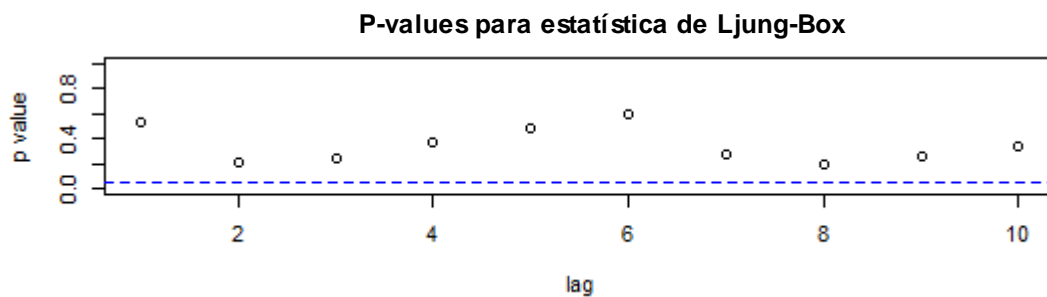


Figura 4.32 – *P-values* para teste Ljung-Box

Fonte: Desenvolvido pelo autor

A Figura 4.33 apresenta o QQ-Plot dos resíduos para o Produto 3. Os parâmetros do teste de normalidade Shapiro-Wilk apresentaram como resultado  $W = 0,9639$  e  $p\text{-value} = 0,349$ , ou seja, os resíduos apresentam distribuição normal, com nível de significância de 5%.

Tabela 4.12 – Teste de normalidade Shapiro-Wilk dos resíduos do Produto 3

<b>Produto 3</b>	<b>Teste de normalidade</b>	
	Shapiro-Wilk	
	<b>W</b>	<b>p-value</b>
	0,9639	0,349

Fonte: Desenvolvido pelo autor

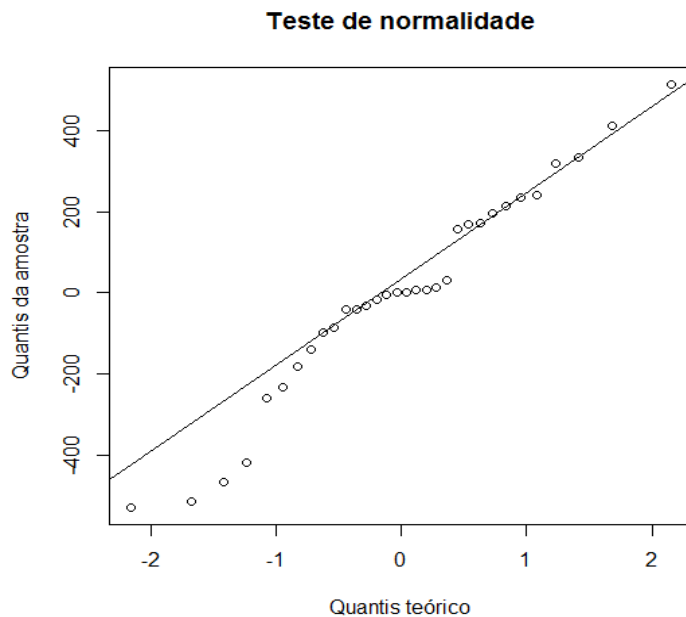


Figura 4.33 – QQ-Plot dos resíduos para o Produto 3

Fonte: Desenvolvido pelo autor

A Figura 4.34 ilustra o gráfico comparativo entre a série temporal real analisada e a série temporal predita pelo *software* R do Produto 3.

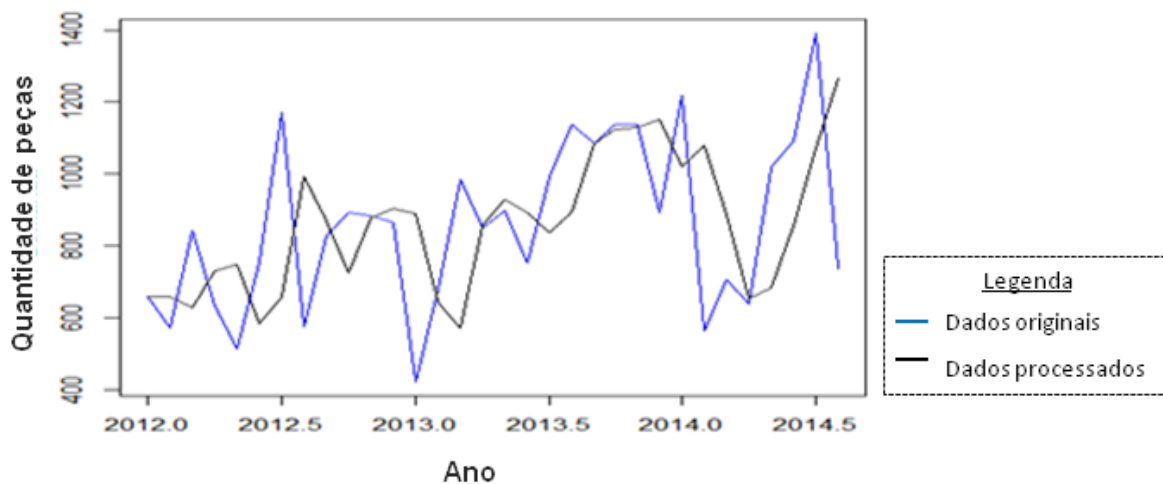


Figura 4.34 – Comparativo entre a demanda real e predita do Produto 3

Fonte: Desenvolvido pelo autor

Conforme a Figura 4.35, verifica-se que o resultado da previsão apresenta características semelhantes à previsão gerada pelo Produto 1, com pequenas variações nos primeiros períodos e posteriormente uma tendência constante nos demais meses preditos.

A partir dos resultados previstos pela metodologia de Box-Jenkins no *software* R, fez-se a validação de desempenho dos erros de previsão RMSE e MAPE e suas estatísticas de ajuste do modelo de previsão apresentaram um valor referente à Raiz do Erro Médio Quadrático (RMSE) correspondente a 563,2885 e o MAPE representando 28%.

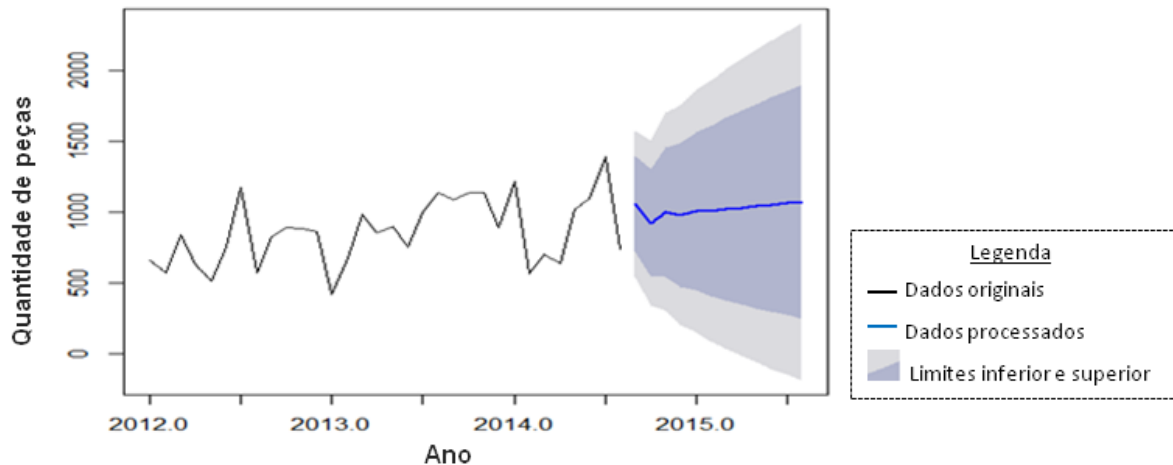


Figura 4.35 – Previsão método Box-Jenkins do Produto 3

Fonte: Desenvolvido pelo autor

#### 4.7 Redes Neurais

Conforme Rojas (1996), sistemas que empregam redes neurais artificiais apresentam uma estrutura parcialmente predeterminada, na qual seus parâmetros modificam as capacidades da rede, tornando necessário encontrar combinações que satisfaçam a solução dos problemas. Faz-se o ajuste dos parâmetros por meio de um algoritmo de aprendizado, o qual não é definido por uma programação explícita, e sim por um método de adaptação automática.

No estudo em questão utilizou-se o para o treinamento da rede neural artificial o algoritmo de aprendizado *Backpropagation*, que corresponde a um método para treinamento do tipo *perceptron* de múltiplas camadas (HAYKIN, 2001).

Na implementação de redes neurais, normalmente utiliza-se a separação dos dados em dois conjuntos: conjunto de treinamento e conjunto de teste. No conjunto de treinamento, tem-se como objetivo principal o treinamento da rede e ajuste dos seus parâmetros por meio do cálculo dos pesos e *bias*. Além disso, é necessário conter uma amostra de dados estatisticamente significativa, de forma a garantir valores representativos para o problema a ser estudado.

Já o conjunto de teste busca analisar a capacidade de generalização da rede sob condições reais de utilização. Nesta etapa não é realizado o ajuste dos seus parâmetros e o erro sobre o conjunto amostral é monitorado.

A base de dados da pesquisa é constituída pelo histórico dos últimos 03 anos da demanda de vendas da empresa analisada, o que corresponde a 36 registros mensais de vendas para cada um dos três produtos selecionados no estudo. Deste total, durante o desenvolvimento da rede neural destinou-se 70% da série temporal para treinamento, totalizando 25 padrões observados. Para a etapa de teste, foram utilizados 20% dos dados com 07 padrões analisados e os 10% restantes, referente aos valores comparados mediante a previsão gerada pela rede neural MLP desenvolvida.

A etapa seguinte foi a normalização estatística dos dados, utilizando o método de normalização sigmoïdal, conforme Equação 27, no qual busca-se colocar uma mesma faixa de valores aos dados, evitando assim que uma dimensão sobreponha a outra e evitando que o aprendizado fique estagnado.

A determinação dos parâmetros principais da rede foi empírica em todos os três produtos implementados.

No modelo de rede neural MLP desenvolvido no *software* Matlab, após várias simulações alterando o número de neurônios da camada de entrada e camada intermediária, além de alterações na taxa de aprendizado e no número de iterações realizadas, buscou-se obter o modelo que apresentasse melhor valor correspondente a Raiz do Erro Quadrático Médio, também conhecido como RMSE. Em todas as análises simuladas obteve-se melhores resultados os modelos com 04 neurônios na camada de entrada, os quais foram utilizados nos três produtos estudados. Além disso, na rede modelada utilizaram-se diferentes faixas de neurônios intermediários entre 04 a 25, com taxas de aprendizado entre 0,3 a 1,0 e número de interações entre 1.000 a 10.000 e posteriormente 100.000.

Para todos os três produtos analisados, utilizou-se a função de aprendizagem baseado no algoritmo de retropropagação, também conhecido como *Backpropagation*. Sua utilização tem papel essencial na identificação das diferenças mínimas entre as saídas desejadas e saídas obtidas pela rede neural, a partir de um erro mínimo, de maneira a ajustar os pesos entre as camadas por meio da retropropagação dos erros encontrados a cada iteração.

Após várias simulações, os resultados gráficos para treinamento, teste e previsão do Produto 1 são mostrados na Figura 4.36.

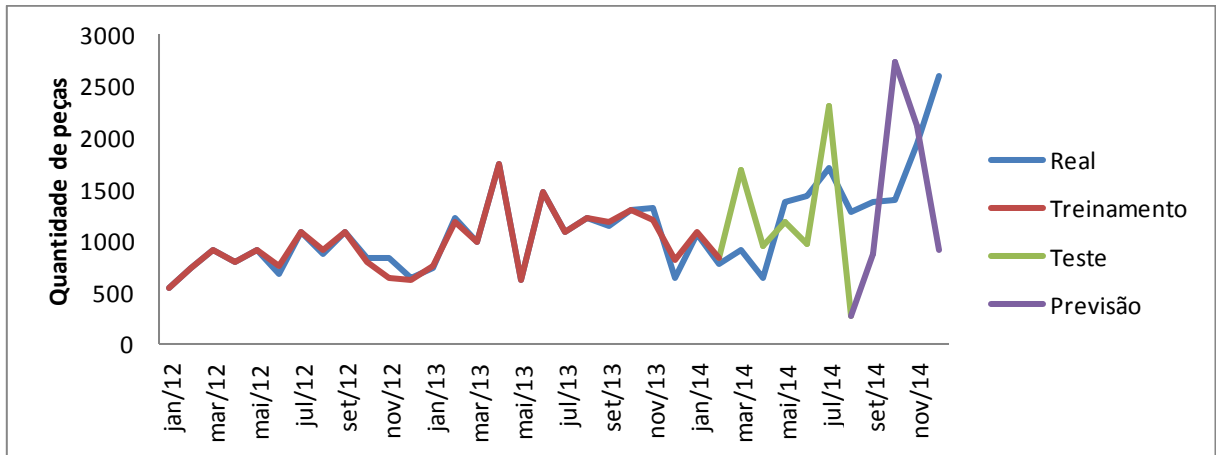


Figura 4.36 – Treinamento, teste e previsão da Rede Neural MLP - Produto 1

Fonte: Desenvolvido pelo autor

Para a série temporal correspondente ao Produto 1, os parâmetros ótimos que garantiram melhores resultados apresentaram na programação da rede neural MLP os seguintes parâmetros:

- 04 neurônios da camada de entrada;
- 04 neurônios intermediários;
- Taxa de aprendizagem: 0,7;
- Número de ciclos para parada do treinamento: 100000 iterações.

O parâmetro de precisão RMSE, encontrado na rede neural MLP desenvolvida apresentou um valor correspondente a 1117,852 e o MAPE corresponde a 52%.

No gráfico da Figura 4.36, na etapa de treinamento tem-se como dados gerados pela rede a série vermelha, enquanto a azul apresenta os valores reais da empresa X. Para o teste, os resultados alcançados são representados pela linha verde, e por fim, o resultado da previsão em roxo. As saídas gráficas do software Matlab geradas, referentes aos três produtos, foram explicitadas separadamente no Apêndice E para cada uma das etapas do método.

Dessa forma, ainda referente ao Produto 1, é possível constatar no resultado do gráfico de treinamento que o comportamento da rede MLP apresenta valores praticamente equivalentes aos valores reais da série. Durante as etapas de teste e previsão, observa-se que a rede consegue replicar a tendência dos dados reais da série, com exceção de alguns meses em que apresentou diferenças comparativas. Em alguns destes meses, apesar de apresentar tendências opostas com relação ao mesmo período e ano, a previsão gerada muitas vezes apresentou a mesma tendência com relação aos períodos correspondentes dos anos anteriores.



Para o Produto 2, os resultados gráficos para treinamento, teste e previsão seguem conforme mostrado na Figura 4.37.

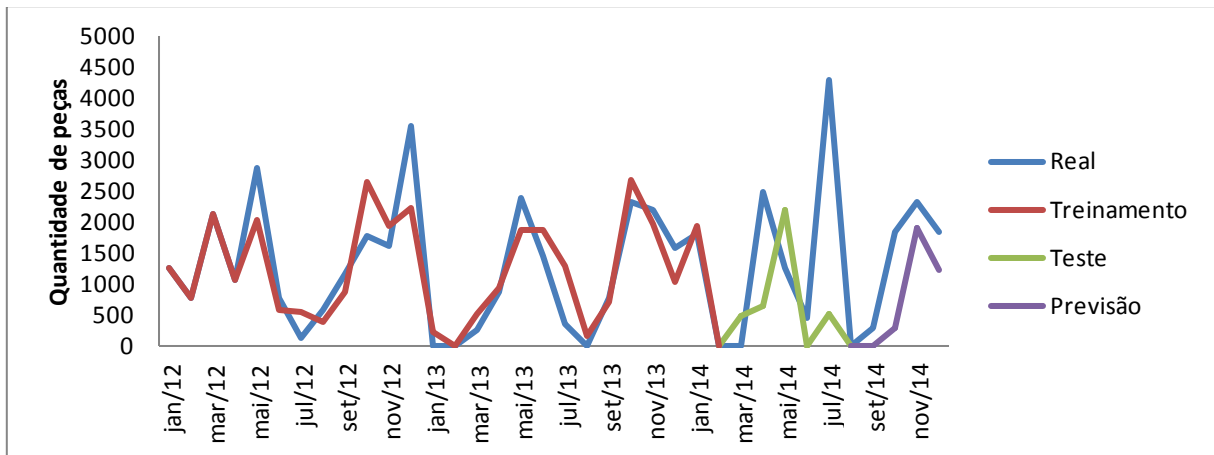


Figura 4.37 – Treinamento, teste e previsão da Rede Neural MLP - Produto 2

Fonte: Desenvolvido pelo autor

A partir da Figura 4.37, visualmente é possível observar que, apesar de na etapa de treinamento os dados processados não estarem totalmente equiparados com os valores reais da série, como obtido durante o modelo observado no Produto 1, nota-se aparentemente, que as oscilações dos dados processados seguem uma tendência de proximidade e equivalência quando associadas aos valores reais da série. De maneira geral, a partir do gráfico de previsão do Produto 2, observa-se que a série acompanha a evolução temporal da série original.

Para a série temporal em questão, referente ao Produto 2, tem-se como parâmetros ótimos apresentados na programação da rede MLP os seguintes parâmetros:

- 04 neurônios da camada de entrada;
- 10 neurônios intermediários;
- Taxa de aprendizagem: 0,7;
- Número de ciclos para parada do treinamento: 10000 iterações.

Neste modelo, o parâmetro de precisão ou RMSE encontrado na rede MLP desenvolvida corresponde a 857,957, enquanto o MAPE é igual a 58%.

O gráfico da Figura 4.38 representa a série temporal do Produto 3, onde são apresentados o comparativo entre os valores reais da demanda (em azul) e os valores processados em cada uma das etapas desenvolvidas para a rede neural MLP, sendo elas, treinamento, teste e previsão.

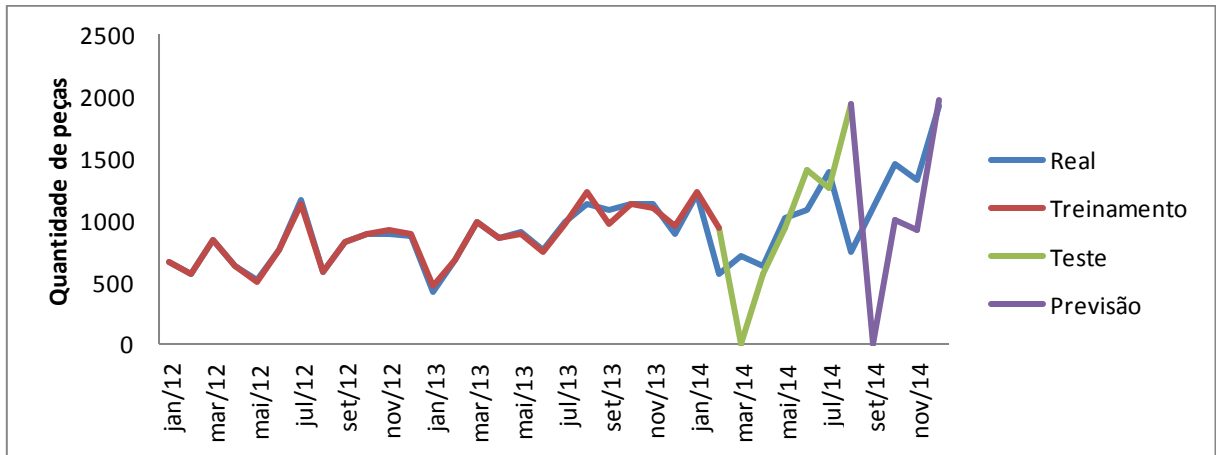


Figura 4.38 – Treinamento, teste e previsão da Rede Neural MLP - Produto 3

Fonte: Desenvolvido pelo autor

Percebe-se, pela análise gráfica, que o comportamento da etapa de treinamento segue basicamente o mesmo traçado dos dados reais da série. Porém, na etapa de teste, os dados processados pela rede apresentam diferentes padrões de tendência em alguns períodos comparados. Observa-se também que, da mesma forma que o Produto 1, alguns meses a série processada pela rede consegue gerar padrões semelhantes aos apresentados durante os meses anteriores equivalentes. Durante a etapa de previsão do Produto 3, observa-se uma defasagem dos valores preditos, porém apresentando uma aproximação do comportamento da série com relação aos dados originais, onde a mesma começa a apresentar picos e vales com uma tendência a equiparar ou atingir os padrões reais.

Os melhores parâmetros utilizados na modelagem da rede para a série de dados referente ao Produto 3 apresenta os seguinte valores:

- 04 neurônios da camada de entrada;
- 04 neurônios intermediários;
- Taxa de aprendizagem: 0,7;
- Número de ciclos para parada do treinamento: 100000 iterações.

O parâmetro RMSE encontrado no modelo correspondente ao Produto 3 da rede MLP desenvolvida corresponde a 630,9923 e o MAPE equivale a 41%.

Devido às séries de dados para todos os três produtos analisados serem avaliadas durante um período consideravelmente pequeno para uma análise de redes neurais, sendo este de 03 anos ou 36 registros, a rede neural conseguiu atingir um intervalo de alcance da previsão correspondente a apenas 04 períodos.

#### 4.8 Análise comparativa dos modelos

Conforme apresentado na seção 1.3, um dos objetivos deste trabalho foi realizar o comparativo entre a metodologia estatística de Box-Jenkins para previsão de demanda, com técnicas de inteligência artificial, particularmente utilizando a rede neural MLP. Após inúmeras observações, modelagens, simulações e previsões obtidas em ambas às técnicas, foi possível realizar o comparativo entre os métodos propostos.

Na análise estatística, identificaram-se padrões no comportamento referente à sazonalidade e oscilação da série. Tanto no modelo estatístico de Box-Jenkins quanto na rede neural desenvolvida, foi possível observar resultados satisfatórios de previsão. Quanto ao comportamento das novas séries, foi possível identificar que a rede neural MLP apresenta um melhor padrão perante aos dados reais especificamente no Produto 2.

A Tabela 4.13 apresenta os resultados das medidas de erros para cada modelo proposto. A validação de ambos os métodos estudados foi calculada a partir do comparativo entre os 04 últimos períodos da série analisada, referente aos meses de setembro até dezembro de 2014, com os valores processados pelos modelos.

Tabela 4.13 – Resumo comparativo de erros dos modelos estudados

Métodos				
Produtos	ARIMA (1,1,0)		Rede Neural MLP	
	MAPE	RMSE	MAPE	RMSE
Produto 1	20%	572,1629	52%	1117,852
Produto 3	28%	563,2885	41%	630,9923
Produto	ARIMA (0,0,0)		Rede Neural MLP	
	MAPE	RMSE	MAPE	RMSE
Produto 2	107%	817,994	58%	857,957

Fonte: Desenvolvido pelo autor

Além disso, conforme a Tabela 4.13 é possível observar que para os produtos 1 e 3 estudados, o ARIMA que melhor explicou a demanda para cada uma das séries foi o ARIMA autorregressivo diferenciado de primeira ordem, ou seja, ARIMA (1,1,0). É interessante observar que a metodologia Box-Jenkins é superior para estes produtos, apresentando melhor desempenho tanto nos resultados do MAPE quanto no RMSE. Por outro lado, o Produto 2 conseguiu alcançar melhores resultados a partir da aplicação da Rede Neural MLP, mostrando vantagens não apenas nas medidas de comparação dos métodos, mas também na similaridade e facilidade de captação do padrão de comportamento da série temporal processada, conforme mostrado na Figura 4.37. Além disso, no Produto 2, o ARIMA (0,0,0) foi o que melhor explicou a série de dados, classificado como modelo de ruído branco, pois não apresenta componente autorregressivo (AR), não envolve nenhuma diferenciação e também não apresenta nenhuma componente de média móvel.

## 5. CONCLUSÃO

Ao longo deste trabalho, buscou-se descrever as principais técnicas para cálculos de previsão de séries temporais, sendo realizado o comparativo entre as técnicas estatísticas e de inteligência artificial para o problema de previsão de demanda dos principais produtos vendidos em uma indústria do ramo metal mecânico.

Constatou-se, por meio da revisão bibliográfica, a existência de várias aplicações tanto utilizando técnicas estatísticas quanto de Redes Neurais, em áreas distintas nos últimos anos. Dessa forma, na presente pesquisa buscou-se apresentar uma abordagem e aplicação voltada para o cálculo de previsões de demanda no setor industrial.

O objetivo principal foi realizar o cálculo de previsão de demanda a partir do desenvolvimento da metodologia estatística de Box-Jenkins e modelagem de uma Rede Neural MLP, a fim de determinar qual dos métodos estudados é mais satisfatório para a determinação de previsões e conseqüentemente para as tomadas de decisão da empresa. Para se validar o desempenho de ambos os métodos propostos realizou-se as análises dos erros de previsão RMSE e MAPE a partir do comparativo dos dados gerados por ambos os modelos e pelos dados históricos conhecidos durante os 04 últimos períodos das séries temporais de cada um dos produtos.

Existem muitas decisões baseadas em processos criados com o objetivo de encontrar soluções para determinar qual o melhor modelo para previsão de demanda, porém a interpretação dos resultados somente pode se dar a partir das condições específicas deste estudo. Dessa forma, inicialmente, observou-se que os resultados nos dois métodos desenvolvidos foram satisfatórios em termos de predição. No entanto, devido à particularidade nas características dos dados de cada uma das séries foi possível observar quais dos métodos melhor se adéquam aos padrões das séries.

Os resultados indicaram que o desempenho das previsões utilizando o método estatístico de Box-Jenkins nos Produtos 1 e 3 foram superiores à aplicação dos modelos de rede neural MLP, visto que além de apresentar uma boa margem de aceitação para as medidas de erros de previsão com relação ao método de redes neurais MLP, também observou-se uma forte relação entre as previsões reais e as processadas pelo método de Box-Jenkins.

As séries temporais dos Produtos 1 e 3 apresentam características semelhantes quanto ao tratamento dos dados e transformações adotadas para eliminação de tendências e *outliers*,

as quais contribuíram na maioria das vezes para a melhora significativa da qualidade das previsões geradas. No entanto, a série temporal do Produto 2 apresentou padrões particulares em que não foi possível adotar tais transformações. Dessa forma, estas características acabaram dificultando os resultados utilizando o método estatístico adotado e trazendo boas vantagens na utilização das redes neurais MLP.

Apesar dos resultados satisfatórios da neural MLP para o Produto 2, há de se ressaltar que a mesma não proporciona conhecimento sobre os dados, visto que não exige a identificação de padrões de comportamento da série temporal, diferente do modelo estatístico. A rede neural é um sistema adaptativo que detém as relações funcionais entre os dados inseridos por meio dos processos de treinamento e testes. Além disso, o desenvolvimento de uma rede neural apresenta dificuldades e limitações, como por exemplo, identificação do número de camadas de entrada, da taxa de aprendizado, definição do número de camadas intermediárias ou ocultas, do número de iterações utilizadas para treinamento, escolha da função de ativação e de toda uma programação do algoritmo.

Outra dificuldade que o método de redes neurais MLP abordou foi a identificação dos principais parâmetros da rede de maneira empírica, visto que, não existe uma metodologia específica para que se possa aplicar um algoritmo de treinamento que encontre um conjunto ideal de pesos.

Quanto às ferramentas de computação, os métodos utilizando redes neurais não estão prontamente disponíveis em *softwares* de fácil utilização e principalmente a um custo acessível. Isso também foi uma dificuldade inicial para a aplicação das previsões utilizando redes neurais, visto que foi necessário realizar testes em outros *softwares* que a princípio pareciam oferecer maior facilidade de desenvolvimento do método, como por exemplo, o *software* SPSS. Porém, o mesmo não garantia recursos suficientes para o entendimento de sua aplicação e resultados.

Visto que, em dois dos três produtos estudados, os resultados estatísticos foram superiores do que o método de redes neurais devido às particularidades dos dados de entrada, uma sugestão para trabalhos futuros seria a incorporação de dados qualitativos as séries temporais. Dessa forma, será possível desenvolver sistemas híbridos, como por exemplo, uma aplicação neuro-fuzzy, onde apresenta características quantitativas e qualitativas em destaque, permitindo aprofundar ainda mais nos padrões das séries e consequentemente garantir uma melhor *performance* nas previsões. Outras opções de sistemas híbridos que podem ser utilizados futuramente seriam sistemas de redes neurais incorporadas a técnicas estatísticas,

de maneira a utilizar a aplicação dos parâmetros da rede neural justificadas nos resultados estatísticos encontrados e garantir maior fundamento no desenvolvimento da rede neural, sem que haja testes empíricos envolvidos.

## REFERÊNCIAS

- APPOLINÁRIO, F. **Metodologia da ciência – filosofia e prática da pesquisa**. São Paulo: Editora Pioneira Thomson Learning, 2006.
- ARENALES, M.; ARMENTANO, V.; MORABITO, R.; YANASSE, H. **Pesquisa Operacional**. Rio de Janeiro: Editora Câmpus - Abepro: 8<sup>a</sup> ed. Rio de Janeiro. 2007.
- BACCI, L. A. **Combinação de Métodos de Séries Temporais para Previsão da Demanda de Café no Brasil**. Dissertação (Mestrado). Universidade Federal de Itajubá, Itajubá, MG, 2007.
- BALLOU, R. H. **Gerenciamento da cadeia de suprimentos/logística empresarial**. 5<sup>a</sup> ed. Porto Alegre: Bookman, 2006.
- BARBOSA, C. M.; CHAVES, C. A. **Um estudo sobre o gerenciamento da demanda nos sistemas de planejamento e controle da engenharia de produção**. VIII Congresso Nacional de Excelência em Gestão. Rio de Janeiro, 08 e 09 de junho de 2012.
- BEALE, R.; JACKSON, T. **Neural computing: an introduction**. Bristol: Adam Hilger, 1990.
- BOCANEGRA, C. W. E. **Procedimento para tornar mais efetivo o uso das redes neurais artificiais em planejamento de transportes**. Tese de mestrado apresentada a Escola de Engenharia de São Carlos, 2002.
- BOX, G. E. P.; JENKINS, G. M. **Time series analysis forecasting and control**. San Francisco: Holden-Day - Edição revisada, 1976.
- BRAGA, A.; CARVALHO, A.; LUDERMIR, T. **Redes Neurais Artificiais: teoria e aplicações**. 1. Ed. Rio de Janeiro: Livros Técnicos e Científicos, 2000.
- BRANCO, S. T. **Especificação de uma ferramenta de previsão de demanda para equipamentos de infra-estrutura de telecomunicações**. Dissertação (Mestrado). PUC Pontifícia Universidade Católica, Curitiba, PR, 2008.
- BRUNI, A. L. **SPSS Guia Prático para pesquisadores**. São Paulo: Editora Atlas S.A., 2012.
- CALDWELL, J. G. **The Box-Jenkins Forecasting Technique**. 2006. Disponível em: <[http://www.foundationwebsite.org/BoxJenkins.htm#\\_Toc145736381](http://www.foundationwebsite.org/BoxJenkins.htm#_Toc145736381)>.



- CARBONNEAU, R.; LAFRAMBOISE, K.; VAHIDOV, R. **Application of machine learning techniques for supply chain demand forecasting**. *European Journal of Operational Research*, v.184, n.3, p.1140-1154, 2008.
- CASTRO, M. C. F. **Predição Não-Linear de Séries Temporais Usando Redes Neurais RBF por Decomposição em Componentes Principais**. Tese (Doutorado). Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas. Campinas, 2001.
- CAVALCANTI-NETTO, M.A. **A previsão com a metodologia de Box-Jenkins**. 2007. Disponível em: <[www.ie.ufrj.br/download/APrevisaoComMetodologiadeBox-Jenkins.pdf](http://www.ie.ufrj.br/download/APrevisaoComMetodologiadeBox-Jenkins.pdf)>
- CONNOR, J. T., MARTIN, R. D. e ATLAS, L.E. **Recurrent neural networks and robust time series prediction**. *IEEE Transactions on Neural Networks*, Vol. 5, NO. 2, 1994.
- DAVIS, Mark M.; AQUILANO, Nicholas J.; CHASE, Richard B. **Fundamentos da Administração da Produção**. 3. ed. Porto Alegre: Bookman, 2001.
- FAVA, V. L. **Manual de econometria**. In: VASCONCELOS, M. A. S.; ALVES, D. São Paulo: Editora Atlas, 2000.
- FIGUEREDO, C. J. **Previsão de Séries Temporais Utilizando a Metodologia Box-Jenkins e Redes Neurais para Inicialização de Planejamento e Controle de Produção**. Dissertação (Mestrado) – UFPR Universidade Federal do Paraná, Curitiba, PR, 2008.
- FLORES, J.H.F. **Comparação de modelos MLP/RNA e modelos Box-Jenkins em séries temporais não lineares**. Dissertação (Mestrado). Programa de Pós-graduação em Engenharia de Produção, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2009.
- GEORGE, E.P. BOX; GWILYM, M. JENKINS. **Time Series Analysis: Forecasting and Control**. San Francisco: Holden Day, 1976.
- GEORGOFF, D. M.; MURDICK, R. G. **Manager's Guide to Forecasting**. *Harvard Business Review*, p.110–120, 1986.
- GRAUPE, D. **Principles of Artificial Neural Networks, Advanced Series on Circuits and Systems** – Vol. 6, World Scientific Publishing, 2nd Edition, 2007.
- GUJARATI, D. N., **Econometria Básica**. 4º Ed. Rio de Janeiro, Elsevier, 2006.
- HAYKIN, S. **Redes Neurais. Princípios e prática**. Porto Alegre, RS: Bookman, 2001.
- HYNDMAN, R. J.; KHANDAKAR, Y. **Automatic time series forecasting: The forecast Package for R**. *Journal of Statistical Software*, Volume 27, Issue 3, 2008.

LABOISSIERE, L. A.; FERNANDES, R. A. S.; LAGE, G. G. **Maximum and minimum stock price forecasting of Brazilian power distribution companies based on artificial neural networks**. Department of Electrical and Computer Engineering, University of São Carlos, São Carlos, SP: Elsevier, 2015.

LEVINE, D. M.; BERENSON, M. L. STEPHAN, D. **Estatística: teoria e aplicações usando Microsoft Excel em português**. Rio de Janeiro. Livros Técnicos e Científicos Editora, 811p. 2000.

LOPES, R. **Previsão de Autopeças: Estudo de Caso em uma Concessionária de Veículos**. Dissertação de Mestrado. UFSC, Florianópolis, 2002.

LUSTOSA, L.J.; MESQUITA, M.A.; QUELHAS, O.L.G.; OLIVEIRA, R.J.; **Planejamento e Controle da Produção**. Rio de Janeiro: Elsevier, 2008.

MAKRIDAKIS, S.; WHEELWRIGHT, S. C.; HYNDMAN, R.J. **Forecasting: Methods and Applications**. John Wiley & Sons, New York, 3<sup>a</sup> Ed. 1998.

MANTOVANI, Almir. **Procedimentos e modelos para previsão de vendas e determinação de quotas na indústria calçadista: Proposta e estudo de caso**. Dissertação (Mestrado). Programa de Pós-graduação em Engenharia de Produção, Universidade Federal de São Carlos. São Carlos, 2011.

MCCABE, B. P. M. e TREMAYNE, A. R. **Testing a time series for difference stationarity**. The Annals of Statistics 23(3), 1015-1028, 1995.

MINE, O. M. **Previsão de demanda de autopeças com Redes Neurais**. Dissertação (Mestrado). Programa de Pós-graduação em Informática, Universidade Federal do Espírito Santo. Vitória – ES, 2010.

MOHAMMADI, R.; FATEMI, S.M.T. e ZEINALI, F. **A new hybrid evolutionary based RBF networks method for forecasting time series: a case study of forecasting emergency supply demand time series**. Engineering Applications of Artificial Intelligence 36:204 – 214, 2014.

MONTGOMERY, D.; JOHNSON, L.; GARDINER, J. **Forecasting and Time Series Analysis**. New York: McGraw-Hill, 1990.

MORAES, C.F.; FERREIRA, J.R.; BALESTRASSI, P.P. **Análise crítica da aplicação de métodos estatísticos em processos definidos por dados que não apresentam distribuição normal**. Revista GEPROS (Gestão da Produção, Operações e Sistemas), UNESP – SP, 2006.

MORETTIN, P. A.; TOLOI, C. M. C. **Previsão de séries temporais**. 2ª ed. Atual Editora, São Paulo, SP, 1986.

MUELLER, A. **Uma aplicação de redes neurais artificiais na previsão do mercado acionário**. Dissertação (Mestrado em Engenharia de Produção) – Programa de Pós-graduação em Engenharia de Produção, UFSC, Florianópolis, 1996.

NARDI, Ivana R.; LIMA, Almir R.; AMORIN, Ariuska K. B.; NERY, Valéria D. **Análise de séries temporais na operação de sistema de tratamento de águas residuárias de abatedouro de frango**. VIII SIBESA, 17 a 22 de setembro. Fortaleza – CE, 2006.

NASCIMENTO, N. O.; NAGHETTINI, M.; HELLER, L.; VON SPERLING, M. **Investigação científica em engenharia sanitária e ambiental. Parte 3: Análise estatística de dados e de modelos**. Engenharia Sanitária e Ambiental (ABES), Ano 1, Vol. 1, No. 4, p.152-168, Out./Dez. 1996.

NETO, R. M.; PUREZA, V. *et al.* **Metodologia de pesquisa em Engenharia de Produção e Gestão de Operações**. Paulo Augusto Cauchick Miguel (organizador). 2.ed. – Rio de Janeiro: Elsevier: ABEPRO, 2012.

PACHECO, A.G.F. **Estudo da influência de variáveis meteorológicas no aparecimento de casos graves de leptospirose em Salvador – BA via modelos de séries temporais**. Ministério da Saúde, Fundação Oswaldo Cruz, Escola Nacional de Saúde Pública, 2001.

PASQUOTTO, J.L.D. **Previsão de Séries Temporais no Varejo Brasileiro: Uma Investigação Comparativa das Redes Neurais Recorrentes de Elman**. Dissertação (Mestrado). Universidade de São Paulo, São Paulo, SP, Brasil, 2010.

PELLEGRINI, Fernando R.; FOGLIATTO, Flávio, S., **Passos para Implantação de Sistemas de Previsão de Demanda – Técnicas e Estudo de Caso**. Revista Produção, v.11 n.1, novembro, 2001.

ROJAS, R.; **Neural Networks: A Systematic Introduction**. Springer Verlag, Berlin 1996.

RUSSELL, S.; NORVIG, P. **Inteligência Artificial**. Editora Campus, 2004.

SANTOS, P. M. **Seleção de modelos de previsão baseada em informações de desempenho**. Dissertação (Mestrado). Programa de Pós-Graduação em Ciência da Computação, Universidade Federal de Pernambuco – UFPE, Recife, 2006.

SARTORIS, A. **Estatística e introdução à econometria**. São Paulo: Saraiva, 2003.

SHAPIRO, S. S.; WILK, M. B. **An analysis of variance test for normality (complete samples)**. In: *Biometrika*, 52: 3-4, 591–611, 1965.

SHARDA, R. **Neural Networks for the MS/OR Analyst: An Application Bibliography**. *Interfaces*, v. 24, p. 116 - 130, 1994.

SILVA, E.; OLIVEIRA, A. C. **Dicas para configuração de Redes Neurais**. Núcleo de Computação Eletrônica, UFRJ, Rio de Janeiro, RJ, 2002.

SILVA, Ivan Nunes da, SPATTI, D. H.; FLAUZINHO, R. A. **Redes Neurais Artificiais para engenharia e ciências aplicadas: Curso Prático**. São Paulo: Artliber, 2010.

SLACK, Nigel; CHAMBERS, Stuart; JOHNSTON, Robert. **Administração da Produção**. 3.ed. São Paulo, Atlas, 2009.

SOARES, A. da S. **Predição de Séries Temporais Econômicas por Meio de Redes Neurais Artificiais e Transformada Wavelet: Combinando Modelo Técnico e Fundamentalista**. Dissertação (Mestrado). Programa de Pós-Graduação em Engenharia Elétrica, Universidade de São Paulo – São Carlos, 2008.

SOUSA, A.P. **Análise comparativa de métodos de previsão de séries temporais através de modelos estatísticos e rede neural artificial**. Programa de Mestrado em Engenharia de Produção e Sistemas, Pontifícia Universidade Católica de Goiás (PUC-GOÍÁS), 2012.

SOUZA, F. M. **Modelos Box & Jenkins aplicados a previsão de demanda de leitos hospitalares**. Programa de Pós-Graduação em Estatística e Modelagem Quantitativa, Universidade Federal de Santa Maria (UFSM), RS, 2006.

STAUDT, Francielly H., **Estudo de métodos de previsão de demanda com incorporação de julgamentos**. Dissertação (Mestrado). Programa de Pós-graduação em Engenharia de Produção, Universidade Federal de Santa Catarina. Santa Catarina, 2011.

SZOPLIK, Jolanta. **Forecasting of natural gas consumption with artificial neural networks**. Department of Chemical Engineering, West Pomeranian University of Technology, Energy 85, Al. Piastow 42, PL-71-065, Szczecin, Poland, 2015.

TAFNER, M. A.; XEREZ, M.; FILHO, I. R. **Redes neurais artificiais: introdução e princípios de neurocomputação**. Blumenau: EKO, 1996.

TUBINO, Dalvio Ferrari. **Planejamento e Controle da Produção**. 2.ed. São Paulo, Atlas, 2009.

TURRIONI, J. B.; MELLO, C.H.P. **Metodologia de pesquisa em Engenharia de Produção**. Programa de Pós-graduação em Engenharia de Produção. Curso de Especialização em Qualidade & Produtividade. Universidade Federal de Itajubá – UNIFEI, 2012.

VIGLIONI, G. M. C. **Comparação entre Redes Neurais e Técnicas Clássicas para Previsão de Demanda de Transporte Ferroviário**. XX Congresso da Associação Nacional de Pesquisa e Ensino de Transportes (ANPET), 2006.

VILLAMAGNA, M. R. **Seleção de Modelos de Séries Temporais e Redes Neurais Artificiais na Previsão de Consumo e Demanda de Energia Elétrica**. Programa de Pós-graduação em Estatística e Experimentação Agropecuária, Universidade Federal de Lavras, UFLA, 2013.

WERNER, L.; RIBEIRO, J. L. D. **Previsão de Demanda: Uma aplicação dos modelos Box-Jenkins**. *Gestão & Produção*, v10, n.1, p,47-60, 2003.

WERBOS, P.J. **Generalization of Backpropagation With Application to a Recurrent Gas Market Model**. In: *Neural Networks*, n. 1, pp. 339 - 356, 1988.

ZHANG, G.P. **Time series forecasting using a hybrid ARIMA and neural network model**. Department of Management, J. Mack Robinson College of Business, Georgia State University, University Plaza, Atlanta, GA 30303, USA, 2001.

ZOUCAS, F.A.M. **Aplicação de redes neurais para previsão de séries temporais em segmentos industriais**. Programa de Mestrado de Engenharia Mecânica com ênfase em Produção. Centro Universitário da FEI, 2009.

## APÊNDICE A – DEMANDA DE PRODUTOS

Produto 1				Produto 2			Produto 3			
Período	Demanda real	Demanda real após tratamento outliers	Demanda normalizada após tratamento outliers	Período	Demanda real	Demanda normalizada	Período	Demanda real	Demanda real após tratamento outliers	Demanda normalizada após tratamento outliers
jan/12	539	539	0,21	jan/12	1250	0,49	jan/12	659	659	0,30
fev/12	729	729	0,29	fev/12	770	0,38	fev/12	571	571	0,24
mar/12	910	910	0,39	mar/12	2140	0,69	mar/12	842	842	0,43
abr/12	795	795	0,33	abr/12	1050	0,44	abr/12	632	632	0,28
mai/12	904	904	0,38	mai/12	2890	0,81	mai/12	513	513	0,21
jun/12	668	668	0,27	jun/12	780	0,38	jun/12	756	756	0,37
jul/12	1076	1076	0,48	jul/12	130	0,25	jul/12	1171	1171	0,69
ago/12	1436	871 (1)	0,37	ago/12	580	0,34	ago/12	573	573	0,24
set/12	1089	1089	0,49	set/12	1150	0,47	set/12	1573	828 (1)	0,42
out/12	833	833	0,35	out/12	1760	0,61	out/12	893	893	0,47
nov/12	830	830	0,34	nov/12	1620	0,58	nov/12	884	884	0,47
dez/12	647	647	0,26	dez/12	3570	0,89	dez/12	864	864	0,45
jan/13	731	731	0,29	jan/13	0	0,23	jan/13	422	422	0,16
fev/13	1215	1215	0,56	fev/13	0	0,23	fev/13	676	676	0,31
mar/13	995	995	0,43	mar/13	260	0,28	mar/13	984	984	0,55
abr/13	1738	1738	0,81	abr/13	880	0,41	abr/13	851	851	0,44
mai/13	615	615	0,24	mai/13	2400	0,74	mai/13	898	898	0,48
jun/13	1471	1471	0,70	jun/13	1450	0,54	jun/13	752	752	0,36
jul/13	1092	1092	0,49	jul/13	360	0,30	jul/13	992	992	0,55
ago/13	1223	1223	0,57	ago/13	0	0,23	ago/13	1792	1136 (2)	0,67
set/13	1152	1152	0,52	set/13	770	0,38	set/13	1085	1085	0,63
out/13	1291	1291	0,60	out/13	2340	0,72	out/13	1597	1136 (2)	0,67
nov/13	3744	1326 (1)	0,62	nov/13	2200	0,70	nov/13	2693	1136 (1)	0,67
dez/13	647	647	0,26	dez/13	1590	0,57	dez/13	891	891	0,47
jan/14	1075	1075	0,48	jan/14	1800	0,62	jan/14	1217	1217	0,72
fev/14	774	774	0,32	fev/14	0	0,23	fev/14	564	564	0,24
mar/14	920	920	0,39	mar/14	0	0,23	mar/14	704	704	0,33
abr/14	640	640	0,25	abr/14	2480	0,75	abr/14	637	637	0,28
mai/14	1377	1377	0,65	mai/14	1260	0,49	mai/14	1018	1018	0,58
jun/14	1431	1431	0,68	jun/14	450	0,31	jun/14	1090	1090	0,63
jul/14	1706	1706	0,80	jul/14	4310	0,94	jul/14	1390	1390	0,82
ago/14	1286	1286	0,60	ago/14	0	0,23	ago/14	736	736	0,35
set/14	1379	1379	0,65	set/14	300	0,29	set/14	1103	1103	0,64
out/14	1398	1398	0,66	out/14	1830	0,62	out/14	1453	1453	0,85
nov/14	1935	1935	0,87	nov/14	2310	0,72	nov/14	1320	1320	0,78
dez/14	2608	2608	0,97	dez/14	1829	0,62	dez/14	1919	1919	0,96

## APÊNDICE B – CÓDIGOS NO MATLAB

### Modelo de Rede Neural MLP

#### Treinamento da rede neural

```

clear all;clc; %limpar a área de trabalho e zerar qualquer tipo de variável

amountInputNeurons=4;% inserir quantidade de entradas da rede

matrix=xlsread('dados-rafael.xls');%carregar arquivo de dados

matrixsize=length(matrix);%determinar a quantidade de valores armazenados pela matriz

traincoef=0.7;%porcentagem destinada ao treinamento da rede MLP

testcoef=0.2; %porcentagem destinada ao teste da rede MLP

trainingsize=floor(matrixsize*traincoef);%realiza o arredondamento para baixo da amostra
selecionada para treinamento

testsize=floor(matrixsize*(traincoef+testcoef));%realiza o arredondamento para baixo da
amostra selecionada para teste

1º Laço % Definir os quatro meses de entrada e o mês de saída para cada época (iterações até o
final dos dados trabalhados)

%Matriz de valores para treinamento

for count=1:(trainingsize-amountInputNeurons)

    auxinput=matrix(1,count:count+amountInputNeurons-1);

    input(count,:)=auxinput;

    output(count,1)=matrix(count+amountInputNeurons);

end

%Matriz de valores para teste

for count=(trainingsize-amountInputNeurons+1):(testsize-amountInputNeurons)

    auxinput=matrix(1,count:count+amountInputNeurons-1);

    inputtest(count-trainingsize+amountInputNeurons,:)=auxinput;

    outputtest(count-
trainingsize+amountInputNeurons,1)=matrix(count+amountInputNeurons);

```

end

%Quantidade de neurônios de entrada

qtyInput = length(input(1,:));%conta a quantidade colunas da matriz input

%Quantidade de neurônios na camada escondida

hidden = 4;

%Quantidade de neurônios na camada de saída

qtyOut = 1;

% Total de neurônios para ter os pesos calculados (escondidos + saída)

hiddenOut = hidden+qtyOut;

% Inicializar o bias

bias=ones(1,hiddenOut)\*-1;

% Coeficiente de aprendizado

coeff = 0.7;

% Pesos calculados aleatoriamente usando semente

rand('state',sum(100\*clock));

% Pesos para cada neurônio escondido

weightsHidden = -1 +2.\*rand(hidden,qtyInput+1);

% Pesos para cada neurônio de saída

weightsOut = -1 +2.\*rand(qtyOut,hidden+1);

iteracoes=0;% inicializar no 0

% Cria a matrix de saídas. 'qtyout' resultados de saída por evento

% exemplo: para 4 eventos, 4 resultados de saídas

out = zeros(size(output));%cria uma matriz out (obtidos) com o mesmo tamanho da matriz output (esperados) com apenas 0, para inicializar a matriz

3° Laço



```

while iteracoes<100000

    difference=0;

    % Quantidade de eventos de entrada

    numIn = length(input(:,1));%definindo quanto eventos terão

    % Começa o loop para calcular os pesos de cada evento

    % Se existir 10 eventos, iremos calcular 10 loops

    for j = 1:numIn

        % cria um vetor sigma. Cada posição guarda um sigma (saída do neurônio escondido ou
        neurônio de saída) para cada neurônio

        x2=zeros(1,hidden);%Cria a variável x2 (vetor sigma) e inicializa a variável

        %um loop para calcular o sigma de cada neurônio escondido

        for k=1:(hidden)%para cada neurônio da camada escondida

            % Soma( $W_i * X_i$ )->  $W_i$ : pesos dos neurônios

            %  $X_i$ : valor do neurônio de entrada

            % H recebe o bias * peso para cada neurônio escondido

            H(k) = bias(k)*weightsHidden(k,1);%soma da multiplicação dos pesos pelos valores de
            entrada

            % O mesmo é realizado para linha acima, mas agora para cada neurônio de entrada

            for l=1:(qtyInput)

                H(k) = H(k)+(input(j,l)*weightsHidden(k,l+1));

            end

            %H(k)

            % x2 guarda o sigma para cada neurônio escondido

            x2(k)= 1/(1+exp(-H(k)));%aplica a função de ativação

        end

        % Camada de saída – o mesmo feito para os neurônios escondidos, mas agora para os
        neurônios de saída
    end
end

```

```

%x3=zeros(1,qtyOut);

for k=1:(qtyOut)

    H = bias(hidden+k)*weightsOut(k,1);

    for l=1:(hidden)

        H = H+(x2(l)*weightsOut(k,l+1));

    end

    out(j,k)=1/(1+exp(-H));

    difference=difference+(((out(j,k))-(output(j,k)))^2);

end

% Ajuste dos valores do delta dos pesos para a camada de saída

%  $\delta(w_i) = x_i \cdot \delta$ , %delta do peso para cada neurônio é saída daquele neurônio x o
delta

%  $\delta = (1 - \text{actual output}) \cdot (\text{desired output} - \text{actual output})$  %delta do neurônio

delta3=zeros(1,qtyOut);

delta2=zeros(1,hidden);

for k=1:qtyOut

    delta3(k) = out(j,k)*(1-out(j,k))*(output(j,k)-out(j,k));

    for l=1:hidden

        % Retropropaga os pesos o delta dentro da camada escondida

        delta2(l) = x2(l)*(1-x2(l))*weightsOut(k,l+1)*delta3(k);

    end

end

% Adicionar as alterações dos pesos aos pesos originais e usar os novos pesos gerados
para repetir o processo

%  $\delta \text{ weight} = \text{coeff} \cdot x \cdot \delta$ 

for k = 1:hidden

    for l = 1:(qtyInput+1)

```

```

    if(l==1)
        weightsHidden(k,1) = weightsHidden(k,1) + coeff*bias(k)*delta2(k);
    else
        weightsHidden(k,1) = weightsHidden(k,1) + coeff*input(j,l-1)*delta2(k);
    end
end
end
for k = 1:qtyOut
    for l = 1:(hidden+1)
        if(l==1)
            weightsOut(k,1) = weightsOut(k,1) + coeff*bias(hidden+k)*delta3(k);
        else
            weightsOut(k,1) = weightsOut(k,1) + coeff*x2(l-1)*delta3(k);
        end
    end
end
end
difference=difference+((out(j)-output(j))^2);
end
mseerror=difference/numIn;% numIn (quantidade de épocas)
iteracoes=iteracoes+1;
end
mseerror
rmseerror=sqrt(mseerror)
save('pesos.mat');
plot(output,'blue');
hold;

```

```
plot(out,'red');
```

### Teste da rede MLP

```
clear all;clc;
```

```
load('pesos.mat');
```

```
outtraining=out;
```

```
figure('Position',[10 10 1400 900]);
```

```
subplot(1,2,1);
```

```
xlabel('Mes');
```

```
ylabel('Valor');
```

```
plot(output,'red');
```

```
hold;
```

```
plot(outtraining,'blue');
```

```
%%%%%%%% INÍCIO DO TESTE %%%%%%%%%%
```

```
inputtest=inputtest;
```

```
outputtest=outputtest;
```

```
numIn=length(inputtest(:,1));
```

```
out=zeros(numIn,qtyOut);
```

```
for j = 1:numIn
```

```
    % Cria um vetor sigma onde cada posição guarda um sigma para cada neurônio
```

```
    x2=zeros(1,hidden);
```

```
    % Um loop para calcular o sigma e para cada neurônio escondido
```

```
    for k=1:(hidden)
```

```
        % Soma( $W_i * X_i$ )->  $W_i$ : pesos dos neurônios
```

```
        %  $X_i$ : valor do neurônio de entrada
```

```
        % H recebe o bias * peso para cada neurônio escondido
```

```

    H(k) = bias(k)*weightsHidden(k,1); %soma da multiplicação dos pesos pelos valores de
entrada

    % O mesmo da linha acima, mas para cada neurônio de entrada
    for l=1:(qtyInput)
        H(k) = H(k)+(inputtest(j,l)*weightsHidden(k,l+1));
    end

    % x2 guarda o sigma para cada neurônio escondido
    x2(k)=1/(1+exp(-H(k)));
end

% Camada de saída – o mesmo referente aos neurônios escondidos, mas para os neurônios
de saída
for k=1:(qtyOut)
    H = bias(hidden+k)*weightsOut(k,1);
    for l=1:(hidden)
        H = H+(x2(l)*weightsOut(k,l+1));
    end
    out(j,k)=1/(1+exp(-H));
end
end

title('Training Results');
subplot(1,2,2);
xlabel('Mes');
ylabel('Valor');
plot(outputtest,'green');
hold;
plot(out,'black');
title('Tests Results');

```

```
save('resulteste.mat');
```

### **Previsão da MLP**

```
clear all;clc;
```

```
load('resulteste.mat');
```

```
figure('Position',[10 10 1400 900]);
```

```
subplot(2,2,1);
```

```
hold;
```

```
xlabel('Mes');
```

```
ylabel('Valor');
```

```
plot(output,'red');
```

```
plot(outtraining,'blue');
```

```
title('Treinamento');
```

```
subplot(2,2,2);
```

```
hold;
```

```
xlabel('Mes');
```

```
ylabel('Valor');
```

```
plot(outputtest,'green');
```

```
plot(out,'black');
```

```
title('Teste');
```

```
%%%%%%%% INICIA PREVISAO %%%%%%%%%%
```

```
firstinputs(1,:)=matrix(length(matrix)-amountInputNeurons+1:length(matrix));
```

```
firstinputs=firstinputs;
```

```
numIn=12;
```

```
out=zeros(numIn,qtyOut);
```

```
for j = 1:numIn
```

```

% Cria um vetor sigma onde cada posição guarda um sigma para cada neurônio
x2=zeros(1,hidden);

% Um loop para calcular o sigma e para cada neurônio escondido
for k=1:(hidden)

    % Soma(Wi*Xi)-> Wi: pesos dos neurônios
    % Xi: valores neurônios de entrada
    % H recebe o bias * peso para cada neurônio escondido
    H(k) = bias(k)*weightsHidden(k,1); %soma da multiplicação dos pesos pelos valores de
entrada

    % O mesmo da linha acima, mas para cada neurônio de entrada
    for l=1:(qtyInput)
        H(k) = H(k)+(firstinputs(j,l)*weightsHidden(k,l+1));
    end

    %H(k)

    % x2 guarda o sigma para cada neurônio escondido
    x2(k)=1/(1+exp(-H(k)));

end

% Output layer - same as hidden neurons, but for output neurons

% Camada de saída – o mesmo referente aos neurônios escondidos, mas para os neurônios
de saída

%x3=zeros(1,qtyOut);

for k=1:(qtyOut)

    H = bias(hidden+k)*weightsOut(k,1);

    for l=1:(hidden)
        H = H+(x2(l)*weightsOut(k,l+1));
    end
end

```

```
    out(j,k)=1/(1+exp(-H));  
end  
  
    firstinputs(j+1,1:amountInputNeurons-  
qtyOut)=firstinputs(j,qtyOut+1:amountInputNeurons);  
  
    firstinputs(j+1,amountInputNeurons-qtyOut+1:amountInputNeurons)=out(j,1:qtyOut);  
end  
  
% out=round(out)  
  
subplot(2,2,3);  
xlabel('Mes');  
ylabel('Valor');  
plot(out,'magenta');  
hold;  
title('Previsao');
```



## APÊNDICE C – CURVA AB DOS PRODUTOS

Produtos	Total	Curva	Produtos	Total	Curva	Produtos	Total	Curva	Produtos	Total	Curva
Produto 1	42.899,00	A	Produto 65	4.802,00	A	Produto 129	2.570,00	A	Produto 193	1.616,00	A
Produto 2	46.509,00	A	Produto 66	4.787,00	A	Produto 130	2.559,00	A	Produto 194	1.590,00	A
Produto 3	36.715,00	A	Produto 67	4.760,00	A	Produto 131	2.528,00	A	Produto 195	1.584,00	A
Produto 4	27.827,00	A	Produto 68	4.667,00	A	Produto 132	2.509,00	A	Produto 196	1.571,00	A
Produto 5	27.944,00	A	Produto 69	4.512,00	A	Produto 133	2.508,00	A	Produto 197	1.564,00	A
Produto 6	39.215,00	A	Produto 70	4.390,00	A	Produto 134	2.486,00	A	Produto 198	1.555,00	A
Produto 7	23.874,00	A	Produto 71	4.365,00	A	Produto 135	2.464,00	A	Produto 199	1.541,00	A
Produto 8	22.650,00	A	Produto 72	4.335,00	A	Produto 136	2.464,00	A	Produto 200	1.536,00	A
Produto 9	21.892,00	A	Produto 73	4.333,00	A	Produto 137	2.438,00	A	Produto 201	1.531,00	A
Produto 10	20.637,00	A	Produto 74	4.287,00	A	Produto 138	2.425,00	A	Produto 202	1.520,00	B
Produto 11	17.776,00	A	Produto 75	4.279,00	A	Produto 139	2.402,00	A	Produto 203	1.517,00	B
Produto 12	17.525,00	A	Produto 76	4.270,00	A	Produto 140	2.396,00	A	Produto 204	1.512,00	B
Produto 13	16.540,00	A	Produto 77	4.268,00	A	Produto 141	2.390,00	A	Produto 205	1.500,00	B
Produto 14	14.725,00	A	Produto 78	4.230,00	A	Produto 142	2.390,00	A	Produto 206	1.500,00	B
Produto 15	13.685,00	A	Produto 79	4.172,00	A	Produto 143	2.364,00	A	Produto 207	1.496,00	B
Produto 16	13.098,00	A	Produto 80	4.135,00	A	Produto 144	2.351,00	A	Produto 208	1.489,00	B
Produto 17	12.630,00	A	Produto 81	4.108,00	A	Produto 145	2.323,00	A	Produto 209	1.478,00	B
Produto 18	12.527,00	A	Produto 82	4.077,00	A	Produto 146	2.317,00	A	Produto 210	1.477,00	B
Produto 19	11.300,00	A	Produto 83	3.989,00	A	Produto 147	2.259,00	A	Produto 211	1.466,00	B
Produto 20	10.945,00	A	Produto 84	3.977,00	A	Produto 148	2.240,00	A	Produto 212	1.463,00	B
Produto 21	10.737,00	A	Produto 85	3.944,00	A	Produto 149	2.228,00	A	Produto 213	1.458,00	B
Produto 22	10.555,00	A	Produto 86	3.846,00	A	Produto 150	2.198,00	A	Produto 214	1.442,00	B
Produto 23	10.540,00	A	Produto 87	3.786,00	A	Produto 151	2.188,00	A	Produto 215	1.442,00	B
Produto 24	10.279,00	A	Produto 88	3.742,00	A	Produto 152	2.169,00	A	Produto 216	1.435,00	B
Produto 25	9.958,00	A	Produto 89	3.700,00	A	Produto 153	2.140,00	A	Produto 217	1.425,00	B
Produto 26	9.875,00	A	Produto 90	3.696,00	A	Produto 154	2.100,00	A	Produto 218	1.417,00	B
Produto 27	9.803,00	A	Produto 91	3.641,00	A	Produto 155	2.098,00	A	Produto 219	1.412,00	B
Produto 28	9.748,00	A	Produto 92	3.501,00	A	Produto 156	2.085,00	A	Produto 220	1.411,00	B
Produto 29	9.684,00	A	Produto 93	3.500,00	A	Produto 157	2.084,00	A	Produto 221	1.404,00	B
Produto 30	9.500,00	A	Produto 94	3.476,00	A	Produto 158	2.082,00	A	Produto 222	1.400,00	B
Produto 31	9.141,00	A	Produto 95	3.449,00	A	Produto 159	2.074,00	A	Produto 223	1.400,00	B
Produto 32	8.891,00	A	Produto 96	3.413,00	A	Produto 160	2.020,00	A	Produto 224	1.379,00	B
Produto 33	8.847,00	A	Produto 97	3.392,00	A	Produto 161	2.019,00	A	Produto 225	1.374,00	B
Produto 34	8.420,00	A	Produto 98	3.392,00	A	Produto 162	2.017,00	A	Produto 226	1.367,00	B
Produto 35	8.414,00	A	Produto 99	3.392,00	A	Produto 163	2.000,00	A	Produto 227	1.360,00	B
Produto 36	8.393,00	A	Produto 100	3.363,00	A	Produto 164	1.994,00	A	Produto 228	1.337,00	B
Produto 37	8.357,00	A	Produto 101	3.308,00	A	Produto 165	1.970,00	A	Produto 229	1.333,00	B
Produto 38	8.145,00	A	Produto 102	3.288,00	A	Produto 166	1.968,00	A	Produto 230	1.331,00	B
Produto 39	8.137,00	A	Produto 103	3.278,00	A	Produto 167	1.953,00	A	Produto 231	1.326,00	B
Produto 40	7.719,00	A	Produto 104	3.196,00	A	Produto 168	1.944,00	A	Produto 232	1.320,00	B
Produto 41	7.521,00	A	Produto 105	3.129,00	A	Produto 169	1.944,00	A	Produto 233	1.319,00	B
Produto 42	7.024,00	A	Produto 106	3.125,00	A	Produto 170	1.935,00	A	Produto 234	1.313,00	B
Produto 43	7.018,00	A	Produto 107	3.098,00	A	Produto 171	1.897,00	A	Produto 235	1.292,00	B
Produto 44	6.997,00	A	Produto 108	3.095,00	A	Produto 172	1.890,00	A	Produto 236	1.283,00	B
Produto 45	6.902,00	A	Produto 109	3.088,00	A	Produto 173	1.872,00	A	Produto 237	1.267,00	B
Produto 46	6.570,00	A	Produto 110	3.041,00	A	Produto 174	1.865,00	A	Produto 238	1.256,00	B
Produto 47	6.524,00	A	Produto 111	3.023,00	A	Produto 175	1.845,00	A	Produto 239	1.241,00	B
Produto 48	6.451,00	A	Produto 112	3.004,00	A	Produto 176	1.832,00	A	Produto 240	1.240,00	B
Produto 49	6.423,00	A	Produto 113	3.004,00	A	Produto 177	1.831,00	A	Produto 241	1.205,00	B
Produto 50	6.404,00	A	Produto 114	3.000,00	A	Produto 178	1.808,00	A	Produto 242	1.202,00	B
Produto 51	6.142,00	A	Produto 115	2.949,00	A	Produto 179	1.790,00	A	Produto 243	1.189,00	B
Produto 52	6.115,00	A	Produto 116	2.846,00	A	Produto 180	1.778,00	A	Produto 244	1.185,00	B
Produto 53	6.090,00	A	Produto 117	2.795,00	A	Produto 181	1.768,00	A	Produto 245	1.175,00	B
Produto 54	5.746,00	A	Produto 118	2.760,00	A	Produto 182	1.761,00	A	Produto 246	1.168,00	B
Produto 55	5.633,00	A	Produto 119	2.678,00	A	Produto 183	1.741,00	A	Produto 247	1.160,00	B
Produto 56	5.629,00	A	Produto 120	2.674,00	A	Produto 184	1.740,00	A	Produto 248	1.160,00	B
Produto 57	5.621,00	A	Produto 121	2.673,00	A	Produto 185	1.733,00	A	Produto 249	1.157,00	B
Produto 58	5.378,00	A	Produto 122	2.638,00	A	Produto 186	1.732,00	A	Produto 250	1.155,00	B
Produto 59	5.220,61	A	Produto 123	2.615,00	A	Produto 187	1.729,00	A	Produto 251	1.139,00	B
Produto 60	5.175,00	A	Produto 124	2.613,00	A	Produto 188	1.720,00	A	Produto 252	1.138,00	B
Produto 61	4.903,00	A	Produto 125	2.607,00	A	Produto 189	1.675,00	A	Produto 253	1.107,00	B
Produto 62	4.889,00	A	Produto 126	2.575,00	A	Produto 190	1.669,00	A	Produto 254	1.099,00	B
Produto 63	4.876,00	A	Produto 127	2.572,00	A	Produto 191	1.663,00	A	Produto 255	1.098,00	B
Produto 64	4.812,00	A	Produto 128	2.570,00	A	Produto 192	1.658,00	A	Produto 256	1.056,00	B

## APÊNDICE C – CURVA AB DOS PRODUTOS (Continuação)

Produtos	Total	Curva	Produtos	Total	Curva	Produtos	Total	Curva	Produtos	Total	Curva
Produto 257	1.055,00	B	Produto 321	731,00	B	Produto 385	546,00	B	Produto 445	395,00	B
Produto 258	1.047,00	B	Produto 322	727,00	B	Produto 386	545,00	B	Produto 446	394,00	B
Produto 259	1.018,00	B	Produto 323	725,00	B	Produto 387	542,00	B	Produto 447	388,00	B
Produto 260	997,00	B	Produto 324	722,00	B	Produto 388	540,00	B	Produto 448	384,00	B
Produto 261	993,00	B	Produto 325	712,00	B	Produto 389	535,00	B	Produto 449	384,00	B
Produto 262	981,00	B	Produto 326	706,00	B	Produto 390	529,00	B	Produto 450	383,00	B
Produto 263	972,00	B	Produto 327	705,00	B	Produto 391	525,00	B	Produto 451	381,00	B
Produto 264	964,00	B	Produto 328	705,00	B	Produto 392	523,00	B	Produto 452	379,00	B
Produto 265	960,00	B	Produto 329	703,00	B	Produto 393	519,00	B	Produto 453	379,00	B
Produto 266	957,00	B	Produto 330	700,00	B	Produto 394	518,00	B	Produto 454	377,00	B
Produto 267	955,00	B	Produto 331	698,00	B	Produto 395	516,00	B	Produto 455	374,00	B
Produto 268	939,00	B	Produto 332	694,00	B	Produto 396	512,00	B	Produto 456	370,00	B
Produto 269	937,00	B	Produto 333	693,00	B	Produto 397	505,00	B	Produto 457	370,00	B
Produto 270	936,00	B	Produto 334	693,00	B	Produto 398	502,00	B	Produto 458	370,00	B
Produto 271	936,00	B	Produto 335	691,00	B	Produto 399	500,00	B	Produto 459	370,00	B
Produto 272	927,00	B	Produto 336	686,00	B	Produto 400	498,00	B	Produto 460	370,00	B
Produto 273	926,00	B	Produto 337	681,00	B	Produto 401	497,00	B	Produto 461	367,00	B
Produto 274	921,00	B	Produto 338	680,00	B	Produto 402	490,00	B	Produto 462	365,00	B
Produto 275	910,00	B	Produto 339	679,00	B	Produto 403	490,00	B	Produto 463	365,00	B
Produto 276	907,00	B	Produto 340	678,00	B	Produto 404	490,00	B	Produto 464	361,00	B
Produto 277	898,00	B	Produto 341	674,00	B	Produto 405	488,00	B	Produto 465	360,00	B
Produto 278	894,00	B	Produto 342	674,00	B	Produto 406	486,00	B	Produto 466	359,00	B
Produto 279	888,00	B	Produto 343	663,00	B	Produto 407	479,00	B	Produto 467	357,00	B
Produto 280	881,00	B	Produto 344	662,00	B	Produto 408	479,00	B	Produto 468	354,00	B
Produto 281	881,00	B	Produto 345	652,00	B	Produto 409	469,00	B	Produto 469	353,00	B
Produto 282	877,00	B	Produto 346	642,00	B	Produto 410	468,00	B	Produto 470	353,00	B
Produto 283	872,00	B	Produto 347	641,00	B	Produto 411	465,00	B	Produto 471	352,00	B
Produto 284	857,00	B	Produto 348	638,00	B	Produto 412	465,00	B	Produto 472	352,00	B
Produto 285	854,00	B	Produto 349	638,00	B	Produto 413	461,00	B	Produto 473	350,00	B
Produto 286	853,00	B	Produto 350	638,00	B	Produto 414	461,00	B	Produto 474	349,00	B
Produto 287	851,00	B	Produto 351	633,00	B	Produto 411	465,00	B	Produto 475	345,00	B
Produto 288	846,00	B	Produto 352	632,00	B	Produto 412	465,00	B	Produto 476	345,00	B
Produto 289	843,00	B	Produto 353	631,00	B	Produto 413	461,00	B	Produto 477	342,00	B
Produto 290	840,00	B	Produto 354	627,00	B	Produto 414	461,00	B	Produto 478	342,00	B
Produto 291	837,00	B	Produto 355	627,00	B	Produto 415	460,00	B	Produto 479	340,00	B
Produto 292	822,00	B	Produto 356	618,00	B	Produto 416	454,00	B	Produto 480	338,00	B
Produto 293	818,00	B	Produto 357	618,00	B	Produto 417	452,00	B	Produto 481	335,00	B
Produto 294	808,00	B	Produto 358	618,00	B	Produto 418	450,00	B	Produto 482	333,00	B
Produto 295	803,00	B	Produto 359	617,00	B	Produto 419	450,00	B	Produto 483	333,00	B
Produto 296	801,00	B	Produto 360	616,00	B	Produto 420	450,00	B	Produto 484	332,00	B
Produto 297	792,00	B	Produto 361	611,00	B	Produto 421	449,00	B	Produto 485	331,00	B
Produto 298	789,00	B	Produto 362	610,00	B	Produto 422	449,00	B	Produto 486	330,00	B
Produto 299	785,00	B	Produto 363	608,00	B	Produto 423	448,00	B	Produto 487	330,00	B
Produto 300	785,00	B	Produto 364	607,00	B	Produto 424	440,00	B	Produto 488	329,00	B
Produto 301	777,00	B	Produto 365	602,00	B	Produto 425	434,00	B	Produto 489	327,00	B
Produto 302	774,00	B	Produto 366	600,00	B	Produto 426	429,00	B	Produto 490	327,00	B
Produto 303	769,00	B	Produto 367	600,00	B	Produto 427	429,00	B	Produto 491	326,00	B
Produto 304	764,00	B	Produto 368	599,00	B	Produto 428	429,00	B	Produto 492	326,00	B
Produto 305	761,00	B	Produto 369	598,00	B	Produto 429	426,00	B	Produto 493	325,00	B
Produto 306	758,00	B	Produto 370	592,00	B	Produto 430	422,00	B	Produto 494	324,00	B
Produto 307	755,00	B	Produto 371	591,00	B	Produto 431	421,00	B	Produto 495	321,00	B
Produto 308	754,00	B	Produto 372	589,00	B	Produto 432	420,00	B	Produto 496	320,00	B
Produto 309	753,00	B	Produto 373	587,00	B	Produto 433	420,00	B	Produto 497	320,00	B
Produto 310	751,00	B	Produto 374	587,00	B	Produto 434	416,00	B	Produto 498	320,00	B
Produto 311	749,00	B	Produto 375	585,00	B	Produto 435	414,00	B	Produto 499	318,00	B
Produto 312	748,00	B	Produto 376	581,00	B	Produto 436	411,00	B	Produto 500	318,00	B
Produto 313	746,00	B	Produto 377	571,00	B	Produto 437	410,00	B	Produto 501	317,00	B
Produto 314	744,00	B	Produto 378	558,00	B	Produto 438	410,00	B	Produto 502	316,00	B
Produto 315	743,00	B	Produto 379	555,00	B	Produto 439	409,00	B	Produto 503	315,00	B
Produto 316	740,00	B	Produto 380	551,00	B	Produto 440	405,00	B	Produto 504	313,00	B
Produto 317	738,00	B	Produto 381	550,00	B	Produto 441	400,00	B	Produto 505	312,00	B
Produto 318	733,00	B	Produto 382	550,00	B	Produto 442	398,00	B	Produto 506	310,00	B
Produto 319	732,00	B	Produto 383	550,00	B	Produto 443	397,00	B			
Produto 320	732,00	B	Produto 384	547,00	B	Produto 444	396,00	B			

**APÊNDICE D – SÉRIE TEMPORAL REAL DOS PRODUTOS X SÉRIE  
PROCESSADA PELA RNA MLP**

<b>Produto 1 - Rede Neural MLP</b>			
<b>Série temporal normalizada após tratamento outliers</b>	<b>Série temporal normalizada processada pela Rede</b>	<b>Série temporal real após tratamento outliers</b>	<b>Série temporal processada pela Rede Neural Artificial</b>
0,21	0,21	539	539
0,29	0,29	729	729
0,39	0,39	910	910
0,33	0,33	795	795
0,38	0,39	904	919
0,27	0,31	668	761
0,48	0,48	1076	1078
0,37	0,39	871	920
0,49	0,49	1089	1085
0,35	0,32	833	792
0,34	0,25	830	630
0,26	0,24	647	622
0,29	0,31	731	759
0,56	0,55	1215	1189
0,43	0,43	995	989
0,81	0,81	1738	1739
0,24	0,24	615	612
0,70	0,70	1471	1473
0,49	0,49	1092	1091
0,57	0,57	1223	1226
0,52	0,54	1152	1175
0,60	0,61	1291	1299
0,62	0,55	1326	1202
0,26	0,34	647	822
0,48	0,49	1075	1093
0,32	0,34	774	826
0,39	0,79	920	1680
0,25	0,41	640	943
0,65	0,55	1377	1187
0,68	0,42	1431	968
0,80	0,94	1706	2311
0,60	0,13	1286	274
0,65	0,37	1379	873
0,66	0,98	1398	2742
0,87	0,91	1935	2122
0,97	0,38	2608	905

**APÊNDICE D – SÉRIE TEMPORAL REAL DOS PRODUTOS X SÉRIE  
PROCESSADA PELA RNA MLP (Continuação)**

<b>Produto 2 - Rede Neural MLP</b>			
<b>Série temporal normalizada após tratamento outliers</b>	<b>Série temporal normalizada processada pela Rede</b>	<b>Série temporal original após tratamento outliers</b>	<b>Série temporal original processada pela Rede Neural Artificial</b>
0,49	0,49	1250	1250
0,38	0,38	770	770
0,69	0,69	2140	2140
0,44	0,44	1050	1050
0,81	0,67	2890	2048
0,38	0,34	780	595
0,25	0,34	130	559
0,34	0,30	580	394
0,47	0,40	1150	864
0,61	0,78	1760	2639
0,58	0,65	1620	1953
0,89	0,70	3570	2226
0,23	0,27	0	235
0,23	0,19	0	0
0,28	0,33	260	531
0,41	0,42	880	945
0,74	0,63	2400	1880
0,54	0,63	1450	1870
0,30	0,50	360	1283
0,23	0,26	0	158
0,38	0,37	770	695
0,72	0,78	2340	2676
0,70	0,65	2200	1961
0,57	0,44	1590	1033
0,62	0,64	1800	1928
0,23	0,20	0	0
0,23	0,32	0	469
0,75	0,36	2480	657
0,49	0,70	1260	2186
0,31	0,23	450	0
0,94	0,33	4310	516
0,23	0,06	0	0
0,29	0,08	300	0
0,62	0,29	1830	296
0,72	0,64	2310	1913
0,62	0,49	1829	1242

**APÊNDICE D – SÉRIE TEMPORAL REAL DOS PRODUTOS X SÉRIE  
PROCESSADA PELA RNA MLP (Continuação)**

<b>Produto 3 - Rede Neural MLP</b>			
<b>Série temporal normalizada após tratamento outliers</b>	<b>Série temporal normalizada processada pela Rede</b>	<b>Série temporal original após tratamento outliers</b>	<b>Série temporal original processada pela Rede Neural Artificial</b>
0,30	0,30	659	659
0,24	0,24	571	571
0,43	0,43	842	842
0,28	0,28	632	632
0,21	0,20	513	507
0,37	0,37	756	757
0,69	0,67	1171	1136
0,24	0,24	573	574
0,42	0,42	828	828
0,47	0,47	893	885
0,47	0,50	884	927
0,45	0,46	864	882
0,16	0,19	422	474
0,31	0,31	676	678
0,55	0,55	984	984
0,44	0,44	851	855
0,48	0,47	898	894
0,36	0,35	752	738
0,55	0,54	992	976
0,67	0,73	1136	1234
0,63	0,53	1085	965
0,67	0,66	1136	1127
0,67	0,63	1136	1090
0,47	0,53	891	959
0,72	0,73	1217	1234
0,24	0,51	564	932
0,33	0,00	704	0
0,28	0,23	637	557
0,58	0,51	1018	932
0,63	0,83	1090	1407
0,82	0,75	1390	1266
0,35	0,96	736	1932
0,64	0,02	1103	0
0,85	0,56	1453	995
0,78	0,49	1320	915
0,96	0,97	1919	1963



## APÊNDICE E – SAÍDAS GRÁFICAS DOS TRÊS PRODUTOS GERADAS PELO SOFTWARE MATLAB

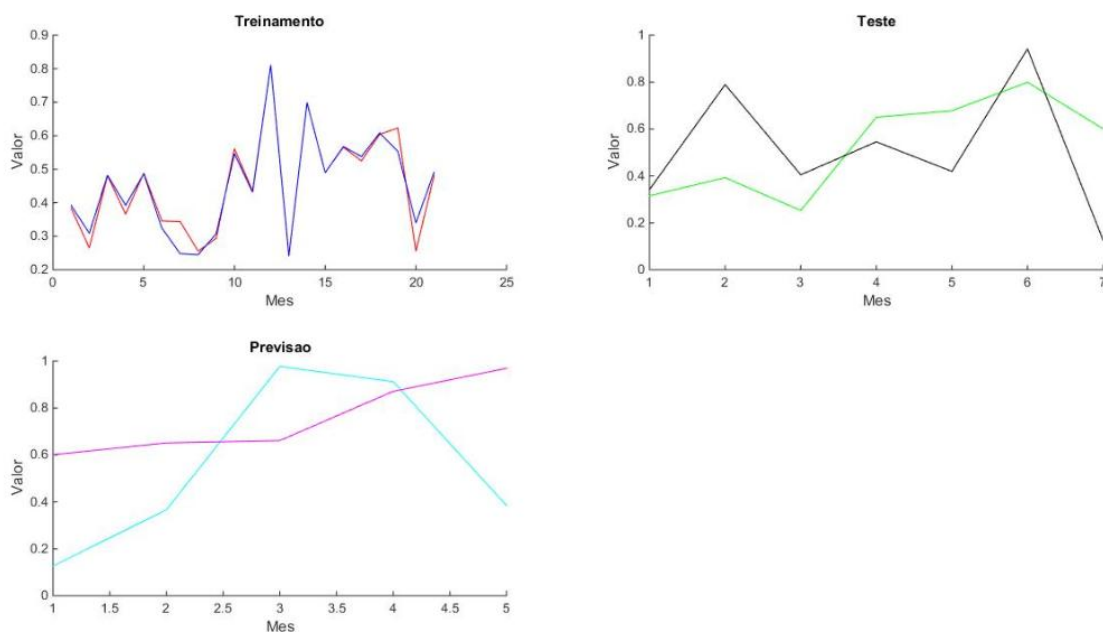


Gráfico 1 – Saída do *software* MATLAB referente ao treinamento, teste e previsão da Rede Neural MLP - Produto 1

Fonte: Desenvolvido pelo autor

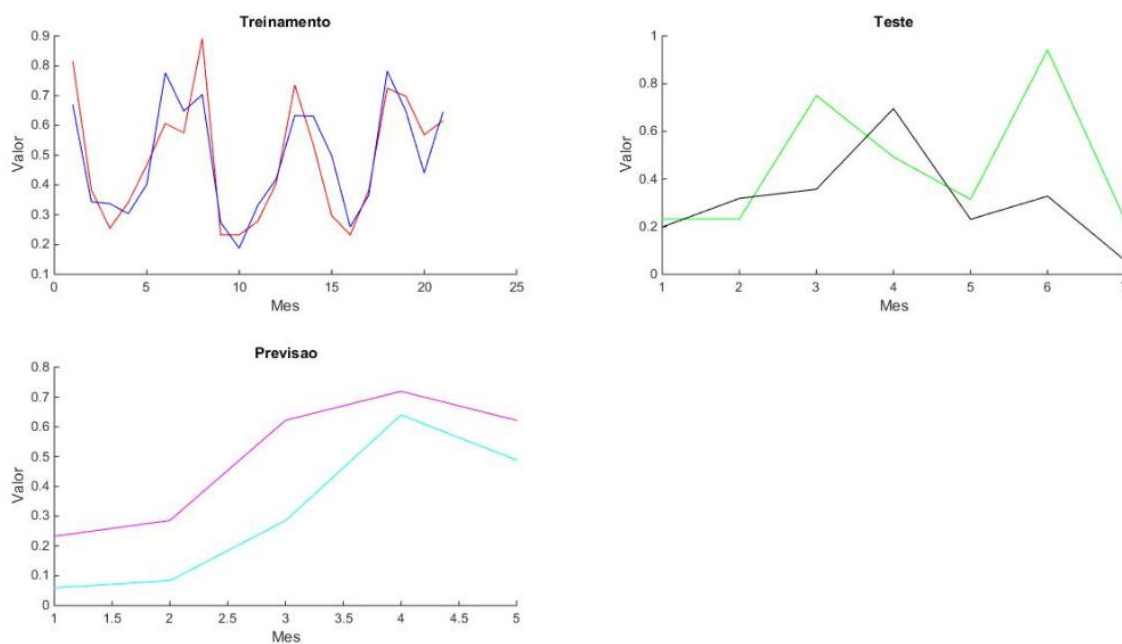


Gráfico 2 – Saída do *software* MATLAB referente ao treinamento, teste e previsão da Rede Neural MLP - Produto 2

Fonte: Desenvolvido pelo autor

## APÊNDICE E – SAÍDAS GRÁFICAS DOS TRÊS PRODUTOS GERADAS PELO SOFTWARE MATLAB (Continuação)

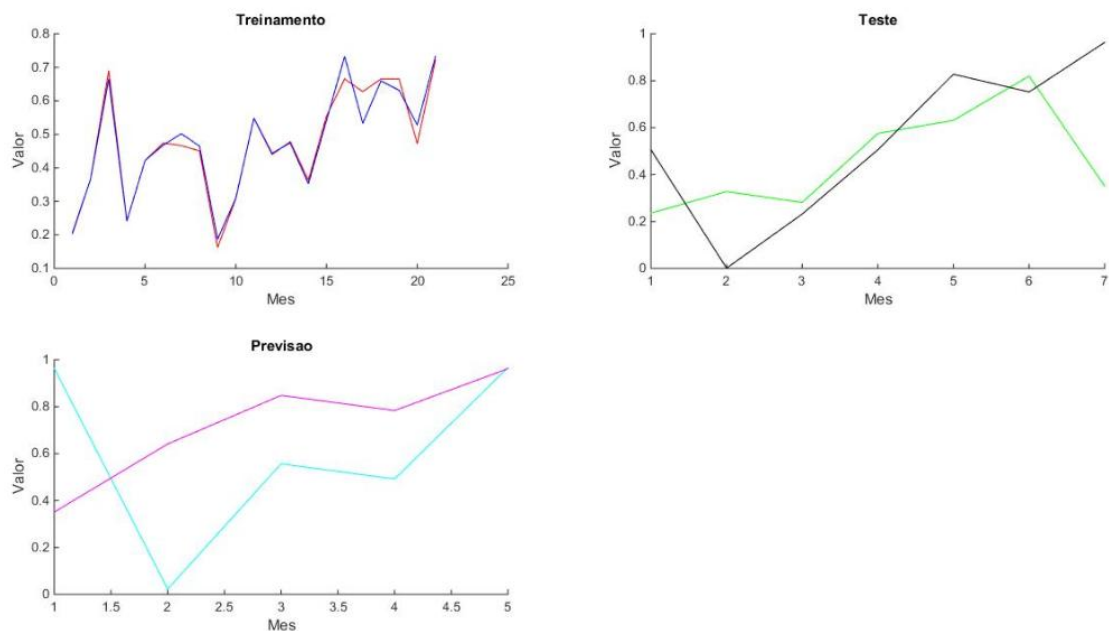


Gráfico 3 – Saída do *software* MATLAB referente ao treinamento, teste e previsão da Rede Neural MLP - Produto 3

Fonte: Desenvolvido pelo autor

## APÊNDICE F – CÓDIGOS NO R

### Método de Box-Jenkins

```
###Pacotes necessários
```

```
library(forecast)
```

```
library(tseries)
```

```
library(fUnitRoots)
```

```
library(fpp)
```

```
library(expsmooth)
```

```
library(mcomp)
```

```
library(fma)
```

```
dados= c(dados da série temporal)
```

```
y= ts(dados, start=c(2012, 1), freq=12)
```

```
summary(y)
```

```
var(y)
```

```
sd(y)
```

```
lag.plot(y, lags=6, do.lines=FALSE)
```

```
acf(y, ylab="Função de autocorrelação", xlab="Defasagem, em anos", main="(a)")
```

```
pacf(y, ylab="Função de autocorrelação parcial", xlab="Defasagem, em anos", main="(b)")
```

```
library(forecast)
```

```
arimafit<-auto.arima(y)
```

```
arimafit
```

```
summary(arimafit)
```

```
residuos<-c(arimafit$residuals)
```

```
residuos
```

```
tsdiag(arimafit)
```

```
Box.test(residuos, lag=1, type='Ljung-Box')
```



```
Box.test(residuos, lag=2, type='Ljung-Box')  
shapiro.test(residuos)  
plot(forecast(arimafit, h= 12))  
forecast(arimafit)  
plot(y,col=4)  
lines(fitted(arimafit))  
qqnorm(residuos,xlab="Quantis teórico", ylab="Quantis da amostra", main="Teste de  
normalidade")  
qqline(residuos)
```