

MARCO AURÉLIO SANTILI

Métodos de previsão Auto-Regressivo aplicado a uma série de volume de produção de caminhões.

Marco Aurélio Santili

Métodos de previsão Auto-Regressivo aplicado a uma série de volume de produção de caminhões.

Trabalho de Graduação apresentado ao Conselho de Curso de Graduação em Engenharia de Produção Mecânica da Faculdade de Engenharia do Campus de Guaratinguetá, Universidade Estadual Paulista, como parte dos requisitos para obtenção do diploma de Graduação em Engenharia Mecânica.

Orientador: Dr - Aneirson Francisco da Silva

Guaratinguetá
2015

S235m

Santili, Marco Aurélio

Métodos de previsão Auto-Regressivo aplicado a uma série de volume de produção de caminhões / Marco Aurélio Santili – Guaratinguetá, 2015.
42 f. : il.

Bibliografia : f. 37-42

Trabalho de Graduação em Engenharia Mecânica – Universidade Estadual Paulista, Faculdade de Engenharia de Guaratinguetá, 2015.

Orientador: Prof. Dr. Aneirson Francisco da Silva

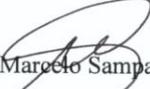
1. Previsão com Metodologia de Box-Jenkins 2. Controle de produção
3. Previsão de vendas I. Título

CDU 658.5

Marco Aurélio Santili

ESTE TRABALHO DE GRADUAÇÃO FOI JULGADO ADEQUADO COMO
PARTE DO REQUISITO PARA A OBTENÇÃO DO DIPLOMA DE
“GRADUADO EM ENGENHARIA MECÂNICA”

APROVADO EM SUA FORMA FINAL PELO CONSELHO DE CURSO DE
GRADUAÇÃO EM ENGENHARIA MECÂNICA


Prof. Dr. Marcelo Sampaio Martins
Coordenador

BANCA EXAMINADORA:


Prof. Dr. Aneirson Francisco da Silva
Orientador/UNESP-FEG


Prof. Msc. Erica Ximenes Dias
UNESP-FEG


Prof. Dr. José Roberto Dale Luche
UNESP-FEG

Dezembro de 2015

SANTILI, M. A. **Métodos de previsão Auto-Regressivo aplicado a uma série de volume de produção de caminhão.** 2015. 45f. Trabalho de Graduação (Graduação em Engenharia Mecânica) – Faculdade de Engenharia do Campus de Guaratinguetá, Universidade Estadual Paulista, Guaratinguetá, 2015.

RESUMO

A frota nacional de caminhões sofreu uma expansão forte nas últimas décadas. Contudo, devido à variação da demanda que o mercado se expõe, necessitou-se de tomadas de decisões estratégicas mais efetivas das montadoras. Essas decisões são tomadas após uma garantida avaliação de previsões de vendas. Este trabalho tem como objetivo gerar uma previsão anual de produção de caminhão por meio da metodologia Box e Jenkins. Foram utilizados dados anuais para modelagem da previsão referente a partir do ano de 1957 a 2014, que foram obtidos pela Associação Nacional de Fabricantes de veículos Automotores (ANFAVEA). O modelo utilizado foi Autorregressivo Integrado de Média Móvel (ARIMA) e foi possível escolher o melhor modelo para a série em estudo, sendo o ARIMA (2,1,3) o mais representativo para realização de previsão de produção de caminhão.

PALAVRASCHAVE: ARIMA. Previsão de produção de caminhão.

ABSTRACT

The national truck fleet has expanded strongly in recent decades. However, due to fluctuations in the demand that the market is exposed, it needed up making more effective strategic decisions of automakers. These decisions are made after an evaluation of guaranteed sales forecasts. This work aims to generate an annual forecast of truck production by Box and Jenkins methodology. They used annual data for referring forecast modeling from the year 1957 to 2014, which were obtained by the National Association of Motor Vehicle Manufacturers (Anfavea). The model used was Autoregressive Integrated Moving Average (ARIMA) and can choose the best model for the series under study, and the ARIMA (2,1,3) as representative for conducting truck production forecast.

KEYWORDS: ARIMA. Truck PRODUCTION FORECAST.

SUMÁRIO

1 INTRODUÇÃO	6
1.1 CONTRIBUIÇÃO DA PESQUISA	6
1.2 OBJETIVOS	8
1.2.1 Objetivos específicos.....	8
1.3 DELIMITAÇÃO E QUESTÃO DE PESQUISA.....	8
1.4 JUSTIFICATIVA.....	8
1.4 MATERIAIS E MÉTODOS	9
1.5 ESTRUTURA DO TRABALHO	10
2 FUNDAMENTAÇÃO TEÓRICA.....	11
2.1 O MERCADO AUTOMOBILÍSTICO BRASILEIRO.....	11
2.2 MODELO DE PREVISÃO CLÁSSICA.....	12
2.2.1 Modelos Auto-regressivos Lineares Estacionários	13
2.2.1.2 Modelo Médias Móveis (MA (q))	14
2.2.1.3 Modelo Médias Móveis Auto Regressivas ARMA (p,q).....	15
2.2.2 Modelos Auto-regressivos Lineares não Estacionários.....	16
2.2.2.1 O Modelo Auto-Regressivo Integrado Médias Móveis (ARIMA(p,d,q)).....	16
2.2.2.2 Modelo Médias Móveis Integradas Auto-regressivas Sazonais SARIMA (p,d,q) x (P,D,Q).....	17
2.3 ERROS DE PREVISÃO	18
2.3.1 Erro Absoluto Médio – MAD	18
2.3.2 Erro Percentual Absoluto Médio – MAPE.....	19
2.3.3 Erro Quadrático Médio – EQM	19
2.3.4 Raiz Unitária.....	19
3 DESCRIÇÃO E MODELAGEM DO PROBLEMA.....	22
3.1 ETAPAS PARA O DESENVOLVIMENTO DO MODELO DE PREVISÃO AUTO-REGRESSIVO	22
3.1.2 Identificação do modelo.....	22
3.1.3 Estimação.....	22
3.1.4 Previsão.....	23
3.2 ANÁLISE DA SÉRIE.....	23
3.3 IDENTIFICAÇÃO DO MODELO ARIMA	28
3.4 VALIDAÇÃO DO MODELO	30
4 CONCLUSÕES E RECOMENDAÇÕES PARA FUTURAS PESQUISAS	35
4.1 VERIFICAÇÃO DOS OBJETIVOS.....	35
4.2 RECOMENDAÇÕES PARA FUTURAS PESQUISAS	35
REFERÊNCIAS BIBLIOGRÁFICAS	37
ANEXO A: DADOS REAIS DE PRODUÇÃO NACIONAL DE CAMINHÕES	41
ANEXO B: VALIDAÇÃO DO MODELO DE PREVISÃO.....	42

1 INTRODUÇÃO

1.1 CONTRIBUIÇÃO DA PESQUISA

A Associação Nacional dos Fabricantes de Veículos Automotores (ANFAVEA) fundada em 15 de maio de 1956, reúne empresas fabricantes de auto veículos (automóveis, comerciais leves, caminhões, ônibus) e máquinas agrícolas automotrizes (tratores de rodas e de esteiras, colheitadeiras e retroescavadeiras) com instalações industriais no Brasil. A entidade estuda temas da indústria e do mercado de veículos e máquinas agrícolas automotrizes, além de coordenar interesses coletivos das empresas associadas.

Segundo a ANFAVEA, o mercado de caminhões se divide da seguinte forma:

- Semileves: $3,5 \text{ ton} < \text{PBT} < 6 \text{ ton}$;
- Leves: $6 \text{ ton} \leq \text{PBT} < 10 \text{ ton}$;
- Médio: $10 \text{ ton} \leq \text{PBT} < 15 \text{ ton}$;
- Semi Pesados: $\text{PBT} \geq 15 \text{ ton}$.

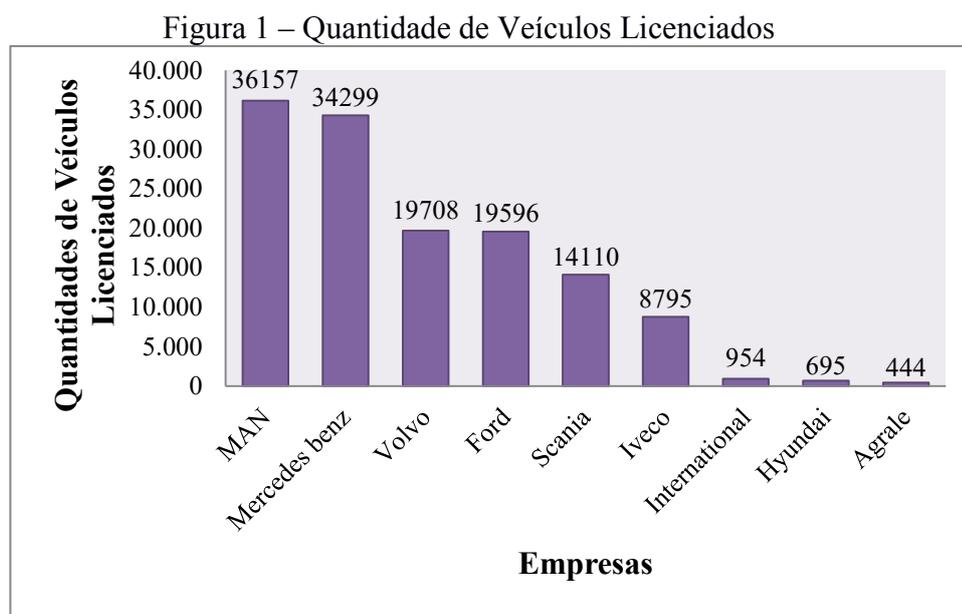
O peso bruto total (PBT) é o peso máximo transmitido ao pavimento pela combinação de um caminhão-trator mais seu semi-reboque ou do caminhão mais seus reboques. Um caminhão-trator é um veículo destinado a tracionar ou arrastar outro e um semi-reboque é um veículo de um ou mais eixos que se apoia na sua unidade tratora ou é ligado por meio de articulação.

Segundo dados da ANFAVEA em seu balanço anual, as montadoras de caminhões que operam no Brasil fabricaram 25% menos veículos em 2014 do que em 2013. Com o decréscimo, a produção de caminhão no mercado brasileiro somou 143.660 unidades em 2014 e 190.962 unidades em 2013. Entretanto, as vendas somaram 134.998 unidades vendidas no mercado interno, ou seja, quase 94% da produção. Outras 2.061 unidades negociadas no mercado internacional, gerando um estoque de 6.601 caminhões.

Dos caminhões licenciados do ano de 2014 os que mais se destacam são semipesados e os pesados, pois exerceram uma marca de 45.236 e 47.428 veículos respectivamente, representando 31,48% e 33,02% do total saído da linha de produção nacional. Conforme os dados da ANFAVEA, a liderança nas vendas de caminhões se manteve com a MAN (que comercializa a marca Volkswagen e da própria marca) pelo décimo ano consecutivo. Nesse ano de 2014 a empresa vendeu 36.157 veículos, com aproximadamente 26,78% do mercado

nacional. Logo em seguida, Mercedes-Benz, que encerrou o ano com 34.299 veículos comercializados. A Volvo aparece em terceiro com 19.708, em quarto lugar a Ford com 19.596, em seguida a Scania com 14.110 caminhões no ano. A sexta colocada no *ranking* de vendas do mercado nacional ficou a Iveco com 8.795 unidades. As vendas restantes se dividiram entre as marcas Internacional, Hyundai e Agrale.

A relação dos licenciamentos totais obtidos no ano de 2014, a Figura 1 aborda o fechamento do ano de 2014:



Fonte: Anuário da Indústria Brasileira 2015, ANFAVEA

Neste contexto, os modelos de previsão são utilizados para ajudar as organizações a dimensionarem as quantidades de produtos que irão produzir ou até mesmo serviços que algumas prestarão, de tal maneira, que as organizações venham a atender o mercado do melhor modo possível (TUBINO, 2007).

Segundo Tubino (2007), a previsão da demanda é parte integrante do processo decisório, devido a sua importância na decisão de recursos necessários, que são as bases do planejamento estratégico e da gestão empresarial. É fundamental para o sucesso de qualquer negócio obter uma previsão confiável e com um bom dimensionamento, proporcionando benefícios para fabricantes, fornecedores e clientes (FORMIGONI, 2013).

Segundo Werner e Ribeiro (2003), uma das técnicas quantitativas mais difundidas para realização de previsões é a metodologia Box & Jenkins, essa metodologia foi descrita na década de 70. Os modelos Box & Jenkins iniciam de uma ideia de que cada valor da série

temporal pode ser explicado por valores prévios, a partir do uso de correlações temporais que frequentemente há nos valores da série.

Uma vez que o setor automotivo é de suma importância para o segmento industrial e econômico nacional, este trabalho tem como principal contribuição analisar dados reais e gerar previsão na produção de caminhões, utilizando a metodologia Box & Jenkins (GUARNIERI; SOUZA; ALMEIDA, 2009).

1.2 OBJETIVOS

O objetivo deste trabalho foi prever a demanda anual de produção de caminhões. Por meio de modelos Auto-Regressivos.

1.2.1 Objetivos específicos

Utilizar um modelo Auto-Regressivo integrado de médias móveis-ARIMA na previsão anual de produção de caminhões.

Validar o modelo por meio dos cenários elaborados por meio de economistas.

1.3 DELIMITAÇÃO E QUESTÃO DE PESQUISA

O presente trabalho restringe-se a modelagem de séries temporais lineares, desta forma, não serão estudados modelos de séries temporais não lineares.

O modelo ARIMA gerará uma previsão de demanda confiável vinculada a produção anual de caminhões?

1.4 JUSTIFICATIVA

O estudo de previsão de demanda por séries temporais consiste em utilizar dados passados e presentes para se obter dados futuros, por meio de modelos descritos e variáveis que estão sendo estudadas (SANTOS, 2012). A previsão de demanda é uma atividade de grande importância para a competitividade das empresas, uma vez que auxilia a prever o que

deverá ser produzido ou comprado, evitando perdas desnecessárias ou o acúmulo de estoques (WERNER, 2009).

Roldan e Miyake (2004) relatam que, quanto mais próxima da realidade for a previsão, maior será o grau de atendimento da demanda pela produção programada. A indústria automobilística e sua cadeia de suprimentos é um bom exemplo da dependência da previsão de demanda para o sucesso operacional (ROLDAN e MIYAKE, 2004).

As indústrias tem respondido apressadamente a demanda do mercado tendo como base o gerenciamento de cadeias e nos conceitos de manufatura enxuta (SOMBRIO; NEUMANN; WENER, 2010). Porém, segundo o modelo tradicional dessas indústrias, o planejamento e a programação dos processos produtivos e o fluxo na cadeia são baseados na previsão de demanda (ROLDAN; MIYAKE, 2004). As previsões acuradas impactam diretamente em melhorias no desempenho de custo da organização, visto que, viabilizam a redução de perdas (ALMEIDA; WERNER, 2015).

Com as informações subsidiadas pela previsão de demanda é possível, por meio da atividade de planejamento, estabelecer uma utilização otimizada de recursos, que por consequência fornecem maior competitividade à organização (DANESE e KALCHSCHIMIDT, 2011; LI *et al.*, 2013).

1.4 MATERIAIS E MÉTODOS

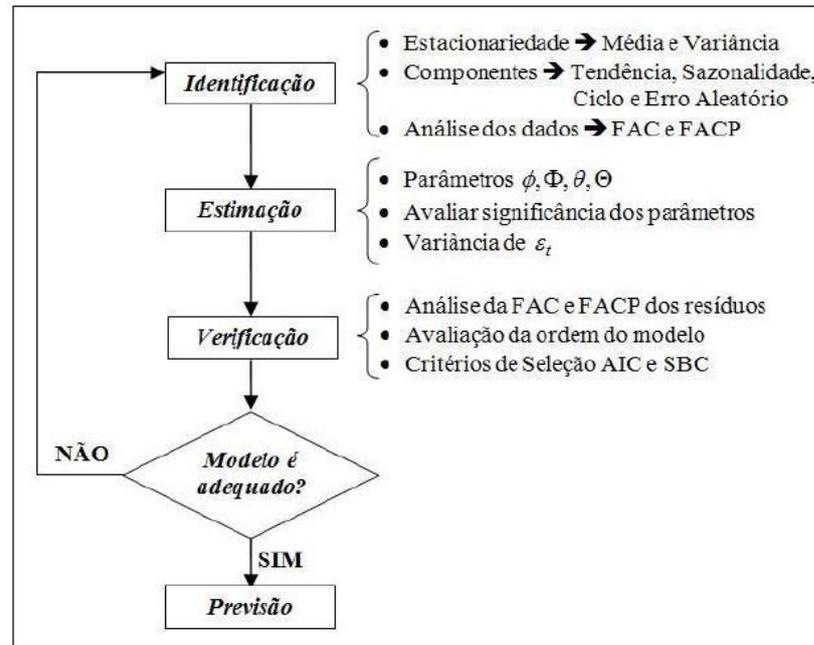
Segundo Bertrand e Fransoo (2002), essa pesquisa é classificada com uma pesquisa aplicada, pois visa proporcionar melhorias para a literatura atual. Os objetivos são empíricos descritivos, uma vez que busca desenvolver um modelo que favoreça a compreensão das relações causais que existem em problemas reais. A forma de abordar o problema é quantitativa, sendo o método de pesquisa a modelagem.

Sobre os materiais, essa pesquisa investigará modelos de previsão auto-regressivos para séries temporais lineares, como, por exemplo, os métodos AR – Auto-Regressivo, ARMA – Médias Móveis Auto-Regressivas, ARIMA – Médias Móveis Integradas Auto-Regressivas, por meio do *software* GRETL.

Segundo Moretin e Tolo (2006), a construção do modelo Box & Jenkins é baseada em um ciclo iterativo, no qual a escolha do modelo é feita com os próprios dados. A metodologia de Box & Jenkins é composta de três etapas, conforme a Figura 2. Estas etapas serão

obedecidas para encontrar um modelo de previsão que descreve um comportamento da série de produção de caminhões, possibilitando construir uma série com menor erro.

Figura 2 – Quadro Representativo dos Estágios do Ciclo Iterativo de Box-Jenkins



Fonte: Adaptado de Marchezan (2007)

1.5 ESTRUTURA DO TRABALHO

O presente trabalho está estruturado em 4 capítulos, incluindo esse capítulo introdutório. No capítulo 2 encontra-se a fundamentação teórica que aborda os modelos de previsão clássica, erros de previsão, modelos Auto – Regressivos e formulações matemáticas tais como AR, ARMA, ARIMA e SARIMA.

O capítulo 3 aborda a descrição e modelagem do problema, apresentando identificação do modelo, estimação, previsão e validação do modelo. Já o capítulo 4, estão as conclusões e recomendações para futuras pesquisas, seguidas pelas referências bibliográficas e os anexos.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 O MERCADO AUTOMOBILÍSTICO BRASILEIRO

Com o surgimento da indústria de automóveis em série, no início do século XX, por Henry Ford, o mercado automotivo apresentou uma posição de destaque no setor industrial do mundo inteiro. No Brasil, em meados de 1950, a indústria de veículos teve seu início com incentivo de capital estrangeiro no país proposto por Juscelino Kubitschek, pelo fato de o setor automobilístico apresentar um multiplicador de investimento muito elevado.

Para se produzir um carro, necessita-se dos seus acessórios, tais como rodas, pneus, motores, amortecedores, vidro e faróis, gerando incentivo para criação de outras empresas. Isso fez com que a indústria automobilística criasse um estímulo para a criação das indústrias de autopeças.

No ano de 2005, segundo a Federação Nacional da Distribuição de Veículos Automotores (FENABRAVE), a indústria de produção de veículos alcançou a marca de 2 milhões de veículos produzidos. Em 2007, esse feito foi ainda maior, pois a frota nacional ocupou a 9ª posição, com 2.232.192 veículos, onde o 1ª foi ocupado pelos Estados Unidos com 16.121.776 veículos.

No início do ano de 2012 passou a vigorar a sétima edição do Proconve (Programa de Controle de Poluição do Ar por Veículos Automotores). Esse programa foi criado em 1986 pelo CONAMA (Conselho Nacional Do Meio Ambiente) com os objetivos de reduzir as emissões de veículos novos, desenvolver tecnologias nacional e melhorar a qualidade dos combustíveis. É aplicado para veículos leves e pesados, como os caminhões, enfrentaram ainda mais desafios afim de atender as novas legislações e se manter competitivos.

Para o atendimento da resolução do Proconve 7, as montadoras implementaram novos sistemas de motorização, combustão e exaustão dos combustíveis em seus veículos, conhecidos como Euro 5. Devido a isso, teve-se um ganho tecnológico e ambiental agregado no produto final, resultando em um custo de 10% a 20% dos veículos. Com a chegada da nova legislação Proconve 7 e o aumento no preço dos veículos devido à motorização Euro 5, a retração no mercado foi eminente. Isso ocorreu pelo fato dos caminhões serem bens de produção e a disponibilidade dos veículos Euro 3, mais poluentes e mais baratos, no mercado fez-se com que o fenômeno do *pre-buying* antecipou a renovação da frota, motivo pelo qual se torna muito importante conhecer a previsão demanda do volume total de produção de caminhões (SANTOS, 2012).

Segundo a Anfavea (2015), em 2013 o setor automobilístico apresentou um faturamento de US\$ 110,9 bilhões, possuindo uma participação de 23% no PIB nacional, ocupando o 7º lugar como produtor mundial e o 4º lugar no mercado interno. Com isso, registrou-se 31 fabricantes de veículos e 5.386 concessionárias além as 64 unidades industriais distribuídas no país, demonstrando ser um setor de muita importância para o país.

Na década de 2010, a indústria automobilística vem passando por um processo de transformação para acompanhar as mudanças e exigências econômicas de um mercado cada vez mais competitivo. Essas mudanças englobam a necessidade de novos investimentos para modernizar o parque industrial, novas exigências dos consumidores quanto à qualidade do produto, estratégias locais orientadas pelas mundiais, entre outros pontos que redirecionaram as perspectivas do mercado automobilístico brasileiro, delineando novos rumos para o mercado, impactando na competitividade setorial (RIBEIRO, 2014).

2.2 MODELO DE PREVISÃO CLÁSSICA

De acordo com Box e Luceño (1997), Morettin e Toloí (2004) e outros autores, uma série temporal, também podendo ser definida por uma série histórica, é uma sequência de dados obtidos em intervalos regulares de tempo, durante um período específico, e que apresentam uma dependência seriada. Para representar a série temporal é utilizada a notação Z , e no momento t será Z_t , em que $t = 1, 2, \dots, n$ indica o tamanho da série.

Os processos estocásticos são utilizados para descrever modelos de séries temporais, isto é, processos controlados por leis probabilísticas em uma série de dados distribuídos no tempo (ESPINDOLA, 2013). É uma maneira mais formal utilizada para dizer que uma série temporal é realização de um processo estocástico (GUARNIERI; SOUZA; ALMEIDA, 2009).

Segundo Morettin e Toloí (2004), uma das suposições mais frequentes que se faz a respeito de uma série temporal é que ela seja estacionária, ou seja, desenvolve-se no tempo, aleatoriamente ao redor de uma média constante, refletindo alguma forma de equilíbrio estável. Entretanto, a maior parte das séries que é encontrado na prática apresenta alguma forma de não-estacionariedade, que geralmente é justificada devido a presença de quatro componentes:

- *Tendência*: verifica o deslocamento da série ao longo do tempo.
- *Sazonalidade*: movimento ondulatório de curta duração, em geral, inferior a um ano; geralmente associada a mudança climática.

- *Ciclo*: movimento ondulatório que ao longo de vários anos tende a ser periódico.
- *Ruído Aleatório*: compreende a variedade intrínseca aos dados e não pode ser modelado.

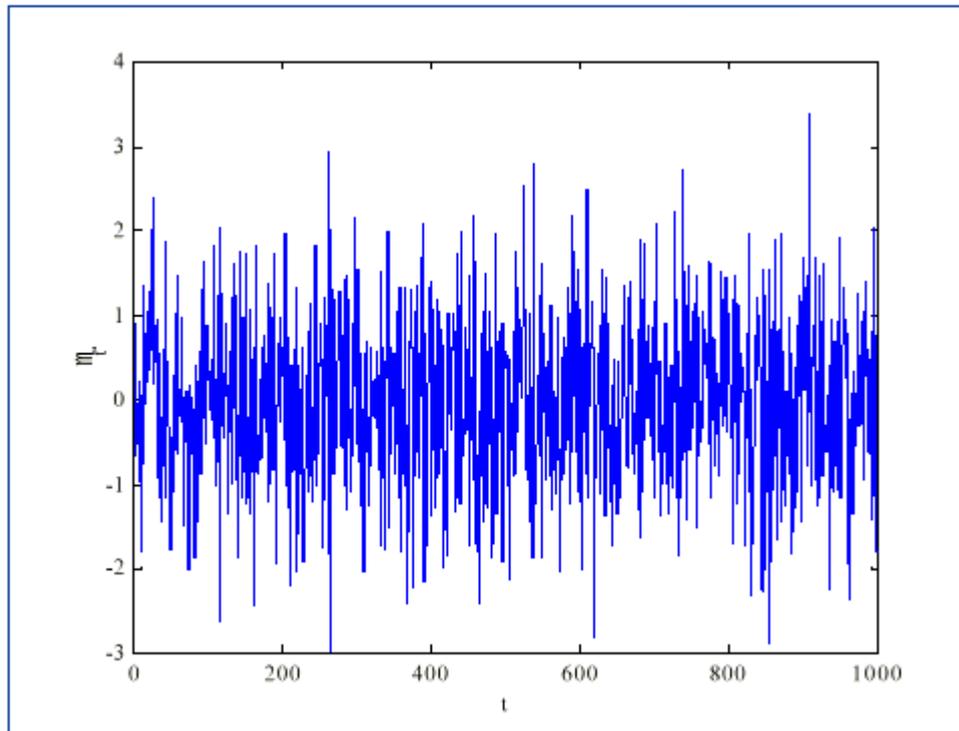
Os métodos clássicos de séries temporais são: Média Móvel, Ajustamento Exponencial, Tendência Linear, Tendência Não Linear (MORETTIN, 1987). As formas de previsão de séries temporais são baseadas na extrapolação de características transferidas entre seus relacionamentos possibilitando previsões atiladas caso o comportamento futuro se assemelhe com o passado (WHEEL, WRIGHT, 1985, GOOIJER, 2006).

Grande parte dos métodos utilizados em séries temporais, estão ligados a linearidade e a estacionaridade, como, por exemplo, os modelos Auto-Regressivos (AR) e médias móveis Auto-Regressivas (ARMA) (MORETTIN, 2006). Neste contexto, um processo é estacionário quando, distribuições unidimensionais não variam ao passar do tempo, sendo a média e a variância constantes (MORETTIN, 2006).

2.2.1 Modelos Auto-regressivos Lineares Estacionários

Modelos estacionários são aqueles em que sua média, variância e autocovariância são invariáveis em relação ao tempo. São classificados em: Auto-regressivos (AR); Médias Móveis (MA) e Auto-regressivo de Médias Móveis (ARMA) (TUBINO; LOPES; ZANINI, 2010). A Figura 3 exibe um ruído branco, processos em que todas as suas variáveis aleatórias são expressas por meio de uma distribuição normal de média 0 com variância constante e covariância nula, que é estacionário.

Figura 3 - Ruído Branco



Fonte: Morettin e Toloi (2006).

2.2.1.1 Modelo Auto-Regressivo (AR(p))

Chama-se de auto-regressivo de ordem p $AR(p)$ a Equação 1.

$$Z_t = \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + \dots + \varphi_p Z_{t-p} + a_t \quad (1)$$

Onde os $Z_{t-1}, Z_{t-2}, \dots, Z_{t-p}$ são termos que independem de a_t . A série Z_t possui valores que são combinações lineares dos p valores passados mais um termo a_t , no qual incorpora coisas na série até o tempo t que não é explicado pelos valores passados.

2.2.1.2 Modelo Médias Móveis (MA (q))

Tomando a série Z_t e fazendo os $\psi^0 s = -\theta^0 s$ no processo linear geral e considerando a série como finita, chama-se de médias-móveis de ordem q ($MA(q)$.) a Equação 2.

$$Z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (2)$$

Um vez em que Z_t é obtido aplicando os pesos $1, -\theta_1, -\theta_2, \dots, -\theta_q$, às variáveis $a_t, a_{t-1}, a_{t-2}, \dots, a_{t-q}$, então, movendo-se os mesmos pesos 1 unidade do tempo a frente assim sendo aplicado a eles as somas $a_{t+1}, a_t, a_{t-1}, \dots, a_{t-q+1}$ para obter Z_{t+1} .

2.2.1.3 Modelo Médias Móveis Auto Regressivas ARMA (p,q)

Considerando-se uma série descrita pelas partes auto regressiva e médias móveis, tem-se um modelo mais genérico de séries temporais apresentado na Equação 3.

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (3)$$

Diz-se que Z_t é um processo auto-regressivo médias-móveis de ordens p e q , respectivamente, com parâmetros ϕ 's e θ 's, ou abreviadamente ARMA(p, q).

Para um melhor entendimento, é necessário saber sobre Fator de Autocorrelação e Fator de Autocorrelação Parcial que alguns autores relatam. Segundo Souza e Camargo (1996), a Autocorrelação mede a correlação da série com ela mesma defasada de k unidades de tempo. Pode-se dizer esta função estabelece o quanto um valor no tempo t depende do valor no tempo $t-1$, em $t-2$ em $t-3$ e assim sucessivamente.

Souza e Camargo (1996) comentam que, a Função de Autocorrelação Parcial a extensão da Autocorrelação, ou seja, se for medido a correlação entre duas observações $Z(t)$ e $Z(t+k)$ eliminando a dependência dos termos intermediários, $Z(t+1)$, $Z(t+2)$, $Z(t+k-1)$, tem-se o que se denomina Autocorrelação Parcial, assim como na Equação 4.

$$\text{Cor} = [Z(t), Z(t+k) / Z(t+1), \dots, Z(t+k-1)] \quad (4)$$

Para o modelo geral ARMA (p, q), dado que a_t é independente de $Z_t - 1, Z_t - 2, \dots$, uma solução estacionária para Z_t existe, somente se, as raízes da equação característica AR, $\phi(x) = 0$, excederem uma unidade em módulo. Ao passo em que condições de estacionariedade são satisfeitas ($\psi_0=1, \psi_1 = -\theta_1 + \phi_1, \psi_2 = -\theta_2 + \phi_2 + \phi_1 \psi_1 \dots$), o modelo pode ser escrito como um processo linear geral com os pesos ψ_j determinados como na Equação 5.

$$\psi_j = -\theta_j \phi_p \psi_{j-p} + \dots + \phi_1 \psi_{j-1} \quad (5)$$

onde $\psi_j = 0$ para $j > 0$ e $\theta_j = 0$ para $j > q$.

No momento em que se assume a estacionariedade pode-se mostrar que a FAC (Fator de Autocorrelação) satisfaz a Equação 6.

$$\rho_k = \varphi_1 \rho_{k-1} + \varphi_2 \rho_{k-2} + \dots + \varphi_p \rho_{k-p}, \quad k > q \quad (6)$$

2.2.2 Modelos Auto-regressivos Lineares não Estacionários

Modelos não estacionários são aqueles que apresentam uma média e variância que variam com o tempo. São classificados em: Auto-regressivo Integrado de Médias Móveis (ARIMA) e Auto-regressivos Sazonais Integrados de Médias Móveis (SARIMA).

2.2.2.1 O Modelo Auto-Regressivo Integrado Médias Móveis (ARIMA(p,d,q))

Para as séries Z_t , com um número finito de diferenças, d , que convertem-se em estacionárias, são chamadas não-estacionárias homogêneas se obedecerem a Equação 7.

$$W_t = \Delta_d Z_t \quad (7)$$

Representa-se W_t por um modelo ARMA(p,q) como na Equação 8.

$$\varphi(X)W_t = \theta(X)a_t \quad (8)$$

Se W_t é uma diferença de Z_t , então Z_t é uma integral de W_t , portanto pode-se dizer que Z_t possui um modelo auto-regressivo-integrado-médias-móveis, ARIMA(p,d,q), apresentado na Equação 9.

$$\varphi(X)\Delta^d Z_t = \theta(X)a_t \quad (9)$$

De ordem (p,d,q), se p e q são as ordens de $\varphi(X)$ e $\theta(X)$, respectivamente.

No modelo, todas as raízes de $\varphi(X)$ estão fora do círculo unitário. Entretanto deve-se considerar a equivalência demonstrada na Equação 10,

$$\zeta(X)Z_t = \theta(X)a_t \quad (10)$$

donde $\zeta(X)$ é um operador auto-regressivo não-estacionário, de ordem $p+d$, com d raízes iguais a 1 (sobre o círculo unitário) e as restantes p estão fora do círculo unitário, logo deve-se representar como na Equação 11.

$$\zeta(X) = \varphi(X)\Delta^d = \varphi(X)(1 - X)^d \quad (11)$$

2.2.2.2 Modelo Médias Móveis Integradas Auto-regressivas Sazonais SARIMA (p,d,q) x (P,D,Q)

A metodologia Box-Jenkins investiga a autocorrelação entre os valores da série em diferentes instantes de tempo sucessivos. Os padrões de autocorrelação, em geral, possibilitam identificar um ou vários modelos possíveis para a série temporal (KHASHEI; BIJARI, 2010). Ao observar a autocorrelação dentro do período de um ano, considera-se uma estação de sazonalidade e as séries originais podem ser ajustadas por um modelo ARIMA sazonal (MARTINS; WERNER, 2014; MORETTIN; TOLOI, 2006).

Os modelos ARIMA sazonal (SARIMA), de ordem $(p, d, q) \times (P, D, Q)$ são constituídos por uma parte não-sazonal (p, d, q) e outra sazonal (P, D, Q) . Esses modelos são os mais requisitados para descrição de séries temporais sazonais, demonstrando sucesso em suas aplicações nos últimos 30 anos (CHEN; WANG, 2007).

Quando $\{Z_t\}$ apresenta um comportamento sazonal determinístico com período 12, um modelo que pode ser útil é apresentado na Equação 12.

$$Z_t = \mu_t + N_t, \quad (12)$$

Sendo que μ_t é uma função determinística periódica, ($\mu_t - \mu_{t-12} = 0$, ou $(1 - B^{12})\mu_t = 0$) e N_t é um processo estacionário que pode ser modelado por um ARIMA (p,q) .

Dessa maneira, N_t satisfaz a Equação 13.

$$\phi(B)N_t = \theta(B)a_t, \quad (13)$$

Sendo a_t é ruído branco e μ_t tem solução dada pela Equação 14.

$$\mu_t = \mu + \sum_{j=1}^6 \alpha_j \left[j \cos \frac{(2\pi jt)}{12} + \beta_j \sin \frac{(2\pi jt)}{12} \right] \quad (14)$$

Com $\mu, \alpha_j, \beta_j, j=1, \dots, 6$, constantes desconhecidas.

Assim, para um modelo sazonal determinístico, aplicando a diferença sazonal $(1 - B^{12})$ à Equação (14), obteve-se a Equação 15.

$$(1 - B^{12})Z_t = (1 - B^{12})\mu_t + (1 - B^{12})N_t \quad (15)$$

E de acordo com (13), tem-se

$$(1 - B^{12})Z_t = (1 - B^{12})N_t. \quad (16)$$

Substituindo (15) em (16), obteve-se

$$\phi(B)(1 - B^{12})Z_t = \theta(B)(1 - B^{12})a_t, \quad (17)$$

$$\phi(B)W_t = \theta(B)(1 - B^{12})a_t \quad (18)$$

Sendo $W_t = (1 - B^{12})Z_t$.

A obra de Morettin e Toloí (2006) é recomendada para caso de aprofundamento com mais detalhes.

2.3 ERROS DE PREVISÃO

A precisão da previsão é referida à quão perto as previsões chegam dos dados reais. Os indicadores são utilizados para verificar medidas de desempenho do nível de precisão que o modelo de previsão possui. Quando as previsões estão muito próximas dos dados reais, significam erros de precisão baixos, logo são mais aceitáveis. Quando os erros de precisão são maiores é sinal que o modelo de previsão deve ser alterado ou ajustado.

Considere um erro de previsão entre X real e o valor previsto \hat{X} representado por:

$$e = X - \hat{X}.$$

Para analisar e decidir o melhor modelo de previsão que se ajusta aos dados deve-se levar em conta os indicadores listados nas seções seguintes.

2.3.1 Erro Absoluto Médio – MAD

O erro absoluto médio MAD (*mean absolute deviation*) é a soma dos desvios absolutos dos períodos dividida pelo número de períodos. É representado pela fórmula:

$$MAD = \frac{\sum_{i=1}^n |X - \widehat{X}_i|}{n} \quad (19)$$

Se MAD for pequeno, os dados reais seguem estreitamente as previsões da variável dependente e o modelo de previsão fornece previsões acuradas.

2.3.2 Erro Percentual Absoluto Médio – MAPE

O erro percentual absoluto médio MAPE (*mean absolute percentage error*) representa a média percentual da divisão entre erro de previsão e o valor real.

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{X_i - \widehat{X}_i}{X_i} \right| (100)}{n} \quad (20)$$

Valores pequenos para o MAPE determinam precisão nos dados previstos.

2.3.3 Erro Quadrático Médio – EQM

O erro quadrático médio é dado pela expressão:

$$EQM = \frac{\sum_{i=1}^n (X_i - \widehat{X}_i)^2}{n} \quad (21)$$

De uma forma geral, quando os valores do erro forem pequenos, os dados reais seguem estreitamente as previsões da variável dependente e o modelo de previsão fornece previsões acuradas

2.3.4 Raiz Unitária

Segundo Fava e Cati (1995) a origem da discussão sobre a raiz unitária nas séries está no debate sobre a estacionaridade ou não da tendência, sendo que grande parte dos dados utilizados na análise empírica é em forma de série temporal. Uma série com uma tendência estocástica se diferencia de outra com uma tendência determinística, pois as mudanças na

mesma deixam de ter um caráter transitório e passam a apresentar um caráter permanente (GUJARATI, 2000). Portanto, a presença da raiz unitária é relevante para o processo de verificação de várias teorias.

A utilização dos modelos de regressão envolvendo as séries temporais não estacionárias pode conduzir ao problema que se convencionou chamar de espúria, isto é, quando se tem um alto R^2 sem uma relação significativa entre as variáveis (HARRIS, 1995). Isto ocorre devido ao fato de que a presença de uma tendência, decrescente ou crescente, em ambas as séries leva a um alto valor do R^2 , mas não necessariamente, a presença de uma relação verdadeira entre séries (GUJARATI, 2000).

Basicamente, a presença de raiz unitária na série temporal conduz a resultados viesados, invalidando os pressupostos da estatística clássica de que a média e a variância são constantes ao longo do tempo, e, com isto, mascarando o relacionamento entre duas, ou mais, variáveis. Detectada a presença de raiz unitária, então se deve trabalhar com as séries temporais diferenciadas e não em nível, ou seja, a tendência precisa ser removida.

Seja um processo estocástico de raiz unitária:

$$Y_t = \rho Y_{t-1} + \varepsilon_t, \quad -1 \leq \rho \leq 1,$$

onde ε_t é um termo de erro do ruído branco.

Quando $\rho = 1$ (raiz unitária), o processo gerador da série Y_t é o passeio aleatório e a série é não – estacionária.

Subtraindo-se, Y_{t-1} dos dois lados da igualdade, escreve-se:

$$\begin{aligned} Y_t - Y_{t-1} &= \rho Y_{t-1} - Y_{t-1} + \varepsilon_t \\ \therefore \Delta Y_{t-1} &= \delta Y_{t-1} + \varepsilon_t, \end{aligned} \quad (22)$$

sendo $\delta = \rho - 1$.

Um procedimento de teste da raiz unitária pode ser aplicado à equação (21). Dessa forma, estima-se os parâmetros da equação (21) e testa-se se $\delta = 0$. Se $\delta = 0$, $\rho = 1 \Rightarrow$ raiz unitária e a série original é não-estacionária. Entretanto, a estatística *t Students* não se aplica nesse caso (GUJARATI, 2006).

Usa-se o teste de Dickey-Fuller, que avalia se o valor da estatística *t* estimado para o parâmetro Y_{t-1} segue a distribuição da estatística τ . Dessa forma, o teste de Dickey-Fuller verifica a hipótese nula ($\delta = 0$) é rejeitada ou não em determinados níveis de significância estatística, conforme valores tabulados (GUJARATI, 2006). Se o valor em módulo de $t < |\text{estatística } \tau|$, não é possível rejeitar a hipótese nula ou seja, a hipótese de não-estacionariedade.

À estimação da equação (21), devem se juntar a estimação das variantes possíveis do processo de passeio aleatório e respectivos testes de hipóteses sobre $\delta = 0$:

$$\Delta Y_t = \beta_1 + \delta Y_{t-1} + \varepsilon_t \quad (23)$$

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \varepsilon_t \quad (24)$$

sendo t é o tempo ou a variável de tendência.

Nessas duas variantes [(23) e (24)], $\delta = 0$ significa que há raiz unitária ou seja, a série Y_t é não-estacionária. A hipótese alternativa é que $\delta < 0$, significando $\rho - 1 < 0$ ou seja $\rho < 1$ (observe que $\rho > 1$ não é aceita por implicar em explosão de valores da série).

Segundo Gujarati (2006), se a hipótese nula for rejeitada, isso significa que:

- Y_t é uma série temporal estacionária com média zero: equação (22)
- Y_t é uma série temporal estacionária com média diferente de zero: equação (23)
- Y_t é uma série estacionária em torno de tendência determinística: equação (24)

3 DESCRIÇÃO E MODELAGEM DO PROBLEMA

Primeiramente, observou-se uma série de dados de produção anual de caminhões no território nacional, no período de 1957 a 2014, fornecidos pela ANFAVEA (Anexo A). Com o auxílio do *software* GRETL foi possível abordar o problema de forma quantitativa e usando a modelagem como método de pesquisa. Sendo assim, torna-se uma pesquisa aplicada visando o proporcionar melhorias para a literatura atual.

3.1 ETAPAS PARA O DESENVOLVIMENTO DO MODELO DE PREVISÃO AUTO-REGRESSIVO

Segundo Morettin e Tolo (1986; 2006), o desenvolvimento de um modelo de previsão deve seguir (respectivamente) as etapas: Identificação do modelo, estimação e a previsão.

3.1.2 Identificação do modelo

A identificação de modelos realiza-se em dois passos:

1. Obteve-se estimativas preliminares, α , β_j , de μ , α_j , $e\beta_j$, $j = 1, \dots, 6$, por meio de uma análise de regressão de Z_t sobre 1, $\text{sen}\frac{2\pi jt}{12}$ e $\text{cos}\frac{2\pi jt}{12}$, $j = 1, \dots, 6$.
2. Calcula-se os resíduos, Equação 25.

$$N_t = Z_t - \mu - \sum_{j=1}^6 [\alpha_j \cos\frac{2\pi jt}{12} + \beta_j \text{sen}\frac{2\pi jt}{12}] \quad (25)$$

e examina-se as funções de auto correlação e auto correlação parcial para identificar um modelo ARIMA (p,q) para N_t .

3.1.3 Estimação

A estimação de máxima verossimilhança dos parâmetros μ , α_j , β_j , $\theta_j e \phi_k$, $i = 1, \dots, 6$, $j = 1, \dots, q$ e $k = 1, \dots, p$, foi obtida na estimação dos parâmetros de um modelo inicial ARIMA(2,1,4). Após a estimação e o ajuste do modelo, chegou-se ao que possui a melhor aderência: ARIMA (2,1,3).

3.1.4 Previsão

As previsões de valores futuros Z_{t+h} , dados Z_1, \dots, Z_t , são obtidos notando a igualdade entre as Equações 26 e 27.

$$Z_t(h) = E[Z_{t+h}|Z_{t-1}, \dots] \quad (26)$$

$$\mu_{t+h} + E[n_t \vee Y_t, Y_{t-1}, \dots] \quad (27)$$

Observa-se então a Equação 28.

$$Z_t(h) = \mu_{t+h} + N_t(h), \quad (28)$$

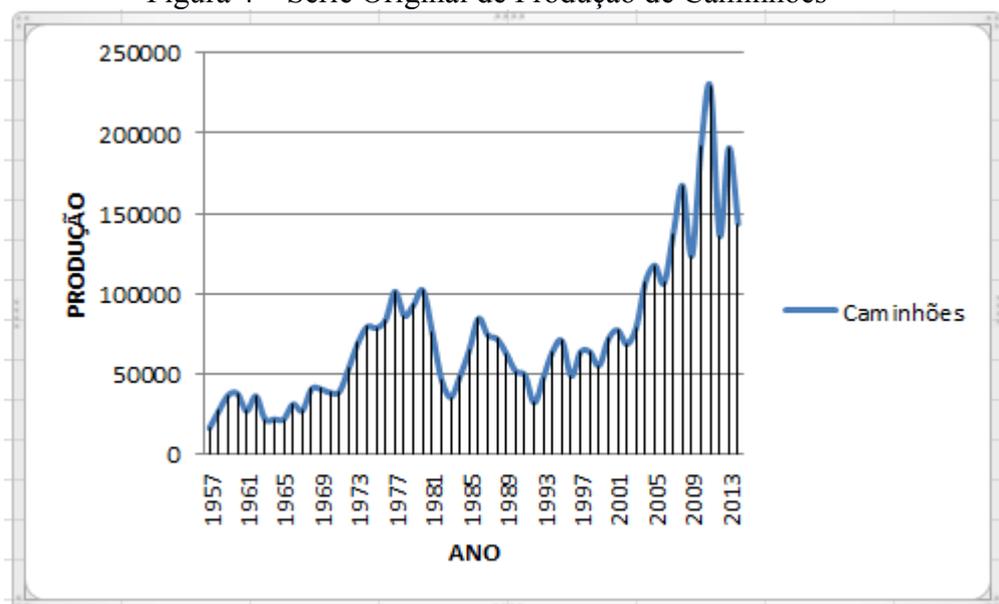
Sendo que μ_{t+h} e $N_t(h)$ são cálculos utilizando os modelos (26) e (27), respectivamente.

Na prática, esses modelos não tem parâmetros conhecidos, então utiliza-se μ_t no lugar de μ_{t+h} , que é obtido por meio da equação (28), substituindo cada um dos parâmetros por seus estimadores de mínimos quadrados e substituindo N_t por $N_t = Z_t - \mu_t$.

3.2 ANÁLISE DA SÉRIE

A primeira etapa da metodologia de Box e Jenkins consiste em identificar os modelos apropriados. Dessa forma, o primeiro passo será a análise do gráfico de tempo da série de produção de caminhões, do ano de 1957 a 2014, encontrado na Figura 4 e a Tabela 1 com os valores está apresentada no Anexo A. A análise deste gráfico indica uma linha de tendência crescente, pois há alteração na média e na variância, revelando que a série é não-estacionária.

Figura 4 – Série Original de Produção de Caminhões



Fonte: Anuário 2015 ANFAVEA.

Como a análise visual não é comprobatória, recorreu-se a análise do teste da raiz unitária, que debate a estacionariedade ou não da tendência. No teste da raiz unitária de Dickey – Fuller, é observado os valores assintóticos acima de 5% no teste com constante e também no com constante e tendência, ou seja, para o teste com constante o p-valor é de 0,6828, já para com constante e tendência o p-valor é de 0,1569. Desta forma, com esses resultados, a série indica tendência e, portanto, não-estacionária. No Quadro 1 encontra-se o resultado do teste da raiz unitária para a série original com os p-valores em negrito.

Quadro 1 –Teste da Raiz Unitária

Teste Aumentado de Dickey-Fuller para caminhões
incluindo 4 defasagens de (1-L)caminhões (o máximo foi 10)
dimensão de amostragem 53
hipótese nula de raiz unitária: $a = 1$

teste com constante

modelo: $(1-L)y = b_0 + (a-1)*y(-1) + \dots + e$

coeficiente de 1ª ordem para e: -0,008

diferenças defasadas: $F(4, 47) = 5,541 [0,0010]$

valor estimado de $(a - 1)$: -0,106735

estatística de teste: $\tau_c(1) = -1,18617$

p-valor assintótico 0,6828

com constante e tendência

modelo: $(1-L)y = b_0 + b_1*t + (a-1)*y(-1) + \dots + e$

coeficiente de 1ª ordem para e: 0,085

diferenças defasadas: $F(3, 48) = 5,825 [0,0018]$

valor estimado de $(a - 1)$: -0,329557

estatística de teste: $\tau_{ct}(1) = -2,917$

p-valor assintótico 0,1569

Fonte: Elaborado pelo Autor

Foi feita a análise da função de Autocorrelação (FAC) e função de Autocorrelação Parcial (FACP) da série original, que também auxiliou na verificação da estacionariedade e na definição do modelo mais apropriado. De acordo com a literatura, foi calculado um valor de referência utilizando a seguinte fórmula $c = \mp \frac{2}{\sqrt{n}}$, sendo n é quantidade da amostra e c o valor absoluto de referência. A série que está sendo alvo do estudo contém $n = 57$, e portanto $c = 0,264906$.

Observando a FAC, referente a série original de produção de caminhões, é notável a não-estacionariedade da série, pois na defasagem 11 o valor do FAC praticamente zerou, e depois esse valor não poderia ter aumentado como pode ser observado a partir da defasagem 14.

No Quadro 2 encontra-se os resultados da Função de Autocorrelação para a série de caminhões, e em negrito a defasagem 11 e 14.

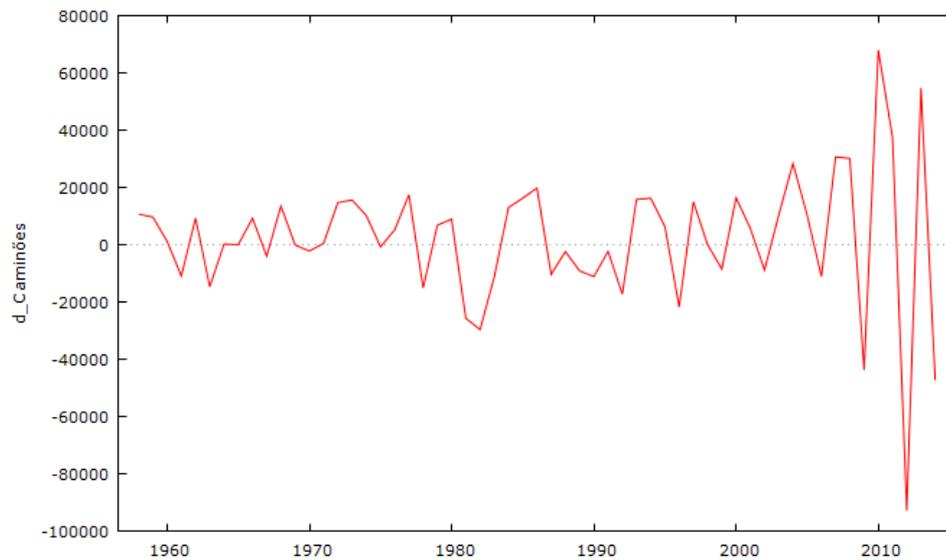
Quadro 2 – Teste de Autocorrelação para Série Original Caminhões

Função de Autocorrelação para Caminhões			
***, **, * indicam significância aos níveis de 1%, 5% e 10%			
usando erro padrão $1/T^{0,5}$			
Defas.	FAC	FACP	Estat. Q [p-valor]
1	0,8324 ***	0,8324 ***	42,3007 [0,000]
2	0,7361 ***	0,1407	75,9690 [0,000]
3	0,6944 ***	0,1687	106,4764 [0,000]
4	0,5121 *** -	0,4214 ***	123,3758 [0,000]
5	0,4094 ***	0,0465	134,3808 [0,000]
6	0,3660 ***	0,0879	143,3432 [0,000]
7	0,2622 **	0,0195	148,0326 [0,000]
8	0,1691	-0,1768	150,0223 [0,000]
9	0,1112	-0,0803	150,9000 [0,000]
10	0,0519	0,0742	151,0953 [0,000]
11	-0,0233	-0,0396	151,1356 [0,000]
12	-0,0639	-0,0281	151,4445 [0,000]
13	-0,0713	0,0387	151,8377 [0,000]
14	-0,1106	-0,0210	152,8058 [0,000]
15	-0,1394	-0,0592	154,3780 [0,000]
16	-0,1313	-0,0045	155,8056 [0,000]
17	-0,1346	0,0732	157,3436 [0,000]
18	-0,1305	0,0496	158,8260 [0,000]

Fonte: Elaborado pelo Autor

Para eliminar a tendência, recorreu-se a transformação da série original, por meio de uma diferença, a qual foi suficiente para a série temporal ficar estacionária. Observa-se que, a série passou a oscilar em torno de uma média constante, conforme a Figura 5.

Figura 5 – Série diferenciada de Produção de Caminhões



Fonte: Elaborado pelo autor

O teste da raiz unitária foi recorrido e em seus resultados é notável que a série original (não-estacionária) se tornou estacionária por meio de apenas uma diferenciação. O teste da raiz unitária mostrou que, ambos valores dos p-valores ficaram menores que 5%, comprovando que a série se tornou estacionária. No Quadro 3 encontra o resultado do teste da raiz unitária para a série diferenciada e os seus p-valores em negrito.

Quadro 3 – Teste da Raiz Unitária para Série Diferenciada

Teste Aumentado de Dickey-Fuller para $d_{\text{caminhões}}$
incluindo 3 defasagens de $(1-L)d_{\text{caminhões}}$ (o máximo foi 10)
dimensão de amostragem 53
hipótese nula de raiz unitária: $a = 1$
teste com constante
modelo: $(1-L)y = b_0 + (a-1)*y(-1) + \dots + e$
coeficiente de 1ª ordem para e: -0,024
diferenças defasadas: $F(3, 48) = 4,762 [0,0055]$
valor estimado de $(a - 1)$: -1,47177
estatística de teste: $\tau_c(1) = -3,75537$
p-valor assintótico 0,003416
com constante e tendência
modelo: $(1-L)y = b_0 + b_1*t + (a-1)*y(-1) + \dots + e$
coeficiente de 1ª ordem para e: -0,025
diferenças defasadas: $F(3, 47) = 4,594 [0,0067]$
valor estimado de $(a - 1)$: -1,51591
estatística de teste: $\tau_{ct}(1) = -3,59122$
p-valor assintótico 0,0305

Fonte: Elaborado pelo Autor

O próximo procedimento é a identificação dos modelos significativos que representam o comportamento da série. Morretin e Tolo (2004) sugerem que na etapa de identificação do modelo, um procedimento que muitas vezes é utilizado é identificar não só um único modelo, mas alguns modelos que serão então estimados e verificados. Se o propósito é previsão, escolher-se-á entre os modelos ajustados o melhor, ou seja, o que fornecer menor erro quadrático médio de previsão.

3.3 IDENTIFICAÇÃO DO MODELO ARIMA

Para se identificar o modelo ARIMA (p,d,q) é necessário definir os filtros p, d e q. Como a série se tornou estacionária com apenas uma diferenciação, tem-se $d=1$ e os filtros p e q são definidos com o teste de Função de Autocorrelação da primeira diferença da série.

Partiu-se de um pressuposto com o valor de referência de $c = \frac{2}{\sqrt{n}}$, onde $n=56$, pois agora a série foi diferenciada uma vez. Obteve-se o valor de referência $c = 0,2672$ para análise da série estacionária. Segundo a literatura, para se determinar o filtro Auto-Regressivo (p) e o filtro de Médias-Móveis(q), observa-se o primeiro valor abaixo do valor de referência, e depois, utiliza-se os valores da posição a qual o valor foi encontrado. É importante salientar que trabalha-se com o valor absoluto, sendo assim, para o FACP o valor se encontra na segunda posição e para o FAC a quarta posição. Sendo assim, tem-se o valor de $p=2$ e $q=4$.

Uma observação importante é que para o fator q analisado, o lag (posição) 2 é o primeiro abaixo do valor de referência deveria ser escolhido para atribuir valor ao filtro q, porém isso não ocorreu. Optou-se a escolha do lag 4, devido que o valor voltou a subir no lag 3. No entanto, após o lag 4 o valor tendeu-se a zero e não voltou a subir, por isso foi escolhido o lag 4 para assumir o valor do filtro de Médias – Móveis (q).

Para melhor visualização os valores selecionados com o teste de Função de Autorocorelação para a série diferenciada estão em negrito no Quadro 4. Com isso, tomou-se um ARIMA (2,1,4) de início e para melhor abordagem e aderência ao modelo, foi feito algumas iterações no modelo escolhido e se obteve um ARIMA (2,1,3) modelo mais significativo para a série em estudo.

Quadro 4 – Função de Autocorrelação da Série Diferenciada

Função de Autocorrelação para d caminhões			
Defas.	FAC	FACP	Estat. Q [p-valor]
1	-0,3674 ***	-0,3674 ***	8,1075 [0,004]
2	-0,0750	-0,2427 *	8,4512 [0,015]
3	0,3140 **	0,2341 *	14,5899 [0,002]
4	-0,2503 *	-0,0636	18,5648 [0,001]
5	-0,0632	-0,1591	18,8230 [0,002]
6	0,0932	-0,1147	19,3961 [0,004]
7	0,0175	0,1162	19,4167 [0,007]
8	-0,0721	0,0210	19,7735 [0,011]
9	-0,0102	-0,1100	19,7809 [0,019]
10	0,0690	-0,0411	20,1219 [0,028]
11	-0,0481	0,0424	20,2909 [0,042]

Fonte: Elaborado pelo Autor

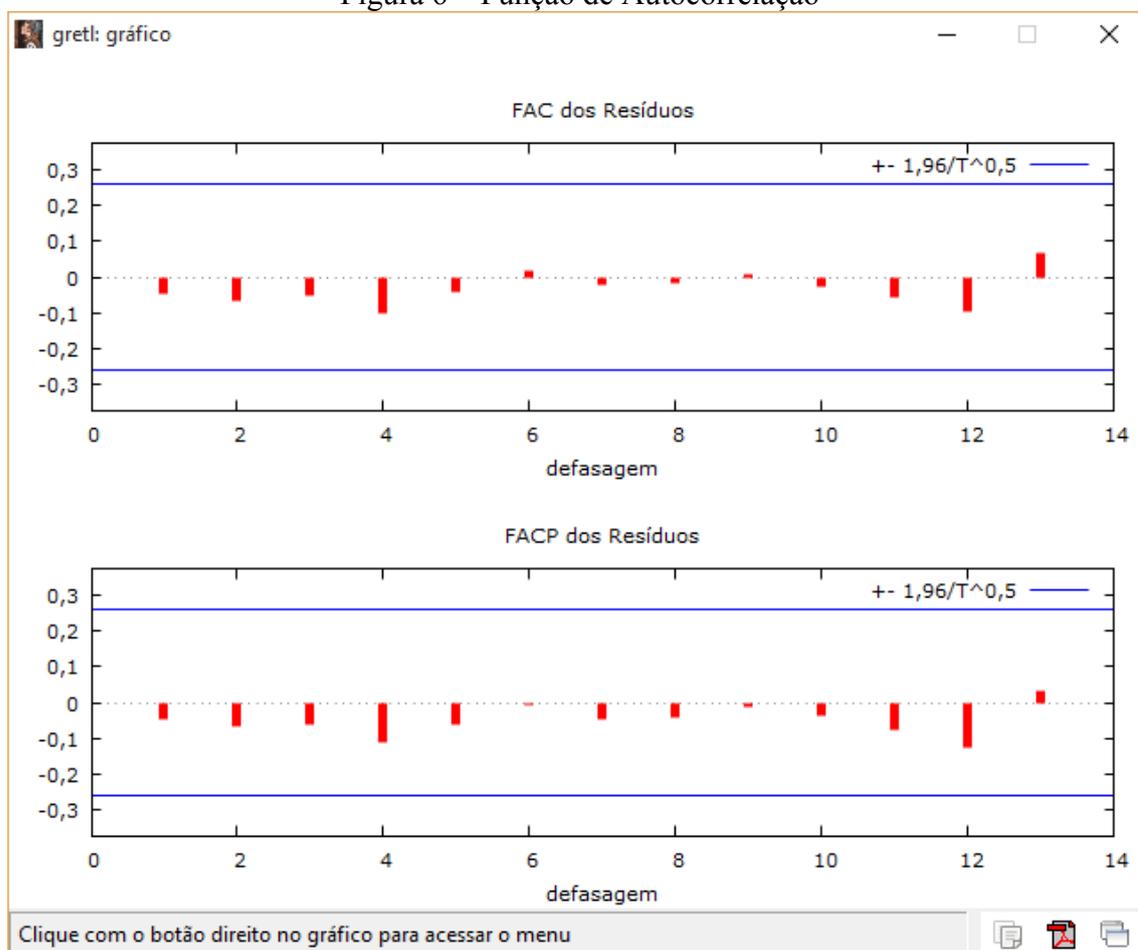
3.4 VALIDAÇÃO DO MODELO

Para Fava (2000), as formas de verificação frequentemente utilizadas são: análise dos resíduos e avaliação da ordem do modelo. Após a identificação dos parâmetros do modelo, foi possível analisar os resíduos para aderência e em seguida realizar a previsão.

3.4.1 Análise de Resíduos

Segundo Werner e Ribeiro (2003), as estimativas devem apresentar-se de maneira tal, com que, suas auto-correlações não podem demonstrar significância, ou seja, seus resíduos devem possuir comportamentos aleatórios, caracterizando a adequação do modelo. Para atender aos pressupostos da metodologia, procedeu-se ao teste para a verificação da condição de Autocorrelação dos resíduos gerados pelo ARIMA (2,1,3). Para tanto, a Figura 6 apresenta o correlograma da FAC e FACP dos resíduos do modelo.

Figura 6 – Função de Autocorrelação



Fonte: Elaborado pelo autor

É percebido pela Figura 6, que não há coeficientes de Autocorrelação fora do intervalo de aceitação. A existência de coeficientes fora do intervalo de aceitação é considerada uma situação indesejada na modelagem de séries temporais utilizando os modelos ARIMA (p,d,q), pois determinados padrões existentes na série não teriam sido capturados pelo modelo. Este teste comprovou que o modelo desenvolvido poderá ser utilizado. De um outro modo, para se comprovar que o modelo poderá ser utilizado para fazer a previsão, é verificar o p-valor na análise dos resíduos.

Se o p-valor é maior que 5% H_0 é aceito, se não H_1 é aceito, sendo H_0 : resíduo é normal e H_1 : resíduo não é normal. A Figura 9 e a Figura 10 no Anexo B, encontra-se o histograma da normalidade dos resíduos com o valor de Qui-quadrado e o gráfico dos resíduos, respectivamente. No Quadro 5 estão os resultados da análise dos resíduos, nota-se que, o valor do p-valor que está em negrito é maior que 5%, comprovando o aceite da normalidade.

Quadro 5 – Teste da Normalidade dos Resíduos

Distribuição de frequência para uhat17, observações 2-58				
número de classes = 7, média = 1489,91, desvio padrão = 20056,6				
intervalo	pt. médio	freqüência	rel.	acum.
< -42763,	-50375,	2	3,51%	3,51% *
-42763, -	-27540, -35152,	2	3,51%	7,02% *
-27540, -	-12317, -19928,	6	10,53%	17,54% ***
-12317, -	2906,3 -4705,2	19	33,33%	50,88% *****
2906,3 -	18129, 10518,	18	31,58%	82,46% *****
18129, -	33353, 25741,	7	12,28%	94,74% ****
>= 33353,	40964,	3	5,26%	100,00% *

Teste para a hipótese nula de distribuição normal:
 Qui-quadrado(2) = 3,051 com **p-valor 0,21747**

Fonte: Elaborado pelo Autor

3.4.2 Avaliação da ordem do modelo

Para a validação da ordem do modelo tem-se que principalmente verificar seus parâmetros em excesso, em outras palavras, sua parcimoniosidade. Segundo Werner e Ribeiro (2003), se o valor absoluto for pequeno comparado ao seu erro padrão, este pode ser descartado, pois não é significativo. Sua verificação é realizada tendo como base o erro padrão dos coeficientes, sendo phi o coeficiente relacionado ao filtro p (Auto – Regressivo) e theta o coeficiente que relaciona o filtro q (Média – Móveis). No Quadro 6 tem-se os resultados do ARIMA (2,1,3), verificou-se que os valores de p-valores que estão em negrito são menores que 0,05 (nível de significância), sendo estatisticamente significativos.

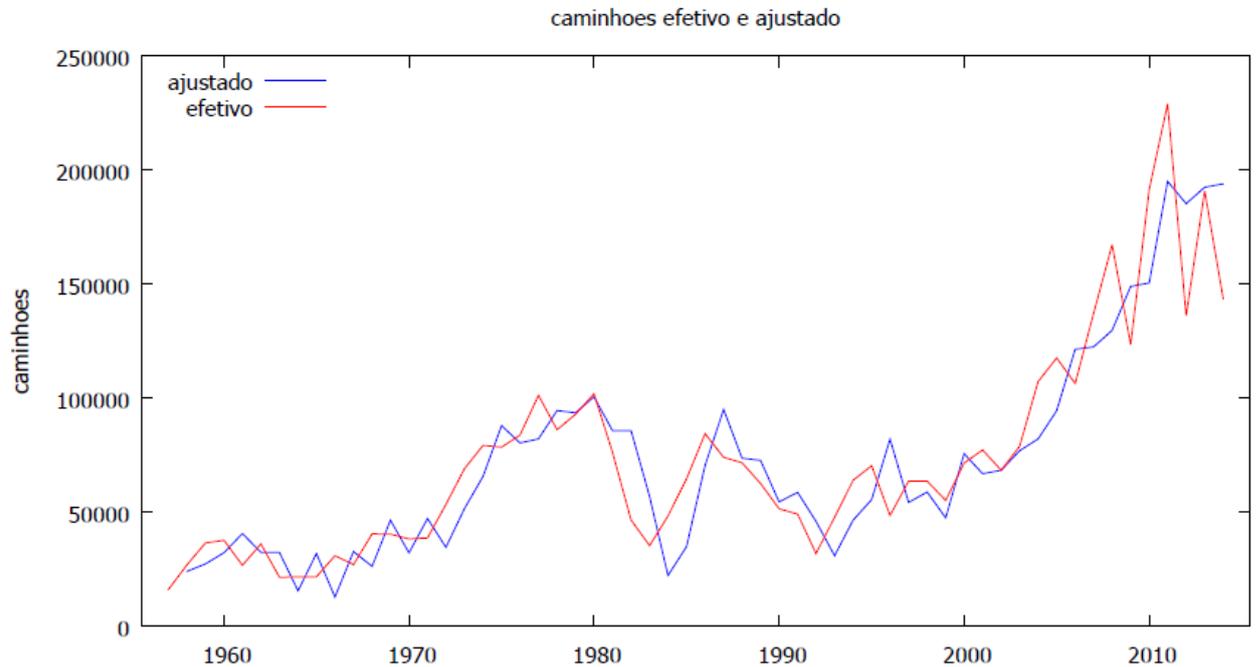
Quadro 6 – Resultados do ARIMA(2,1,3)

	<i>Coefficiente</i>	<i>Erro Padrão</i>	<i>z</i>	<i>p-valor</i>	
phi_1	-0,900626	0,180728	-4,9833	<0,00001	***
phi_2	-0,756319	0,201279	-3,7576	0,00017	***
theta_1	0,774618	0,207791	3,7279	0,00019	***
theta_2	0,634473	0,208067	3,0494	0,00229	***
theta_3	0,427324	0,145627	2,9344	0,00334	***

Fonte: Elaborado pelo Autor

Como o modelo ARIMA (2,1,3) está com uma boa aderência de acordo com a literatura, com o auxílio do *software* Gretl foi gerado um gráfico de ajuste com o tempo, esse gráfico é exibido na Figura 7.

Figura 7 – Gráfico Ajuste com o tempo



De forma visual, Figura 7, observa-se a aderência do modelo em função dos dados preditos. Com isso, realizou-se a previsão para os próximos 10 anos, e a matriz com seus valores preditos são mostrados no Quadro 7.

Quadro 7 – Valores de Produção Preditos pelo Gretl.

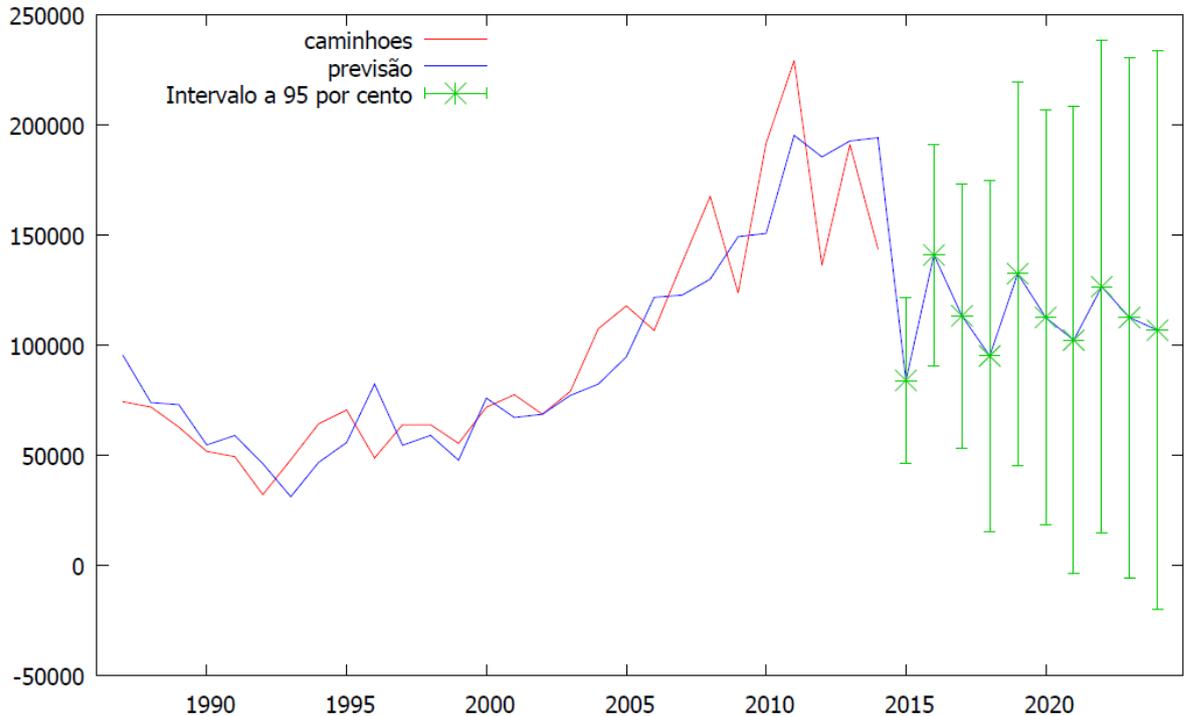
Para intervalos de confiança de 95%, $z(0,025) = 1,96$

Obs.	caminhões	Previsão	erro padrão	intervalo a 95%
2015	indefinido	83880,6	19277,8	(46096,8, 121664,)
2016	indefinido	140845,	25603,0	(90664,5, 191026,)
2017	indefinido	113227,	30561,1	(53328,6, 173126,)
2018	indefinido	95017,3	40718,7	(15210,2, 174824,)
2019	indefinido	132306,	44449,3	(45186,7, 219425,)
2020	indefinido	112495,	48059,7	(18300,2, 206691,)
2021	indefinido	102135,	54053,3	(-3807,38, 208078,)
2022	indefinido	126449,	57156,5	(14424,1, 238474,)
2023	indefinido	112387,	60307,3	(-5813,21, 230587,)
2024	indefinido	106663,	64653,1	(-20055,1, 233380,)

Fonte: Elaborado pelo Autor

O modelo se mostrou aderente em função dos dados preditos e seus valores estão dentro dos intervalos de confiança. Para melhor visualização, a Figura 8 mostra um gráfico dos valores reais e os valores preditos.

Figura 8 – Gráfico de Previsão



Fonte: Elaborado pelo autor

Com a crise econômica atual essa previsão apresenta uma lógica confiável, pois economistas comentam que a produção total de caminhões em 2015 deve recuar em 46%, ou seja, terá uma produção de 77 mil unidades. Segundo a Anfavea (2015), de janeiro até outubro deste ano foram produzidos 66.130 caminhões, portanto ao término do ano, o valor de produção de caminhões será próximo ao predito pelos economistas.

O modelo previu uma produção de caminhões para 2015 de aproximadamente 83.880 unidades, com intervalo de confiança de 95%, ou seja, esse valor apresentou-se no intervalo entre 46.096 a 121.664 unidades produzidas.

Por se tratar de um bem de produção, o volume de produção é altamente relacionado a fatores econômicos e políticos responsáveis pela variação de produção no decorrer dos anos. No entanto, o modelo se mostrou capaz de realizar previsões confiáveis em um período de incerteza e instabilidade.

4 CONCLUSÕES E RECOMENDAÇÕES PARA FUTURAS PESQUISAS

Este capítulo aborda a verificação dos objetivos e recomendações para as futuras pesquisas.

4.1 VERIFICAÇÃO DOS OBJETIVOS

O objetivo do trabalho foi apresentar uma previsão para a série de volume de produção de caminhões no Brasil. Para esta tarefa, foram realizados os testes contidos neste trabalho para alcançar o modelo mais adequado. O problema foi abordado de forma quantitativa, sendo o método da pesquisa a modelagem, utilizando o modelo de previsão Box – Jenkins.

É notável que os objetivos gerais e os específicos foram completamente atendidos, pois o modelo ARIMA (2,1,3) resultou-se em uma excelente aderência quando comparado aos dados reais, como é apresentado na Figura 8. É importante ressaltar que, há maior dispersão de erro do predito comparado com o real no ano de 2012. Isso ocorreu devido a entrada da nova legislação Proconve 7 e o aumento no preço dos veículos devido à motorização Euro 5. Pelo fato dos caminhões serem bens de produção e a disponibilidade dos veículos Euro 3, mais baratos e mais poluentes, fez-se no mercado o fenômeno *pre-buying* antecipando a renovação da frota. Devido a isso, a produção de caminhões em 2012 sofreu uma grande queda.

Conclui-se que, os resultados desse trabalho indicaram que a metodologia Box e Jenkins é um modelo apropriado para previsão de demanda de produção de caminhões na indústria brasileira. Portanto, é uma importante ferramenta gerencial para tomada de decisões, tanto para indústria como para outros setores, para realizações de previsões.

4.2 RECOMENDAÇÕES PARA FUTURAS PESQUISAS

- Recomenda-se utilizar a combinação de múltiplos métodos de previsão afim de uma melhor robustez do dado predito em relação ao real.
- Aplicar o modelo auto-regressivo em outras séries temporais, tais como, veículos leves e ônibus.

- Realizar previsões para produção de automóveis utilizando uma análise de intervenção, a qual, segundo Morretin e Toloí (2004), tem por objetivo avaliar o impacto de algum evento atípico no comportamento da série.
- Realizar um estudo num período de incertezas e instabilidade, incluindo o efeito da crise na modelagem.

REFERÊNCIAS BIBLIOGRÁFICAS

ANFAVEA, Associação Nacional dos Fabricantes de Veículos Automotores. São Paulo, 2015: Anuário da Indústria Automobilística Brasileira, 2015. Disponível em: <<http://www.anfavea.com.br/anuario.html>>. Acesso em: 21 Agost. 2015.

AUTOMOTIVEBUSINES- Disponível em: <<http://www.automotivebusiness.com.br/noticia/23050/fenabreve-projeta-2016-pior-que-2015>> . Acesso em : 25 Novembro de 2015.

BOX, G.; LUCEÑO, A. **Statistical control by monitoring and feedback adjustment**. New York: Wiley Interscience, 1997.

CHEN, K. Y.; WANG, C. H. A hybrid SARIMA and support vector machines in forecasting the production values of the machinery industry in Taiwan. **Expert Systems with Applications**, v. 32, n. 1, p. 254-264, 2007.

CONAMA - CONSELHO NACIONAL DO MEIO AMBIENTE - Disponível em: <http://www.mma.gov.br/port/conama/legiano1.cfm?codlegitipo=3&ano=1986> Acesso em : 17 Nov. 2015.

CONSONI, F. L.; QUADROS CARVALHO, R. Desenvolvimento de produtos na indústria automobilística brasileira: perspectivas e obstáculos para a capacitação local. **Revista de Administração Contemporânea**, v. 6, n. 1,p. 39-62, 2002.

DANESE, P.; KALCHSCHMIDT, M. The impact of forecasting on companies' performance : Analysis in a multivariate setting. **International Journal of Production Economics**, v. 133 (1), p. 458–469, 2011b. <http://dx.doi.org/10.1016/j.ijpe.2010.04.016>.

ECONOMIA E MERCADO – Disponível em : <<https://economiaemercado.wordpress.com/2015/09/21/piora-de-cenario-atinge-a-expectativa-para-o-setor-automotivo-em-2016/>>_ Acesso em: 25 de nov. de 2015.

FAVA, V.L. Análise de Séries de Tempo. In: Vasconcellos, M.A.S.; Alves, D. **Manual de econometria**: nível intermediário. São Paulo: Atlas, 2000.

FENABRAVE – FEDERAÇÃO NACIONAL DA DISTRIBUIÇÃO DE VEÍCULOS AUTOMOTORES. Disponível em: <<http://www.fenabrave.com.br/>>. Acesso em: 13 nov. 2015.

WALTER, O. M. F. C.; HENNING, E.; MORO, G.; WAYNE SAMOHYL, R. W. - **Aplicação de um modelo SARIMA na previsão de vendas de motocicletas**
Exacta, São Paulo, v. 11, n. 1, p. 77-88, 2013.

ESPÍNDOLA, A. M. S. 2013. **Inteligência competitiva e modelos de séries temporais para previsão de consumo**: o estudo de uma empresa do setor metalúrgico. 2013. 134 f. Dissertação (Mestrado em Administração – Administração da produção) – Universidade Caxias do Sul, Caxias do Sul, 2013.

GUJARATI, D.M. **Econometria básica**. 3. ed. São Paulo: Makron books, 2000. 846p.

GUARNIERI J, P.; SOUZA, A, M.; ALMEIDA, S, G. Análise do impacto da crise econômica na previsão de produção de automóveis no mercado brasileiro, via modelos box & jenkins. In: ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO, 29., 2009, Salvador, **Anais...** Salvador: ABEPRO, 2009. Disponível em: <www.abepro.org.br/.../enegep2009_TN_STO_107_713_12958.pdf> Acesso: 10 de out. 2015.

GOOIJER, J.G.; HYNDMAN R.J. 25 years of time series forecasting. **International journal of forecasting**, v. 22, p. 443–473, 2006.

HARRIS, R.I.D. **Using cointegration analysis in econometric modelling**. London: Prentice Hall, 1995, 176p. Disponível em : < <http://piobookreca.com/dafyd953449.pdf>> Acesso: 10 out. 2015.

KHASHEI, M.; BIJARI, M. **An artificial neural network (p, d, q) model for timeseries forecasting**. Expert Systems with Applications, v. 37, p. 479-489, 2010. Disponível em: <

<http://webcache.googleusercontent.com/search?q=cache:VeOt4fwpercJ:citeseerx.ist.psu.edu/viewdoc/download%3Fdoi%3D10.1.1.475.3147%26rep%3Drep1%26type%3Dpdf+&cd=1&hl=pt-BR&ct=clnk&gl=br>> Acesso: 6 set. 2015.

LUTOSA, L.; MESQUITA, M. A.; QUELHAS, O.; OLIVEIRA, R. **Planejamento e controle da produção**. Rio de Janeiro: Elsevier, 2008. 380p.

MAKRIDAKIS, S.; WHEEL WRIGHT, S.; HYNDMAN, R. J., **Forecasting methods and applications**, John Wiley & Sons, EUA, 1998.

MARTINS, V.L.M; WERNER, L. Comparação de previsões individuais e suas combinações: um estudo com séries industriais. **Prod.**, São Paulo, v. 24, n. 3, Setembro. 2014. Disponível em <<http://www.scielo.br/scielo.php?pid=S0103-65132014000300011>>. Acesso: 6 set. 2015

MORETTIN, P. A; TOLOI, C. M. C. **Análise de séries temporais** - 2. ed. São Paulo: Edgar Blucher, 2006. 544p.

MORETTIN, P. A.; TOLOI, C. M. **Previsão de séries temporais**. 2.ed, São Paulo: Atual, 1987. 520p.

MORETTIN, P. A. & TOLOI, C. M. **Análise de séries temporais**. 1.ed. São Paulo: Edgar Blücher, 2004. 550p.

NEGRI, J.A. A cadeia de valor global da indústria automobilística no Brasil. In: A inserção da América Latina nas Cadeias Globais de Valor. 2010.

MERCADO AUTOMOBILÍSTICO NACIONAL: COMPETIVIDADE, INOVAÇÕES TECNOLÓGICAS E PERSPECTIVAS, 10, CONGRESSO NACIONAL DE EXCELÊNCIA EM GESTÃO, 10, 2014, Rio de Janeiro. **Anais...** Niteroi: Escola de Engenharia da Universidade Federal Fluminense, 2014. 420 p.

ROLDAN, F.; MIYAKE, D. **Mudanças de forecast na indústria automobilística: Iniciativas para a estruturação dos processos de tomada de decisão e processamento da**

informação. v. 11, n. 3, p. 413-427, 2004. Disponível em : <
<http://www.scielo.br/pdf/gp/v11n3/a13v11n3>> Acesso: 10 ago. 2015.

LOURENÇO, I, S; NASCIMENTO, L. O. **Métodos de previsão aplicados a uma série de volume de produção de caminhões**. 2009. 63 f. Trabalho de Graduação (Graduação em Métodos Estatísticos Computacionais) - Universidade Federal de Juiz de Fora, Juiz de Fora, 2009.

SILVA, F. V. V. **Análise das estratégias competitivas na indústria automobilística**. 2007. 90 f. Dissertação (Mestrado em Administração de Empresas) – Pontifícia Universidade Católica (PUC/RJ), Rio de Janeiro, 2007.

SOMBRIO, H, B.; NEUMMAN, S, R.; WERNER, L. Aplicação de método de previsão de demanda em empresa automobilística. In: SIMPÓSIO BRASILEIRO DE PESQUISAS OPERACIONAIS, 42, 2010, Bento Gonçalves. Anais... Bento Gonçalves: SBPO, 2010. Disponível em: < www.din.uem.br/sbpo/sbpo2010/pdf/72170.pdf > Acesso em: 10 set. 2015.

SOUZA, R.C., CAMARGO, M.E. **Análise e previsão de séries temporais**. Ijuí: ed. Unijuí, 1996. 220 p.

TUBINO, D. F. **Planejamento e controle da produção**. Sao Paulo: Atlas, 2007. 190 p.

WANG, Q. Artificial neural networks as cost engineering methods in a collaborative manufacturing environment. **International Journal of Production Economics**, v. 109, n. 1-2, p. 53–64, 2007. Disponível em: <<http://dx.doi.org/10.1016/j.ijpe.2006.11.006> > Acesso: 20 ago. 2015.

WERNER, L.; RIBEIRO, J.L. Previsão de Demanda: uma aplicação dos modelos Box-Jenkins na área de assistência técnica de computadores pessoais. **Gestão & Produção**, v.10, n.1, p.47-67, abr.2003.

ANEXO A: DADOS REAIS DE PRODUÇÃO NACIONAL DE CAMINHÕES

Tabela 1 – Dados Produção de Caminhões

Ano	Caminhões	Ano	Caminhões
1957	16259	1986	84544
1958	26998	1987	74205
1959	36657	1988	71810
1960	37810	1989	62699
1961	26891	1990	51597
1962	36174	1991	49295
1963	21556	1992	32025
1964	21790	1993	47876
1965	21828	1994	64137
1966	31098	1995	70495
1967	27141	1996	48712
1968	40642	1997	63744
1969	40569	1998	63773
1970	38388	1999	55277
1971	38868	2000	71686
1972	53557	2001	77431
1973	69202	2002	68558
1974	79413	2003	78960
1975	78688	2004	107338
1976	83891	2005	117693
1977	101368	2006	106601
1978	86269	2007	137281
1979	93051	2008	167406
1980	102017	2009	123633
1981	76350	2010	191621
1982	46698	2011	229083
1983	35487	2012	136264
1984	48497	2013	190962
1985	64769	2014	143.660

Fonte: Anuário da Indústria Brasileira 2015, ANFAVEA

