UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
COORDENAÇÃO DE ENGENHARIA DE PRODUÇÃO
CURSO DE GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

LUCAS MARUJO

COMPARAÇÃO DE MÉTODOS ESTATÍSTICOS PARA PREVISÃO DA DEMANDA DE CAFÉ NO BRASIL

TRABALHO DE CONCLUSÃO DE CURSO

MEDIANEIRA
2017

LUCAS MARUJO

COMPARAÇÃO DE MÉTODOS ESTATÍSTICOS PARA PREVISÃO DA DEMANDA DE CAFÉ NO BRASIL

TRABALHO DE CONCLUSÃO DE CURSO

Trabalho de Conclusão de Curso apresentado junto ao Curso de Graduação em Engenharia de Produção, da Universidade Tecnológica Federal do Paraná – UTFPR, como requisito parcial à obtenção do título de Bacharel em Engenharia de Produção.

Orientadora: Profa. Ma. Neusa Ideck Scherpinsk

Co-Orientadora: Profa. Dra. Carla Adriana Pizarro Schmidt

MEDIANEIRA

2017

MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
CAMPUS MEDIANEIRA

Diretoria de Graduação
Coordenação de Engenharia de Produção
Curso de Graduação em Engenharia de Produção

TERMO DE APROVAÇÃO

COMPARAÇÃO DE MÉTODOS ESTATÍSTICOS PARA PREVISÃO DA DEMANDA DE CAFÉ NO BRASIL

Por

LUCAS MARUJO

Este de trabalho de conclusão de curso foi apresentado às 15h50 do dia 22 de novembro de 2017 como requisito parcial para aprovação na disciplina de TCC 2, da Universidade Tecnológica Federal do Paraná, Campus Medianeira. O candidato foi arguido pela Banca Examinadora composta pelos professores abaixo assinados. Após deliberação, a Banca Examinadora considerou o trabalho de diplomação aprovado.

_____ Profa. Me. Neusa Idick Scherpinsk Universidade Tecnológica Federal do Paraná Orientadora	_____ Profa. Dra. Carla Adriana Pizarro Schmidt Universidade Tecnológica Federal do Paraná Co-Orientadora
_____ Prof. Me. Edson Hermenegildo Pereira Junior Universidade Tecnológica Federal do Paraná	_____ Profa. Me. Priscila Pigatto Gasparin Universidade Tecnológica Federal do Paraná

_____
Prof. Dr. Sérgio Adelar Brun
Universidade Tecnológica Federal do Paraná
Coordenador do Curso

- O Termo de Aprovação assinado encontra-se na Coordenação do Curso -

Às memórias de meu avô, Luiz, por todo seu amor.
À minha avó, Ana, exemplo de bondade.

AGRADECIMENTOS

Agradeço a Deus, pelos singelos milagres do dia a dia. Aos meus avós, Luiz e Ana, pela sua incondicional fé em mim, seu constante apoio e por ser uma fonte de inspiração e exemplo para minha vida. Aos meus familiares, pelo apoio e compreensão pelos meus momentos de ausência. Agradeço muito às professoras, Orientadora e Co-Orientadora, por compartilharem o seu conhecimento e tempo, e estarem sempre dispostas a me ajudar, além de terem depositado confiança e apoio em todos os momentos que se fizeram necessários. Aos amigos, colegas e professores, aos quais, por estarem presentes no mesmo cotidiano acadêmico, entendem e compartilham das mesmas dificuldades e satisfações, manifestam apoio e incentivo constante. Agradeço, com o mais sincero apreço, a todos que, com boa intenção, de alguma forma colaboraram para a realização e finalização deste trabalho.

RESUMO

MARUJO, Lucas. **COMPARAÇÃO DE MÉTODOS ESTATÍSTICOS PARA PREVISÃO DA DEMANDA DE CAFÉ NO BRASIL:** 2017. 96 p. Monografia (Bacharel em Engenharia de Produção) - Universidade Tecnológica Federal do Paraná. Medianeira, 2017.

O Brasil é o maior produtor e exportador mundial e o segundo no *ranking* referente ao consumo, e esse produto é de grande importância para o país, em termos históricos, econômicos e culturais. Entende-se ainda, a necessidade da previsão de demanda como ferramenta de auxílio aos gestores em relação a estruturação do planejamento estratégico, PCP, *marketing*, vendas, além de diversas outras áreas, de modo a ser utilizada de forma direta ou indireta em praticamente toda a cadeia produtiva do café. Este estudo teve como principal intuito, realizar previsões para a demanda de café no Brasil por intermédio da confecção, interpretação e avaliação de diversos modelos matemáticos. Os modelos foram feitos a partir de metodologias estatísticas exploradas na literatura e aplicadas a série histórica de consumo de café no Brasil entre os anos de 1990 e 2016. A avaliação se deu por meio da comparação direta entre seus MAD (*Mean Absolute Deviation*), além da interpretação do coeficiente de U de Theil, para selecionar quais métodos se sobressaíram em relação a previsão trivial. Por meio dessas medidas de acurácia e as análises de coerência ao comportamento natural dos dados, escolheu-se o modelo ARIMA (1,1,5), com o qual, gerou-se projeções para os 5 períodos subsequentes ao último da série real de consumo. Nota-se, como mais relevante a projeção para o ano imediatamente posterior, devido as características anuais dos dados e a possíveis reajustes com o valor do ano atual nos modelos para o próximo ano. Dessa forma, a previsão para o ano de 2017 com o modelo ajustado e intervalo de confiança de 95%, foi de 21.438.900 sacas de 60 kg, o que foi considerado condizente com as estimativas das organizações de referência na área.

Palavras-chave: Holt-Winters; Modelagem; ARIMA; Séries Temporais; Regressão.

ABSTRACT

MARUJO, Lucas. **COMPARISON OF STATISTICAL METHODS FOR THE FORECASTING OF COFFEE DEMAND IN BRAZIL:** 2017. 96 p. Monograph (Graduate in Production Engineering) - Federal Technological University of Paraná. Medianeira, 2017.

The Brazil is the world's largest producer and exporter and the second in terms of consumption and this product is of great importance for the country in historical, economic and cultural terms. It is also understood the need to forecast demand as a tool to help managers in relation to the structuring of strategic planning, PCP, marketing, sales, as well as several other areas to be used directly or indirectly in practically all the coffee production chain. The main purpose of this study was to carry out forecasts for coffee demand in Brazil through the making, interpretation and evaluation of several mathematical models. The models were based on statistical methodologies explored in the literature and applied the historical series of coffee consumption in Brazil between 90s and 2016s. The evaluation was done by means of a direct comparison between their Mean Absolute Deviation (MAD) and the interpretation of the Theil U coefficient to select which methods excelled in relation to the trivial prediction. By means of these measures of accuracy and the coherence analysis to the natural behavior of the data, the ARIMA model (1,1,5) was chosen, with which, it was generated projections for the 5 periods after the last of the real series of consumption. It should be noted that, as the projection for the following year is more relevant, due to the annual characteristics of the data and possible readjustments with the value of the current year in the models for the next year. Thus, a forecast for the 2017s with the adjusted model and 95% confidence interval was 21,438,900 bags of 60 kg, which was considered as estimates of the reference organizations in the area.

Keywords: Holt-Winters; Modeling; ARIMA; Time Series; Regression.

LISTA DE FIGURAS

Figura 1 - Gráfico da evolução das exportações em alguns países exportadores selecionados entre 2015 e 2017. 17
Figura 2 - Evolução da produtividade total de café no Brasil em relação ao comportamento bienal entre 2003 e 2017. 19
Figura 3 - Evolução da produção total de café no Brasil em relação ao comportamento bienal entre 2003 e 2017. 20
Figura 4 - Comparativo da produção de café no Brasil por estados em 2017. 20
Figura 5 - Exemplo de ciclo de vida de produto. 22
Figura 6 - Exemplos de comportamentos de séries temporais. 23
Figura 7 - Classificação de alguns métodos de previsão estudados. 24
Figura 8 - Fluxograma do processo de escolha da abordagem de previsão. 25
Figura 9 - Exemplo de ajustamento linear. 37
Figura 10 - Significado dos parâmetros da reta de regressão. 38
Figura 11 - Ilustração do método dos mínimos quadrados (MMQ). 39
Figura 12 - Exemplo de ajustamento não-linear. 40
Figura 13 - Fluxograma do processo de escolha de um modelo para previsão de demanda. 48
Figura 14 - Gráfico da evolução da série histórica do consumo de café no Brasil de 1990 a 2016. 57
Figura 15 - Histograma de consumo com sobreposição da curva normal. 58
Figura 16 - Boxplot da série de consumo. 59
Figura 17 - Teste de normalidade para a série de consumo. 60
Figura 18 - Curvas de tendências ajustadas a série de consumo. 64
Figura 19 - Gráfico comparativo entre a série de consumo e a do modelo polinomial de sexto grau ajustado com extrapolações para 5 períodos. 65
Figura 20 - Gráfico comparativo entre a série de MMS ajustada com n = 2 e a série de consumo. 67
Figura 21 - Gráfico comparativo entre a série do modelo DES ajustado e a série de consumo. 70
Figura 22 - Diagrama da série histórica de consumo diferenciada uma vez. 71
Figura 23 - Diagrama da série histórica de consumo diferenciada duas vezes. 72
Figura 24 - Correlograma da função de autocorrelação parcial para a série de consumo. 73
Figura 25 - Correlograma da função de autocorrelação para a série de consumo. 74
Figura 26 - Gráfico comparativo entre a série do modelo de regressão múltipla ajustado e a série de consumo. 79
Figura 27 - Diagrama de dispersão entre os ajustes e os resíduos do modelo escolhido. 81
Figura 28 - Correlograma da função de autocorrelação para para os 81
Figura 29 - Correlograma da função de autocorrelação parcial para os resíduos do modelo escolhido. 82
Figura 30 - Teste de normalidade para os resíduos do modelo escolhido. 82
Figura 31 - Run Chart dos resíduos do modelo escolhido. 83
Figura 32 - Gráfico comparativo entre os valores reais e previstos para o consumo de café com previsões para 5 períodos e intervalo de confiança de 95%. 84

LISTA DE QUADROS

Quadro 1 - Comportamento teórico da FAC e FACP para alguns modelos. 36
Quadro 2 - Principais equações das curvas de ajustamento. 41
Quadro 3 - Critérios para a interpretação dos valores de r. 43
Quadro 4 - Classificação das pesquisas científicas. 51
Quadro 5 - Classificação de dados de pesquisa documental.....52
Quadro 6 - Formulário dos critérios para comparação dos métodos.....54

LISTA DE TABELAS

Tabela 1 - Relação de oferta e demanda de café no mundo de 2013 a 2016. 18
Tabela 2 - Evolução da série histórica de consumo de café no Brasil em milhares de sacas de 60kg. 56
Tabela 3 - Estatística descritiva dos dados da série de consumo em milhares de sacas de 60 kg. 57
Tabela 4 - Teste de Dickey-Fuller Aumentado para série de consumo. 61
Tabela 5 - Teste de Phillips-Perron para série de consumo. 62
Tabela 6 - Teste KPSS para série de consumo. 62
Tabela 7 - Equações e R² das curvas de tendência ajustadas. 63
Tabela 8 - Medidas de acurácia para as análises de tendência. 65
Tabela 9 - Testes de sazonalidade para a série de consumo. 66
Tabela 10 - Medidas de acurácia para os modelos de MMS ajustados. 67
Tabela 11 - Medidas de acurácia para o modelo de 70
Tabela 12 - Teste de Dickey-Fuller Aumentado para série diferenciada uma vez. 71
Tabela 13 - Teste de Phillips-Perron para série de consumo diferenciada uma vez. 71
Tabela 14 - Teste KPSS para série de consumo diferenciada uma vez. 72
Tabela 15 - Teste de Dickey-Fuller Aumentado para série de consumo diferenciada duas vezes. 72
Tabela 16 - Teste de Phillips-Perron para série de consumo diferenciada duas vezes. 72
Tabela 17 - Teste KPSS para série de consumo diferenciada duas vezes. 73
Tabela 18 - Medidas de acurácia para os modelos ARIMA ajustados. 75
Tabela 19 - Correlação entre os fatores causais aplicados na análise de regressão múltipla. 76
Tabela 20 - Relação de equações de regressão múltipla e $\boldsymbol{Ra2}$ 77
Tabela 21 - Medidas de acurácia para Regressão Múltipla. 78
Tabela 22 - Comparação entre os métodos aplicáveis. 80
Tabela 23 - Previsões para a o consumo de café no Brasil para os anos 84

LISTA DE SIGLAS E ACRÔNIMOS

ABIC	Associação Brasileira da Industria de Café
ADF	*Augmented Dickey-Fuller*
AIC	*Akaike Information Criterion*
ARIMA	*Auto Regressive Integrated Moving Averages*
BIC	*Bayesian Information Criterion*
Cecafé	Conselho dos Exportadores de Café do Brasil
CONAB	Companhia Nacional de Abastecimento
DSE	Dupla Suavização Exponencial
Embrapa	Empresa Brasileira de Pesquisa Agropecuária
FAC	Função de Autocorrelação
FACP	Função de Autocorrelação Parcial
HWA	Holt-Winters Aditivo
HWM	Holt-Winters Multiplicativo
IBGE	Instituto Brasileiro de Geografia e Estatística
MAD	*Mean Absolute Deviation*
MAPE	*Mean Absolute Percentage Error*
MSD	*Mean Squared Deviation*
MMP	Média Móvel Ponderada
MMS	Média Móvel Simples
NIST	*National Institute of Standards and Technology*
OIC	Organização Internacional do Café
OMS	Organização Mundial da Saúde
PCP	Planejamento e Controle da Produção
SSE	Simples Suavização Exponencial
SQE	Soma dos quadrados dos erros

SUMÁRIO

1 INTRODUÇÃO 13
1.1 Objetivo Geral 14
1.2 Objetivos Específicos 15
2 REVISÃO DE LITERATURA 16
2.1 CAFÉ NO MUNDO E A IMPORTÂNCIA DO BRASIL 16
2.2 CAFÉ NO BRASIL 18
2.3 PREVISÂO DE DEMANDA 21
2.4 PREVISÕES BASEADAS EM MÉTODOS QUALiTATÍVOS 26
2.5 PREVISÕES BASEADAS EM MÉTODOS DE SÉRIES TEMPORAIS 27
2.5.1 Previsão Ingênua 28
2.5.2 Média Móvel Simples (MMS) 29
2.5.3 Média Móvel Ponderada (MMP) 30
2.5.4 Simples Suavização Exponencial (SSE) 31
2.5.5 Dupla Suavização Exponencial (DSE) ou Método de Holt 32
2.5.6 Método de Holt-Winters 32
2.5.7 ARIMA 34
2.6 PREVISÕES BASEADAS EM MÉTODOS CAUSAIS 36
2.6.1 Regressão Linear Simples 37
2.6.2 Regressão Não-Linear 40
2.6.3 Regressão Múltipla 41
2.6.4 Coeficientes de correlação e de determinação 42
2.7 MONITORAMENTO DO MODELO E ERROS DE PREVISÃO 45
2.7.1 Análise de resíduos 47
3 MATERIAL E MÉTODOS 48
3.1 CLASSIFICAÇÃO DA PESQUISA 50
3.2 COLETA DE DADOS 53
3.3 PROCEDIMENTOS METODOLÓGICOS 53
4 RESULTADOS E DISCUSSÕES 56
4.1 SÉRIE HISTÓRICA DE CAFÉ E ESTATÍSTICA DESCRITIVA 56
4.2 TESTES DE ESTACIONARIEDADE, TENDÊNCIA E SAZONALIDADE 60
4.2.1 Teste de Estacionariedade 61
4.2.2 Análise de Tendência 62
4.2.3 Teste de Sazonalidade 66
4.3 MODELO DE MÉDIA MÓVEL SIMPLES 67
4.4 MODELO DE MÉDIA MÓVEL PONDERADA 68
4.5 MODELO DE SUAVIZAÇÃO EXPONENCIAL SIMPLES 69
4.6 MODELO DE HOLT (DUPLA SUAVIZAÇÃO EXPONENCIAL (DSE)) 69
4.7 MODELO DE HOLT-WINTERS 70
4.8 MODELO ARIMA 71
4.9 MODELO DE REGRESSÃO MULTIPLA 76
4.10 COMPARAÇÔES ENTRE OS MODELOS APLICÁVEIS 79
4.11 VALIDAÇÃO DO MODELO E PREVISÔES 80
4.11.1 Independência e autocorrelação dos resíduos 80
4.11.2 Teste de Normalidade dos resíduos 82
4.11.3 Homocedasticidade dos resíduos 83
4.11.4 Previsões 84

5 CONSIDERAÇÕES FINAIS 86
6 SUGESTÕES PARA TRABALHOS FUTUROS 87
REFERÊNCIAS 88

1 INTRODUÇÃO

O café é um dos produtos de maior relevância no Brasil há séculos, segundo dados da Organização Internacional do Café - OIC (2017) e da Associação Brasileira da Indústria de Café – ABIC (2017). O país é o maior produtor de café no mundo, produz anualmente, um valor equivalente a soma dos outros seis países que mais produzem, tem também o título de maior exportador e o segundo maior no *ranking* de consumo.

A importância do Brasil para o mercado de café no mundo e sua relevância na economia nacional, é a principal motivação em relação ao estudo da demanda desse produto. Motivado também pelo gosto pessoal ao tema, com o propósito de estudar seu comportamento, analisar de forma correlacionada com sua produção, seu preço, no contexto econômico envolvido, e o aumento contínuo da população do país. Com o uso de ferramentas estatísticas, pode-se analisar esses dados de consumo e efetuar previsões para essa demanda.

Considera-se que, em qualquer ambiente competitivo, ter uma compreensão de eventos futuros pode ser um imenso diferencial, a busca pela previsão, se torna então, uma atividade natural, em que existe a perspectiva da elaboração de metodologias que venham a evoluir tal propósito. Para Fernandes e Godinho Filho (2010, p.17) as previsões servem “[...] como guia para o planejamento estratégico da produção, finanças e vendas de uma empresa”.

Existem diversos métodos para executar tais projeções, de ordem qualitativa e quantitativa. Em uma abordagem qualitativa, investigam-se opiniões de especialistas ou informações de questionários com métodos técnicos para tal. Um estudo de previsão quantitativo, tem uma abordagem matemática, baseia-se na ideia de estimar valores futuros a partir de dados coletados dos eventos passados (KRAJEWSKI *et al.,* 2009).

Com o uso de um planejamento adequado, as previsões atuam como um meio regulador para a administração dos setores, Martins e Laugeni (2005, p. 226), no âmbito da previsão, afirmam que

> [...] a previsão de vendas é importante para utilizar as máquinas de maneira adequada, para realizar a reposição dos materiais no momento e na quantidade certa, e para que todas as demais atividades necessárias ao processo industrial sejam adequadamente programadas.

Conforme Tubino (2009, p.15), dentro das indústrias, as previsões são usadas pelo setor de planejamento e controle da produção (PCP) para planejar o sistema produtivo e o uso desse sistema. Ballou (2006 p. 241) afirma que, o planejamento da cadeia de suprimentos e a logística dependem das projeções feitas através de previsões em determinados setores como planejamento, *marketing* ou mesmo o departamento logístico.

A previsão de demanda não se limita apenas as empresas de manufatura, no varejo ela é útil para prever reposição de estoque e planejamento de compras, por exemplo. Assim como, em empresas prestadoras de serviço, em que, proporciona planejamento de seus recursos disponíveis para atender as necessidades do cliente. Neste caso se houver erros acentuados de previsão poderão ter como consequência ociosidade de recursos, tempo longo de espera por atendimento, que, de certa forma, é uma característica da demanda para esse tipo de sistema (LUSTOSA *et al*., 2008 p. 52).

Com base no processo de previsão de demanda, analisam-se outras motivações para o trabalho. A previsão de demanda tem como propósito auxiliar ao gestor no processo decisório de uma organização, assim como no entendimento do mercado, para que o mesmo possa ter mais segurança em seu planejamento. Corrêa e Corrêa (2012, p. 236) afirmam que para um gestor tomar uma boa decisão “[...] é necessário que ele tenha uma visão, a mais clara possível, de futuro, para que a decisão tomada hoje seja adequada, não ao presente, mas ao momento no futuro em que a decisão de fato tomar efeito”.

1.1 OBJETIVO GERAL

O objetivo geral a ser alcançado nesse trabalho consiste em elaborar uma previsão confiável para a demanda de café no Brasil, por intermédio da comparação de diversos métodos estatísticos de previsão.

1.2 OBJETIVOS ESPECÍFICOS

Para alcançar o objetivo principal, constituíram-se os seguintes objetivos específicos:

a) Elaborar modelos matemáticos para a previsão com o uso de métodos estatísticos, a fim de fazer estimativas para a demanda.
b) Testar os modelos ajustados em relação aos seus desempenhos estatísticos por meio de seus erros.
c) Avaliar estatisticamente as influências de fatores causais, tais como população, preço, produção e exportação de café no Brasil.
d) Identificar quais seriam os métodos mais indicados para previsão da demanda em consideração aos fatores que possam influenciar a tomada de decisão.

2 REVISÃO DE LITERATURA

Aborda-se nesta seção a referência literária utilizada como embasamento para a elaboração do trabalho. Este tópico tem por finalidade auxiliar ao leitor em relação ao entendimento sobre o assunto proposto.

2.1 CAFÉ NO MUNDO E A IMPORTÂNCIA DO BRASIL

A história do café provém de lendas, Rodrigues *et al.* (2015) comentam o fato de existirem relatos que evidenciam a presença da bebida no mundo há mais de mil anos. Segundo a ABIC (2017) o café é originado da Etiópia, onde é tido como vegetação natural, porém foi a Arábia quem o difundiu, e propagou essa cultura para o mundo.

Normalmente tomada no desjejum, por ser uma bebida energética, o café é uma bebida de grande consumo mundial, além de ser considerado útil para ajudar na prevenção de certas doenças. Segundo Avanza (2016), no Jornal da USP, a Organização Mundial da Saúde (OMS) além de reclassificar o café, em 2016, como não cancerígeno, ainda indicou que pode auxiliar na prevenção de certos tumores, a ABIC (2017), destaca ainda, o auxílio contra a depressão, o suicídio, dentre outras doenças cerebrais.

O Brasil é o maior produtor de café há séculos, corresponde no ano de 2017 a 30% do mercado de café, o que é equivalente ao dobro de segundo melhor no *ranking*, ou a soma das produções dos outros seis países melhores colocados (OIC, 2017). Segundo interpretação de Ferreira e Santos (2017) do Consórcio Pesquisa Café da Empresa Brasileira de Pesquisa Agropecuária – Embrapa, “No Brasil cafeicultura corresponderá a 4,2% do Valor Bruto da Produção Agropecuária estimado em R$ 546 bilhões” no ano de 2017, e, graças ao trabalho de profissionais técnicos, assim como uma boa gestão aplicada, em 2016, o país bateu o recorde em produtividade, superando seu próprio recorde anterior em 2012.

Em seu relatório sobre o mercado de café de janeiro de 2017, a OIC (2017), afirma que

> Em dezembro de 2016 as exportações totalizaram 10,2 milhões de sacas, 7% a mais que em dezembro de 2015. Com isso, as exportações do primeiro trimestre do ano cafeeiro de 2016/17 (outubro a dezembro) se elevaram a 29,8 milhões de sacas, 8,3% a mais que no mesmo período de 2015/16.

Também é apresentado, nesse mesmo relatório, um demonstrativo do desempenho dos principais exportadores, assim como a significativa influência do Brasil nesse quesito, conforme pode ser visto na Figura 1.

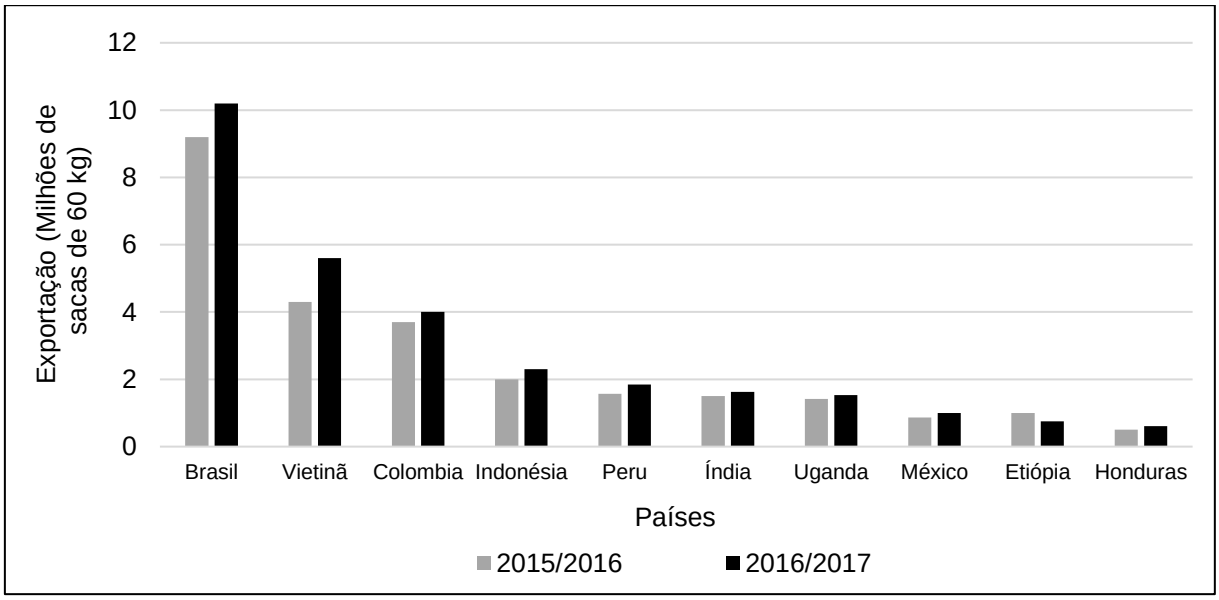


Figura 1 - Gráfico da evolução das exportações em alguns países exportadores selecionados entre 2015 e 2017.
Fonte: Adaptado de OIC (2017).

Já em seu relatório sobre o mercado cafeeiro de abril de 2017, a OIC (2017) informa que “Nos sete primeiros meses do ano cafeeiro de 2016/17 (outubro de 2016 a abril de 2017) elas aumentaram 3,1%, passando a 69,51 milhões de sacas, ante 67,41 milhões em relação às exportações do mesmo período do ano cafeeiro passado”, o que denota um constante e elevado patamar nas exportações.

Apresenta-se ainda, nesse relatório, uma relação entre a produção e o consumo de café no mundo, em que pode ser visto a produção mundial dentre os dois principais tipos de café, a divisão da produção das principais regiões produtoras no mundo e o consumo em relação aos países importadores e exportadores, conforme a Tabela 1,

Tabela 1 - Relação de oferta e demanda de café no mundo de 2013 a 2016.

Ano de começo da safra	2013	2014	2015	2016	% de mudança 2015-2016
Produção (Milhares de sacas)	152130	148724	151438	151624	0,1%
Arábica	90163	86151	88273	95204	7,9%
Robusta	61967	62572	63165	56419	- 10,7%
África	16243	15987	16229	16353	0,8%
Ásia e Oceania	46461	45666	48967	43110	- 12,0%
México e América Central	16598	17116	17291	17740	2,6%
América do Sul	72828	69954	68951	74420	7,9%
Consumo (Milhares de sacas)	152130	151822	15572	155100	- 0,4%
Países Exportadores	46109	47245	48262	48337	0,2%
Países Importadores (Ano cafeeiro)	102931	104577	107450	106763	- 0,6%
África	10595	10739	10745	10774	0,3%
Ásia e Oceania	30714	32602	33665	33669	0,0%
México e América Central	5158	5240	5311	5237	- 1,4%
Europa	50169	50907	51802	51544	- 0,5%
América do Norte	27714	27372	28875	28535	- 1,2%
América do Sul	24682	24962	25313	25341	0,1%
Balanço	3098	- 3098	- 4274	- 3476	- 18,7%

Fonte: Adaptado de OIC (2017).

Nota-se no balanço, considerando a divisão em relação as principais regiões consumidoras, o fato do consumo estar ultrapassando a produção nos últimos anos, esses valores negativos são justificados pela própria OIC (2017), devido a transações não documentadas.

2.2 CAFÉ NO BRASIL

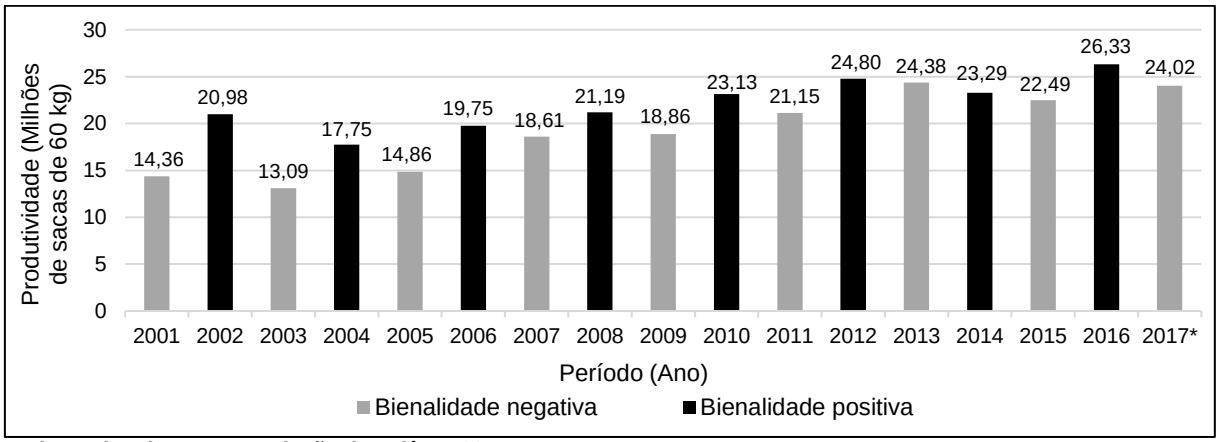
De acordo com Faria e Manolescu (2004) o café chegou ao Brasil por volta do ano de 1700, oriundo dos países da Guiana Francesa e Holandesa. Devido as condições climáticas favoráveis do Brasil a planta se dispersou através de vários estados, entre eles, Maranhão, Bahia, São Paulo, Pará, Rio de Janeiro, Paraná e Minas Gerais. Em pouco tempo se tornaria o principal produto econômico do Brasil, o que desenvolveu as áreas da agricultura e indústria, e tornou o café um dos principais produtos do país.

Nesse período, a produção no Brasil se firmou graças ao trabalho escravo que os grandes proprietários das fazendas exploravam, obtveve lucros altos, o que gerou enormes fortunas. Graças a essa produção a economia do Brasil cresceu rapidamente, por meio do incentivo ao investimento em infraestrutura para atender a demanda, como por exemplo a construção da estrada de ferro para o porto de Santos (RODRIGUES *et al*., 2015).

Segundo dados da Conab (2017), em seu relatório de acompanhamento da safra brasileira de maio de 2017, o Brasil apresentou nesse período uma estimativa de área total cultivada de café de 2,212 milhões de hectares (341,4 mil hectares em formação e 1,87 milhões de hectares em produção), e essa área apresenta diminuição, devido ao aumento da produtividade em função do aperfeiçoamento do processo e da tecnologia envolvida.

Pode-se observar na Figura 2, o aumento da produtividade ao longo do tempo, também, de maneira destacada, ilustra-se o comportamento característico de bienalidade do café e a importância climática para a produção. A Conab (2017) entende que

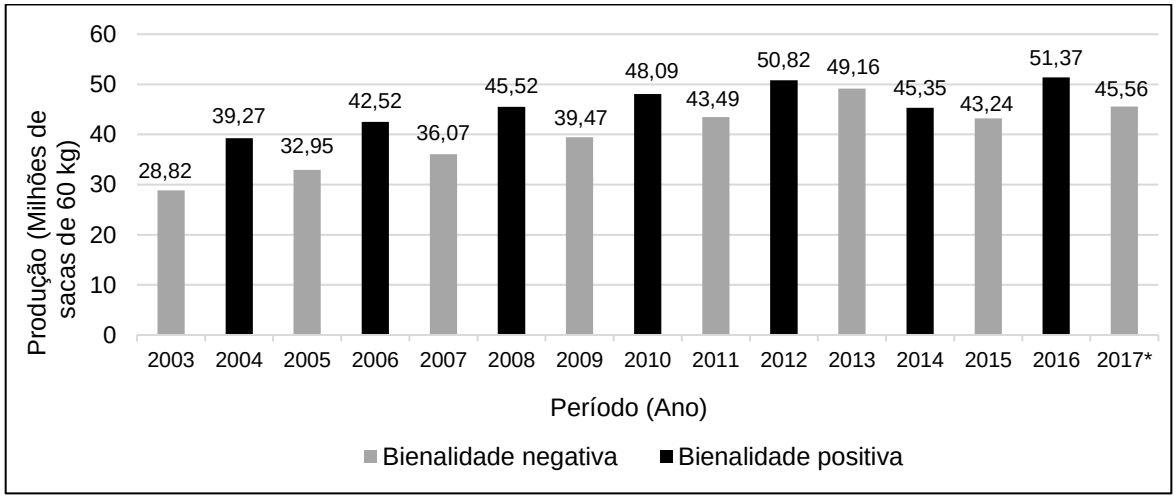
> O ciclo bienal é, portanto, característica do cafeeiro. Contudo, uma adversidade climática pode alterar o ciclo bienal, como ocorreu em 2014 que, apesar de ser uma safra positiva, a forte restrição hídrica fez com que a produtividade fosse inferior ao ano anterior.



***Valor estimado para a produção de café em 2017.**
Figura 2 - Evolução da produtividade total de café no Brasil em relação ao comportamento bienal entre 2003 e 2017.
Fonte: Conab (2017).

Ainda nesse relatório, é apresentada a produção estimada para o ano de 2017, de 45,56 milhões de sacas de 60 kg de café beneficiado, 11,3% menor em relação a 2016, que foi de 51,37 milhões de sacas. Explica-se através de fatores

econômicos e climáticos, além da relação natural da bienalidade, como pode ser visto na Figura 3.



***Valor estimado para a produção de café em 2017.**

Figura 3 - Evolução da produção total de café no Brasil em relação ao comportamento bienal entre 2003 e 2017.
Fonte: Adaptado de Conab (2017).

Segundo a ABIC (2017), os principais estados produtores de café no Brasil são: Minas Gerais, Espirito Santo, São Paulo, Bahia, Rondônia, Paraná e outros com menor expressão, como, Rio de Janeiro, Goiás, Pará, Mato Grosso e Amazonas. Uma comparação entre essas produções pode ser observada na Figura 4.

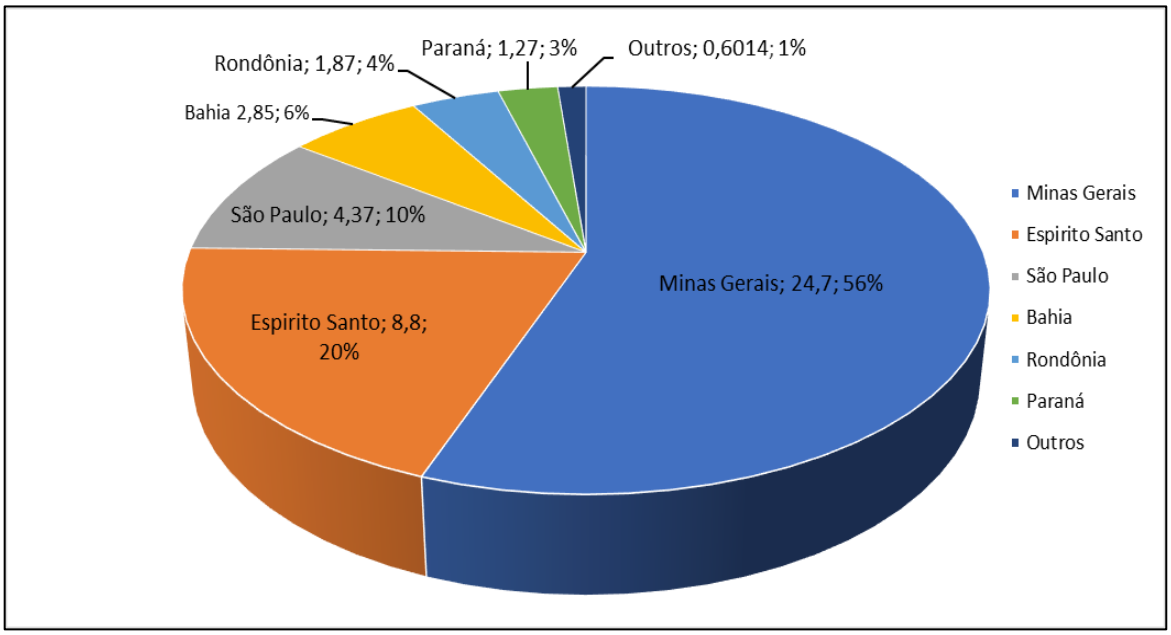


Figura 4 - Comparativo da produção de café no Brasil por estados em 2017.
Fonte: Adaptado de Conab (2017).

Conforme dados de análise por produção entre estados do relatório de maio de 2017 da Conab (2017), é possível notar a importância do estado de Minas Gerais, corresponde a mais da metade de toda produção do país, compõe assim, a principal parcela na produção. Se consideramos a espécie de café produzido, têm-se o estado do Espirito Santo com grande relevância no tipo Canilon, enquanto, Minas Gerais produz o tipo Arábica.

2.3 PREVISÂO DE DEMANDA

Determinados fenômenos de várias naturezas, podem ser eventualmente previstos com um certo grau de incerteza mediante a elaboração de modelos que se ajustem aos dados coletados, de modo a representá-los o mais realisticamente possível, considerando as relações entre os fatores que influenciam esses dados através de modelos matemáticos (CARDOSO, 2011, p.7).

Martins e Laugeni (2005, p. 226) definem previsão como “[...] o processo metodológico para determinação de dados futuros baseados em modelos estatísticos, matemáticos ou econométricos ou ainda modelos subjetivos apoiados em uma metodologia de trabalho clara e previamente definida”. Para Fernandes e Godinho Filho (2010, p. 17) “[...] os métodos de previsão geralmente pressupõem que o comportamento causal que existiu no passado continuará a existir no futuro” e ainda, é possível “[...] através da matemática e/ou estatística gerar uma série de informações que ainda vai passar por várias análises” (OLIVEIRA; LIMA, 2016, p.13).

Para Lustosa *et al.* (2008) demanda é “[...] a disposição dos clientes ao consumo de bens e serviços ofertados por uma organização”, em que essa é influenciada por fatores que se estendem desde macroeconomias até a parte operacional, como disponibilidade de produto e o preço do produto final na prateleira.

O mesmo autor considera ainda, duas situações para a demanda: a pontual e a repetitiva. Demanda pontual se caracteriza por um pico concentrado e após um período curto de tempo a demanda diminui gradativamente caracterizando um período sazonal, tem-se como exemplo, o aumento de locações de casas em uma praia no verão. Percebe-se que nesse caso os erros de previsão podem resultar em excesso

ou falta de produto e muitas vezes a dificuldade em prever esse tipo de demanda está na falta de dados históricos. A demanda repetitiva, pode ser classificada em dependente ou independente, em que a primeira associada a demanda de um outro produto, por exemplo, matéria-prima está correlacionada a produtos acabados, enquanto que a demanda independente está associada ao “ciclo de vida de produto”, conforme Figura 5.

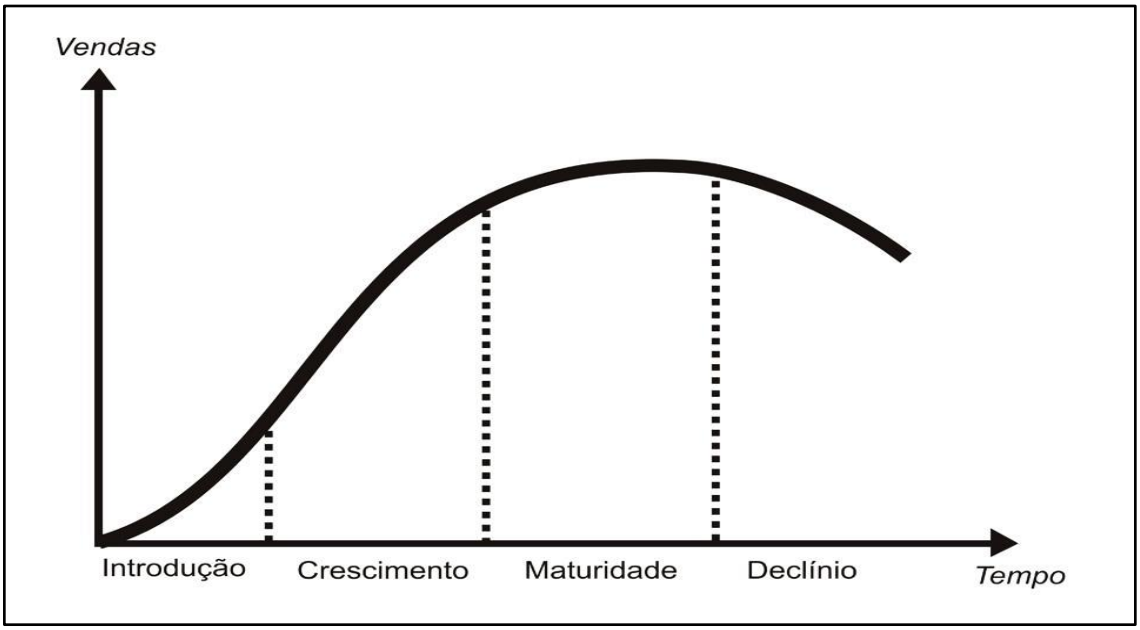


Figura 5 - Exemplo de ciclo de vida de produto.
Fonte: Adaptado de Lustosa *et al*. (2008).

Podem ter áreas em que a demanda se apresenta de forma estacionária, refere-se a fase de maturidade do produto com variações nas vendas. Em um patamar constante de demanda, apresenta certa tendência de venda com crescimento ou redução sistemática, normalmente obtida na fase de crescimento ou declínio das vendas. Pode apresentar ainda, uma terceira fase de sazonalidade, uma oscilação repetitiva na demanda ao longo do tempo, a qual, também melhor definida na seção 2.5, referente a séries temporais (LUSTOSA *et al*., 2008).

A demanda nesse ciclo, ou seja, na situação de demanda repetitiva, pode assumir alguns padrões, aos quais, podem constituir características clássicas de séries temporais, que serão melhores abordadas na seção 2.5. Alguns exemplos desses padrões em disposições de séries temporais, estão representados, conforme Figura 6.

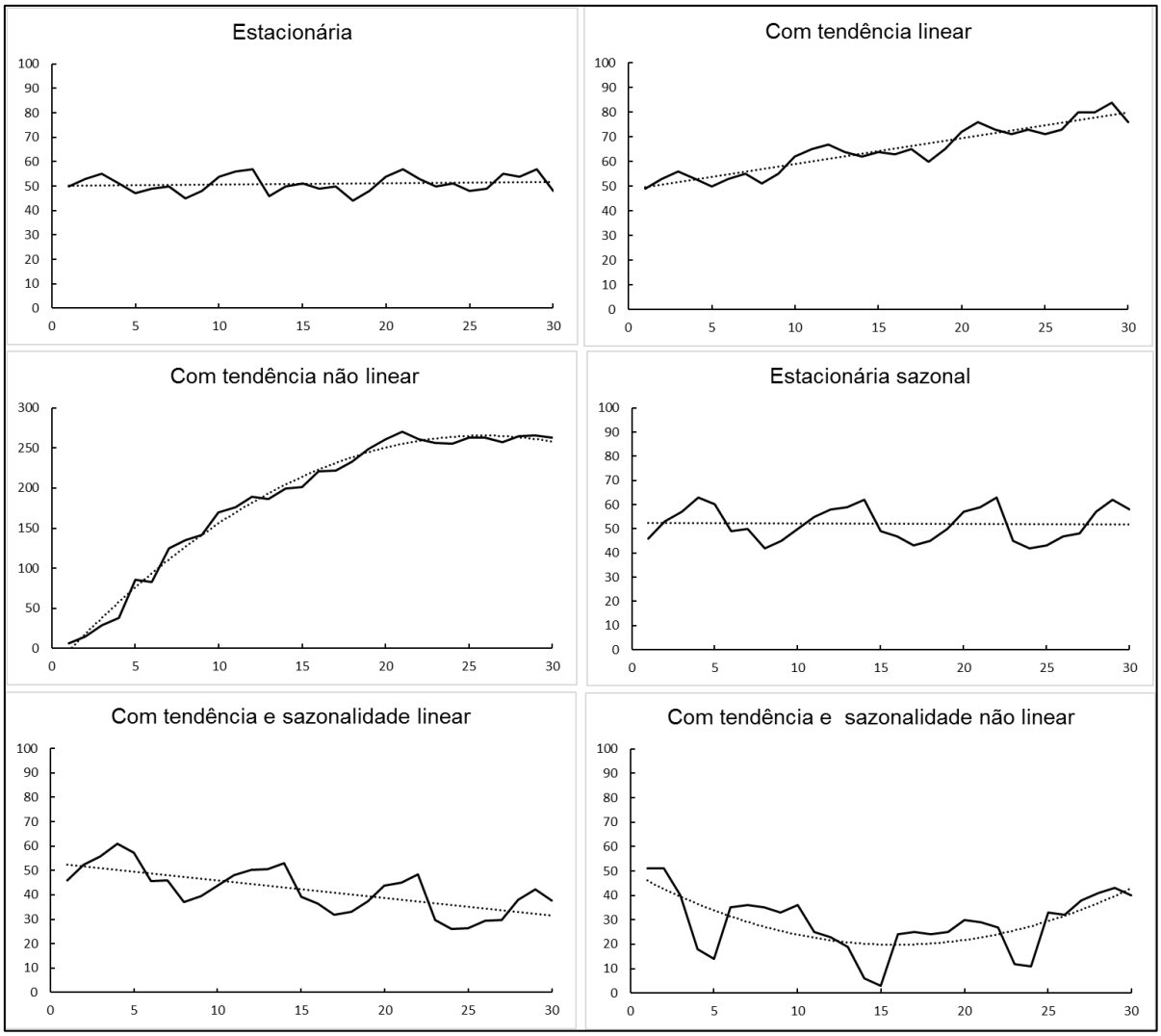


Figura 6 - Exemplos de comportamentos de séries temporais.
Fonte: Adaptado de Lustosa *et al*. (2008), Ballou (2006) e Martins e Laugeni (2005).

Ballou (2006 p. 242) divide demanda como irregular ou regular, classificadas também quanto a ser independente ou dependente. A demanda irregular é caracterizada por uma grande inconstância, existe um padrão intermitente aleatório e incerto. Enquanto na demanda regular pode-se associar a padrões de demanda já estudados e com métodos já conhecidos.

O autor define ainda demanda independente como aquela gerada por diversos clientes, em que muitos efetuam compras esporádicas e em pouco volume, sugere que para esse tipo de demanda “[...] os procedimentos de previsão estatística dão bons resultados”. Define demanda dependente, da mesma forma que os demais autores, como aquela gerada a partir de exigências, ou predefinições correlacionadas com algum processo produtivo.

Morettin e Toloi (1981, p. 10) sugerem que na economia se definem duas

grandes vertentes de estudo, a econometria e análise de séries temporais. Enquanto o primeiro utiliza-se de bases primordialmente teóricas o segundo pressupõe um estudo caráter matemático, nos quais, “[...] os dados falam por si só”. Já para Fischer (1982, p. 17)

> A Econometria é o ramo da Ciência Econômica que, através da análise estatística da realidade, busca estabelecer proposições econômicas, de caráter quantitativo, que possibilitam não só compreender e analisar o comportamento das variáveis sob uma determinada base econômica teórica, como também permitem a previsão de seus comportamentos futuros.

Tubino (2009) considera, de forma geral, que as previsões de demanda são divididas em duas abordagens metodológicas principais, abordagem qualitativa e abordagem quantitativa, pode-se ainda ser subdividida em métodos de decomposição de series temporais e métodos causais.

Segundo Peinado e Graeml (2007, p. 333), “[...] os modelos qualitativos são, essencialmente, subjetivos”. Podem ser utilizados geralmente na ausência de dados históricos ou em conjunto, dependem de especialistas experientes, abordados mais frequentemente pela área de marketing. Já as técnicas quantitativas possuem em sua base a análise numérica, usa-se modelos matemáticos para realizar as previsões futuras. Pode se dividir ainda as técnicas quantitativas, em séries temporais e séries baseadas em correlações. Lustosa *et al*. (2008, p. 57) e Corrêa e Corrêa (2012, p. 248) ilustram essas abordagens conforme a Figura 7.

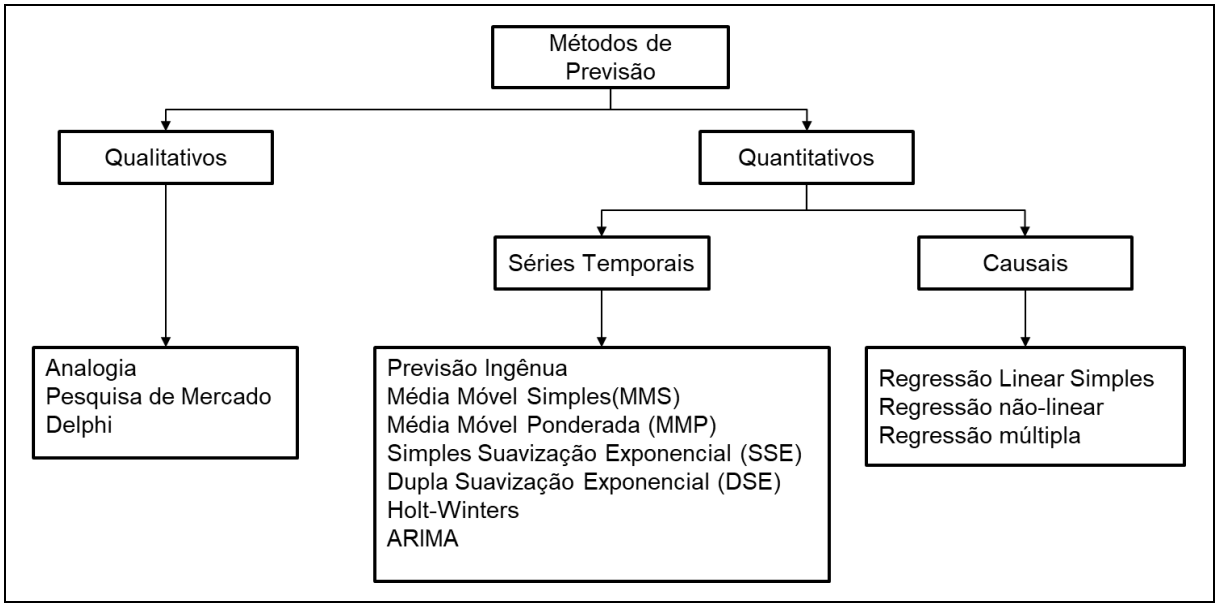


Figura 7 - Classificação de alguns métodos de previsão estudados.
Fonte: Adaptado de Lustosa *et al*. (2008, p. 57) e Corrêa e Corrêa (2012, p. 248).

Para Lustosa *et al*. (2008), qualquer técnica de previsão quantitativa deve utilizar parâmetros para definir o modelo, em que, “[...] valores iniciais e constantes de suavização constituem os parâmetros do modelo a serem estimados”, podem ser baseados nos dados históricos, simulações e os que apresentam o menor índice de erro, ver seções 2.6.4 e 2.6.5. A escolha de abordagem pode ser ilustrada de maneira sucinta, conforme a Figura 8.

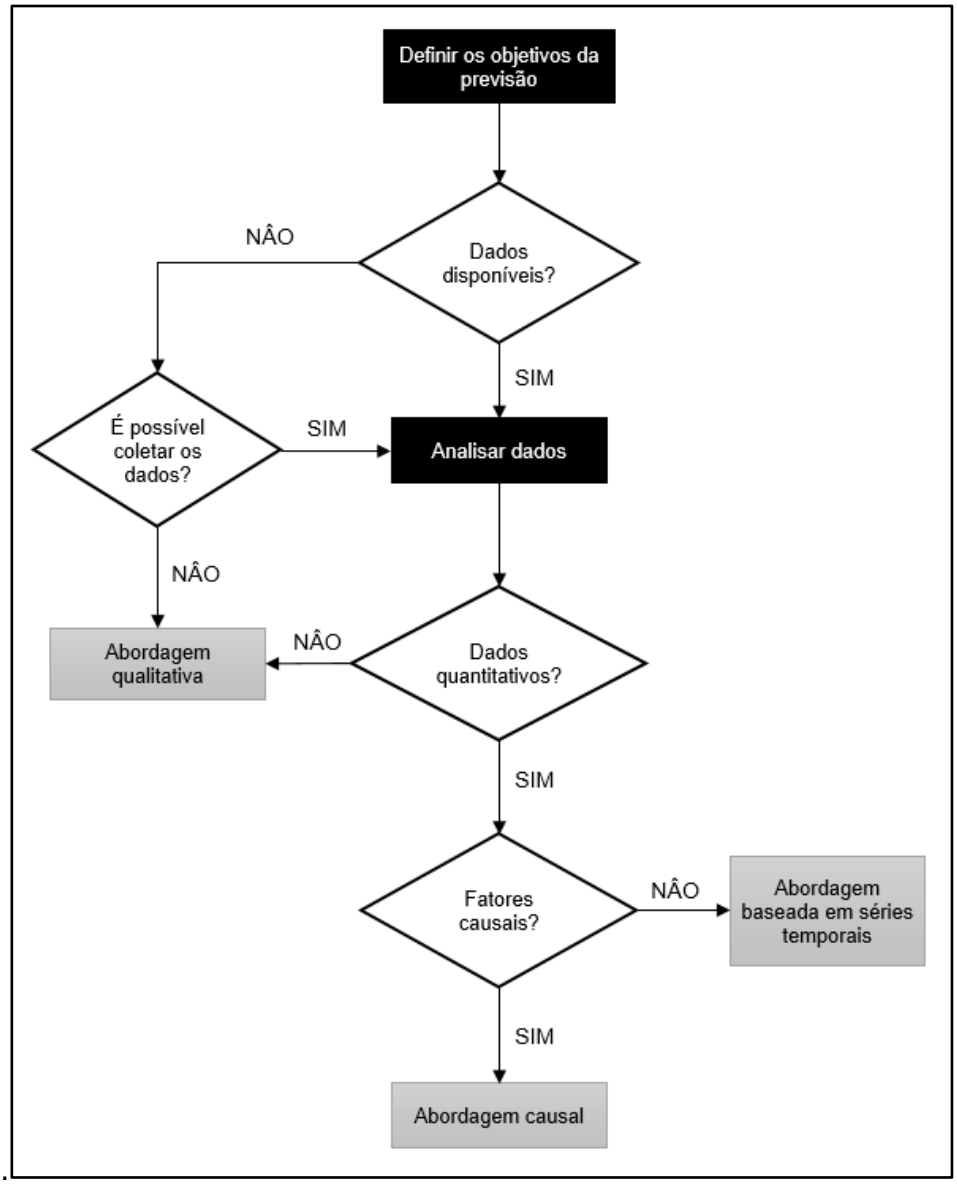


Figura 8 - Fluxograma do processo de escolha da abordagem de previsão.
Fonte: Adaptado de Sipper e Bulfin (1997) *apud* Fernandes e Godinho Filho (2010).

Na interpretação de Tubino (2009), existem algumas similaridades entre as várias técnicas disponíveis, essas são levadas em consideração, assim como, fatores comuns a todas, como por exemplo, se irão se comportar da mesma maneira

no futuro, considerar que as previsões não são perfeitas, existem variações e ocorrências aleatórias durante sua aplicação, que sua acuracidade diminui de acordo com o período que é estimada a previsão, e que as previsões de grupos de produto possuem melhores resultados de estimativas que produtos individuais, pois nos grupos os erros se minimizam.

2.4 PREVISÕES BASEADAS EM MÉTODOS QUALITATÍVOS

Os métodos com ênfase na abordagem qualitativa, são analisados de forma subjetiva, são aqueles que, de modo geral, não podem ser quantificados, e sim qualificados, ou seja, atribui-se identificações em termos de características específicas (ROSSETTO *et al.*, 2011, p. 2).

Esses métodos podem ser empregados em situações de ausência de registros históricos, por exemplo, ao introduzir um novo produto no mercado a ser estudado pelo departamento de marketing, também no caso em que existam dados, mas não são tão adequados ao uso, ou, ainda, para auxiliar e agregar em análises quantitativas, faz-se necessário a utilização de métodos de previsão de cunho qualitativo (KRAJEWSKI *et al.*, 2009).

Segundo Bonotto e Fogliatto (2015), esses métodos tendem a ser utilizados de forma combinada entre si ou mesmo de forma estruturada em junção aos métodos quantitativos, como suporte para a tomada de decisões. Dessa forma, faz-se a analise numérica, com suas projeções quantitativas e posteriormente essas são analisadas e ajustadas por especialistas, que fazem uso de métodos qualitativos.

Makridakis *et al*. (1998), afirmam que os principais métodos utilizados são: Pesquisa de Mercado; Grupos focados; Analogia; e Delphi. Aos quais não são o foco deste trabalho, porém devem estar envoltos no contexto de interpretações extrínsecas que compõem a análise econômica, social e cultural que envolvem o comportamento da série estudada, no caso, as características do mercado onde o café está inserido.

2.5 PREVISÕES BASEADAS EM MÉTODOS DE SÉRIES TEMPORAIS

As séries temporais relacionam dados históricos obtidos dos produtos com a variável tempo, a partir de uma base de dados históricos e retira-se pontos extremos, com a consideração de tendências e sazonalidades, e faz-se uma projeção futura, na ausência de influência de outras formas de variáveis, caracteriza-se por um modelo mais simples e usual e quando bem realizado oferece ótimos resultados (PEINADO; GRAEML, 2007).

Para Farias e Trevisan (2013, p. 2),

> O objetivo da análise de uma série temporal consiste em elaborar um modelo estatístico que descreva adequadamente a procedência de dita série, de maneira que as implicações teóricas do modelo resultem compatíveis com as pautas de amostras observadas nas séries temporais. Depois o modelo elaborado a partir da série temporal pode ser utilizado para prever a evolução futura da série ou explicar a relação entre os distintos componentes do modelo.

Para Fischer (1982, p. 19) “A grande maioria dos estudos econômicos que se valem do tratamento de séries temporais utiliza séries discretas, onde as observações são geradas em um intervalo de tempo com amplitude constante’’. São também classificadas entre determinísticas e estocásticas, na qual a primeira pode ser determinada com precisão através de uma relação de função matemática. Já na segunda a relação envolve outros fatores além da relação matemática ao longo do tempo.

Segundo Morettin e Toloi (1981) a primeira suposição geralmente feita, é a de que ela seja estacionária, ou seja, sua variação ao longo do tempo é em torno de uma média constante. A grande maioria apresenta algum tipo de não estacionariedade, apresentam alguma tendência, comumente uma tendência aproximada para uma reta (tendência linear).

Existem diversos métodos para testar a estacionariedade de uma série, a maioria utiliza-se de um teste de hipóteses chamado de teste de raiz unitária, em que a hipótese nula é a de que a série possui ao menos uma raiz unitária, nesse caso, não é considerada estacionária, consequentemente a hipótese alternativa é de não possuir raiz unitária, o que indica uma tendência depende apenas do tempo. O principal método utilizado é o teste de Dickey-Fuller Aumentado (ADF-*Augmented Dickey-Fulle*r) (HYNDMAN; ATHANASOPOULOS, 2017).

Uma análise clássica pressupõe que uma série temporal possa apresentar quatro características principais: nível, tendência, sazonalidade e aleatoriedade. Nível se refere a nada mais do que o valor do eixo de vendas (y) da série, tendência é o quanto os valores dos dados tendem a aumentar ou decrescer ao longo do tempo (x), “[...] podem ajustar-se a uma reta (tendência de variação linear) ou a alguma outra curva (exponencial, por exemplo)” (CORRÊA; CORRÊA, 2012, p. 251). A sazonalidade remete a padrões de oscilações na demanda que caracteriza ciclos, “[...] é sempre um período fixo e conhecido” (HYNDMAN; ATHANASOPOULOS, 2017) Existem ainda fatores aleatórios independentes, que devem ser considerados como influentes, esses dados não podem ser previstos, portanto “[...] devem ser retirados da série histórica e substituídos pela média” (TUBINO, 2009, p. 19).

Para Morettin e Toloi (2004, p. 63), para analisar a tendência, sazonalidade, variabilidade ou observações discrepantes de uma série, é primordial a análise gráfica, embora possam se realizar testes de hipóteses estatísticos como o teste de sequências (Wald-Wolfowitz) para tendência e o teste de Kruskal-Wallis para a sazonalidade. Dessa forma constitui-se então “[...] que toda série temporal pode ser analisada e decomposta em uma parte sistemática, composta por nível, tendência e sazonalidade, e outra parte não sistemática composta pela aleatoriedade” (PEINADO; GRAEML, 2007).

2.5.1 Previsão Ingênua

O método de previsão ingênua ou trivial, é o mais simples dos métodos, resume-se na utilização do último valor da série como previsão para o valor sucessor. “Apesar de muito simples, esse método é bastante utilizado e realmente pode ser eficaz para um processo constante” (FERNANDES; GODINHO FILHO, 2010).

Esse método, pode servir de parâmetro para comparação em relação ao coeficiente U de Theil, utilizado como análise de decisão sobre a adequação de um modelo em relação aos seus erros (SOUZA; SAMOHYL, 2005). “Um exemplo clássico é a previsão do preço de uma ação – geralmente a melhor previsão para o preço de amanhã é o preço de hoje” (BARROS, 2003).

2.5.2 Média Móvel Simples (MMS)

Segundo Bertolde e Xavier Junior (2013), o modelo de média móvel “[...] consiste em calcular a média aritmética de uma série de n observações”, utiliza os dados mais recentes da série e ao avançar dos períodos vão se substitui-se os dados mais antigos, ou seja, retira-se o último valor com a inserção de um novo.

Para Tubino (2009, p. 20) “[...] o número de períodos incluídos no cálculo da média móvel determina sua sensibilidade com relação aos dados mais recentes”, dessa forma, percebe-se que ao trabalhar com pequenos períodos tem-se uma reação mais rápida a oscilações de demanda, já em períodos mais longos essa resposta é menor, na qual, existe uma relação mais homogênea.

Moreira (2011, p. 312) considera que a MMS produz bons resultados quando aplicada em uma série estacionária, uma vez que ao ser aplicada em uma série com tendência, fornecerá previsões defasadas em relação aos dados reais. Da mesma forma quando houver sazonalidade, também não é uma boa aplicação, pois pode-se não captar as variações e ocultar o padrão sazonal nas previsões. Lustosa *et al*. (2008, p. 61) afirma que nesses casos “[...] uma alternativa melhor são os métodos de suavização”.

Hyndman e Athanasopoulos (2017) afirmam ainda que “[...] é possível aplicar uma média móvel a uma média móvel”, principalmente com um intuito de simetria em médias móveis de ordem par. Segundo Tubino (2009), a média móvel pode ser obtida através da Equação 1.

$$MMS_n = \frac{\sum_{i=1}^{n} Di}{n} \quad (1)$$

Em que:

$MMS_n =$ Média móvel de n períodos;

D_i = Demanda ocorrida no período i;

n = Número de períodos;

i = índice do período (i=1, 2, 3, ...).

2.5.3 Média Móvel Ponderada (MMP)

De acordo com Davis *et al*. (1999), a média móvel ponderada possui similaridades com a média móvel simples, pois utilizam um número de n valores reais que são retirados dos dados históricos da demanda, a diferença entre MMP e MMS é o peso que os valores recebem, enquanto que na simples os valores recebem pesos iguais ou simplesmente não é atribuído peso, nas ponderadas se atribuem pesos relativos "[...] geralmente refletindo uma maior importância dada aos valores mais recentes da demanda" (MOREIRA, 2011, p. 312). Fernandes e Godinho Filho (2010) apresentam o cálculo dessa média conforme a Equação 2.

$$MMP_T = \frac{\sum_{i=T-n+1}^{T} W_i D_i}{n} \quad (2)$$

Em que:

$MMP_T =$ Média móvel ponderada para o período T;

D_i = Demanda ocorrida no período i;

W_i = Peso atribuídos aos valores da demanda ocorrida no período i;

n = Número de períodos;

i = índice do período (i=1, 2, 3, ...).

Ainda para os mesmos autores, é importante ressaltar que o somatório dos pesos W_i deve ser igual a um, ou seja: $\sum_{i=T-n+1}^{T} W_i = 1$.

Para Moreira (2011) a média ponderada possui a vantagem de que é possível observar uma certa tendência em cima dos dados recentes da demanda e por isso recebem um peso maior, em que se varia entre dados antigos e recentes. Mas há uma restrição, pois, quanto maior for o número de dados, mais a previsão suavizará os efeitos sazonais, e, como consequência, responderá mais lentamente a variações que a demanda possa sofrer, Davis *et al.* (1999) restringe ainda, que quem cria o modelo terá que definir os pesos a serem utilizados.

2.5.4 Simples Suavização Exponencial (SSE)

Modelos de suavização exponencial se determinam por aplicarem pesos diferentes aos dados históricos de forma com que a relevância dos dados decaia exponencialmente dos mais recentes aos mais antigos. Existem parâmetros de suavização que precisam ser determinados, pois eles que definem a divisão dos pesos a serem atribuídos a cada dado histórico (MAKRIDAKIS *et al*., 1998).

A suavização exponencial consiste basicamente em "[...] uma média ponderada que atribui pesos decrescentes de forma exponencial para observações mais antigas" (BERTOLDE; XAVIER JUNIOR, 2013) e pode ser representada pela Equação 3.

$$\hat{y}_{t+1} = \hat{y}_t + \alpha(y_t - \hat{y}_t) = \alpha y_t + (1 - \alpha)\hat{y}_t \quad (3)$$

Em que:

$\hat{y}_{t+1}$ = Previsão de demanda para o período de tempo t+1;

y_t = Valor real de demanda em um período de tempo t;

$\hat{y}_t$ = Previsão de demanda para um período de tempo t;

α = Constante de suavização, (0<α<1).

A previsão pode ser obtida através de uma soma ponderada entre o valor de x_t e a previsão F_t, descrito de modo recursivo através da Equação 4, em que fica claro a acentuação do peso para o valor x_t e o decrescimento exponencial para os demais (NOGUEIRA, 2017).

$$F_{t+1} = \alpha x_t + \alpha(1 - \alpha)F_{t-1} + (1 - \alpha)^2 F_{t-2} + \cdots \quad (4)$$

2.5.5 Dupla Suavização Exponencial (DSE) ou Método de Holt

É um método similar ao SSE, difere-se no sentido de que, agora, utiliza-se a previsão de 1ª Ordem obtida na suavização simples, sem tendência ou sazonalidade, como base para a próxima suavização, dessa forma, caracteriza uma suavização de se 2ª Ordem, em que se considera além da variável de nível, uma para a tendência nos cálculos (MOREIRA, 2011).

O método de suavização exponencial simples, tende a não apresentar bons resultados quando se trata de uma série com tendência, aplica-se nesses casos a suavização dupla, pois essa "[...] usa duas constantes e é melhor para se manipular tendências (BERTOLO, 2017).

Representadas, segundo Bertolde e Xavier Junior (2013), pelas Equações 5, 6 e 7.

$$L_t = \alpha y_t + (1-\alpha)(L_{t-1} + T_{t-1}) \quad (5)$$

$$T_t = \gamma(L_t - L_{t-1}) + (1-\gamma)T_{t-1} \quad (6)$$

$$\hat{Y}_{t+1} = L_t + pT_t \quad (7)$$

Em que:

$\hat{Y}_{t+1}$ = Previsão de demanda para o período de tempo $t+p$;

y_t = Valor real de demanda em um período de tempo t;

L_t = Valor suavizado de x_t no período de tempo t;

T_t = Estimativa da tendência no período de tempo t;

α = Constantes de suavização para os dados da série, ($0<\alpha<1$);

γ = Constantes de suavização para a estimativa da tendência, ($0<\gamma<1$).

2.5.6 Método de Holt-Winters

O modelo Holt-Winters é uma extensão do método de Holt, indicado quando a série de dados apresenta, além de tendência, também sazonalidade. Utiliza-

se de três equações, “Uma para o nível, outra para tendência e outra para sazonalidade”. A sazonalidade pode ter dois efeitos em cima dos dados históricos, efeito multiplicativo ou aditivo, o que gera duas variações do método (SAMOHYL *et al*., 2001).

A partir do pressuposto que a amplitude da variação da sazonalidade permanece constante no período de tempo t (PELLEGRINI, 2000), de forma que a variável tendência independa das demais, considera-se o método aditivo o mais indicado. Já o método multiplicativo é mais apropriado quando existe uma variação sazonal proporcional ao longo da série (HYNDMAN; ATHANASOPOULOS, 2017).

Samohyl *et al*. (2001) apresentam quatro equações, as três já descritas e uma para a previsão, para os dois tipos de modelo, Holt-Winters Aditivo (HWA) e Holt-Winters Multiplicativo (HWM). As Equações 8, 9, 10 e 11 representam o modelo HWA e o modelo WHM é representado pelas Equações 12, 13, 14 e 15.

Nível	$L_t = \alpha(Y_t - S_{t-s}) + (1 - \alpha)(L_t - 1 + T_{t-1})$	(8)
Tendência	$T_t = \gamma(L_t - L_{t-1}) + (1 - \gamma)T_{t-1}$	(9)
Sazonalidade	$S_t = \delta(Y_t - L_t) + (1 - \delta)S_{t-s}$	(10)
Previsão	$\hat{Y}_{t+m} = L_t + T_{t-m} + S_{t-s+m}$	(11)

Nível	$L_t = \alpha(\frac{Y_t}{S_t} - s) + (1 - \alpha)(L_{t-1} + T_{t-1})$	(12)
Tendência	$T_t = \gamma(L_t - L_{t-1}) + (1 - \gamma)T_{t-1}$	(13)
Sazonalidade	$S_t = \delta(\frac{Y_t}{L_t}) + (1 - \delta)S_{t-s}$	(14)
Previsão	$\hat{Y}_{t+m} = (L_t + T_{t-m})S_{t-s+m}$	(15)

Em que:

s = Comprimento da sazonalidade;

L_t = Nível da série;

T_t = Tendência;

S_t = Componente sazonal;

$\hat{Y}_{t+m}$ = Previsão para o período m adiante;

Y_t = Valor observado;

α, γ e δ são parâmetros exponenciais alisadores, do nível, da tendência e da sazonalidade, respectivamente.

2.5.7 ARIMA

Em teoria o modelo ARIMA (*Auto Regressive Integrated Moving Averages*), ou metodologia Box-Jenkins, referente aos estatísticos que o elaboraram, é um dos modelos mais gerais quando comparado aos outros tipos de previsão de série temporal, indicado para ser usado em séries com tendência e sazonalidade. Seu ajuste leva em consideração acrescentar as defasagens das séries diferenciadas e dos erros de previsão para a equação de previsão para poder retirar autocorrelação dos erros (BERTOLDE; XAVIER JUNIOR, 2013).

Os modelos de suavização exponencial e ARIMA são os mais utilizados para a previsão, ambos diferem na sua abordagem. Enquanto a metodologia de suavização considera fatores de tendência e sazonalidade nos cálculos, o ARIMA implica em analisar as autocorrelações existentes entre os dados após estabilizá-los, ou transformar a série em um modelo ARMA. (HYNDMAN; ATHANASOPOULOS, 2017).

O mesmo autor explica que a modelagem ARIMA utiliza o conceito de série estacionária, de forma sucinta, que não possui tendência ou sazonalidade, dessa forma, o método compreende ajustar modelos que estabilizem os dados para que se possa aplicar as previsões, em termos práticos, diferencia-se os dados até que a série se torne estacionária, o que corresponde ao “I” do modelo, acrescentado ao modelo ARMA estacionário. Isso permite que uma estimativa de y_t, possa ser explicada por valores defasados da própria variável y em termos dos seus erros.

O intuito é escolher modelos ARIMA que melhor se ajustem a série de dados estudada e posteriormente, realimentar tais modelos, para que esses, de certa forma, possam aprender. Deve-se também, prosseguir de forma sucessiva e constante essa iteração, com o propósito de monitoramento e minimização dos erros (QUEIROZ *et al.,* 2012).

Segundo Bertolde e Xavier Junior (2013), dentro do método existem, então, classificações sobre o tipo de modelo a ser utilizado, baseadas na escolha de

parâmetros do tipo p, d, q que minimizem os erros, ou que melhor representem a série, para que se utilize em previsões. Em que p, d, q são respectivamente, o número de termos autorregressivos, o número da ordem de diferenciação necessária para estabilizar a série e o número de erros de previsões defasadas na equação de predição. Esses parâmetros devem ser os mais simples possíveis, ou seja, valores baixos, para que o modelo tenha um baixo número de termos, considera-se sempre a minimização do erro (SMITH, 2015).

A metodologia Box & Jenkins, aborda critérios para identificação e adequação do modelo, na qual exigem maior experiência do analista ao analisar as funções de autocorrelação (FAC) e funções de autocorrelação parciais (FACP). Embora existam outros métodos de estimação, como AIC (*Akaike Information Criterion*) e BIC (*Bayesian Information Criterion*), mais necessários para uma maior quantidade de dados, ou uma série longa (NIST, 2017). Segundo Fischer (1982), a estrutura geral do método ARIMA (p, d, q), pode ser representado conforme a Equação 16.

$$\Phi_{(\mathrm{B})}\Delta^{\mathrm{d}}Z_t = \Theta_{(B)}a_t \tag{16}$$

Em que:

$\Phi_{(B)} =$ Operador auto-regressivo de ordem p;

$\theta_{(B)} =$ Operador de médias móveis de ordem q;

$a_t =$ Ruído branco d representa o número de diferenças;

$\nabla = 1 - B$ = Operador diferença.

Segundo Franco (2011), de modo geral, obtém-se o parâmetro d, a partir da quantidade de diferenciações necessárias para estacionar a série. A escolha de p e q, derivam do comportamento das funções FAC e FACP dos dados, em comparação à modelos teóricos. Geralmente em um modelo estacionário $p + d \leq 2$. No Quadro 1, podem ser vistos alguns modelos conhecidos, e seus comportamentos esperados.

Modelo	FAC	FACP
MA (1)	1 pico no *lag* 1	Decrescimento exponencial
AR (1)	Decrescimento exponencial	1 pico no *lag* 1
MA (2)	1 pico no *lag* 1 e 1 pico no *lag* 2	Mistura de exponenciais e ondas senóides amortecidas
AR (2)	Mistura de exponenciais e ondas senóides amortecidas	1 pico no *lag* 1 e 1 pico no *lag* 2
ARIMA (1,1)	Decrescimento exponencial	Decrescimento exponencial

Quadro 1 - Comportamento teórico da FAC e FACP para alguns modelos.
Fonte: Adaptado de Franco (2011, p. 45).

Tem-se ainda, a observação de que os componentes do modelo ARIMA, p, d e q, tem relação direta com os termos vistos no Quadro 1. Os modelos AR se relacionam com o termo p (termo autorregressivo), o I é o filtro de integração que caracteriza o modelo ARIMA, tem como coeficiente o d, ou seja, o número de diferenciações, e o modelo de MA, componente de médias móveis, tem relação com o coeficiente q. Dessa forma, cada componente pode ser um modelo isolado e a junção desses filtros fornece o modelo ARIMA (SANTOS, 2014).

2.6 PREVISÕES BASEADAS EM MÉTODOS CAUSAIS

Segundo Tubino (2009), os métodos causais baseiam-se em relacionar fatores que possam influenciar na demanda, para agregar à previsão. Lustosa *et al*. (2008) afirma que a mesma está correlacionada a fatores independentes, com graus de previsão como simples e múltipla, leva-se em consideração o número de variáveis que se utilizam na previsão.

Os métodos casuais, correlação ou regressão são uma das técnicas mais usuais e sua aplicabilidade está associada a técnicas quantitativas, por meio de uma correlação dos dados ao longo do tempo, subentende-se, que uma série temporal pode ser predita através de um modelo de tendência ou mesmo no caso específico de ausência de tendência, ou seja, estacionária (FRANCO, 2016).

2.6.1 Regressão Linear Simples

Em um diagrama de dispersão pode-se observar com certa frequência que o agrupamento dos dados tende a um tipo de curva. Se o ajustamento desses dados tender para uma linha reta como na Figura 9, determina-se que existe uma relação linear entre as variáveis (SPIEGEL, 1993, p.330).

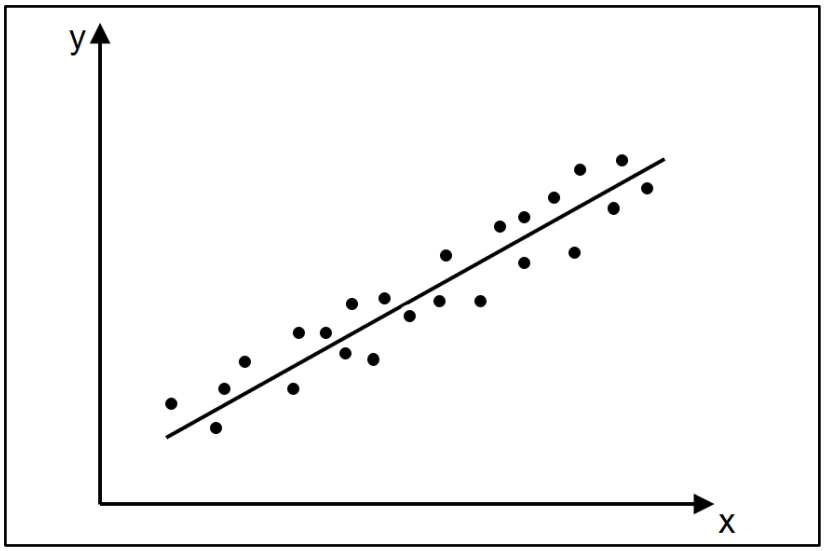


Figura 9 - Exemplo de ajustamento linear.
Fonte: Adaptado de Spiegel (1993).

A regressão linear simples tem por característica o coeficiente angular dado pela tangente do ângulo formado pela figura plotada dos dados da série, em que o coeficiente angular pode ser tanto negativo quanto positivo, conforme Figura 10, porém, o método gráfico pode conduzir a erros, “[...] já que a colocação da reta no papel depende do julgamento humano. Duas pessoas podem obter valores diferentes para os parâmetros procurados”. Existem outros métodos, o método dos mínimos quadrados (MMQ) é o mais usual deles (MOREIRA, 2011).

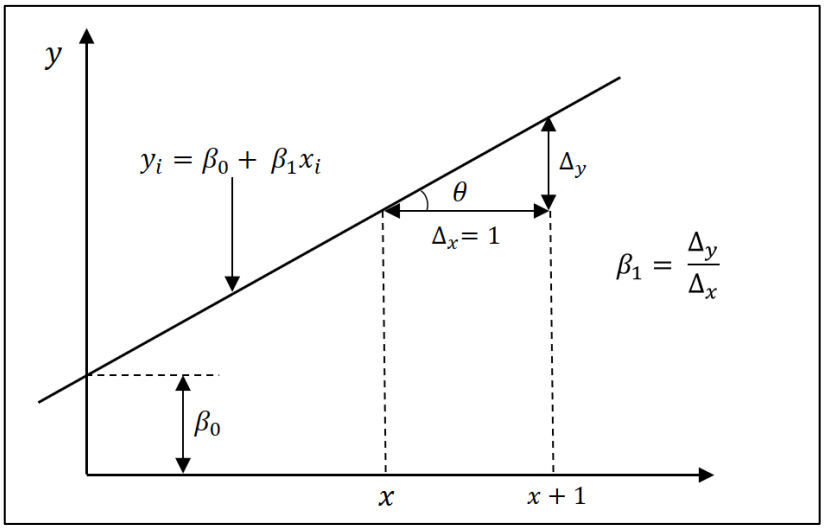


Figura 10 - Significado dos parâmetros da reta de regressão.
Fonte: Adaptado de Moreira (2011).

Devore (2013) define a equação do modelo de regressão linear simples conforme relacionada na Equação 17.

$$y = \beta_0 + \beta_1 x + \varepsilon \tag{17}$$

Em que:

y = Variável dependente;

x = Variável independente;

β_0 = Corresponde ao período inicial da série histórica;

β_1 = Inclinação da linha de tendência;

ε = Variável aleatória do erro.

Segundo Anderson *et al*. (2011, p. 429) “O Termo de erro é responsável pela variabilidade em y que não pode ser explicada pela relação linear entre x e y”. Os parâmetros β_0 e β_1 da equação, geralmente não são conhecidos, dessa forma, na prática, procura-se uma estimativa através de dados amostrais que represente os parâmetros populacionais. São esses denotados por b_0 e b_1. A equação de regressão estimada, então, é obtida por meio da substituição desses parâmetros na Equação 17, conforme apresentado na Equação 18.

$$\hat{y}_i = b_0 + b_1 x \tag{18}$$

Em que, $\hat{y}_i$ é o valor estimado de y.

De acordo com Moreira (2011), a Equação 18, portanto, resultará em uma previsão de como se comportarão os valores de $\hat{y}_i$, ou seja, da demanda, e não os valores reais que ela apresentará, a diferença entre o valor real e previsto ($y_i - \hat{y}_i$), em que i = 1, 2, 3... n) será o erro da previsão. O somatório do erro de cada previsão terá uma tendência à zero, quando elevado ao quadrado os números resultantes sempre serão positivos e podemos assim calcular o erro quadrático total, ou em termos mais matemáticos, segundo Devore (2013, p. 445), a "[...] soma dos quadrados dos erros (que equivale a soma dos quadrados dos resíduos)", representado por SQE, de acordo com a Equação 19.

$$SQE = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{19}$$

Utiliza-se o método dos mínimos quadrados (MMQ) para encontrar uma equação de reta que melhor se ajuste aos valores da série. O método consiste em encontrar a reta que minimize o somatório das distâncias entre o valor real da série e os pontos da própria reta, conforme podem ser exemplificados na Figura 11 (PEINADO; GRAEML, 2007).

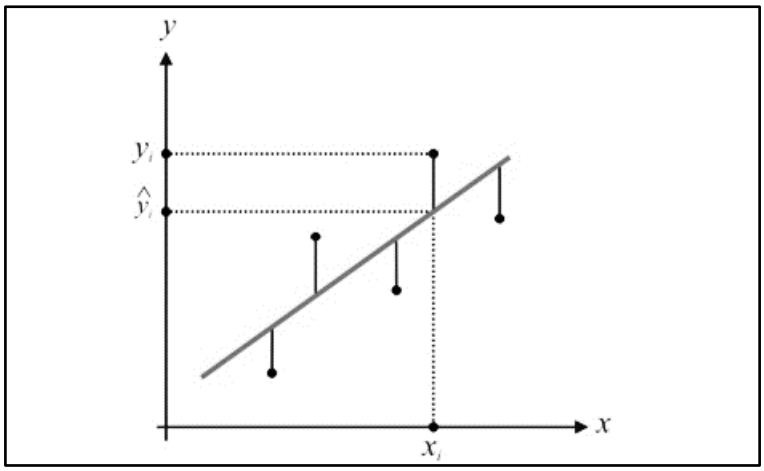


Figura 11 - Ilustração do método dos mínimos quadrados (MMQ).
Fonte: Adaptado de Moreira (2011).

O critério do método MMQ é, portanto, encontrar os parâmetros b_0 e b_1, que minimizem a Equação 19, para isso normalmente utiliza-se cálculo diferencial, chegando-se de forma sucinta nas Equações 20 e 21 (ANDERSON *et al.*, 2011).

$$b_1 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \quad (20)$$

$$b_0 = \bar{y} - b_1\bar{x} \quad (21)$$

Em que:

y_i = Valor da variável dependente para a i- ésima observação;

x_i = Valor da variável independente para a i- ésima observação;

$\bar{y}$ = Valor médio da variável dependente, dado por: $\bar{y} = \sum\frac{y_i}{n}$;

$\bar{x}$ = Valor médio da variável independente, dado por: $\bar{x} = \sum\frac{x_i}{n}$;

n = número total de observações.

2.6.2 Regressão Não-Linear

Considera-se que o ajustamento de dados agrupados em um diagrama de dispersão pode não seguir um padrão linear, conforme o diagrama exemplificado na Figura 12. Ajusta-se esses dados, portanto, a outros tipos de curvas, denominadas relações não-lineares (SPIEGEL, 1993, p.330).

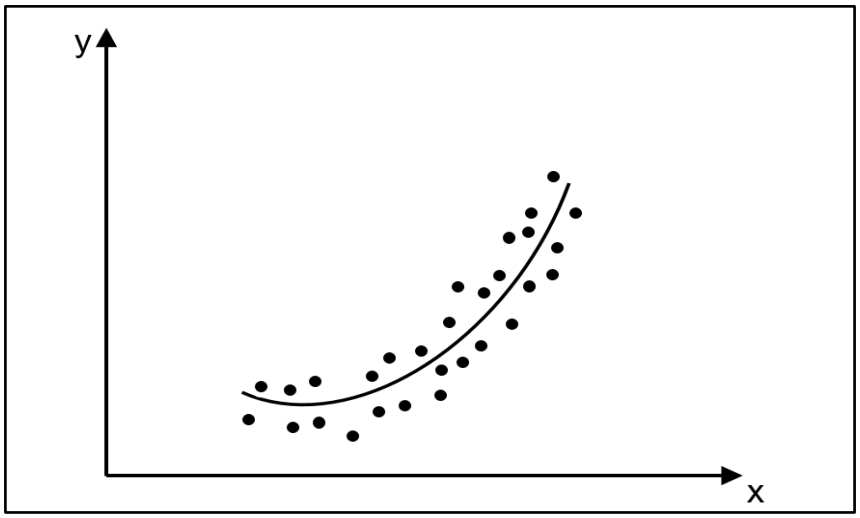


Figura 12 - Exemplo de ajustamento não-linear.
Fonte: Adaptado de Spiegel (1993).

Dessa forma, o intuito é adequar, através de diferentes métodos, outros tipos de equações para representar os dados analisados. As principais formas de curvas apresentadas por Spiegel (1993), incluindo a equação linear, podem ser representadas de forma simplificada, conforme o Quadro 2.

Equações	Nome
$y = a_0 + a_1 x$	Linha reta ou Linear
$y = a_0 + a_1 x + a_3 x^2$	Curva do 2º grau ou Quadrática
$y = a_0 + a_1 x + a_3 x^2 + a_4 x^3$	Curva do 3º grau ou Cúbica
$y = a_0 + a_1 x + a_3 x^2 + \cdots + a_n x^n$	Curva de grau n
$y = \frac{1}{a_0 + a_1 x}$	Hipérbole
$y = ab^x$	Curva exponencial
$y = a \ln x + b$	Curva Logarítmica
$y = ax^b$	Curva geométrica
$y = ab^x + g$	Curva de Gompertz
$y = a_0 + a_1(\log x) + a_2(\log x)^2$	Curva Logística

Quadro 2 - Principais equações das curvas de ajustamento.
Fonte: Adaptado de Spiegel (1996).

2.6.3 Regressão Múltipla

Segundo Devore (2013), pode-se generalizar o modelo de regressão linear simples, substituindo seus parâmetros lineares, por uma função de regressão não-linear de x que compreenda mais de uma variável independente. O autor ressalta que "[...] antes de usar esses métodos, entretanto, o analista de dados primeiro verifica a validade do modelo escolhido". Interpreta-se que em certos momentos "A complexidade de muitos dos mecanismos científicos é tamanha que, para prevermos uma importante resposta, precisamos de um modelo de regressão múltipla*"* (WALPOLE *et al.,* 2009, p. 284).

Anderson *et al.* (2011, p. 488), definem regressão múltipla como "[...] o estudo de como a variável dependente y se relaciona com duas ou mais variáveis independentes". Sua utilização se dá de forma semelhante as regressões simples. Toma-se como exemplo a curva linear, pode-se entender melhor e generalizar para as demais. Uma regressão linear múltipla, pode ser estruturada com base na

suposição do modelo de regressão, conforme Equação 22, bem como a equação para a estimativa da regressão, exibida na Equação 23.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + \varepsilon \quad (22)$$

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_p x_p + \varepsilon \quad (23)$$

2.6.4 Coeficientes de correlação e de determinação

O conceito de correlação vem da interpretação da relação existente entre as variáveis, ou, basicamente o quanto a variável dependente depende da variável independente. Dessa relação também pode-se tirar o conceito de que, quando não existe relação entre duas variáveis, ou seja, sua correlação for nula, essas são independentes uma da outra (BISQUERRA *et al*., 2004).

Pode-se procurar entender de modo qualitativo, quão bem correlacionadas estão as variáveis ou o quanto a equação da curva ajustada representa essa relação, apenas com a observação do gráfico de dispersão entre elas. Mas se for preciso tratar de maneira quantitativa, se faz necessário o uso de medidas de correlação (SPIEGEL, 1993, p. 360). Segundo Walpole *et al.* (2009, p. 275) “A *análise de correlação* tenta medir a força de tais relações entre duas variáveis por meio de um único número”. Moreira (2011) propõe que o comportamento de ordem linear das variáveis x e y, obtido por meio da análise da plotagem da série dos dados históricos, pode ser medida por meio de um coeficiente de correlação.

A medida mais comum utilizada para estabelecer correlações lineares na estatística, é o coeficiente de correlação linear de Pearson (r), ele mostra o quão linearmente os dados estão relacionados, ou de forma sucinta, indica a possibilidade em representar um conjunto de dados por meio de uma reta (GLEN, 2012).

Segundo Anderson *et al*. (2011, p. 102), esse coeficiente pode ser obtido através da Equação 24.

$$r_{x,y} = \frac{s_{x,y}}{S_x S_y} \quad (24)$$

Em que:

$r_{x,y}$ = Coeficiente de correlação da amostra;

$S_{x,y}$ = Covariância da amostra, dada por: $S_{x,y}=\sum \frac{(x_i-\bar{x})(y_i-\bar{y})}{n-1}$;

S_x = Desvio padrão da amostra de x: $S_x=\sqrt{\sum \frac{(x_i-\bar{x})^2}{n-1}}$;

S_y = Desvio padrão da amostra de y: $S_y=\sqrt{\sum \frac{(y_i-\bar{y})^2}{n-1}}$.

Devore (2013, p. 470) atribui 5 propriedades como as mais importantes para o coeficiente de correlação de Pearson. As quais são essas:

1. O valor de r não depende de qual das duas variáveis é adotada como dependente ou independente.
2. O valor de r independe da unidade das variáveis.
3. O valor de r está entre -1 e 1 ($-1 \leq 0 \leq 1$).
4. $r = -1$, se e somente se, todos os pontos estiverem alinhados e com o mesmo coeficiente angular negativo e $r = 1$, se e somente se, todos os pontos estiverem alinhados e com o mesmo coeficiente angular positivo.
5. O quadrado de r fornece o valor do coeficiente de determinação (r^2) que resulta do modelo de regressão linear simples.

Bisquerra *et al.* (2004, p. 147) sugere os seguintes critérios para a interpretação do r, conforme Quadro 3.

Coeficiente	Interpretação
$r = 1$	Correlação perfeita
$0{,}80 < r < 1$	Muito alta
$0{,}60 < r < 0{,}80$	Alta
$0{,}40 < r < 0{,}60$	Moderada
$0{,}20 < r < 0{,}40$	Baixa
$0 < r < 0{,}20$	Muito baixa
$r = 0$	Nula

Quadro 3 - Critérios para a interpretação dos valores de r.
Fonte: Adaptado de Bisquerra *et al*. (2004).

Por meio do ajuste de uma curva nos dados plotados, tira-se um coeficiente para auxiliar a interpretação da qualidade desse ajuste, de forma similar ao coeficiente linear de Pearson (r), o coeficiente de determinação (R^2), é utilizado para uma quantificação do quão bem ajustados ou representativos são os parâmetros ajustados (QUININO *et al*., 2011).

Moreira (2011) apresenta o R^2 como um valor compreendido no intervalo entre 0 e 1, justamente por ser o quadrado do r, o uso desse coeficiente em vez do de correlação é aconselhável pelo autor, pois apresenta valores mais conservadores e uma interpretação mais interessante. O autor explica ainda, que o coeficiente de determinação R^2 "[...] é interpretado como a proporção de variância comum entre y e x " dessa forma, se o valor de R^2 for de 0,25, por exemplo, interpreta-se que 25% da variação de y é explicado pela variação do valor de x.

Segundo Anderson *et al*. (2011), o cálculo do coeficiente de determinação pode feito através dos cálculos apresentados na Equação 25.

$$R^2 = \frac{SQR}{SQT} = 1 - \frac{SQE}{SQT} = \frac{\Sigma(\hat{y}_i - \bar{y})^2}{\Sigma(y_i - \bar{y})^2} \quad (25)$$

Em que:

R^2 = Coeficiente de determinação;

SQR = Soma dos quadrados da regressão;

SQT = Soma dos quadrados totais;

SQE = Soma dos quadrados dos erros (resíduos).

Segundo Anderson et al. (2011, p. 443), com base no pressuposto de que já se tenha feito a análise de regressão e obtido o valor de r², pode-se obter o coeficiente de correlação da amostra conforme Equação 26.

$$r_{x,y} = (sinal\ de\ b_1)\sqrt{R^2} \quad (26)$$

Vale ressaltar que tanto o coeficiente de correlação de Pearson, quanto para a regressão linear, cabe a aplicação de teste de hipótese, se satisfeitas as condições para tal, assim como intervalo de confiança, o R^2 pode é aplicado também

as regressões múltiplas, seção 4.6.3 (MARTINS, 2002).

Considera-se ainda, no caso da regressão múltipla, a utilização do coeficiente de determinação ajustado (R^2_a), para evitar "[...] uma superestimação do impacto de se adicionar uma variável independente à quantidade de variabilidade explicada pela equação de regressão estimada" (ANDERSON *et al.* 2011, p. 498). Dado pela Equação 27.

$$R_a^2 = 1 - (1 - R^2)\frac{n-1}{n-p-1} \quad (27)$$

Em que:

R_a^2 = Coeficiente de determinação ajustado;

R^2 = Coeficiente de determinação;

n = Número de variações;

p = Número de variáveis independentes.

2.7 MONITORAMENTO DO MODELO E ERROS DE PREVISÃO

Para Ehlers (2007) "[...] ao comparar sequencialmente as previsões com os valores observados pode-se julgar a adequação relativa de modelos alternativos com base em sua performance preditiva". O suporte do Minitab (2017) e o Portal Action (2017), utilizam em seus *softwares*, três medidas principais para informar os erros relativos aos métodos de previsão. São essas, as estatísticas MAPE (*Mean Absolute Percentage Error* ou Erro Percentual Absoluto Médio), MAD (*Mean Absolute Deviation* ou Desvio Absoluto Médio) e MSD (*Mean Squared Deviation* ou Desvio Quadrado Médio).

Segundo Devore (2013), essas medidas são uteis para a comparação dos métodos em função da qualidade do ajuste aos dados. Nessas três formas de medir os erros, quanto menor o valor menor o erro. Para Tubino (2009 p.16) os erros de previsão possuem ainda relação com o tipo de sistema que está é utilizado, por exemplo, nos sistemas em massa ou contínuos, têm-se um erro menor, pois a demanda é mais estável e geralmente possuem uma grande quantidade de dados disponíveis sobre certo produto sem grandes mudanças em sua estrutura.

Jacobs *et al*. (2014) e Bacci (2007) utilizam tais erros em seus trabalhos. Segundo Morettin e Toloi (1985), os erros podem ser calculados, conforme as Equações 28, 29 e 30.

$$MAPE = \frac{100}{n}\sum_{t=1}^{n}\left|\frac{(y_t-\hat{y}_t)}{y_t}\right| \quad ,para\ y_t \neq 0 \tag{28}$$

$$MAD = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t| \tag{29}$$

$$MSD = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t|^2 \tag{30}$$

Em que, $(y_t - \hat{y}_t)$ é o erro de previsão.

Utiliza-se também o índice U de Theil, em que é possível efetuar uma comparação dos resultados em relação a previsão ingênua, pois pressupõe–se que qualquer modelo de previsão de série temporal, no mínimo, deva apresentar um desempenho superior ao método trivial para que sua utilização seja considerada viável (GUTIÉRREZ, ROMANEL, 2003).

Segundo Lobo (2007) e Silva e Dantas (2009), o coeficiente pode ser obtido através da Equação 31.

$$UTheil = \frac{\sqrt{\sum_{j=1}^{N}(A_t-F_t)^2}}{\sqrt{\sum_{j=1}^{N}(A_t-A_{t-1})^2}} \tag{31}$$

Em que:

A_t = Valor atual;

F_t = Valor predito.

Interpreta-se:

U = 1, significa que o modelo é equivalente à previsão ingênua;

U > 1, significa o modelo é pior que a previsão ingênua;

U < 1, significa o modelo é melhor que a previsão ingênua.

U = 0, significa modelo é perfeito para a previsão.

2.7.1 Análise de resíduos

Para validação do modelo ajustado, ou seja, para que esse possa ser efetivamente utilizado para previsões, além de representar que o analista efetuou as avaliações com considerações éticas. A aplicação desse modelo “[...] depende do bom cumprimento de um conjunto de pressupostos” (LEVINE *et al.,* 2000, p. 541).

Montgomery e Runger (2012, p. 278), ressaltam que se faz importante a analises dos resíduos ($y_i - \hat{y}_i$), para atestar e validar “[...] a suposição de que os erros sejam distribuídos de forma aproximadamente normal, com variância constante”, conforme as suposições iniciais do modelo, somente assim, o que possibilita seu uso para projeções.

De forma geral, Walpole (2009), Navidi (2012) e Devore (2013), definem os como pressupostos para os resíduos que:

a) Sejam independentes ou que possuam ausência de autocorrelação
b) Sigam uma distribuição normal
c) Sejam homocedásticos ou possuam uma dispersão regular em torno de uma média constante.

3 MATERIAL E MÉTODOS

Nesta seção são explicados os detalhes referentes a metodologia da pesquisa cientifica a qual pretende-se efetuar, para que se caracterize o método cientifico, pois esse “[...] é fundamental para validar as pesquisas e seus resultados serem aceitos” (OLIVEIRA, 2011, p. 8).

O modelo seguido para este trabalho pode ser definido conforme as dez etapas propostas por Russel e Taylor (2011, p. 502), representadas na Figura 13.

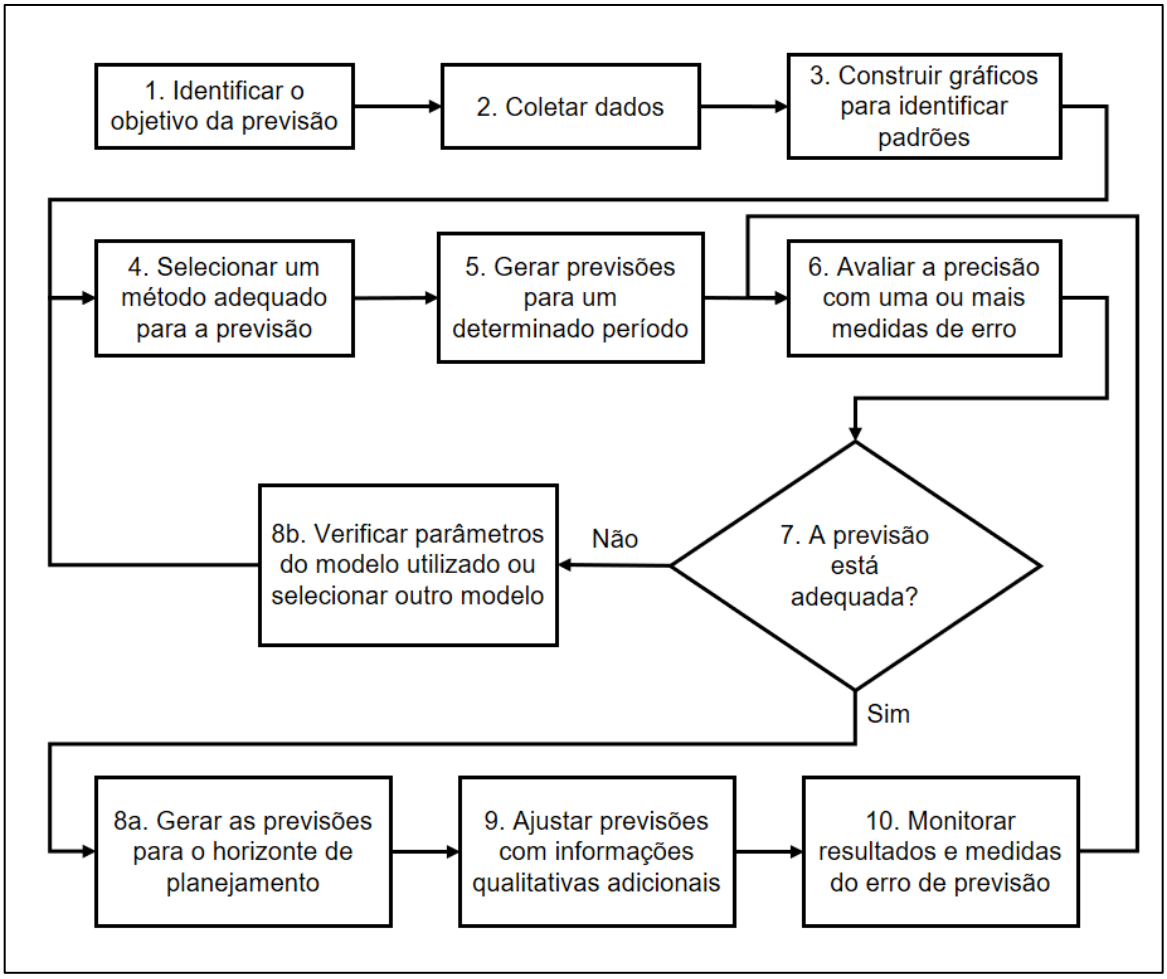


Figura 13 - Fluxograma do processo de escolha de um modelo para previsão de demanda. Fonte: Adaptado de Russel e Taylor (2011).

Na primeira etapa do modelo deve-se definir o objeto de estudo, ou objetivo, também se caracteriza os fatores que fazem essa previsão ser necessária, seu grau de acuracidade, em relação a importância. Para este trabalho o objeto subentende-se como a demanda de café no Brasil, devido a sua grande importância,

explicada nos itens introdutórios que justificam a pesquisa, bem como em sua revisão literária.

Na segunda etapa o objetivo é realizar a coleta e análise dos dados. Os dados foram coletados conforme explica-se na seção 3.2, de modo que se possibilitou fazer a análise, a qual, subentende-se como o principal quesito da pesquisa, pois é ela que visa dar suporte ao propósito já estipulado na primeira etapa, por meio da comparação dos métodos.

Com os dados devidamente organizados, confecciona-se o gráfico da série histórica e sua estatística descritiva, para entender o comportamento da série, o que compete a etapa 3. De forma a identificar possíveis fatores de tendência ou sazonalidade, que serão utilizados para os ajustamentos dos modelos matemáticos.

Para as etapas 4 e 5, efetua-se comparações, foram elaborados modelos de previsão com os métodos referidos, o que permitiu a execução de previsões para os períodos existentes, e também que se aplique a etapa 6, na qual os modelos foram avaliados quanto aos critérios propostos. Explica-se os detalhes desses critérios no item 3.3.

Na etapa 7, tem-se uma decisão em relação as avaliações feitas na etapa anterior, aqui foram testadas as precisões dos métodos e quais podem ser utilizados ou não, em relação aos fatores de relevância estipulados na etapa 1. Caso atenda aos critérios, o método é utilizado conforme a etapa 8a, caso contrário, procura-se um método substituto que melhor se adeque, liga-se a etapa 8b à etapa 4, para retomar o processo.

Aos modelos que seguem adiante no teste, para a etapa 8a, quando as previsões são efetuadas conforme desejado e seguem para a etapa 9, com o objetivo de melhorá-las com informações qualitativas, das quais, apesar de não serem o foco desse trabalho, são úteis como complemento para os métodos quantitativos empregados, principalmente para a possibilidade de aumentar o período de previsão.

Por fim, na etapa 10, monitora-se as previsões efetuadas com base nos erros vistos nas seções 2.6.4, 2.6.5 e 3.3, para interpretar seus dados a fim de realimentar o modelo com informações e testá-los para refinar os resultados.

3.1 CLASSIFICAÇÃO DA PESQUISA

Antes de caracterizar a pesquisa, apresentam-se algumas definições propostas na literatura para facilitar posterior entendimento. Para Gil (2010, p. 1) “Pode-se definir pesquisa como o procedimento racional e sistemático que tem como objetivo proporcionar respostas aos problemas que são propostos”. Busca-se assim, relacionar um problema ou questionamento inicialmente proposto, para que a pesquisa faça a junção entre esse problema, as informações encontradas e os resultados ou respostas obtidas.

De forma similar, Köche (2011, p. 106) sugere que “A ciência pode ser encarada como um processo de investigação que se interessa em descobrir a relação existente entre aspectos que envolvam fatos, situações, acontecimentos, fenômenos ou coisas [...]”.

Para encaixar o conceito, conforme a ideia de Prodanov e Freitas (2013, p. 14), definem metodologia como ”[...] a aplicação de procedimentos e técnicas que devem ser observados para construção do conhecimento, com o propósito de comprovar sua validade e utilidade nos diversos âmbitos da sociedade”. Dessa forma, pode-se entender que a ciência exige a utilização de métodos adequados para investigar e explicar um determinado problema, e que se traduzem em técnicas e procedimentos através da metodologia, a qual, identifica, explica e analisa tais métodos.

Percebe-se a necessidade de estabelecer padrões e definir as características que envolvem o trabalho, para que assim seja claro o entendimento da sua utilidade, as áreas de abrangência do conhecimento envolvido e os métodos utilizados. Existem inúmeras maneiras de se classificar uma pesquisa, em relação aos critérios estabelecidos, desde que coerentes com a clareza necessária para a identificação da pesquisa e suas explicações (GIL, 2010, p. 25).

Prodanov e Freitas (2013), abordam a classificação de uma pesquisa científica quanto a metodologia, de maneira geral, seguiu-se os critérios conforme vistos no Quadro 4.

Natureza	Abordagem	Objetivos	Procedimentos
Básica Pratica	Quantitativa Qualitativa	Exploratória Descritiva Explicativa	Pesquisa Bibliográfica Pesquisa Documental Pesquisa Experimental Pesquisa Operacional Estudo de caso Pesquisa Ex-Post-Facto Pesquisa de Levantamento Pesquisa Participante Pesquisa-Ação

Quadro 4 - Classificação das pesquisas científicas.
Fonte: Adaptado de Prodanov e Freitas (2013).

Para este trabalho, de forma sucinta, definiu-se alguns métodos para previsão de demanda empregados dentro dos padrões de métodos científicos, utilizou-se o procedimento de escolha de modelos indicados para realizar a previsão da demanda, conforme Figura 13 e o método estatístico. Pois, com a utilização desse método, segundo Gil (2008, p. 17) “[...] torna-se possível determinar, em termos numéricos, a probabilidade de acerto de determinada conclusão, bem como a margem de erro de um valor obtido”.

Pode-se dizer que entre os modelos ajustados foi aplicado o método comparativo, no qual busca-se, segundo Prodanov e Freitas (2013, p. 38), “[...] estudar semelhanças e diferenças, esse método realiza comparações com o objetivo de verificar semelhanças e explicar divergências”.

Por meio da linha clássica de classificação, pode-se dizer que a pesquisa proposta possui natureza básica no intuito de procurar fornecer uma previsão confiável para a demanda de café no Brasil, pois visa gerar modelos de previsão e testá-los de maneira a garantir confiabilidade para a utilização de tal método. Conforme Silva e Menezes (2005, p. 25) pesquisa de natureza básica é aquela que “[...] objetiva gerar conhecimentos novos úteis para o avanço da ciência sem aplicação prática prevista”.

Trata-se de uma pesquisa de abordagem quantitativa, uma vez que com o uso de ferramentas estatísticas buscará traduzir as informações sobre o assunto abordado de forma numérica e explicativa (KAUARK *et al*., 2010, p. 26). Possui objetivo exploratório, pois além de buscar a criação de modelos para a inferência dos dados, também visa aprofundar o conhecimento em diferentes áreas, tais como previsão de demanda, estatística, agroindústria, setor de café, entre outros. Gil (2010, p. 27) afirma que

> [...] a maioria das pesquisas realizadas com propósitos acadêmicos, pelo menos num primeiro momento, assume o caráter de pesquisa exploratória, pois nesse momento é pouco provável que o pesquisador tenha uma definição clara do que irá investigar.

Possui aspectos operacionais quanto ao seu procedimento, com a proposta de elaboração de modelos e o uso de simulação dos resultados para fins científicos. Para Cardoso (2011, p. 4) a pesquisa do tipo operacional utiliza-se de modelagem matemática "[...] com a finalidade de prever e comparar estratégias ou decisões alternativas, cujo objetivo é dar suporte à definição de políticas e determinação de ações".

Outro procedimento adotado refere-se a coleta de dados, por esse parâmetro, pode-se caracterizar a pesquisa como documental, a partir de documentação indireta, com a premissa de que os dados retirados de fonte primária serão compilados e analisados pelo autor, conforme explicados na seção 3.2.

Marconi e Lakatos (2013, p. 49), classificam a pesquisa documental de maneira estruturada, conforme mostrado no Quadro 5, com o intuito de esclarecer possíveis confusões em relação a definição da origem dos dados.

		ESCRITOS		OUTROS	
		PRIMÁRIOS	SECUNDÁRIOS	PRIMÁRIOS	SECUNDÁRIOS
CONTEMPORÂNEOS		Compilados na ocasião pelo autor	Transcritos de fontes primárias contemporâneas	Feitos pelo autor	Feitos por outros
		Exemplos Documentos de arquivos públicos Publicações parlamentares e administrativas Estatísticas (censos) Documentos de arquivos privados Cartas Contratos	**Exemplos** Relatórios de pesquisa baseados em trabalho de campo de auxiliares Estudo histórico recorrendo aos documentos originais Pesquisa estatística baseada em dados do recenseamento Pesquisa usando a correspondência de outras pessoas	**Exemplos** Fotografias Gravações em fita magnética Filmes Gráficos Mapas Outras ilustrações	**Exemplos** Material cartográfico Filmes comerciais Rádio Cinema Televisão
RETROSPECTIVOS		Compilados após o acontecimento pelo autor	Transcritos de fontes primárias retrospectivas	Analisados pelo autor	Feitos por outros
		Exemplos Diários Autobiografias Relatos de visitas a instituições Relatos de viagens	**Exemplos** Pesquisa recorrendo a diários ou autobiografias	**Exemplos** Objetos Gravuras Pinturas Desenhos Fotografias Canções Folclóricas Vestuário Folclore	**Exemplos** Filmes comerciais Rádio Cinema Televisão

Quadro 5 - Classificação de dados de pesquisa documental.
Fonte: Marconi e Lakatos (2013, p. 49).

3.2 COLETA DE DADOS

A coleta de dados se deu de maneira documental, extraiu-se informações dos bancos de dados de organizações sugestivas. Para o consumo, produção, preço em sacas de 60 kg e exportação de café no Brasil, coletou-se os valores da OIC (2017); dados na unidade de escala de milhares de sacas de 60 kg (1:1000); observou-se ainda, que a safra acontece entre um ano e outro, assim como, para nomear e facilitar a interpretação, adotou-se o ano de início como referência. Por exemplo, a safra 1990/91, nomeou-se para 1990, dessa forma, obtida uma série desde 1990 a 2016, com total de 27 anos.

Para os dados referentes à série histórica de preços do café no varejo no Brasil, buscou-se os dados da ABIC (2017), na qual encontram-se dados mensais desde 1996 a 2017. Para equalizar os dados calculou-se uma média anual desses preços no varejo. Em relação ao fator população, retirou-se as informações no banco de dados do IBGE (2017).

Com base no fato de que os dados coletados de forma indireta, seguem os parâmetros necessários para o estudo, buscou-se em fontes alternativas na literatura, comparações para confrontar e validar a coerência dos dados. Conforme sugere Ferrari (2004, p. 2), “[...] os dados obtidos devem ser criticados à procura de falhas sistemáticas no planejamento, aquisição e armazenamento dos dados”

3.3 PROCEDIMENTOS METODOLÓGICOS

Descreve-se nesta seção os procedimentos considerados para a composição da análise e estruturação dos resultados obtidos.

Inicialmente organizou-se os dados coletados em planilhas do *software* Microsfoft Excel® por meio do qual foi confeccionado o gráfico da série, para possibilitar a interpretação da disposição dos dados, bem como, a existência de estacionariedade, tendências ou padrões sazonais.

Realizou-se também uma análise de estatística descritiva dos dados, com os principais testes de significância e os parâmetros de testes utilizados pelos

softwares. Para normalidade utilizou-se o teste de Anderson-Darling. Para a estacionariedade da série, o teste mais relevante utilizado foi o de Dickey-Fuller Aumentado (ADF). Para a tendência, fez-se várias interpretações gráficas e ajustes de tendências. A qualidade do ajuste de regressão foi analisada pelos critérios comumente utilizados para a regressão e resíduos, comparadas por critérios técnicos, matemáticos e estatísticos, além do coeficiente de determinação (R^2). Ainda, para a sazonalidade, além da observação gráfica, utilizou-se o teste de Kruskal-Wallis.

Posteriormente, organizou-se os dados em planilhas, de forma que fosse possivel efetuar a aplicação dos métodos de séries temporais e causais sugeridos na literatura para previsão, seguiu-se basicamente a sequência disposta nas seções 2.5 e 2.6.

Para os métodos de séries temporais e causais, foram utilizados os *softwares* Minitab®, versão 17 e Action Stat®, versão 3.0. Outras análises foram concebidas a partir do próprio Excel®, versão 2016, devido a sua praticidade de utilização para a regressão e manipulação das fórmulas. Observa-se que, para efeito de teste, comparação e segurança, pode-se efetuar os mesmos testes em diferentes *softwares.*

Os modelos foram então testados, selecionados a partir da comparação dos critérios de precisão, Desvio Absoluto Médio (MAD - *Mean Absolute Deviation*) e coeficiente U de Theil. Estudou-se também, fatores causais interpretados pelas análises de regressão múltipla. Esses fatores foram analisados por critérios de interpretação crítica e pelo coeficiente de determinação ajustado (R_a^2). As fórmulas utilizadas podem ser visualizadas de forma resumida no Quadro 6 , e foram melhor esclarecidas na seção 2.6.4 e 2.6.5.

Critério	Equação
Desvio Absoluto Médio (MAD)	$MAD = \frac{1}{n}\sum_{t=1}^{n} \lvert y_t - \hat{y}_t \rvert$
Coeficiente U de Theil	$UTheil = \frac{\sqrt{\sum_{j=1}^{N}(A_t - F_t)^2}}{\sqrt{\sum_{j=1}^{N}(A_t - A_{t-1})^2}}$

Quadro 6 - Formulário dos critérios para comparação dos métodos.
Fonte: Autoria própria.

O critério MAD é utilizado como parâmetro pelos softwares Action Stat® e Minitab®, que foram empregados nas análises dos métodos, assim como, sua rápida confecção no Excel, tornou-se mais fácil a comparação entre métodos e *softwares*. Ao buscar métodos de comparação na literatura, observa-se uma grande utilização de tal critério; sua escala de grandeza assume a do produto medido, e os dados coletados estão representados em milhares de sacas de 60 kg, o MAD está na mesma escala de 1:1000.

Em relação ao coeficiente U de Theil, utilizou-se, devido a sua interpretação simples e prática, a qual permite uma noção de forma rápida sobre a eficiência do ajuste do método analisado em comparação a previsão mais simples possível, a ingênua ou trivial.

A fim de avaliar a qualidade de aderência dos modelos ajustados, foram efetuados os testes de qualidade do ajuste convencionalmente aplicados nas análises causais. Avaliou-se critérios tais como: ANOVA, coeficiente de determinação, coeficiente de determinação ajustado, teste de normalidade, teste de homocedásticidade,

Para a validação dos modelos ajustados efetuou-se para cada modelo análise de resíduos. Sendo efetivamente apresentada apenas a análise dos resíduos do modelo escolhido, uma vez, que foi à esse atribuído o peso de mais representativo, e é o modelo de fato utilizado para as previsões da demanda.

4 RESULTADOS E DISCUSSÕES

Nesta seção busca-se apresentar os principais resultados obtidos perante às análises feitas em torno da série histórica de dados. Também possui o propósito de explicar as aplicações dos conceitos definidos na revisão de literatura.

4.1 SÉRIE HISTÓRICA DE CAFÉ E ESTATÍSTICA DESCRITIVA

Os dados da série histórica de café no Brasil, coletados conforme descritos no item 3.2, estão dispostos na Tabela 2. Inicialmente confeccionou-se o diagrama da evolução da série ao longo do tempo, como visto na Figura 14, para observar, conforme descrito por Lustosa *et al*. (2008), se a demanda é estacionária ou apresenta possíveis tendências ou sazonalidades. Aplicou-se também demais testes estatísticos com a utilização dos *softwares*, conforme seção 3.3, bem como, toda a estatística descritiva de interesse para à análise.

Tabela 2 - Evolução da série histórica de consumo de café no Brasil em milhares de sacas de 60kg.

Ano	Consumo	Ano	Consumo	Ano	Consumo
1990	8.200	1999	12700	2008	17660
1991	8.500	2000	13200	2009	18390
1992	8.900	2001	13590	2010	19132
1993	9.100	2002	13750	2011	19720
1994	9.300	2003	14200	2012	20330
1995	10100	2004	14946	2013	20085
1996	11000	2005	15540	2014	20333
1997	11500	2006	16331	2015	20500
1998	12200	2007	17125	2016	20500

Fonte: Adaptado de OIC.

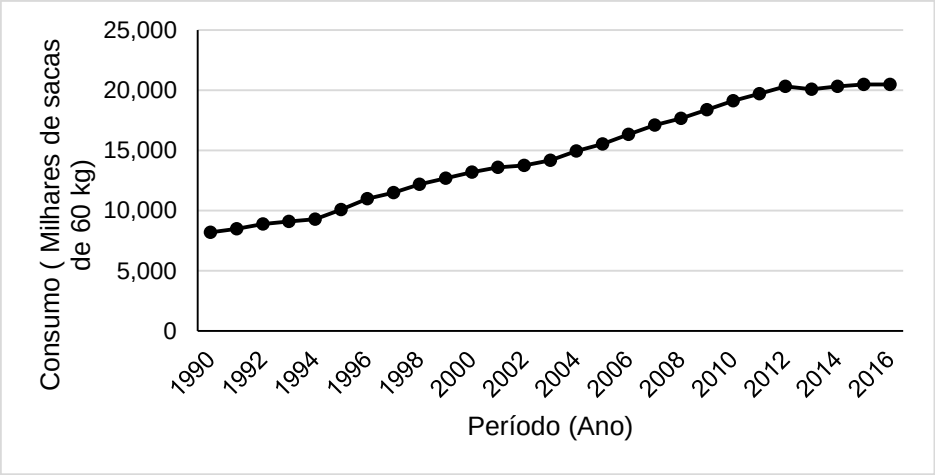


Figura 14 - Gráfico da evolução da série histórica do consumo de café no Brasil de 1990 a 2016.
Fonte: Autoria própria.

Diante da simplicidade ao se interpretar dados puramente numéricos, destaca-se a importância da análise gráfica, na qual, nota-se claramente uma tendência ascendente ao longo do tempo. Com base nessa visível tendência, pode-se presumir uma não estacionariedade, conforme Morettin e Toloi (1981), sugerem como primeira suposição para uma série temporal a de que essa seja, ou não, estacionária. Essa tendência de crescimento, ou a não estacionariedade podem ser percebidas também pela análise descritiva dos dados, conforme Tabela 3 .

Tabela 3 - Estatística descritiva dos dados da série de consumo em milhares de sacas de 60 kg.

Medida	Valor
1º Quartil	11.000
Média	14.697,48
Erro padrão	819,16
Mediana	14.200
Moda	20.500
3º Quartil	19.132
Desvio padrão	4.256,48
Variância da amostra	18.117.610,34
Coeficiente de variação	28,96%
Curtose	0,35
Assimetria	0,117
Amplitude	12.300
Mínimo	8.200
Máximo	20.500
Soma	396.832
Contagem	27
Nível de confiança (95,0%)	1.683,81

Fonte: Autoria própria.

Nota-se na análise da estatística descritiva, um consumo médio anual de café de 14.697 no período estudado, o que, entende-se como levemente superestimado, uma vez que os dados tendem à um aumento ao longo do tempo, esses valores podem causar uma certa defasagem em relação ao centro, conforme Reis e Reis (2002, p. 34), “[...] valores extremos “puxam” o valor da média em direção a si, deslocando também a representação do centro”.

Essa suposição é reforçada pelo seu coeficiente de variação de 28,96 %, que se caracteriza um pouco acima dos parâmetros de homogeniedade, além de sua assimetria positiva, em que, o valor da média é maior que o da mediana. Embora exista uma pequena distorção, causada pela tendência, suas características parecem atender os padrões de normalidade

Analisou-se também o histograma e *boxplot* da série, para auxiliar no entendimento dos dados através de análise gráfica. No histograma, visto na Figura 15, considerou-se a distribuição de classes proposta pelo *software* Minitab, que é a mesma para o Action Stat. Segundo Anderson *et al*. (2011, p. 32), “Uma das utilidades mais importantes de um histograma é fornecer informações sobre a forma, ou formato, de uma distribuição”. Essa forma está relacionada com a simetria do conjunto, assim como sua curtose.

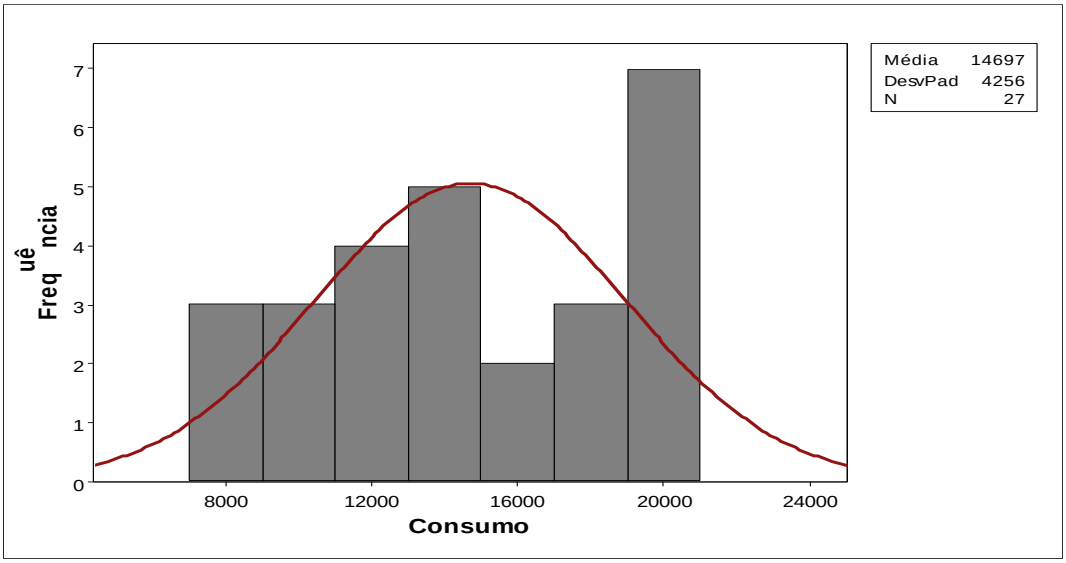


Figura 15 - Histograma de consumo com sobreposição da curva normal.
Fonte: Autoria própria.

Na série analisada, pode-se perceber a tendência de crescimento, e ainda, que a maior frequência se encontra junto aos maiores valores de consumo, nota-se,

incialmente essa apresentava uma maior homogeneidade. Com o auxílio da curva normal, percebe-se essa defasagem ascendente dos dados.

Assim como no histograma, o *boxplot* na Figura 16, permite uma interpretação visual mais prática sobre os dados, entendeu-se melhor sua assimetria, distribuição dos dados e a interpretação das medidas de tendência central. Com o auxílio do símbolo da média, fica claro a interpretação da assimetria positiva, pois a média é maior que a mediana, essa interpretação vai de encontro a lógica de crescimento, ou, ao menos, a representação de que a maioria dos dados coletados se encontram acima da mediana.

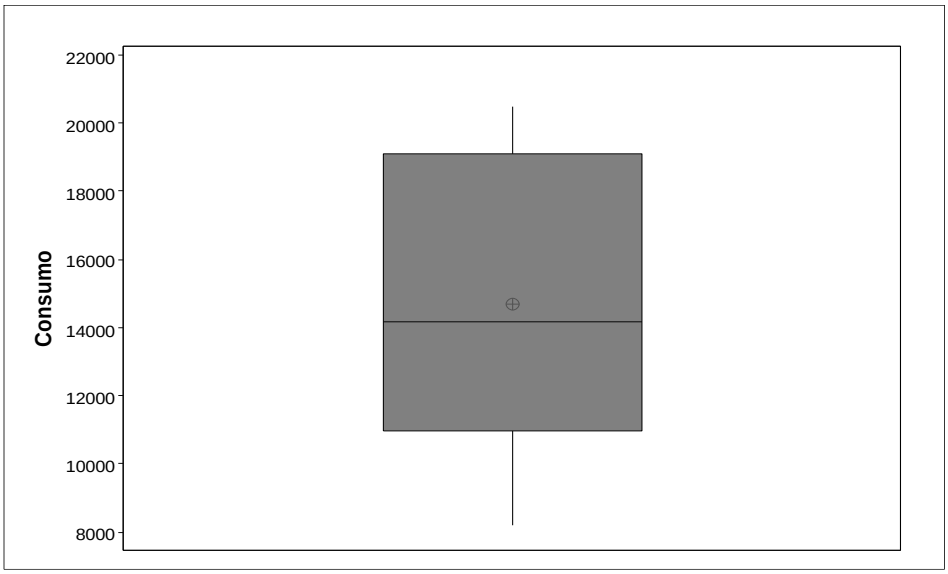


Figura 16 - Boxplot da série de consumo.
Fonte: Autoria própria.

Outra avaliação importante a ser feita nesse gráfico é a presença de pontos discrepantes (*Outliers*). Morettin e Bussab (2011, p. 49), conceituam estatisticamente que, “[...] um *outlier* pode ser produto de um erro de observação ou de arredeondamento”. Observou-se que não existem valores que possam ser considerados discrepantes ou *outlier*s, o que indica um comportamento normal dos dados.

Para garantia de que os dados seguem uma distribuição normal, executou-se um teste de normalidade. O teste padrão realizado pelo *software* Minitab é o de Anderson-Darling, com um intervalo de confiança de 95%, ou 5% de significância, conforme Figura 17. Como o teste e os parâmetros foram satisfatórios, optou-se por não realizar outros testes do mesmo *software,* tomando como averiguação as demais

análises da estatística descritiva.

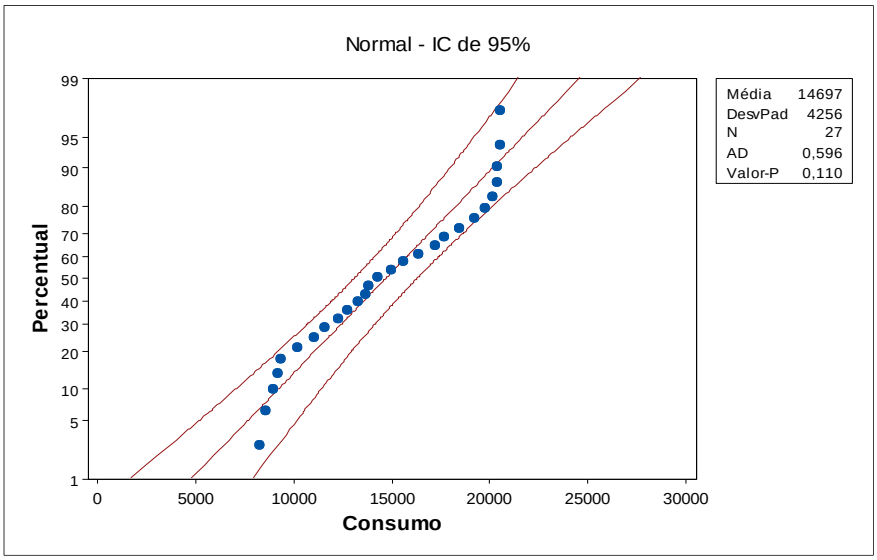


Figura 17 - Teste de normalidade para a série de consumo.
Fonte: Autoria própria.

Constatou-se que a hipótese, na qual, a série segue uma distribuição normal, ou seja, não se rejeita a hipótese nula, devido a interpretação do p-valor de 0,11 ou 11%, suficientemente maior que a significância de 5%. Para Groppo (2005, p. 25) “[...] a distribuição normal é verificada pela qualidade do ajuste das variáveis com a linha reta”. Na análise gráfica, nota-se que todos os pontos estão entre as linhas do intervalo de confiança.

4.2 TESTES DE ESTACIONARIEDADE, TENDÊNCIA E SAZONALIDADE

Observa-se que entender o comportamento natural dos dados e sua análise é fundamental para entender quais métodos melhor se aplicam, antes mesmo de comparar suas medidas de acurácia. Dessa forma, utilizam-se testes para avaliar a série quanto sua estacionariedade, tendência e sazonalidade antes da aplicação dos métodos de fato.

Devido ao fato de que a série não é estacionária, por exemplo, métodos como Média Móvel Simples ou Média Móvel Ponderada, não são tão aplicáveis quanto os métodos de Holt-Winters ou ARIMA. Segundo Moreira (2011, p. 312) o método de MMS “[...] pode ser um método eficiente quando a demanda é estacionária”, enquanto,

segundo Lustosa *et al*. (2008), os métodos de suavização, ou ainda, os de decomposição, são mais indicados quando uma série apresenta tendência ou sazonalidade.

Por meio da análise do gráfico da evolução histórica de consumo e das demais ferramentas da estatística descritiva, pode-se perceber que se trata de uma série não estacionária, com tendência de crescimento, e ainda, que não aparenta sazonalidade. Embora a produção do café obedeça a períodos sazonais da agricultura, e possua a característica de bienalidade, o que se analisa em questão é a série de consumo. No entanto, busca-se confirmação matemática, para tal, com a utilização do *software* Action Stat, realizou-se diversos testes relacionados à essas características.

4.2.1 Teste de Estacionariedade

Os principais testes utilizados para avaliar a estacionariedade de uma série analisam a presença de raízes unitárias classificadas através de um teste de hipóteses se se faz necessário a utilização de diferenciações, ou transformações para estacionar a série, características do método ARIMA. Cavalcanti e Mello (2014, p.44) fazem a ressalva de é preciso observar o fato de que a rejeição da hipótese de raiz unitária, não caracteriza por si só uma série estacionária, e sim que essa tem uma tendência determinística.

O *software* Action Stat, faz a análise considerando três testes a 95% de confiança: Dickey-Fuller Aumentado (ADF), Phillips-Perron (PP) e KPSS. Nos dois primeiros as hipóteses nulas são a de possuir raiz unitária e no terceiro se inverte. Os resultados são apresentados respectivamente nas Tabelas 4, 5 e 6.

Tabela 4 - Teste de Dickey-Fuller Aumentado para série de consumo.

Medida	Valor
P-Valor	0,76
Tamanho da Amostra	27
Hipótese Nula	Existe pelo menos uma raiz unitária
Hipótese Alternativa	Não existe raiz unitária

Fonte: Autoria própria.

Tabela 5 - Teste de Phillips-Perron para série de consumo.

Medida	Valor
P-Valor	0,69
Tamanho da Amostra	27
Hipótese Nula	Existe pelo menos uma raiz unitária
Hipótese Alternativa	Não existe raiz unitária

Fonte: Autoria própria.

Tabela 6 - Teste KPSS para série de consumo.

Medida	Valor
P-Valor	0,01
Tamanho da Amostra	27
Hipótese Nula	A série temporal é estacionária
Hipótese Alternativa	A série temporal apresenta raiz unitária

Fonte: Autoria própria.

Pode-se interpretar os testes através de cada p-valor em termos das hipóteses propostas. Nota-se que os dois primeiros testes apresentaram um p-valor acima da significância de 0,05 ou 5%, o p-valor foi de 0,76 (76%) no teste de ADF e 0,69 (69%) no teste de PP. O que caracteriza uma não rejeição da hipótese nula, ou seja, de possuir raiz unitária.

Para o teste KPSS a lógica se inverte, mas o resultado se mantém. O seu p-valor de 0,01 (1%), ficou abaixo da significância, dessa forma, rejeita-se a hipótese nula, a de que a série apresenta ao menos uma raiz unitária, portanto, não é estacionária. Sob o pressuposto de não estacionariedade da série, espera-se uma melhor adequação dos métodos de decomposição em relação aos demais.

4.2.2 Análise de Tendência

Para interpretar a existência de tendência ou qual curva de tendência melhor representa os dados, construiu-se através do Excel, diversos gráficos com ajustes de curvas na série de consumo. Para algumas tendências, como a linear, seria possível utilizar-se do eixo x (Período) entre 1 a 27, ao invés de 1990 a 2016, pois a variável de tendência não se altera, conforme Franco (2016, p. 13), se “[...]

esta não muda com a definição da variável tempo, podemos usar as duas formas", o que simplificaria as equações, e é útil também, para o método de Holt (DSE).

Para uma primeira interpretação utilizou-se como parâmetro o coeficiente de determinação (R^2) de cada equação ajustada das principais curvas de tendência. Seus coeficientes e as suas equações podem ser visualizados, na Tabela 7.

Tabela 7 - Equações e R^2 das curvas de tendência ajustadas.

Tendência	Equação	R^2
Linear	$y = 533{,}269841269841x - 1.053.442{,}01058201$	0,98896
Exponencial	$y = 8.10^{-30}e^{0{,}0382167059578275x}$	0,97582
Logarítmica	$y = 1.068.155{,}84515602ln(x) - 8.105.843{,}86638197$	0,98891
Quadrática	$y = -1{,}2222895878067x^2 + 5.429{,}76193002348x - 5.957.204{,}6852338$	0,98911
Potencial	$y = 2.10^{-249}x^{76{,}5649038270653}$	0,97627
Polinomial de sexto grau	$y = 0{,}0011742284466294x^6 - 14{,}1214884859805x^5 + 70.761{,}1244505169x^4 - 189.106.003{,}723554x^3 + 189.106.003{,}723554x^2 - 227.909.869.478.173x + 76.133.639.633.286.872$	0,99816

Fonte: Autoria própria.

Observa-se, pela interpretação dos coeficientes de determinação, que as principais curvas propostas apresentaram um ótimo desempenho na explicação da série. Antecipadamente já se pressupunha um comportamento de característica logarítmica, pois, observa-se que os dados apresentam crescimento ao longo do tempo e tendem a se estabilizar, na qual aproxima-se uma assíntota nos últimos anos.

A segunda opção seria uma curva linear, mesmo que está tenda a crescer infinitamente no tempo, para previsões de curto prazo se adequaria muito bem, além de ter uma notável vantagem pela simplicidade da sua equação característica.

A curva quadrática tende a decair a longo do tempo, o que não seria natural de se esperar da série, ao se aumentar o grau do polinômio, sua aderência melhora gradualmente. As pequenas oscilações foram melhor explicadas por um polinômio de sexto grau, pelo critério dos coeficientes de determinação.

Mesmo que a série aparentemente não apresente sazonalidade, o que é sugerido pelo comportamento oscilatório característico da curva, o modelo polinomial de sexto grau satisfaz as suposições. Contudo, para Morettin e Toloi (2001, p. 19) é preciso tomar cuidado ao fazer estimativas com polinômios, pois "[...] embora ele

possa se ajustar bem ao conjunto de valores observados, extrapolações futuras podem ser bastante ruins".

Na tendência exponencial, pressupõe-se um crescimento repentino, o que também não aparenta ser viável. Porém, como o maior interesse está em analisar previsões para o período imediatamente seguinte, ou seja, curto prazo. Da mesma forma a curva potencial, assim como a linear, tende a um crescimento anormal no final da série, o que não é condizente com o comportamento dos dados.

As distribuições das séries em comparação ao consumo medido, podem ser visualizadas na Figura 18, na qual, pode-se notar também, de forma discreta ou implícita, os padrões de comportamento de cada curva de tendência.

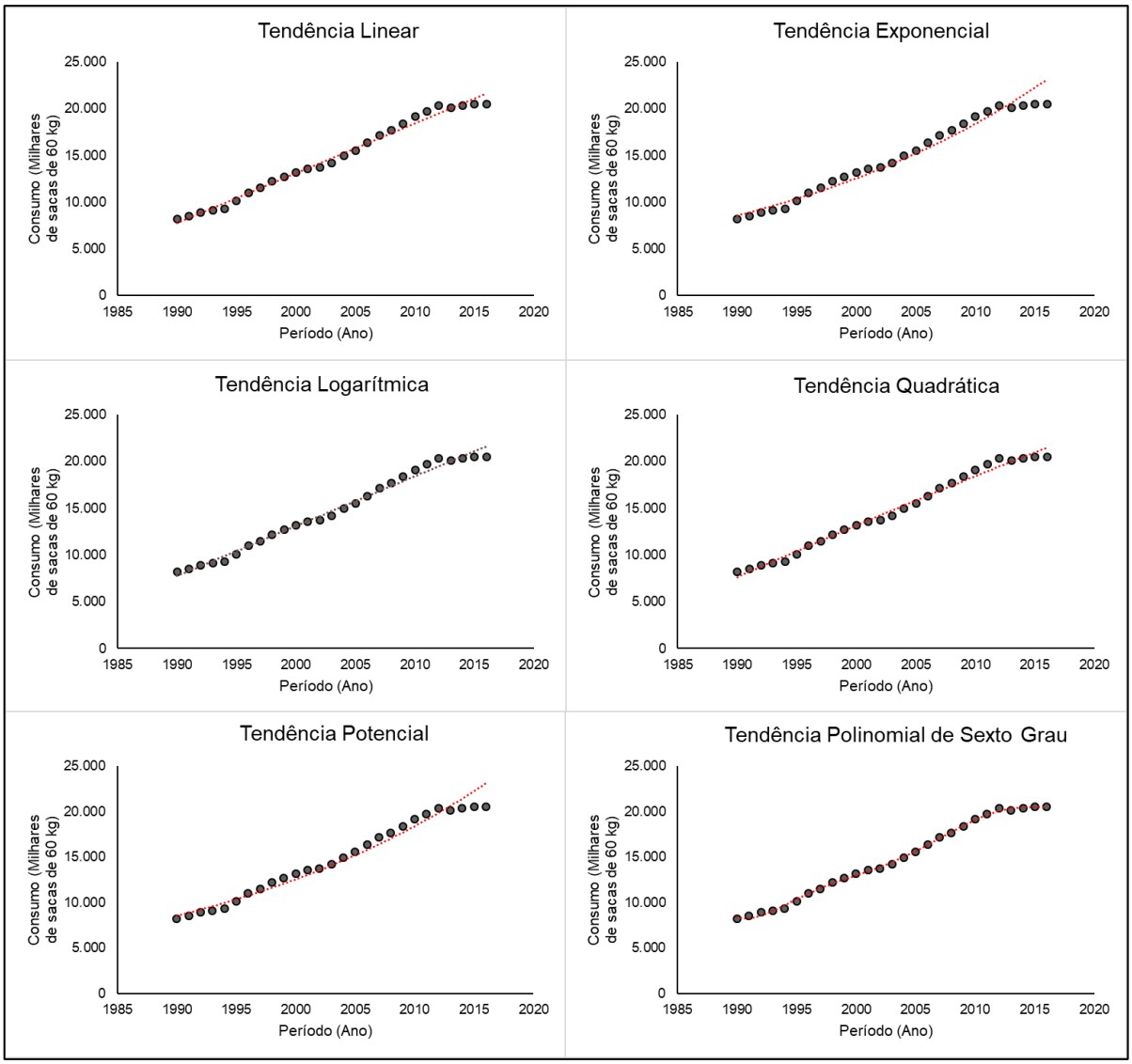


Figura 18 - Curvas de tendências ajustadas a série de consumo.
Fonte: Autoria própria.

Com as equações modeladas, buscou-se calcular as medidas de acurácia estabelecidas para comparação, Desvio Médio Absoluto (MAD) e U de Theil “[...] que compara o desempenho da regressão com a previsão ingênua (GIRARDI; ISHIKAWA, 2006, p. 10)”. As medidas estão dispostas na Tabela 8.

Tabela 8 - Medidas de acurácia para as análises de tendência.

Tendência	MAD	U de Theil
Linear	341,68	0,8053
Exponencial	635,90	1,4851
Logarítmica	341,57	0,8023
Quadrática	343,98	0,7850
Potencial	578,74	1,2140
Polinomial de sexto grau	146,22	0,3331

Fonte: Autoria própria.

Pode-se notar que as características esperadas são confirmadas, a curva polinomial de sexto grau, de fato, apresenta o menor dos erros. Apesar desse modelo atender as suposições da análise de regressão, entende-se que o seu comportamento característico não se adequa devidamente ao dos dados. Para averiguar e justificar, testou-se o comportamento da curva após cinco anos, conforme visto na Figura 19.

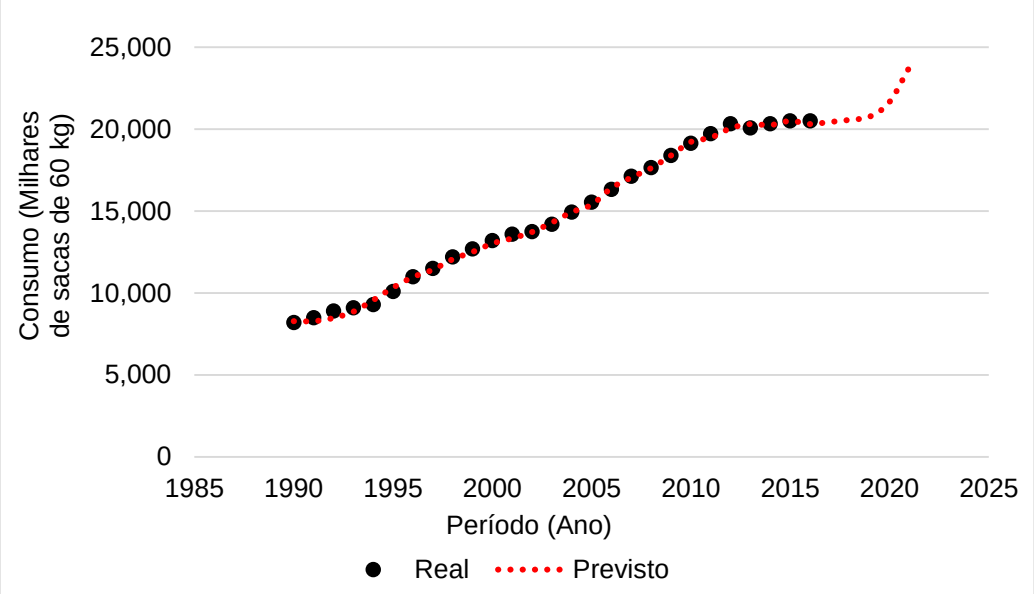


Figura 19 - Gráfico comparativo entre a série de consumo e a do modelo polinomial de sexto grau ajustado com extrapolações para 5 períodos.
Fonte: Autoria própria.

Dessa forma, a logarítmica, é, por padrão, a mais representativa entre as tendências. Os erros dessa curva justificam sua escolha, foram melhores em relação aos modelos linear e quadrático. Enquanto que as curvas exponencial e potencial, são descartadas por apresentarem um MAD elevado, considerada pela interpretação do coeficiente U de Theil, pior que a previsão ingênua, uma vez que essa apresentou coeficiente maior que 1.

Ciente de que essa análise de tendências releva previsões para curto prazo, ou mesmo, para o ano seguinte, todos os modelos analisados serviriam de forma satisfatória. Reforçado pelo fato de que, no período de um ano qualquer modelo seria realimentado e reajustado com o dado do ano atual, devido as características do ambiente econômico cafeeiro de safra anual.

A partir da melhor compreensão da série estuda em relação a sua estacionariedade, tendência e sazonalidade, busca-se formular os demais modelos propostos, com o interesse em constatar suas adequações e identificar os mais precisos para previsão.

4.2.3 Teste de Sazonalidade

Dentre as análises de interesse, testou-se os dados para identificar possíveis oscilações sazonais ou cíclicas. Pois essa interpretação é determinante na utilização de alguns métodos como o de Holt-Winters.

Com a análise do gráfico da série, pode-se notar a ausência de períodos sazonais, na qual apresenta um crescimento relativamente uniforme ao longo do intervalo de tempo estudado. Contudo, buscou-se verificar essas hipóteses com testes específicos em diversos períodos sazonais. Os testes realizados pelo *software* Action Stat, são: teste de Kruskal-Wallis e o Teste de Friedman, conforme, Tabela 9.

Tabela 9 - Testes de sazonalidade para a série de consumo.

	Teste de Kruskal-Wallis		**Teste de Friedman**	
Comprimento Sazonal	12	2	12	2
P-Valor	0,0703	0,2702	0,0571	0,7815
Tamanho da Amostra	27	27	27	27
Hipótese Nula	Não há sazonalidade determinística			
Hipótese Alternativa	Há sazonalidade determinística			

Fonte: Autoria própria.

Pode-se constatar pela interpretação do p-valor nos dois testes referidos, considerados todos os períodos sazonais analisados, que não há sazonalidade na série de consumo, dessa forma, descarta-se os métodos aplicáveis que consideram essa ponderação nos procedimentos.

4.3 MODELO DE MÉDIA MÓVEL SIMPLES

A análise da série por meio de médias móveis tem relação com a quantidade de anos observados. Com essa concepção, calculou-se as previsões e as medidas de erro para n períodos, a partir de n=2, por ser o menor possível, no qual, segundo Tubino (2009, p. 20) a média móvel tem sua maior sensibilidade. Os resultados de acurácia e a comparação das previsões da MMS com n=2 e a série original, estão respectivamente na Tabela 10 e na Figura 20.

Tabela 10 - Medidas de acurácia para os modelos de MMS ajustados.

Período	MAD	U de Theil
2	726,00	1,4486
3	985,18	1,8870
9	2651,28	4,1194
13	3767,01	5,1048

Fonte: Autoria própria.

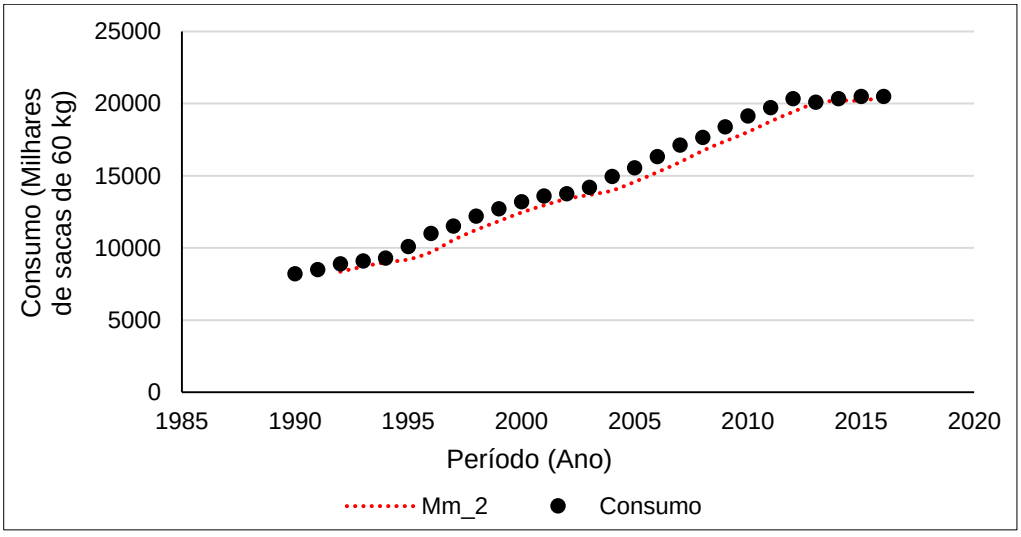


Figura 20 - Gráfico comparativo entre a série de MMS ajustada com n = 2 e a série de consumo.
Fonte: Autoria própria.

Constatou-se, porém, que desde seu menor n, esse método não apresentou um resultado satisfatório. Corroborando para a análise prévia feita referente à estacionariedade, na qual, o método MMS não se aplica de maneira adequada a séries não estacionárias.

Mesmo na MMS com n=2 que foi a mais próxima dos dados reais nesse método, é possível ver claramente a defasagem sugerida por Moreira (2011, p. 312), causada pela tendência crescente dos dados, o que indica que tais modelos de médias móveis simples, não se fazem muito uteis para previsões da série temporal estudada.

4.4 MODELO DE MÉDIA MÓVEL PONDERADA

No mesmo princípio de MMS com n=2, buscou-se, conforme sugerido por Peinado e Graeml (2011, p. 20) “[...] utilizar programação linear (ferramenta Solver do Excel) para determinar quais os valores de ponderação produzem o menor desvio padrão para a previsão”.

Porém, no presente caso o interesse está na minimização do erro, o MAD, ou seja, minimizar a função da média do módulo dos erros. Considerou-se também que o modelo a ser minimizado não atende os padrões de linearidade exigidos pela ferramenta, utilizou-se as opções não lineares. A opção indicada para modelos não suaves apresentou melhor desempenho.

Testou-se, assim como na MMS, vários intervalos para ponderação. Em todos os períodos a solução ótima consiste em atribuir todo o peso, para o último dado, ou seja, transformando-a em uma previsão ingênua. Dessa forma, esse método não conseguiu se mostrar eficiente para a série em questão, em que foi equivalente a previsão trivial, ou, teve U de Theil igual a 1 e MAD = 518,57.

4.5 MODELO DE SUAVIZAÇÃO EXPONENCIAL SIMPLES

Para o método SSE, também se fez uso da ferramenta Solver para a otimização da constante de suavização (α). A ferramenta sugere o parâmetro máximo, $\alpha = 1$, "[...] transformando o modelo de suavização exponencial em um modelo de média móvel simples" (PEINADO; GRAEML, 2011, p. 26).

Esse modelo ajustado, assim como no de médias ponderadas, igualou-se ao coeficiente 1 de U de Theil, o que fez com que se equivalesse à previsão ingênua. Seu erro médio absoluto foi levemente menor, MAD = 499,6.

4.6 MODELO DE HOLT (DUPLA SUAVIZAÇÃO EXPONENCIAL (DSE))

Percebe-se que, a priori, o modelo de Holt apresentaria um bom resultado quando aplicado à série, uma vez que, a equação de tendência linear apresentou um bom ajuste. Devido ao comportamento da série, esse método leva uma natural vantagem em relação aos anteriormente vistos e até mesmo o de Holt-Winters, aplicado a séries sazonais.

Comparou-se os resultados nos diferentes *softwares*, além da técnica minimização de pesquisa operacional do Solver, para obter o erro minimizado. O ajuste em relação aos dados reais pode ser visto na Figura 21. Os coeficientes de nível e tendência que melhor se ajustaram e suas medidas de acurácia estão dispostos na Tabela 11.

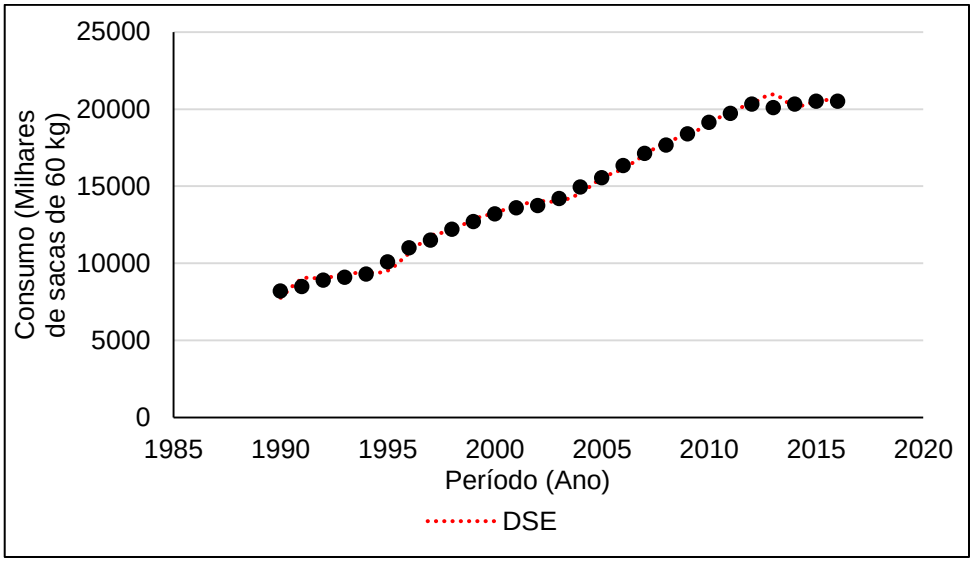


Figura 21 - Gráfico comparativo entre a série do modelo DES ajustado e a série de consumo.
Fonte: Autoria própria.

Tabela 11 - Medidas de acurácia para o modelo de DSE ajustado.

α	Υ	MAD	U de Theil
0,970709	0,6314904	199,69	0,5136

Fonte: Autoria própria.

É possível perceber que o modelo apresentou um bom desempenho, em critério do MAD e que parece ser adequado em relação ao comportamento dos dados. Dessa forma, é viável considerar sua utilização para previsões.

4.7 MODELO DE HOLT-WINTERS

Por meio das análises nos testes de sazonalidade, entende-se que o método do Holt-Winters não é viável para os dados em questão, uma vez, que esses não apresentam nenhum padrão sazonal. Dessa forma o método aditivo para padrões sazonais constantes e o multiplicativo para padrões crescentes, tem sua constante de sazonalidade (δ) desconsiderada, o que a torna equivalente à uma suavização exponencial dupla. Na qual, com os mesmos coeficientes otimizados, encontra-se o mesmo valor de erro.

4.8 MODELO ARIMA

O método ARIMA baseia-se em estacionar séries com tendência, como a do presente estudo. Com base nesse pressuposto, utilizou-se de diferenciações em busca de estacioná-la, e testou-se a série diferenciada através dos critérios já utilizados na seção 4.2.1.

Após a primeira diferenciação, obteve-se a série distribuída na Figura 22. Sua estacionariedade foi então avaliada, os resultados podem ser visualizados nas Tabelas 12, 13 e 14.

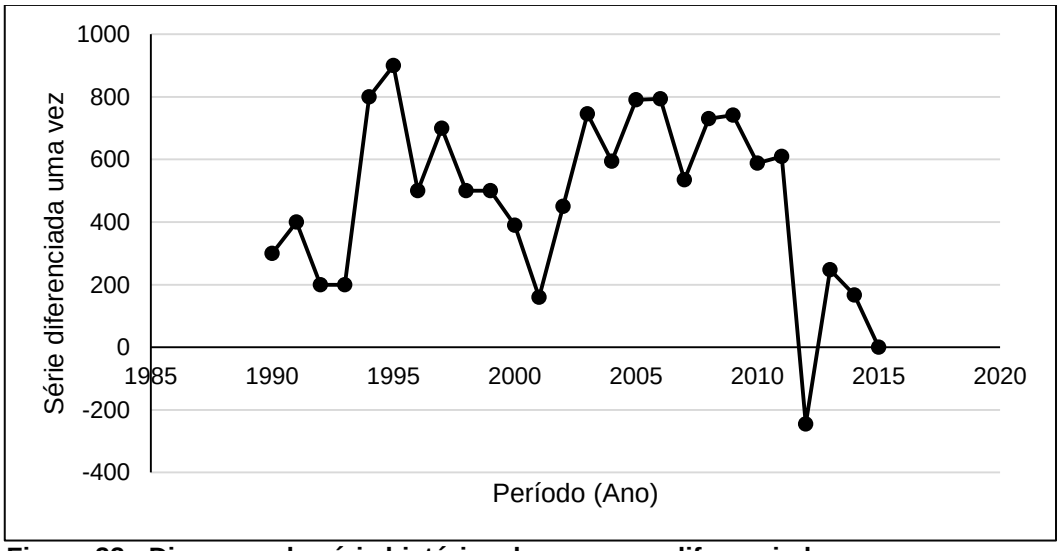


Figura 22 - Diagrama da série histórica de consumo diferenciada uma vez.
Fonte: Autoria própria.

Tabela 12 - Teste de Dickey-Fuller Aumentado para série diferenciada uma vez.

Medida	Valor
P-Valor	0,56
Tamanho da Amostra	26
Hipótese Nula	Existe pelo menos uma raiz unitária
Hipótese Alternativa	Não existe raiz unitária

Fonte: Autoria própria.

Tabela 13 - Teste de Phillips-Perron para série de consumo diferenciada uma vez.

Medida	Valor
P-Valor	0,29
Tamanho da Amostra	26
Hipótese Nula	Existe pelo menos uma raiz unitária
Hipótese Alternativa	Não existe raiz unitária

Fonte: Autoria própria.

Tabela 14 - Teste KPSS para série de consumo diferenciada uma vez.

Medida	Valor
P-Valor	0,1
Tamanho da Amostra	26
Hipótese Nula	A série temporal é estacionária
Hipótese Alternativa	A série temporal apresenta raiz unitária

Fonte: Autoria própria.

Constata-se que apenas no teste KPSS a série pode ser considerada estacionária, dessa forma, decidiu-se elaborar uma segunda diferenciação para avaliar novamente, com o interesse de garantir a estacionariedade da série, conforme Figura 23, confirmada pelos testes das Tabelas 15, 16 e 17.

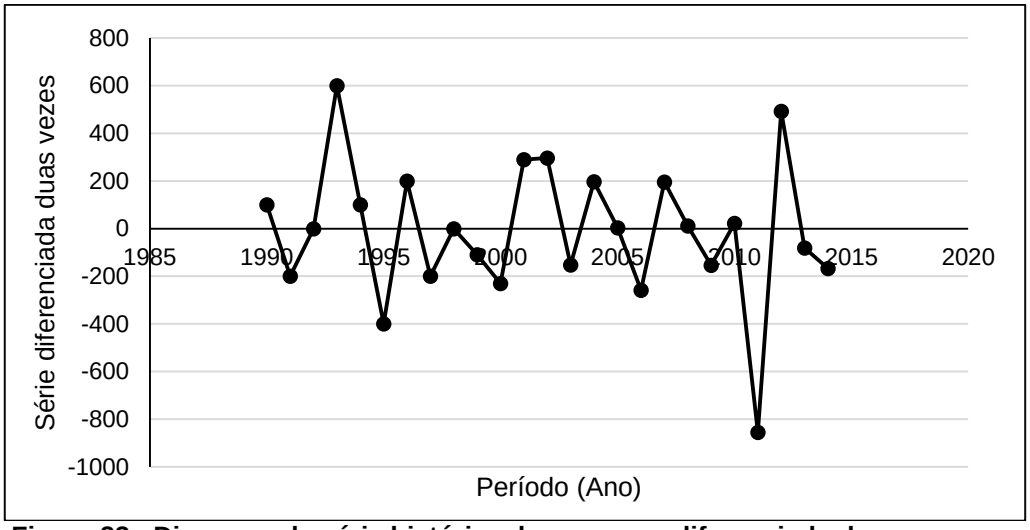


Figura 23 - Diagrama da série histórica de consumo diferenciada duas vezes.
Fonte: Autoria própria.

Tabela 15 - Teste de Dickey-Fuller Aumentado para série de consumo diferenciada duas vezes.

Medida	Valor
P-Valor	0,01
Tamanho da Amostra	25
Hipótese Nula	Existe pelo menos uma raiz unitária
Hipótese Alternativa	Não existe raiz unitária

Fonte: Autoria própria.

Tabela 16 - Teste de Phillips-Perron para série de consumo diferenciada duas vezes.

Medida	Valor
P-Valor	0,01
Tamanho da Amostra	25
Hipótese Nula	Existe pelo menos uma raiz unitária
Hipótese Alternativa	Não existe raiz unitária

Fonte: Autoria própria.

Tabela 17 - Teste KPSS para série de consumo diferenciada duas vezes.

Medida	Valor
P-Valor	0,1
Tamanho da Amostra	25
Hipótese Nula	A série temporal é estacionária
Hipótese Alternativa	A série temporal apresenta raiz unitária

Fonte: Autoria própria.

Assegurou-se a estacionariedade da série a partir da segunda diferenciação, pode-se presumir como parâmetro $d = 2$, o que corresponde a principal característica do modelo ARIMA. Resta agora, estimar os demais parâmetros de um modelo ARMA (estacionário) através da análise da função de autocorrelação (FAC) e função de autocorrelação parcial (FACP) da série original. Conforme Kirchner (2006, p. 28), "[...] o analista de séries temporais identifica a estrutura comparando o correlograma da função de autocorrelação e o da função de autocorrelação parcial com os correlogramas teóricos das diversas estruturas ARMA (p, q)".

Em contexto a esses critérios, confeccionou-se os correlogramas de interesse a partir da série de consumo, conforme as Figuras 24 e 25. Utiliza-se a FACP para encontrar o termo p, ou o modelo AR equivalente, e, a FAC para estimar o modelo MA que compõe o coeficiente q do modelo ARIMA.

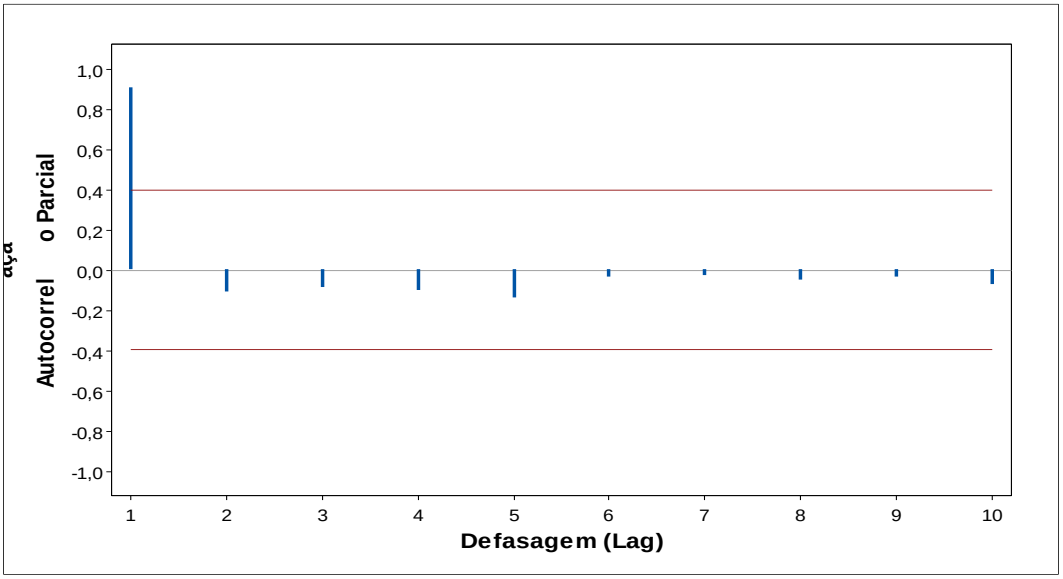


Figura 24 - Correlograma da função de autocorrelação parcial para a série de consumo.
Fonte: Autoria própria.

Pode-se notar um comportamento característico de decaimento exponencial, comportamento característico ao modelo AR (1), observa-se também, que, “[...] o correlograma da FACP, apresenta a primeira defasagem (lag) diferentes de zero significativamente. Assim, há uma indicação de que a ordem do modelo autoregressivo é p=1” (SANTOS, 2014).

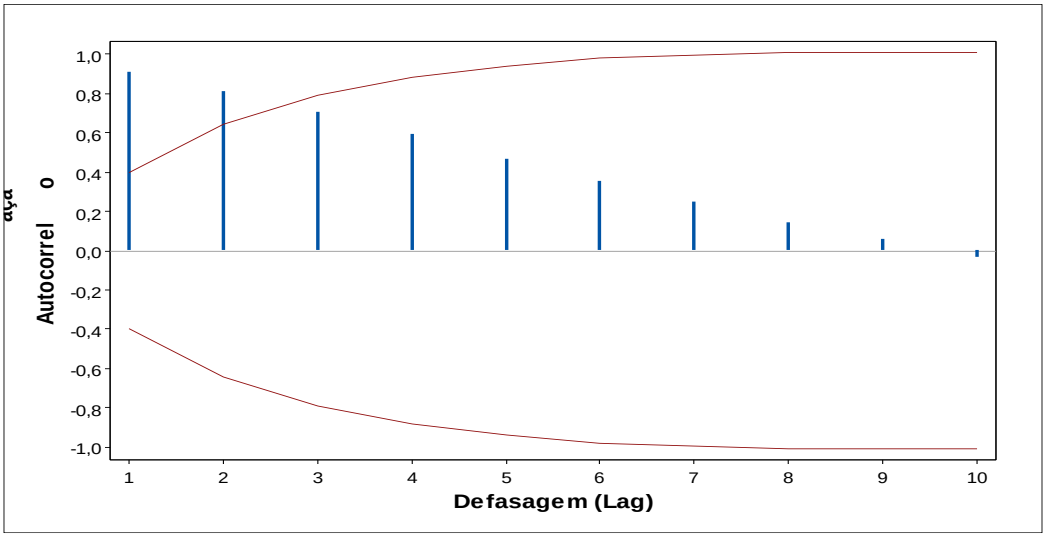


Figura 25 - Correlograma da função de autocorrelação para a série de consumo.
Fonte: Autoria própria.

Da mesma forma, fez-se uso dos critérios no correlograma da FACP, nota-se, além do decaimento característico do modelo, conforme Franco (2011). Pode-se perceber ainda, que existem duas defasagens acima do limite de aceitação, portanto, a série começa a ser estatisticamente igual a zero, depois do terceiro *lag*. Com foco na simplicidade do modelo e na minimização do MAD, tem-se como uma boa escolha para o grau $q = 2$. Dessa forma, com os coeficientes estimados, subentende-se que os filtros, componham um modelo ARIMA (1,2,2), AR=1, I=2 e MA= 2.

No entanto, devido a relação interpretativa do método, e a praticidade de utilização dos *softwares*, na medida da coerência estatística, buscou-se outros modelos que minimizassem o erro. Foram avaliados modelos com uma diferenciação e outras ponderações dos graus de $p\ e\ q$, para que se possa analisar, sua estacionariedade com uma diferenciação e se o modelo inicial esta subespecificado ou superespecificado. As medidas de precisão do modelo (1,2,2) indicado e dos demais ajustados podem ser vistas na Tabela 18.

Tabela 18 - Medidas de acurácia para os modelos ARIMA ajustados.

Modelo	MAD	U de Theil
1,2,2	187,26	0,4675
1,1,1	186,77	0,4513
1,1,5	147,21	0,3576
1,1,2	204,85	0,4513
1,2,3	194,73	0,4644
2,2,2	180,58	0,4609
2,2,1	192,73	0,4222
2,2,4	174,66	0,3876
1,2,5	182,79	0,4154
1,2,1	192,73	0,4222

Fonte: Autoria própria.

Destaca-se alguns modelos, principalmente o ARIMA (1,1,5) ($p = 1,\ d = 1\ e\ q = 5$), ao qual apresentou um erro bastante abaixo dos demais, percebe-se que o aumento do termo MA do modelo, tende a aumentar sua aderência, o que indica que o modelo (1,2,2) estava subdimensionado em termos de MA e o grau d, com duas diferenciações, estava superdimensionado. Pressupondo essa interpretação e por atender as suposições de ajuste, é viável considerar o modelo ARIMA (1,1,5) para previsões.

Mesmo que esse modelo tenha sido encontrado através de oscilações estratégicas dos graus $p, d\ q$ do modelo original, considera-se que, para esse estudo o interesse é avaliar os métodos em si, de maneira técnica e interpretativa, para analisar a aderência pratica para o posterior uso por gestores, economistas, estatísticos, entre outros. Substitui-se o modelo ARIMA (1,2,2) ajustado pela metodologia inicial, pelo modelo ARIMA (1,1,5) que com melhor representação.

Para um estudo similar, com uma série histórica de consumo de café entre 1975 e 2004, Bacci (2007), ajustou um modelo ARIMA (2,2,5) para comparação entre os demais modelos. O MAD obtido por esse modelo, em termos das devidas proporções e diferenças das suas séries, foi de 265, 387 milhares de sacas de 60 kg, ao se considerar que ambos os modelos atendem as especificações estatísticas necessárias, é possível avaliar que há coerência entre as análises.

4.9 MODELO DE REGRESSÃO MULTIPLA

Com o interesse em avaliar as influências de fatores causais, essas variáveis são: produção, exportação, população, preço no varejo (preço_varejo), preço para sacas de 60 kg em U$ cents/lb (preço_sacas) e a variável tempo expressa em período de anos. Efetuou-se uma análise de regressão múltipla que envolvesse os referidos fatores.

Para possibilitar a regressão, nivelou-se os dados de todas as variáveis. Para o modelo com o preço no varejo, considerou-se valores de 1996 a 2016, e, para o que leva em consideração o preço das sacas, utilizou-se dados de 1991 a 2015, no total de 25 anos. Essa diferença entre a quantidade dos dados, faz com que sejam feitas duas analises em paralelo, em que se avaliam as duas variáveis separadamente. Por fim, considerou-se ambas as variáveis ao modelo, niveladas entre 1996 a 2015, período de 20 anos.

Essa análise em duas cotações de preço, se deve à tentativa em acoplar ao modelo, o preço que é repassado para o consumidor final (preço_varejo), em uma média nacional, embora, pela escala das demais componentes, percebeu-se a necessidade em adequar um modelo com uma variável de preço equivalente em escala (preço_sacas). O estudo das variáveis em conjunto, buscou encontrar as correlações e as significâncias entre as variáveis que possam melhor explicar o comportamento dos dados de consumo.

Inicialmente avaliou-se as correlações existentes entre os fatores, para entender a importância de cada fator para a equação de regressão proposta, bem como, selecionar quais seriam mais relevantes para compor à análise. A Tabela 19, demonstra os coeficientes de correlação de Pearson (r) para cada variável.

Tabela 19 - Correlação entre os fatores causais aplicados na análise de regressão múltipla.

	Consumo	Produção	Exportação	População	Preço_sacas	Preço_varejo
Produção	0,801					
Exportação	0,616	0,965				
População	0,990	0,773	0,582			
Preço_sacas	0,509	0,404	0,309	0,417		
Preço_varejo	0,953	0,745	0,562	0,925	0,574	
Período	0,948	0,782	0,593	0,996	0,443	0,948

Fonte: Autoria própria.

Têm-se, a quesito de maior relevância, as correlações das variáveis com o consumo e o tempo (Período), exceto a variável preço_sacas, que apresenta uma correlação baixa ou moderada com as demais, segundo os critérios dados por Bisquerra *et al.* (2004), o que se caracteriza como uma possível eliminação dessa variável na equação de regressão.

Posteriormente, ajustou-se diversos modelos os parâmetros envolvidos, com o intuito de encontrar uma equação que melhor representa a demanda de café em termos desses diversos fatores. Na

Tabela 20, estão os modelos testados e suas equações com seus respectivos coeficientes de determinação ajustados (R_a^2).

Tabela 20 - Relação de equações de regressão múltipla e R_a^2

Modelo	Equação	R_a^2
M1	Consumo = 0 + 1,000 Produção - 1,000 Exportação - 0 População – 0 Preço_sacas - 0 Preço_varejo - 0 Período	1,0000
M2	Consumo = 375.100 + 0,0276 Produção + 0,000286 População + 182 Preço_varejo - 206 Período	0,9843
M3	Consumo = 358.933 + 0,0241 Exportação + 0,000290 População + 185 Preço_varejo - 199 Período	0,9834
M4	Consumo = - 563.646 + 0,0265 Produção + 0,000100 População + 5,20 Preço_sacas + 279 Período	0,9942
M5	Consumo = - 590.784 + 0,0249 Exportação + 0,000101 População + 5,30 Preço_sacas + 293 Período	0,9938
M6	Consumo = 58.114 + 0,02160 Produção + 0,000218 População + 5,55 Preço_sacas + 149,5 Preço_varejo - 42 Período	0,9944
M7	Consumo = - 23.474 + 0,02112 Produção + 0,000202 População + 5,63 Preço_sacas + 134,6 Preço_varejo	0,9959
M8	Consumo = 50.713 + 0,02079 Exportação + 0,000221 População + 5,69 Preço_sacas + 151,4 Preço_varejo - 39 Período	0,9942
M9	Consumo = - 24.016 + 0,02034 Exportação + 0,000207 População + 5,76 Preço_sacas + 137,8 Preço_varejo	0,9945
M10	Consumo = - 52.5694 + 0,0279 Produção + 0,000109 População + 260 Período	0,9903
M11	Consumo = - 56.0819 + 0,0240 Exportação + 0,000109 População + 277 Período	0,9898

Fonte: Autoria própria.

Inicialmente montou-se um modelo com todos os fatores, denominado M1. Constatou-se que as variáveis exportação e produção apresentaram coeficientes iguais a 1, e os demais fatores, coeficientes próximo ou igual a zero. Esse fenômeno

pode ser explicado de maneira simples, uma vez que, o consumo tem relação direta com esses valores. A OIC (2017), considera que, o que é produzido menos o que for consumido é o que é exportado, portanto, explicam todo o fenômeno.

Para uma análise mais clara da influência de cada fator, então, se faz necessário a retirada de um dos dois componentes, produção ou exportação, da fórmula, em função dos modelos com a variável preço_varejo, expressos em M2 e M3 e preço_sacas, em M4 e M5.

Nota-se dois comportamentos não condizentes nessas equações, o sinal positivo dos preços e o negativo do período. Esses sinais não seguem a lógica, pois com o aumento de preço, espera-se uma queda no consumo, ao contrário do que descreve a equação, na mesma ideia, pelas análises feitas anteriormente, consegue-se ver um crescimento da demanda ao passar do tempo e não o decréscimo caracterizado pelo coeficiente negativo do termo.

Devido a essa distorção, buscou-se então ajustar os dois fatores de preço e a variável período na equação, nos modelos M6 a M9. Após alternar-se as variáveis percebeu-se que os parâmetros de comparação faziam mais sentido, sem a consideração dos preços, as suposições continuam atendidas, e seus R_a^2, não apresentaram uma piora significativa, conforme M10 e M11, por isso, foi considerada a mais representativa entre elas, aquela que é constituída com o fator produção ao invés de exportação.

Percebeu-se que os modelos melhores ajustados são os com a variável produção, considerou-se apenas os modelos que a contem, para a análise de erros, conforme Tabela 21.

Tabela 21 - Medidas de acurácia para Regressão Múltipla.

Modelo	MAD	U de Theil
M2	301,99	0,6780
M4	222,54	0,5013
M6	164,28	0,3710
M7	165,31	0,3719
M10	298,78	0,7191

Fonte: Autoria própria.

Apesar dos erros do modelo M10 não serem os melhores encontrados nesse método (MAD = 298,78 e Coeficiente de U de Theil = 0,7191), ela é a mais relevante para representação da série, referente ao método de relação causal, devido

a sua coerência com o comportamento lógico dos dados em comparação aos demais modelos de regressão múltipla ajustados. A comparação entre real e previsto deste modelo, pode ser visto na Figura 26.

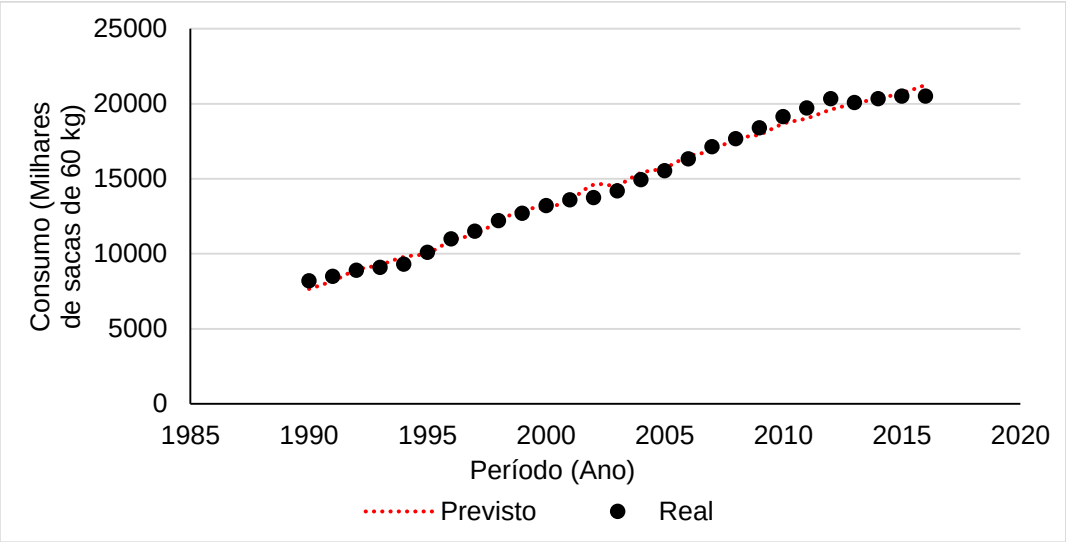


Figura 26 - Gráfico comparativo entre a série do modelo de regressão múltipla ajustado e a série de consumo.
Fonte: Autoria própria.

Economicamente essa equação representa, em termos práticos, que para a demanda de café, ou seja, o consumidor brasileiro, o mais importante é a existência do produto café, e não o preço em si. Essa ocorrência pode ser justificada pelo fato do café ser um produto considerado culturalmente popular no país ou item básico de consumo.

Nota-se ainda, uma forte relação com a produção e principalmente, a variável população, essa ideia pode ser reforçada com base na presença dos critérios de produção, população e período na equação final de regressão, o que indica que o aumento do consumo de café está diretamente ligado com o aumento da produção e da população ao longo do tempo.

4.10 COMPARAÇÔES ENTRE OS MODELOS APLICÁVEIS

Relacionou-se todos os métodos abordados em uma comparação direta, com o interesse de definir, através não só do critério de menor erro, como também

de interpretações relevantes anteriormente feitas à cada modelo. Dessa forma, os melhores modelos de cada método aplicável e suas medidas de acurácia, estão relacionados, do mais indicado para o menos indicado, conforme Tabela 22.

Tabela 22 - Comparação entre os métodos aplicáveis.

Modelo	MAD	U de Theil
ARIMA (1,1,5)	147,21	0,3576
DSE (Holt)	199,69	0,5136
Regressão Múltipla	298,78	0,7191
Tendência Logarítmica	341,57	0,8023

Fonte: Autoria própria.

Com a ausência dos modelos polinomial de sexto grau e alguns de regressão, por critérios técnicos e matemáticos, fica evidente o desempenho avantajado do modelo ARIMA (1,1,5), o qual, foi considerado o mais apropriado para realizar previsões para a demanda. Embora seu ajuste tenha partido de um critério relativamente empírico, nota-se que, mesmo o modelo originalmente ajustado ARIMA (1,2,2) com MAD = 187,26 e U de Theil = 0,4675, em uma comparação com os demais, seria o mais indicado para previsão.

4.11 VALIDAÇÃO DO MODELO E PREVISÔES

Após identificar o modelo mais apropriado por meio de comparação, em relação ao MAD, é preciso validar o modelo para sua efetiva aplicação. Para isso, foi feita a análise dos resíduos do modelo, seguindo os pressupostos descritos no item 2.6.6. Nos quais, os resíduos devam ser independentes ou livres de autocorrelação, seguir uma distribuição normal e serem homocedásticos.

4.11.1Independência e autocorrelação dos resíduos

Para atestar-se de que os resíduos não têm correlação entre si, Navidi

(2012), sugere que se faça um gráfico de dispersão entre os resíduos e as variáveis de ajuste, ou seja, as previsões para os períodos existentes feitas com o modelo escolhido. Conforme Figura 27.

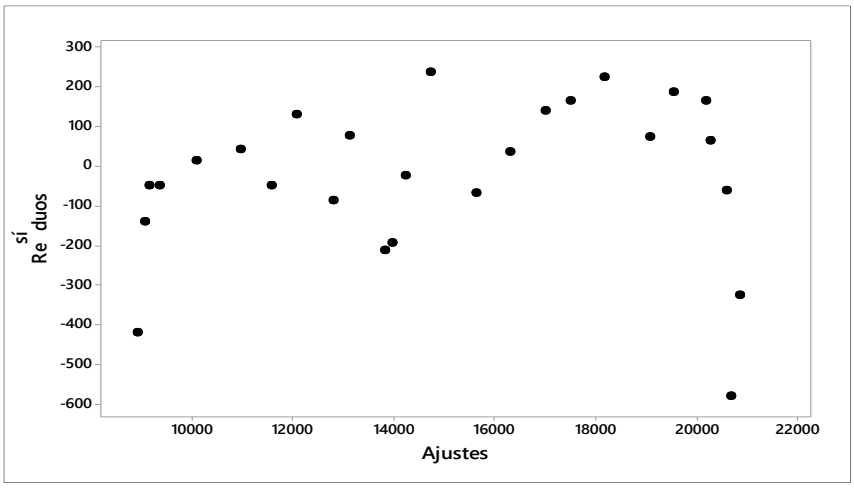


Figura 27 - Diagrama de dispersão entre os ajustes e os resíduos do modelo escolhido.
Fonte: Autoria própria.

Nota-se, uma dispersão não tão bem caracterizada, o que pode acentuar uma leve tendência. Para averiguar de maneira mais profunda, com o intuito de garantir que de fato os resíduos não estejam autocorrelanados, utilizou-se as análises da função de autocorrelação (FAC) e função de autocorrelação parcial (FACP) dos resíduos. Os correlogramas das FAC e FACP com 95% de confiança, estão respectivamente representados nas Figuras 28 e 29.

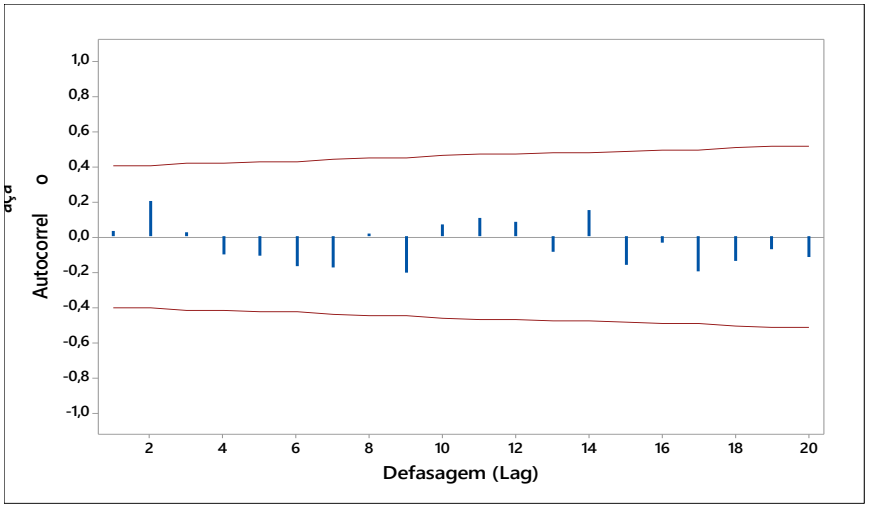


Figura 28 - Correlograma da função de autocorrelação para para os resíduos do modelo escolhido.
Fonte: Autoria própria.

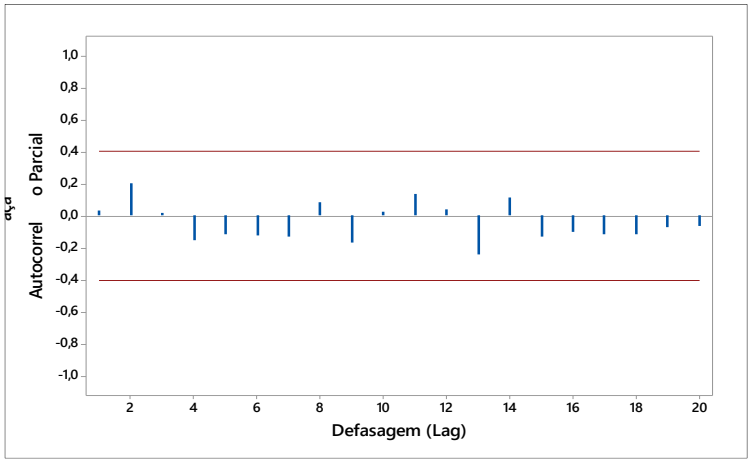


Figura 29 - Correlograma da função de autocorrelação parcial para os resíduos do modelo escolhido.
Fonte: Autoria própria.

A partir das análises das FAC e FACP, tem-se que os resíduos não apresentam nenhuma autocorrelação significativa a 95% de confiança. Uma vez que mesmo no correlograma com 20 defasagens, todas as autocorrelações ficaram dentro dos limites de significância.

4.11.2 Teste de Normalidade dos resíduos

Para atender ao pressuposto de que os resíduos sigam uma distribuição normal, realizou-se o teste de normalidade padrão *software* Minitab, de forma análoga ao realizado para os dados de consumo no item 4.1.

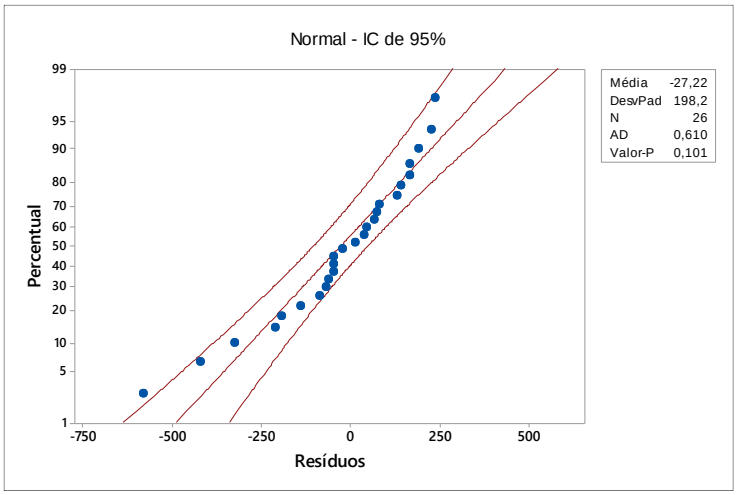


Figura 30 - Teste de normalidade para os resíduos do modelo escolhido.
Fonte: Autoria própria.

Como o p-valor é maior que 0,05, têm se, com 95% de confiança que os resíduos são normais, ou seja, atendem a uma distribuição normal de dados.

4.11.3 Homocedasticidade dos resíduos

Para caracterizar a homocedasticidade dos resíduos, executou-se a ferramenta *Run Chart* do *software* Minitab, que gera um ensaio com os dados por meio de um gráfico de controle. Dessa forma, é possível avaliar a estabilidade dos dados entorno de uma média constante. O gráfico pode ser visto na Figura 31.

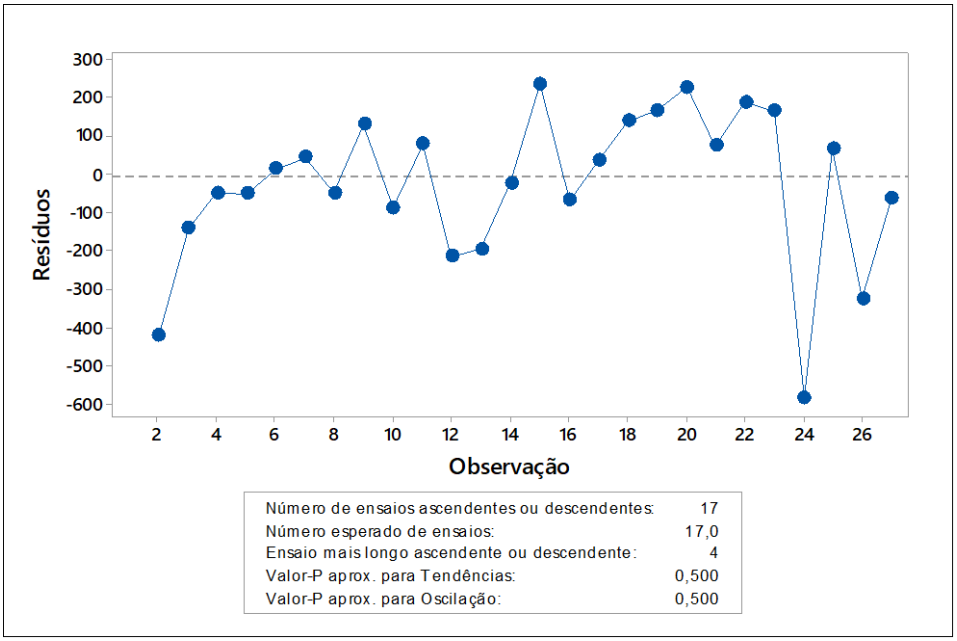


Figura 31 - Run Chart dos resíduos do modelo escolhido.
Fonte: Autoria própria.

Pode-se notar uma distribuição aleatória dos dados, apesar de possuir uma oscilação acentuada na observação 24, tem-se, pela análise dos parâmetros de tendência e oscilação, no qual ambos obtiveram p-valor maior que 0,05, que o processo está estável e que sua variância é constante a 95% de confiança. Portanto os resíduos são homocedásticos.

4.11.4 Previsões

Após a validação do modelo têm-se o principal objetivo do trabalho, dessa forma, simulou-se previsões para se analisar o comportamento futuro da demanda, ou seja, quais os prováveis dados de consumo para os anos seguintes. As previsões para os próximos cinco anos com intervalo de confiança de 95%, podem ser vistas na Tabela 23, e de forma conjunta com a evolução da série histórica de café na Figura 32.

Tabela 23 - Previsões para a o consumo de café no Brasil para os anos de 2017 a 2021 com intervalo de confiança de 95%.

Período	Previsão	Limite Inferior	Limite Superior

2017	21.438,9	20.988,8	21.889,0
2018	22.179,9	21.419,9	22.939,8
2019	22.776,7	21.829,4	23.723,9
2020	23.578,4	22.428,0	24.728,8
2021	24.043,7	22.851,9	25.235,4

Fonte: Autoria própria.

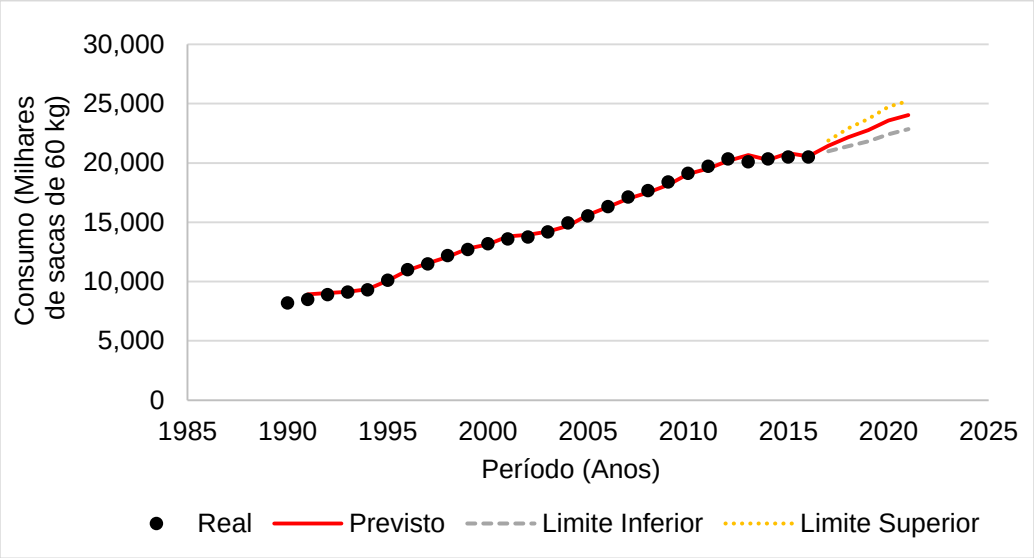


Figura 32 - Gráfico comparativo entre os valores reais e previstos para o consumo de café com previsões para 5 períodos e intervalo de confiança de 95%.
Fonte: Autoria própria.

O modelo ARIMA (1,1,5) possui desvio médio absoluto (MAD) de 147.210 sacas de 60 kg e a previsão para o período de 2017 é de 21,44 milhões de sacas de 60 kg, com arredondamento e conversão de escala. A estimativa da ABIC (2017) para o ano de 2017 foi de aproximadamente 22 milhões de sacas de 60 kg, cerca de 3% a 3,5% em relação a 2016. O Cecafé (2017) projetou um consumo de 24 milhões de sacas até 2019, diante de 22,78, previstos pelo método, o que caracteriza uma leve tendência a subestimar as previsões, embora, a própria ABIC (2017) considera essa previsão uma “[...] hipótese otimista”, em relação a fatores econômicos adversos,

Observa-se ainda, que a previsão mais relevante é a do ano imediatamente posterior ao último medido, uma vez que as demais previsões estão sujeitas à reajustes com as informações futuramente coletadas, além da inclusão de parâmetros qualitativos de análises econômicas feita por especialistas. Contudo, é notável a aderência da curva prevista em relação a real, que consequentemente possui uma projeção coerente para os 5 períodos futuros, o que leva a ter intervalos de confiança relativamente menores ou mais precisos.

5 CONSIDERAÇÕES FINAIS

Este estudo teve como principal intuito realizar previsões confiáveis para a demanda de café no Brasil, com base nisso, buscou-se por meio da confecção, interpretação e avaliação de diversos modelos matemáticos, atender a esse propósito. Os modelos, foram feitos a partir de metodologias estatísticas anteriormente exploradas de forma teórica na literatura e aplicadas a série histórica de consumo café no Brasil entre 1990 e 2016.

Por meio dos parâmetros estabelecidos para as análises, considera-se que o objetivo foi atendido de forma satisfatória em termos estatísticos, uma vez, que o modelo ajustado escolhido para as previsões apresentou um desvio pequeno em comparação a série real e mostra-se confiável para a execução de previsões.

De forma análoga ao objetivo geral, todos os objetivos específicos estipulados foram atendidos. Isso gerou a cada etapa considerações e aprendizado agregado, o que pode ser considerado como um dos objetivos implícitos ao trabalho. Assim como absorver o conhecimento com o estudo e a busca na literatura, tem-se a importância em gerar conhecimento por intermédio da aplicação da pesquisa realizada.

Durante a formulação e análises dos modelos, pode-se compreender diversos critérios e conceitos estatísticos que auxiliaram nas interpretações e decisões intrínsecas no processo de modelagem. Com a adição dos fatores de influência, pode-se entender em termos estatísticos as características econômicas do mercado de café no Brasil. Interpretou-se esse produto, como item básico de consumo, a julgar por um caráter popular e cultural, na qual, sua demanda tende a se manter, mesmo que seu preço eventualmente oscile.

Por fim atendeu-se também a motivação para esse trabalho. Na qual atribuiu-se a utilização da previsão de demanda como ferramenta no auxílio à tomada de decisão para todo o âmbito de planejamento estratégico. Por meio da modelagem matemática e análise estatística, permitiu-se a aplicação dos modelos comparados por gestores, para entender e antecipar-se as características do mercado consumidor de café no Brasil.

6 SUGESTÕES PARA TRABALHOS FUTUROS

Como sugestão para futuros trabalhos, indica-se a combinação dos melhores métodos através de média simples ou mesmo, ponderando-se pesos mais elevados em ordem de colocação dos métodos, como por exemplo, no trabalho de Bacci (2006), em que se compara os métodos estatísticos conhecidos com uma combinação entre os melhores modelos ajustados.

Outra possível vertente de estudo é a de inclusão de outros métodos de algoritmos computacionais, como redes neurais e lógica Fuzzy, para a análise comparativa. Como exemplo, o trabalho de Miranda *et al*. (2013), no qual, aborda-se como principal objetivo, comparar modelos econométricos e "[...] identificar qual modelo tem o poder de previsão mais robusto e dinâmico em relação à rede neural no mercado futuro do café arábica".

No trabalho de Campos (2008), no qual utilizou-se da comparação de métodos, tais quais, ARIMA, Redes Neurais e Neuro-Fuzzy para prever o consumo de energia elétrica. Ainda, os métodos testados por Faustino (2011), que correspondem, com base na classificação do trabalho, à métodos estatísticos, não-lineares e por lógica nebulosa, além de testes com combinações entre métodos. Esses trabalhos denotam a relevância da abordagem estatística, operacional e computacional, agregam ideias e corroboram para o estudo dessa presente pesquisa.

Têm-se também a importância da análise qualitativa, principalmente em casos de previsões para prazos maiores. Os métodos quantitativos funcionam muito bem no curto prazo, mas ao passar do tempo, vão se aglomerando fatores econômicos, sociais e técnicos que são difíceis a qualquer modelo considerar. Por isso, se faz necessário a complementação com informações qualitativas, geralmente oriunda da análise de especialistas no ramo.

Por fim, a aplicação dos métodos estudados em outras séries de consumo ou de qualquer outro contexto que se almeje obter previsões. Com a consideração das características específicas de cada série, pode-se utilizar os conhecimentos obtidos e apresentados nesse trabalho para agregar a outros estudos subsequentes.

REFERÊNCIAS

ABIC – Associação Brasileira da Industria de Café, 2017. Disponível em: <www.abic.com.br/>. Acesso em: 16 de maio de 2017.

ANDERSON, David R.; SWEENEY, Dennis J.; WILLIAMS, Thomas A. **Estatística aplicada a administração e economia.** 2. ed. São Paulo: Cengage Learning, 2011. 597 p.

AVANZA, Marcia. **Café não provoca câncer, concluí OMS.** Jornal da USP, 2016. Disponível em: < http://jornal.usp.br/atualidades/cafe-nao-provoca-cancer-conclui-oms/>. Acesso em: 12 de junho de 2017.

ARAUJO, Marco Antonio de. **Administração de produção e operações:** uma abordagem prática. Rio de Janeiro: Brasport, 2009. 393 p.

BACCI, Lívio Agnew. **Combinação de métodos de séries temporais para previsão da demanda de café no Brasil.** Itajubá: Universidade Federal de Itajubá, 2007. Disponível em: <http://saturno.unifei.edu.br/bim/0031728.pdf>. Acesso em: 28 de abril de 2017.

BALLOU, Ronald H. **Gerenciamento da cadeia de suprimentos/logística empresarial.** Tradução Raul Rubenich. 5. ed. Porto Alegre: Bookman, 2006. p. 241-267.

BARROS, Mônica. **Processos estocásticos.** Rio de Janeiro: Publit, 2003. p. 202-229.

BERTOLDE, Adelmo Inácio, XAVIER JUNIOR, Walter Paganucci. **Uma comparação de métodos de previsão de demanda de peças de reposição:** uma aplicação ao transporte ferroviário**.** Revista Gestão Industrial. Ponta Grossa: UTFPR, 2013. 17 p. Disponível em: <https://periodicos.utfpr.edu.br/revistagi/article/view/1267/1096>. Acesso em: 25 de maio de 2017.

BERTOLO, Luiz Antonio. **Métodos Quantitativos**. 2017. Disponível em: <http://www.bertolo.pro.br/MetodosQuantitativos/Previsao/pmc4.htm>. Acesso em: 04 de novembro de 2017.

BISQUERRA, Rafael; SARRIERA, Jorge Costellá; MARTÍNEZ, Francesc.

Introdução estatística: enfoque informático com o pacote estatístico SPSS. Porto Alegre: Artmed, 2004. 256 p.

BONOTTO, Giulia; FOGLIATTO, Flávio Sanson. **Previsão de demanda a partir de métodos quantitativos aplicada ao setor varejista.** Porto Alegre: Universidade Federal do Rio Grande do Sul-UFRGS, 2015. Disponível em: <https://www.lume.ufrgs.br/bitstream/handle/10183/147496/000999425.pdf?sequence=1>. Acesso em: 05 de julho de 2017.

CAMPOS, Roger Júnio. **Previsão de séries temporais com aplicações a séries de consumo de energia elétrica.** Universidade Federal de Minas Gerais-UFMG. Belo Horizonte, 2008. 110 p. Disponível em: < https://www.ppgee.ufmg.br/documentos/Defesas/752/Dissertacao_Mestrado_Roger.pdf>. Acesso em: 14 de novembro de 2017.

CONAB – Companhia Nacional do Abastecimento, 2017. Disponível em: <http://www.conab.gov.br/>. Acesso em: 12 de junho de 2017.

CARDOSO, Andréa. **Fundamentos da pesquisa operacional.** UNIFAL-MG. 2011. 102 p. Disponível em: <http://moodle.utfpr.edu.br/pluginfile.php/208143/mod_resource/content/1/po.pdf>. Acesso em: 23 de abril de 2017.

CAVALCANTI, Marco Antônio Freitas de Hollanda; MELLO, Fabrício. **Modelos ARIMA**: teoria e aplicações. Técnicas de Pesquisa em Economia. Departamento de Economia. Pontifícia Universidade Católica do Rio de Janeiro. Rio de Janeiro, 2014.

CECAFÉ - Conselho dos Exportadores de Café do Brasil, 2017. Disponível em: http://www.cecafe.com.br/sobre-o-cafe/consumo/>. Acesso em: 17 de novembro de 2017.

CORRÊA, Henrique L.; CORRÊA, Carlos A. **Administração de produção e operações:** manufatura e serviços: uma abordagem estratégica. 3. ed. São Paulo: Atlas, 2012. p. 235-269.

DEVORE, Jay L. **Probabilidade e estatística:** para engenharia e ciências. São Paulo: Cengage Learning, 2013. 685 p.

DAVIS, Mark M.; AQUILANO, Nicholas J.; CHASE, Richard B. **Fundamentos da administração da produção.** 3. ed. São Paulo: Bookman, 1999. p. 263-266.

EHLERS, Ricardo Sandes. **Análises de séries temporais.** 4. ed. Departamento de estatística, UFPR, 2007. Disponível em: julho de 2017.

FARIA, Alexandre Carlos dos Santos; MANOLESCU, Friedhilde M. K. **A produção de café no Brasil.** VIII Encontro Latino Americano de Iniciação Cientifica e IV Encontro Latino Americano de Pós-Graduação. UNIVAP, 2004. Disponível em: <http://www.inicepg.univap.br/cd/INIC_2004/trabalhos/inic/pdf/IC6-8.pdf>. Acesso em: 19 de maio de 2017.

FARIAS, Jaiana Brunhauser; TREVISAN, Laís Viera. **Séries temporais.** Santa Maria: Universidade Federal de Santa Maria-UFSM, 2013.

FAUSTINO, Cláudio Paulo. **Previsão de séries temporais via modelos baseados em regras.** Universidade Federal De Itajubá – UNIFEI. Itajubá, 2011. 128 p. Disponível em: < https://saturno.unifei.edu.br/bim/0038068.pdf>. Acesso em: 14 de novembro de 2017.

FERRARI, Fabricio. **Estatística básica.** Universidade Federal do Rio Grande – FURG: Rio Grande, 2004. 33 p. Disponível em: <http://www.ferrari.pro.br/home/documents/FFerrari-Curso-Estatistica-Basica.pdf>. Acesso em: 13 de novembro de 2017.

FERNANDES, Flavio Cesar Faria; GODINHO FILHO, Moacir. **Planejamento e controle da produção:** dos fundamentos ao essencial. São Paulo: Atlas, 2010. p. 17-46.

FERREIRA, Lucas Tadeu; SANTOS, Jamilsen. **Consumo de café superou produção mundial em 2016 pelo terceiro ano consecutivo.** Brasilia: Embrapa, 2017. Disponível em: <https://www.embrapa.br/busca-de-noticias/-/noticia/21131685/consumo-de-cafe-superou-producao-mundial-em-2016-pelo-terceiro-ano-consecutivo>. Acesso em: 05 de junho de 2017.

FISCHER, Séiergio. **Séries univariantes de tempo** – metodologia de Box & Jenkins. Porto Alegre: FEE, 1982. 186 p.

FRANCO, Glaura da Conceição. **Apostila de modelos lineares em séries temporais.** Belo Horizonte: Universidade Federal de Minas Gerais – UFMG, 2016. 144 p. Disponível em: <ftp://est.ufmg.br/pub/glaura/MLST/Modelos%20Lineares%20em%20S%E9ries%20Temporais.pdf>. Acesso em: 03 de novembro de 2017.

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

GIL, Antonio Carlos. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Atlas, 2008.

GIRARDI, Thaís Rohling; ISHIKAWA, Gerson. **Previsão De Curto Prazo De Carga De Energia Elétrica Do Estado De Santa Catarina.** Bauru. XIII Simpósio de Engenharia de Produção – SIMPEP, 2006. Disponível em: < http://www.simpep.feb.unesp.br/anais/anais_13/artigos/478.pdf>. Acesso em: 31 de outubro de 2017.

GLEN, Stephanie. **Pearson Correlation:** definition and easy steps for use. Statistics How To, 2012. Disponível em: < http://www.statisticshowto.com/what-is-the-pearson-correlation-coefficient/>. Acesso em: 15 de junho de 2017.

GROPPO, Juliano Daniel. **Estudo de tendências nas séries temporais de qualidade da água de rios do Estado de São Paulo com diferentes graus de intervenção antrópica.** Piracicaba, 2005. 72 p.

GUTIÉRREZ, José Luís Carrasco; ROMANEL, Celso. **Aplicação de redes neurais na previsão de vazão através da fundação da ombreira esquerda da barragem Corumbá-I.** Pontifícia Universidade Católica do Rio De Janeiro - Departamento de Engenharia Civil. Salvador: Comitê Brasileiro de Barragens - XXV Seminário Nacional de Grandes Barragens, 2003. Disponível em: <http://www.cbdb.org.br/documentos/site/92/9223.pdf>. Acesso em: 12 de junho de 2017.

HYNDMAN, Rob J.; ATHANASOPOULOS, George. **Forecasting:** principles and practice. Texts Online, OPEN-ACESS TEXTBOOKS. 2017. Disponível em: <https://www.otexts.org/fpp>. Acesso em: 30 de maio de 2017.

IBGE - Instituto Nacional de Geografia e Estatística, 2017. Disponível em: <www.ibge.gov.br/>. Acesso em: 07 de maio de 2017.

JACOBS, William; ZANINI, Roselaine Ruviaro; COSTA, Manfred. **Estudo comparativo de séries temporais para previsão de vendas de um produto.** Iberoamerican Journal of Industrial Engineering, Florianopolis, SC, Brazil, v. 6, n. 12, p. 112-133, 2014. Disponível em: <http://incubadora.periodicos.ufsc.br/index.php/IJIE/article/view/2599/pdf_63>. Acesso em: 20 de maio de 2017.

KAUARK, Fabiana; MANHÃES, Fernanda Castro; MEDEIROS, Carlos Henrique. **Metodologia da pesquisa:** um guia prático. Itabuna: Via Litterarum, 2010. 88 p.

KIRCHNER, Rosane Maria. **Modelagem e previsão por meio de metodologia Box & Jenkins:** uma ferramenta de gestão. Universidade Federal de Santa Maria – UFSM: Santa Maria, 2006. 63 p. Disponível em: < http://cascavel.ufsm.br/tede/tde_arquivos/12/TDE-2007-05-28T081855Z-603/Publico/Kirchner.pdf>. Acesso em: 08 de novembro de 2017.

KÖCHE, José Carlos. **Fundamentos da metodologia científica:** teoria da ciência e iniciação à pesquisa. 29. ed. Petrópolis: Vozes, 2011.

KRAJEWSKI, Lee J.; RITZMAN, Larry; MALHOTRA, Manoj. **Administração de produção e operações.** Tradução Mirian Santos Ribeiro de Oliveira; revisão técnica André Luís de Castro Moura Duarte e Susana Carla Farias Pereira. São Paulo: Pearson Prentice Hall, 2009. p.435-473.

LEVINE, David M.; BERENSON, Mark L.; STEPHAN, David. **Estatística:** teoria e aplicações usando o Microsoft® Excel em Português. Rio de Janeiro: LTC, 2000.

LOBO, Bernardo Almada. **Análise de erros.** Departamento de Engenharia Mecânica e Gestão Industrial - DEMEGI. Faculdade de Engenharia da Universidade do Porto – FEUP. Porto, 2007. Disponível em: < https://sigarra.up.pt/feup/pt/conteudos_service.conteudos_cont?pct_id=34385&pv_cod=1112qaTawyc8>. Acesso em: 31 de outubro de 2017.

LUSTOSA, Leonardo; MESQUITA, Marco A.; QUELHAS, Osvaldo; OLIVEIRA, Rodrigo. **Planejamento e controle da produção.** 10. ed. Rio de Janeiro: Elsevier, 2008.

MAKRIDAKIS, Spyros; WHEELWRIGT, Steven. C.; HYNDMAN, Rob. J. **Forecasting:** methods and applications. 3. ed. New York: John Wiley & Sons, 1998.

MARCONI, Mariana de Andrade; LAKATOS, Eva Maria. **Técnicas de pesquisa: planejamento e execução de pesquisas, amostragens e técnicas de pesquisa, elaboração, análise e interpretação de dados**. 7. ed. 7. reimpr. São Paulo: Atlas, 2013. 277 p.

MARTINS, Gilberto de Andrade. **Estatística geral e aplicada.** 2. ed. São Paulo: Atlas, 2002. 417 p.

MIRANDA, André Pacheco; CORONEL, Daniel Arruda; VIEIRA, Kelmara Mendes. **Previsão do mercado futuro do café arábica utilizando redes neurais e métodos econométricos.** Revista Estudos do CEPE, n38. Santa Cruz do Sul, 2013. p. 66-98. Disponível em: < https://online.unisc.br/seer/index.php/cepe/article/download/3329/3122>. Acesso em: 14 de novembro de 2017.

MARTINS, Petrônio Garcia; LAUGENI, Fernando Piero. **Administração da produção**. 2. ed. São Paulo: Saraiva, 2005. p. 225-258.

MONTGOMERY, Douglas C.; RUNGER, George C. **Estatística aplicada e probabilidade para engenheiros.** 5. ed. Tradução e revisão técnica Verônica Calado, Rio de Janeiro: LTC, 2012. 521 p.

MOREIRA, Daniel Augusto. **Administração da produção e operações**. 2. ed. rev. E ampl. São Paulo: Cengage Learning, 2011.

MORETTIN, Pedro Alberto; BUSSAB, Wilton O. **Estatística básica.** 7. ed. São Paulo: Saraiva, 2011. p. 35-56.

MORETTIN, Pedro Alberto; TOLOI, Clélia Maria de Castro. **Modelos para previsão de séries temporais**. Instituto de Matemática Pura e Aplicada. Rio de Janeiro, 1981. 350 p.

MORETTIN, Pedro Alberto; TOLOI, Clélia Maria de Castro. **Séries temporais -** métodos quantitativos. São Paulo: Atual, 1985.

MORETTIN, Pedro Alberto; TOLOI, Clélia Maria de Castro. **Análise de séries temporais**. São Paulo: Edgard Blucher, 2004. 535 p.

NAVIDI, Willian. **Probabilidade e estatística para ciências exatas.** Tradução: José Lucimar do Nascimento; revisão técnica: Antonio Pertence Júnior. Porto Alegre: AMGH, 2012. 604 p.

NIST – *National Institute of Standards and Technology.* ***e-Handbook of Statistical Methods.*** Gaithersburg. 2017. Disponível em: < http://www.itl.nist.gov/div898/handbook/>. Acesso em: 09 de novembro de 2017.

NOGUEIRA, Fernando. **Modelagem e simulação:** Modelos de previsão. Notas de aula, 2017. Disponível em: <http://www.ufjf.br/epd042/files/2009/02/previsao1.pdf>. Acesso em: 31 de maio de 2017.

OIC - Organização Internacional Do Café, 2017. Disponível em: <www.ico.org/>. Acesso em: 28 de abril de 2017.

OLIVEIRA, Maxwell Ferreira de. **Metodologia científica:** um manual para a realização de pesquisas em Administração. Catalão: UFG, 2011. 72 p.

OLIVEIRA, Pedro; LIMA, Vladimir de. **Previsão de demanda:** o básico que você precisa saber. 1. ed. São Paulo: Baraúna, 2016.159 p.

PEINADO, Jurandir; GRAEML, Alexandre Reis. **Administração da produção:** operações industriais e de serviços. Curitiba: UnicenP, 2007. 750 p.

PELLEGRINI, Fernando Rezende**. Metodologia para implementação de sistemas de previsão de demanda.** Mestrado em Engenharia de Produção, Escola de Engenharia. UFRGS, 2000. Disponível em: < http://www.producao.ufrgs.br/arquivos/publicacoes/Fernando%20R%20Pellegrini.pdf >. Acesso em: 25 de maio de 2017.

PORTAL ACTION, Séries Temporais: Suavização Exponencial de Holt, 2017. Disponível em: <http://www.portalaction.com.br/series-temporais/33-suavizacao-exponencial-de-holt-seh>. Acesso em: 05 de junho de 2017.

PRODANOV, Cleber Cristiano; FREITAS, Ernani Cesar de. **Metodologia do trabalho científico:** métodos e técnicas da pesquisa e do trabalho acadêmico. 2. ed. Novo Hamburgo: Feevale, 2013. 276 p.

QUEIROZ, Fernanda Cristina Barbosa Pereira; HÉKIS, Hélio Roberto; ANDRADE, Dalliane Vanessa Pires; QUEIROZ, Jamerson Viegas; MACÊDO, Danielle Moraes de. **Previsão e séries temporais para tomada de decisão empresarial em uma indústria moveleira da região de Criciúma–SC.** Revista Catarinense da Ciência Contábil – CRCSC – Florianópolis, v. 11, n. 32, p. 26-42, abr./jul. 2012. Disponível em: <http://www.redalyc.org/pdf/4775/477548341003.pdf>. Acesso em: 05 de junho de 2017.

QUININO, Roberto C.; REIS, Edna A.; BESSEGATO, Lupércio, F. **O coeficiente de determinação r² como instrumento didático para avaliar a utilidade de um modelo de regressão linear múltipla.** Belo Horizonte: UFMG, 2011. Disponível em:

< http://www.est.ufmg.br/portal/arquivos/rts/PD_28102011_Final.pdf>. Acesso em: 12 de junho de 2017.

REIS, Edna. A., REIS Ilka. A. **Análise descritiva de dados.** Relatório técnico do departamento de estatística da UFMG, 2002. 64 p. Disponível em: <http://www.est.ufmg.br/portal/arquivos/rts/rte0202.pdf.>. Acesso em: 10 de novembro de 2017.

RODRIGUES, Helena Leandro; DIAS, Frederico Divino; TEIXEIRA, Natália de Carvalho. **A origem do café no Brasil: A semente que veio para ficar.** Revista Pensar Gastronomia, v.1, n.2, 2015. Disponível em: <http://revistapensar.com.br/gastronomia/pasta_upload/artigos/a44.pdf>. Acesso em: 20 de maio de 2017.

RUSSELL, Roberta S; TAYLOR, Bernard W. **Operations management:** creating value along the supply chain. 7. ed. Wiley, New York, 2011. p. 495- 552.

ROSSETTO, Marta; DEIMLING, Moacir Francisco; ZANIN, Antonio; RODRIGUES, Márcio da Paixão; ROCHA NETO, Anselmo. **Técnicas qualitativas de previsão de demanda:** um estudo multicasos com empresas do ramo de alimentos. Resende: VIII Simpósio de Excelência em Gestão e Tecnologia- SEGeT, 2011. Disponível em: < https://www.aedb.br/seget/arquivos/artigos11/55814697.pdf>. Acesso em: 17 de novembro de 2017.

SAMOHYL, Robert Wayne, ROCHA, Rubson, MATTOS, Viviane Leite D. de. **Utilização do método de holt-winters para previsão do leite entregue às indústrias catarinenses.** Salvador: XXI Encontro Nacional de Engenharia de Produção - ENEGEP, 2001. Disponível em: < http://www.abepro.org.br/biblioteca/ENEGEP2001_TR66_0742.pdf>. Acesso em: 25 de maio de 2017.

SANTOS, Verena. **Séries Temporais: ARIMA.** Belém: Universidade Federal do Pará - Instituto de Ciências Exatas e Naturais - Faculdade de Estatística, 2014. Disponível em: <http://www.ufpa.br/heliton/arquivos/aplicada/seminarios/M1_08_ARIMA_Verena .pdf>. Acesso em: 12 de junho de 2017.

SILVA, Edna Lúcia da; MENEZES, Estera Muszkat. **Metodologia da pesquisa e elaboração de dissertação.** 4. ed. Florianópolis: UFSC, 2005. 138 p.

SILVA, Lincoln Teixera da; DANTAS, Tiago Mendes. **Previsão de seguros de automóveis utilizando modelagem estrutural.** Rio de Janeiro: Escola Nacional de

Ciências Estatísticas, 2009. Disponível em: <http://www2.ime.unicamp.br/sinape/sites/default/files/PREVIS%C3%83O%20DE%20SEGUROS%20DE%20AUTOM%C3%93VEIS%20UTILIZANDO%20MODELAGEM%20ESTRUTURAL_0.pdf>. Acesso em: 12 de junho de 2017.

SMITH, Michael J. de. ***Statistical Analysis Handbook*****:** *a comprehensive handbook of statistical concepts, techniques and software tools.* STATSREF, 2015. 677 p. Disponível em: <http://www.statsref.com/HTML/index.html>. Acesso em: 11 de novembro de 2017.

SOUZA, Gueibi Peres; SAMOHYL, Robert W. **Combinação de previsões em séries temporais do consumo industrial de energia elétrica em Santa Catarina.** Porto Alegre: XXV ENEGEP, 2005. Disponível em: < http://www.abepro.org.br/biblioteca/enegep2005_enegep0202_0502.pdf>. Acesso em: 31 de maio de 2017.

SPIEGEL, Murray R. **Estatística.** Tradução e revisão técnica Pedro Consentino. São Paulo: Pearson Makron Books, 1993. 643 p.

TUBINO, Dalvio Ferrari. **Planejamento e controle da produção:** teoria e prática. 2. ed. São Paulo, 2009. p. 16-34.

WALPOLE, Ronald E.; MYERS, Raymond H.; MYERS, Sharon L.; YE, Keying. **Probabilidade e estatística para engenharia e ciências.** 8. Ed. São Paulo: Pearson Prentice Hall, 2009. 491 p.