

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

CAIQUE LUCHINI

**APLICAÇÃO DE MODELOS ESTOCÁSTICOS PARA PREVISÃO DE DEMANDA
DE UM PRODUTO AGROQUÍMICO**

LONDRINA

2022

CAIQUE LUCHINI

**APLICAÇÃO DE MODELOS ESTOCÁSTICOS PARA PREVISÃO DE DEMANDA
DE UM PRODUTO AGROQUÍMICO**

Stochastic Models For Demand Forecasting Of A Crop Protection Product

Trabalho de conclusão de curso apresentado como requisito para obtenção do título de Bacharel em Engenharia de Produção da Universidade Tecnológica Federal do Paraná (UTFPR) do campus Londrina.
Orientador(a): Prof. Dr. Rogério Tondato.
Coorientador(a):

LONDRINA

2022



[4.0 Internacional](https://creativecommons.org/licenses/by-nc-sa/4.0/)

Esta licença permite remixe, adaptação e criação a partir do trabalho, para fins não comerciais, desde que sejam atribuídos créditos ao(s) autor(es) e que licenciem as novas criações sob termos idênticos. Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

CAIQUE LUCHINI

**APLICAÇÃO DE MODELOS ESTOCÁSTICOS PARA PREVISÃO DE
DEMANDA DE UM PRODUTO AGROQUÍMICO**

Trabalho de conclusão de curso apresentado
como requisito para obtenção do título de
Bacharel em Engenharia de Produção da
Universidade Tecnológica Federal do Paraná
(UTFPR) do campus Londrina.

Orientador(a): Prof. Dr. Rogério Tondato.

Coorientador(a):

Data de aprovação: 07/junho/2022

Rogério Tondato

Doutor

Universidade Tecnológica Federal do Paraná

Silvana Rodrigues Quintilhano

Doutora

Universidade Tecnológica Federal do Paraná

José Ângelo Ferreira

Doutor

Universidade Tecnológica Federal do Paraná

LONDRINA

2022

Este trabalho é em dedicação aos meus pais e
irmão, a vocês todo meu esforço e carinho.

LISTA DE FIGURAS

Figura 1 – Esquema sistema dinâmico	11
Figura 2 – Série temporal do preço real do Frango x Tempo	19
Figura 3 – Comparativo dos indicadores para os modelos testados	19
Figura 4 – Resultados considerando o MOD 2 para 2004.....	20
Figura 5 – Resultados considerando o MOD 2 para 2005.....	20
Figura 6 – Número de atendimentos para a primeira categoria	21
Figura 7 – Comparação entre os modelos testados para a primeira categoria	22
Figura 8 – Previsão de demanda para a primeira categoria	22
Figura 9 – Previsão de demanda para a segunda categoria.....	23
Figura 10 – Previsão de demanda para a terceira categoria	23
Figura 11 – Demanda total da empresa	24
Figura 12 – Resultado de vendas segmentado para o herbicida	26
Figura 13 – Sobreposição das vendas do herbicida	26
Figura 14 – Série temporal das vendas do herbicida	27
Figura 15 – ACF para a série original	28
Figura 16 – PACF da série original	29
Figura 17 – ACF e PACF dos resíduos do Modelo ARIMA(1,1,0)	30
Figura 18 – Comparação entre os modelos testados para o herbicida	31
Figura 19 – Previsão da demanda para os próximos dois ciclos	31
Figura 20 – Comparação do modelo escolhido com a série de resíduos	32

RESUMO

Previsão de demanda no contexto de mercado atual é ferramenta quase que imprescindível. A competitividade da contemporaneidade implica na imposição de processos cada vez mais robustos e com menor subjetividade para manutenção e sobrevivência das organizações. Nesse espaço, o presente trabalho, através de uma pesquisa quantitativa, explicativa e experimental, tem por objetivo avaliar o comportamento dos algoritmos de previsão de séries temporais na previsão de demanda de um produto da indústria agroquímica. Para tanto, a literatura correlata é apresentada no contexto do mercado agroindustrial para citar e introduzir os métodos e algoritmos com maior grau de utilização. Além disso, ele discorre sobre a previsão de demanda no setor agroquímico, apresentando e aplicando a metodologia ARIMA para um conjunto de dados de vendas de um produto herbicida fornecidos por uma empresa multinacional de grande atuação no território brasileiro. É possível notar que o modelo utilizado teve pouca representatividade na expectativa de previsão da demanda futura dado o comportamento dos resíduos. No entanto, o método científico estabelecido e a estruturação do processo são, notadamente, resultados positivos consequentes do trabalho.

Palavras-chave: Séries temporais; ARIMA; Agronegócio.

ABSTRACT

Demand forecasting in the current market context is an almost essential tool. The competitiveness of contemporaneity implies the imposition of increasingly robust processes, with less subjectivity for the maintenance and survival of organizations. In this space, the present work, through quantitative, explanatory, and experimental research, aims to evaluate the behavior of time series forecasting algorithms in forecasting demand for a product of the agrochemical industry. To this end, the related literature is presented in the context of the agro-industrial market to cite and introduce the methods and algorithms with the highest degree of usage. In addition, it discusses demand forecasting in the agrochemical sector, presenting and applying the ARIMA methodology to a set of sales data for an herbicide product provided by a multinational company with a large presence in Brazil. It is possible to notice that the model used had little representation in the expectation of forecasting future demand given the behavior of waste curves. However, the established scientific method and the structuring of the process are, notably, positive results from this work.

Keywords: Time series; ARIMA; Agribusiness.

SUMÁRIO

1	INTRODUÇÃO	9
1.1	Objetivos	10
1.2	Justificativa	11
2	REFERENCIAL TEÓRICO	12
2.1	Séries Temporais	12
2.2	Sistemas Estacionários	13
2.3	Modelos Paramétricos	14
2.4	Modelos de erro ou de regressão	15
2.5	Modelos ARIMA (Autorregressivo Integrado de Média Móvel)	15
3	MÉTODOS E TÉCNICAS DE PESQUISA	18
4	TRABALHOS CORRELATOS	20
4.1	Aplicação de Modelo ARIMA para previsão do preço do frango inteiro resfriado no grande atacado do Estado de São Paulo	20
4.2	Previsão de demanda: Uma aplicação dos modelos de Box-Jenkins na área de assistência técnica de computadores pessoais	22
5	AQUISIÇÃO E TRATAMENTO DOS DADOS HISTÓRICOS	27
6	APLICAÇÃO DO MÉTODO BOX-JENKINS	29
6.1	Identificação	29
6.2	Estimativa	31
6.3	Verificação	32
6.4	Previsão	33
7	CONCLUSÃO	35
	REFERÊNCIAS	37

1 INTRODUÇÃO

Previsão de demanda é, de maneira geral, tarefa difícil no setor agroquímico. A variedade de aplicações, clientes, tipos de cultura, regionalizações, interação dos diferentes ativos, bem como a atuação de diversas organizações concorrendo pelo mercado final são apenas o começo do desafio em encontrar, de forma concreta, a demanda de cada produto.

Não só no setor agroquímico, mas em grande parte das organizações, a previsão de demanda é parte de um programa de negócios chamado de S&OP – *Sales and Operations Planning* (Planejamento de Vendas e Operações). Tal programa se dá de forma exaustiva e subjetiva, baseada na experiência dos profissionais envolvidos, levando em consideração, quase que exclusivamente, a percepção da realidade observada pelos executores.

O início da previsão de demanda, normalmente, se baseia no histórico, no planejado e no realizado do ano anterior. Não há, em muitos casos, projeções estatísticas ou aplicações de modelos teóricos de como esses dados irão se comportar. Isso implica em maior empenho de recursos para se alcançar uma assertiva previsão.

Nesse sentido, a análise de séries temporais para a previsão de demanda é uma metodologia amplamente utilizada no mercado atual, diversos modelos estocásticos são desenvolvidos e aplicados com o anseio de mapear a curva de demanda de um produto, ou antever a volatilidade de custos. A aplicação de métodos estocásticos para a previsão da demanda incorre na diminuição da subjetividade da projeção. No entanto, adequar o modelo, levando em consideração os diversos fatores influenciadores, para a observação da realidade prática é, na maior parte dos casos, o grande desafio.

1.1 Objetivos

O objetivo geral desse trabalho é avaliar o comportamento dos algoritmos de previsão de séries temporais na previsão de demanda de um produto da indústria agroquímica.

Os objetivos específicos desse trabalho são:

- Realizar uma revisão teórica sobre os principais modelos de previsão de demanda.
- Definir e coletar dados de demanda histórica do produto a ser aplicada a previsão de demanda futura.
- Aplicar e comparar diferentes tipos de modelos estocásticos para os dados históricos capturados.

1.2 Justificativa

O ponto de partida para uma previsão de demanda tem fundamental importância para a quantidade de recursos que deverão ser alocados para a realização desse processo em uma S&OP. De maneira geral, se utiliza como um ponto de partida básico o histórico realizado no período correspondente, ou seja, se a previsão de demanda está sendo elaborada a fim de encontrar a demanda do mix de produtos para o ano seguinte, o realizado do ano atual será a base para sua construção.

Ao empregar recursos para elaborar um ponto de partida mais próximo da demanda dos produtos, se utilizando para tanto de recursos estatísticos e modelos de previsão de demanda teóricos bem estabelecidos, espera-se economizar tempo e minimizar a subjetividade dos envolvidos, justificando o desenvolvimento da presente proposta de pesquisa.

De modo geral, o setor agroquímico pouco investe em tecnologia e conhecimento científico na área de análise de demanda. Testar modelos teóricos bem definidos em dados reais pode qualificar tais modelos na área aplicada introduzindo métodos científicos em aplicabilidades reais do mercado. Isso não só beneficia e valoriza a comunidade científica e suas pesquisas bem como ao setor produtivo.

Além do exposto, há grande desenvolvimento profissional e visibilidade para o pesquisador caso haja implementação da pesquisa na organização. Isso se amplia caso de fato haja melhora no processo com tamanha relevância para empresa.

2 REFERENCIAL TEÓRICO

2.1 Séries Temporais

Modelar matematicamente o comportamento de um fenômeno físico é ideia muito bem estabelecida, ou seja, é possível calcular a trajetória de um lançamento conhecendo sua velocidade, direção e as leis físicas que regem o movimento. O modelo seria determinístico caso fosse um cálculo exato para a posição do objeto após determinado tempo, porém, como há impactos de fatores não conhecidos – ou exatos – como a variação da atuação do vento, por exemplo, se faz necessário a utilização do cálculo de probabilidade de o evento futuro ocorrer, nesse caso a posição do objeto. O objetivo do modelo probabilístico é, portanto, se aproximar do modelo determinístico, que, para situações reais, nada mais é que a observação da série temporal após sua ocorrência (Box *et. al.*, 1994).

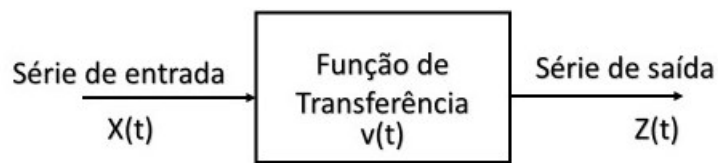
Séries temporais, por sua vez, são definidas por Morettin (2006) como sendo um conjunto de observações discretas ordenadas no tempo. Há dois grandes conjuntos de modelos definidos para séries temporais, paramétricos e não paramétricos. Eles se diferenciam pelo seu domínio, enquanto o primeiro se refere ao tempo e apresenta, portanto, número finito de parâmetros o segundo se refere à frequência, surgindo a análise espectral.

O estudo de uma série temporal tem por objetivo um ou mais dos seguintes três fatores:

- i. Fazer previsões de valores futuros;
- ii. Descrever o comportamento da série;
- iii. Encontrar e descrever o mecanismo gerador da série;

Modelos de séries temporais, tanto probabilísticos quanto estocásticos, são, por diversas vezes, representados por sistemas dinâmicos, constituídos por uma série de entrada, uma função transferência e uma série de saída. A figura 1 apresenta a dada representação:

Figura 1 – Esquema Sistema Dinâmico



Fonte: Morettin (2006) adaptado

A modelagem de uma série temporal deve seguir o princípio da Parcimônia, isso é, o modelo que representa mais adequadamente a série temporal é aquele que se utiliza do menor número de parâmetros possível (Cryer D. J. 2008). De modo geral uma série temporal é dada por vetor $Z(t)$ de ordem $r \times 1$, $r \in N$.

$$Z(t) = [Z_1(t), Z_2(t), Z_3(t), \dots, Z_r(t)]$$

2.2 Sistemas Estacionários

Modelos de séries temporais estacionários são uma classe muito especial chamada de processos estocásticos estacionários, são os quais há estado de equilíbrio estatístico. Um processo estritamente estacionário é aquele em que suas propriedades não se alteram com uma mudança no tempo de observação (Box et. al. 1994).

Morettin (2006) apresenta a definição formal de um processo estocástico estacionário:

$Z = \{Z(t), t \in \tau\}$ é estritamente estacionário se todas as distribuições finito-dimensionais permanecem as mesmas sob translações no tempo, ou seja,

$$F(Z_1, \dots, Z_n; t_1 + \tau, \dots, t_n + \tau) = F(Z_1, \dots, Z_n; t_1, \dots, t_n)$$

Para quaisquer $t_1, \dots, t_n, \tau \in \tau$

A suposição de sistemas estacionários implica em distribuição de probabilidade da função temporal constante em todos os tempos t , portanto $p(Z_t)$ pode ser escrito como $p(Z)$. Isso leva a uma média também constante para esses casos:

$$\mu = E[Z_t] = \int_{-\infty}^{\infty} zp(z)dz$$

A variância também é constante e dada por:

$$\sigma_z^2 = E[(Z_t - \mu)^2] = \int_{-\infty}^{\infty} (z - \mu)^2 p(z) dz$$

Portanto, a média e a variância podem ser estimados, respectivamente, como a média e a variância da amostra:

$$\bar{Z} = \frac{1}{N} \sum_{t=1}^N Z_t$$

$$\hat{\sigma}_Z^2 = \frac{1}{N} \sum_{t=1}^N (Z_t - \bar{Z})^2$$

Isso simplifica consideravelmente a complexidade do estudo das séries temporais, no entanto, a maior parte das séries temporais encontradas na prática apresentam alguma forma de não estacionariedade. Para esses casos, pode-se utilizar um dos dois seguintes procedimentos:

- i. Suposição de estacionariedade e seguinte transformação dos dados originais, se estes não formam uma série estacionária.
- ii. Aplicação de métodos não estacionários

2.3 Modelos Paramétricos

Os modelos paramétricos, ou seja, no domínio do tempo, mais frequentemente utilizados, segundo Morettin (2006) são: os modelos de erro, – também de chamados de regressão – os modelos autorregressivos e de médias móveis (ARMA), os modelos autorregressivos integrados e de médias móveis (ARIMA), modelos de memória longa (ARFIMA), modelos estruturais e não lineares.

A função $Z(t)$ pode ser escrita, além da forma definida anteriormente, como:

$$Z_t = f(t) + a_t, t = 1, \dots, N$$

Onde $f(t)$ é nomeada sinal e a_t ruído.

2.4 Modelos de erro ou de regressão

Nesse tipo de modelo a função sinal é totalmente determinada (parte determinística) e o ruído é uma sequência aleatória independente do sinal, com média zero e variância constante. De tal forma, os efeitos no tempo não alteram a função ruído.

I. Média constante

$$Z_t = Q + a_t, t = 1, \dots, N$$

Modelo normalmente utilizado para representar o erro de mensuração, nesse caso o Q é constante.

II. Modelo de tendência linear

$$Z_t = \alpha + \beta t + a_t, t = 1, \dots, N$$

Nesse caso $f(t)$ é uma função linear dos parâmetros.

III. Modelo de regressão

$$Z_t = \alpha + \beta x_t + a_t, t = 1, \dots, N$$

Novamente $f(t)$ é uma função linear, nesse caso os parâmetros são encontrados, normalmente, usando o método de mínimos quadrados.

IV. Modelo de curva de crescimento

$$Z_t = \alpha \cdot e^{\beta t + a_t}, t = 1, \dots, N$$

Como fica evidente pela expressão, nesse modelo $f(t)$ não é função linear em relação aos parâmetros.

2.5 Modelos ARIMA (Autorregressivo Integrado de Média Móvel)

Nos modelos anteriormente apresentados existe a hipótese de erros não correlacionados, o que limita a validade do modelo. Em situações as quais são

observados erros autocorrelacionados e que influenciam na evolução do processo, modelos ARIMA, comumente chamados de metodologia de Box-Jenkins, são normalmente utilizados.

A estratégia para construção de tais modelos é baseada em um ciclo iterativo no qual a escolha da estrutura do modelo é baseada nos próprios dados. Os modelos ARIMA resultam da combinação de três componentes também denominados de filtros: AR – Autorregressivo; I – Integração; MA – Média Móveis.

Esse tipo de modelo é usualmente representado por ARIMA(p,d,q), no qual p, d, q representam a ordem do modelo

Os modelos ARIMA podem ser descritos por três classes de processos estocásticos:

- I. Processos lineares estacionários, representados pela forma:

$$Z_t - \mu = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots = \sum_{k=0}^{\infty} \psi_k a_{t-k}, \quad \psi_0 = 1$$

- II. Processos lineares não estacionários homogêneos

Se caracterizam pela generalização dos processos lineares estacionários que supõem que o mecanismo gerador da série produz erros auto correlacionados e que as séries sejam não estacionárias em inclinação ou em nível.

- III. Processos de memória longa

Consistem em processos estacionários que possuem função de autocorrelação com decaimento muito lento. Nesse caso, a análise se faz necessária com uma diferença fracionária.

São quatro os estágios do ciclo iterativo de um modelo ARIMA:

- a) *Especificação* de uma classe geral para análise.
- b) *Identificação* de um modelo orientando-se pela análise de autocorrelações e autocorrelações parciais, entre outros critérios.

Morettin (2004) propõe a sequência seguinte para a determinação dos filtros (p, d, q): Classificar a série quanto a sua estacionariedade, se ela for estacionária o filtro d, ou ordem de integração, é zero. O filtro d, portanto, apresenta o número de diferenciações necessárias para tornar a série estacionária. Para o processo de identificação de estacionariedade, se utiliza a função de autocorrelação FAC e sua representação gráfica (correlograma), ou o teste da raiz unitária, sendo o teste de Dickey-Fuller (DF) o mais comumente utilizado.

- c) *Estimação* dos parâmetros do modelo. Nessa etapa são utilizados métodos de mínimos quadrados, máxima verossimilhança ou mínimos quadrados não-lineares (MA diferente de 0) a depender da série.
- d) *Verificação* do modelo ajustado analisando os resíduos. A fim de verificar se o modelo contém parâmetros em excesso ou menos que o necessário, Fava (2000) sugere:
- Se o valor de um coeficiente estimado, em relação a seu erro-padrão, for pequeno em demasia, isso poderá ser indicativo da sua não significância estatística e provável superespecificação. Se for o coeficiente de maior ordem, deve-se suprimi-lo, se for um de menor ordem, convém analisar sua correlação com os demais coeficientes do modelo, se há alta correlação indica que um dos dois pode ser redundante.
 - Introduzir parâmetros adicionais e analisar sua significância estatística é uma alternativa para encontrar subespecificação.

3 MÉTODOS E TÉCNICAS DE PESQUISA

Quanto à natureza da pesquisa, foi quantitativa pois foram utilizadas capturas de dados de vendas e disponibilidade dos produtos da empresa alvo para a testagem de modelos de previsão de demanda. Isso é justificado conforme Fonseca (2002) *apud* Gerhardt e Silveira (2009), a pesquisa quantitativa tem sua base no positivismo, centrada na objetividade, ela recorre de recursos matemáticos para descrever ou explicar um fenômeno.

Quanto aos objetivos da pesquisa apontou-se a pesquisa explicativa uma vez que houve coleta de dados, tratamento, testagem para modelos de análise de demanda, interpretar seus resultados os comparando com a previsão de demanda estruturada pela organização. Em concordância, Prodanov (2013) que afirma que a necessidade do aprofundamento da realidade, se baseando na manipulação e controle de variáveis a fim de estudar, determinar e explicar um fenômeno alvo é o objetivo primordial da pesquisa explicativa.

Quanto aos métodos da pesquisa o adotado é do tipo experimental. Na pesquisa não se pretendeu modelar matematicamente o comportamento da demanda futura, ela, no entanto, se propos a coletar os dados históricos de vendas de produtos, manipulá-los para que fosse aplicável em modelos teóricos de previsão de demanda, avaliando o comportamento da resposta e conflitando com os resultados em aplicações de diferentes modelos entre si e com a previsão de demanda estabelecida pelo processo atual da organização.

A pesquisa experimental consiste em determinar um objeto de estudo, selecionar as variáveis que seriam capazes de influenciá-lo, definir as formas de controle e de observação dos efeitos que a variável produz (GIL 2008). Para Fiorentini (2009) a pesquisa experimental é caracterizada pela realização de “experimentos” com a finalidade de verificar hipóteses em relação a um fenômeno ou problema podendo ser ou não realizada em laboratório. Manipulando certas variáveis e verificando seus efeitos no comportamento de outras.

Quanto às etapas metodológicas, esse trabalho foi dividido em três etapas. O referencial teórico sobre previsão de demanda foi descrito na primeira

etapa apresentando os principais modelos de previsão de demanda, abordando aqueles que mais são utilizados no setor agroquímico e, por consequência, a definição dos principais modelos que se pretende testar. Enquanto, na segunda etapa, foi desenvolvida a coleta e tratamento dos dados. Já na terceira etapa, os dados tratados foram utilizados nos modelos de previsão de demanda e houve a análise dos resultados com a comparação entre os modelos.

4 TRABALHOS CORRELATOS

A aplicação de séries temporais na previsão de demanda é tema recorrente na literatura. Nesse tópico serão citados dois trabalhos os quais se utilizaram do método de *Box-Jenkins* para a previsão de pontos futuros de uma série temporal. Com isso pretende-se, além de apresentar um comparativo dos resultados entre os trabalhos, apontar alguns estudos de casos a fim de ilustrar sua vasta aplicação tanto na academia quanto no setor produtivo.

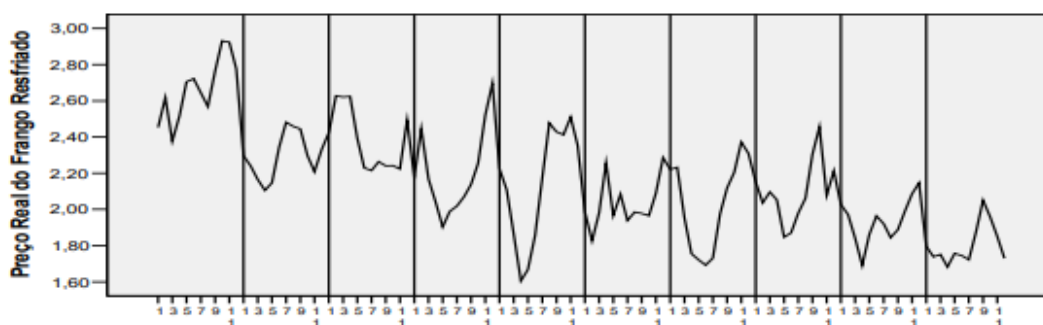
4.1 Aplicação de Modelo ARIMA para previsão do preço do frango inteiro resfriado no grande atacado do Estado de São Paulo

Nesse artigo os autores fazem uso do método de Box-Jenkins (ARIMA sem variáveis explicativas – uni variada), os dados históricos são os preços médios, na grande São Paulo, segmentados por meses, de frangos inteiros no intervalo de 1996 a 2005. Os dados foram coletados a partir de relatórios de uma consultoria especializada, nominada JOX Assessoria Agropecuária e normalizados utilizando a correção de preço a partir do IGP-DI (Índice Geral de Preços e Disponibilidade Interna) da Fundação Getúlio Vargas.

Foi utilizado, para a análise e aplicação da modelagem, o *software* estatístico SPSS. Os autores adotaram os três seguintes indicadores para aferição do desempenho do modelo: EPAM (Erro Percentual Absoluto Médio), REQM (Erro Quadrático Médio e U de Theil).

No processo de Identificação os autores mostram a representação gráfica dos dados afim de avaliar a estacionariedade da série. Notadamente, há a tendência de queda ao longo do tempo, sendo essa, uma característica intrínseca de funções não estacionárias.

Figura 2 - Série temporal do preço real do Frango x Tempo



Fonte – Campos (2006)

Os autores, então, aplicam a Função de Autocorrelação (FAC) e pela análise do Correlograma confirmam que a série é de fato não estacionária e apresenta uma sazonalidade. É realizada a diferenciação sequencial e sazonal da série para que seja apontada uma série estacionária, na qual é possível aplicar o método.

Novamente, para o teste de estacionariedade o artigo apresenta o Correlograma, porém, também aplica o teste de Dickey-Fuller e a série se apresenta como estacionária de primeira ordem. Com a análise do Correlograma da série diferenciada sugere-se uma mistura de processo autorregressivo e de médias móveis, portanto, o modelo parte da configuração básica: $ARIMA(1,1,1),(0,1,0)$, os autores variam os parâmetros e a partir dos indicadores definidos apresentam os resultados dos 4 melhores modelos testados:

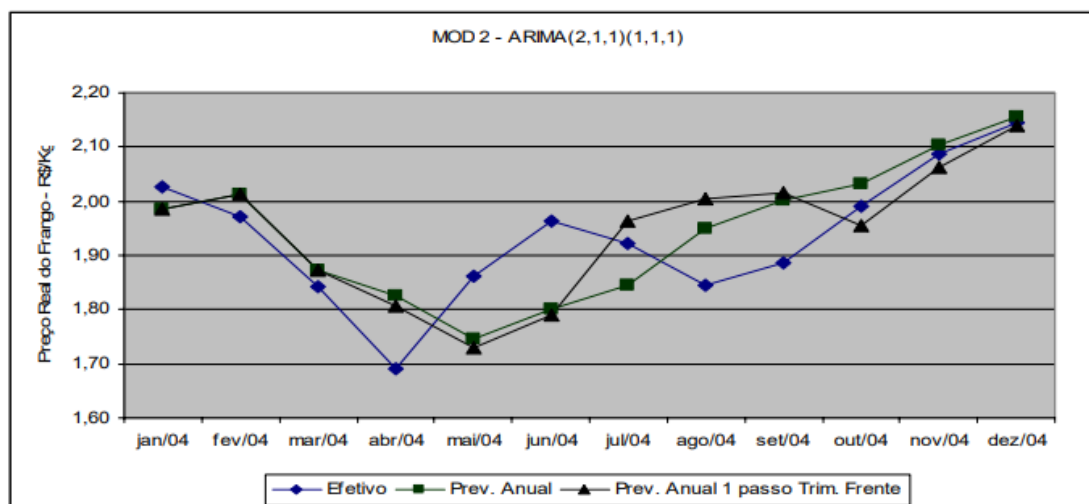
Figura 3 – Comparativo dos indicadores para os modelos testados

ARIMA	MOD 1 (2,1,1)(0,1,1)	MOD 2 (2,1,1)(1,1,1)	MOD 3 (2,1,2)(0,1,1)	MOD 4 (2,1,2)(1,1,1)
Previsão Anual – 2004				
EPAM	4,13%	3,93%	4,63%	3,99%
REQM	0,0937	0,0884	0,1095	0,0952
U	0,0240	0,0227	0,0278	0,0243
Previsão Anual – 2005				
EPAM	5,62%	5,42%	5,67%	6,01%
REQM	0,1354	0,1299	0,1344	0,1390
U	0,0367	0,0353	0,0365	0,0377

Fonte – Campos (2006)

O artigo apresenta o comparativo do realizado com a previsão encontrada ao se utilizar o segundo modelo (MOD 2) para os anos de 2004 e 2005 considerando como dados históricos até 2003.

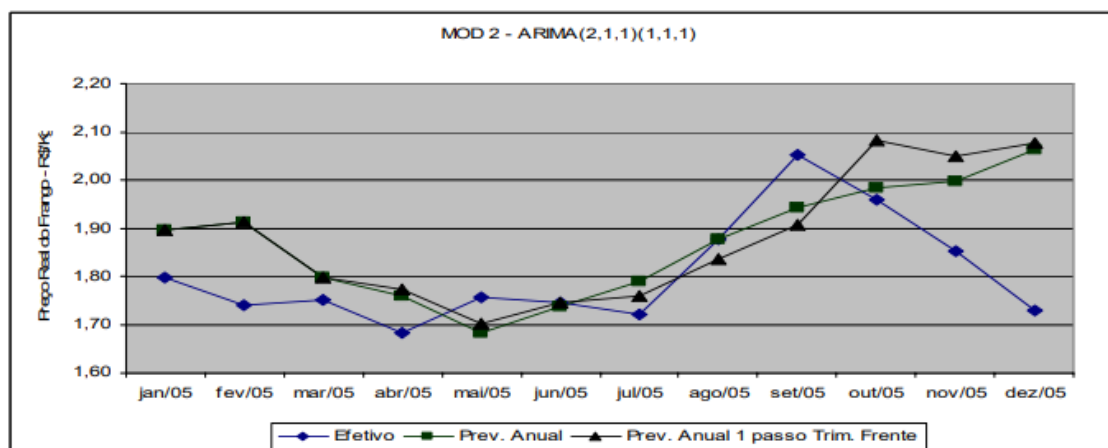
Figura 4 – Resultados considerando o MOD 2 para 2004



Fonte – Campos (2006)

É possível observar boa aderência do modelo para os dois anos, com exceção do último trimestre de 2005 no qual o modelo não previu a abrupta queda:

Figura 5 – Resultados considerando o MOD 2 para 2005



Fonte – Campos (2006)

4.2 Previsão de demanda: Uma aplicação dos modelos de Box-Jenkins na área de assistência técnica de computadores pessoais

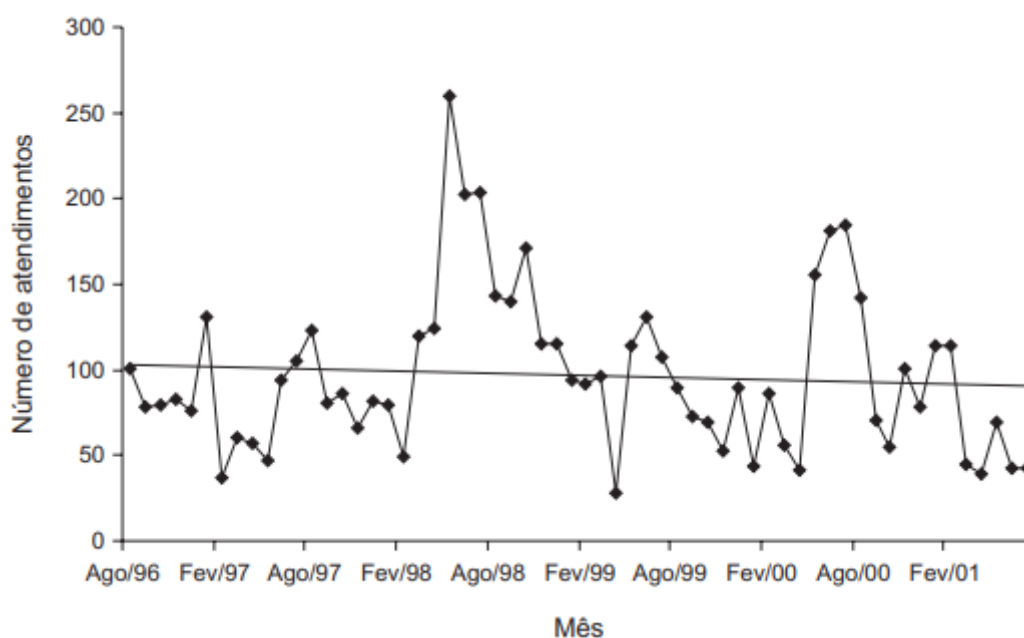
É apresentado, nesse trabalho, uma análise de previsão de demanda para uma empresa que presta serviços de assistência técnica em equipamentos de informática. Para tal, é realizada a segmentação dos clientes da empresa em três categorias: clientes que assinam contratos de manutenção; clientes que

utilizam do serviço no tempo de garantia do equipamento; e os clientes avulsos, de livre demanda. Assume-se que a demanda total da organização é composta pela soma da demanda das três categorias.

Os dados históricos utilizados são de agosto de 1996 a julho de 2001. Foi utilizado o *software* SPSS para manipulação e análise dos dados, bem como para a aplicação do método.

A série temporal para a primeira categoria é assumida como estacionária, visto a sua quase nula tendência:

Figura 6 – Número de atendimentos para a primeira categoria



Fonte – Werner (2003)

Observando as funções de correlações para essa série é possível observar que se trata de uma série com auto-regressão de primeira Ordem - AR(1). Como se trata de uma série estacionária, a ordem de Integração é zero, portando $d = 0$. O modelo AR(1) foi comparado com outros dois modelos apresentando sua eficiência considerando desvio padrão residual e variância residual.

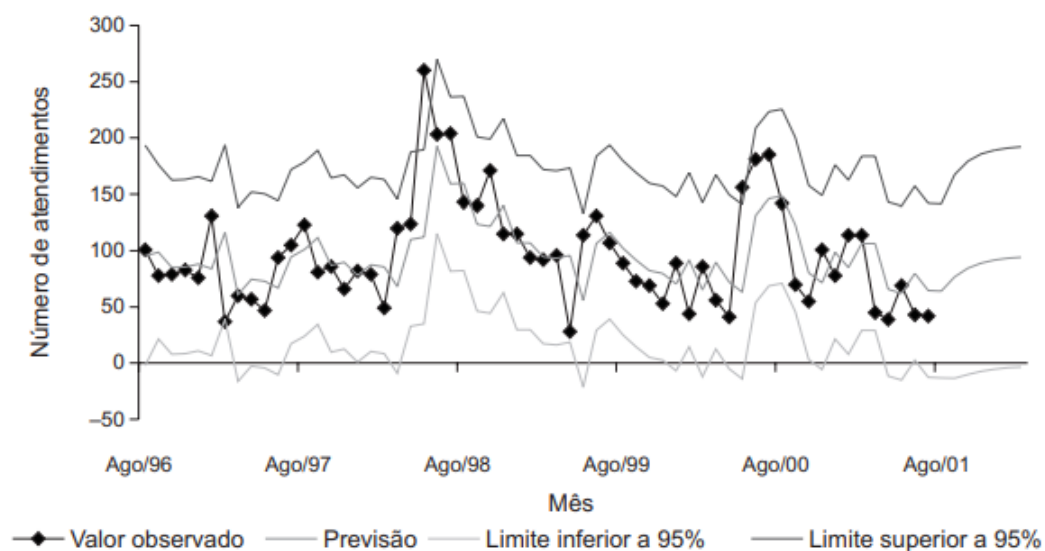
Figura 7 – Comparação entre os modelos testados para a primeira categoria

Modelo	<i>p-value</i> dos coeficientes	Crítério AIC	Crítério SBC	Desvio-padrão residual	Variância residual
AR(1)	$\phi_1: 0,000$ nível do processo: 0,000	610,12 (T = 60)	614,31 (T = 60)	38,29	1.466,46
AR(2)	$\phi_1: 0,000$ $\phi_2: 0,661$ nível do processo: 0,000	611,95 (T = 60)	618,23 (T = 60)	38,56	1.486,84
ARIMA(1,1,0)	$\phi_1: 0,056$ nível do processo: 0,832	609,47 (T = 59)	613,63 (T = 59)	41,62	1.732,37

Fonte – Werner (2003)

Os autores, então, fazem a previsão de demanda até fevereiro de 2002 e mostram na figura a seguir:

Figura 8 – Previsão da demanda para a primeira categoria

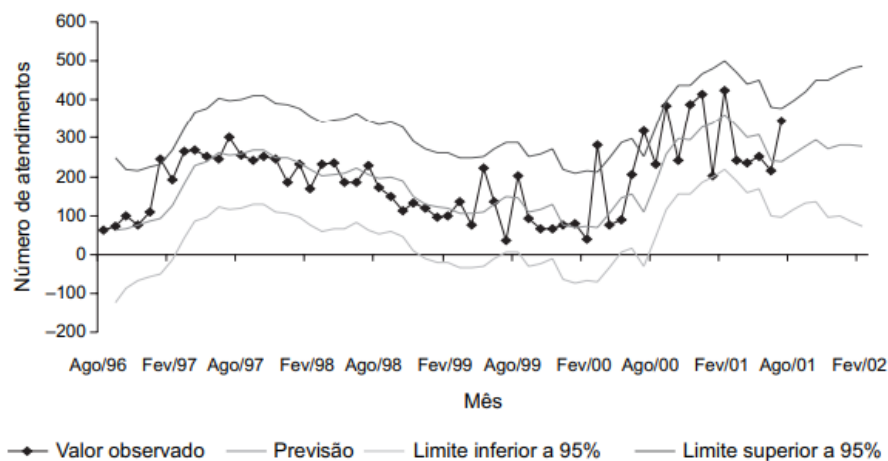


Fonte – Werner (2003)

O mesmo procedimento é realizado para 2 outras categorias de clientes. Para a categoria de garantia foi utilizado o modelo ARIMA (2,1,0). Enquanto para a categoria de clientes avulsos utilizou-se SARIMA (0,1,0)(0,1,1)₁₂.

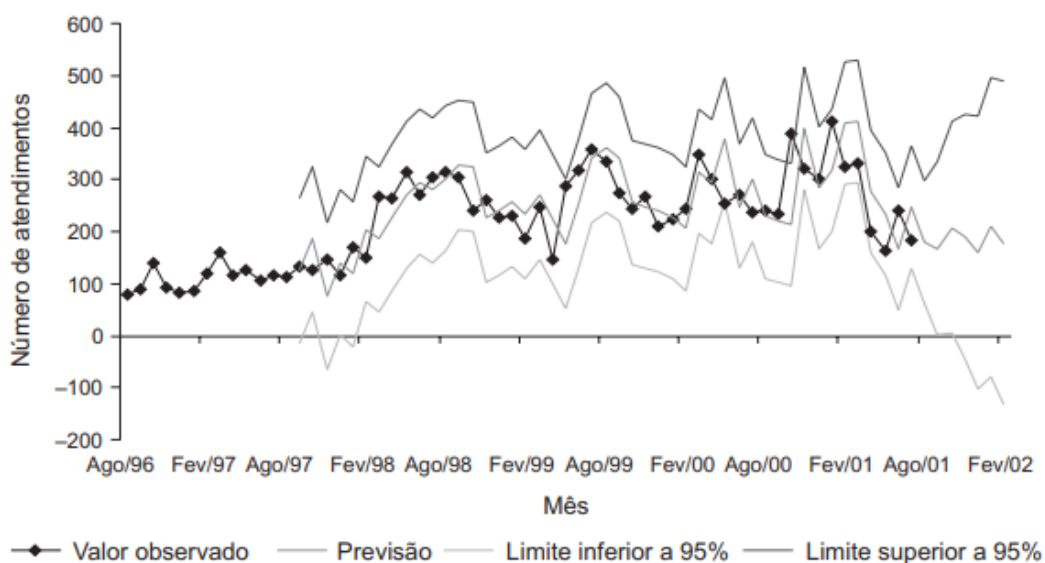
As previsões para as duas categorias, bem como a comparação com os dados históricos, estão dispostas nas figuras a seguir.

Figura 9 – Previsão da demanda para a segunda categoria



Fonte – Werner (2003)

Figura 10 – Previsão da demanda para a terceira categoria



Fonte – Werner (2003)

Por fim, os autores sumarizam as três categorias a fim de apontar a demanda total da empresa no período futuro aos dados, nesse caso, de agosto de 2001 a fevereiro de 2002:

Figura 11 – Demanda total da empresa

Mês	Previsão de demanda			Total
	Contrato	Garantia	Avulsos	
Agosto/2001	64	257	181	502
Setembro/2001	77	277	166	520
Outubro/2001	85	294	205	584
Novembro/2001	89	273	188	550
Dezembro/2001	92	283	157	532
Janeiro/2002	94	283	206	583
Fevereiro/2002	95	279	175	549

Fonte – Werner (2003)

5 AQUISIÇÃO E TRATAMENTO DOS DADOS HISTÓRICOS

A base histórica apresentada nesse trabalho se trata das vendas faturadas por uma multinacional do setor de defensivos agrícolas segmentados por mês de um herbicida comercializado em todo território brasileiro. A identidade e localização da organização não serão apontadas nesse texto para garantir a privacidade e não divulgação de dados sensíveis.

Os dados disponibilizados pela organização estavam em 3 diferentes bancos de dados e não possuíam tabulação uniforme. Portanto, a primeira etapa do desenvolvimento desse trabalho concentrou esforços na consolidação e preparação dos dados. Essa etapa foi realizada em planilhas eletrônicas no *software* Excel.

Por motivos de segurança comercial e privacidade de dados alguns métodos de modificação e camuflagem foram aplicados aos dados originais de tal forma que não se alterassem as principais características da série temporal:

1. Apesar de possuir um período específico e serem em anos consecutivos, nesse trabalho, os anos serão apresentados como: 1, 2, 3, 4 e 5.
2. Os dados de vendas por mês foram multiplicados por um fator k constante.

Os dados encontrados estão dispostos na tabela 3:

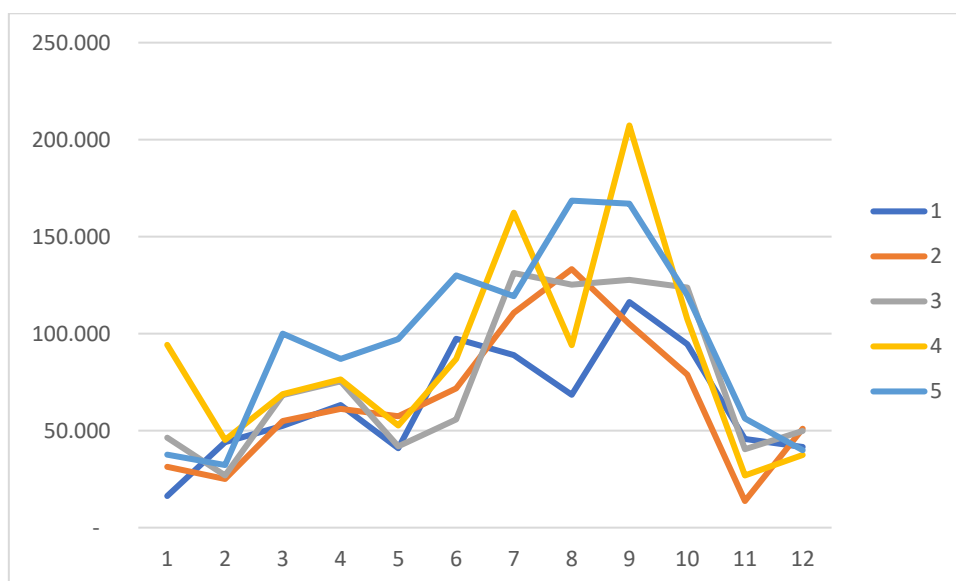
Figura 12 – Resultado de vendas segmentado para o herbicida

Ano Mês	1	2	3	4	5	Total
1	16.297	31.300	46.302	94.151	37.653	225.703
2	44.098	25.034	26.897	45.291	32.261	173.581
3	52.435	54.958	68.260	68.888	100.044	344.585
4	63.217	61.145	75.456	76.356	86.894	363.069
5	40.929	57.372	41.953	52.711	97.191	290.156
6	97.424	71.768	55.698	86.907	130.074	441.871
7	88.872	110.802	131.238	162.366	119.325	612.602
8	68.425	133.253	125.228	93.989	168.567	589.462
9	116.296	104.950	127.773	207.325	167.077	723.422
10	94.481	78.868	123.789	107.476	119.866	524.481
11	45.660	13.711	40.354	26.891	56.311	182.926
12	41.552	51.004	49.838	37.439	39.894	219.726
Total Geral	769.685	794.166	912.785	1.059.789	1.155.158	4.691.583

Fonte – Autor

Na sequência, é apresentado o gráfico 1 dividindo a série para que seja possível perceber a diferença nos valores de vendas durante os meses. Na figura 13 também é possível observar a tendência de crescimento nas vendas com o passar dos anos:

Figura 13 – Sobreposição das vendas do herbicida



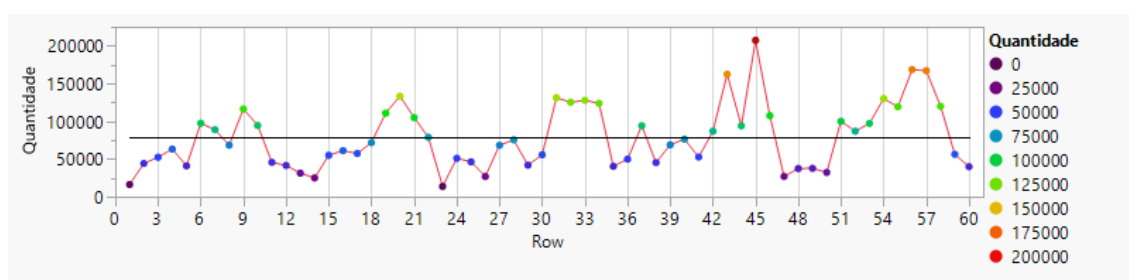
Fonte – Autor

6 APLICAÇÃO DO MÉTODO BOX-JENKINS

6.1 Identificação

A primeira etapa do método ARIMA consiste na caracterização e determinação da estacionariedade da série temporal. Na figura 8 é possível observar, claramente, que não se trata de uma série temporal estacionária, uma vez que há diferenças nas variâncias e a presença de sazonalidade, essa última característica é comum em séries econômicas mensais. Há também, tendência, sendo essa característica pouco visível pela representação gráfica da série.

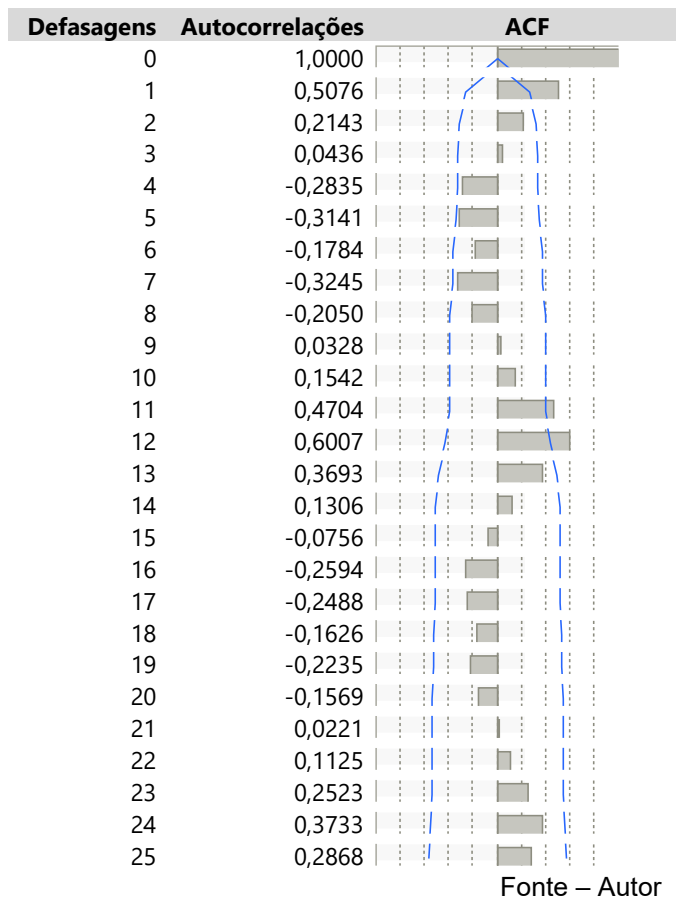
Figura 14 – Série temporal das vendas do herbicida



Fonte – Autor

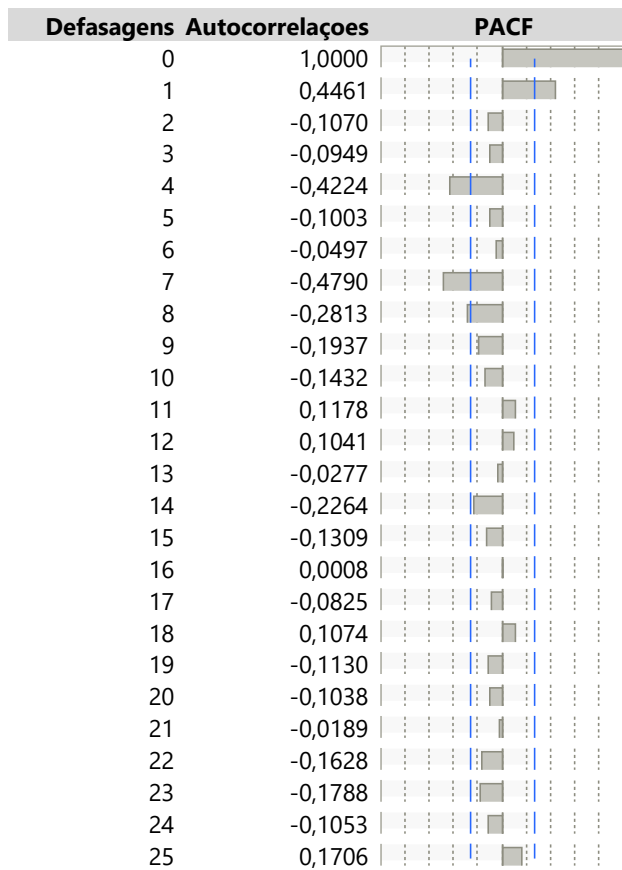
Na sequência, se faz a análise das funções de autocorrelações (ACF) e autocorrelações parciais (PACF) com a finalidade de demonstrar a não estacionariedade sugerida anteriormente pela análise do gráfico de dispersão.

Figura 15 – ACF para a série original



É possível observar, a partir da análise do gráfico anterior, que a ACF apresenta autocorrelações com decaimento próximo ao exponencial, natural de processo autorregressivo. Enquanto, pela análise do gráfico da PACF, gráfico 3, encontra-se os possíveis valores de $p = 1, 4, 7$ ou 8 .

Figura 16 – PACF da série original

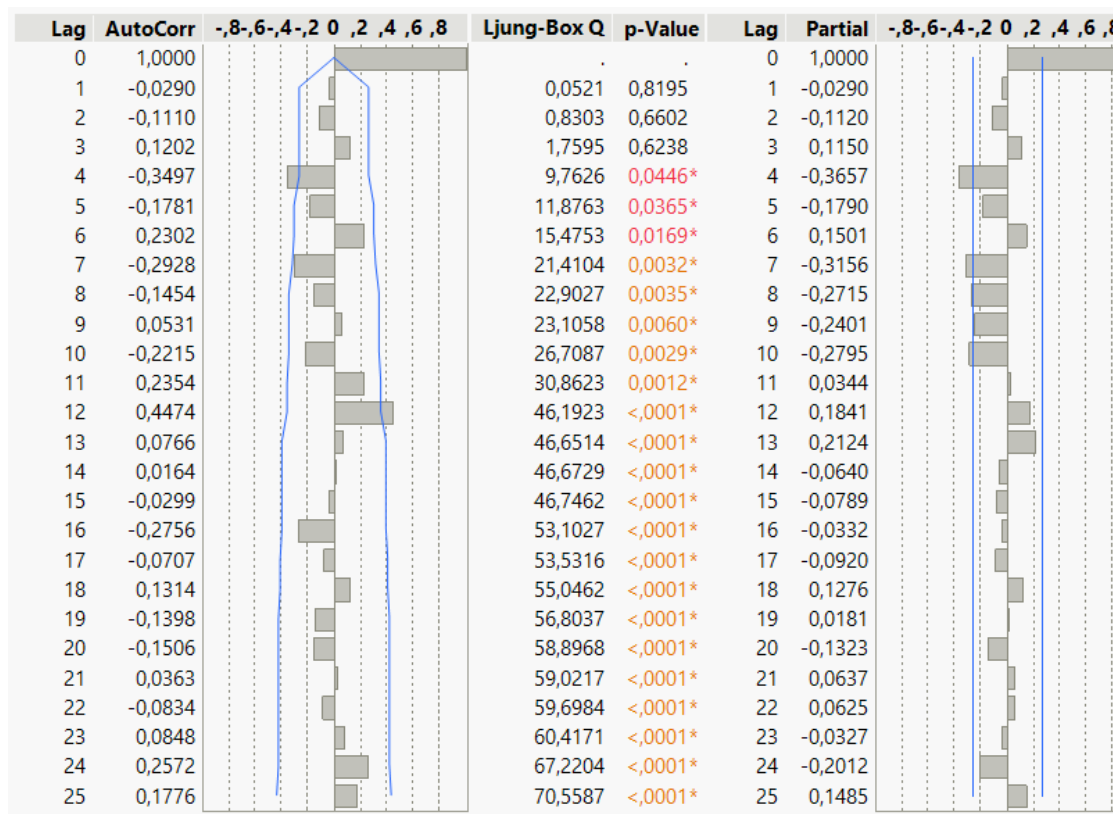


Fonte – Autor

6.2 Estimativa

Para a etapa de estimativa foi utilizado o *software JMP Statiscal Discovery- SAS- versão 16.2.0*. A série foi diferenciada haja visto a sua não estacionariedade e aplicado o modelo $ARIMA(1,1,0)$. Porém, claramente os resíduos desses modelos não são *White noise*.

Figura 17 – ACF e PACF dos resíduos do Modelo ARIMA(1,1,0)



Fonte – Autor, *output* do software JMP

6.3 Verificação

A partir do gráfico 4 e análise da PACF dos resíduos do modelo ARIMA(1,1,0) proposto pela análise das curvas ACF e PACF da série original apresentadas na primeira etapa, fica clara a relevância da consideração da sazonalidade, haja visto o comportamento dos resíduos não se aproximar ao comportamento de ruído branco, apresentando autocorrelação significativa na 10^o defasagem.

Portando, a calibração do modelo se faz necessária para consideração da sazonalidade, caracterizando, portanto, o método como SARIMA. Por se tratar de dados mensais a segmentação do parâmetro de sazonalidade será 12.

Alguns modelos foram testados, na tabela abaixo é possível analisar a comparação entre os que possuíam coeficientes significativos e os critérios utilizados para analisá-los.

Figura 18 – Resultado de vendas segmentado para o herbicida

Modelo	Critério AIC	Critério SBC	Desvio padrão residual	Variância estimada	R ²	MAPE
ARIMA (1,1,0)	1.422	1.426	40.827	1.666.844.236	0,0645	52,57
SARIMA (1,1,0),(0,1,1)12	1.119	1.114	26.719	713.887.106	0,4743	39,62
SARIMA (4,1,0),(0,1,1)12	1.111	1.123	26.774	716.823.494	0,5662	39,57
SARIMA (4,0,2),(1,2,2)12	877	893	24.522	601.305.586	0,3994	41,39

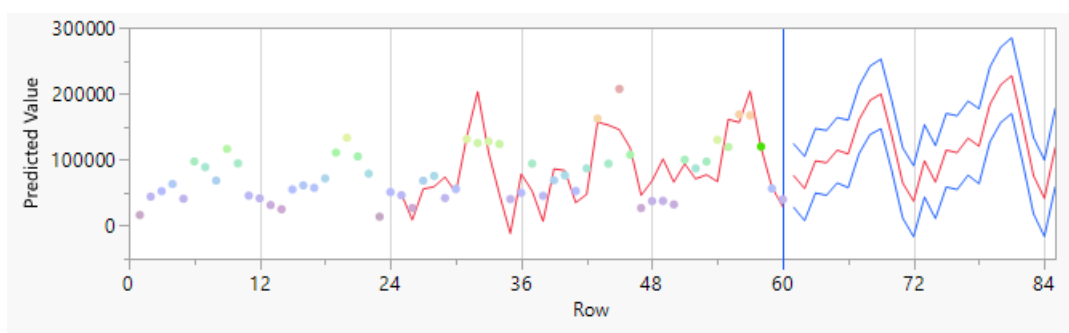
Fonte – Autor

Pelos indicadores apontados o modelo escolhido foi o SARIMA(4,0,2),(1,2,2) realizado na série diferenciada em uma ordem.

6.4 Previsão

Para a previsão foi utilizado o modelo escolhido e calculado em dois próximos ciclos, ou seja, até T=84. O gráfico 5 apresenta um intervalo de confiança de 95% para as previsões.

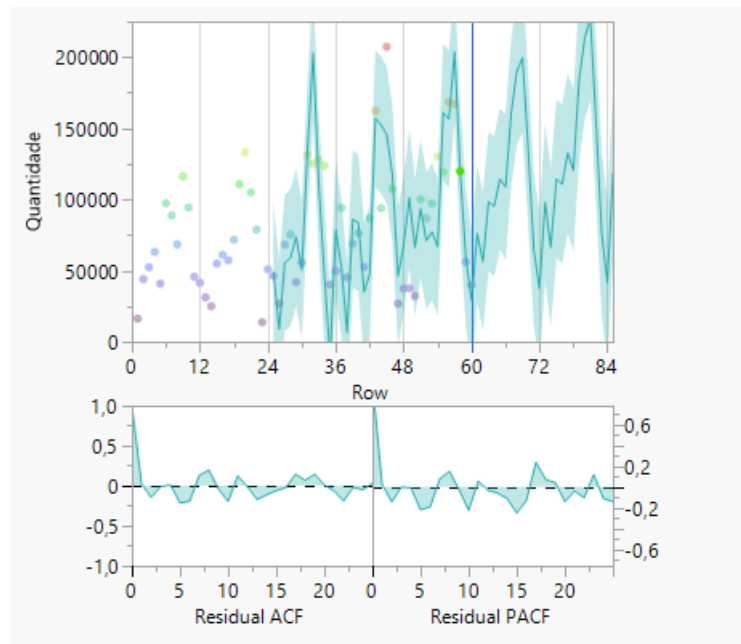
Figura 19 –Previsão da demanda para os próximos 2 ciclos



Fonte – Autor

É possível perceber que o modelo é pouco aderente apesar de ser o que apresentou os melhores resultados nos modelos testados. Ele é significativo, porém apresenta altos resíduos como é possível observar no gráfico 6.

Figura 20 – Comparação do Modelo escolhido com a série e resíduos



Fonte – Autor

É comum, ao se aplicar o método de Box-Jenkins, testar o modelo com a base histórica. Ou seja, não inserir a parte final da série na modelagem para que ela seja comparada com o modelo escolhido e seja avaliado o erro encontrado entre o modelo e o realizado. Essa prática não foi utilizada no presente trabalho por dois motivos principais: a série disponibilizada era relativamente pequena com apenas 60 pontos; importantes *outliers* estavam presentes nos últimos pontos da série.

7 CONCLUSÃO

Modelos de previsão de demanda são, atualmente, essenciais para que as organizações consigam planejar as produções, compras, avaliar possíveis investimentos e construir o cenário de vendas para o próximo ciclo comercial. Muitas vezes eles são base para a construção do planejamento estratégico das empresas e influenciam diretamente nos planos de ações comerciais e produtivos dessas.

A complexidade, no entanto, na construção desses modelos está diretamente relacionada com a quantidade de variáveis independentes e de eventos perturbantes (*disturbance*).

No trabalho apresentado foi testado o método de Box-Jenkins para uma série temporal que apresenta as vendas de um herbicida de uma empresa do setor agroquímico. Como é de se esperar, esse tipo de produto apresenta Sazonalidade. Porém, ela é diminuída visto que o produto é utilizado tanto na safra de verão quanto na safra de inverno (popularmente conhecida como safrinha).

A série também apresentou tendência de crescimento, fato que contribuiu para a complexidade da modelagem. Esse fator foi bem apontado e absorvido pelo modelo e método utilizado.

Os resultados da aplicação do método, no entanto, foram aquém do esperado, uma vez que apesar de apresentarem boa adequação à série original, o modelo apresentou significativos resíduos, caracterizando o modelo como pouco representativo para aplicação de previsão.

O método de Box-Jenkins é comumente utilizado para curtos períodos de previsão e requisita quantidades relevantes de base histórica. Isso pode ser um grande desafio na sua utilização no setor agroquímico. Os ativos e formulações estão em constante mudança e a longevidade dos produtos está se reduzindo ano após ano.

Ademais, sugere-se a testagem de outros modelos estocásticos presentes na literatura para os dados coletados afim de se encontrar um modelo

mais bem representativo. Alguns métodos se destacam na literatura, como é o caso das redes neurais e *Machine learning*.

Por fim, apesar dos resultados encontrados não atenderem diretamente às necessidades da organização se considera que o trabalho representa um avanço na construção de ferramentas quantitativas para representação da demanda futura para a empresa.

A construção do processo científico de análise permite ao autor e à empresa que a mesma metodologia seja aplicada para diferentes produtos, diferentes períodos de vendas até para diferentes métodos de previsão de demanda. Portanto, os objetivos definidos no início do presente texto foram alcançados.

REFERÊNCIAS

- BOX, G. E. P., JENKINS, G. M. and Reinsel, G. (1994). Time Series Analysis: Forecasting and Control. Third Edition. Englewood Cliffs: Prentice Hall.
- CAMPOS, P. A. C., CORDEIRO, A. A. L. Aplicação do Modelo ARIMA para a previsão do preço do frango inteiro resfriado no grande atacado do estado de São Paulo. Belo Horizonte, 2006
- FAVA, V. L., Metodologia de Box-Jenkins para modelos univariados. In: VASCONSELOS, M. A.S.; ALVES, Denisard (Org.). Manual de Econometria: nível intermediário. São Paulo: Atlas, 2000b. p. 205-231.
- FIorentini, D.; LOrenzato, S. Investigação em educação matemática: percursos teóricos e metodológicos. 2ed. Campinas: Autores Associados, 2009. 240 p.
- FONSECA, J. J. S. Metodologia da pesquisa científica. Fortaleza: UEC, 2002. Apostila
- LEE, C. K.; SONG, H. J.; MJELDE, J. W. (2008). The forecasting of International Expo tourism using quantitative and qualitative techniques. Tourism Management, v.29, p.1084–1098, 2008;5
- MORETTIN, P. A.; TOLOI, C. Análise de séries temporais. Blucher, 2006.
- PRODANOV, C. C. Metodologia do trabalho científico: Métodos e Técnicas da Pesquisa e do Trabalho Acadêmico. Nova Hamburgo, 2013 ebook
- WERNER, L.; RIBEIRO, J. L. D. Previsão de demanda: Uma boa aplicação dos modelos Box-Jenkins na área de assistência técnica de computadores pessoais, v10, n. 1, p. 47-67 Porto Alegre, 2003