

# Aplicação da metodologia Box & Jenkins para previsão de vendas de emulsificante

Eduardo Campana Barbosa<sup>1</sup>

Carlos Henrique Osório Silva<sup>2</sup>

**Resumo:** Utilizou-se a metodologia Box & Jenkins para previsão da demanda de um emulsificante utilizado para panificação. Foram comparados 6 modelos estatísticos segundo os critérios de Informação AIC (*Akaike Information Criterion*), SBC (*Schwarzs Bayesian Criterion*) e pela Variância Residual. Além destes foram verificadas as pressuposições de resíduos com distribuição normal, independentes e homocedásticos. O modelo selecionado foi o SARIMA (2,1,2) x (1,1,0)<sub>12</sub> com o qual foram estimadas as vendas para os 6 primeiros meses de 2013.

**Palavras-chave:** SARIMA, Emulsificante, AIC, SBC.

## 1 Introdução

O constante desenvolvimento tecnológico e as frequentes mudanças de valores e do mercado influenciam para que o ambiente empresarial se torne cada vez mais competitivo. Este cenário obriga as organizações a permanecerem atentas e preparadas para adaptações, buscando flexibilidade e menor probabilidade de risco. Para minimizar estes problemas as empresas utilizam algumas estratégias competitivas, dentre as quais destacamos a previsão de demanda via análise de séries temporais (MAKRIDAKIS; WHEELWRIGHT; HYNDMAN, 1998).

No presente trabalho ajustou-se modelos Box & Jenkins do tipo SARIMA (*Seasonal Auto Regressive Integrated Moving Average*) para previsão da demanda de um emulsificante fornecido por uma empresa do interior de Minas Gerais, com o intuito de estimar as vendas para os 6 primeiros meses de 2013 e conseqüentemente auxiliar na gestão econômica.

## 2 Materiais e Métodos

Foi analisado o volume mensal de vendas de emulsificante dos 4 últimos anos (Janeiro de 2009 a Dezembro de 2012). A metodologia de análise seguiu as 3 etapas usuais: identificação,

---

<sup>1</sup> UFV - Universidade Federal de Viçosa. Email: eduardo.barbosa@ufv.br

<sup>2</sup> UFV - Universidade Federal de Viçosa. Email: chos@ufv.br

estimação e verificação do modelo, conforme em Makridakis *et al.* (1982), Box & Jenkins (1970) e Fava (2000). As análises estatísticas foram implementadas com o pacote *forecast* do R por máxima verossimilhança.

## 2.1 Modelos Box & Jenkins do tipo SARIMA (p,d,q) x (P,D,Q)<sub>s</sub>

Segundo Box & Jenkins (1970) os modelos SARIMA baseiam-se em hipóteses de que os componentes da série temporal possuem distribuição de probabilidade aleatória (processo estocástico). São utilizados quando existe autocorrelação periódica e inferior a um ano, o que caracteriza o efeito sazonal. Os modelos SARIMA contém além dos 3 componentes principais dos modelos ARIMA (p,d,q) os 3 componentes sazonais (P,D,Q)<sub>s</sub>, o que resulta no modelo multiplicativo SARIMA (p,d,q) x (P,D,Q)<sub>s</sub> com a seguinte equação:

$$\phi_p(B)\Phi_p(B^s)W_t = \theta_q(B)\Theta_q(B^s)\varepsilon_t$$

$\Phi_p(B^s)$ ,  $\phi_p(B)$  são os coeficientes auto-regressivos sazonais e não sazonais,  $\Theta_q(B^s)$  e  $\theta_q(B)$  são os coeficientes de média móvel sazonais e não sazonais,  $B$  é o operador de retardo,  $s$  a periodicidade sazonal,  $W_t$  é a série após  $d$  diferenças de *lag* 1 e/ou  $D$  diferenças de *lag*  $s$  e  $\varepsilon_t$  é o erro aleatório associado ao modelo.

## 3 Resultados

Foi constatada a presença dos efeitos de tendência e sazonalidade nos dados, portanto aplicou-se uma diferença de *lag* 1 e uma diferença de *lag* 12 para estabilizar a média e a variância. A Figura 1 apresenta os resultados após a aplicação destas diferenças.

Verifica-se que a série se tornou estacionária, com os *lags* 1 e 3 (ACF) e os *lags* 1, 3 e 8 (PACF) significativos. Foi proposto um modelo SARIMA (2,1,2) x (1,1,0)<sub>12</sub>. Segundo Makridakis *et al.* (1982) a complexidade desta etapa esta relacionada ao grande número de modelos que podem ser adequados para o estudo da série, logo, estes autores sugerem que sejam testados outros modelos inserindo e/ou retirando valores para os componentes "p" e "q", mantendo "P" e "Q" assumindo valores entre 0 e 2. Portanto, como critério de comparação, foram propostos 6 modelos, que podem ser visualizados na Tabela 1.

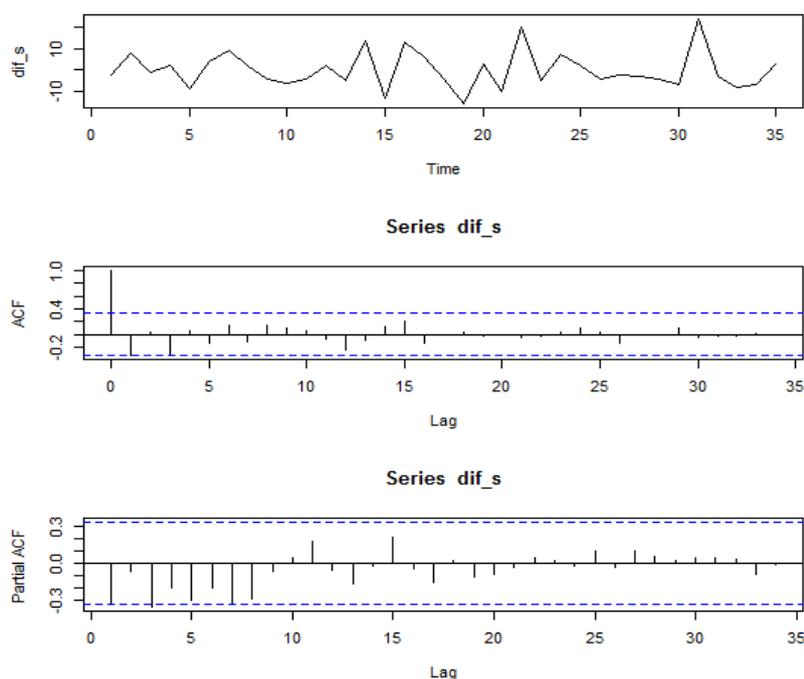


Figura 1 – Série, ACF e PACF após a aplicação das diferenças. Fonte: *Software R*

Tabela 1 – Comparação de Modelos

Modelo SARIMA	AIC <sup>3</sup>	SBC	<i>p-value</i>	<i>p-value</i>	Variância Residual	Conclusão
			Ljung-Box	Shapiro-Wilk		
(2,1,2) x (1,1,0) <sub>12</sub>	239,73	249,06	0,2268	0,1376	31,44	Aprovado
(2,1,1) x (1,1,0) <sub>12</sub>	240,10	247,88	0,0427	0,1235	34,63	Reprovado
(2,1,0) x (1,1,0) <sub>12</sub>	245,24	251,46	0,0733	0,099	43,82	Aprovado
(1,1,0) x (1,1,0) <sub>12</sub>	243,72	248,38	0,1103	0,1628	45,22	Aprovado
(1,1,0) x (1,0,1) <sub>12</sub>	321,19	328,59	0,3500	0,8245	43,27	Aprovado
(1,1,2) x (0,1,1) <sub>12</sub>	241,10	248,88	0,0231	0,2337	36,20	Reprovado

Fonte: Elaborado pelo Autor

Dentre os 6 modelos propostos 2 foram reprovados no teste Ljung-Box (para  $k=15$ ), visto que o *p-value* foi inferior a 0,05, o que leva a rejeição da hipótese nula de Independência dos erros. Dentre os 4 aprovados o melhor modelo foi o SARIMA (2,1,2) x (1,1,0)<sub>12</sub>, obtendo menor AIC, SBC e variância residual, além de resíduos com distribuição normal, independentes e com mesma variância (ruído branco). A Figura 2 corrobora esta conclusão.

<sup>3</sup> Mais detalhes sobre o cálculo dos testes da Tabela 1 em Ehlers (2003) e Pindyck e Rubinfeld (2004)

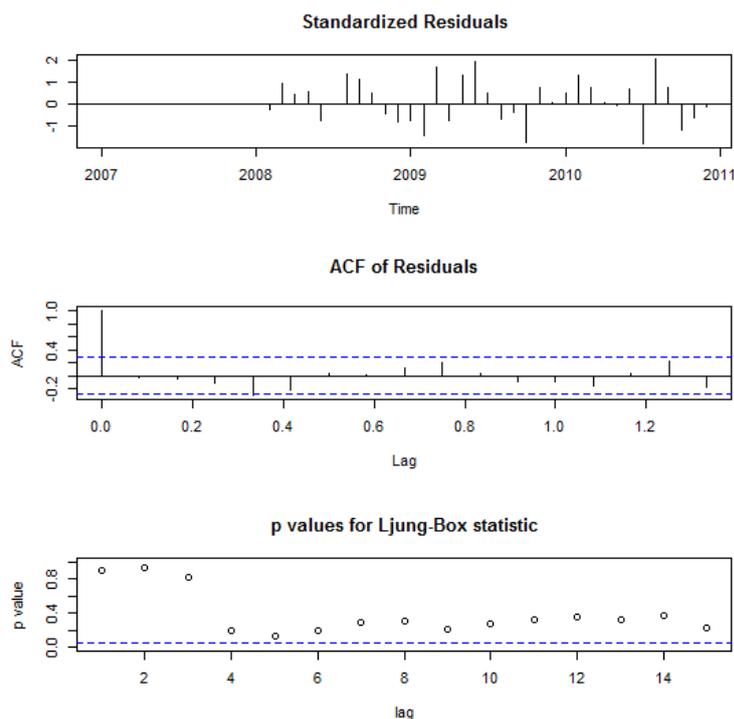


Figura 2 - Pressuposições para o Modelo SARIMA (2,1,2) x (1,1,0)<sub>12</sub>. Fonte: *Software R*

Destaca-se que na metodologia Box-Jenkins é adequado que os coeficientes do modelo sejam estatisticamente diferentes de zero de acordo com o teste t de *student*, porém, neste trabalho essa premissa não foi uma condição fundamental para aprovação de modelos, e sim os testes que avaliam as características de ruído branco dos resíduos. A Tabela 2 apresenta os coeficientes do modelo e suas respectivas estatísticas.

Tabela 2 – Modelo Estimado

	<b>AR(1)</b>	<b>AR(2)</b>	<b>MA(1)</b>	<b>MA(2)</b>	<b>SAR(12)</b>
<b>Coefficientes</b>	-0,7830	0,1927	-0,0938	-0,9061	-0,5163
<b>Erro-Padrão</b>	0,2354	0,1936	0,2076	0,2024	0,1632
<b>Teste t de Student</b>	-3,3263	0,9954	-0,4518	-4,4768	-3,1636
<b>p-value</b>	0,0023	0,3275	0,6546	0,0001	0,0036

Fonte: Elaborado pelo Autor

A Tabela 3 apresenta as estimativas para as vendas de emulsificante nos 6 primeiros meses de 2013, além dos valores reais e resíduos. Foi também analisado o intervalo de predição ao nível de 95% de confiança, a fim de verificar se o mesmo é capaz de conter os valores reais da série.

Tabela 3 – Estimativas das vendas de Emulsificante para os 6 primeiros meses de 2013

Meses	Limite Inferior (5%)	Previsto	Limite Superior (95%)	Valor Real	Resíduos	Dentro do Intervalo de Predição?
Janeiro	25	36	48	19	-17	Não
Fevereiro	13	24	38	37	13	Sim
Março	32	42	54	37	-5	Sim
Abril	26	37	49	40	3	Sim
Mai	35	46	58	33	-13	Não
Junho	28	39	51	43	4	Sim

Fonte: Elaborado pelo autor

## 4 Conclusão

Verificou-se que o SARIMA (2,1,2) x (1,1,0)<sub>12</sub> obteve os melhores resultados e inclusive apresentou resíduos com características de ruído branco, isto é, com distribuição normal, independentes e mesma variância.

## Referências

- [1] BOX, G.E.P., Jenkins, G.M. *Time Series Analysis, Forecasting and Control*. San Francisco, Holden-Day, 1970.
- [2] EHLERS, R. S. *Análise de Séries Temporais*. Desenvolvimento de material didático ou instrucional – Apostila, 2003.
- [3] FAVA, V. L. *Metodologia de Box-Jenkins para modelos univariados*. In: Vasconcellos, M. A. S.; Alves, D. Manual de econometria: nível intermediário. São Paulo: Atlas, 2000.
- [4] MAKRIDAKIS, S. et al. *The accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition*. Journal of Forecasting. v. 1, p. 111-153, 1982.
- [5] MAKRIDAKIS, S.; WHEELWRIGHT, S. & HYNDMAN, R.J. *Forecasting: Methods and Applications*. John Wiley & Sons. 3a Edição. New York, 1998.
- [6] PINDYCK, R. S.; RUBINFELD, D. L. *Econometria: modelos e previsões*. Ed. 4. São Paulo: Atlas, 2004.
- [7] PRIESTLEY, M. B. *Spectral Analysis and Time Series*. London: Academic Press, 1981.