

XII Escola de Modelos de Regressão, Fortaleza-CE, 13-16 Março 2011  
**Análise de modelos de previsão de preços de Uva Itália: uma aplicação do modelo SARIMA**

João Ricardo F. de Lima<sup>1</sup>, Luciano Alves de Jesus Júnior<sup>2</sup>

1 Pesquisador - Embrapa Semiárido. Email: joao.ricardo@cpatsa.embrapa.br

2 Analista – Embrapa Semiárido. Email: luciano.alves@cpatsa.embrapa.br

**Abstract:** The objective of this paper was to evaluate a price forecasting model of “Itália” Grape (*Vitis vinifera* L.) sold in the CEAGESP-São Paulo, using the Box and Jenkins (1994) methodology. The data are monthly, from February 1994 to July 2009. The results shows that the appropriate model for forecasting was a SARIMA (2,1,1)x(0,1,1)<sub>12</sub>, with the dynamic forecasting option.

**Keywords:** Time Series, SARIMA, Price Forecasting.

## 1. INTRODUÇÃO

No setor agropecuário os riscos de investimentos são mais elevados em comparação com outros ramos da economia. Outra questão importante a considerar é que neste setor o produtor incorre em despesas durante todo o ciclo produtivo e, apenas após a comercialização do produto, percebe as receitas. Se o preço recebido na hora da venda for abaixo da expectativa, o produtor pode incorrer em grandes perdas. O bom entendimento do comportamento de uma série de preços é de grande valia por ser uma fonte de informações para, por exemplo, auxiliar os produtores no planejamento de longo prazo de seus negócios; melhor alocar os recursos produtivos, selecionando o que, quanto e como produzir; indicar, no caso dos especuladores que operam nos mercados futuros, a melhor hora de entrar ou sair de um mercado; ou ainda, pelos órgãos governamentais, para definir mercados e produtos a serem enfatizados por políticas agrícolas (LAMOUNIER, 2001). Assim, o objetivo principal deste trabalho é estimar um modelo de previsão de preço de Uva Itália (*Vitis vinifera* L.) comercializada no CEAGESP em São Paulo.

## 2. METODOLOGIA

Podem-se retirar informações importantes sobre o comportamento (componentes) de uma série histórica: tendência, ciclos, sazonalidade e volatilidade (variações irregulares). Inicialmente, para análise da série, verifica-se a estacionariedade, ou seja, o comportamento da média, da variância e da autocovariância ao longo do tempo. O problema relativo às séries não estacionárias relaciona-se com a limitação para se realizar previsões.

Neste trabalho, o teste de raiz unitária utilizado para identificar se uma série é ou não estacionária é o Dickey Fuller-Mínimos Quadrados Generalizados (DF-GLS). É considerado um teste de raiz unitária de segunda geração, que possui maior potência (BAUM, 2001). Existem duas possíveis hipóteses alternativas:  $y_t$  é estacionário sobre uma tendência linear ou  $y_t$  é estacionário sem a tendência linear. Considerando a primeira hipótese alternativa, o teste DF-GLS é realizado inicialmente estimando o intercepto e a tendência via Mínimos Quadrados Generalizados (MQG). A estimação por MQG é feita através da inclusão de novas variáveis  $\tilde{y}_t$ ,  $x_t$  e  $z_t$ , na qual

$$\tilde{y}_1 = y_1$$

$$\tilde{y}_t = y_t - \alpha^* y_{t-1} \quad t=2, \dots, T$$

$$x_1 = 1$$

$$x_t = 1 - \alpha^* \quad t=2, \dots, T$$

$$z_1 = 1$$

$$z_t = t - \alpha^*(t-1)$$

$$\text{e } \alpha^* = 1 - (13.5/T)$$

Faz-se então uma regressão por Mínimos Quadrados Ordinários (MQO):

$$\tilde{y}_t = \delta_0 x_t + \delta_1 z_t + \epsilon_t$$

Sendo que os estimadores  $\hat{\delta}_0$  e  $\hat{\delta}_1$  são usados para se retirar a tendência de  $y_t$ . O passo seguinte é gerar  $y^*$

$$y_t^* = y_t - (\hat{\delta}_0 + \hat{\delta}_1 t)$$

Finalmente, o teste DF-GLS envolve a estimação do ADF, com a substituição das variáveis do GLS modificada por  $y_t^*$  e uma regressão por MQO

$$\Delta y_t^* = \alpha + \beta y_{t-1}^* + \sum_{j=1}^k \zeta \Delta y_{t-j}^* + \epsilon_t$$

(1)

e então testar a hipótese nula de  $\beta = 0$ .

Com relação a segunda hipótese alternativa, o procedimento é semelhante, mas define-se  $\alpha^* = 1 - (7/T)$ , eliminando  $z$  da regressão por MQG, computando  $y_t^* = y_t - \hat{\delta}_0$  e estimando a regressão do ADF sobre as novas variáveis transformadas. A hipótese nula do teste é  $\beta = 0$ .

Existem diversos modelos usados para descrever o processo gerador de uma série temporal, denominados AR (auto regressivos), MA (médias móveis), ARMA (auto regressivos com médias móveis), ARIMA (auto regressivos-integrado-médias móveis) e suas variantes sazonais (SARIMA). Estes visam captar a autocorrelação entre os valores da série temporal e, com base esse comportamento, realizar previsões futuras. De acordo com Lutkepohl e Kratzig (2004), um modelo sazonal multiplicativo geral, denominado SARIMA (p,d,q)x(P,D,Q)<sub>s</sub>, é definido por,

$$\alpha(L)A(L^s)\Delta_s^D \Delta^d y_t = m(L)M(L^s)u_t$$

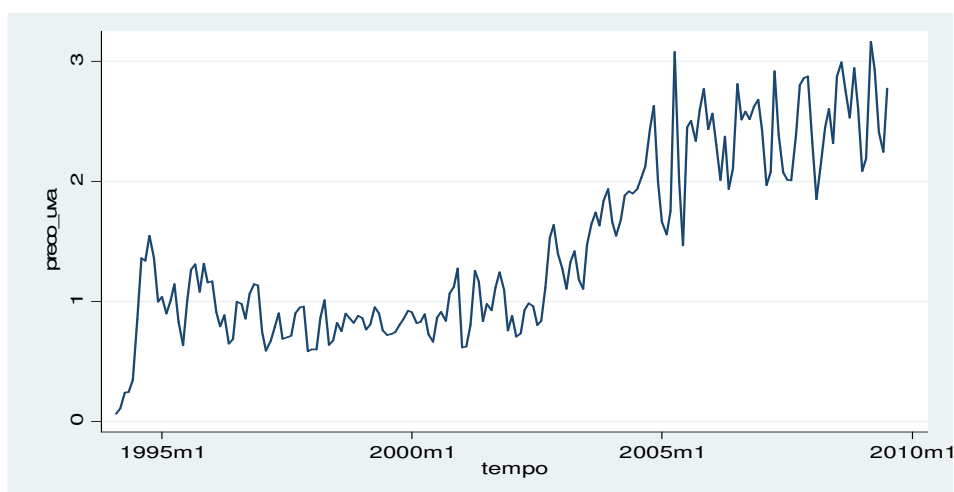
(2)

onde os termos  $d$  e  $D$  representam, respectivamente, as ordens de diferenciação não sazonal ( $\Delta^d y_t = y_t - y_{t-1}$ ) e sazonal ( $\Delta^D y_t = y_t - y_{t-s}$ ) que são necessárias para tornar a série estacionária e  $L$  é o operador de defasagem,  $A$  e  $M$  representam, respectivamente, os parâmetros sazonal autoregressivo (SAR) e sazonal de médias móveis (SMA);  $P$  e  $Q$  referem-se, respectivamente, às ordens autorregressiva e de média móvel sazonais; e  $s$  é o período sazonal. Se o modelo estiver bem ajustado, pode ser utilizado para previsão. A pesquisa utiliza a metodologia de Box-Jenkins (1994), considerando as etapas de identificação, estimação, verificação e previsão (FAVA, 2000).

Foram utilizados dados mensais secundários dos preços reais de Uva Itália comercializada no CEAGESP-SP de fevereiro de 1994 a julho de 2009. Na estimação do modelo serão usados os dados até dezembro de 2008, deixando os demais para comparar com os valores previstos. Os dados foram obtidos em anuários estatísticos da produção agrícola brasileira (Agrianual) publicados pela FNP. O *software* utilizado foi o Stata 11.1.

### 3. RESULTADOS E DISCUSSÃO

A Figura 1 demonstra a evolução dos preços da uva Itália comercializada no CEAGESP-SP entre o período de fevereiro de 1994 a julho de 2009. Pela observação gráfica, pode ser observada uma tendência crescente de preços, no período. Fica evidente que os preços oscilam bastante, com diferentes médias e variâncias entre os períodos e que os preços crescem fortemente a partir de meados do ano de 2003.



Fonte: Agrianual (1999, 2003, 2007 e 2010)

Figura 1 – Evolução do Preço da Uva Itália (R\$/kg) comercializada no CEAGESP-SP entre fevereiro de 1994 e julho de 2009.

O teste DF-GLS, considerando a opção do teste estatístico que retira a tendência linear, indica que a série é estacionária considerando o valor crítico de 1% e com 2 lags definido pelos critérios de Schwarz. A estatística de teste é -3,720 e o valor crítico (1%) é -3,477. A série foi pré-branqueada retirando a tendência fazendo a diferenciação da série.

Foi realizada, então, uma análise sobre o componente ciclo e sazonalidade. A estimação do periodograma não demonstra picos mais significativos nas frequências elevadas, relativo a ciclos de longo prazo. Com relação à Sazonalidade, esta foi analisada pelo método da regressão com uso de *dummies*. Considerando 10% de significância, apenas três meses se apresentaram não significativas, agosto, setembro e novembro, respectivamente. Contudo, considerando o princípio da parcimônia, na modelagem do modelo SARIMA é considerado apenas uma diferenciação sazonal de 12 meses, dado que o volume comercializado em dezembro é duas vezes superior a média dos demais meses do ano. Diversos modelos foram testados e para escolher entre eles, os critérios escolhidos foram os valores dos testes de Akaike e Schwarz. O modelo melhor ajustado está reproduzido na Tabela 1.

Tabela 1 – Resultado do modelo SARIMA estimados para os preços da Uva Itália (R\$/kg) comercializada no CEAGESP-SP entre fevereiro de 1994 e dezembro de 2008.

Modelo	Variável	Coeficientes	Valores de Probabilidade	Akaike	Schwarz
I	AR(1)	0,4184	0,000		
SARIMA	AR(2)	-0,2479	0,005	-9,5599	6,0000
(2,1,1)x(0,1,1) <sub>12</sub>	MA(1)	-0,7149	0,000		
	SAR(1)	-0,8568	0,000		

Fonte: Resultados da Pesquisa.

A análise dos resultados dos coeficientes estimados no modelo indica grande significância estatística. Uma questão com relação à verificação do modelo é observar se os resíduos gerados possuem média zero e se não são correlacionados. Em ambos os casos os resíduos possuem média zero e os erros são ruído branco. Assim, o modelo II foi escolhido para se fazer a previsão dos preços para o período de janeiro a julho de 2009. A Tabela 2 demonstra os resultados considerando dois tipos de previsões: um passo a frente e dinâmica (previsões que se baseiam em valores previstos pelo modelo em períodos anteriores).

Tabela 2 – Resultado das previsões *one-step* e dinâmica com o modelo SARIMA para os preços da Uva Itália (R\$/kg) comercializada no CEAGESP-SP entre janeiro e agosto de 2009.

Período	Preço Obs.	Preço_OS	Lim. Sup*	Lim. Inf*	Preço_Dyn	Lim. Sup*	Lim. Inf*
Janeiro/09	2,0830	2,4564	2,8827	2,0301	2,6697	3,0960	2,2434
Fevereiro/09	2,1880	2,0314	2,4577	1,6051	2,3946	2,8209	1,9683
Março/09	3,1590	2,3894	2,8156	1,9632	2,4788	2,9050	2,0526
Abril/09	2,9250	3,3004	3,7266	2,8742	2,8933	3,3195	2,4671
Mai/09	2,4080	2,4611	2,8873	2,0350	2,6523	3,0785	2,2262
Junho/09	2,2440	2,3362	2,7624	1,9101	2,4973	2,9235	2,0712
Julho/09	2,7700	2,6942	3,1204	2,2681	2,8169	3,2431	2,3908
Agosto/09	-	2,8365	3,2627	2,4104	2,8490	3,2752	2,4229

Fonte: Dados da Pesquisa

\* Os limites inferior e superior do Intervalo com 95% de Confiança.

Considerando a previsão um passo a frente (*One Step-OS*), apenas no mês de março de 2009 o valor observado ficou acima do limite superior do intervalo de confiança da previsão. Já no caso da previsão dinâmica (*dyn*), além do mês de março, o mês de janeiro esteve abaixo do limite inferior do intervalo. Para a previsão em Agosto de 2009, o Preço previsto pela método dinâmico é ligeiramente superior (2,8490) ao *one step* (2,8365).

#### 4. CONCLUSÕES

Com base nos resultados obtidos, o modelo que melhor ajustou a série de preços de Uva Itália comercializada no CEAGESP – SP é o SARIMA (2,1,1)x(0,1,1)<sub>12</sub>, o qual se mostrou adequado para realização da previsão de preços no período, principalmente no caso da previsão dinâmica em que todos os valores observados estão dentro do intervalo de confiança dos valores previsto.

#### REFERÊNCIAS BIBLIOGRÁFICAS

- AGRIANUAL. *Anuário estatístico da produção agrícola*. São Paulo: FNP, 1999, 2003, 2007 e 2010.
- BAUM, C. (2000). *Tests for stationarity of a time series*. Stata Technical Bulletin, StataCorp LP, vol. 10(57). p.36-39.
- BOX, G. E. P., JENKINS, G.M., REINSEL, G.C. (1994). *Time series analysis: forecasting and control*. New Jersey: PRENTICE HALL, 1994. 598p.

XII Escola de Modelos de Regressão, Fortaleza-CE, 13-16 Março 2011

FAVA. V.L (2000). Metodologia de Box-Jenkins para modelos univariados. In: VASCONCELOS. M.A.S.. ALVES. D. *Manual de Econometria - nível intermediário*. São Paulo: Atlas. 308p.

LAMOUNIER, Wagner Moura. *Comportamento dos preços no mercado "spot" de café do Brasil: análise nos domínios do tempo e da frequência*. Viçosa: UFV. 2001. Tese (Doutorado em Economia Aplicada) - Universidade Federal de Viçosa, 2001.

LÜTKEPOHL. H.; KRÄTZIG. M (2004). *Applied time series econometrics*. New York: Cambridge University Press, 2004. 323p.