

## Previsão do preço dos principais grãos produzidos no Rio Grande do Sul

### Forecasting the price of major grains produced in Rio Grande do Sul

Analice Marchezan<sup>I</sup> Adriano Mendonça Souza<sup>II</sup>

#### RESUMO

Objetivou-se realizar previsões para o ano de 2007, referente ao preço das principais culturas das lavouras temporárias desenvolvidas no Estado do Rio Grande do Sul. Os dados-base da quantidade produzida dos principais grãos analisados foram a média anual, de 1995 a 2006, as previsões de preços, de janeiro de 1995 a dezembro de 2006. Para realização das previsões do preço das culturas de soja, arroz e milho no Estado do Rio Grande do Sul foi utilizada Metodologia Box & Jenkins. Para as culturas destacadas, a metodologia captou vários modelos significativos, mas por meios dos Critérios Penalizadores, AIC, SBC e Erro Quadrado Médio, escolheu-se o melhor modelo para cada série, sendo ideal o modelo ARIMA (0,1,2)(0,1,0) para a soja; o modelo SARIMA(1,1,0)(1,0,1) para o arroz e o modelo ARIMA (2,1,1)(0,1,0) para o milho.

**Palavras-chave:** previsão, grãos; preço, arroz; soja, milho.

#### ABSTRACT

The aim of this research was to make predictions for 2007, referring to the price of temporary crops developed in the Rio Grande do Sul state. The data-base of the amount produced of the major grain analyzed were the annual average 1995 to 2006 and price produced, from January 1995 to December 2006. To accomplish the forecasts in relation to prices of soybean, rice and sweetcorn in RS state, it was used the Box & Jenkins methodology. To such cultivations, the methodology showed many significant models, however, through the penalizing criterias, AIC and SBC, in addition to the square average error, it was possible to choose the best model for each serie, being the model ARIMA (0,1,2)(0,1,0) to soybean; the model SARIMA (1,1,0)(1,0,1) to rice and the model ARIMA (2,1,1)(0,1,0) to sweetcorn.

**Key words:** forecast, grains, prices, soybean, rice, sweetcorn.

#### INTRODUÇÃO

A diversidade, a qualidade e a quantidade do desenvolvimento agrícola gaúcho ocorrem por haver, no Estado do Rio Grande do Sul, um clima diversificado e representativos avanços tecnológicos. Aliado a isso tem-se incentivos agrícolas, boas máquinas, irrigação e combate a pragas, a fim de se obter uma boa safra. Mas, para se ter acesso a todos esses implementos, os gaúchos necessitam organizar-se e prever quais serão os seus benefícios, avaliando os lucros e prejuízos.

Dentre as ferramentas que visam auxiliar indiretamente os agricultores estão as análises estatísticas, que mostram uma síntese da realidade da agricultura e apontam aspectos importantes na tomada de decisão dos agricultores, das empresas e do próprio Estado. Tais decisões podem ser embasadas em análises descritivas e previsões futuras. Deve-se destacar que a análise descritiva gera uma síntese dos dados que se deseja analisar, enquanto que uma previsão é uma manifestação relativa a sucessos desconhecidos em um futuro determinado. Para MORETTIN (1989), a previsão não constitui um fim em si, mas um meio de fornecer informações e subsídios para uma conseqüente tomada de decisão, visando a atingir determinados objetivos.

A realização de previsões é destaque atribuído ao modelo de George Box e Gwilyn Jenkins, devido, principalmente, a sua fundamentação teórica,

<sup>I</sup>Instituto Federal Farroupilha (IFFarroupilha), Campus Santa Rosa, Rua Uruguai, 1675, Bairro Central, 98900-000, Santa Rosa, RS, Brasil. E-mail: analicemarchezan@yahoo.com.br. Autor para correspondência.

<sup>II</sup>Departamento de Estatística, Universidade Federal de Santa Maria (UFSM), Santa Maria, RS, Brasil.

sendo a princípio capaz de manipular séries temporais de qualquer natureza. O método de Box & Jenkins consiste na busca de um modelo ARIMA - Auto-Regressivos Integrados de Médias Móveis, que representam o processo estocástico gerador da série temporal.

Portanto, os modelos ARIMA são modelos matemáticos que almejam captar o comportamento da correlação seriada, ou autocorrelação entre os valores da série temporal, e, com base nesse comportamento, realizar previsões futuras.

Diante disso, objetivou-se utilizar a metodologia Box & Jenkins, para realizar previsões, para o ano de 2007, referente ao preço dos principais grãos produzidos no Estado do Rio Grande do Sul no período de Janeiro de 1995 a Dezembro de 2006, referente às lavouras temporárias, que são aquelas culturas que precisam ser replantadas após cada colheita, por possuírem um curto ciclo vegetativo.

Estas previsões visam auxiliar os agricultores gaúchos na tomada de decisões e na criação de políticas de incentivo, bem como, podem servir de suporte para a organização do próprio governo que, ao detectar necessidades no Estado, poderá desenvolver os incentivos necessários para sanar problemas enfrentados pelos empreendedores rurais.

## MATERIAL E MÉTODOS

Nesta pesquisa analisou-se 24 culturas pertencentes a lavouras temporárias, por meio das médias anuais da quantidade produzida, em toneladas, no período de 1995 a 2006. Esses dados foram coletados no site do IBGE – Instituto Brasileiro de Geografia e Estatística. Identificaram-se as três culturas mais produzidas no Rio Grande do Sul, que são: o milho, a soja e o arroz, através da estatística descritiva.

As culturas selecionadas as mais produzidas no Estado e possuem alta demanda, pois são utilizadas como fonte de alimentação humana e animal, bem como fonte de produção de energia, como o biodiesel. Desta forma sabe-se que o preço destas culturas são o que norteiam a sua produção e a busca pela produtividade.

Identificadas as culturas mais produzidas no Estado, utilizaram-se os preços de cotações mensais de janeiro de 1995 a dezembro de 2006, fornecidos pela Emater/RS-ASCAR, para analisar o comportamento destas séries, visto que se almeja estudar o comportamento dos preços dos principais grãos do Estado para assim realizar as previsões para o ano de 2007. A análise do comportamento bem como os valores previstos foram determinados por meio da Metodologia Box & Jenkins.

## Metodologia Box & Jenkins

A previsão é uma das principais razões da popularidade desta metodologia e, em muitos casos, as previsões, principalmente de curto prazo, são melhores que as obtidas com base nos modelos econométricos tradicionais.

Esta metodologia é composta pelos Modelos ARIMA, que são modelos estatísticos lineares para análise de Séries Temporais. Esses modelos são mais sofisticados por usarem a correlação entre as observações. No modelo ARIMA, a Série Temporal é gerada por um processo estocástico, cuja natureza pode ser representada através de um modelo. Segundo MORETTIN & TOLOI (2004), SOUZA & CAMARGO (1996), BARROS (2006), uma Série Temporal, também pode ser definida por uma Série Histórica, que é uma seqüência de dados obtidos em intervalos regulares de tempo, durante um período específico, e que apresentam uma dependência seriada. A notação utilizada para representar uma Série Temporal  $Z$ , no momento  $t$  será  $Z_t$ , em que  $t=1,2,K, n$  indica o tamanho da Série.

A notação utilizada para designar o modelo ARIMA é *ARIMA* ( $p, d, q$ ), em que  $p$  é o número de termos auto-regressivos;  $d$ , o número de diferenciações para que a Série torne-se estacionária e  $q$ , o número de termos de médias móveis. Os termos  $p, d$  e  $q$  são todos inteiros maiores ou iguais a zero. São casos particulares: o modelo *ARMA* ( $p, q$ ), o modelo auto-regressivo *AR* ( $p$ ) e o modelo de médias móveis *MA* ( $q$ ), sendo os três modelos utilizados para Séries Temporais estacionárias ( $d=0$ ).

## Modelos estacionários

Interpretam a Série Temporal como uma realização de um vetor aleatório do tamanho das Séries disponíveis, seguindo o princípio da parcimônia e da construção de modelos, através de um ciclo iterativo. Os modelos estacionários são:

Modelos auto-regressivos de ordem  $p$  – *AR* ( $p$ ): corresponde à defasagem da série  $Z_t$ , denotado por *AR* ( $p$ ), em que  $p$  indica a ordem do modelo, isto é, o número de defasagens. A notação é:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t$$

A identificação do modelo adequado começa pelos valores dos parâmetros, ou seja, o modelo será considerado estacionário se  $|\phi_i| < 1$ , satisfazendo a condição da estacionaridade. A condição de estacionaridade do *AR* ( $p$ ) estabelece que todas as raízes da equação devem cair fora do círculo unitário.

Modelo de médias móveis de ordem  $q$  – *MA* ( $q$ ): é a combinação do ruído branco  $a_t$ , do período

atual com os ocorridos anteriormente. A notação do modelo é  $MA(q)$ , na qual  $q$  indica a ordem do modelo. A notação é:  $Z_t = \mu + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$ .

Dessa forma, identificando a condição de estacionariedade, o processo apresenta ordem finita, portanto, será sempre estacionário.

Quanto à condição de invertibilidade, é imposta a restrição de que o parâmetro em módulo é menor que um, ou seja,  $|\theta_i| < 1$ .

Modelo auto-regressivo de médias móveis –  $ARMA(p, q)$ : Esse modelo é a união dos modelos vistos anteriormente, os auto-regressivos e o de médias móveis. Ele é utilizado em Séries encontradas na prática, para as quais não se deseja números muito grandes de parâmetros.

A notação do modelo  $ARMA(p, q)$  é:

$$Z_t = \underbrace{\phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p}}_{AR(p)} + a_t - \underbrace{\theta_1 a_{t-1} - \dots - \theta_q a_{t-q}}_{MA(q)}$$

Percebe-se que as condições de estacionariedade de um processo  $ARMA(p, q)$  são as mesmas que as de um processo  $AR(p)$ , e as condições de invertibilidade são as mesmas que as de um processo  $MA(q)$ .

#### Modelos Não-Estacionários

Séries não-estacionárias são aquelas que não se desenvolvem ao longo do tempo ao redor de uma média constante. Mas, essas Séries podem tornar-se estacionárias, ao recorrer-se ao processo de diferenciação.

Modelo auto-regressivo integrado de médias móveis- $ARIMA(p, d, q)$ : Os modelos estacionários e os não-estacionários diferenciam-se pelo número de diferença necessária para a estacionarização de uma série, representada pela letra  $d$ . A estratégia utilizada para a construção de modelos, pela metodologia Box & Jenkins, é baseada no ciclo iterativo.

Ciclo iterativo é composto pela identificação do modelo, estimação e diagnóstico. Na identificação do modelo analisa-se a estacionariedade da série através da média e variância, em seguida analisa-se as componentes tendência, sazonalidade e ciclo chegando a análise de dados através da função de autocorrelação e função de autocorrelação parcial.

A estimação analisa-se a significância dos parâmetros estimados sejam eles modelos estacionários e modelos não-estacionários.

Após as etapas de identificação e estimação, chega-se a adequação do modelo, através de critérios de seleção do modelo adequado podendo-se realizar as previsões. Caso o modelo se apresentar inadequado retornar a etapa da estimativa de novos parâmetros.

## RESULTADOS E DISCUSSÃO

Analisando as médias das quantidades produzidas das lavouras temporárias, no período de 1995 a 2006 no Estado do Rio Grande do Sul, identificou-se as três culturas que apresentaram o maior produção em toneladas, sendo elas: milho, soja e arroz.

1º Lugar: milho, com produção mensal média de 203.354.148 toneladas;

2º Lugar: soja, com produção mensal média de 174.086.546 toneladas;

3º Lugar: arroz, com produção mensal média de 74.519.863 toneladas.

Nas posições subsequentes observou-se a mandioca (61.460.898 toneladas); trigo (36.753.201 toneladas); batata-inglesa (13.245.288 toneladas) e o feijão (10.725.246 toneladas).

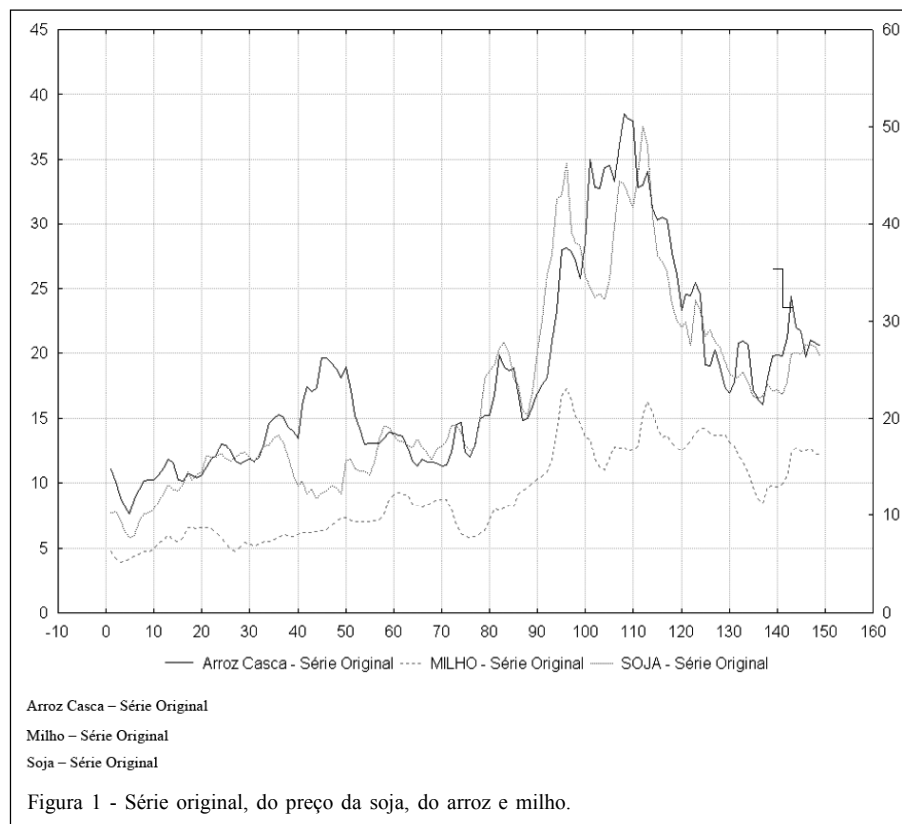
Para analisar o comportamento da Série do preço do milho, da soja e do arroz, de janeiro de 1995 a dezembro de 2006, no Rio Grande do Sul, desenvolveu-se um Gráfico da Série Original, que possibilitou observar o comportamento ao longo do período de análise. Na figura 1, foi possível observar que as Séries não assumem um comportamento estacionário por apresentar uma tendência crescente e picos. Como a análise visual não é comprobatória, recorreu-se a análise das Funções de Autocorrelação de cada Série Original, e aos testes de significância, para se concluir se as Séries apresentam tendência.

Ao observar as funções de autocorrelação da série original referentes ao preço da soja, arroz e milho, respectivamente, é notável a não-estacionariedade da Série, uma vez que a FAC apresenta decaimento muito lento.

Como a série original e as funções de autocorrelação indicaram uma série com tendência, sua significância foi apresentada por meio do Teste do Sinal (Cox-Stuart), para testar as hipóteses:

$H_0$  não existe tendência, ou seja a série é estacionária para  $\alpha=5\%$ , ou seja,  $(P > \alpha)$ .

$H_1$  existe tendência, ou seja a série é não estacionária para  $\alpha=5\%$ , ou seja,  $(P < \alpha)$ .



Como soja / arroz / milho apresentaram  $P=0,000$  ( $P < \alpha$ , ou seja,  $P < 0,05$ ), portanto a série é não estacionária.

Além da componente tendência as séries indicam sazonalidade, a qual pode ser verificada através do Teste não-paramétrico de Kruskal-Wallis, verificando as hipóteses:

$H_0$  não existe sazonalidade determinística para  $\alpha = 5\%$ , ou seja, ( $P > \alpha$ );

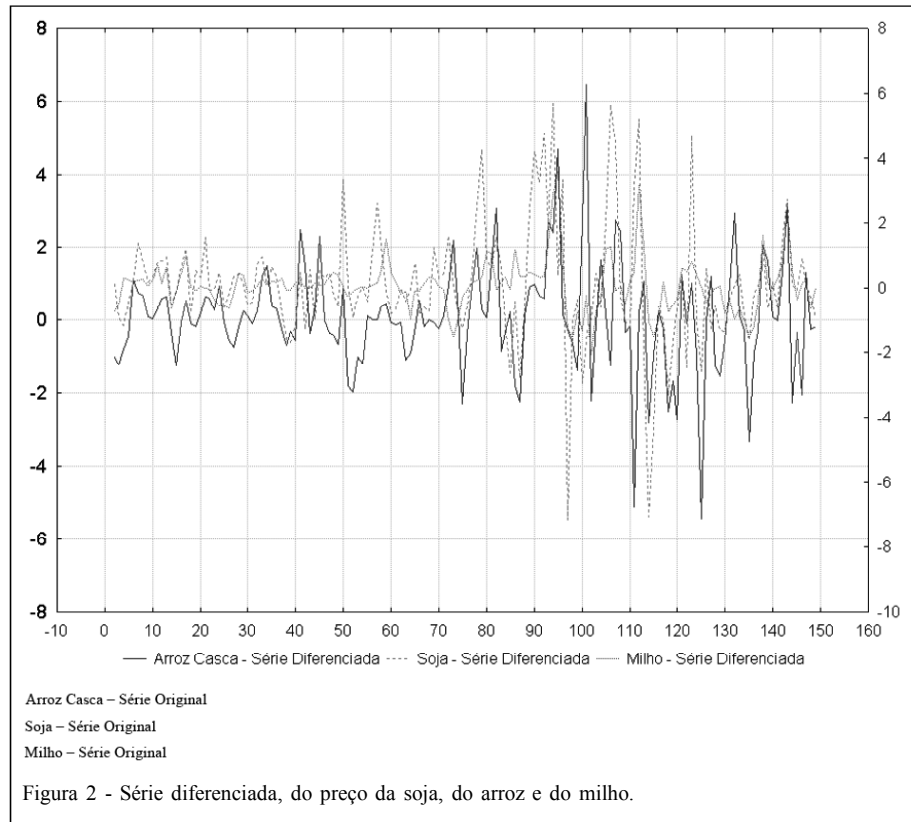
$H_1$  existe sazonalidade determinística para  $\alpha = 5\%$ , o seja, ( $P < \alpha$ ).

Como soja / arroz / milho apresentaram  $P=0,000$  ( $P < \alpha$ , ou seja,  $P < 0,05$ ), portanto a série é sazonal.

Verificado que a não estacionariedade das Séries é significativa, o procedimento adequado foi aplicar sucessivas diferenças nas Séries Originais, até estacionarizá-las. Por meio de uma única diferença, é possível observar que as Séries Diferenciadas, figura 2, apresentaram um comportamento estacionário, ou seja, oscila em torno de uma média, não sendo necessárias mais diferenciações.

Analisadas as séries diferenciadas, o próximo procedimento é identificação dos modelos significativos que representam o comportamento do Preço das culturas analisadas. Através da análise da tabela 1, observou-se os modelos significativos encontrados bem como seus parâmetros. Como regra de decisão para optar pelo melhor modelo, tem-se os Critérios Penalizadores, o AIC e o SBC, que definem os modelos parcimoniosos, por meio dos menores valores encontrados. Ainda como critério complementar, utilizou-se o Erro Quadrado Médio (EQM), que quanto menor o valor encontrado, mais condizente esse modelo está com a realidade da Série. Tais valores encontram-se na tabela 1.

Encontrados os modelos significativos, o próximo procedimento é a verificação das Funções de Autocorrelação (FAC) e Autocorrelação Parcial (FACP) de cada modelo verificando se os resíduos não se apresentam autocorrelacionados. As FAC e FACP satisfizeram esta condição e encontraram-se dentro dos limites de confiabilidades., portanto os modelos representativos de cada série são:



*Soja*: o modelo  $ARIMA(0,1,2)(0,1,0)_{12}$ , representado pela equação:

$$Z_t = a_t - 0,6414a_{t-1} - 0,358a_{t-2}$$

*Arroz*: o modelo  $SARIMA(1,1,0)(1,0,1)_{12}$ , representado pela equação:

$$Z_t = (1+0,19492B^1)(1-B)(1-0,8401B^{12})(1-B)^{12}Z_t = (1-0,97B^{12})a_t$$

*Milho*: o modelo  $ARIMA(2,1,1)(0,1,0)_{12}$ , representado pela equação:

$$Z_t = 0,70394Z_{t-1} + 0,2065Z_{t-2} + a_t - 0,95807a_{t-1}$$

Satisfeitas as condições de estacionariedade das séries, encontrados os modelos significativos e seus respectivos parâmetros, passa-se para a última etapa do Ciclo Iterativo que é a realização das previsões para cada série através dos modelos encontrados, atingindo o objetivo da pesquisa, conforme a tabela 2.

## CONCLUSÃO

A análise descritiva das lavouras temporárias destacou que as maiores produções de grãos no período de 1995 a 2006 no Rio Grande do Sul foram a soja, o arroz e o milho.

A metodologia Box & Jenkins apresentou-se eficiente para representar as séries em estudo, captando os efeitos sazonais ocorridos nos preços e descrevendo o comportamento da série. Salienta-se aqui o binômio soja-milho que são duas culturas que são produzidas alternadamente, por este motivo quando uma cultura aparece em deatque de produção a outra também é cultivada.

Foram encontrados os seguintes modelos para ass dos preços dos principais grãos: para a soja o modelo  $ARIMA(0,1,2)(0,1,0)$  com uma diferença sazonal de tamanho 12; para o arroz foi o modelo sazonal  $SARIMA(1,1,0)(1,0,1)$  e para o milho o modelo  $ARIMA(2,1,1)(0,1,0)$  com uma diferenciação sazonal de ordem 12. Observando que o os valores previstos por tais modelos foram representativos da realidade o que pode ser confirmado pelo Erro Quadrado Médio.

## AGRADECIMENTOS

Adriano Souza agradece a bolsa Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Processo BEX- 1784/09-9.

Tabela 1 - Modelos significativos para a Série do Preço da Soja, do Arroz e do Milho, respectivamente, no Estado do Rio Grande do Sul.

Modelos	Estimativas dos parâmetros do modelo	EQM	AIC	SBC
-----Soja-----				
ARIMA (2,1,0)(0,1,0) <sub>12</sub>	$\phi_1 = 0,3859 ; \phi_2 = 0,1938$	39,91	40,67	46,59
SARIMA (2,1,1)(1,1,0) <sub>12</sub>	$\phi_1 = 0,2816 ; \phi_2 = 0,1960$ $\theta_1 = 0,9999 ; \Phi_1 = 0,2605 \phi$	8,45	26,09	37,92
ARIMA (2,1,2) <sub>12</sub>	$\phi_1 = 0,9379 ; \phi_2 = 0,5042$ $\theta_1 = 0,8540 ; \theta_2 = 0,7648$	4,51	367,13	378,95
SARIMA (2,1,2)(0,0,1) <sub>12</sub>	$\phi_1 = 0,6495 ; \phi_2 = 0,8420$ $\theta_1 = 0,8612 ; \theta_2 = 0,9964$ $\Theta_1 = 0,2828$	11,01	355,71	370,52
SARIMA (0,1,1)(0,1,1) <sub>12</sub>	$\theta_1 = 0,9999 ; \Theta_1 = 0,2678$	3,43	347,29	353,20
SARIMA (0,1,2)(0,1,0) <sub>12</sub>	$\theta_1 = 0,6414 ; \theta_2 = 0,358$	2,12	25,62	31,54
-----Arroz-----				
SARIMA (1,1,0)(1,0,1) <sub>12</sub>	$\phi_1 = 0,19492$ $\Phi_1 = -0,84 ; \Theta_1 = -0,977$	4,22	283,38	292,27
ARIMA (1,1,0)(0,1,0) <sub>12</sub>	$\phi_1 = -0,2767$	24,56	-336,05	-333,10
ARIMA (0,1,1)(0,1,0) <sub>12</sub>	$\theta_1 = 0,90134$	17,25	348,69	351,65
ARIMA (0,1,1)(1,1,1) <sub>12</sub>	$\theta_1 = 0,91879$ $\Phi_1 = -0,76 ; \Theta_1 = -0,909$	5,03	348,38	357,25
-----Milho-----				
ARIMA (1,1,1)(0,1,0) <sub>12</sub>	$\phi_1 = 0,599 ; \theta_1 = 0,969$	8,00	337,23	343,14
SARIMA (2,1,1)(0,1,0) <sub>12</sub>	$\phi_1 = 0,70394 ; \phi_2 = 0,2065$ $\theta_1 = 0,95807$	0,64	323,55	332,42

Tabela 2 - Previsão em reais, do preço da Soja, Arroz e Milho, para o Ano de 2007.

Meses	-----Soja-----			-----Arroz-----			-----Milho-----		
	Previsões	Valores reais	Erro padrão	Previsões	Valores reais	Erro padrão	Previsões	Valores reais	Erro padrão
Jan/07	26,52	26,58	1,826	20,98	21,75	1,52	16,89	16,55	0,69
Fev/07	26,62	27,49	3,081	20,92	19,71	2,59	16,76	16,68	1,38
Mar/07	26,72	27,63	3,956	20,86	21,01	3,39	16,69	16,92	2,00
Abr/07	26,82	27,33	4,671	20,80	20,74	4,09	16,69	16,24	2,52
Mai/07	26,93	26,39	5,291	20,74	20,56	4,73	16,73	16,22	2,98
Jun/07	27,03		5,846	20,68		5,34	16,78		3,40
Jul/07	27,13		6,353	20,62		5,93	16,82		3,79
Ago/07	27,23		6,822	20,57		6,50	16,86		4,17
Set/07	27,33		7,262	20,51		7,06	16,91		4,53
Out/07	27,43		7,677	20,45		7,61	16,95		4,89
Nov/07	27,53		8,071	20,39		8,16	16,99		5,25
Dez/07	27,63		8,447	20,33		8,71	17,03		5,61

**REFERÊNCIAS**

BARROS, M.C.L. **Introdução às séries temporais e aos modelos ARIMA**. Cap.5. Disponível em: <http://www.mbarros.com/sitebuildercontent/sitebuilderfiles/LivroProcessosCap5parte.pdf>. Acesso em: 30 out. 2006.

EMATER/RS - Associação Riograndense de Empreendimentos de Assistência Técnica e Extensão Rural. Disponível em: <http://www.emater.tche.br>. Acesso em: 02 mar. 2007.

IBGE – Instituto Brasileiro de Geografia e Estatística. Disponível em: <http://www.ibge.com.br>. Acesso em: 05 out. 2006.

MORETTIN, P.A.; TOLOI, C.M.C. **Modelos de função de transferência**. São Paulo: ABE, 1989. V.1, 166p.

MORETTIN, P.A.; TOLOI, C.M. **Análise de séries temporais**. São Paulo: E. Blücher, 2004. 535p.

SOUZA, R.C.; CAMARGO, M.E. **Análise e previsão de séries temporais: os modelos ARIMA**. Ijuí, RS: Sedigraf, 1996. 242.