

IDENTIFICANDO E MODELANDO A SAZONALIDADE E O PROCESSO GERADOR DA SÉRIE DE TEMPO DO EMPREGO RURAL NO ESTADO DE SÃO PAULO

Carlos Eduardo Fredo¹
Mario A. Margarido²

RESUMO: Esse trabalho utilizou o método X-12, teste de raiz unitária sazonal e Modelo Auto-regressivo Integrado de Médias Móveis (ARIMA) para identificar e modelar o processo gerador da série temporal do emprego rural no estado de São Paulo no período de janeiro de 1996 a dezembro de 2006. Os resultados mostram que há forte sazonalidade na demanda por mão-de-obra rural, principalmente temporária, no período de abril a agosto em função da época da colheita nas diversas culturas. O teste de raiz unitária sazonal detectou a presença de raiz unitária sazonal, confirmando a sazonalidade na série de emprego rural. O modelo ARIMA captou o dinamismo do emprego rural de trabalhadores temporários, pois foi necessária inclusão de parâmetros de médias móveis de pequena ordem, além de um parâmetro sazonal de média móvel e um parâmetro auto-regressivo o qual captou a contratação de trabalhadores permanentes.

Palavras-chave: Emprego rural, sazonalidade, modelo ARIMA, método X-12.

IDENTIFICATION AND MODELING OF THE SEASONAL PATTERNS AND THE GENERATION PROCESS FROM RURAL LABOR IN THE STATE OF SÃO PAULO

ABSTRACT: This paper used the X-12 method, seasonal unit root and Autoregressive Integrated Moving Average Model to identify and to model the generating process of rural labor in the state of São Paulo in the period from January 1996 to December 2006. The results show that there is strong seasonal demand for occasional rural labor from April to August for several main harvests. The seasonal unit root test detected the presence of seasonal unit root. This result confirms the presence of a seasonal pattern in the rural labor time series. The ARIMA model caught the rural labor dynamism for occasional workers. It was necessary to add three moving average parameters of small orders, beyond a seasonal moving average parameter, and an autoregressive parameter which represents the hiring of permanent workers.

Key-words: Rural labor, seasonality, ARIMA model, X-12 method.

INTORDUÇÃO

O processo de modernização do setor agrícola brasileiro se iniciou na década de 60 e teve como base o denominado modelo neoclássico, o qual baseia-se na formulação pelo

¹ Engenheiro de Computação e Pesquisador Científico de Instituto de Economia Agrícola. E-mail: cfredo@iea.sp.gov.br

² Economista, MS em Economia de Empresas, Dr. em Economia Aplicada e Pesquisador do Instituto de Economia Agrícola. E-mail: mamargarido@iea.sp.gov.br

estado de uma política de estímulo à elevação da produtividade agrícola via choques tecnológicos. Em linhas gerais, conforme MARTINS, MARGARIDO e BUENO (2007, p.31) o “processo de modernização do setor agrícola no Brasil consistiu basicamente em dois pilares: oferta e demanda. O primeiro, do lado da oferta com a concessão de estímulos para a implementação de indústrias visando à produção doméstica, ou então, com subsídios para a importação de fertilizantes, defensivos, tratores, etc., além de investimentos públicos em pesquisa e difusão para utilização desses insumos modernos. Pelo lado da demanda, o principal estímulo para os produtores utilizarem esses insumos, foi à constituição do programa de crédito agrícola subsidiado, o qual perdurou até meados dos anos 80s”³.

O avanço tecnológico observado nas últimas décadas alterou significativamente os processos dos sistemas produtivos no setor agropecuário em especial as relações no mercado de trabalho. Em busca por maior produtividade, agilidade, rapidez e minimização de custos, as máquinas e equipamentos foram inseridos em operações como plantio e colheita em detrimento da força de trabalho humana.

No Estado de São Paulo, várias culturas apresentam-se quase em sua totalidade sob trabalho mecânico como, por exemplo, a soja ou amendoim⁴. Outras estão em processo de mecanização, como a cana-de-açúcar que atualmente possui 40% da área cultivada, já é mecanizada. Nas culturas perenes como café e laranja, a mecanização é menos acentuada, o trabalho ainda é realizado manualmente, com eventual auxílio de equipamentos.

É inevitável que mais e mais pessoas a cada safra agrícola sejam desligadas de suas atividades e uma pequena parcela (mediante qualificação profissional) é que poderão ser absorvidas em outras funções no próprio setor.

O cultivo da cana-de-açúcar, em especial, é a atividade agrícola mais importante para o Estado de São Paulo responsável por 45,5% de todo o valor da produção gerado no setor rural em 2006 (IEA). Atividade que mais emprega mão-de-obra no campo tem sido palco para os debates sobre o mercado de trabalho.

Diferentemente de outros mercados de trabalho, o mercado de trabalho rural é caracterizado por dois aspectos intrínsecos a esse segmento, quais sejam, o elevado nível de incerteza e a sazonalidade. Em relação à incerteza, um evento exógeno climático, como por exemplo, o fenômeno *El Niño*, pode retardar o período de chuvas, e como resultado, afetar a quantidade a ser colhida e também a demanda por mão-de-obra e fixação do nível de salários nesse setor. Quanto a sazonalidade, apesar de seus efeitos declinarem ao longo do tempo, em função, principalmente, da introdução de novas tecnologias, tais como, o processo de irrigação, o qual permite múltiplas safras ao longo do ano, ainda assim, seus efeitos permanecem consideráveis sobre o comportamento da produção agropecuária, alternando-se períodos de safra e entressafra para diversos produtos agropecuários, pois, em inúmeros casos, apesar da tecnologia estar disponível, sua utilização pode não ser economicamente viável em função de seus elevados custos de implantação, os quais, muitas vezes não podem ser repassados para os preços do consumidor final em função da elevada elasticidade-preço desses produtos agropecuários. Portanto, dadas a incerteza que permeia a produção agropecuária e a sazonalidade, o mercado de trabalho rural apresenta comportamento distinto dos demais mercados de trabalho. Conforme apresentado em RAY (1998), no caso do setor rural não há um ponto de equilíbrio entre a oferta e demanda de

³ Detalhes sobre o processo de modernização do setor agrícola no Brasil podem ser encontrados em Pastore (1973), Pastore, Alves e Rizzieri (1974) e Santos (1986).

⁴ Especialistas apontam que o Índice de Mecanização da Soja e do Amendoim estão acima de 90%.

mão-de-obra, mas, sim, um intervalo em que ocorrerá esse equilíbrio em função da incerteza que incide sobre o processo de produção do setor agropecuário.

OBJETIVOS

Para fins de formulação de políticas públicas é importante conhecer os determinantes da demanda de mão-de-obra rural a fim de se minimizar as conseqüências sociais devido à rápida introdução de novas tecnologias poupadoras de mão-de-obra. Dessa maneira, este trabalho objetiva determinar o padrão sazonal da demanda por mão-de-obra rural e seu respectivo processo gerador da série de pessoal admitido no segmento rural para o Estado de São Paulo.

MATERIAL E MÉTODOS

Material

Os dados utilizados foram obtidos no Cadastro Geral de Empregados e Desempregados (CAGED) do Ministério do Trabalho e Emprego disponível em CD-Rom e na Internet⁵. Esta base consolida mensalmente informações sobre número de admissões e desligamentos celetistas em todos os grandes setores econômicos brasileiros, no caso, o setor agropecuário no Estado de São Paulo. Dada a obrigatoriedade dos estabelecimentos prestarem declarações sobre o fluxo de movimentação de mão-de-obra, o CAGED pode ser considerado um censo sobre o mercado de trabalho formal no Brasil. A variável escolhida foi número de admissões no período de Janeiro de 1996 a Dezembro de 2006.

Métodos

Foram utilizados métodos relacionados com modelos de séries de tempo, tais como: método da decomposição, também denominado de método X-12 para a obtenção dos índices sazonais para a variável número de pessoas admitidas, teste de raiz unitária sazonal para tornar essa série estacionária e modelo Auto-regressivo Integrado de Médias Móveis (ARIMA) para determinar o processo gerador dessa série.

Método da decomposição

Conforme ARMSTRONG (2000), em linhas gerais, uma série de tempo caracteriza-se pela coleção de valores que são observados seqüencialmente ao longo do tempo. Desta maneira, dados de séries de tempo descrevem o movimento de determinada variável, evento ou fenômeno em relação a um determinado período de tempo.

Basicamente, uma série de tempo pode ser desagregada em quatro componentes: ciclo, tendência, sazonalidade e componente aleatório. Matematicamente, uma série de tempo pode ser multiplicativa ou aditiva:

$$\hat{Y}_t = \hat{T}_t * \hat{S}_t * \hat{C}_t * \hat{I}_t \quad (1)$$

⁵ Para ter acesso aos dados é necessário fazer parte do Programa de Disseminação das Estatísticas do Trabalho (PDET).

ou

$$\hat{Y}a_t = \hat{T}_t + \hat{S}_t + \hat{C}_t + \hat{I}_t \quad (2)$$

onde: $\hat{Y}m_t$ = série de tempo classificada como multiplicativa; $\hat{Y}a_t$ = série de tempo classificada como aditiva; \hat{T}_t = Tendência; \hat{S}_t = Sazonalidade; \hat{C}_t = Ciclo e \hat{I}_t = Componente Aleatório.

A tendência capta elementos de longo prazo relacionados com a série de tempo. Pode ser determinística, ou seja, é uma função matemática, ou então, estocástica⁶, isto é, resultado de um processo aleatório.

O componente sazonal capta os padrões regulares da série de tempo, tais como mudanças de temperatura, índice pluviométrico, safra ou entressafra de produtos agropecuários, vendas da indústria, vendas do varejo, entre outros.

Os ciclos são caracterizados por longas ondas, mais ou menos regulares, em torno de uma linha de tendência. O interesse pelos ciclos está diretamente relacionado com seus pontos de mudanças, também denominados de pontos de inflexão⁷, duração, frequência, etc.

O componente aleatório capta todos os efeitos que não foram incorporados pela série de tempo via os três componentes anteriormente citados, ou seja, é o resíduo.

De acordo com YAFFEE e MCGEE (2000, p.66): “o programa do método X-12 foi desenvolvido pelo *U.S. Census* e contém algumas inovações em relação ao método X-11 e sua posterior atualização em 1988, o X-11 ARIMA, desenvolvido por E.B. Dagum *et al.* Dagum introduziu o X-11 ARIMA para reduzir a viés de previsão para o final das séries. O novo X-12 contém uma nova sistemática, focada no diagnóstico para avaliar a qualidade de ajustamentos sazonais”.

O método X-12 tem uma variedade de filtros que permitem escolher e extrair a tendência e os padrões sazonais, além de mais um conjunto de filtros assimétricos para serem utilizados no final das séries. Alguns diagnósticos avaliam a estabilidade dos componentes extraídos das séries.

Para calcular a amplitude sazonal em termos percentuais foi calculado o Coeficiente de Amplitude (CA), conforme apresentado FREITAS, *et al.* (1998). Este coeficiente é calculado pela seguinte fórmula:

$$C.A.\% = \frac{(\text{Indice max imo} - \text{Indice min imo})}{(\text{Indice max imo} + \text{Indice min imo})} * 2 * 100 \quad (3)$$

O pacote econométrico utilizado para a obtenção dos índices sazonais foi o *Statistical Analysis Software (SAS)* versão 9.1, conforme descrito em *U.S. Bureau of the Census* (1999).

Modelo auto-regressivo integrado de médias móveis

⁶ Neste caso, a tendência estocástica, em economia, mostra alterações no nível da série de tempo, as quais podem ser resultado de mudanças sociais, tecnológicas, condições de mercado, meio ambiente, entre outros.

⁷ Em economia, o interesse em conhecer adequadamente este componente da série de tempo está relacionado com os ciclos econômicos, visando determinar períodos de prosperidade, recessão, entre outros.

Os modelos Auto-regressivos Integrados de Médias Móveis (ARIMA) têm como base o trabalho de BOX; JENKINS; REINSEL (1994). De acordo com esse método, uma série temporal pode ser explicada em parte por ela mesma⁸ tendo como base seus respectivos valores passados, também denominados de parâmetros auto-regressivos (AR) e/ou próprios erros presente e passados (parâmetros de médias móveis (MA)). Basicamente, o modelo ARIMA tem a seguinte representação:

$$\tilde{y}_t = \frac{\theta(B)\Theta(B^s)}{\phi(B)\Phi(B^s)} a_t \quad (4)$$

onde \tilde{y}_t é a variável y_t diferenciada⁹ e centrada em relação à sua própria média, enquanto que $\theta(B)$ e $\phi(B)$ são polinômios que representam o operador de média móvel de ordem q e o operador auto-regressivo de ordem p respectivamente, e $\Theta(B^s)$ e $\Phi(B^s)$ representam os polinômios de médias móveis sazonal e auto-regressivo sazonal, respectivamente¹⁰. A ordem de integração de uma variável é o número de vezes que essa variável deve ser diferenciada para se tornar estacionária¹¹. O operador de diferença regular (∇^d) é representado da seguinte forma:

$$\nabla X_t = X_t - X_{t-1} \quad (5)$$

onde d representa a ordem da diferença. De acordo com VANDAELE (1983), o operador diferença também pode ser utilizado na remoção da sazonalidade das séries. Nesse caso, tem-se que D -ésima é a diferença sazonal de ordem s e é representada por $\nabla_s^D X_t$, onde:

$$\nabla_s^D X_t = X_t - X_{t-s} \quad (6)$$

“A utilidade da aplicação dos operadores de diferença reside no fato de que eles são capazes de deixar as séries estacionárias, o que significa que esses operadores não somente estabilizam a variância, como também removem a tendência que está por trás das séries originais, tornando-as estáveis” (MARGARIDO, 1998, p. 76).

MILLS (1990), enfatiza que transformação de uma variável utilizando o operador diferença reveste-se de suma importância para quem trabalha com dados econômicos. Uma combinação particularmente relevante é a seguinte:

$$\nabla \log X_t = \log X_t - \log X_{t-1} = \log \frac{X_t}{X_{t-1}} \approx \frac{X_t}{X_{t-1}} - 1 = \frac{X_t - X_{t-1}}{X_{t-1}} \quad (7)$$

Uma vez que, a relação X_t/X_{t-1} é relativamente pequena, isto é, tomando-se a diferença dos logaritmos equivale a utilizar taxas de crescimento.

Ao se trabalhar com modelos ARIMA, o procedimento é composto de três fases: Identificação, Estimação e Verificação (MARGARIDO; KATO; UENO, 1994). Analisando-se com mais detalhes, inicialmente, filtra-se à série original visando descobrir e identificar seu respectivo processo gerador, de modo a torná-la estacionária. A fase de estimação é realizada visualizando-se as respectivas funções de autocorrelação regular e parcial, verifica-se se o processo é auto-regressivo e/ou de médias móveis, e, também, qual(is) é(são) a(s) sua(s)

⁸O modelo ARIMA também é denominado de univariado.

⁹No caso que a variável necessite ser diferenciada para se tornar estacionária, essa diferença pode ser somente regular, somente sazonal, ou então, ambas conjuntamente.

¹⁰Maiores detalhes sobre os modelos ARIMAs podem ser encontrados em BOX; JENKINS; REINSEL (1994), MILLS (1990) e VANDAELE (1983), entre outros.

¹¹Questões relacionadas com a ordem de integração das variáveis e estacionariedade serão abordadas mais adiante.

respectiva(s) ordem(ns). Finalmente, após a construção do(s) filtro(s) e submeter à série a este(s) filtros, analisa-se os resíduos para a verificação da efetividade da filtragem”. A hipótese implícita é de que a passagem da(s) série(s) pelo(s) filtros lineares resultará(ão) em um resíduo denominado de ruído branco, isto é, normal e independentemente distribuído com média zero e variância constante¹².

Para verificar a efetividade da filtragem será utilizado o teste Ljung-Box (1978), o qual é dado pela seguinte fórmula:

$$Q = T(T+2) \sum_{k=1}^s r_k^2 / (T-k) \quad (8)$$

onde: T = número de observações; r_k = autocorrelação na defasagem k . Q tem distribuição χ^2 com k graus de liberdade. A série será ruído branco quando Q assumir valor igual a zero, ou seja, não há presença de autocorrelação nos resíduos da série. O teste Ljung-Box é superior ao tradicional teste de Box-Pierce¹³, pois possui performance superior para casos com pequenas amostras relativamente ao teste Box-Pierce.

Teste de raiz unitária sazonal

Uma série de tempo é estacionária quando preencher três requisitos. Mais precisamente, conforme FREITAS; MARGARIDO; BARBOSA; FRANCA (2001, p.4), um “processo estocástico é dito ser estacionário, ou mais especificamente possuir fraca estacionariedade, quando preencher três requisitos básicos. Em primeiro lugar, sua média tem de ser constante ao longo do tempo ($E(y)$ é constante para todo t). Outra condição é que a sua variância também seja invariante ao longo do tempo ($Var(y)$ é constante para todo t). Finalmente, a sua covariância indica que a autocorrelação entre dois valores de y tomados a partir de dois períodos de tempo distintos depende somente do intervalo do tempo entre esses dois valores e não da sua data ($Cov(y_t, y_{t+s})$ e é constante para todo t que não seja igual a s ”.

De acordo com RAO (1994, p.2) os “*métodos padrões de estimação, que são rotineiramente utilizados em trabalhos de aplicação econométrica, têm como base o pressuposto de que a média e a variância das variáveis são constantes e independentes do tempo. Contudo, a aplicação dos testes de raiz unitária tem mostrado que esses pressupostos não são válidos para um grande número de séries de tempo macroeconômicas*¹⁴. *Variáveis cujas médias e variâncias mudam ao longo do tempo são denominadas não estacionárias ou variáveis com raiz unitária*”. Ainda segundo RAO (1994), a utilização de métodos tradicionais para estimar o relacionamento entre variáveis que tenham raiz unitária, tais como, o método de mínimos quadrados ordinários, fornece inferências viesadas. Mais precisamente, dado que, diante da presença de raiz unitária às médias e variâncias das variáveis não sejam invariantes

¹²VANDAELE (1983) chama a atenção para o fato de que a hipótese de normalidade não é realmente necessária, mas, ao adotar tal hipótese em conjunto com as demais, o resíduo é denominado ruído branco gaussiano.

¹³BOX e PIERCE (1970).

¹⁴“*Isto porque se a hipótese de raiz unitária for verdadeira para uma série, os choques aleatórios que ela sofresse gerariam na mesma um efeito permanente. As flutuações não seriam transitórias, derrubando, por exemplo, as teorias de que os ciclos econômicos seriam flutuações temporárias em torno de uma tendência*” (ALENCAR, 1998, p. 171). O trabalho utilizado como referencial para o fato de que a maioria das séries econômicas possuem raiz unitária é o de NELSON; PLOSSER (1982).

ao longo do tempo, isto implica que as estatísticas obtidas a partir desses métodos tradicionais falham em convergir para seus respectivos valores verdadeiros quando amplia-se o tamanho da amostra. Como resultado, obtém-se relacionamento espúrio entre as variáveis conforme apresentado em GRANGER e NEWBOLD (1974) e PHILLIPS (1986). Assim, a determinação da ordem de integração das variáveis torna-se imprescindível.

Para identificar a ordem de integração das variáveis, geralmente, se utiliza o teste de raiz unitária convencional do tipo Dickey-Fuller Aumentado (DICKEY e FULLER, 1979 e 1981), o qual, basicamente, parte da seguinte auto-regressão:

$$\nabla y_t = \alpha + \beta t + (\rho_1 - 1)y_{t-1} - \sum_{j=1}^{p-1} \rho_{j+1} \nabla y_{t-j} + e_t \quad (9)$$

onde: α é o intercepto; t é a tendência, ∇ é o operador diferença ($\nabla y_t = Y_t - Y_{t-1}$) e ∇_{t-j} é a própria variável dependente diferenciada e defasada, sendo que o número de defasagens é determinado pelo menor valor do Critério de Schwarz (SBC), e visa eliminar a autocorrelação dos resíduos e, finalmente, e_t representa a estrutura de erro, a qual assume-se ser idêntica e independentemente distribuída.

O teste de raiz unitária¹⁵ testa a hipótese nula de presença de raiz unitária ($\rho = 1$) contra a hipótese alternativa de que a série é estacionária ($\rho < 1$), considerando a presença ou não da constante e/ou da tendência (denominadas de testes τ , τ_μ , τ_τ). Os valores críticos para os testes individuais encontram-se em MACKINNON (1991), enquanto que os valores tabelados para os testes conjuntos encontram-se em DICKEY e FULLER (1981).

A sazonalidade é um aspecto relevante para a explicação de determinadas séries econômicas. Os testes de raiz unitária convencionais não podem ser utilizados quando a série econômica apresenta padrão sazonal, dado que, somente a aplicação de uma diferença de ordem um não é suficiente para tornar a série estacionária.

DICKEY; HASZA; FULLER (1984) desenvolveram teste para detectar a presença de raiz unitária sazonal, o qual baseia-se no teste Dickey-Fuller Aumentado (ADF). Nesse caso, a ordem da diferença, é denominada de diferença sazonal e vai depender de sua regularidade. Para séries mensais com sazonalidade anual, a diferença é de ordem 1 ($d=12$); para séries mensais com padrão semestral, a diferença é de ordem 2 ($d=6$); para séries mensais com padrão trimestral, a diferença é de ordem 4 ($d=4$).

O teste Dickey-Hasza-Fuller (DHF) tem como base a seguinte auto-regressão:

$$y_t = \sum_{s=1}^4 \mu_s D_{st} + \beta t + \rho_4 y_{t-4} + u_t \quad (10)$$

onde D_{st} são variáveis binárias (*dummies*) sazonais, t é a tendência determinística. A hipótese nula (H_0) é que $\rho_4 = 1$.

Outra forma de escrever o modelo acima é a seguinte:

$$\nabla^4 y_t = \alpha y_{t-4} + d_1 S_{1t} + d_2 S_{2t} + d_3 S_{3t} + d_4 S_{4t} + \beta t + \sum_{i=1}^k \phi_i \nabla^4 y_{t-i} + u_t \quad (11)$$

sendo que, S_{it} ($i = 1, 2, 3, 4$) são variáveis *dummies* sazonais, t é a tendência determinística, enquanto que $\nabla^4 y_t = y_t - y_{t-4}$ é a diferença sazonal. A hipótese nula é que $\alpha = 0$.

¹⁵Maiores detalhes sobre o procedimento operacional passo a passo para a execução dos testes de raiz unitária dos tipos Dickey-Fuller (DF), Dickey-Fuller Aumentado (ADF) e o teste não paramétrico Phillips-Perron (PP) podem ser encontrados em MARGARIDO; ANEFALOS (1999).

Finalmente, o termo $\sum_{i=1}^k \phi_i \nabla^4 y_{t-i}$, o qual representa a variável diferenciada e defasada, é utilizado para eliminar a autocorrelação dos resíduos. O número de defasagens pode ser determinado pela utilização de algum critério de informação (AKAIKE (AIC) ou SCHWARZ (SBC)), ou então, pelo critério apresentado em NG e PERRON (1995).

ANÁLISE DE RESULTADOS

Em função das características inerentes ao próprio ciclo de produção da agropecuária, a demanda por mão-de-obra no segmento rural concentra-se em determinados períodos do ano. O menor nível de emprego tende a se concentrar entre os meses outubro a março, período que inicia a entressafra, e como consequência, a redução da demanda por mão-de-obra no campo. O mês com menor demanda por mão-de-obra no campo se refere a dezembro, sendo seu índice sazonal igual a 40,749. O maior volume de pessoas empregadas no campo tende a ocorrer de abril a setembro (período de safra). Já, o pico pela mão-de-obra no campo ocorre em maio e corresponde a um índice sazonal igual a 167,61 (Figura 1).

O coeficiente de amplitude é igual a 121,7708, enquanto que a média de todos os índices sazonais é aproximadamente igual a 100 (99,99708). Esse resultado mostra que há expressiva variação da demanda por mão-de-obra rural ao longo do ano, em torno de 21,77%. Apesar das inovações tecnológicas introduzidas no meio rural, como por exemplo, a irrigação, ainda assim, a questão da sazonalidade da produção agrícola continua a exercer forte pressão na condução do processo de produção agropecuária. Também, foram calculadas as médias dos índices sazonais do período de entressafra (74,4675) e safra (125,5267), conforme pode ser visualizado na Figura 2.

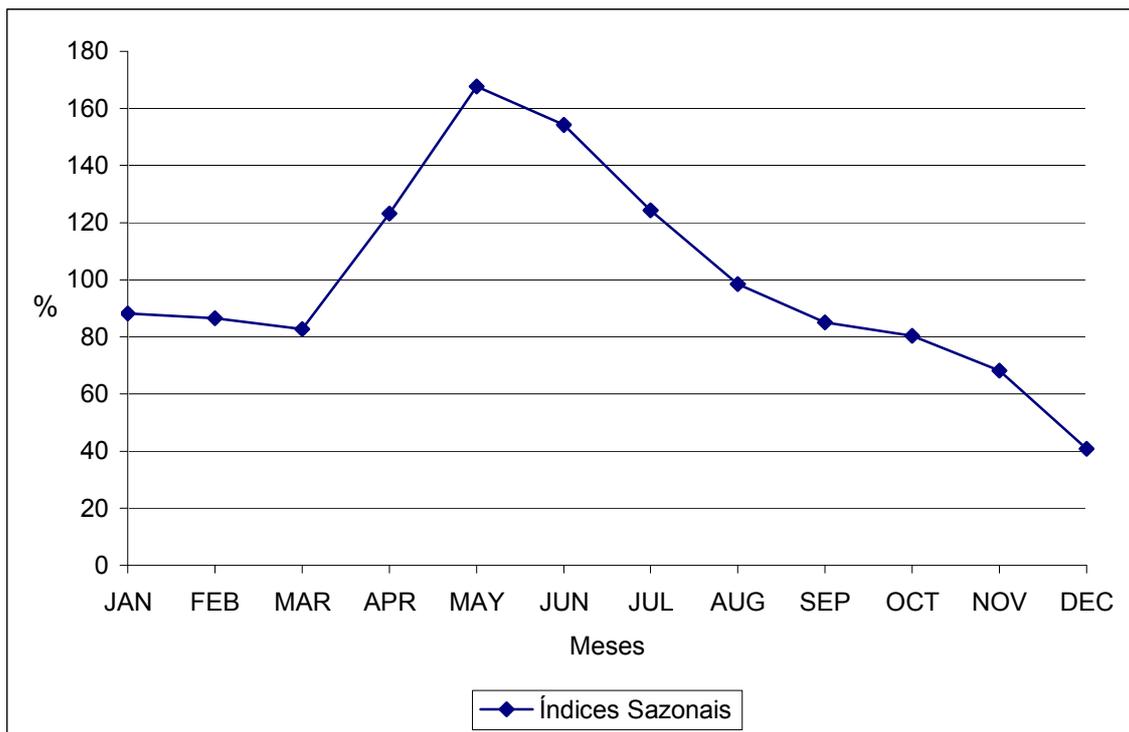


Figura 1 - Índices Sazonais, Pessoal Admitido, 1996-2006.

Fonte: Elaborada pelos autores a partir de dados básicos do CAGED.

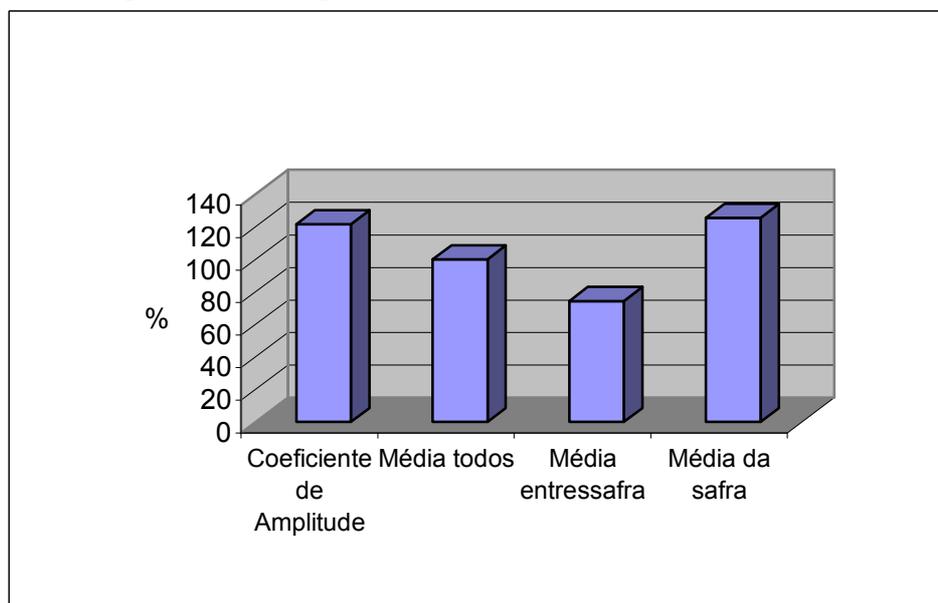


Figura 2 - Coeficiente de Amplitude, Média de Todos Índices Sazonais, Média dos Índices Sazonais da Entressafra e Média dos Índices Sazonais da Safra.

Fonte: Elaborada pelos autores a partir de dados básicos do CAGED.

O primeiro passo consistiu na identificação do modelo ARIMA com a variável número de pessoas admitidas em nível¹⁶. A visualização do respectivo correlograma mostra que o processo dessa variável tem memória longa, ou seja, decai lentamente ao longo do tempo. Isso é um indicativo da presença de raiz unitária. Outro aspecto observado foi que o correlograma apresenta “picos” nas defasagens 12 e 24, indicando a presença de sazonalidade nessa série.

Figura 3.- Função de Autocorrelação da variável Número de Pessoas Admitidas em nível.

Autocorrelations																									
Lag	Covariance	Correlation	-1	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	1	Std Error	
0	0.227694	1.00000																						0	
1	0.170637	0.74941																							0.087039
2	0.125889	0.55288																							0.126827
3	0.092793	0.40753																							0.143933
4	0.047022	0.20651																							0.152424
5	0.0026797	0.01177																							0.154529
6	-0.011052	-0.04854										*													0.154536
7	0.0020293	0.00891																							0.154651
8	0.046395	0.20376																							0.154655
9	0.086030	0.37783																							0.156676
10	0.111519	0.48977																							0.163433
11	0.145435	0.63873																							0.174198
12	0.187085	0.82165																							0.191119
13	0.140469	0.61692																							0.216230
14	0.102029	0.44810																							0.229176
15	0.073095	0.32102																							0.235720
16	0.028618	0.12568																							0.239009
17	-0.013269	-0.05828										*													0.239510
18	-0.026799	-0.11770										**													0.239617
19	-0.017990	-0.07901										**													0.240055
20	0.023583	0.10357										**													0.240251
21	0.060055	0.26375																							0.240589
22	0.083166	0.36525																							0.242770
23	0.109883	0.48259																							0.246898
24	0.143680	0.63102																							0.253944

"." marks two standard errors

Fonte: Elaborada pelos autores a partir de dados básicos do CAGED.

Em função dos fatos de terem sido constatados que a respectiva série em nível não é estacionária e que possui sazonalidade, foi realizado o teste de raiz unitária sazonal proposto por DICKEY, HASZA e FULLER (1984). O teste de raiz unitária sazonal permite testar a presença de raiz unitária sazonal para dois modelos. O primeiro é o modelo que não utiliza parâmetro relacionado com tendência determinística, porém, utiliza uma constante (ou intercepto). Nesse caso, a estatística utilizada é denominada de τ_{μ} . O segundo modelo não utiliza nem tendência determinística nem constante e utiliza-se a estatística τ . Para ambos testes, a hipótese nula é de que há raiz unitária sazonal, enquanto que, a hipótese alternativa pressupõe que não há raiz unitária sazonal.

No caso do modelo somente com constante, a hipótese nula pode ser rejeita, pois há somente 2,51% de probabilidade de rejeitar essa hipótese e ela ser verdadeira. O mesmo resultado não se aplica ao modelo sem constante e sem tendência, dado que, há 99,99% de se cometer o Erro Tipo I, ou seja, há 99,99% de probabilidade de se rejeitar a hipótese nula de que há raiz unitária sazonal e essa hipótese ser verdadeira. Portanto, apesar desse conflito entre os resultados dos testes de raiz unitárias sazonais, ainda assim, pode-se

¹⁶ Nesse ponto é necessário realçar que está sendo utilizado o logaritmo da variável pessoal admitido em todos os estágios relacionados com o modelo ARIMA.

15	0.0042324	0.13147	.	***	.	0.157151
16	-0.0013673	-.04247	.	*	.	0.158065
17	-0.0043580	-.13537	.	***	.	0.158160
18	-0.0046599	-.14475	.	***	.	0.159123
19	-0.0064542	-.20049	.	****	.	0.160216
20	-0.0043673	-.13566	.	***	.	0.162294
21	-0.0031438	-.09765	.	**	.	0.163236
22	0.00077583	0.02410	.	.	.	0.163722
23	-0.0011646	-.03618	.	*	.	0.163752
24	-0.0059363	-.18440	.	***	.	0.163818

"." marks two standard errors

Fonte: Elaborada pelos autores a partir de dados básicos do CAGED.

Sendo assim, uma vez detectada que ainda há raiz unitária, foi aplicada uma diferença de ordem 1, a qual tornou a série estacionária (Figura 5).

Figura 5.- Função de Autocorrelação da variável Número de Pessoas Admitidas diferenciada de ordens 12 e 1.

Autocorrelations																								
Lag	Covariance	Correlation	-1	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	1	Std Error
0	0.028349	1.00000												*****										0
1	-0.0062472	-.22037									****	.		.										0.091670
2	-0.0043573	-.15370									***	.		.										0.096019
3	-0.0028531	-.10064									**	.		.										0.098064
4	-0.0009410	-.03319									*	.		.										0.098928
5	-0.0020006	-.07057									*	.		.										0.099022
6	-0.0023079	-.08141									**	.		.										0.099444
7	0.0065422	0.23078									.		****	*										0.100002
8	0.0046200	0.16297									.		***	.										0.104382
9	-0.0020690	-.07298									.	*	.	.										0.106498
10	-0.0025295	-.08923									.	**	.	.										0.106918
11	0.0037384	0.13187									.		***	.										0.107542
12	-0.0097896	-.34533									*****	.	.	.										0.108892
13	-0.0007714	-.02721									.	*	.	.										0.117736
14	0.0014476	0.05106									.		*	.										0.117789
15	0.0090031	0.31758									.		*****	.										0.117975
16	-0.0026023	-.09179									.	**	.	.										0.124953
17	-0.0027073	-.09550									.	**	.	.										0.125518
18	0.0014880	0.05249									.		*	.										0.126127
19	-0.0038681	-.13645									.	***	.	.										0.126311
20	0.00082842	0.02922									.		*	.										0.127543
21	-0.0026955	-.09508									.	**	.	.										0.127599
22	0.0058875	0.20768									.		****	.										0.128193
23	0.0028431	0.10029									.		**	.										0.130990
24	-0.0050390	-.17775									.	****	.	.										0.131634

"." marks two standard errors

Fonte: Elaborada pelos autores a partir de dados básicos do CAGED.

Uma vez determinada a identificação da ordem de integração da variável, o próximo passo consistiu na estimação do modelo ARIMA.

A elaboração do modelo ARIMA para tornar a série ruído branco (*white noise*) exigiu além de duas diferenças, sendo uma de ordem um e outra de ordem 12, mais quatro parâmetros sendo dois de médias móveis, um de média móvel sazonal e um auto-regressivo. A estimativa do parâmetro de médias móveis de ordem um é igual a 0,50709. Isso quer dizer que, na média, os erros da variável pessoal admitido são corrigidos em torno de 50,71% no período atual relativamente a seu próprio valor do mês anterior. O mesmo tipo de análise vale para o parâmetro de médias móveis de ordem 2. Nesse caso, na média, o desequilíbrio no número de pessoal admitido é corrigido em torno de 24,18% no período atual tendo como base seu respectivo valor passado de dois meses. Aparentemente, esses dois parâmetros de médias móveis de ordem pequena captam o dinamismo (rotatividade) do mercado de trabalho por mão-de-obra rural no curto prazo (Tabela 4).

Dadas às características inerentes da produção agrícola, alterando períodos de safra e entressafra, isso induz a um comportamento expressivamente sazonal da demanda por mão-de-obra. Apesar da necessidade de se aplicar uma diferença sazonal para tornar a série estacionária, ainda assim, esse procedimento não foi suficiente para remover toda sazonalidade da série, pois foi necessária a inclusão de um parâmetro sazonal de ordem um. O valor para a estimativa desse parâmetro sazonal mostra que, na média, os erros da variável pessoal admitido são corrigidos em 78,33% (Tabela 4).

Finalmente, houve a necessidade de adicionar um parâmetro auto-regressivo, cujo valor da estimativa corresponde a 0,2390. Isso quer dizer que o número de pessoas admitidas no período atual (t) e influenciada pelo número de pessoas admitidas de sete meses atrás ($t-7$) em cerca de 23,91% (Tabela 4).

Antes de entender este parâmetro é necessário ter em mente que basicamente, há dois tipos de contratos de trabalho no setor agropecuário. O primeiro envolve trabalhadores que são contratados para a execução de tarefas no curto prazo, como por exemplo, somente no período de colheita. É este grupo que engrossa o volume de contratações realizadas no ano e se concentram em cinco meses (abril a agosto). Conforme RAY (1998), os trabalhadores contratados pelo regime temporal de curto prazo, geralmente, são alocados em tarefas onde a supervisão de suas respectivas funções é relativamente fácil de acompanhar, tais como, colheita e semeadura, etc, e conseqüentemente, seu desligamento do processo produtivo é mais simples relativamente aos empregados permanentes.

A Tabela 3 exemplifica o curto período de tempo em que estes trabalhadores permanecem em suas funções. Do total de demissões realizadas no ano de 2006 no setor, 44,6% permaneceram até seis meses nas atividades agropecuárias.

Tabela 3 – Distribuição de Desligamentos por Faixa de Tempo de Permanência, Setor Agropecuário, Estado de São Paulo, 2006.

Períodos	Número de demissões	% de demissões
De 1,0 a 2,9 meses	80.063	22,8
De 3,0 a 5,9 meses	80.278	22,8
De 6,0 a 11,9 meses	122.528	34,9
De 12,0 a 23,9 meses	24.165	6,9
De 24,0 a 35,9 meses	10.938	3,1
De 36,0 a 59,9 meses	10.195	2,9
De 60,0 a 119,9 meses	7.585	2,2
120 meses ou mais	3.260	0,9
Ignorado*	12.558	3,6
Total	351.570	100,0

* O atributo “Ignorado” refere-se aos postos de trabalho sem informação (ou informação duvidosa) quanto ao tempo de permanência no emprego.

Fonte: CAGED, MTE.

O segundo tipo é composto por aqueles trabalhadores cujos contratos são de longo prazo (trabalhadores permanentes) e alocados em tarefas mais específicas, ou seja, exigem certa especialização e suas respectivas atividades são mais complexas de se monitorar, tais como, aplicações de fertilizantes, defensivos, utilização de irrigação, etc. No entanto, esses empregados podem, eventualmente, participar do processo de colheita, também. Por

exemplo, no cultivo do café, uma colheita seletiva significa num produto com maior qualidade e maior aceitação no mercado.

Quando a demanda por trabalhadores temporários diminui, a partir de setembro e até março, a principal movimentação do trabalho é por parte dos permanentes destacando-se, portanto num período que compreende sete meses do ano, comprovando assim o parâmetro fornecido.

Tabela 4. – Resultados dos Parâmetros Estimados para o Modelo ARIMA, Pessoal Admitido, Janeiro de 1996 a Dezembro de 2006

Parâmetro	Estimativa	Desvio-Padrão	t	Pr > t	Defasagem
MA (1)	0.50709	0.09058	5.60*	<.0001	1
MA (2)	0.24183	0.09179	2.63*	0.0096	2
MA (12)	0.78331	0.07051	11.11*	<.0001	12
AR (7)	0.23906	0.09279	2.58**	0.0112	7

*Significativo em nível de 1,0%, **Significativo em nível de 5,0%.

Fonte: Elaborada pelos autores a partir de dados básicos do CAGED.

Matematicamente, esse modelo ARIMA é escrito da seguinte forma:

$$\nabla_{s=1}^{12} \nabla ladm_t = \frac{(1-0,5070B-0,2418B^2)(1-0,7833B^{12})a_t}{(1-0,2390B)}$$

Compactamente, esse modelo univariado é escrito como sendo um modelo ARIMA (7, 1, 2)(0,1,1)₁₂.

Após a identificação e estimação dos parâmetros a última fase para análise do modelo ARIMA é a sua verificação, isto é, conferir se o processo de filtragem da variável foi efetiva de modo a tornar os resíduos não correlacionados, também denominados de ruído branco (*white noise*). Para verificar se os resíduos são não correlacionados utilizou-se a o teste Ljung-Box. Esse teste tem como hipótese nula se o grupo de autocorrelações é significativamente diferente de zero. Conforme os resultados obtidos, a hipótese nula de que os resíduos são diferentes de zero não pode ser rejeitada para nenhuma das 24 defasagens conforme os respectivos valores das probabilidades. Isso confirma a efetividade da filtragem conforme pode ser visualizado na tabela 4.

Tabela 4. – Resultados do teste Ljung-Box, Pessoal Admitido, Janeiro de 1996 a Dezembro de 2006

Defasagem	χ^2	Graus de Liberdade	P -value	Autocorrelações					
			χ^2						
6	0,32	2	0,8534	0,005	-0,011	0,008	0,033	0,031	0,017
12	3,36	8	0,9101	0,012	0,104	-0,053	-0,068	0,030	0,062
18	12,99	14	0,5272	0,133	-0,016	0,203	-0,096	-0,025	-0,026
24	20,77	20	0,4111	-0,085	-0,019	-0,137	0,051	0,107	-0,110

Fonte: Elaborada pelos autores a partir de dados básicos do CAGED.

CONSIDERAÇÕES FINAIS

A forte sazonalidade observada no mercado de trabalho rural conforme vista nesse trabalho, com altos picos de admissões no período da colheita sofrerá mudanças num futuro próximo por diversas razões. Uma delas é o alto custo de se ter um trabalhador temporário por apenas alguns meses. A fiscalização das delegacias regionais do trabalho tem pressionado os produtores a cada safra para que formalizem o registro em carteira desses trabalhadores. Conforme REZENDE (2006) analisou, para o empregador recai uma série de encargos sociais decorrentes da CLT para ter esse trabalhador com carteira assinada, portanto em termos de se minimizar o custo, é preferível a substituição de oitenta trabalhadores por uma máquina e um tratorista permanente (estimativa para o cultivo de cana) e, diga-se de passagem, mais qualificado que os demais.

Dessa maneira, se a mecanização é uma solução mais atrativa sob o ponto de vista do proprietário rural (agilidade, rapidez e aumento da produtividade) ela se torna mais acelerada com a Lei no. 11.241/2002 cuja finalidade principal é mitigar a emissão de gases de efeito estufa, proibindo gradativamente a queima da cana. Até 2021, em áreas possíveis de serem mecanizadas a queima da cana deverá estar totalmente extinta, e até 2031 para o restante das áreas cultivadas. Nada impede que se aproveite ainda o corte manual que deverá ser realizado com a cana crua (evitando a queima), porém haverá um maior desgaste físico, as condições de trabalho serão piores e cairá a produtividade desses trabalhadores, encarecendo ainda mais os custos do produtor, inviabilizando, portanto, esta alternativa.

Ainda que o artigo dez dessa Lei preveja a criação de mecanismos para realocação da mão-de-obra ou que o impacto gerado pela mecanização não seja tão grande no quesito desemprego, pouco se tem feito para diminuir os impactos sociais desse processo.

Há uma mudança estrutural no mercado de trabalho agropecuário com menor demanda no uso da mão-de-obra no período da colheita, aumento do desemprego, diminuição dos picos de admissão e conseqüente alteração na sazonalidade do emprego. Faz-se necessárias políticas públicas de combate ao desemprego no setor canavieiro, como por exemplo, qualificando os trabalhadores reinserindo-os em outras funções ou em outras culturas.

LITERATURA CITADA

ALENCAR, Leonardo S. de. Raízes unitárias e cointegração: uma introdução. *Boletim do Banco Central do Brasil*, Brasília, p. 171-210, abr. 1998.

ARMSTRONG, J.S. The forecasting dictionary. In: _____ (Ed.). *Principles of forecasting: a handbook for researchers and practitioners*. Norwell, MA: Kluwer Academic Publishers, 2000.

BOX, George E. P.; JENKINS, Gwilym M.; *Time series analysis: forecasting and control*. San Francisco: Holden-Day, 1976. 375p.

BOX, G.E.P. and PIERCE, D.A. Distribution of residual autocorrelations in autoregressive moving average time series models. *Journal of the American Statistical Association*, 65, p,1509-26. 1970.

_____ ; _____; REINSEL, Gregory C. _____. 3rd. New Jersey: Prentice Hall, 1994. 598p.

BRASIL. Ministério do Trabalho e Emprego. CADASTRO GERAL DE EMPREGADOS E DESEMPREGADOS (CAGED). Disponível em <<http://www.mte.gov.br/Empregador/CAGED/Estatísticas>> ou em CD-ROM. Acesso em 21/01/2007.

DICKEY, David A.; HASZA, D.P.; FULLER, W.A. Testing for unit roots in seasonal time series. *Journal of the American Statistical Association*, v.79, p.355-67, 1984.

DICKEY, David A; FULLER, Wayne A. Distribution of the estimators for autoregressive time series with units root. *Journal of the American Statistical Association*, Whashington, v.74, n.366, p.427-31, Jun. 1979.

_____; _____. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, Chicago, v.49, p 1057-72, Jul. 1981.

FREITAS, S.M. de; MARGARIDO, Mario A; BARBOSA, Marisa Z.; FRANCA, Teresinha J.F. Análise da Dinâmica de Transmissão de Preços no Mercado Internacional de Farelo de Soja, 1990-99. *Agricultura em São Paulo*, v.48(1):1-20. 2001.

FREITAS, S.M. de; FERREIRA, C.R.R.P.T.; BARBOSA, M.Z. Oportunidades e entraves à expansão de dendeicultura brasileira. *Agricultura em São Paulo*, São Paulo, v. 45, t.2, p.1-16, abr. 1998.

GRANGER,C.W.J.; NEWBOLD, P. Spurious regressions in econometrics. *Journal of Econometrics*, Nottingham, v.2, p. 111-20, Jul. 1974.

LJUNG, G.M. and BOX, G.E.P. On a measure of lack of fit in time series models. *Biometrika*, 66, 67-72. 1978.

MACKINNON, James G. Critical values for cointegration tests. In: ENGLE, Robert F. ; GRANGER, W. J. *Long-run economic relationships: readings in cointegration*. New York: Oxford University Press, 1991. p.267-76.

MARGARIDO, Mário A.; ANEFALOS, Lilian C. Testes de raiz unitária e o *software* SAS. *Agricultura em São Paulo*, São Paulo, v. 46, t.2, p.19-45, 1999.

MARGARIDO, Mario A. *Transmissão de preços internacionais de suco de laranja para preços ao nível de produtor de laranja no Estado de São Paulo*. São Paulo: IEA. 1998. 127p. (Coleção Estudos Agrícolas, 6/98).

MARGARIDO, Mario A.; KATO, Heitor T.; UENO, Lídia H. Aplicação da metodologia Box-Jenkins na análise da transmissão de preços no mercado de tomate no Estado de São Paulo. *Estudos Econômicos*, São Paulo, v.24, n. 3, p. 405-32, set./dez. 1994.

MILLS, Terence C. *Time series techniques for economists*. New York: Cambridge University, 1990. 377p.

NELSON, Charles R.; PLOSSER, Charles I. Trends and random walks in macroeconomic time series. *Journal of Monetary Economics*, v.10, p.139-62, 1982.

Ng S. PERRON P. Unit root test in ARMA models with data-dependent methods for the selection of the truncation lag. *Journal of the American Statistical Association*, v.90, p.268-281. 1995.

PASTORE, Affonso C. A oferta de produtos agrícolas no Brasil. In: PASTORE, José. *Agricultura e desenvolvimento*. S.n.t., 1973. p.113-49.

PASTORE, Affonso C.; ALVES, E.; RIZZIERI, J. *A inovação induzida e os limites à modernização na agricultura brasileira*. Brasília: EMBRAPA, 1974. (Teoria e Metodologia, v.2).

PHILLIPS, Peter C.B. Understanding spurious regressions in econometrics. *Journal of Econometrics*, North-Holland, v. 33, n.3, p.311-40, Dec. 1986.

RAO, B. Bhaskara. (Ed.). The econometrics of disequilibrium models. In: _____. *Cointegration for the applied economist*. New York: ST. Martin's Press, 1994. p.1-8.

RAY, Debraj. *Development Economics*. United States of America: Princeton University Press, 1998. 848p.

REZENDE, Gervásio Castro de. Políticas trabalhista, fundiária e de crédito agrícola no Brasil. *Rev. Econ. Sociol. Rural*, Jan./Mar. 2006, vol.44, no.1, p.47-78. ISSN 0103-2003.

SANTOS, Robério F. dos. *Presença de vieses e mudança técnica na agricultura brasileira*. São Paulo: IPE/USP. 1986. (Ensaio Econômico, 63).

VANDAELE, Walter. *Applied time series and Box-Jenkins models*. New York: Academic Press, 1983. 417p.

YAFEE, Robert; McGEE, Monnie. *Introduction to time series analysis and forecasting: with applications of SAS and SPSS*. United States of America: Academic Press, 2000. 528p.