

TCC/UNICAMP

B288m

1290003106/IE

nelcio alexandre



1290003106

Klaus Ribeiro Soares Saad Barufatti

*Modelos de Sazonalidade Aplicados para
Séries de Comércio Exterior no Brasil*

Orientador:

Prof. Dr. Emerson Fernandes Marçal

MONOGRAFIA EM CIÊNCIAS ECONÔMICAS
INSTITUTO DE ECONOMIA
UNIVERSIDADE ESTADUAL DE CAMPINAS

Campinas - SP

Dezembro / 2006

CEBOC/IE

Sumário

Lista de Figuras

Lista de Tabelas

1	Capítulo 1: Descrição do Problema e Metodologias	p. 6
1.1	Descrição do Problema	p. 6
1.1.1	O que é a sazonalidade?	p. 7
1.2	Motivações	p. 7
1.3	Metodologias	p. 7
1.3.1	O modelo mais geral de sazonalidade	p. 9
1.3.1.1	Filtros	p. 9
1.3.2	Box-Jenkins	p. 10
1.3.2.1	Sazonalidade no modelo Box-Jenkins	p. 11
1.3.3	Modelo com <i>dummies</i> sazonais	p. 11
1.3.4	Seasonal ARMA	p. 11
1.3.4.1	Seasonal ARIMA	p. 13
1.3.5	X-11	p. 14
1.3.6	Modelo Estrutural	p. 15
1.3.7	O teste de HEGY para sazonalidade	p. 16
1.4	Critério de seleção dos modelos	p. 17
1.5	Breve descrição dos dados	p. 18
2	Apresentação dos Resultados (exportações)	p. 21

2.1	Para os dados Trimestrais	p. 21
2.1.1	O Teste HEGY	p. 21
2.2	Para dados Mensais	p. 23
2.2.1	O teste HEGY	p. 23
2.2.2	O ajustamento pelo X-12-ARIMA	p. 24
2.3	SARIMA	p. 25
2.4	Modelo Estrutural	p. 27
3	Apresentação dos Resultados (importações)	p. 30
3.1	Para os dados Trimestrais	p. 30
3.1.1	O teste HEGY	p. 30
3.2	Para os dados Mensais	p. 32
3.2.1	O Teste HEGY	p. 32
3.2.2	O ajustamento pelo X-12-ARIMA	p. 33
3.3	SARIMA	p. 34
3.4	Modelo Estrutural	p. 36
	Referências	p. 39

Lista de Figuras

1	A metodologia do Box-Jenkins para modelos ARIMA	p. 10
2	Evolução das Exportações. De 1959 até 2006	p. 18
3	Evolução das Importações. De 1959 até 2006	p. 18
4	Correlograma das Exportações em nível	p. 19
5	Correlograma das Importações em nível	p. 19
6	Correlograma das Exportações em primeira diferença	p. 19
7	Correlograma das Importações em primeira diferença	p. 19
8	Série de Exportações Ajustada pelo X-12 e I.C. das Previsões	p. 26
9	Série de Exportações Ajustada pelo SARIMA	p. 27
10	Componente de sazonalidade do modelo estrutura para exportações	p. 28
11	Componente de tendência do modelo estrutura para exportações	p. 28
12	Coefficiente das Dummies Mensais no tempo, para o Modelo Estrutural das Exportações	p. 29
13	Série de Importações Ajustada pelo X-12 e I.C. das Previsões	p. 35
14	Série de Importações Ajustada pelo SARIMA	p. 36
15	Componente de sazonalidade do modelo estrutura para importações	p. 37
16	Componente de tendência do modelo estrutura para importações	p. 37
17	Coefficiente das Dummies Mensais no tempo, para o Modelo Estrutural das Importações	p. 38

Lista de Tabelas

1	Simulação de valores para demonstrar propriedade do filtro $\Delta\Delta_4$	p. 14
2	Verificação do teste de raiz unitária sazonal para dados trimestrais, filtros a serem utilizados	p. 17
3	Estatística descritiva. Exportações de 1959:01–2006:04 (mensal)	p. 19
4	Estatística descritiva. Importações de 1959:01–2006:04 (mensal)	p. 19
5	Resumo dos valores críticos obtidos com o teste HEGY para dados trimestrais de exportações	p. 22
6	Resumo dos valores críticos obtidos com o teste HEGY para dados mensais de exportações	p. 24
7	Coefficientes dos modelos X-12-ARIMA	p. 24
8	Previsão pelo X-12 com base no modelo $(2\ 1\ 2)(1\ 1\ 1)$, intervalo de confiança de 5%.	p. 25
9	Modelo $(2\ 1\ 2)(1\ 1\ 1)$, com <i>dummies</i> sazonais e de dia útil	p. 26
10	Ajustamento da série de Importações pelo SARIMA	p. 27
11	Resumo dos valores críticos obtidos com o teste HEGY para dados trimestrais de importações	p. 31
12	Resumo dos valores críticos obtidos com o teste HEGY para dados mensais de importações	p. 33
13	Coefficientes dos modelos X-12-ARIMA	p. 33
14	Previsão com base no modelo $(2\ 1\ 2)(1\ 1\ 1)$, intervalo de confiança de 5%	p. 34
15	Modelo $(2\ 1\ 2)(1\ 1\ 1)$, com <i>dummies</i> e dia útil	p. 35
16	Ajustamento da série de Importações pelo SARIMA	p. 36

1 Capítulo 1: Descrição do Problema e Metodologias

1.1 Descrição do Problema

A idéia desse trabalho é modelar de forma adequada os efeitos sazonais presentes nas séries de exportação e importação brasileiras, com base nos dados mensais fornecidos pelo Banco Central do Brasil, que contemplam dados para quase 50 anos.

Os níveis de exportação e importação são indicadores importantes do estado cíclico da economia, bem como fazem parte da demanda agregada. A série mensal dessas variáveis é “contaminada” por distúrbios, como variações sazonais, efeitos de calendário e feriado. Os métodos tradicionais de ajustamento sazonal não capturam de forma rigorosa essas variações. Como, por exemplo, os métodos conhecidos por X-11 e X-12, que são amplamente utilizados pelas organizações de estudos estatísticos, como o IBGE.¹

Nesse sentido, o trabalho almeja apresentar metodologias alternativas de ajuste sazonal, que tratam os dados de forma muito mais adequada e flexível. Uma das vantagens desse tipo de ajustamento, além da “suavização” da série, possibilita auferir previsões de mais longo prazo, e entender alguns dos comportamentos dos agentes.

Estudos mais recentes, como os de FRANCES (1996), mostram que a eficiência dos modelos tradicionais de desazonalização pode ser contestada através dois aspectos fundamentais:

1. As flutuações sazonais não são constantes ao longo do tempo, tanto para ajustes mensais quanto trimestrais.
2. Mais importante é que nas séries macroeconômicas de modo geral, existem poucas evidências de as flutuações sazonais e não sazonais sejam independentes. Ou seja,

¹(U.S. CENSUS BUREAU, 2002)

observa-se padrões de comportamento sazonal diferentes entre os momentos de ciclos expansivos e recessivos.

1.1.1 O que é a sazonalidade?

A sazonalidade é o movimento intra-anual sistemático, embora não necessariamente regular, causado por variações do tempo ou das estações, como na produção de bebidas e vestuário. Variações do calendário (feriados), também da ocasião de tomadas de decisão de consumo e produção pelos agentes econômicos. É comum ainda, perceber a existência da sazonalidade em produtos agropecuários, que é associada ao próprio ciclo do produto (safra e entressafra).

É importante reconhecer a presença do componente sazonal, estimá-lo e removê-lo (tal processo é conhecido como ajustamento sazonal), para então analisar o comportamento da tendência de longo prazo da série.

“Seasonality is the systematic, although not necessarily regular, intra-year movement caused by changes of the weather, the calendar, and timing of decisions, directly or indirectly through the production and consumption decisions made by the agents of the economy. These decisions are influenced by the endowments, the expectations and the preferences of agents, and the production techniques available in the economy.”²

1.2 Motivações

Dentre as motivações desse projeto, a principal delas é o aprendizado. A Econometria é um instrumental interessante e poderoso, embora limitado. O presente trabalho é uma forma de buscar contato mais estreito com a área de Econometria, bem como aprofundar em alguns temas que foram estudados durante a graduação, como por exemplo, a modelagem ARIMA, modelo estrutural e padrões sazonais.

1.3 Metodologias

Nesta seção vamos discutir quais as metodologias econométricas que serão utilizadas para tratar o problema da sazonalidade, bem como uma breve explicação de cada uma.

²(HYLLEBERG, 1992 apud FRANSES, 1996)

Estarei trabalhando com a modelagem do tipo ARIMA, que utiliza o conceito de estrutura autoregressiva e média móvel, podendo ser ou não integrado.

Muitos modelos macroeconômicos tem a característica de que os choques se acumulam, alterando de forma significativa o nível ou tendência da série. Conhecidos por “random walk”, ou passeio aleatório, podendo ter tendência determinística (*random walk com drift*) ou não (*random walk sem drift*). Esses processos podem ser integrados (ou seja, tomar alguma diferença) para que se tornem estacionários. Um exemplo simples de processo integrado de ordem um, $I(1)$, é o $AR(1)$ a seguir:

$$y_t = y_{t-1} + \varepsilon_t \quad , \text{ com } \varepsilon_t \Rightarrow \text{niid } (0, \sigma^2) \quad (1.1)$$

Podemos tomar a primeira diferença de um modelo como esse para torná-lo estacionário, ou seja, para fazer com que os choques não se acumulem mais. Podemos reescrever o processo acima, através de modificações recursivas (obtendo-se um processo $MA(1)$) onde fica mais fácil perceber que os choques se acumulam:

$$y_t = y_0 + \sum_{i=1}^t \varepsilon_i \quad (1.2)$$

Basta notar que os erros se somam. Pode ser obtido facilmente também que: $E(y_t^2) = t\sigma^2$, onde a variância depende do tempo, claro sinal de que o processo é não estacionário. Mais, $E(y_t y_{t-k}) = \frac{(t-k)\sigma^2}{t}$ para $k > 0$, portanto as covariâncias também são função do tempo, gerando um correlograma que não tende à zero.

Como já mencionado, o processo de “random walk com drift”, em que a diferença é a presença de um termo determinístico que introduz, por exemplo, tendência de crescimento. As séries macroeconômicas agregadas costumam apresentar esse tipo de comportamento, por exemplo, a série de Produto e mesmo as séries de Exportação e Importação.

Outro tratamento dado ao problema da Integração é associado à definição de raiz unitária. Se diz que um processo integrado de ordem d tem d raízes unitárias (que é também o número de diferenças que deve ser tomada). Em geral, as séries econômicas quando são integradas, são de ordem 1 (PIB, por exemplo) e algumas vezes de ordem 2 (há evidência para a inflação em alguns períodos muito inflacionários). Explicando a idéia de forma simplificada, vamos reescrever o processo $AR(1)$ dado por $Y_t = \rho Y_{t-1} + u_t$ como:

$$\Delta Y_t = (\rho - 1)Y_{t-1} + u_t \quad (1.3)$$

Podendo ser $(\rho - 1)$ representado por δ . Caso $\hat{\delta} = 0$, esse processo tem uma raiz unitária e é não estacionário. Infelizmente, a estatística de teste para essa situação tem distribuição não-padrão.³ Uma das formas de testar raiz unitária é utilizando a estatística τ , proposta por Dickey-Fuller, e o teste ADF para raiz unitária. A hipótese nula é de que $\delta = 0$, ou ainda, $\rho = 1$, portanto há raiz unitária. Caso a nula seja rejeitada, o processo é $I(0)$, vale dizer, estacionário.

1.3.1 O modelo mais geral de sazonalidade

De modo sintético, os estudos tradicionais sobre sazonalidade tentam decompor o fenômeno em três partes, através de:

$$y_t = \text{tendência} + \text{sazonalidade} + \text{irregularidade} \quad (1.4)$$

Na qual o componente de tendência representa o movimento de longo prazo. O irregular reflete movimentos não-sistemáticos e residuais. Isolando os efeitos da tendência e dos erros não-sistemáticos, podemos então captar os efeitos sazonais.

O modelo acima é adequado quando se supõe que a sazonalidade independe da tendência. Por outro lado, se as amplitudes sazonais variam com a tendência, o modelo multiplicativo, dado por: $y_t = \text{tendência} \times \text{sazonalidade} \times \text{irregularidade}$ é o mais apropriado para representar este comportamento. Caso queiramos voltar ao modelo aditivo, podemos fazê-lo simplesmente aplicando o logaritmo (\log), isto é, $\log(y_t) = \log(\text{tendência}) + \log(\text{sazonalidade}) + \log(\text{irregularidade})$.

1.3.1.1 Filtros

Os filtros são técnicas, das mais diversas, adotadas para alisar uma série. Por exemplo, no caso de um processo $I(1)$, aplica-se o filtro Δ_1 (ou seja, tomar a primeira diferença), para torná-lo estacionário. Os filtros tradicionalmente utilizados no estudo de sazonalidade são aqueles que trabalham com médias móveis. A idéia básica é a de tentar “diluir” uma pancada recente, digamos em dados mensais, somando esse valor com outros anteriores e calculando a média.

³(DICKEY; FULLER, 1979)

1.3.2 Box-Jenkins

A abordagem da metodologia Box-Jenkins é uma das mais utilizadas no estudo das séries temporais. Por motivos tais como sua generalidade (pode ser aplicada para a maioria das séries sendo elas estacionárias ou não, contendo elementos sazonais ou não) e também por estar implementada em diversos *softwares* econométricos. Essa metodologia é dividida em cinco etapas, explicitadas a seguir, e esquematizadas na Figura 1. (1) Diferenciar a série para conseguir estacionariedade, (2) identificação de um modelo experimental, (3) estimação do modelo, (4) checagem de adequação (sendo inadequado, voltar à segunda etapa) e (5) modelo pode ser usado para inferência e previsão.

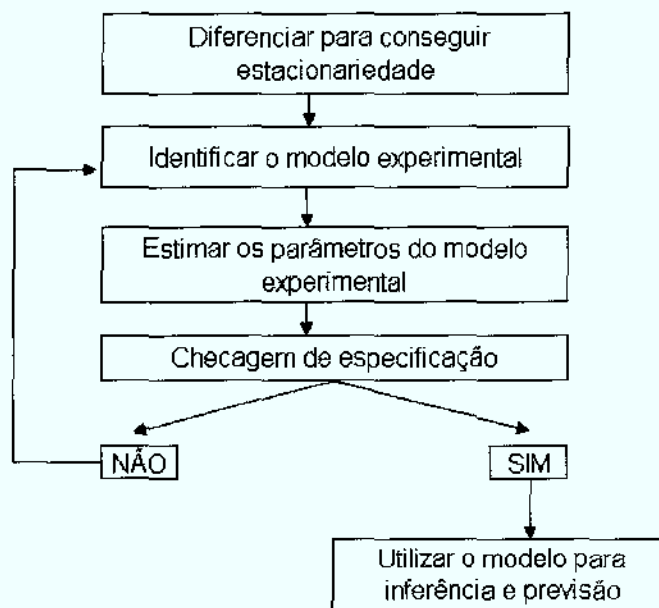


Figura 1: A metodologia do Box-Jenkins para modelos ARIMA

1. A diferenciação da série para obter estacionariedade já foi discutida brevemente na seção sobre integração. Uma outra forma de ver se a série é estacionária consiste em observar o seu correlograma. Se o correlograma da série tende rapidamente para zero, então a série é estacionária, se não, alguma diferença deve ser tomada para fazê-la.
2. Olhando também para o correlograma, podemos ter idéia da ordem dos processos Autoregressivo e Média-móvel que geraram a série. O processo MA de ordem m domina as m primeiras autocorrelações e o processo AR domina as outras a partir desta. A partir daí podemos supor as ordens para o processo ARIMA e ver como se adequam à metodologia Box-Jenkins.

3. O próximo passo é estimar os parâmetros do modelo especificado no item 2.
4. Existem diversos testes para medir a adequação do modelo. O teste do Multiplicador de Lagrange (LM) é o mais utilizado na metodologia do Box-Jenkins.

1.3.2.1 Sazonalidade no modelo Box-Jenkins

Se a tendência sazonal for determinística, podemos ter uma boa estimativa através de um modelo com *dummies* sazonais. Quando a tendência sazonal é estocástica, o fenômeno pode ser estudado tomando as diferenças. A idéia das diferenças pode ser trabalhada no modelo Box-Jenkins, quando se tem séries mensais, é bem provável que exista sazonalidade na frequência anual, pois observamos que os meses de Novembro e Dezembro são sempre mais altos em relação aos outros. Esse fato faz com que a 12ª autocorrelação seja sempre significativa. O que pode ser feito nesse caso é tomar a diferença sazonal, ou seja, (Janeiro de 2005 - Janeiro 2004), (Dezembro 2005 - Dezembro 2004) e assim por diante. Sendo $\Delta = (1 - L)$ o operador de defasagens, essa diferença pode ser expressa por: $\Delta_{12} = (1 - L^{12})$. Portanto, $\Delta_{12}y_t = y_t - y_{t-12}$. Se as séries são trimestrais, costuma ser usada a quarta diferença, ou seja, a diferença trimestral que é expressa pelo filtro: $\Delta_4 = (1 - L^4)$.

1.3.3 Modelo com *dummies* sazonais

Quando o processo sazonal é *puramente determinístico*, podemos utilizar o modelo de *dummies* sazonais para descrevê-lo. Supondo que a série seja trimestral teríamos o seguinte modelo:

$$x_t = \alpha_0 + \alpha_1 S_{1t} + \alpha_2 S_{2t} + \alpha_3 S_{3t} + \varepsilon_t \quad (1.5)$$

em que, S_1 , S_2 e S_3 são as *dummies* sazonais; α_1 até α_4 são os coeficientes dessas *dummies* e α_0 é o intercepto. Caso deseje-se estudar o modelo com as quatro variáveis binárias, isso é possível omitindo-se o intercepto (para evitar o problema de multicolinearidade perfeita). Note que os choques de um passado muito distante estão incorporados no valor dos coeficientes (choques se “somam”).

1.3.4 Seasonal ARMA

Imagine um simples caso em que a sazonalidade de uma série trimestral pode ser capturada tomando a observação do mesmo trimestre do ano anterior, como por exemplo:

$$y_t = \phi_4 y_{t-4} + \varepsilon_t, \quad |\phi_4| < 1 \quad (1.6)$$

note que esse modelo pode ser interpretado também como sendo um AR(4), com a condição de que $\phi_1 = \phi_2 = \phi_3 = 0$. Quanto mais perto $|\phi_4|$ estiver de 1, mais forte é a sazonalidade. Porém, desde que esse valor seja menor que a unidade, o efeito sazonal é não-determinístico, uma vez que as autocorrelações tenderão a zero conforme o tempo aumentar.

O modelo (1.6) pode ser extrapolado para um ARMA qualquer de s lags sazonais. A formulação geral é a que se segue:

$$\Phi(L^s)y_t = \Theta(L^s)\zeta_t \quad (1.7)$$

em que:

$$\Phi(L^s) = 1 - \Phi_1(L^s) - \dots - \Phi_P L^{Ps} \quad (1.8)$$

$$\Theta(L^s) = 1 + \Theta_1(L^s) + \dots + \Theta_Q L^{Qs} \quad (1.9)$$

e ζ_t é um ruído branco. Os valores de s serão quatro ou doze, para séries trimestrais e mensais respectivamente.

O modelo (1.7) é um ARMA puramente sazonal, de ordem $(P, Q)_s$. A função de autocorrelação terá “buracos” nos lags não-sazonais.⁴ Não se trata de um modelo muito adequado empiricamente, pois é de se supor que uma observação em Maio esteja relacionado com a observação em Maio de períodos anteriores, mas também com a observação de Abril do mesmo ano.

Há duas formas de construir modelos que permitem movimentos sazonais e não-sazonais. O primeiro, onde a sazonalidade é não-multiplicativa, pode ser construído, simplesmente adicionando algum lag no modelo (1.6). Por exemplo:

$$y_t = \phi_1 y_{t-1} + \phi_4 y_{t-4} + \varepsilon_t \quad (1.10)$$

A segunda forma é em geral preferível, e modela sazonalidade do tipo multiplicativa. Podemos obtê-la trabalhando melhor o termo de ruído branco presente no modelo (1.7), a fim de incorporar um fator ARMA (p, q) :

$$\phi(L)\zeta_t = \theta(L)\varepsilon_t \quad (1.11)$$

⁴A exemplo do que ocorre em (1.6), onde a única autocorrelação positiva será para lags múltiplos de 4, o resto será zero.

substituindo (1.11) em (1.7) conseguimos:

$$\Phi(L^s)\phi(L)y_t = \Theta(L^s)\theta(L)\varepsilon_t \quad (1.12)$$

essa é a forma multiplicativa de um ARMA sazonal, no formato $(p, q) \times (P, Q)_s$. Para ilustrar, considere um ARMA(1, 0)(1, 0)₄. De modo que:

$$(1 - \phi L)(1 - \Phi L^4) = \varepsilon_t \quad (1.13)$$

podendo ser reescrita como:

$$y_t = \phi_1 y_{t-1} + \phi_4 y_{t-4} + \phi_5 y_{t-5} + \varepsilon_t \quad (1.14)$$

sendo $\phi_1 = \phi, \phi_4 = \Phi, \phi_5 = -\phi\Phi$.

Podemos observar que a diferença entre os modelos (1.14) e (1.10) está no termo $-\phi\Phi y_{t-5}$, que é a propriedade multiplicativa incorporada pelo segundo modelo.

1.3.4.1 Seasonal ARIMA

Como discutido anteriormente, a idéia de modelos ARIMA é que uma série com tendência não estacionária pode ser estudada pelo modelo ARMA, em que a variável a ser explicada é a série original em primeira diferença. O primeiro passo é generalizar (1.7) para:

$$\Phi(L^s)\Delta_s^D y_t = \Theta(L^s)\zeta_t \quad (1.15)$$

onde D é a ordem de diferença sazonal (raramente para dados econômicos assume valor superior a uma unidade).

O ruído branco presente em (1.15) pode ser substituído por um ARIMA qualquer, (p, d, q) , similarmente ao que fizemos em (1.12):

$$\begin{aligned} \Phi(L^s)\phi(L)\Delta^d\Delta_s^D y_t &= \theta_0 + \Theta(L^s)\theta(L)\varepsilon_t \\ t &= d + sD + 1, \dots, T \end{aligned} \quad (1.16)$$

que é conhecido como modelo SARIMA multiplicativo, de ordens $(p, d, q) \times (P, D, Q)_s$.

Note o seguinte filtro: $(1-L)(1-L^s)y_t$,⁵ quando $s = 4$ esse filtro representa a variação do trimestre no tempo t em relação a variação do mesmo trimestre em $t-1$; se os valores estiverem em LN, isso representa a variação na taxa de crescimento do trimestre atual em

⁵Ou $\Delta\Delta_s y_t$.

relação ao crescimento do mesmo trimestre do período anterior. Essa dupla diferenciação da série é comumente aplicada ao modelo Box-Jenkins descrito nesse capítulo.

Podemos perceber o que foi dito anteriormente na Tabela 1. Fiz o cálculo para dados fictícios. Podemos perceber que quando o crescimento no primeiro trimestre é maior (menor) do que o crescimento no segundo trimestre, o valor do filtro $\Delta\Delta_4$ é positivo (negativo). E a diferença entre as duas taxas de crescimento é, aproximadamente, o valor do filtro.

Trimestre	Valor	Crescimento em t	Crescimento em $t-1$	$\Delta\Delta_4 y_t$
2000:01	10	0	0	0
2000:02	12	0	0	0
2000:03	15	0	0	0
2000:04	23	0	0	0
2001:01	24	0	0	0
2001:02	27	0,13	0,20	-0,06
2001:03	25	-0,07	0,25	-0,30
2001:04	31	0,24	0,53	-0,21
2002:01	32	0,03	0,04	-0,01
2002:02	35	0,09	0,13	-0,03
2002:03	35	0,00	-0,07	0,08
2002:04	34	-0,03	0,24	-0,24
2003:01	37	0,09	0,03	0,05
2003:02	41	0,11	0,09	0,01
2003:03	42	0,02	0,00	0,02
2003:04	43	0,02	-0,03	0,05
2004:01	48	0,12	0,09	0,03
2004:02	50	0,04	0,11	-0,06
2004:03	52	0,04	0,02	0,02

Tabela 1: Simulação de valores para demonstrar propriedade do filtro $\Delta\Delta_4$

1.3.5 X-11

A metodologia mais difundida para tratar o problema de sazonalidade, inclusive por organismos reconhecidos de pesquisa como o IBGE, é o X-11.⁶

A versão aditiva do X-11 envolve duas etapas principais. A primeira é a aplicação sequencial de diversos filtros média-móveis, a fim de caracterizar as tendências e as flutuações sazonais. Em seguida, realiza-se a remoção dos “out liers”, ponderando seus valores de forma especial (menos peso).

⁶(IBGE, 2004)

A principal crítica ao método X-11, com relação à estimação dos componentes não-observáveis, está no fato de que as estimativas dos anos mais recentes não possuem o mesmo grau de confiabilidade se comparadas com as observações do centro da série. Trata-se de uma limitação inerente aos procedimentos de alisamento que utilizam filtros lineares, já que as primeiras e as últimas observações não podem ser suavizadas com o mesmo conjunto de pesos simétricos aplicados às observações centrais. Por este motivo, as estimativas para as observações correntes são revistas à medida que novos dados são adicionados à série observada.

1.3.6 Modelo Estrutural

Um padrão sazonal fixo apresenta a seguinte propriedade: o efeito sazonal no tempo t , digamos γ_t , é igual ao efeito sazonal no tempo $t-s$, γ_{t-s} . Uma forma de modelar um efeito sazonal que varie seria:

$$\gamma_t = \gamma_{t-s} + \omega_t \quad (1.17)$$

sendo ω_t um termo de ruído branco, e o modelo:

$$y_t = \gamma_t + u_t \quad (1.18)$$

Esta especificação tem problemas, uma vez que o operador $(1 - L^s)$ mistura nível com componente sazonal, já que $(1 - L^s) = (1 - L)(1 + L + L^2 + \dots + L^{s-1})$ e o primeiro operador extrai a tendência e o segundo tem a propriedade de somar zero a cada s períodos que é uma característica do componente sazonal. O segundo problema com o operador $(1 - L^s)$ é não somar zero a cada s períodos, o que faz com que esta especificação não tenha as características desejáveis para um componente sazonal.

Uma forma de restringir os fatores a somarem zero a cada s períodos é especificar o componente sazonal da seguinte forma:

$$\gamma_t + \gamma_{t-1} + \dots + \gamma_{t-s} = 0 \quad (1.19)$$

para permitir uma variação suave em (1.19), podemos fazer com que essa soma seja igual à um termo de erro aleatório:

$$\sum_{j=0}^{s-1} \gamma_{t-j} = \omega_t \Leftrightarrow \gamma_t = - \sum_{j=0}^{s-1} \gamma_{t-j} + \omega_t \quad (1.20)$$

pela forma estocástica de sazonalidade, podemos elaborar que:

$$y_t = \mu_t + \gamma_t + u_t, \quad t = 1, \dots, T \quad (1.21)$$

em que μ_t que representa o nível pode também mudar ao longo do tempo se o tratarmos como se fosse um *random walk*. Diferentemente de (1.18), os componentes de nível e sazonal estão separados em (1.21).

1.3.7 O teste de HEGY para sazonalidade

O teste de HEGY (nome dado pelos seus autores: Hylleberg, Engle, Granger e Yoo) serve para identificar se existe raiz sazonal numa série, e também determinar qual o tipo de sazonalidade presente (puramente determinística, estacionária ou integrada). Há fortes evidências⁷ que o teste de HEGY é robusto o bastante para esse fim. Conforme (MADDALA; KIM, 1998), o teste pode ser realizado através da estimação do seguinte modelo por M.Q.O.:

$$\Delta_4 y_t = \sum_{s=1}^4 \alpha_s D_{st} + \gamma T_t + \pi_1 y_{1,t-1} + \pi_2 y_{2,t-1} + \pi_3 y_{3,t-2} + \pi_4 y_{3,t-1} + \sum_{i=1}^k \phi_i \Delta_4 y_{t-i} + \varepsilon_t$$

Em que:

$$y_{1t} = (1 + L + L^2 + L^3)y_t$$

$$y_{2t} = -(1 - L + L^2 - L^3)y_t$$

$$y_{3t} = -(1 - L^2)y_t$$

T é um termo de tendência

D_{st} representam as *dummies* sazonais.

As estatísticas t 's dos parâmetros π_1 , π_2 , π_3 e π_4 devem ser comparados à uma tabela de valores críticos, e a análise desses valores leva à conclusão de qual o tipo de sazonalidade e a frequência dos ciclos.

Se $\pi_1 = 0$, a série contém uma raiz unitária na frequência zero, ou seja, a série tem uma tendência estocástica não-sazonal (deve-se operar com as primeiras diferenças). Quando $\pi_2 = 0$, há dois ciclos por ano. Se $\pi_3 = \pi_4 = 0$, existe uma raiz unitária na frequência sazonal. Os filtros que deverão ser utilizados estão representados pela Tabela 2.

⁷(MONTANES; LOPES, 2005)

Teste	Filtro a ser utilizado
$\pi_1 = 0$	$(1 - L)$
$\pi_2 = 0$	$(1 + L)$
$\pi_3 = \pi_4 = 0$	$(1 + L^2)$

Tabela 2: Verificação do teste de raiz unitária sazonal para dados trimestrais, filtros a serem utilizados

O método de HEGY pode ser também elaborado para trabalhar com dados mensais. Conforme apresentado por (MADDALA; KIM, 1998) a formulação é a que se segue:

$$\begin{aligned} \varphi^*(L)y_{8,t} = & \pi_1 y_{1,t-1} + \pi_2 y_{2,t-1} + \pi_3 y_{3,t-2} + \pi_4 y_{3,t-1} + \pi_5 y_{4,t-2} + \pi_6 y_{4,t-1} \\ & + \pi_7 y_{5,t-2} + \pi_8 y_{5,t-1} + \pi_9 y_{6,t-2} + \pi_{10} y_{6,t-1} + \pi_{11} y_{7,t-2} + \pi_{12} y_{7,t-1} + \mu_t + \varepsilon_t \end{aligned}$$

em que $\varphi^*(L)$ é uma função polinomial de L , μ_t é um componente determinístico e:

$$\begin{aligned} y_{1,t} &= (1 + L)(1 + L^2)(1 + L^4 + L^8)y_t \\ y_{2,t} &= -(1 - L)(1 + L^2)(1 + L^4 + L^8)y_t \\ y_{3,t} &= -(1 - L^2)(1 + L^4 + L^8)y_t \\ y_{4,t} &= -(1 - L^4)(1 - \sqrt{3}L + L^2)(1 + L^2 + L^4)y_t \\ y_{5,t} &= -(1 - L^4)(1 + \sqrt{3}L + L^2)(1 + L^2 + L^4)y_t \\ y_{6,t} &= -(1 - L^4)(1 - L^2 + L^4)(1 - L + L^2)y_t \\ y_{7,t} &= -(1 - L^4)(1 - L^2 + L^4)(1 + L + L^2)y_t \\ y_{8,t} &= -(1 - L^{12})y_t \end{aligned}$$

Caso exista raiz sazonal, os respectivos π_i s são iguais a zero. Devido ao fato de que há pares complexos, só existe raiz sazonal se o par de π s forem iguais a zero simultaneamente.⁸ Não haverá raiz sazonal se π_2 até π_{12} forem estatisticamente diferentes de zero. Se $\pi_1 = 0$, a série tem uma raiz unitária. Quando $\pi_1 = 0$ e π_2 até π_{12} são diferentes de zero, a sazonalidade pode ser modelada através da primeira diferença e por *dummies* sazonais. Se todos os π_i forem iguais a zero, é conveniente a aplicação do filtro Δ_{12} .

1.4 Critério de seleção dos modelos

O marco metodológico a ser utilizado para seleção de modelos no decorrer deste trabalho é o de que nenhum modelo é considerado verdadeiro, ou preferido *a priori*, por tanto,

⁸As raízes i e $-i$ só existem se $\pi_3 = \pi_4 = 0$.

todos os modelos “potenciais” serão igualmente comparados usando teoria estatística. O que implica em selecionar o modelo que minimize certo valor de um critério de informação, que combinará tanto o ajustamento obtido quanto o número de parâmetros utilizados. Está de acordo com as sugestões apresentadas em (GRANGER *et al.*, 1995).

Primeiro serão realizados testes de especificação para o modelo. Em seguida, será aplicado os critérios de informação (que avaliam ajuste e parcimônia). O teste F e de razão de verossimilhança para modelos de encaixe. E teste para modelos não encaixados.

Uma função de informação bem conhecida é o critério de Akaike (AIC), calculado através de:

$$AIC(k) = n \log \hat{\sigma}_{ML}^2 + 2k$$

onde $\hat{\sigma}_{ML} = SQR/n$, sendo SQR o somatório dos resíduos ao quadrado. Existem outros critérios de informações parecidos, como por exemplo o critério de Schwarz,⁹ entre outros.

1.5 Breve descrição dos dados

Os dados foram obtidos no *site* do Banco Central do Brasil, nas freqüências mensais e trimestrais. Nessa seção será feita uma breve apresentação dos dados de estatística descritiva, para importações e exportações.

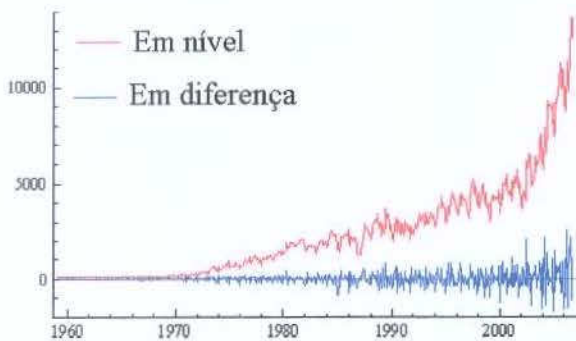


Figura 2: Evolução das Exportações. De 1959 até 2006

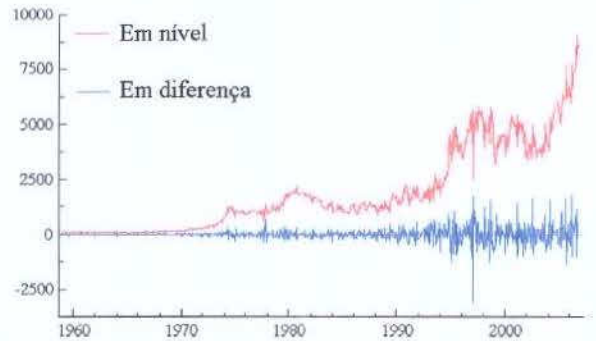


Figura 3: Evolução das Importações. De 1959 até 2006

Observado o gráfico, a possibilidade de variações sazonais é bem clara. Além disso, há tendência de crescimento e a série é claramente não-estacionária. Na série de importações houve, aparentemente, uma mudança muito acentuada, refletida no salto de nível e de volatilidade a partir de 1994 (talvez uma “quebra estrutural” introduzida pelo Plano

⁹Que penaliza mais fortemente o aumento do número de parâmetros.

Real).

Média	Mediana	Desv. Pad.
2.258,40	1.754,00	2.306,20
Skewness	Ex. kurtosis	C. V.
1,485	2,449	1,021

Tabela 3: Estatística descritiva. Exportações de 1959:01–2006:04 (mensal)

Média	Mediana	Desv. Pad.
1.834,00	1.231,80	1.796,70
Skewness	Ex. kurtosis	C. V.
1,034	-0,043	0,980

Tabela 4: Estatística descritiva. Importações de 1959:01–2006:04 (mensal)

Construindo o correlograma para as séries nas Figuras 4 e 5, percebemos que a memória é alta e cai muito lentamente. Embora um teste formal seja necessário, representa um forte indício de que a série é integrada.

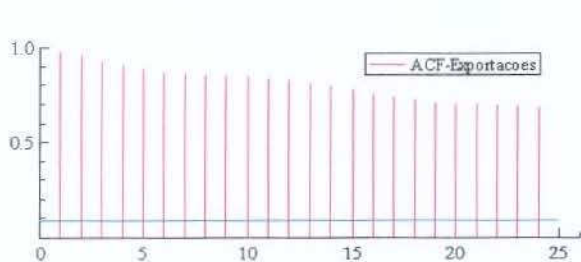


Figura 4: Correlograma das Exportações em nível

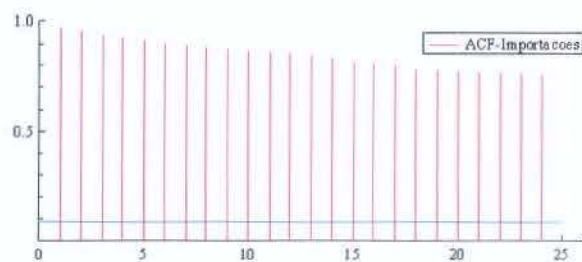


Figura 5: Correlograma das Importações em nível

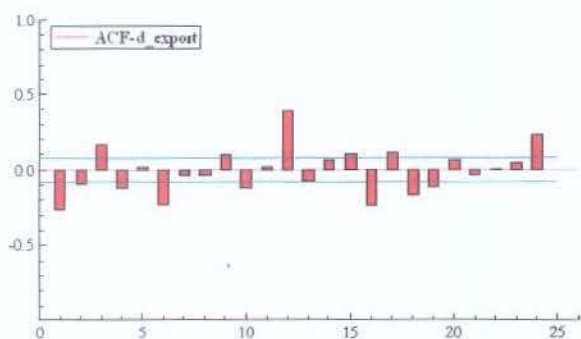


Figura 6: Correlograma das Exportações em primeira diferença

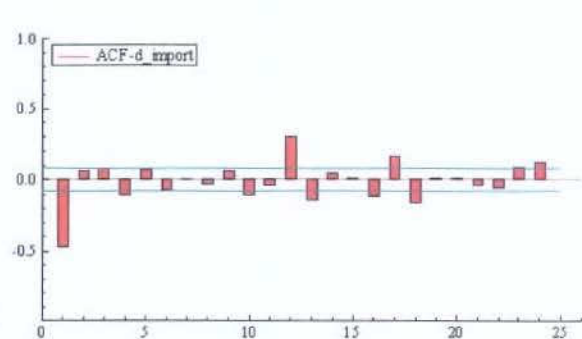


Figura 7: Correlograma das Importações em primeira diferença

As Figuras 6 e 7 temos o correlograma para a série em primeira diferença. Há algumas coisas interessantes a serem observadas, quando tomamos a diferença da série, o correlograma se torna próximo à zero muito rapidamente (novamente, evidência de que é integrada de ordem 1). Além disso, algumas autocorrelações voltam a ser significativas depois de passado alguns meses, notavelmente aqueles dados que representam o mesmo mês nos anos anteriores, ou seja, a décima segunda e a vigésima quarta, para ambas as séries. Fato que pode ser causado pela presença de fatores sazonais.

2 Apresentação dos Resultados (exportações)

2.1 Para os dados Trimestrais

2.1.1 O Teste HEGY

Abaixo é apresentado o modelo de regressão auxiliar para o teste HEGY. Vimos na primeira seção deste trabalho que pode ser utilizado para avaliar se a série é integrada sazonalmente, e em quais freqüências, indicando quais diferenças devem ser tomadas em caso afirmativo.

$$\begin{aligned} \widehat{\Delta_4 y_t} = & -0,00853689 y_{1,t-1} - 0,336891 y_{2,t-1} - 0,190675 y_{3,t-1} - 0,381238 y_{3,t-2} \\ & (-1,499) \quad (-3,813) \quad (-2,276) \quad (-4,790) \\ & + 0,000849629 trend + 0,114253 Dq_1 + 0,292168 Dq_2 + 0,298402 Dq_3 \\ & (1,416) \quad (0,867) \quad (2,242) \quad (2,267) \\ & + 0,204802 Dq_4 + 0,361089 \Delta_4 y_{t-1} - 0,100643 \Delta_4 y_{t-2} - 0,114298 \Delta_4 y_{t-3} \\ & (1,543) \quad (3,229) \quad (-0,985) \quad (-1,148) \\ & + 0,0198064 \Delta_4 y_{t-4} + 0,0767104 \Delta_4 y_{t-5} \\ & (0,204) \quad (1,036) \\ N = 182 \quad \bar{R}^2 = 0,6238 \quad F(14,168) = 24,09 \quad \hat{\sigma} = 0,10447 \\ & \text{(estatísticas-}t \text{ entre parênteses)} \end{aligned}$$

em que *trend* é um termo de tendência; Dq_i são dummies sazonais (para os respectivos trimestres); $y_{1,t}$ até $y_{3,t}$ representam as variáveis independentes que são utilizadas para testar a existência de raiz unitária na freqüência sazonal; e $y_{4,t-i}$ são valores defasados da variável a ser explicada.

O valor- t para π_1 foi de -1,49 contra um valor crítico¹ de -3,49. Vale dizer, não

¹Sempre ao nível de significância de 5%. Os valores críticos seguem distribuição não-padrão e podem ser consultados em: (MADDALA; KIM, 1998)

rejeitamos a hipótese nula de que $\pi_1 = 0$. O valor-t para π_2 foi de -3,80 (contra -2,91); para π_3 de -4,78 (contra -3,41) e para π_4 de -2,27 (contra -2,27). O teste-F conjunto para π_3 e π_4 foi de 14,08 (contra 6,57). Sendo assim, π_2 , π_3 e π_4 são estatisticamente diferentes de zero, assim como a hipótese conjunta para $\pi_3 = \pi_4 = 0$ também foi rejeitada.

A única hipótese $\pi_i = 0$ aceita foi quando $i = 1$. Isso implica que a série tem uma raiz unitária na frequência zero (tendência estocástica *não-sazonal*). O filtro apropriado nesse caso é $(1 - L)$, ou seja, apenas a primeira diferença (o fato de que $\pi_1 = 0$ indica que a série é integrada de ordem 1). Na metodologia do X-11, o filtro $\Delta_1 y_t$ já é o suficiente (não sendo necessário aplicar o segundo filtro $\Delta_4 y_t$).

Coeficiente	Valor auferido	Valor crítico	$\pi_i = 0$
π_1	-1,50	-3,49	Aceita
π_2	-3,81	-2,91	Rejeita
π_3	-4,79	-3,41	Rejeita
π_4	-2,28	-2,27	Rejeita
π_3, π_4	14,08	6,57	Rejeita

Tabela 5: Resumo dos valores críticos obtidos com o teste HEGY para dados trimestrais de exportações

2.2 Para dados Mensais

2.2.1 O teste HEGY

$$\begin{aligned}
\hat{y}_8 = & -0,00167115 y_{1,t-1} - 0,135054 y_{2,t-1} - 0,0483334 y_{3,t-1} - 0,127782 y_{3t-2} \\
& \quad (-1,473) \quad (-5,973) \quad (-2,306) \quad (-6,288) \\
& - 0,184169 y_{4,t-1} - 0,180643 y_{4,t-2} + 0,0111933 y_{5,t-1} - 0,0607087 y_{5,t-2} \\
& \quad (-6,860) \quad (-6,894) \quad (0,737) \quad (-3,911) \\
& - 0,146651 y_{6,t-1} - 0,0701468 y_{6,t-2} + 0,0848202 y_{7,t-1} - 0,360368 y_{7,t-2} \\
& \quad (-7,014) \quad (-3,249) \quad (2,746) \quad (-8,690) \\
& + 0,000170635 trend - 0,0542600 d_{jan} + 0,0428707 d_{fev} + 0,166184 d_{mar} \\
& \quad (1,430) \quad (-0,828) \quad (0,658) \quad (2,563) \\
& + 0,156524 d_{abr} + 0,172529 d_{mai} + 0,166623 d_{jun} + 0,171468 d_{jul} \\
& \quad (2,411) \quad (2,653) \quad (2,574) \quad (2,637) \\
& + 0,138260 d_{ago} + 0,110179 d_{set} + 0,0666891 d_{out} + 0,0504791 d_{nov} \\
& \quad (2,116) \quad (1,682) \quad (1,017) \quad (0,770) \\
& + 0,135353 d_{dez} + 0,0841553 y_{8,t-1} + 0,0491967 y_{8,t-2} + 0,0324386 y_{8,t-3} \\
& \quad (2,068) \quad (1,729) \quad (1,008) \quad (0,682) \\
& + 0,0654313 y_{8,t-4} - 0,0721275 y_{8,t-5} \\
& \quad (1,389) \quad (-1,709) \\
N = 557 \quad \bar{R}^2 = 0,6735 \quad F(29, 527) = 40,55 \quad \hat{\sigma} = 0,11034 \\
& \quad \text{(estatísticas-}t \text{ entre parênteses)}
\end{aligned}$$

em que *trend* é um termo de tendência; $d_{mês}$ são dummies sazonais (para os respectivos meses); $y_{1,t}$ até $y_{7,t}$ representam as variáveis independentes que são utilizadas para testar a existência de raiz unitária na frequência sazonal; e $y_{8,t-i}$ são valores defasados da variável a ser explicada.

Percebemos pela Tabela 6, o teste-t para os parâmetros rejeita a maioria das hipóteses nulas (de que $\pi_i = 0$), exceto para dois casos, a saber: π_1 e π_4 . O teste-F conjunto para esses dois parâmetros também não consegue rejeitar a hipótese nula.

Como $\pi_1 = 0$, a série é integrada de ordem 1, ou seja, deve ser modelada utilizando a primeira diferença. Caso todas as outras fossem diferentes de zero, seria adequado modelar pela primeira diferença e *dummies* sazonais. Como π_4 não foi significativo, podemos dizer que existe pelo menos uma raiz unitária sazonal.

Coefficiente	Valor auferido	Valor crítico	$\pi_i = 0$
π_1	-1,47	-3,24	Aceita
π_2	-5,97	-2,65	Rejeita
π_3	-6,29	-1,71	Rejeita
π_4	-2,31	-3,12	Aceita
π_5	-6,89	-2,99	Rejeita
π_6	-6,86	-3,12	Rejeita
π_7	-3,91	+0,12	Rejeita
π_8	+0,74	-0,43*	Rejeita
π_9	-3,25	-2,54	Rejeita
π_{10}	-7,01	-3,07	Rejeita
π_{11}	-8,69	-0,73	Rejeita
π_{12}	+2,75	-0,44*	Rejeita
π_1, π_4	3,83	5,63	Aceita
π_5, π_6	25,58	5,84	Rejeita
π_1, π_5	24,85	5,90	Rejeita
π_9, π_{10}	24,62	5,71	Rejeita
π_{11}, π_{12}	47,16	5,84	Rejeita

Tabela 6: Resumo dos valores críticos obtidos com o teste HEGY para dados mensais de exportações

2.2.2 O ajustamento pelo X-12-ARIMA

O ajustamento foi feito pelo método multiplicativo, e calculado no "GiveWin", com o módulo X-12. Além dos ARIMA's padrões, estes outros modelos foram calculados: (1 1 1); (2 1 2); (2 1 1); (2 1 3); (3 1 1); (3 1 2); (3 1 3); (2 1 2)(1 0 1); (2 1 3)(1 0 1); (3 1 2)(1 0 1); (1 1 2)(1 0 1); (2 1 2)(1 1 1); (1 1 3)(1 0 1).

Coeficientes	Modelos		
	(2 1 2)(1 1 1)	(2 1 2)(1 0 1)	(1 1 2)(1 0 1)
AR não-sazonal (Lag 1)	-0,379	-0,401	-0,535
(Lag 2)	-0,190	-0,171	∅
MA não-sazonal (Lag 1)	0,104	0,072	-0,040
(Lag 2)	0,112	0,153	0,386
AR sazonal	0,135	0,980	0,979
MA sazonal	0,891	0,820	0,819
AICC	-817,97	-835,7820	-835,7823
BIC	-787,86	-805,5240	-809,8254

Tabela 7: Coeficientes dos modelos X-12-ARIMA

A grande maioria dos modelos foi rejeitada por apresentar a estatística Ljung-Box menor do que 5%, (ou seja, evidência de autocorrelação serial), ou por evidência de

super-diferenciação. Os três modelos que passaram por todos os testes foram: $(1\ 1\ 2)(1\ 0\ 1)$; $(2\ 1\ 2)(1\ 0\ 1)$ e $(2\ 1\ 2)(1\ 1\ 1)$, e os resultados estão sumarizados na Tabela 7.

Os testes F para presença de sazonalidade (paramétrico, não-paramétrico e “moving seasonality”), acusam sazonalidade com significância menor que 1%. Todos os M’s foram aceitos, com exceção do M1.

O modelo selecionado pelo programa foi $(2\ 1\ 2)(1\ 1\ 1)$, o que foi um resultado interessante, pois ele está de acordo com as conclusões tiradas pelo teste de HEGY para dados mensais.² Os coeficientes estão apresentados na Tabela 7.

Data	Lim. Inferior	Previsão	Lim. Superior
2006.Nov	9.492	11.867	14.835
2006.Dez	9.644	12.411	15.973
2007.Jan	8.304	10.847	14.168
2007.Fev	8.006	10.719	14.353
2007.Mar	9.275	12.673	17.317
2007.Abr	8.851	12.318	17.142
2007.Mai	9.392	13.308	18.855
2007.Jun	9.421	13.577	19.566
2007.Jul	9.996	14.640	21.442
2007.Ago	9.983	14.850	22.090
2007.Set	9.544	14.411	21.760
2007.Out	9.343	14.313	21.926

Tabela 8: Previsão pelo X-12 com base no modelo $(2\ 1\ 2)(1\ 1\ 1)$, intervalo de confiança de 5%.

O quadro 8 mostra a estimativa para o futuro com base no modelo selecionado, e os respectivos intervalos de confiança. Em seguida temos esses valores representados graficamente, na figura 8, bem como todos os valores da série ajustados.

2.3 SARIMA

Como o teste de HEGY apontou que a série poderia ser estudada em primeira diferença e com o auxílio de *dummies* sazonais, resolvi rodar o mesmo modelo selecionado pelo programa X-12-ARIMA, incluindo as variáveis *dummies* para cada mês e também uma variável *DU* para dias úteis, que foi calculada através da diferença entre os dias úteis do mês e a média de dias úteis daquele ano.

²Onde foi visto que há evidências de que a série de exportações é integrada na frequência sazonal e na frequência zero. Portanto o modelo esperado seria no formato $(p\ 1\ q)(P\ 1\ Q)$

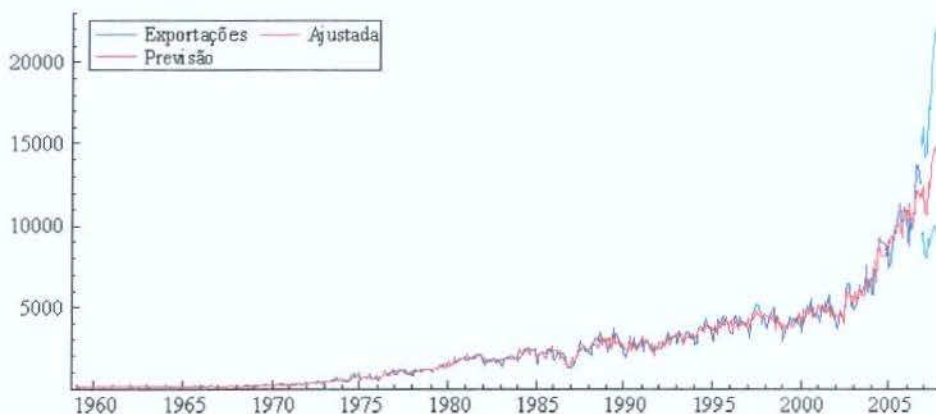


Figura 8: Série de Exportações Ajustada pelo X-12 e I.C. das Previsões

Variável	Coefficiente	estatística-t	Variável	Coefficiente	estatística-t
ϕ_1	-0,4494	-3,371*	d_{mar}	0,0639	2,719*
ϕ_2	-0,1672	-2,329*	d_{abr}	0,0893	3,782*
Φ_1	0,2900	2,724*	d_{mai}	0,0664	3,471*
θ_1	-0,0217	-0,162	d_{jun}	0,0602	3,347*
θ_2	-0,1373	-1,435	d_{jul}	0,0222	1,159
Θ_1	-0,1001	-0,854	d_{ago}	0,0094	0,425
DU	0,0302	5,329*	d_{set}	0,0100	0,679
d_{jan}	-0,1546	-5,886*	d_{out}	-0,0595	-4,164*
d_{fev}	-0,0036	-0,121	d_{nov}	-0,0088	-0,474
AIC / BIC	-882,30	-795,78	d_{dez}	0,0230	1,406

Tabela 9: Modelo (2 1 2)(1 1 1), com *dummies* sazonais e de dia útil

os valores-t com um * indicam que o respectivo teste foi significativo ao nível de 5%. É curioso notar que a introdução das *dummies* sazonais torna não-significantes os coeficientes de média-móvel do modelo original (θ_1, θ_2 e Θ_1). A variável nova introduzida, *DU*, foi bastante significativa e apresentou o coeficiente de 0,0302; vale dizer, para cada dia útil a mais (menos) que o mês tem em relação à média do ano, o valor das exportações será aproximadamente 3% maior (menor) que a média.

A tabela 10 e a figura 9 (destaque para o momento mais recente, de 1980 até 2006) mostram a qualidade do ajustamento. Para o período recente, o erro girou em torno de 4% do valor esperado (para mais ou para menos).

Visualmente, o ajuste parece ser melhor do que aquele conseguido pelo modelo X-12-ARIMA.

2.4 Modelo Estrutural

O modelo estrutural estimado a partir da equação abaixo, cuja metodologia foi explicada no primeiro Capítulo,

$$y_t = \text{tendência} + \text{sazonal}(\text{dummy}) + \text{irregular}$$

Data	Observado	Estimado	Data	Observado	Estimado
2004.Set	8.923	9.474	2005.Out	9.903	9.901
2004.Out	8.843	8.371	2005.Nov	10.790	9.750
2004.Nov	8.159	8.209	2005.Dez	10.896	11.328
2004.Dez	9.195	9.186	2006.Jan	9.271	8.888
2005.Jan	7.444	7.236	2006.Fev	8.750	9.287
2005.Fev	7.756	7.379	2006.Mar	11.366	10.952
2005.Mar	9.251	9.245	2006.Abr	9.804	10.541
2005.Abr	9.202	8.953	2006.Mai	10.275	11.378
2005.Mai	9.818	9.956	2006.Jun	11.435	11.441
2005.Jun	10.206	10.904	2006.Jul	13.622	11.552
2005.Jul	11.061	10.381	2006.Ago	13.642	13.561
2005.Ago	11.346	11.549	2006.Set	12.548	12.664
2005.Set	10.635	11.096	2006.Out	12.661	11.983

Tabela 10: Ajustamento da série de Importações pelo SARIMA

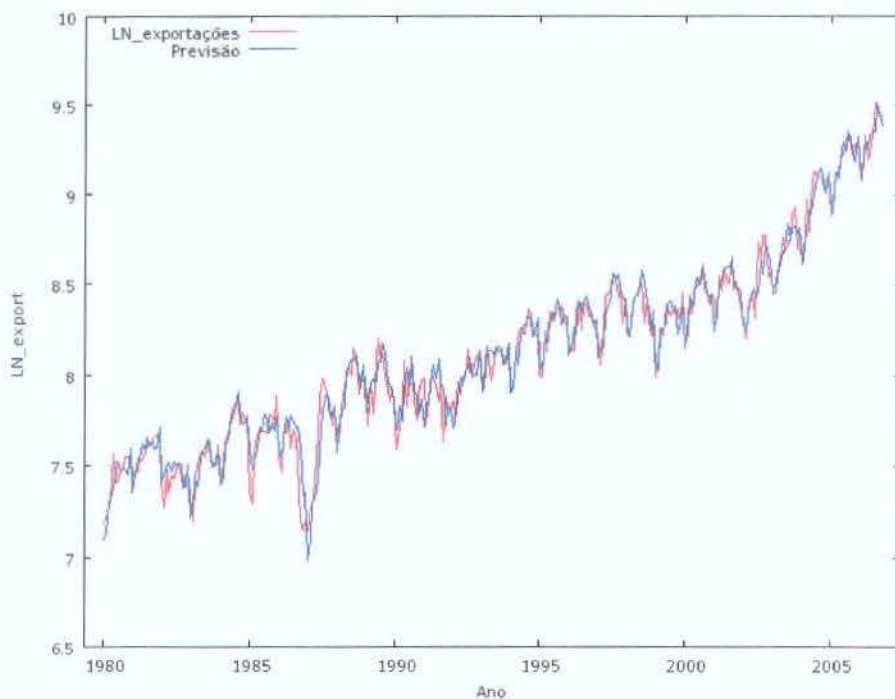


Figura 9: Série de Exportações Ajustada pelo SARIMA

os gráficos que ilustram o comportamento do modelo para os componentes sazonais e de tendência estão representados nas figuras 10 e 11, respectivamente. Podemos perceber que o padrão de sazonalidade muda sensivelmente a partir da década de 80, se tornando menos intenso do que era antes. Talvez possa ser entendido a partir de uma mudança da composição da pauta exportadora (mais diversificada em manufaturados), ou da introdução de novas tecnologias e deslocamento para novas áreas produtivas, principalmente nos setores que dependem de fatores sazonais por definição (por exemplo, agricultura).

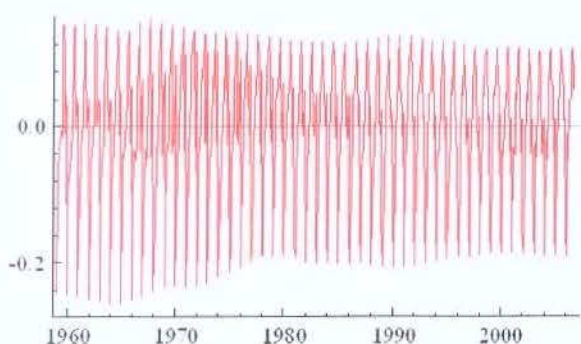


Figura 10: Componente de sazonalidade do modelo estrutura para exportações

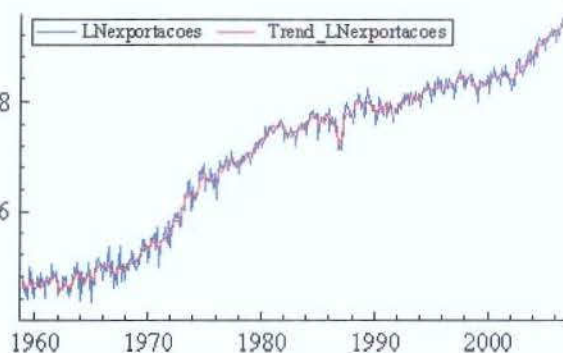


Figura 11: Componente de tendência do modelo estrutura para exportações

O fato de que o teste HEGY para dados trimestrais não aponta raiz unitária sazonal parece contradizer o teste para dados mensais, que aponta para a existência da raiz unitária na frequência sazonal. Isso pode ser devido à alterações dentro do trimestre, informação essa que passa despercebida quando se observa os dados agregados de três em três meses.

Para tentar perceber quais foram essas mudanças, se é que existiram, chegou-se na Figura 12, que mostra a evolução do valor das *dummies* sazonais. De modo geral, quando o a reta cresce (decresce), significa que o mês em questão elevou (reduziu) sua importância/contribuição.³

Podemos observar através da Figura 12 que o padrão sazonal das exportações medido através de *dummies* é bastante instável. Todos os trimestres apresentaram mudanças importantes em sua composição interna. Historicamente, o trimestre que apresenta a maior importância (em termos de volume) é o terceiro, e continua sendo. Entretanto, o segundo trimestre ganha peso considerável, especialmente após a década de 80 e principalmente em função dos meses de Maio e Junho. Pode ser atribuída à essa constatação, a mudança

³Se cada mês fosse responsável por $\frac{1}{12}$ da exportação anual, as retas seriam horizontais no valor zero. Como há meses em que, historicamente, se exporta mais (ou menos), o valor da *dummy* nesse período será maior (ou menor) do que zero. Por se tratar de um modelo estrutural, o coeficiente da *dummy* não é fixo e podemos, portanto, estudar a sua evolução.

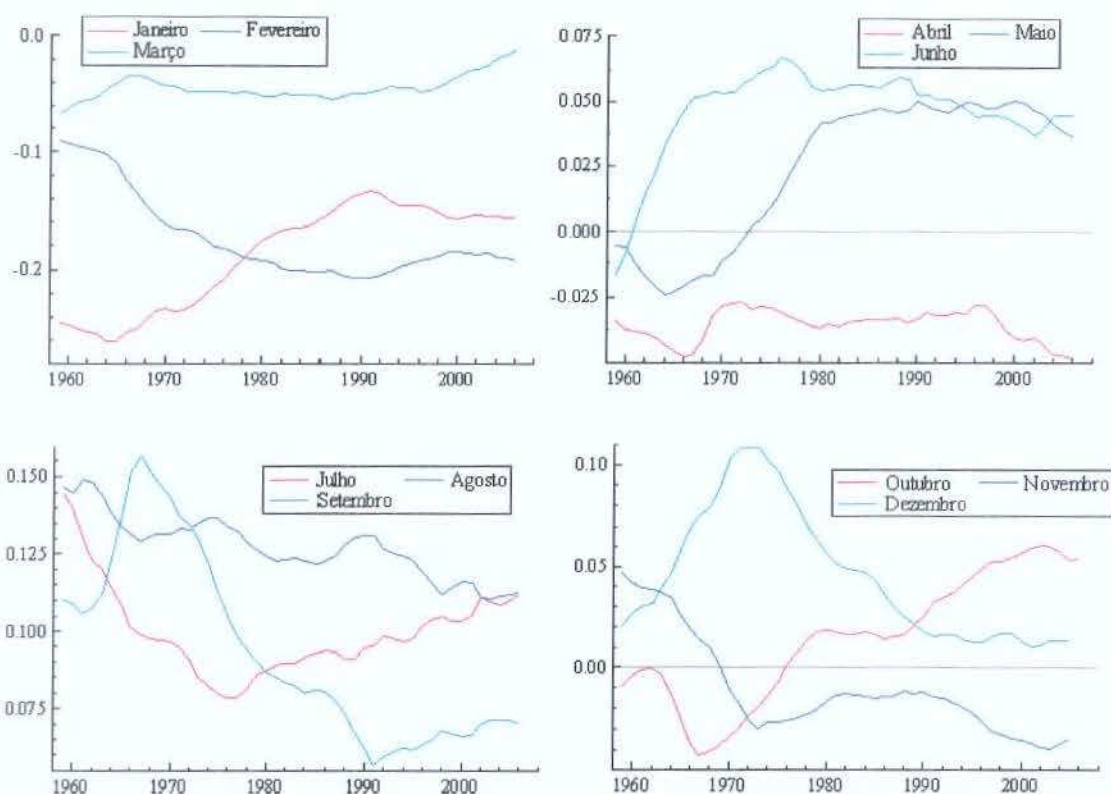


Figura 12: Coeficiente das Dummies Mensais no tempo, para o Modelo Estrutural das Exportações

captada no padrão sazonal como um todo, que foi observado anteriormente com a Figura 10.

Há no entanto mudanças importantes no primeiro e no último trimestres. Fevereiro perde importância e Janeiro ganha, apesar do semestre como um todo continuar sendo o pior em termos de volume médio. No último trimestre, houve redução expressiva no volume médio exportado no mês de Dezembro, mas que parece ser compensado dentro do trimestre pela importância cada vez maior do mês de Outubro. Essas mudanças todas foram mais do que suficientes para tornar a série de exportações integrada na frequência sazonal.

3 Apresentação dos Resultados (importações)

3.1 Para os dados Trimestrais

3.1.1 O teste HEGY

Como apresentado na Metodologia, o teste HEGY avalia se a série é integrada, e em quais frequências, indicando quais diferenças devem ser tomadas em caso afirmativo.

O modelo obtido para os dados trimestrais de importação é o que se segue:

$$\begin{aligned} \widehat{\Delta_4 y_t} = & -0,00966352 y_{1,t-1} - 0,323884 y_{2,t-1} - 0,405825 y_{3,t-1} - 0,415829 y_{3,t-2} \\ & \quad \quad \quad (-2,170) \quad \quad \quad (-3,290) \quad \quad \quad (-3,916) \quad \quad \quad (-4,037) \\ & + 0,000915716 trend + 0,124074 Dq_1 + 0,266325 Dq_2 + 0,300780 Dq_3 \\ & \quad \quad \quad (2,022) \quad \quad \quad (1,206) \quad \quad \quad (2,660) \quad \quad \quad (2,976) \\ & + 0,234553 Dq_4 + 0,286526 y_{4,t-1} - 0,0830304 y_{4,t-2} + 0,100835 y_{4,t-3} \\ & \quad \quad \quad (2,280) \quad \quad \quad (2,250) \quad \quad \quad (-0,771) \quad \quad \quad (0,937) \\ & - 0,177221 y_{4,t-4} + 0,148088 y_{4,t-5} \\ & \quad \quad \quad (-1,661) \quad \quad \quad (1,969) \end{aligned}$$

$$N = 182 \quad \bar{R}^2 = 0,7702 \quad F(13, 168) = 47,66 \quad \hat{\sigma} = 0,10079$$

(estatísticas-*t* entre parênteses)

em que *trend* é um termo de tendência; Dq_i são dummies sazonais (para os respectivos trimestres); $y_{1,t}$ até $y_{3,t}$ representam as variáveis independentes que são utilizadas para testar a existência de raiz unitária na frequência sazonal; e $y_{4,t-i}$ são valores defasados da variável a ser explicada.

O valor-*t* para π_1 foi de -2,17 contra um valor crítico¹ de -3,49. Vale dizer, não rejeitamos a hipótese nula de que $\pi_1 = 0$. O valor-*t* para π_2 foi de -3,29 (contra -2,91);

¹Sempre ao nível de significância de 5%. Os valores críticos seguem distribuição não-padrão e podem ser consultados em: (MADDALA; KIM, 1998)

para π_3 de -4,04 (contra -3,41) e para π_4 de -3,92 (contra -2,27). O teste-F conjunto para π_3 e π_4 foi de 15,90 (contra 6,57). Sendo assim, π_2 , π_3 e π_4 são estatisticamente diferentes de zero, assim como a hipótese conjunta para $\pi_3 = \pi_4 = 0$ também foi rejeitada.

O fato de aceitarmos a hipótese $\pi_1 = 0$ implica que a série tem uma raiz unitária na frequência zero (tendência estocástica *não-sazonal*). O filtro apropriado nesse caso é $(1 - L)$, ou seja, apenas a primeira diferença (a aceitação dessa hipótese indica que a série é integrada de ordem 1). As conclusões foram semelhantes aquelas obtidas para a série de exportação.

Coeficiente	Valor auferido	Valor crítico	$\pi_i = 0$
π_1	-2,17	-3,49	Aceita
π_2	-3,29	-2,91	Rejeita
π_3	-4,04	-3,41	Rejeita
π_4	-3,92	-2,27	Rejeita
π_3, π_4	15,90	6,57	Rejeita

Tabela 11: Resumo dos valores críticos obtidos com o teste HEGY para dados trimestrais de importações

3.2 Para os dados Mensais

3.2.1 O Teste HEGY

$$\begin{aligned}
\hat{y}_8 = & -0,00171386 y_{1,t-1} - 0,0944327 y_{2,t-1} - 0,0405074 y_{3,t-1} - 0,212184 y_{3,t-2} \\
& \quad (-1,864) \quad (-5,393) \quad (-1,412) \quad (-7,794) \\
& - 0,166849 y_{4,t-1} - 0,187636 y_{4,t-2} + 0,0187007 y_{5,t-1} - 0,0778603 y_{5,t-2} \\
& \quad (-6,554) \quad (-7,468) \quad (1,101) \quad (-4,474) \\
& - 0,182437 y_{6,t-1} - 0,114244 y_{6,t-2} + 0,104459 y_{7,t-1} - 0,425437 y_{7,t-2} \\
& \quad (-7,979) \quad (-4,823) \quad (3,258) \quad (-9,681) \\
& + 0,000160396 trend - 0,00469782 d_{jun} - 0,0226334 d_{fev} + 0,138548 d_{mar} \\
& \quad (1,724) \quad (-0,086) \quad (-0,419) \quad (2,571) \\
& + 0,106073 d_{abr} + 0,147499 d_{mai} + 0,116313 d_{jun} + 0,171955 d_{jul} \\
& \quad (1,966) \quad (2,740) \quad (2,188) \quad (3,232) \\
& + 0,167504 d_{ago} + 0,0938359 d_{set} + 0,116989 d_{out} + 0,0934299 d_{nov} \\
& \quad (3,122) \quad (1,738) \quad (2,165) \quad (1,723) \\
& + 0,141250 d_{dez} + 0,0673080 y_{8,t-1} + 0,0267789 y_{8,t-2} + 0,0855158 y_{8,t-3} \\
& \quad (2,602) \quad (1,355) \quad (0,542) \quad (1,749) \\
& + 0,0271266 y_{8,t-4} - 0,0442583 y_{8,t-5} \\
& \quad (0,580) \quad (-1,031) \\
N = 557 \quad \bar{R}^2 = 0,7645 \quad F(29, 527) = 63,25 \quad \hat{\sigma} = 0,11217 \\
& \quad \text{(estatísticas-}t \text{ entre parênteses)}
\end{aligned}$$

em que *trend* é um termo de tendência; $d_{mês}$ são dummies sazonais (para os respectivos meses); $y_{1,t}$ até $y_{7,t}$ representam as variáveis independentes que são utilizadas para testar a existência de raiz unitária na frequência sazonal; e $y_{8,t-i}$ são valores defasados da variável a ser explicada.

Observando a Tabela 12, o teste-t para os parâmetros rejeita a maioria das hipóteses nulas (de que $\pi_i = 0$), exceto para dois casos, a saber: π_1 e π_4 . O teste-F conjunto para esses dois parâmetros também não consegue rejeitar a nula.

Como $\pi_1 = 0$, a série é integrada de ordem 1, ou seja, deve ser modelada utilizando a primeira diferença. Caso todas as outras fossem diferentes de zero, seria adequado modelar pela primeira diferença e *dummies* sazonais. Como π_4 não foi significativo (bem como o teste conjunto $\pi_1 = \pi_4 = 0$), podemos dizer que existe pelo menos uma raiz unitária sazonal.

Coefficiente	Valor auferido	Valor crítico	$\pi_i = 0$
π_1	-1,86	-3,24	Aceita
π_2	-5,39	-2,65	Rejeita
π_3	-7,79	-1,71	Rejeita
π_4	-1,41	-3,12	Aceita
π_5	-7,47	-2,99	Rejeita
π_6	-6,55	-3,12	Rejeita
π_7	-4,47	+0,12	Rejeita
π_8	+1,10	-0,43*	Rejeita
π_9	-4,82	-2,54	Rejeita
π_{10}	-7,98	-3,07	Rejeita
π_{11}	-9,68	-0,73	Rejeita
π_{12}	+3,26	-0,44*	Rejeita
π_1, π_4	2,82	5,63	Aceita
π_5, π_6	27,95	5,84	Rejeita
π_1, π_5	29,61	5,90	Rejeita
π_9, π_{10}	32,48	5,71	Rejeita
π_{11}, π_{12}	58,66	5,84	Rejeita

Tabela 12: Resumo dos valores críticos obtidos com o teste HEGY para dados mensais de importações

3.2.2 O ajustamento pelo X-12-ARIMA

O ajustamento foi feito pelo método multiplicativo, e calculado no "GiveWin", com o módulo X-12. Além dos ARIMA's padrões, estes outros modelos foram calculados: (1 1 1); (2 1 2); (2 1 1); (2 1 3); (3 1 1); (3 1 2); (3 1 3); (2 1 2)(1 0 1); (2 1 3)(1 0 1); (3 1 2)(1 0 1); (2 1 2)(1 1 1); (1 1 2)(1 0 1); (1 1 3)(1 0 1); (0 0 0)(2 1 2); (0 0 0)(1 1 1); (0 0 0)(1 1 1)₁₂; (0 0 0)(2 1 2)₁₂; (0 0 0)(1 0 1)₁₂.

Coeficientes	Modelos			
	(2 1 0)(0 1 1)	(3 1 2)(1 0 1)	(2 1 3)(1 0 1)	(2 1 2)(1 1 1)
AR não-sazonal (Lag 1)	-0,576	-1,415	-0,749	-0,635
(Lag 2)	-0,212	-0,957	-0,565	-0,220
(Lag 3)	∅	-0,272	∅	∅
MA não-sazonal (Lag 1)	∅	-0,848	-0,179	-0,061
(Lag 2)	∅	-0,296	-0,252	0,033
(Lag 3)	∅	∅	0,187	∅
AR sazonal	∅	0,969	0,972	0,131
MA sazonal	0,898	0,834	0,846	0,944
AICC	-805,35	-830,89	-832,23	-806,36
BIC	-788,11	-796,34	-797,68	-776,25

Tabela 13: Coeficientes dos modelos X-12-ARIMA

Alguns modelos foram rejeitada por apresentarem a estatística Ljung-Box menor do que 5%, (ou seja, evidência de autocorrelação serial), ou por evidência de super-diferenciação. Os modelos que passaram por todos os testes foram: $(2\ 1\ 0)(0\ 1\ 1)$, $(3\ 1\ 2)(1\ 0\ 1)$, $(2\ 1\ 3)(1\ 0\ 1)$ e $(2\ 1\ 2)(1\ 1\ 1)$.

Os testes F para presença de sazonalidade (paramétrico, não-paramétrico e “moving seasonality”), acusam sazonalidade com significância menor que 1%.

O modelo selecionado pelo programa foi $(2\ 1\ 2)(1\ 1\ 1)$, o mesmo escolhido para a série de exportações. Novamente, foi um resultado bastante interessante, pois ele está de acordo com as conclusões tiradas pelo teste de HEGY para dados mensais.² Os coeficientes estão apresentados na Tabela 13.

Data	Lim. Inferior	Previsão	Lim. Superior
2006.Nov	6.886	8.610	10.766
2006.Dez	6.827	8.704	11.096
2007.Jan	5.928	7.772	10.191
2007.Fev	5.444	7.361	9.952
2007.Mar	6.209	8.587	11.877
2007.Abr	5.739	8.124	11.500
2007.Mai	6.018	8.704	12.589
2007.Jun	5.921	8.738	12.894
2007.Jul	6.246	9.397	14.137
2007.Ago	6.406	9.815	15.039
2007.Set	5.998	9.353	14.586
2007.Out	6.195	9.826	15.586

Tabela 14: Previsão com base no modelo $(2\ 1\ 2)(1\ 1\ 1)$, intervalo de confiança de 5%

O quadro 14 mostra a estimativa para o futuro com base no modelo selecionado, e os respectivos intervalos de confiança. Em seguida temos esses valores representados graficamente, na figura 13, bem como todos os valores da série ajustados.

3.3 SARIMA

Como o teste de HEGY apontou que a série de importações poderia ser estudada em primeira diferença e com o auxílio de *dummies* sazonais, resolvi rodar o modelo selecionado pelo programa X-12-ARIMA, mas adicionando as variáveis *dummies* para cada mês e

²Onde foi visto que há evidências de que a série de importações é integrada na frequência sazonal e na frequência zero. Portanto o modelo esperado seria no formato $(p\ 1\ q)(P\ 1\ Q)$

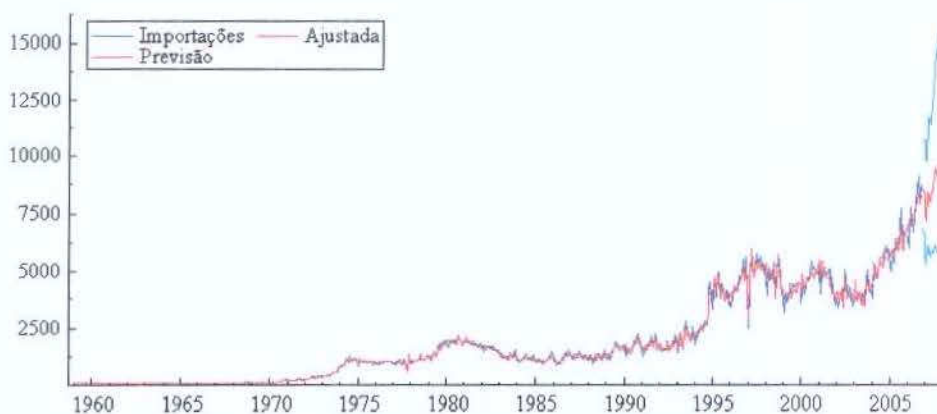


Figura 13: Série de Importações Ajustada pelo X-12 e I.C. das Previsões

também uma variável DU para dias úteis, que foi calculada através da diferença entre os dias úteis do mês e a média de dias úteis daquele ano.

Variável	Coefficiente	estatística-t	Variável	Coefficiente	estatística-t
ϕ_1	-0,9006	-6,441*	d_{mar}	-0,0110	-1,923
ϕ_2	-0,2722	-3,356*	d_{abr}	0,0288	4,800*
Φ_1	-0,0498	-1,101	d_{mai}	-0,0064	-1,340
θ_1	0,3089	2,125*	d_{jun}	0,0071	1,316
θ_2	-0,1022	-1,165	d_{jul}	-0,0321	-4,469*
Θ_1	-0,8049	-24,952*	d_{ago}	-0,0281	-3,631*
DU	0,0276	4,853*	d_{set}	0,0078	1,615
d_{jan}	-0,0138	-3,196*	d_{out}	-0,0099	-2,471*
d_{fev}	0,0486	5,029*	d_{nov}	0,0287	4,187*
AIC / BIC	-792,84	-706,76	d_{dez}	-0,0149	-3,268*

Tabela 15: Modelo (2 1 2)(1 1 1), com *dummies* e dia útil

os valores-t com um * indicam que o respectivo teste foi significativo ao nível de 5%. No entanto, o coeficiente autoregressivo sazonal não foi significativo, bem como a média móvel de segunda ordem. A variável nova introduzida, DU , foi bastante significativa e apresentou o coeficiente de 0,0276; vale dizer, para cada dia útil a mais (menos) que o mês tem em relação à média do ano, o valor das exportações será aproximadamente 2,8% maior (menor) que a média.

A tabela 16 e a figura 14 (destaque para o momento mais recente, de 1980 até 2006) mostram a qualidade do ajustamento. Para o período recente, o erro girou em torno de 5,5% do valor esperado (para mais ou para menos).

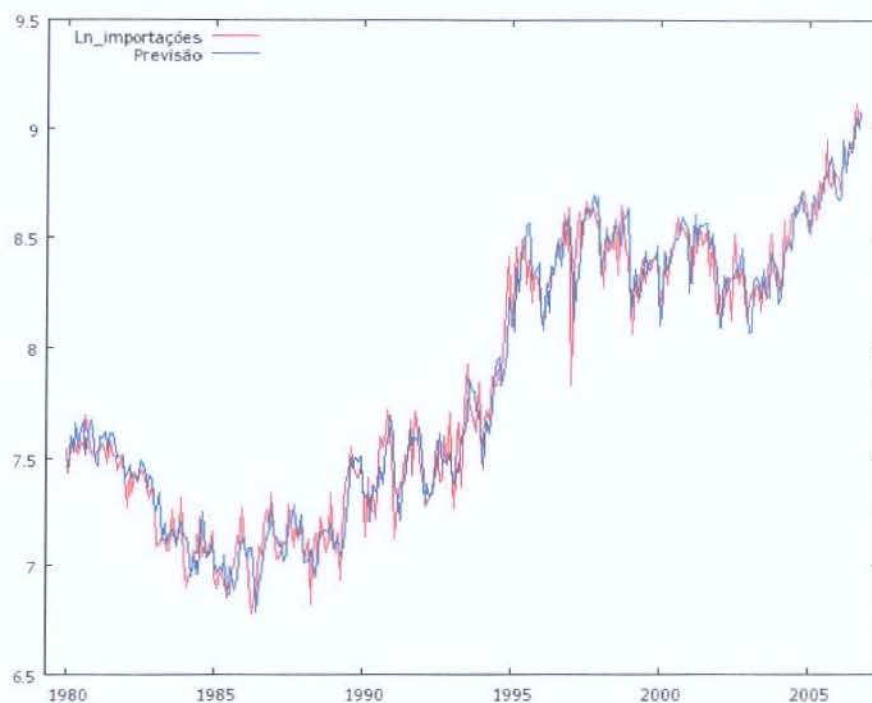


Figura 14: Série de Importações Ajustada pelo SARIMA

3.4 Modelo Estrutural

O modelo estrutural estimado a partir da equação abaixo pode ser representado como:

$$y_t = \text{tendência} + \text{sazonal}(\text{dummy}) + \text{irregular}$$

Data	Observado	Estimado	Data	Observado	Estimado
2004:09	5.753	5.646	2005:10	6.222	7.088
2004:10	5.840	6.072	2005:11	6.703	6.406
2004:11	6.083	5.555	2005:12	6.556	6.040
2004:12	5.686	5.543	2006:01	6.451	5.836
2005:01	5.257	5.052	2006:02	5.950	5.965
2005:02	4.980	5.160	2006:03	7.732	7.643
2005:03	5.909	6.016	2006:04	6.729	6.647
2005:04	5.332	5.779	2006:05	7.272	7.652
2005:05	6.369	5.639	2006:06	7.363	7.259
2005:06	6.171	5.947	2006:07	7.989	7.490
2005:07	6.057	6.515	2006:08	9.129	8.539
2005:08	7.690	6.517	2006:09	8.121	8.114
2005:09	6.314	6.793	2006:10	8.745	8.754

Tabela 16: Ajustamento da série de Importações pelo SARIMA

os gráficos que ilustram o comportamento do modelo para os componentes sazonais e de tendência estão representados nas figuras 15 e 16, respectivamente. Na década de 70 se inicia um movimento onde o padrão sazonal se torna menos intenso, mas que é sofre retrocesso no período atual.

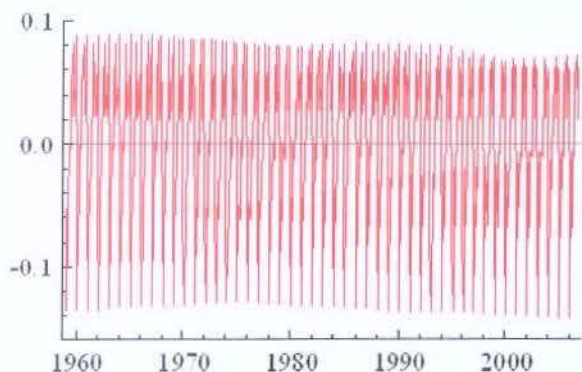


Figura 15: Componente de sazonalidade do modelo estrutura para importações

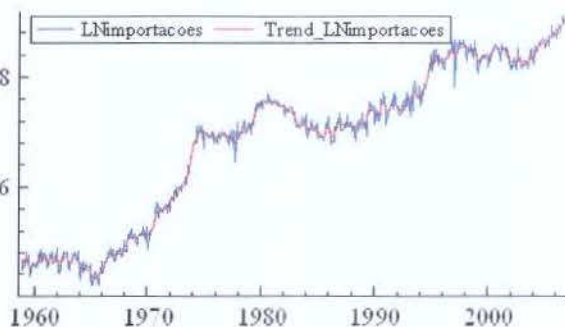


Figura 16: Componente de tendência do modelo estrutura para importações

Analogamente ao que aconteceu com o estudo para as exportações, o teste HEGY para dados trimestrais não apontou raiz unitária sazonal mas para os dados mensais sim. Isso pode ser devido à alterações da composição dentro do trimestre, informação essa que passa despercebida quando se observa os dados agregados de três em três meses.

Para tentar perceber quais foram essas mudanças, se é que existiram, chegou-se na Figura 17, que mostra a evolução do valor das *dummies* sazonais. De modo geral, quando o a reta cresce (decrease), significa que o mês em questão elevou (reduziu) sua importância/contribuição.³

Contrariamente ao que acontece com as exportações, o padrão sazonal das importações medido através de *dummies* é bastante estável, especialmente no caso do segundo e terceiro trimestres. Há duas mudanças importantes, uma no primeiro trimestre, em que o mês de Março ganha cada vez mais importância e o mês de Janeiro perde, praticamente na mesma proporção. A outra mudança importante ocorre no quarto trimestre, percebemos que o mês de Dezembro sofre uma redução expressiva, contrabalançado pelo aumento do mês de Outubro. Apesar de sutis, essas oscilações foram suficientes para tornar a série integrada na frequência sazonal.

³Se cada mês fosse responsável por $\frac{1}{12}$ da importação anual, as retas seriam horizontais no valor zero. Como há meses em que, historicamente, se importa mais (ou menos), o valor da *dummy* nesse período será maior (ou menor) do que zero. Por se tratar de um modelo estrutural, o coeficiente da *dummy* não é fixo e podemos, portanto, estudar a sua evolução.

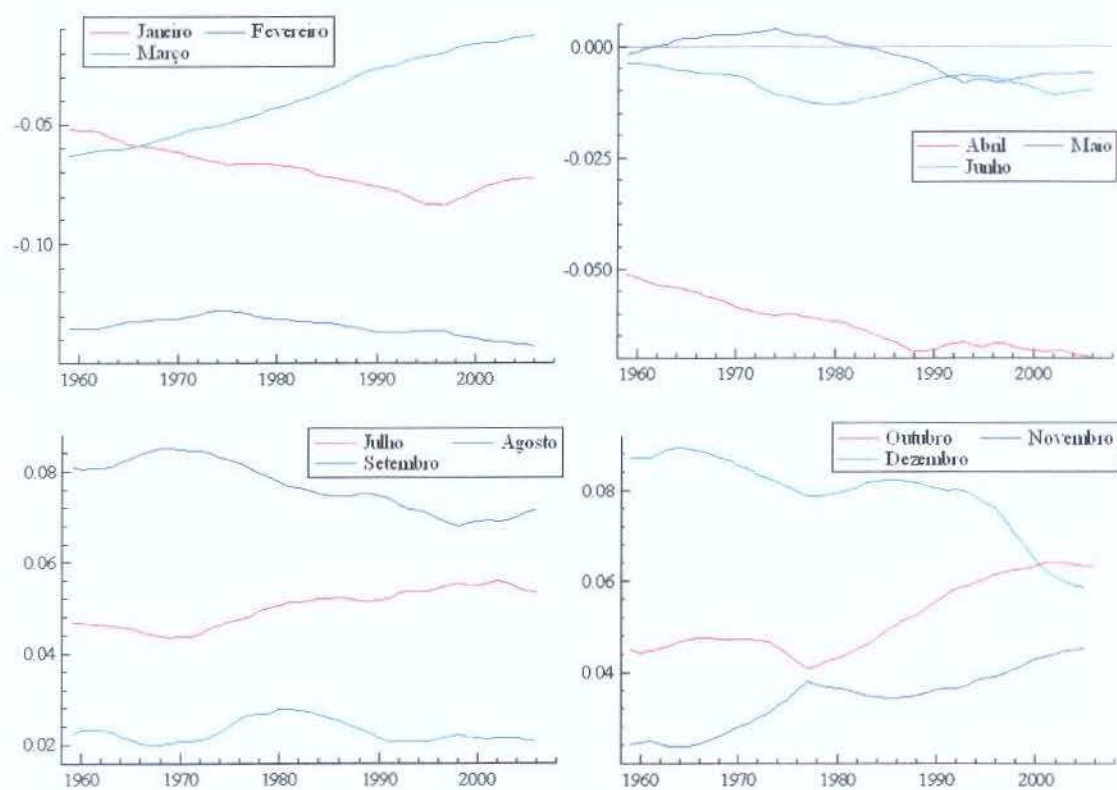


Figura 17: Coeficiente das Dummies Mensais no tempo, para o Modelo Estrutural das Importações

Referências

- DICKEY, D. A.; FULLER, W. A. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, v. 74, p. 427–431, 1979.
- ENDERS, W. *Applied Econometric Time Series*. [S.l.]: John Wiley & Sons.
- FRANSES, P. H. *Periodicity and Stochastic Trends In Economic Time Series*. Oxford: Oxford University Press, 1996.
- GRANGER, C. W. J.; KING, M. L.; WHITE, H. Comments on testing economic theories and the use of model selection criteria. *Journal of Econometrics*, Elsevier Science, v. 61, n. 1, p. 173–187, May 1995.
- GREENE, W. *Econometric Analysis*. Fifth. New Jersey: Pearson Education, 2003.
- GUJARATI, D. N. *Econometria Básica*. Terceira. São Paulo: Makron Books, 2004.
- HARVEY, A. *Time Series Models*. Cambridge: MIT Press, 1993.
- HYLLEBERG, S. *Modelling Seasonality*. Oxford: Oxford University Press, 1992.
- IBGE. *Série de Relatórios Metodológicos, Indicadores Conjunturais da Indústria*. Rio de Janeiro, 2004. v. 31.
- MADDALA, G. S. *Introduction to Econometrics*. Segunda. Englewood Cliffs: Prentice Hall, 1992.
- MADDALA, G. S.; KIM, I.-M. *Unit Roots, Cointegration, and Structural Change*. Cambridge: Cambridge University Press, 1998.
- MONTANES, A.; LOPES, A. C. B. da S. The behavior of hegy tests for quarterly time series with seasonal mean shifts. *Econometric Reviews*, Taylor & Francis, Inc., p. 83–108, 2005.
- PINYCK, R.; RUBINFELD, D. *Econometria - Modelos e Previsões*. [S.l.]: Elsevier, 2004.
- U.S. CENSUS BUREAU. *X-12-ARIMA Reference Manual*. Washington, 2002. Disponível em: <<http://www.census.gov/srd/www/x12a/>>.