

**UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE CIÊNCIAS NATURAIS E EXATAS
CURSO DE ESPECIALIZAÇÃO EM ESTATÍSTICA
E MODELAGEM QUANTITATIVA**

**ESTUDO DA OFERTA INTERNA DE ENERGIA E DO
CRESCIMENTO ECONÔMICO NO BRASIL
DE 1940 A 2009**

MONOGRAFIA DE ESPECIALIZAÇÃO

Lizandra Salau da Rocha

Santa Maria, RS, Brasil

2012

**ESTUDO DA OFERTA INTERNA DE ENERGIA E DO
CRESCIMENTO ECONÔMICO NO BRASIL
DE 1940 A 2009**

Lizandra Salau da Rocha

Monografia apresentada ao Curso de Especialização em Estatística e Modelagem Quantitativa, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do grau de **Especialista em Estatística e Modelagem Quantitativa**.

Orientador Prof. Dr. Adriano Mendonça Souza

Santa Maria, RS, Brasil

2012

Universidade Federal de Santa Maria
Centro de Ciências Naturais e Exatas
Curso de Especialização em Estatística e Modelagem Quantitativa

A Comissão Examinadora, abaixo assinada,
aprova a Monografia de Especialização

**Estudo da Oferta Interna de Energia e do Crescimento
Econômico no Brasil de 1940 a 2009**

elaborada por
Lizandra Salau da Rocha

como requisito parcial para obtenção do grau de
Especialista em Estatística e Modelagem Quantitativa

COMISSÃO EXAMINADORA:

Adriano Mendonça Souza, Dr.
Presidente/ Orientador

Roselaine Ruviaro Zanini, Dr. (UFSM)

Luciane Flores Jacobi, Dr. (UFSM)

Santa Maria, 28 de fevereiro de 2012.

RESUMO

Monografia de Especialização
Curso de Especialização em Estatística e Modelagem Quantitativa
Universidade Federal de Santa Maria

ESTUDO DA OFERTA INTERNA DE ENERGIA E DO CRESCIMENTO ECONÔMICO NO BRASIL DE 1940 A 2009

AUTORA: Lizandra Salau da Rocha

ORIENTADOR: Adriano Mendonça Souza

Data e Local da Defesa: Santa Maria, 28 de fevereiro de 2012.

A necessidade do uso de energia para mover os diversos setores dos países como o setor de transportes (aéreo, marítimo, terrestre), residencial, comercial e industrial evidencia uma relação direta entre o Crescimento Econômico e a Oferta Interna de Energia (OIE). Nesse contexto o presente estudo tem como objetivo medir as taxas de crescimento destas séries no período de 1940 a 2009 no Brasil, e analisar o comportamento delas por meio dos modelos Box-Jenkins. Para comparar as taxas de variação das séries no período estudado foram utilizados dois modelos de regressão simples (que medem a taxa de crescimento ou decréscimo) as quais se efetua o cálculo do teste de paralelismo de retas. Realizou-se também a modelagem Box-Jenkins para verificar o processo gerador de cada série. Os resultados mostram que o PIB teve uma taxa de crescimento no período analisado maior do que OIE, este resultado está de acordo com o esperado já que em meados de 2001 e 2002 houve déficit na oferta de energia no país. Além disso, verificou-se que o modelo mais adequado para explicar a OIE foi um ARIMA (1,1,0) e o PIB um ARIMA (1,1,1) estes modelos mostram os processos geradores das séries, indicando como elas devem ser analisadas ao longo do tempo para este conjunto de dados. Logo, estes resultados evidenciam que o Governo deve estar atento a infraestrutura do setor energético brasileiro evitando novos déficits na oferta de energia propiciando o equilíbrio entre a Oferta Interna de Energia e o Crescimento Econômico.

Palavras Chaves: Modelos Box-Jenkins. Oferta Interna de Energia. Produto Interno Bruto. Teste de Paralelismo.

ABSTRACT

Monograph of specialization
course in statistics and Quantitative Modeling
Universidade Federal de Santa Maria

STUDY OF INTERNAL SUPPLY OF ENERGY AND ECONOMIC GROWTH IN BRAZIL IN 1940 2009

AUTHOR: Lizandra Salau da Rocha

Advisor: Adriano Souza Mendonça

Date and place of defence: Santa Maria, February 28, 2012.

The need to use energy to move the various sectors of countries such as the transportation sector (air, sea, land), residential, commercial and industrial shows a direct relationship between economic growth and internal supply of energy (OIE). In this context, the present study aims to measure the growth rates of these series in 1940 to 2009 in Brazil, and analyze the behavior of them through Box-Jenkins models. To compare the rates of change of the series in the period studied were used two simple regression models (which measure the rate of increase or decrease) which performs the calculation of the test of parallel lines. There was also the modeling Box-Jenkins to verify the process each generator series. The results show that the GDP had a growth rate in the period considered larger than the OIE, this result is as expected since in mid-2001 and 2002 there was deficit in energy supply in the country. Furthermore, it was found that the model best suited to explain the OIE was an ARIMA (1, 1, 0) and the GDP an ARIMA (1, 1, 1) these models show the processes Series generators, indicating how they should be analyzed over time to this DataSet. Soon, these results show that the Government should be aware of the energy sector's infrastructure and avoiding new deficits in power supply providing the balance between internal supply of energy and economic growth.

Keywords: Box-Jenkins Models. Internal supply of energy. Gross Domestic Product. Parallelism test.

LISTA DE FIGURAS

Figura 1- Filtro linear	17
Figura 2- Comportamento das funções de autocorrelação (FAC) e autocorrelação parcial (FACP) de um modelo AR (1).	21
Figura 3- Comportamento da função de autocorrelação parcial e autocorrelação de um modelo MA.....	23
Figura 4- Representação da decomposição das Somas de Quadrados	36
Figura 5-Fluxograma do Método de Box-Jenkins.....	39
Figura 6- Fluxograma do teste de paralelismo.	40
Figura 7- Comportamento da Oferta Interna de Energia (1940 a 2009).	43
Figura 8- Comportamento da Série temporal PIB (1940 a 2009).....	43
Figura 9 - Correlograma da série em nível OIE.	44
Figura 10- Correlograma da série em nível PIB.....	45
Figura 11- Correlograma da série OIE em primeira diferença.....	46
Figura 12- Série do OIE em primeira Diferença	46
Figura 13 - Correlograma da série em primeira diferença PIB	47
Figura 14- Série do PIB em primeira diferença.....	47
Figura 15- FAC e FACP dos resíduos do modelo ARIMA (1,1,0).	50
Figura 16- FAC e FACP dos resíduos ao quadrado do modelo ARIMA (1,1,0).	50
Figura 17- Histograma e estatística descritiva dos resíduos do modelo ARIMA (1,1,0) para a série da Oferta Interna de Energia (OIE).....	51
Figura 18- Teste de Heterocedasticidade Condicional ARCH-LM para a Série OIE.	52
Figura 19- FAC e FACP dos resíduos do modelo ARIMA (1,1,1).	52
Figura 20- FAC e FACP dos resíduos ao quadrado do modelo ARIMA (1,1,1).	53
Figura 21- Histograma e estatística descritiva dos resíduos do modelo ARIMA (1,1,1) para a série da Produto Interno Bruto (PIB).....	53
Figura 22- Teste de Heterocedasticidade Condicional ARCH-LM para a Série PIB.	54

LISTA DE TABELAS

Tabela 1- Resultado dos testes de raiz unitária (ADF e KPSS) para as séries em nível.	45
Tabela 2- Resultado dos testes de raiz unitária (ADF e KPSS) para as séries em primeira diferença.	48
Tabela 3- Modelos concorrentes para a série da OIE que tiveram seus coeficientes significativos presença de ruído branco.	49

LISTA DE ABREVIATURAS E SIGLAS

AR	Modelo Autoregressivo
ARMA	Modelo Autoregressivo e de Médias Móveis
ARIMA	Modelo Autoregressivo Integrado de Médias Móveis
AIC	Critério de Informação Akaike
BIC	Bayesian Information Criterion
GEE	Gases do efeito estufa
NEEE	Núcleo de Estudos Estratégicos de Energia
OIE	Oferta Interna de Energia
PIB	Produto Interno Bruto

SUMÁRIO

1 INTRODUÇÃO	9
1.1 Tema da pesquisa	11
1.2 Justificativa e importância da pesquisa.....	11
1.3 Objetivos.....	12
1.3.1 Objetivo Geral	12
1.3.2 Objetivos Específicos	12
2 REFERENCIAL TEÓRICO	14
2.1 Análise de Séries Temporais.....	14
2.1.1 Modelos Autorregressivos e de Médias móveis	15
2.1.2 Testes de diagnóstico.....	27
2.2 Teste para igualdade das inclinações dos modelos	33
2.3 Comentários gerais do capítulo	37
3 MATERIAL E MÉTODOS	38
3.1 Comentários gerais do capítulo.....	40
4 ANÁLISE E DISCUSSÃO DOS RESULTADOS	42
4.1 Análise do comportamento das Séries Temporais: OIE e PIB.....	42
4.2 Modelagens das séries: OIE e PIB e seus respectivos testes	44
4.2.1 Modelagem das séries.....	48
4.2.2 Validação dos modelos elegidos.....	49
4.3 Teste para igualdade das inclinações	54
4.4 Comentários gerais do capítulo	57
5 CONSIDERAÇÕES FINAIS.....	59
REFERÊNCIAS BIBLIOGRÁFICAS	60

1 INTRODUÇÃO

A partir da Revolução Industrial evidencia-se que os recursos fósseis tais como, carvão mineral, gás natural e petróleo, dentre outros, formam a base da matriz energética no mundo. Pesquisas recentes demonstram que a queima destes combustíveis causam danos ao meio ambiente. Dentre os diferentes tipos de danos ambientais, salienta-se a alteração climática (efeito estufa) que tem sua causa na emissão de gases poluentes lançados na atmosfera a partir da queima de fontes energéticas poluidoras (fontes não renováveis).

As economias contemporâneas necessitam de energia para mover seus diversos setores como o setor de transportes (aéreo, marítimo, terrestre), residencial, comercial e industrial. Logo, evidencia-se uma relação direta entre o Produto Interno Bruto (PIB) e a Oferta Interna de Energia (OIE). Ao considerar o cenário de dependência energética há muitos países que possuem suas matrizes energéticas calcadas em fontes de energia fósseis. De maneira geral, as fontes de energia renováveis são aquelas que não poluem ou possuem um grau de poluição muito pequeno (energia hidráulica, lenha, produtos de cana-de-açúcar, etc.). Já fontes não renováveis ou energias fósseis são aquelas que causam danos significativos ao meio ambiente, em outras palavras, sua queima na obtenção de energia emite gases do efeito estufa (GEE) para a atmosfera.

Ao analisar a energia de forma agregada entende-se que a OIE é toda a energia necessária para movimentar a economia de um país num dado período de tempo. E o Crescimento Econômico habitualmente representado pelo PIB corresponde a toda riqueza gerada num país em determinado período, considerando todos os agentes econômicos (firmas, trabalhadores, empresas estatais, etc.). Assim, uma análise da OIE e do PIB possibilita averiguar a quantificação do grau de dependência da economia mundial em relação ao tipo e a quantidade de energia utilizada.

Conforme o Núcleo de Estudos Estratégicos de Energia¹ - NEEE (2011) o Brasil ocupa a 7ª posição quando o ranking compara o PIB. Em relação à OIE o país ocupa o nono lugar no ranking ficando atrás de países como os Estados Unidos, China, Rússia, Índia, Japão, Alemanha, França e Canadá. Já a geração de energia por meio de fontes renováveis faz com que o país ocupe a 4ª posição no ranking mundial de uso destas fontes na Oferta Interna de

¹ É um núcleo de pesquisas que trabalha junto a Secretaria de Planejamento e Desenvolvimento Energético do Ministério de Minas e Energia (MME). O MME é responsável por desenvolver ações estruturantes de longo prazo e apontar as potencialidades do setor energético para o estabelecimento e implementação de políticas setoriais formulando os princípios básicos e as diretrizes da política energética nacional.

Energia. Em contrapartida, o país está na 15ª posição de produção de energia por meio de fontes fósseis na OIE. Dessa maneira, evidencia-se a vantagem que o país tem em relação a outras economias quando o assunto é geração de energia com baixos índices de poluição (considerando a emissão de CO₂).

Apesar de o Brasil ter vantagem considerando a grande participação de energia renovável na OIE, o mesmo ocupa o 76ª lugar da OIE por habitante evidenciando que o país ainda possui uma distribuição de energia precária, ou seja, há grande parte da população que permanece sem acesso à energia (NEEE, 2011). Esta posição mostra como diferem os conceitos de Crescimento e Desenvolvimento Econômico sendo que o primeiro corresponde ao acúmulo de produção de bens e serviços e da renda e o segundo a distribuição homogênea desses bens e da renda propiciando o bem-estar da população.

Ao analisar o contexto energético e econômico brasileiro verifica-se a cada ano que o PIB está em expansão, necessitando em contrapartida de uma Oferta Interna de Energia maior para movimentar a economia. Esta realidade não se encontra circunscrita ao território brasileiro, em verdade há uma preocupação mundial com a relação aos impactos ambientais e o crescimento econômico principalmente quando se trata de países em desenvolvimento.

Avaliando este contexto, alguns pesquisadores passaram a investigar uma hipótese a qual é chamada de Curva de Kuznets Ambiental (CKA). Esta hipótese tenta mostrar que algumas medidas de danos ambientais aumentam nas primeiras fases do Crescimento Econômico, mas passam a decrescer assim que certo nível de renda for alcançado (GROSMANN; KRUEGER, 1991 e 1995). Há vários fatores que descrevem o formato descrito pela CKA (um “U” invertido) onde a trajetória ascendente dá espaço à trajetória negativamente inclinada. Segundo os autores Selden e Song (1994) a inversão da trajetória na curva pode ocorrer devido a: elasticidade renda positiva para o impacto ambiental, isto é, conforme a renda aumenta, os consumidores tendem a querer mais qualidade ambiental; mudança na composição da produção e do consumo; elevação da consciência das consequências da atividade econômica sobre o meio ambiente; aumento na rigidez dos regulamentos ambientais; e melhores tecnologias.

No entanto, essa hipótese parece se sustentar somente em algumas situações, outros autores como De Bruyn, Van Den Bergh e Opschoor (1998) acreditam que a CKA não se sustenta no longo prazo. Dessa maneira, o formato de “U” invertido seria apenas o início do processo pelo qual a economia passaria considerando a relação do Crescimento Econômico com os danos ambientais. O que estes autores alegam é que a partir de certo nível de renda haveria outro ponto de inflexão no qual a trajetória tornaria a ascender novamente, tomando

um formato de “N”. Assim, estes sugerem que em altos índices de crescimento econômico os danos ambientais voltam a aumentar no longo prazo.

Desse modo, torna-se relevante enfatizar que apesar da folga que o Brasil possua em relação a outros países possuindo sua matriz energética em maior grau de fontes não poluentes deve-se ter em mente que um acelerado aumento na demanda por energia para movimentar a economia pode acarretar no futuro em um aumento no uso de fontes energéticas não renováveis.

Logo, o estudo sobre o comportamento dessas variáveis ao longo do tempo: Oferta Interna de Energia e Produto Interno Bruto pode auxiliar nas decisões dos gestores responsáveis pelas políticas públicas do Brasil, tentando dirimir esforços capazes de manter a produção de energia limpa no futuro. Para tanto, realizar-se-á um estudo quantitativo para avaliar o comportamento dessas variáveis ao longo dos últimos 70 anos descrevendo os processos geradores da OIE e do PIB, bem como medir a variação dessas séries ao longo deste período.

1.1 Tema da pesquisa

Esta pesquisa tem como tema geral a busca da relação das taxas de variação da OIE e do PIB avaliando se há estabilidade econômica em termos de energia para os próximos anos.

1.2 Justificativa e importância da pesquisa

A Oferta Interna de Energia brasileira possui certa vantagem quando comparada com outros países, pois sua matriz energética tem uma participação muito grande de energia hidroelétrica e biomassa tornando-a maior parte formada por fontes renováveis. Contudo, um desequilíbrio entre o Crescimento Econômico e a infraestrutura energética pode levar países em desenvolvimento, como o Brasil a utilizarem maiores quantidades de fontes não renováveis (poluidoras) para suprir a demanda energética na produção de bens e serviços e com isso agravar os impactos ambientais.

Assim, o estudo acerca do comportamento das séries da Oferta Interna de Energia e do Produto Interno Bruto poderá propiciar o conhecimento para serem tomadas medidas no futuro, baseadas em suas taxas de variação ao longo do tempo. Em outras palavras, o conhecimento das taxas de variação das séries, bem como seus respectivos comportamentos poderá mostrar se a Oferta Interna de Energia está acompanhando o Crescimento Econômico ao longo do período estudado. É relevante notar que um desequilíbrio dessas taxas ao longo do tempo pode caracterizar uma possível falta de fornecimento de energia (déficit) e/ou sobra da mesma (superávit), assim as análises dessas taxas podem dar suporte às tomadas de decisões mostrando se as políticas governamentais estão dando suporte ao Crescimento Econômico do país em termos de energia.

1.3 Objetivos

1.3.1 Objetivo Geral

Estudar o comportamento evolutivo e a relação entre a Oferta Interna de Energia e o Produto Interno Bruto brasileiro no período de 1940 a 2009.

1.3.2 Objetivos Específicos

- verificar se a Oferta Interna de Energia apresenta a mesma taxa de variação que o crescimento econômico no período de 1940 a 2009 por meio do teste de paralelismo de retas;
- modelar as séries em estudo por meio dos modelos Box-Jenkins;

O próximo capítulo trará o estudo de séries temporais e da modelagem Box-Jenkins para mostrar a ferramenta de análise de séries temporais que será utilizada na presente pesquisa. No capítulo 3 serão abordados os Materiais e Métodos, mostrando a descrição das variáveis a serem utilizadas, bem como as etapas a serem seguidas para a análise das séries temporais. No capítulo 4 serão apresentados os resultados dos testes estatísticos e da

modelagem das séries temporais (OIE e PIB). E no capítulo 5 constarão as considerações finais da pesquisa.

2 REFERENCIAL TEÓRICO

Este capítulo será dividido em duas partes, a primeira mostrará os estudos de análise de séries temporais e a segunda parte fará referência ao teste de paralelismo de retas.

2.1 Análise de Séries Temporais

Segundo Morettin e Toloí (2004) uma série temporal é qualquer conjunto de observações ordenadas no tempo. Sendo que estas séries podem ser discretas (normalmente é obtida por meio de uma amostragem temporal contínua, ou seja, em intervalos de tempo iguais) ou contínuas (utiliza-se geralmente a amostragem em intervalos de tempo iguais, transformando a série de intervalar para discreta).

A análise de séries temporais propicia a investigação do mecanismo gerador da série temporal descrevendo o comportamento da série ao longo do tempo, realiza previsões de valores futuros, e/ou procura analisar períodos relevantes nos dados (WOLFF et al, 2011).

Em 1976 George Box e Gwilyn Jenkins criaram o método que se tornou conhecido mundialmente pelo Método de Box-Jenkins que é utilizado para manipular qualquer tipo de séries temporais, permitindo conhecer o processo gerador da série em estudo. O estudo da característica da série analisando os seus valores passados explica os valores futuros em função das suas autocorrelações sendo este o seu diferencial em relação a outras análises, como por exemplo, análise de regressão. De fato “a previsão é uma das principais razões da popularidade dos modelos Box e Jenkins, genericamente conhecidos por modelos *Auto Regressive Integrated Moving Averages- ARIMA*” (SOUZA et al, 2011, p. 237).

Com a modelagem das séries pode-se encontrar um modelo AR (Autorregressivo), MA (médias móveis), ARMA (Autorregressivo e de médias móveis), ou ainda, ARIMA (Autorregressivo Integrado e de médias móveis).

A análise de séries temporais vai além da previsão podendo ser utilizada para examinar os seus componentes tendência (T_t), sazonalidade (S_t), ciclo (C_t), *outliers*, quebras estruturais (MORETTIN; TOLOI, 1989):

- i. Tendência: é um movimento suave de crescimento ou decrescimento das observações.

- ii. Sazonalidade: ocorre quando ocorre um movimento ondulatório de curta duração, sendo este inferior a doze meses.
- iii. Ciclo: é um movimento ondulatório que tem longa duração, ou seja, é verificado ao longo de vários anos.

Outra pressuposição para o estudo de uma série temporal é que a série a ser trabalhada deva ser estacionária, isto é, a série deve se distribuir aleatoriamente em torno de uma média constante admitindo certa estabilidade ao longo do tempo. Dessa forma, garante-se que os parâmetros estimados para esta série sejam válidos para todo o período em estudo.

A utilização do método de Box-Jenkins necessita ser realizado em etapas, sendo que estas são altamente dependentes entre si, sendo então imprescindível realizar todos os testes estatísticos e as análises cabíveis ao processo para que o resultado possa ser utilizado para previsões.

1. **Identificação:** encontrar os valores adequados de p (número de parâmetros autorregressivos), q (número de parâmetros de médias móveis) e d (número de diferenciações da série caso seja não estacionária) por meio de correlogramas.
2. **Estimação:** após identificar os valores adequados de p e q estimam-se os parâmetros dos termos autorregressivos e de médias móveis incluídos no modelo.
3. **Verificação de diagnóstico:** depois de escolhido o modelo AR, MA, ARMA, ou ainda ARIMA e de estimados seus parâmetros, deve-se escolher o modelo mais adequado conforme os critérios de seleção e a análise dos resíduos. É exatamente por isso que muitos pesquisadores dizem que a modelagem ARIMA pelo método de Box- Jenkins é mais arte que ciência.

2.1.1 Modelos Autorregressivos e de Médias móveis

Uma série temporal é um conjunto numérico que produz uma sequência de dados distribuídos no tempo. Estes dados devem apresentar-se em tempos equidistantes, apresentando uma forte dependência serial. Devido a esta ordenação cronológica, surge o efeito da correlação entre as observações e deve recorrer à análise da função de autocorrelação e autocorrelação parcial para tratá-la.

Representamos uma série temporal da seguinte forma:

$$Z_1, Z_2, \dots, Z_n \quad \text{ou} \quad Z_t, \quad t = 1, 2, \dots, n.$$

onde n representa o tamanho da série.

Denomina-se um processo estocástico estacionário aquele no qual sua média e variância são constantes ao longo do tempo e quando o valor de sua covariância entre dois períodos de tempo depende apenas da distância da defasagem (GUJARATI, 2006).

$$\text{Média} \rightarrow E(Z_t) = \mu$$

$$\text{Variância} \rightarrow \text{Var}(Z_t) = E(Z_t - \mu)^2 = \sigma^2$$

$$\text{Covariância} \rightarrow \gamma_k = E[(Z_t - \mu)(Z_{t+k} - \mu)]$$

Onde γ_k é a covariância ou autocovariância na defasagem k (covariância entre os valores de Z_t e Z_{t+k}).

Em suma, quando uma série temporal é estacionária sua média, variância e covariância permanecem as mesmas não interessando qual período que forem analisadas, em outras palavras, elas não variam com o tempo. Assim, as flutuações da série estarão em torno da sua média (medidas pela variância) tendo então uma amplitude mais ou menos constante (GUJARATI, 2006).

A descrição dos modelos Box-Jenkins possuem uma notação própria que é utilizada na literatura, conforme se mostra a seguir:

- a. Operador de translação para o passado, denotado por B e definido por:

$$BZ_t = Z_{t-1}, B^m Z_t = Z_{t-m} \quad (1)$$

- b. Operador de translação para o futuro, denotado por F e definido por:

$$FZ_t = Z_{t+1}, F^m Z_t = Z_{t+m} \quad (2)$$

- c. Operador de diferença, já definido antes:

$$\Delta Z_t = Z_t - Z_{t-1} = (1 - B)Z_t \quad (3)$$

Sabendo que:

$$\Delta = 1 - B$$

d. Operador de Soma, denotado por S e definido por:

$$SZ_t = \sum_{j=0}^{\infty} Z_{t-j} = Z_t + Z_{t-1} + \dots = (1 + B + B^2 + \dots)Z_t \quad (4)$$

Sendo que:

$$SZ_t = (1 - B)^{-1}Z_t = \Delta^{-1}Z_t \quad (5)$$

Dessa forma tem-se que $S = \Delta^{-1}$.

Conforme Morettin e Tolo (2004) um processo linear geral pressupõe que uma série temporal é gerada a partir de um filtro linear (ou sistema linear), cuja entrada é um ruído branco, onde este se refere aos termos de resíduos não autocorrelacionados com média zero e variância constante, conforme ilustra Figura 1.

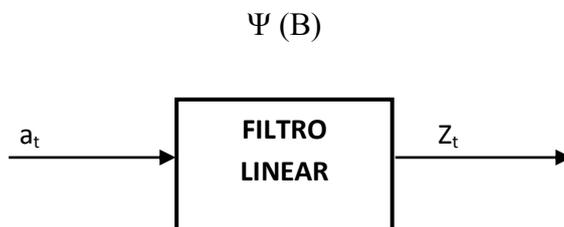


Figura 1- Filtro linear

Fonte: Morettin e Tolo (2004)

Assim tem-se que:

$$Z_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots = \mu + \psi(B)a_t \quad (6)$$

em que

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots \quad (7)$$

Sendo que $\psi(B)$ é chamada de função de transferência do filtro e μ é um parâmetro que determina o nível da série. Sabendo que Z_t descrito pela Equação 6 é um processo linear em que,

$$E(a_t) = 0, \forall t,$$

$$\text{Var}(a_t) = \sigma_a^2, \forall t,$$

$$E(a_t a_s) = 0, s \neq t.$$

Dessa forma chamando $\tilde{Z}_t = Z_t - \mu$, tem-se que:

$$\tilde{Z}_t = \psi(B)a_t \quad (8)$$

Com isso, se a sequência de pesos $\{\psi_j, j \geq 1\}$ for finita ou infinita e convergente, o filtro é estável (somável) e Z_t é estacionária. Neste caso, μ será a média do processo, em caso contrário, Z_t é não estacionária e μ não tem significado específico.

Tem-se então que a média de Z_t é dada conforme a Equação 9 descrita abaixo:

$$E(Z_t) = \mu + E\left(a_t + \sum_{j=1}^{\infty} \psi_j a_{t-j}\right) \quad (9)$$

Como $E(a_t) = 0$, para todo o t , então se tem $E(Z_t) = \mu$ se a série convergir $\sum_{j=1}^{\infty} \psi_j$. Lembrando que a variância de Z_t é dada por,

$$\text{Var}(Z_t) = \sigma_a^2 \sum_{j=0}^{\infty} \psi_j^2 \quad (10)$$

Com $\psi_0 = 1$, se $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ e a função de autocovariância γ_j de Z_t é dada por,

$$\gamma_j = \sigma_a^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+j} \quad (11)$$

Se $\sum_{i=0}^{\infty} \psi_i \psi_{i+j}$ existir, então para todo $j = 0, \pm 1, \pm 2, \dots$. Verifica-se que quando $j=0$ então, $\gamma_0 = \text{Var}(Z_t)$. Dessa forma, tem-se que a média e a variância da série Z_t são constantes e a covariância só vai depender de j , logo Z_t é estacionária (MORETTIN; TOLOI, 1981).

A série \tilde{Z}_t pode ser escrita de uma forma alternativa considerando a soma ponderada de valores passados $\tilde{Z}_{t-1}, \tilde{Z}_{t-2}, \dots$ e mais um ruído a_t :

$$\tilde{Z}_t = \pi_1 \tilde{Z}_{t-1} + \pi_2 \tilde{Z}_{t-2} + \dots + a_t = \sum_{j=1}^{\infty} \pi_j \tilde{Z}_{t-j} + a_t \quad (12)$$

Assim verifica-se que,

$$a_t = \left(1 - \sum_{j=1}^{\infty} \pi_j B^j \right) \tilde{Z}_t \quad \text{ou ainda,} \quad a_t = \pi(B) \tilde{Z}_t \quad (13)$$

Onde $\pi(B)$ é o operador:

$$\pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \dots \quad (14)$$

Observe que nas Equações (13) e (8) tem-se que:

$$a_t = \pi(B) \psi(B) a_t \quad (15)$$

$$\pi(B) = \psi^{-1}(B) \quad (16)$$

A relação descrita acima pode ser utilizada para obter os pesos π_j em função dos pesos de ψ_j e vice-versa.

Os Modelos Autorregressivos Integrados e de Médias Móveis (ARIMA) são lineares e trata-se de modelos univariados em que seus valores correntes são relacionados apenas com seus valores passados e/ou com os valores correntes e passados dos seus erros. Estes podem ser autorregressivos (AR), de médias móveis (MA), autorregressivos e de médias móveis (ARMA), ou ainda processos integrados mistos (ARIMA), estes são “modelos matemáticos que visam captar o comportamento da correlação seriada ou autocorrelação entre os valores da série temporal” (SOUZA et al, 2010, p. 6).

Modelo autorregressivo (AR)

O modelo autorregressivo mostra que o valor previsto de Z no período t é uma proporção de seu valor no período anterior ($t - 1$) acrescido de um choque aleatório no período t . Em outras palavras, um modelo autorregressivo depende de seus valores passados e dos erros aleatórios (HILL; GRIFFITHS; JUDGE, 1999).

Se $\pi_j = 0, j > p$, então se tem um modelo autoregressivo de ordem p , o qual é denotado por $AR(p)$, onde “ p ” indica a ordem do modelo, isto é, o número de defasagens. Na equação abaixo troca-se o nome dos pesos π_j para ϕ_j .

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + a_t \quad (17)$$

Definindo o operador autoregressivo estacionário de ordem p , logo se tem:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad (18)$$

Logo,

$$\phi(B) \tilde{Z}_t = a_t \quad (19)$$

Onde, p é o número de defasagens de Z_t e a_t é um ruído branco, sendo não autocorrelacionados, com média zero e variância constante σ^2 .

Os modelos de ordem 1 e 2 podem ser descritos nas equações abaixo.

$$\text{AR}(1) \rightarrow \tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + a_t \quad (20)$$

$$\text{AR}(2) \rightarrow \tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + a_t \quad (21)$$

Lembrando que $\tilde{Z}_t = Z_t - \mu$, em outras palavras a expressão corresponde aos valores de Z_t subtraídos da média do processo.

Considerando que $\tilde{Z}_t = \psi(B)a_t = (1 + \phi_1 B + \phi_2 B^2 + \dots) a_t$, tem-se então que,

$$\psi(B) = \sum_{j=0}^{\infty} \phi^j B^j = [\phi(B)]^{-1} = (1 - \phi B)^{-1} \quad (22)$$

Conforme a Proposição descrita em Morettin e Toloi (2004, p. 114) o processo será estacionário se $\psi(B)$ convergir para $|B| \leq 1$. Assim deve-se ter $|\phi| < 1$. A condição de estacionariedade do AR(p) estabelece que todas as raízes da equação devam cair fora do círculo unitário. Tendo a raiz da equação $\phi(B) = 1 - \phi B = 0$ e $B = \phi^{-1}$, assim diz-se que a raiz de $\phi(B) = 0$ e deve cair fora do círculo unitário. Em termos gerais, sendo o processo de ordem finita, será sempre inversível.

Portanto, as autocovariâncias não dependem do instante t, e sim da ordem p. Como $|\phi| < 1$ sendo a condição de estacionariedade, quanto maior o valor de p, ou seja, quanto maior a distância entre as observações, menor a autocovariância.

A identificação do modelo será feita através das funções de autocorrelação e autocorrelação parcial, que indicam a ordem do modelo.

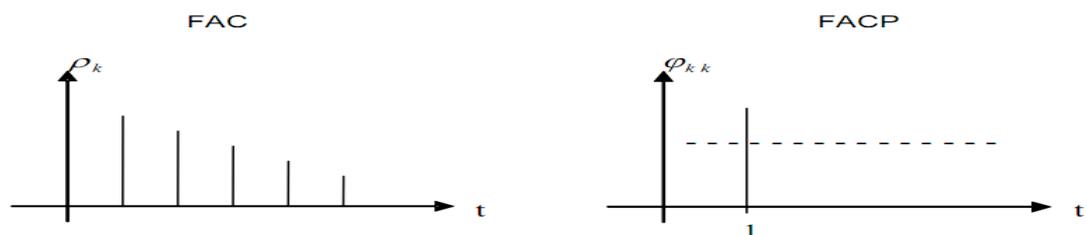


Figura 2- Comportamento das funções de autocorrelação (FAC) e autocorrelação parcial (FACP) de um modelo AR (1).

Fonte: SOUZA, 2006.

A Figura 2 mostra que a função de autocorrelação (FAC) decai exponencialmente (alternando ou não de sinal) e a autocorrelação parcial (FACP) apresenta um corte rápido no lag significativo, indicando a ordem p do modelo.

Processo de Média Móvel (MA)

Segundo Hill, Griffiths e Judge (1999) a média Móvel para um período escolhido consiste numa série de médias aritméticas dos períodos passados. Nestes modelos a variável dependente Z_t é escrita como função linear de um número finito de defasagens dos erros aleatórios não correlacionados. Supondo que $\psi_j = 0, j > q$, tem-se então um processo de médias móveis de ordem q .

$$Z_t = \mu + a_1 + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} \quad (23)$$

Lembrando que $\tilde{Z}_t = Z_t - \mu$, tem-se

$$\tilde{Z}_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t = \theta(B) a_t \quad (24)$$

Onde $\theta(B)$ é o operador de médias móveis de ordem q ,

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_p B^p \quad (25)$$

Logo,

$$\tilde{Z}_t = \theta(B) a_t \quad (26)$$

Onde, q é o número de defasagens das médias móveis e a_t = ruído branco, sendo que possui seus termos não autocorrelacionados, média zero e variância constante.

Os modelos de ordem 1 e 2 podem ser descritos nas equações abaixo.

$$\text{MA}(1) \rightarrow \tilde{Z}_t = a_t + \theta_1 a_{t-1} \quad (27)$$

$$\text{MA}(2) \rightarrow \tilde{Z}_t = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} \quad (28)$$

De acordo com Morettin e Toloi (2004, p. 114) “demonstra-se a proposição que evidencia um processo linear estacionário se $\theta(B) = 1 - \theta B$, tendo $\psi(B) = 1 - \theta B$ que é finito”.

$$a_t = [\theta(B)]^{-1} \tilde{Z}_t = \frac{1}{1 - \theta B} \tilde{Z}_t = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q) \tilde{Z}_t \quad (29)$$

Dessa maneira obtém-se a forma invertível por:

$$\tilde{Z}_t = -\theta \tilde{Z}_t - \theta^2 \tilde{Z}_t - \dots - \theta^q \tilde{Z}_t + a_t \quad (30)$$

Logo se $|\theta| < 1$, ou seja, a série $\pi(B) = \theta^{-1}(B)$ acima vai convergir para $|B| \leq 1$. Sendo equivalente a dizer que os zeros de $\theta(B) = 1 - \theta(B) = 0$ estão fora do círculo unitário.

A identificação do modelo é feita por meio das autocorrelações parciais, sendo que a função de autocorrelação fornece a ordem do modelo:

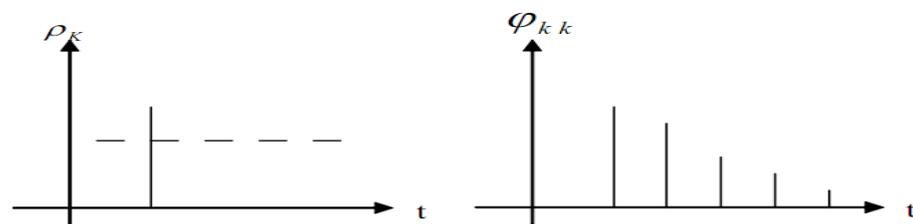


Figura 3- Comportamento da função de autocorrelação parcial e autocorrelação de um modelo MA.

Fonte: SOUZA, 2006.

Como mostra a Figura 3 a autocorrelação apresentará um corte rápido no lag significativo, indicando a ordem do modelo, e a autocorrelação parcial decairá exponencialmente evidenciando assim um processo de médias móveis.

Processo autorregressivo e de médias móveis (ARMA)

Um processo ARMA (p,q), ou seja, modelo autorregressivo e de médias móveis é representado por termos p e q autorregressivos e média móvel, respectivamente. Dessa maneira, o modelo fica mais complexo sendo explicado pelo termo autorregressivo e pelos termos do erro estocástico e de média móvel do termo de erro (MORETTIN & TOLOI, 2004).

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \dots + \phi_p \tilde{Z}_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (31)$$

Então de maneira compacta a equação 31 pode ser escrita como:

$$\phi(B)\tilde{Z}_t = \theta(B)a_t \quad (32)$$

Onde, p é o número de defasagens da parte autoregressiva do modelo e q corresponde ao número de defasagens das médias móveis e a_t são os resíduos aleatórios.

Os modelos de ordem 1 e 2 podem ser descritos nas Equações abaixo.

$$\text{ARMA}(1) \rightarrow \tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + a_t + \theta_1 a_{t-1} \quad (33)$$

$$\text{ARMA}(2) \rightarrow \tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} \quad (34)$$

Segundo Morettin e Toloi (2004) a condição de estacionariedade para um processo ARMA (p,q) é a mesma para um processo AR(p) e a condição de invertibilidade é a mesma que para um processo MA(q), já que conforme a equação (32), pode-se escrever:

$$\tilde{Z}_t = \psi(B)a_t = \theta(B)\theta^{-1}(B)a_t \quad (35)$$

Processo Autorregressivo Integrado e de Médias Móveis

Se uma série não é estacionária, podemos torná-la realizando a sua primeira diferença, ou seja, diferenciá-la até encontrar a estacionariedade “d” vezes. Se essa série for integrada de ordem um, ou seja, I(1) suas primeiras diferenças são I(0), isto é estacionárias. Do mesmo modo, se uma série temporal for I(d) depois de diferenciarmos d vezes obteremos uma série I(0). Dizemos que uma série temporal é ARIMA (p,d,q) isto é, uma série temporal autorregressiva integrada e de médias móveis (MORETTIN, 2006).

$$\Delta^d \tilde{Z}_t = (Z_t - Z_{t-1}) = \phi_1 Z_{t-1} + \theta_2 a_t + \theta_{t-1} \quad (36)$$

Processo estocástico não estacionário

Com frequência no estudo de séries temporais se encontram séries não estacionárias, dentre as quais se tem um exemplo clássico chamado de modelo de passeio aleatório, um exemplo de séries não estacionárias que seguem um modelo de passeio aleatório são os preços das ações e/ou as taxas de câmbio. Existem dois tipos: passeio aleatório sem deslocamento (isto é, sem o termo constante) e passeio aleatório com deslocamento (isto é, existe um termo constante).

Passeio aleatório sem deslocamento

Supondo que a_t é um ruído branco, ou seja, possui média zero e variância constante e os resíduos são não autocorrelacionados. Então se diz que a série Z_t é um passeio aleatório sem deslocamento se:

$$Z_t = Z_{t-1} + a_t \quad (37)$$

Observa-se que o valor de Z no período t é igual ao seu valor no período anterior ($t - 1$) mais um choque aleatório, sendo então um modelo autorregressivo AR(1). Então a partir da Equação 37 pode-se escrever:

$$\begin{aligned} Z_1 &= Z_0 + a_1 \\ Z_2 &= Z_1 + a_2 = Z_0 + a_1 + a_2 \\ Z_3 &= Z_2 + a_3 = Z_0 + a_1 + a_2 + a_3 \end{aligned}$$

De forma geral tem-se que:

$$Z_t = Z_0 + \sum a_t \quad (38)$$

Em consequência,

$$E(Z_t) = E\left(Z_0 + \sum a_t\right) = Z_0 \quad (39)$$

Verificando-se que:

$$Var(Z_t) = t \cdot \sigma^2 \quad (40)$$

Dessa forma, verifica-se que à medida que t aumenta, sua variância aumenta indefinidamente, tendo então a condição de estacionariedade violada. Outra característica deste modelo é que há uma persistência nos choques aleatórios, visto que, dado um choque inicial o mesmo não vai desaparecendo aos poucos, dessa maneira, o passeio aleatório sem deslocamento é comumente chamado de memória infinita.

Passeio aleatório com deslocamento

Supondo que a_t é um ruído branco, então se diz que a série Z_t é um passeio aleatório com deslocamento se:

$$Z_t = \delta + Z_{t-1} + a_t \quad (41)$$

Da mesma forma tem-se nesse modelo um AR(1) sendo que δ é conhecido como parâmetro de deslocamento, onde se verifica que Z_t se desloca para cima ou para baixo dependendo do sinal de δ como se observa a Equação 42.

$$Z_t - Z_{t-1} = \Delta Z_t = \delta + a_t \quad (42)$$

É possível mostrar que um passeio aleatório com deslocamento dá-se da seguinte forma:

$$E(Z_t) = Z_0 + t \cdot \delta \quad (43)$$

$$Var(Z_t) = t \cdot \sigma^2 \quad (44)$$

Verifica-se com as Equações 43 e 44 que tanto a média como as variâncias aumentam com o tempo evidenciando que a série não é estacionária. De forma análoga tem-se que um passeio aleatório sem deslocamento pode ser reescrito da seguinte forma:

$$Z_t = \rho Z_{t-1} + a_t \quad (45)$$

Onde ρ é o coeficiente de correlação entre as duas observações. Sabendo que o mesmo assume valores de $-1 \leq \rho \leq 1$. Se $\rho = 1$ se verifica o problema da presença de raiz unitária, entretanto se $\rho \leq 1$ verifica-se que a série é estacionária. Logo, raiz unitária e passeio aleatório com ou sem deslocamento são considerados sinônimos (GUJARATI, 2006).

2.1.2 Testes de diagnóstico

Existem muitos testes para verificar a estacionariedade de uma série temporal na literatura, no entanto, no presente estudo serão abordados apenas os mais utilizados nas pesquisas, iniciando pela análise gráfica, teste do correlograma e testes de raiz unitária.

Análise gráfica

As séries são plotadas contra o tempo verificando se as mesmas possuem algum tipo de tendência, sazonalidade ou ciclo e se as mesmas se distribuem aleatoriamente em torno de uma média.

Função de Autocorrelação e Correlograma

Este teste utiliza a função da autocorrelação permitindo identificar a ordem q de um processo MA e a função de autocorrelação parcial que propicia identificar a ordem p de um processo AR (BUENO, 2008). Essa função ρ_k é definida por k defasagens onde:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\text{covariância com defasagem } k}{\text{variância}} \quad (46)$$

Como a covariância e a variância têm a mesma unidade então ρ_k trata-se de um número puro e seu gráfico contra k corresponde ao correlograma populacional. No entanto, sabe-se que é muito difícil se trabalhar com a população, mas há a possibilidade de se ter uma amostra e a partir da mesma calcular as covariâncias e a variância amostral. Dessa forma, a função de autocorrelação amostral com defasagem k é:

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} \quad (47)$$

A escolha para o número de defasagem procede de um método empírico onde genericamente se calcula a função de autocorrelação até um terço ou um quarto da amostra, conforme Gujarati (2006). Posteriormente, para se analisar a significância estatística de ρ_k pode ser verificada pelo seu erro-padrão.

Em 1946, Bartlett mostrou que quando uma série temporal qualquer é puramente aleatória os seus coeficientes de autocorrelação da amostra $\hat{\rho}_k$ têm aproximadamente uma distribuição normal com média zero e variância igual a $1/n$. Entretanto, pode-se testar a significância estatística conjunta para os coeficientes de autocorrelação de que até determinada defasagem eles são concomitantemente iguais a zero. A estatística de Ljung e Box é definida por:

$$LB = n(n + 2) \sum_{k=1}^m \left(\frac{\hat{\rho}_k^2}{n - k} \right) \sim \chi^2 m \quad (48)$$

Como se pode notar a estatística de Ljung e Box segue a distribuição Qui-quadrado com m graus de liberdade.

Teste de Raiz Unitária de Dickey e Fuller e KPSS

Considerando o estudo do passeio aleatório descrito anteriormente descreve-se abaixo o princípio do teste de Dickey-Fuller tem-se:

$$Z_t = \rho Z_{t-1} + a_t \quad -1 \leq \rho \leq 1 \quad (49)$$

$$Z_t - Z_{t-1} = \rho Z_{t-1} - Z_{t-1} + a_t \quad (50)$$

Então tem-se que:

$$\Delta Z_t = (\rho - 1)Z_{t-1} + a_t \quad (\rho - 1) = \delta \quad (51)$$

Dessa forma, o teste de Dickey-Fuller (DF) tem as seguintes hipóteses a serem analisadas a hipótese nula a série é não estacionária e em contrapartida, a hipótese alternativa diz que a série é estacionária.

$$\mathbf{H}_0: \delta = 0 \text{ ou } \rho = 1$$

$$\mathbf{H}_1: \delta \neq 0 \text{ ou } \rho \leq 1$$

O teste é simples de ser realizado bastando tomar as primeiras diferenças de Z_t e fazer a regressão destas em relação à Z_{t-1} e verificar se o coeficiente angular é zero ou não. Se o coeficiente for igual a zero, então se tem a presença de uma raiz unitária, ou seja, a série não é

estacionária. Caso contrário, se o coeficiente tiver um valor negativo concluí-se que a série é estacionária.

É relevante notar que este teste não pode utilizar a estatística de “t” de Student para verificar a significância dos parâmetros, pois Z_{t-1} não segue distribuição normal assintótica. Dessa forma, Dickey e Fuller calcularam uma estatística τ (tau), os mesmos autores calcularam os valores críticos com base em simulações de Monte Carlo (GUJARATI, 2006).

Há na literatura uma variação do teste citado acima, pois o mesmo tem o pressuposto de que os erros são não correlacionados, por isso os mesmos autores desenvolveram um teste que considera a presença de autocorrelação entre os resíduos. Este teste é chamado de teste de Dickey-Fuller aumentado (ADF), as hipóteses deste teste são as mesmas do teste DF este consiste em estimar a regressão abaixo:

$$\Delta^d Z_t = \mu + \beta_1 t + \delta Z_{t-1} + \sum_{i=1}^m \alpha_i \Delta Z_{t-i} + a_t \quad (52)$$

Sendo que o número de defasagens é determinado empiricamente, tendo a ideia de incluir o número de termos que seja suficiente para que o resíduo não apresente autocorrelação (GUJARATI, 2006). A decisão do teste continua sendo a mesma do DF basta analisar o coeficiente δ para concluir se a série é ou não estacionária.

Segundo Morettin (2006) há outros testes alternativos para a verificação de raiz unitária já que o teste ADF e DF são de baixo poder, ou seja, há uma elevada probabilidade de se cometer o Erro Tipo II (não rejeitar a hipótese nula quando ela é falsa). Para tentar aumentar o poder dos testes de estacionariedade surge então o teste KPSS desenvolvido por Kwiatkowski, Phillips, Schmidt e Shin em 1992, o mesmo tem como hipótese nula a estacionariedade da série e a hipótese alternativa é de não-estacionariedade, a ideia dos autores é de utilizar este teste como um complemento ao teste ADF, DF ou de Phillips-Perron (BUENO, 2008).

Para entender a mecânica deste teste estima-se a regressão abaixo:

$$Z_t = \varphi_t + a_t \quad (53)$$

Sendo que φ_t é um passeio aleatório sem deslocamento, $\varphi_t = \varphi_{t-1} + a_t$ sabendo que $a_t \sim \text{i.i.d.}(0, \sigma^2)$ e a_t é um processo estacionário. Neste caso a hipótese a ser testada é de que $H_0: \sigma^2 = 0$ e $H_1: \sigma^2 > 0$.

Conforme Bueno (2008) descreve pode-se acrescentar uma tendência determinística no modelo:

$$Z_t = \mu + \beta t + \varphi_t + a_t \quad (54)$$

Então posteriormente o teste utiliza-se de quatro etapas:

1. Estima-se a série em análise contra as variáveis determinísticas;

$$Z_t = \mu + \beta t + a_t \quad (55)$$

2. Calculam-se os resíduos da regressão acima;

$$a_t = \mu + \hat{\beta}t + \hat{\varphi}_t \quad (56)$$

3. Faz-se a soma parcial dos resíduos;

$$S_t = \sum_{j=1}^t \hat{a}_j \quad (57)$$

4. Então se utiliza o teste de multiplicador de Lagrange;

$$KPSS = \sum_{t=1}^T \frac{S_t^2}{T^2 \hat{\sigma}^2} \quad (58)$$

Onde $\hat{\sigma}^2$ é a variância de longo prazo. Dessa forma, se a série Z_t for estacionária, então a soma parcial dos resíduos será não estacionária e o numerador do teste KPSS será um estimador da variância da soma parcial dos resíduos. Entretanto, se Z_t for não estacionária o numerador crescerá sem limites.

Os testes descritos acima são utilizados juntos primeiro calculando-se o teste de Dickey-Fuller e Phillips-Perron, e depois confirmando seus resultados com o teste de KPSS, já que o segundo é mais poderoso que os primeiros.

Realizada essa etapa de verificação da estacionariedade da série e realizando as transformações necessárias para torná-la estacionária quando possível, então se pode encontrar por meio do correlograma os indicativos de qual é o modelo que melhor se ajusta aos dados podendo ser autoregressivo, de médias móveis, ou ainda ambos.

Critério de penalização ou seleção de modelos (ARIMA)

A identificação é a etapa mais difícil da metodologia de Box-Jenkins, visto que há casos em que não se consegue identificar um único modelo, e sim vários modelos concorrentes geradores da série em estudo.

Muitos pesquisadores utilizam critérios de seleção de modelos construídos com base na variância estimada a_t , no tamanho da amostra e nos valores de p e q . Os mais utilizados são o critério Akaike (AIC) e o critério *Bayesian information criterion* (BIC). Em outras palavras, este critério utiliza a variância estimada da amostra e leva em consideração o número de parâmetros utilizados no modelo. Defini-se algebricamente este critério por:

$$AIC = \ln \hat{\sigma}_{p,q}^2 + \frac{2(p+q)}{n} \quad (59)$$

$$BIC = \ln \hat{\sigma}_{p,q}^2 + \frac{(p+q) \ln(n)}{n} \quad (60)$$

Onde p e q são os parâmetros do modelo ARMA (p, q), n é o tamanho da amostra e $\hat{\sigma}_{p,q}^2$ é uma estimativa da variância residual. O critério de escolha entre os modelos concorrentes deve se optar pelo modelo que minimiza estes critérios penalizadores, ou seja, escolhe-se o modelo que tiver o menor valor de AIC e BIC.

2.2 Teste para igualdade das inclinações dos modelos

O teste de igualdade das inclinações permite comparar as taxas de variação, ou as inclinações de dois modelos ou grupos de observações podendo ser utilizada para encontrar quebra estrutural no modelo. Este teste utiliza as variâncias das duas amostras em questão e suas respectivas somas de quadrados dos resíduos (SQE), além da soma dos quadrados dos desvios das variáveis exploratórias (S_{xx}) (CHARNET et al, 1999).

Inicialmente ajusta-se um modelo de regressão linear da variável de interesse em função das demais variáveis explicativas, que podem ser a própria evolução temporal. Deve-se atentar que as variáveis explicativas devem ser iguais para ambos os modelos a serem estudados. O modelo 1 e 2 são apontados nas equações 61 e 62, respectivamente.

$$Z_{i1} = \beta_{01} + \beta_{11}X_{i1} + a_{it} \quad (61)$$

Sendo que $i = 1, 2, 3, \dots, n_1$

$$Z_{i2} = \beta_{02} + \beta_{12}X_{i2} + a_{it} \quad (62)$$

Sendo que $i = 1, 2, 3, \dots, n_2$ e β_0 refere-se ao termo de intercepto da regressão; β_1 refere-se ao coeficiente da tendência das séries e a_t é o resíduo da regressão.

Uma suposição necessária para que se possam comparar dois modelos é que os resíduos referentes aos modelos sejam independentes e homocedásticos, em outras palavras seus erros devem ser independentes e com variância comum para todas as observações (CHARNET, et al, 1999).

Ajustados os modelos descritos anteriormente tem-se o valor do intercepto representativo do modelo descrito na Equação 63 e na Equação 64 o respectivo coeficiente angular.

$$\hat{\beta}_{01} = \bar{Z}_{i1} + \hat{\beta}_{11}X_{i1} \quad (63)$$

$$\hat{\beta}_{i1} = \frac{S_{xy}^1}{S_{xx}^1} \quad (64)$$

E o valor do intercepto representativo do modelo 2 encontra-se na equação 65 e na 66 o coeficiente angular.

$$\hat{\beta}_{02} = \bar{Z}_{i2} + \hat{\beta}_{12}X_{i2} \quad (65)$$

$$\hat{\beta}_{12} = \frac{S_{xy}^2}{S_{xx}^2} \quad (66)$$

Sabendo que SQE corresponde à soma de quadrados dos resíduos. Para os dois modelos tem-se que $k = 1$ e 2 , então:

$\hat{\beta}_{0k}$: corresponde a estimativa do intercepto da reta do grupo k ;

$\hat{\beta}_{1k}$: corresponde a estimativa da inclinação da reta do grupo k ;

\hat{Z}_k : é a média dos valores de Z para o modelo k ;

\hat{X}_k : é a média dos valores de X para o grupo k ;

S_{xz}^k : é a soma dos produtos cruzados dos desvios de X e Z para o grupo k ;

S_{xx}^k : é a soma dos quadrados dos desvios de X para o grupo k ;

Segundo Charnet et al (1999) conforme as suposições dos Modelos de Regressão Linear Simples para $k = 1$ e 2 ,

$$\hat{\beta}_{0k} \sim N\left(\hat{\beta}_{0k}; \sigma^2 \left[\frac{1}{n_k} + \frac{\bar{X}_k^2}{S_{xx}^k} \right]\right) \quad (67)$$

$$\hat{\beta}_{1k} \sim N\left(\hat{\beta}_{1k}; \frac{\sigma^2}{S_{xx}^k}\right) \quad (68)$$

Considerando que as duas amostras são independentes, logo $\hat{\beta}_{0k}$ e $\hat{\beta}_{1k}$ também são independentes, assim como os respectivos coeficientes angulares. Desta forma pode-se realizar um teste para comparação de suas inclinações.

Teste para igualdade das inclinações

Para se comparar as taxas de variação, ou as inclinações nos dois modelos testam-se as seguintes hipóteses:

$$\mathbf{H}_0: \beta_{11} = \beta_{12}$$

$$\mathbf{H}_1: \beta_{11} \neq \beta_{12}$$

Será utilizada a estatística:

$$T_{\beta_1} = \frac{\hat{\beta}_{11} - \hat{\beta}_{12}}{\sqrt{S_{\beta_1}}} \quad (69)$$

Onde S_{β_1} é a variância amostral;

$$S_{\beta_1} = \frac{SQE_{(1)} + SQE_{(2)}}{n_1 + n_2 - 4} \left[\frac{1}{S_{xx(1)}} + \frac{1}{S_{xx(2)}} \right] \quad (70)$$

Lembrando que as somas dos quadrados dos desvios de X são calculadas conforme a seguir (onde x é o valor de cada observação da variável independente):

$$S_{xx} = \sum x^2 - \frac{(\sum x)^2}{n} \quad (71)$$

Já a soma dos produtos cruzados dos desvios de X e Z (S_{xy}) calcula-se da seguinte forma:

$$S_{xz} = \sum x \cdot z - \frac{\sum x \cdot \sum z}{n} \quad (72)$$

A soma dos quadrados dos desvios de Z pode ser calculada pela Equação abaixo:

$$S_{zz} = \sum z^2 - \frac{(\sum z)^2}{n} \quad (73)$$

Para entender a origem da soma de quadrados dos resíduos é necessário entender que a variação de Z (variável dependente) compreende três tipos de desvios:

- i. Desvios totais ($Z_i - \bar{Z}$) o qual corresponde aos valores observados de Z menos a média da variável Z ;
- ii. Desvios explicados ou de regressão ($\hat{Z}_i - \bar{Z}$) correspondendo aos valores estimados de Z menos a média da variável Z ;
- iii. Desvios não explicados ou resíduos ($Z_i - \hat{Z}_i$) correspondem aos valores observados de Z subtraídos dos valores estimados de Z ;

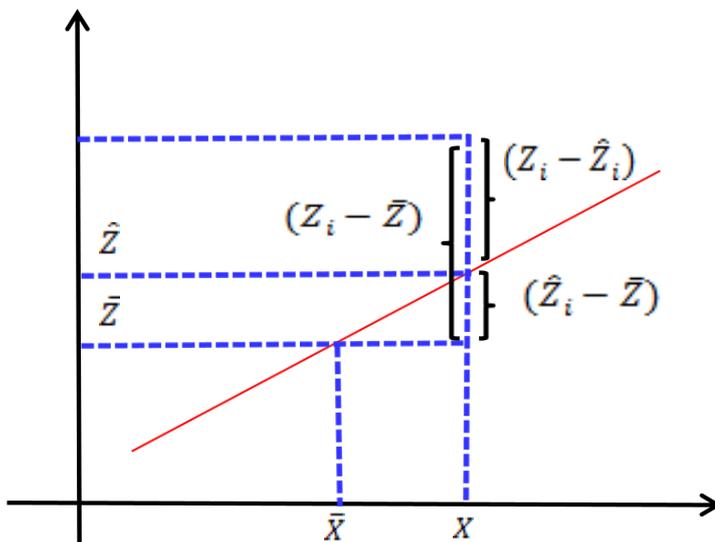


Figura 4- Representação da decomposição das Somas de Quadrados

Assim, entende-se que a soma de quadrado total é dada por:

$$SQ_{\text{Total}} = SQ_{\text{Regressão}} + SQ_{\text{Resíduo}}$$

$$\Sigma(Z_i - \bar{Z})^2 = \Sigma(\hat{Z}_i - \bar{Z})^2 + \Sigma(Z_i - \hat{Z}_i)^2$$

Para entender a origem das somas dos quadrados pode-se encontrar o valor da soma de quadrados dos resíduos pela Equação 73:

$$SQE = S_{zz} - b \cdot S_{xz} \quad (74)$$

Dessa forma, evidencia-se que o teste de igualdade das inclinações mostra se dois modelos possuem o mesmo comportamento ao longo do tempo, isto é, se estes crescem e/ou decrescem juntos durante o mesmo período. Em outras palavras, este teste de igualdade das inclinações mostra se as taxas de variação são iguais para os dois modelos no período analisado.

A estatística utilizada neste teste segue a distribuição t de Student com $(n_1 + n_2 - 4)$ graus de liberdade para a hipótese nula de que há igualdade da inclinação dos dois modelos. Dessa forma, quando a estatística calculada excede a tabulada tem-se que as taxas de variações dos dois modelos no período analisado diferem estatisticamente.

2.3 Comentários gerais do capítulo

Neste capítulo foi exposta a metodologia de séries temporais evidenciando passo a passo a metodologia de Box-Jenkins. Primeiramente, conceituou-se uma série temporal, em seguida foram demonstrados os comportamentos e os pressupostos esperados para a realização de uma série temporal. O estudo do comportamento da série evidencia as características da mesma, bem como as transformações necessárias para que se possa vir a fazer caso sejam necessárias.

Para o estudo das séries temporais há testes estatísticos que dão suporte para a rejeição de modelos e/ou características inapropriadas que são gerados pelos conjuntos de dados. Após a verificação da estacionariedade parte-se então em busca do melhor modelo sendo este capaz de descrever o comportamento mais fidedigno da série no período estudado.

No final do capítulo foi demonstrado como se realiza um teste de inclinação para comparar dois modelos de regressão. O mesmo mostra se os dois modelos têm a mesma taxa de variação durante um período estudado. No próximo capítulo será abordada a metodologia utilizada constando a descrição de como a pesquisa exploratória de dados foi realizada, bem como a descrição das séries em estudo e um fluxograma mostrando as etapas realizadas no presente trabalho na análise das séries temporais elegidas.

3 MATERIAL E MÉTODOS

A realização desta pesquisa conta com dados temporais da série da Oferta Interna de Energia (OIE) que foram obtidas no site do Ministério de Minas e Energia (MME) onde anualmente é publicado o relatório oficial do Balanço Energético Nacional (BEN) destinado a registrar estatísticas da Matriz Energética Nacional. Também foi utilizada a série do Produto Interno Bruto (PIB) a qual está disponível no site do Instituto de Pesquisa Econômica Aplicada (IPEA). As duas séries utilizadas na presente pesquisa têm 70 observações anuais que iniciam no ano de 1940 e vão até o ano de 2009.

A série da OIE tem sua medida conforme consta no Balanço Energético Nacional (BEN, 2010) em tonelada equivalente de petróleo (tep) esta é uma medida de energia definida como o calor liberado na combustão de 1 tonelada de petróleo cru. Já a série do PIB é medida em milhões de reais a preços de 2009.

Segundo o Ministério de Minas e Energia, a OIE é toda a energia necessária para movimentar a economia brasileira num dado período. Esta série é medida anualmente e é composta por recursos fósseis (petróleo, gás natural e seus derivados, carvão mineral e seus derivados, dentre outros) e renováveis (hidráulica, lenha, carvão vegetal e outras fontes renováveis). O Ministério de Minas e Energia também inclui nesta série todo o tipo de energia que o Brasil importa para montar sua Oferta de Energia.

A primeira análise realizada neste estudo foi a observação das séries ao longo do tempo. Para verificação do comportamento das mesmas utilizou-se gráficos das séries contra o tempo e verificou-se se existia tendência ascendente, ou descendente, ou ainda se havia presença de sazonalidade ou ciclo nas séries. Esta etapa de análise do comportamento das séries foi relevante, pois para se realizar a previsão das mesmas faz-se necessário que se conheça as características das séries para que se possa tentar corrigir estes efeitos antes de realizar a previsão.

A realização da análise de séries temporais utilizando as séries da Oferta Interna de Energia (OIE) e do Produto Interno Bruto (PIB) demandou uma sequência lógica necessária para o conhecimento dos processos geradores das séries.

As etapas metodológicas estão descritas por meio de um fluxograma Figura 5, a descrição expandida dessas etapas foram apresentadas no item 2.1 do Capítulo 2.

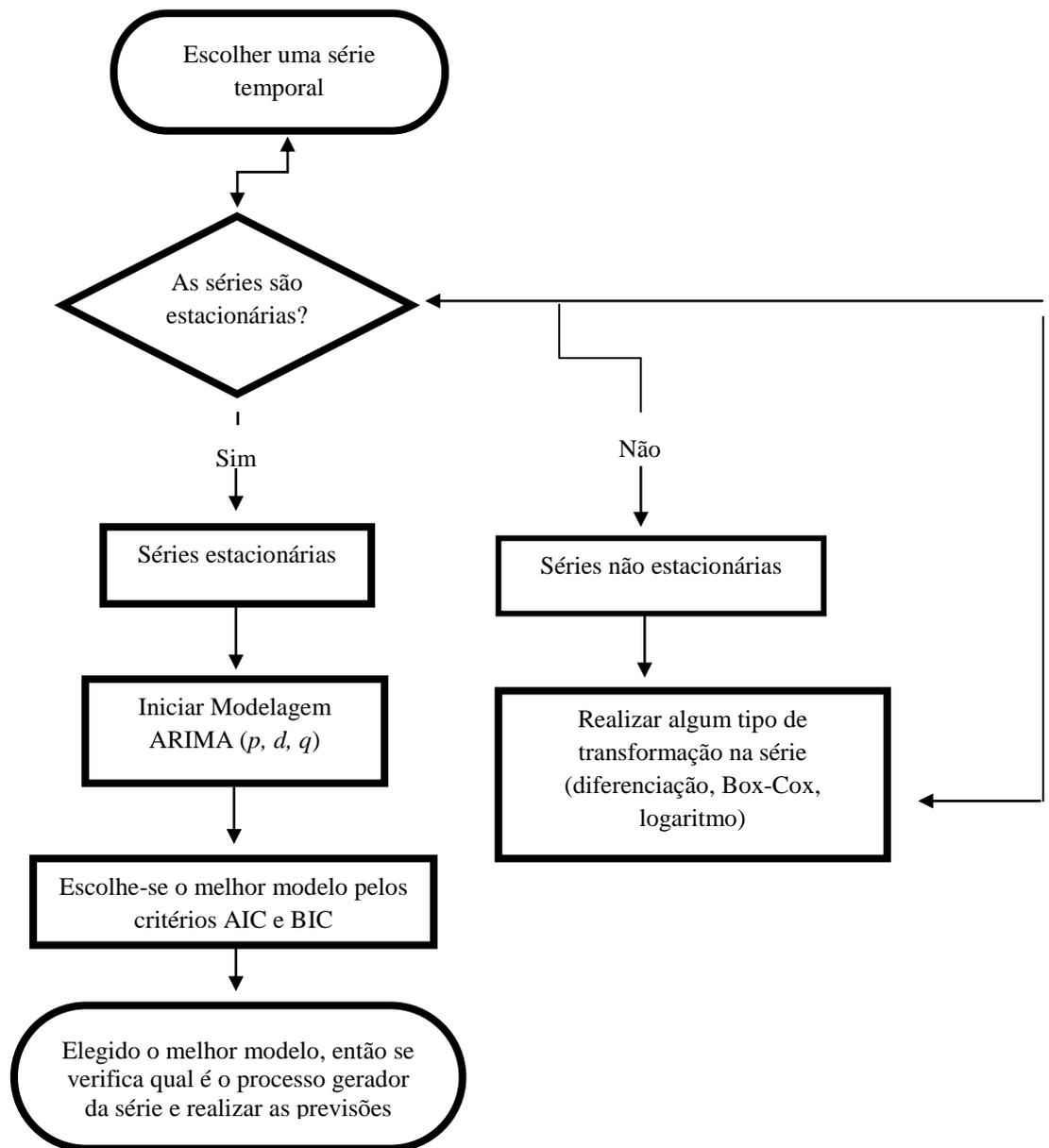


Figura 5-Fluxograma do Método de Box-Jenkins.

A escolha do modelo mais adequado para cada série realizou-se por meio da utilização de alguns critérios utilizados para decidir qual o modelo mais adequado a ser adotado. Foi realizada uma investigação para ver se os resíduos estimados eram ruído branco e após utilizou-se os critérios de AIC e BIC. Foi utilizado o modelo com os menores valores desses critérios seguindo o princípio da parcimônia e a significância dos parâmetros para descrever os comportamentos das séries. Para a análise das séries temporais foi utilizado o *Software Eviews 5.0*.

Em seguida, utilizou-se no estudo o teste de inclinação de dois modelos (teste do paralelismo de retas), para a realização deste teste foi utilizado o Excel para a obtenção das regressões simples. Este teste permite verificar se duas séries ou variáveis possuem as mesmas taxas de variação ao longo do tempo conforme apresentado no item 2.2 do capítulo 2. Para realizar este teste as etapas são descritas no fluxograma abaixo.

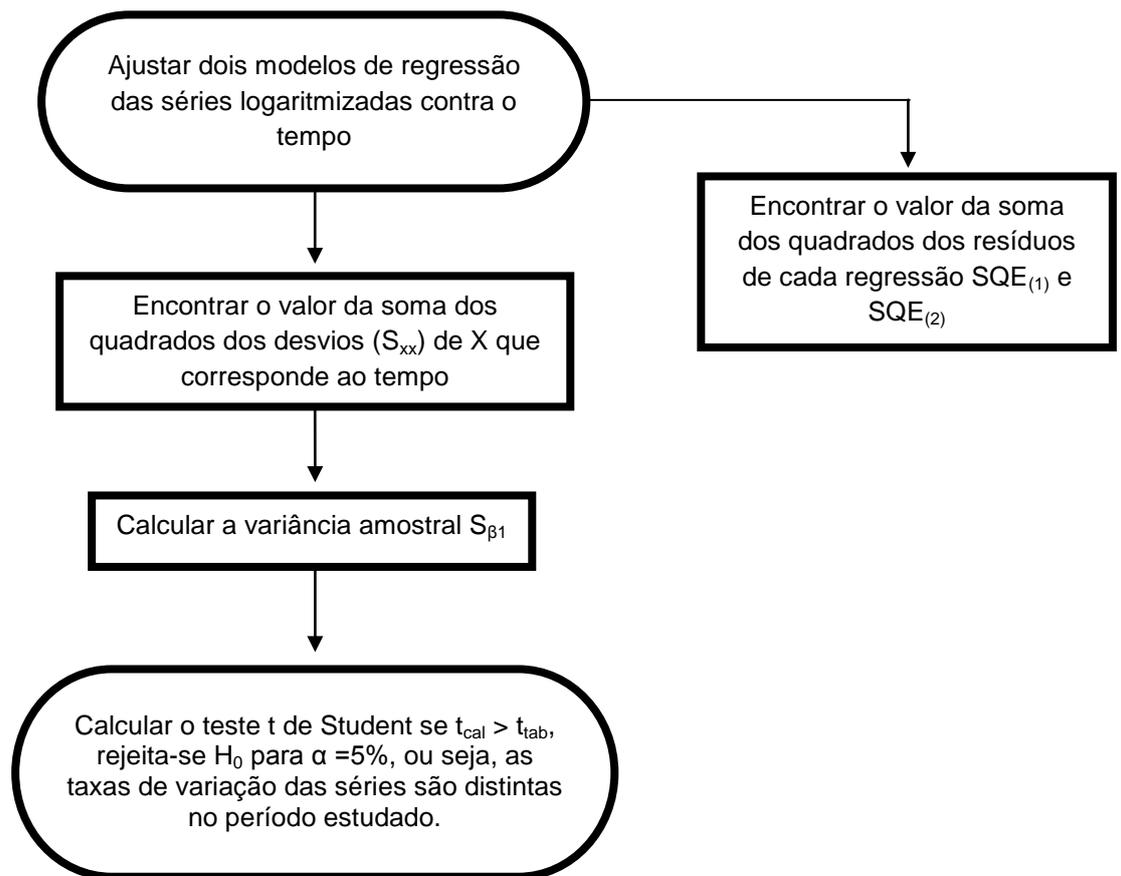


Figura 6- Fluxograma do teste de paralelismo.

3.1 Comentários gerais do capítulo

Foram mostradas as etapas a serem realizadas no estudo, bem como as descrições das séries OIE e PIB, unidades utilizadas, sites onde as séries estão disponíveis. Neste capítulo mostrou-se a ordem da metodologia de Box-Jenkins a qual foi seguida rigorosamente neste

estudo. Começando pela plotagem das séries ao longo período, seguida dos testes de estacionariedade, e a modelagem ARIMA. O próximo capítulo tem os resultados das análises realizadas segundo as etapas expostas neste capítulo.

A realização das modelagens das séries e o teste de inclinação servirão para comparar o comportamento da OIE com o do PIB e ver se as expectativas de uma possível OIE insuficiente ao crescimento econômico são verdadeiras. Em outras palavras, espera-se inferir a cerca do comportamento da OIE e do PIB analisando se haverá equilíbrio entre a taxa de crescimento do PIB e a OIE.

4 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Nesse capítulo são apresentados os resultados encontrados após a aplicação da metodologia adotada conforme explicitada no capítulo 2. A seção 4.1 contém a análise do comportamento das séries OIE e PIB, em estudo. Na seção 4.2 apresentam-se os resultados encontrados na metodologia de Box-Jenkins, bem como os testes estatísticos próprios para tal metodologia. Já na seção 4.3 mostram-se os resultados do teste de inclinação dos modelos.

4.1 Análise do comportamento das Séries Temporais: OIE e PIB

Este item contém os gráficos das séries em estudo contra o tempo e dão uma ideia preliminar do comportamento das séries ao longo do tempo, mostrando visualmente o comportamento destas ao longo do período estudado. A Figura 6 mostra o comportamento da série OIE no Brasil ao longo de 70 anos contando a partir do ano de 1940 até 2009. Vê-se pelo gráfico que a série possui uma tendência ascendente ao longo dos anos, isto é, ela tem um comportamento crescente de 1940 a 2009.

Este aumento da Oferta Interna de Energia no decorrer do período analisado se dá por vários fatores determinantes que fazem parte da evolução do desenvolvimento do Brasil. Dentre estes fatores tem-se o crescimento populacional, acrescido da elevação da população urbana versus população rural. Além disso, tem-se nesse período a industrialização do país passando pouco a pouco de uma economia agrícola para uma economia manufaturada.

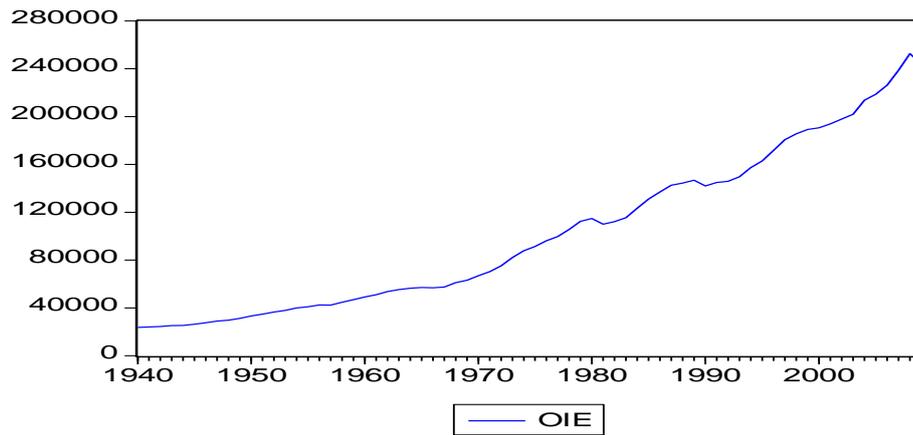


Figura 7- Comportamento da Oferta Interna de Energia (1940 a 2009).

Analisando a série do Produto Interno Bruto (PIB) vê-se que a mesma tem comportamento semelhante ao da OIE, dessa maneira ambas as séries têm uma tendência crescente no período analisado. Também é possível averiguar que as séries têm um crescimento acentuado a partir da década de 1970, sendo que esta característica semelhante pode ser uma evidência de uma relação de longo prazo entre as séries.

O crescimento do PIB no período de 1940 a 2009 pode ser relacionado com a industrialização no país, além de que as políticas econômicas adotadas pelo Governo na tentativa da correção das grandes inflações nas décadas precedentes à década de 90 onde estas permitiram um amadurecimento da economia como um todo, culminando no Plano Real em meados da década de 1990.

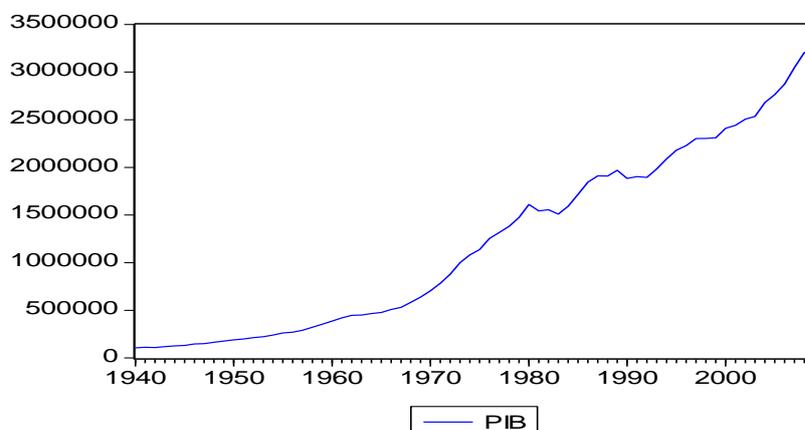


Figura 8- Comportamento da Série temporal PIB (1940 a 2009).

Outro aspecto atrelado aos anteriormente citados é que o aumento do poder aquisitivo da população, derivado da estabilização econômica permitiu aos consumidores elevarem consideravelmente o consumo interno no país. É relevante salientar que o crescimento econômico das últimas décadas no país está acompanhado por melhoras tímidas, mas contínuas nos indicadores sociais, como por exemplo, taxa de mortalidade infantil, coeficiente de Gini, entre outros.

4.2 Modelagens das séries: OIE e PIB e seus respectivos testes

O primeiro passo realizado para a modelagem foi a verificação da estacionariedade das séries. Para tanto se utilizou inicialmente a função de autocorrelação (FAC) e função de autocorrelação parcial (FACP) para verificar a autocorrelação entre as observações das séries.

Os correlogramas das séries temporais OIE e PIB evidenciam que os valores de autocorrelação decrescem lentamente conforme aumenta o número de defasagens, sendo um indicativo de que as séries não são estacionárias e na função de autocorrelação parcial a série é truncada na defasagem 1, indicando que as duas séries podem ser modelos AR(1), ou seja um modelo autorregressivo de ordem um.

Autocorrelation		Partial Correlation		AC	PAC	Q-Stat	Prob	
				1	0.886	0.886	57.364	0.000
				2	0.780	-0.025	102.46	0.000
				3	0.692	0.025	138.45	0.000
				4	0.619	0.027	167.75	0.000
				5	0.554	-0.004	191.53	0.000
				6	0.492	-0.014	210.57	0.000
				7	0.439	0.012	225.99	0.000
				8	0.396	0.018	238.75	0.000
				9	0.360	0.012	249.48	0.000
				10	0.326	-0.005	258.42	0.000
				11	0.290	-0.023	265.60	0.000
				12	0.256	-0.007	271.29	0.000
				13	0.228	0.008	275.90	0.000
				14	0.207	0.013	279.75	0.000
				15	0.186	-0.006	282.91	0.000

Figura 9 - Correlograma da série em nível OIE.

Autocorrelation		Partial Correlation		AC	PAC	Q-Stat	Prob	
				1	0.895	0.895	58.523	0.000
				2	0.801	-0.001	106.09	0.000
				3	0.723	0.029	145.38	0.000
				4	0.658	0.029	178.43	0.000
				5	0.598	0.000	206.20	0.000
				6	0.545	0.005	229.58	0.000
				7	0.494	-0.009	249.12	0.000
				8	0.447	-0.004	265.39	0.000
				9	0.408	0.013	279.12	0.000
				10	0.368	-0.016	290.53	0.000
				11	0.328	-0.027	299.70	0.000
				12	0.288	-0.018	306.90	0.000
				13	0.255	0.008	312.66	0.000
				14	0.226	-0.003	317.25	0.000
				15	0.195	-0.027	320.71	0.000

Figura 10- Correlograma da série em nível PIB

Para confirmar se as séries são estacionárias faz-se necessário aplicar um teste estatístico. No presente estudo foi utilizado o teste ADF e KPSS. Os resultados dos testes estão descritos na Tabela 1. Logo, verifica-se que o teste ADF mostra que as séries em estudo são não estacionárias em nível, onde os p-value > 0,05, ou seja, não há evidências para se rejeitar a hipótese nula a qual se refere que as séries são não estacionárias. Já o teste KPSS que tem as hipóteses invertidas em relação ao ADF mostra que com 5% de significância, rejeita-se H_0 , indicando que as séries são não estacionárias em nível.

Tabela 1- Resultado dos testes de raiz unitária (ADF e KPSS) para as séries em nível.

Ordem de Integração da Série (d= 0)	Teste ADF	Teste KPSS
	$H_0: \delta = 0$ ou $\rho = 1$ $H_1: \delta \neq 0$ ou $\rho \leq 1$	$H_0: \delta \neq 0$ ou $\rho \leq 1$ $H_1: \delta = 0$ ou $\rho = 1$
OIE	$\tau = 3,6134$ (p > 0,05)	LM = 1,0760 (p < 0,05)
PIB	$\tau = 3,3265$ (p > 0,05)	LM = 1,0865 (p < 0,05)

Com a confirmação da não estacionariedade das duas séries em nível realizou-se então uma diferenciação nas séries e após foi realizado os correlogramas das séries diferenciadas como se mostra a seguir.

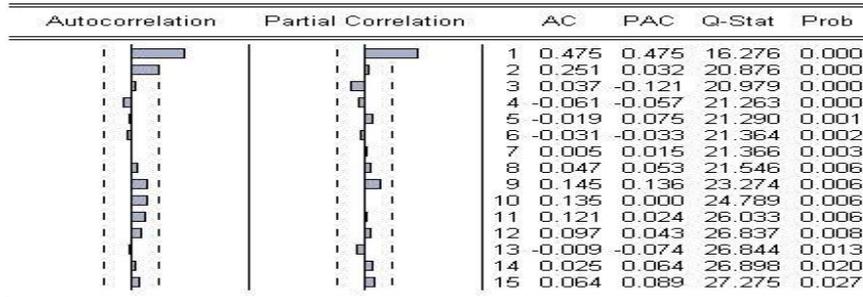


Figura 11- Correlograma da série OIE em primeira diferença.

A verificação da série em primeira diferença também pode ser visualizada pelo gráfico. A Figura 12 contém a série da OIE em nível e em primeira diferença, evidenciando que a série fica mais próxima da média zero, mas com variabilidade a partir da metade da década de 1970 até 2009.

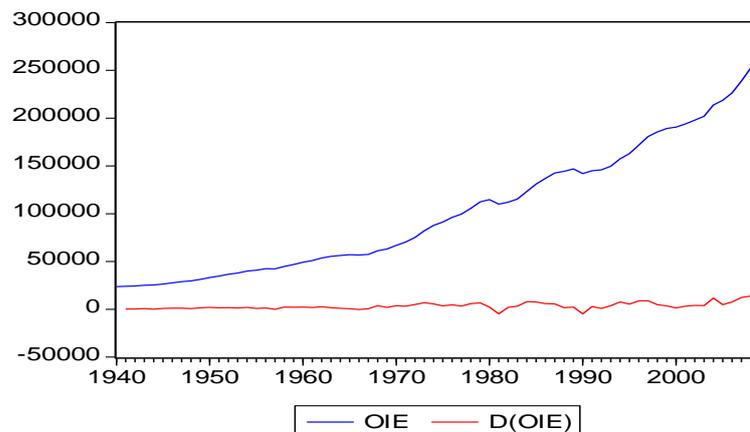


Figura 12- Série do OIE em primeira Diferença

O correlograma da série do PIB em primeira diferença representado pela Figura 13 mostra que todos os coeficientes de autocorrelação estão dentro dos limites de confiabilidade.

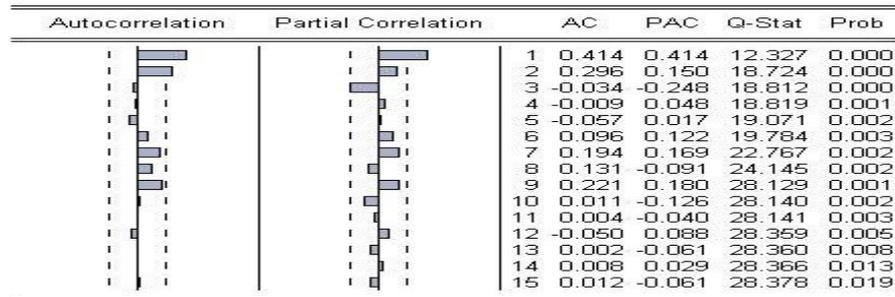


Figura 13 - Correlograma da série em primeira diferença PIB

A série do PIB em primeira diferença pode ser visualizada na Figura 14 a qual mostra a série contra o tempo. O gráfico mostra que com uma diferença a série do PIB torna-se mais próxima da média, porém com variabilidade acentuada a partir da década de 1975 aproximadamente.

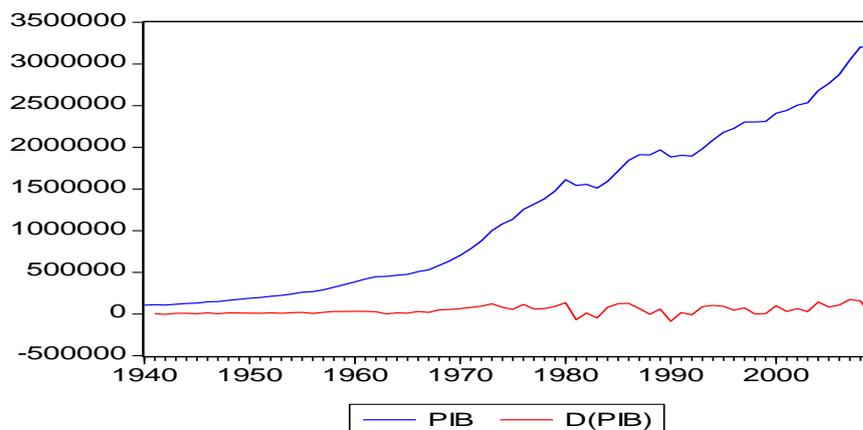


Figura 14- Série do PIB em primeira diferença

Assim, como se pode verificar pelos correlogramas e os gráficos das séries diferenciadas nota-se que com uma diferença nas séries elas aparentemente tornaram-se estacionárias. Assim, realizaram-se novamente os testes estatísticos para a confirmação dos correlogramas. Conforme a Tabela 2 os resultados do teste ADF e KPSS evidenciam que com 5% de significância após uma diferenciação as séries tornaram-se estacionárias.

Tabela 2- Resultado dos testes de raiz unitária (ADF e KPSS) para as séries em primeira diferença.

Ordem de Integração da Série (d= 1)	Teste ADF	Teste KPSS
	$H_0: \delta = 0 \text{ ou } \rho = 1$	$H_0: \delta \neq 0 \text{ ou } \rho \leq 1$
	$H_1: \delta \neq 0 \text{ ou } \rho \leq 1$	$H_1: \delta = 0 \text{ ou } \rho = 1$
ΔOIE	$\tau = 4,8479 \text{ (p} < 0,05)$	$LM = 0,7355 \text{ (p} > 0,05)$
ΔPIB	$\tau = 5,3854 \text{ (p} < 0,05)$	$LM = 0,6161 \text{ (p} > 0,05)$

A estacionariedade diz respeito à estabilidade da série ao longo do tempo, após realizar as transformações das séries verificou-se que as mesmas tornaram-se estacionárias em primeiras diferenças, ou seja, possuem média constante significativa estatisticamente ao longo do tempo, dessa forma estas séries então podem ser modeladas para encontrar seus respectivos processos geradores. Como se notou uma variabilidade nas séries a partir da década de 70 e 80 na OIE e PIB, respectivamente, então se realizou posteriormente o teste de Jarque-Bera e posteriormente o teste ARCH- LM para avaliar a variabilidade das séries.

4.2.1 Modelagem das séries

Após a modelagem na procura do modelo mais adequado para descrever as séries, tem-se nas Tabelas 3 e 4 os modelos concorrentes que tiveram seus coeficientes significativos e os resíduos ruído branco. Por meio da análise dos critérios de seleção AIC e BIC obtiveram-se os melhores modelos que se ajustam aos dados. A Tabela 3 contém os modelos concorrentes para a série da OIE e do PIB com seus respectivos coeficientes (sendo todos significativos a 5% de significância), e os critérios AIC e BIC. É relevante dizer que todos os modelos descritos possuem seus resíduos ruído branco (ou seja, possuem média zero, variância constante e são não-autocorrelacionados).

O modelo escolhido para série OIE é o número 1, onde trata-se de um modelo ARIMA(1,1,0), ou seja, tem um termo autoregressivo no lag 1 e a série foi diferenciada uma vez. Dentre os modelos concorrentes para a série do PIB o modelo elegido para estes dados foi o modelo número 1. Este modelo é um ARIMA (1,1,1) no qual há um termo (p) autoregressivo, a série foi diferenciada uma vez (d) e há um termo de médias móveis (q).

Tabela 3- Modelos concorrentes para a série da OIE que tiveram seus coeficientes significativos presença de ruído branco.

Oferta Interna de Energia				
Nº	Modelo	Coefficientes	Crítério AIC	Crítério BIC
1	ARIMA(1,1,0)	$\mu = 3145,29$ $\phi_1 = 0,3967$	19,1342	19,1995
2	ARIMA(1,1,0)	$\phi_1 = 0,6824$	19,2511	19,2838
3	ARIMA(0,1,1)	$\mu = 3107,15$ $\theta_1 = 0,3691$	19,1461	19,2108
Produto Interno Bruto				
Nº	Modelo	Coefficientes	Crítério AIC	Crítério BIC
1	ARIMA(1,1,1)	$\mu = 45891,21$ $\phi_1 = 0,3268$ $\theta_1 = 0,3242$	24,4234	24,5221
2	ARIMA(1,1,0)	$\mu = 44952,84$ $\phi_1 = 0,3817$	24,4292	24,4945
3	ARIMA(0,1,1)	$\mu = 44204,54$ $\theta_1 = 0,2641$	24,4741	24,5389

Legenda: μ = termo constante do modelo; ARIMA = subdividi-se em: AR = termo autorregressivo (com coeficiente igual a ϕ); I = número de vezes que a variável foi diferenciada; MA = termo de média móvel (com coeficiente igual a θ); AIC = Critério Akaike, BIC = critério Schwarz.

Logo, pela modelagem estabelecida por Box-Jenkins utilizando o critério de AIC e BIC estes são os modelos mais adequados para representar estes conjuntos de dados. A escolha dos modelos para ambas as séries sugerem que há uma captação melhor das características intrínsecas às séries da OIE e PIB.

4.2.2 Validação dos modelos elegidos

Após encontrar os modelos mais adequados para os conjuntos de dados estudados têm-se então os modelos para as séries OIE e PIB respectivamente:

$$\Delta \widehat{OIE}_t = 3145,29 + 0,3967 OIE_{t-1} + \hat{a}_t \quad (75)$$

p-valor (0,0001) (0,0022)

Onde a_t é o resíduo do modelo e Δ (delta) significa que a série foi diferenciada uma vez para se tornar estacionária. O modelo acima demonstra que a Oferta Interna de Energia é uma série autocorrelacionada e com isso esta varia proporcionalmente ao seu valor no período anterior (ano anterior), em outras palavras a OIE terá aproximadamente 0,40 tep explicada pela oferta de energia do período anterior.

Para a validação do modelo foi necessário verificar se os resíduos do mesmo possuíam média zero, variância constante e os resíduos não autocorrelacionados. Conforme as Figuras 15 e 16 que se referem aos resíduos do modelo e os resíduos quadráticos, respectivamente, verificou-se que os coeficientes de autocorrelação estão dentro dos limites de confiabilidade evidenciando que os resíduos são não-autocorrelacionados.

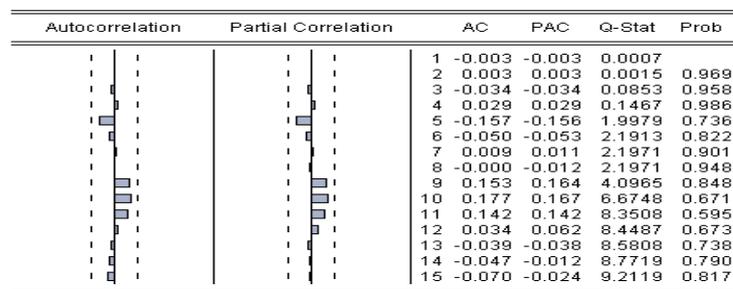


Figura 15– FAC e FACP dos resíduos do modelo ARIMA (1,1,0).

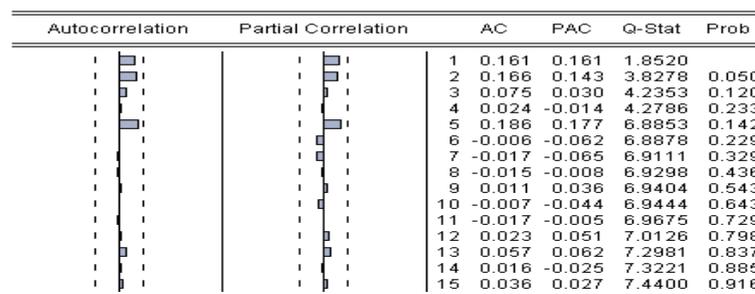


Figura 16- FAC e FACP dos resíduos ao quadrado do modelo ARIMA (1,1,0).

Conforme se verificou variabilidade na série da OIE a partir da década de 1980, realizou-se o teste de Jarque-Bera exposto na Figura 17, o mesmo tem como hipótese nula conjunta que a assimetria é igual a zero e a curtose é igual a 3 evidenciando assim que os dados possuem uma distribuição normal.

Logo, o resultado do teste mostrou que seu p-valor $< 0,05$, rejeitou-se a hipótese nula de que os resíduos seguem uma distribuição normal. Pois, verificou-se que os resíduos da OIE possuem uma assimetria negativa e excesso de curtose, dando uma ideia preliminar de que há heterocedasticidade condicional nos resíduos quadráticos do modelo.

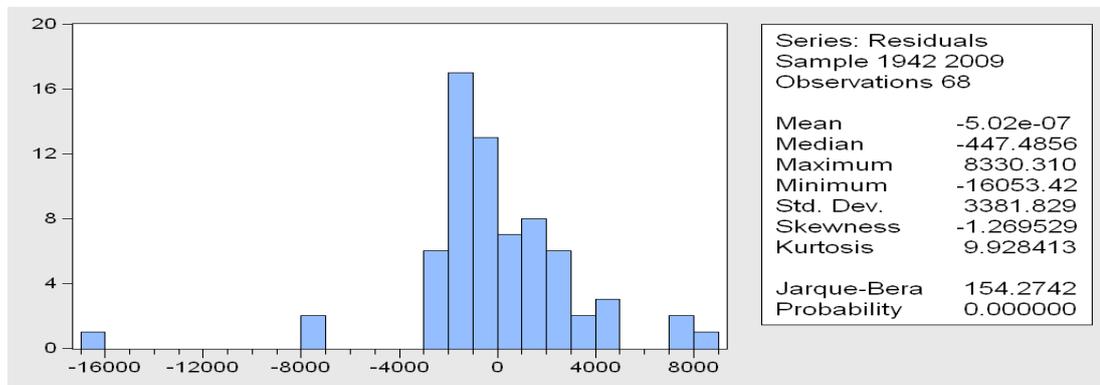


Figura 17– Histograma e estatística descritiva dos resíduos do modelo ARIMA (1,1,0) para a série da Oferta Interna de Energia (OIE).

Para averiguar melhor o comportamento da série realizou-se então o teste de heterocedasticidade condicional ARCH-LM que tem como hipótese nula a ausência de heterocedasticidade. Com um p-valor $< 0,05$ constatou-se que há presença de heterocedasticidade condicional, ou seja há evidências de volatilidade nos resíduos quadráticos do modelo ARIMA (1,1,0). Dessa forma, estudos de modelos que capturem a volatilidade são adequados e tornam-se uma sugestão para trabalhos futuros. O resultado do teste de heterocedasticidade encontra-se na Figura 18.

Heteroskedasticity Test: ARCH				
F-statistic	9.782667	Prob. F(1,65)	0.0026	
Obs*R-squared	8.764580	Prob. Chi-Square(1)	0.0031	
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Date: 11/23/11 Time: 16:01				
Sample (adjusted): 1943 2009				
Included observations: 67 after adjustments				
	Coefficient	Std. Error	t-Statistic	Prob.
C	5264577.	4383198.	1.201081	0.2341
RESID^2(-1)	0.807478	0.258168	3.127726	0.0026
R-squared	0.130815	Mean dependent var		11393583
Adjusted R-squared	0.117443	S.D. dependent var		34161593
S.E. of regression	32092948	Akaike info criterion		37.43557
Sum squared resid	6.69E+16	Schwarz criterion		37.50138
Log likelihood	-1252.091	Hannan-Quinn criter.		37.46161
F-statistic	9.782667	Durbin-Watson stat		1.594563
Prob(F-statistic)	0.002636			

Figura 18- Teste de Heterocedasticidade Condicional ARCH-LM para a Série OIE.

Na Equação 76 o modelo estimado para a série do PIB mostra que o Produto Interno Bruto brasileiro para o período de 1940 a 2009 é explicado por um parâmetro autorregressivo com lag 2 e um de médias móveis no lag 1. O resultado do modelo que melhor se ajusta aos dados para a série do PIB está de acordo com o esperado, pois a priori o PIB depende de muitas variáveis econômicas que são implicitamente correlacionadas além de que algumas são exógenas em relação à política econômica adotada durante determinado período.

$$\Delta \widehat{PIB}_t = 44519,85 + 0,3268PIB_{t-2} + 0,3242a_{t-1} + \hat{a}_t \quad (76)$$

p-valor (0,0002) (0,0127) (0,0157)

As Figuras 19 e 20 mostram a função de autocorrelação e autocorrelação parcial mostrando que todos os coeficientes estão entre os limites de confiabilidade evidenciando assim que os resíduos são não-autocorrelacionados.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.006	0.006	0.0029	
		2 -0.002	-0.002	0.0031	
		3 -0.074	-0.074	0.3998	0.527
		4 -0.042	-0.042	0.5311	0.767
		5 -0.211	-0.212	3.8449	0.279
		6 0.038	0.034	3.9558	0.412
		7 0.180	0.179	6.4466	0.265
		8 0.036	0.005	6.5459	0.365
		9 0.107	0.103	7.4634	0.382
		10 -0.077	-0.103	7.9422	0.439
		11 0.105	0.147	8.8509	0.451
		12 -0.136	-0.058	10.412	0.405
		13 0.078	0.081	10.939	0.448
		14 -0.061	-0.052	11.262	0.507
		15 0.059	0.007	11.574	0.563

Figura 19- FAC e FACP dos resíduos do modelo ARIMA (1,1,1).

	Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1			0.045	0.045	0.1400	
2			0.135	0.133	1.4339	
3			0.092	0.082	2.0402	0.153
4			0.165	0.145	4.0299	0.133
5			0.221	0.200	7.6596	0.054
6			-0.031	-0.084	7.7338	0.102
7			0.043	-0.031	7.8730	0.163
8			0.018	-0.026	7.8991	0.246
9			0.324	0.293	16.253	0.023
10			0.061	0.041	16.558	0.035
11			-0.057	-0.123	16.831	0.051
12			-0.049	-0.137	17.032	0.074
13			-0.023	-0.107	17.079	0.106
14			0.074	-0.012	17.554	0.130
15			-0.058	0.032	17.851	0.163

Figura 20- FAC e FACP dos resíduos ao quadrado do modelo ARIMA (1,1,1).

Analisando a Figura 21 verificou-se pelo teste de Jarque-Bera (onde H_0 : tem distribuição normal nos resíduos) que o p-valor é inferior ao nível de significância de 5%, logo se rejeitou a hipótese nula de que os resíduos seguem uma distribuição normal. Além disso, se observou que os resíduos do modelo (1,1,1) possuem excesso de curtose e assimetria negativa. Assim, este resultado sugere que há presença de heterocedasticidade condicional no modelo.

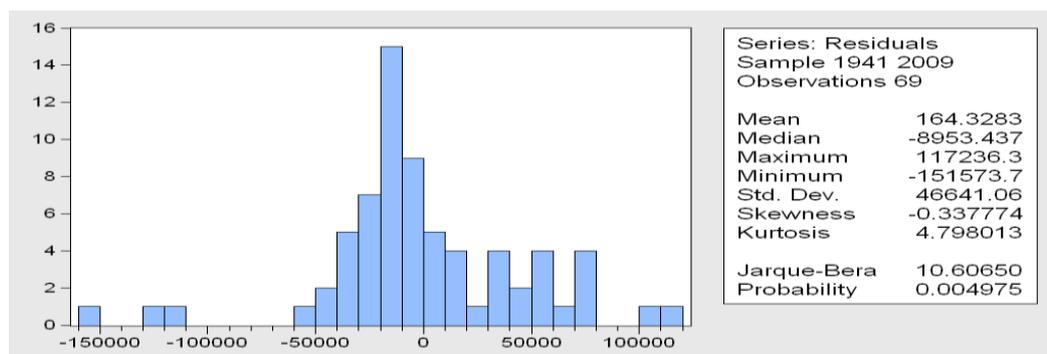


Figura 21- Histograma e estatística descritiva dos resíduos do modelo ARIMA (1,1,1) para a série da Produto Interno Bruto (PIB).

Os resíduos do modelo ARIMA (1,1,1) possuem excesso de curtose, ou presença de caudas pesadas sugerindo que a presença de heterocedasticidade condicional no modelo. Para a verificação da presença de volatilidade nos resíduos foi realizado o teste de heterocedasticidade condicional ARCH-LM o qual tem como hipótese nula a ausência de

heterocedasticidade condicional. O resultado do teste mostrou um p-valor $> 0,05$ evidenciando que a ausência de heterocedasticidade condicional, ou seja, não há evidências de existência de volatilidade nos resíduos quadráticos do modelo ARIMA (1,1,1).

Heteroskedasticity Test: ARCH				
F-statistic	0.155528	Prob. F(1,64)	0.6946	
Obs*R-squared	0.160000	Prob. Chi-Square(1)	0.6892	
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Date: 01/12/12 Time: 14:26				
Sample (adjusted): 1944 2009				
Included observations: 66 after adjustments				
	Coefficient	Std. Error	t-Statistic	Prob.
C	2.09E+09	5.97E+08	3.494957	0.0009
RESID^2(-1)	0.053736	0.136258	0.394371	0.6946
R-squared	0.002424	Mean dependent var	2.19E+09	
Adjusted R-squared	-0.013163	S.D. dependent var	4.31E+09	
S.E. of regression	4.34E+09	Akaike info criterion	47.25042	
Sum squared resid	1.21E+21	Schwarz criterion	47.31678	
Log likelihood	-1557.264	Hannan-Quinn criter.	47.27664	
F-statistic	0.155528	Durbin-Watson stat	1.858088	
Prob(F-statistic)	0.694619			

Figura 22- Teste de Heterocedasticidade Condicional ARCH-LM para a Série PIB.

4.3 Teste para igualdade das inclinações

Para a realização deste teste foi utilizado o modelo proposto pelo item 2.2 no capítulo 2. Foram utilizados modelos de regressão simples para a realização deste teste utilizando as séries temporais OIE e PIB (como variáveis dependentes) e a tendência (tempo) com a variável independente ou regressora.

Na presente estudo foram utilizados dois modelos semilogarítmicos (log-lin, ou seja, calculou-se o logaritmo das variáveis dependentes e mantiveram-se os regressores na forma linear) para que se tornasse factível comparar as taxas de variação das séries ao longo do tempo, já que as mesmas possuem suas unidades de medidas distintas. Assim, os coeficientes passaram a ter a mesma medida a qual corresponde o percentual de crescimento ou de decréscimo ao longo do período estudado (1940 a 2009).

$$\ln \widehat{OIE}_{j1} = 10,05 + 0,0354t_{j1} \quad (77)$$

p-valor (0,0000) (0,0000)

$$\ln\widehat{PIB}_{j2} = 11,72 + 0,0521t_{j2} \quad (78)$$

p-valor (0,0000) (0,0000)

Os modelos expostos nas equações 77 e 78 têm seus coeficientes significativamente diferentes de zero, em outras palavras, os coeficientes angulares e lineares dos modelos abaixo são significativos, pois possuem p-valor < 0,05. Este teste tem como hipótese nula a igualdade dos coeficientes angulares dos dois modelos.

Conforme a apresentação do teste no item 2.2 segue os cálculos necessários para realizar o teste de paralelismo. Primeiramente, obtiveram-se os valores das somas dos quadrados dos desvios das variáveis OIE e PIB e as respectivas somas dos produtos cruzados. Posteriormente, encontraram-se os valores das somas dos quadrados dos resíduos (SQE) os valores seguem abaixo:

$$S_{zz(OIE)} = \sum z^2 - \frac{(\sum z)^2}{n} = 8981,25 - \frac{(791,30)^2}{70} = 36,17 \quad (79)$$

$$S_{xz(OIE)} = \sum x.z - \frac{\sum x \cdot \sum z}{n} = 29104,08 - \frac{791,30 * 2485}{70} = 1.012,93 \quad (80)$$

$$SQE_{(OIE)} = S_{zz} - b * S_{xz} = 36,17 - 0,035448 * 1.012,93 = 0,263657 \quad (81)$$

Onde b corresponde ao coeficiente angular ou de inclinação do modelo de regressão.

$$S_{zz(PIB)} = \sum z^2 - \frac{(\sum z)^2}{n} = 12959,89 - \frac{(949,49)^2}{70} = 80,87 \quad (82)$$

$$S_{xz(PIB)} = \sum x.z - \frac{\sum x \cdot \sum z}{n} = 35194,77 - \frac{949,49 * 2485}{70} = 1.487,87 \quad (83)$$

$$SQE_{(PIB)} = S_{zz} - b * S_{xz} = 80,87 - 0,052067 * 1487,87 = 3,40107 \quad (84)$$

Onde b corresponde ao coeficiente angular ou de inclinação do modelo de regressão.

$$S_{xx} = \sum x^2 - \frac{(\sum x)^2}{n} = 116795 - \frac{(2485)^2}{70} = 28.577,5 \quad (85)$$

$$S_{\beta_1} = \frac{0,263657 + 3,40107}{(70 + 70) - 4} \left[\frac{1}{28577,5} + \frac{1}{28577,5} \right] = 0,00000188586 \quad (86)$$

$$T_{\beta_1} = \frac{0,035448 - 0,052067}{\sqrt{0,00000188586}} = -12,1018 \quad (87)$$

O teste de igualdade de inclinações tem como resultado a rejeição da hipótese nula a qual se refere que os coeficientes angulares dos dois modelos são iguais. Em outras palavras, entende-se que o teste é significativo já que com (n_1+n_2-4) graus de liberdade, tem-se um valor tabelado da estatística t de Student de 1,980 para um nível de significância de 5%, verificando que a estatística calculada é superior à tabelada rejeitando então H_0 . Assim, com 95% de confiança, conclui-se que a taxa de variação para o crescimento do PIB é diferente da taxa de crescimento da OIE.

Este resultado é uma confirmação da atual realidade brasileira quando se compara a oferta de energia e o crescimento econômico, visto que o Governo tem incentivado o aumento da oferta de energia para que não ocorram novos déficits no fornecimento de energia no país como ocorreu nos anos de 2001 e 2002.

Dessa forma, verifica-se que no período de 1940 a 2009 a taxa crescimento das séries ao longo do tempo foi desigual sendo que a taxa de variação para o PIB foi maior ficando em torno de 5,21% e a OIE com 3,54%. Este resultado mostra que ao longo do período estudado o PIB cresceu aproximadamente 1,67 pontos percentuais a mais que a OIE. Neste contexto da disparidade das taxas de crescimento é que o Governo vem incentivando o investimento em infraestrutura no setor de energia, já que uma instabilidade energética pode ser um ponto negativo aos olhos dos investidores estrangeiros e nacionais, bem como para os países que mantêm comércio com o Brasil.

4.4 Comentários gerais do capítulo

O presente capítulo apresentou os resultados encontrados nesta pesquisa. Foi realizada a metodologia de Box-Jenkins para as duas séries em estudo. Os resultados mostraram que para a OIE o melhor modelo que descreve os dados coletados é um ARIMA (1,1,0). Já para a série do PIB o melhor modelo encontrado foi um ARIMA (1,1,1). A segunda análise realizada foi o teste de inclinação de dois modelos que tendo como resultado no período estudado (1940 a 2009) o PIB teve uma taxa de crescimento maior que a OIE.

Os resultados para as duas séries mostram que se deve deixar um alerta aos gestores em relação a possíveis disparidades das taxas de crescimento da OIE e PIB confirmando a necessidade de manter uma oferta de energia proporcionalmente capaz de propiciar a evolução do crescimento econômico. Além do mais, crescimentos exacerbados da demanda por energia, podendo no futuro levar o país a modificar sua matriz energética, devendo trocar a base da matriz energética limpa (utilizando fontes renováveis) para uma matriz energética que utilize na sua maioria fontes não renováveis (poluidoras).

5 CONSIDERAÇÕES FINAIS

A modelagem das séries permitiu verificar seus respectivos processos geradores, onde sabe-se intuitivamente que quando há um aumento no PIB faz-se necessário que aumente a OIE já que esta equivale à energia necessária para mover a economia do país e da mesma forma, quando há um aumento na OIE este possibilita um aumento no nível de produção de bens e serviços conforme a demanda por estes bens.

O resultado do teste de paralelismo de retas evidenciou que a taxa de crescimento do Produto Interno Bruto foi maior do que a Oferta Interna de Energia no período estudado (1940-2009). Logo, comprovou-se que o país cresceu muito economicamente ao longo do tempo ao passo que a Oferta Interna de Energia não acompanhou este rápido crescimento provocando momentos de insuficiência na oferta de energia no país.

Os resultados desta pesquisa evidenciam que há uma tendência de crescimento da série Oferta Interna de Energia e do Crescimento Econômico corroborando com o conhecimento de eventos que estão correlacionados, permitindo aos tomadores de decisões organizarem as políticas a serem adotadas visando melhores estratégias para manter o nível de Oferta Interna de Energia capaz de propiciar um aumento expressivo e contínuo do Produto Interno Bruto. Ademais, verifica-se a importância de o Governo gerar maior infraestrutura no setor energético mantendo a utilização de fontes energéticas renováveis permitindo um crescimento econômico sustentável, estável e duradouro, conforme a política econômica brasileira adotada, bem como a situação da economia mundial.

Dessa forma, entende-se que o país está na direção correta quando se trata de produção de energia baseada em fontes energéticas limpas. Contudo, deve-se ter um planejamento bem estruturado para que a infraestrutura do setor energético no longo prazo continue utilizando na sua maior parte as fontes de energias renováveis (não poluidoras) e não desestabilize esse processo de produção de energia limpa segundo as necessidades da manutenção do crescimento econômico no futuro.

Ao finalizar as considerações deste estudo faz-se relevante apontar sugestões para estudos futuros, como a realização do teste de paralelismo comparando as variações ao longo do tempo da oferta interna de energia com a demanda de energia pelos setores: comercial, residencial, industrial e de transporte, ou ainda comparando as variações destes setores com o crescimento econômico.

REFERÊNCIAS BIBLIOGRÁFICAS

- BUENO, R. L. S.. **Econometria de Séries Temporais**. São Paulo: Cengage Learning, 2008.
- CHARNET, R.; FREIRE, C. A. L.; CHARNET, E. M. R.; BONVINO, H.; **Análise de Modelos de Regressão Linear**. Editora Unicamp, 1999.
- DE BRUYN, S.; VAN DEN BERGH, J. C. M.; OPSCHOOR, J.B. **Economic Growth and Emissions**: reconsidering the empirical basis of environmental Kuznets curves. *Ecological Economics*, Amsterdam, vol. 25, 1998.
- GUJARATI, D. N. **Econometria Básica**. São Paulo: Makron Books, 2006.
- GROSSMAN, G. M.; KRUEGER, A. B. **Environmental Impacts of a North American Free Trade Agreement**, National Bureau of Economic Research Working Paper 3914, NBER, Cambridge, MA, 1991.
- GROSSMAN, G. M.; KRUEGER, A. B. **Economic Growth and the Environment**. *The Quarterly Journal of Economics*, v. 110, n. 2, 1995.
- HILL, R. C., GRIFFITHS, W. E.; JUDGE, G. G.. **Econometria**. São Paulo: Saraiva, 1999.
- INSTITUTO DE PESQUISA ECONÔMICA APLICADA (IPEADATA). Disponível em: <<http://www.ipeadata.gov.br/>>. Acesso em: 20 de mar. 2011.
- MINISTÉRIO DE MINAS E ENERGIA. **Balço Energético Nacional 2009**. MME. Brasília, DF, 2010. Disponível em: <http://www.mme.gov.br/>.> Acesso em mar. 2011.
- MINISTÉRIO DE MINAS E ENERGIA. NÚCLEO DE ESTUDOS ESTRATÉGICOS DE ENERGIA- NEEE. **Ranking Mundial de Energia e Socioeconomia**. Brasília, 2011.
- MORETTIN, P. A.; TOLOI, C. M. C. **Modelos para Previsão de Séries Temporais**. 13º Colóquio Brasileiro de Matemática. v.1, 1981.
- MORETTIN, P. A.; TOLOI, C. M. C., **Modelos de Função de Transferência**, Rio de Janeiro, 3ª Escola e séries Temporais Econometria, 1989.
- MORETTIN, P. A.; TOLOI, C. M. C. **Análise de Séries Temporais**. São Paulo, Edgard Blücher, 2004.
- MORETTIN, P. A. **Econometria Financeira: um curso de séries temporais**. 17º Simpósio Nacional de Probabilidade e Estatística. ABE. São Paulo, 2006.
- SELDEN, T. M.; SONG, D. **Environmental Quality and Development**: Is There a Kuznets Curve for Air Pollution Emissions? *Journal of Environmental Economics and Management*, v. 27, 1994.

SOUZA, F. M. et al. Previsão do consumo de cimento no estado do Rio Grande do Sul. **Pesquisa Operacional para o Desenvolvimento**, v. 2, p. 3-11, 2010.

SOUZA, F. M. et al. Previsão do preço da gasolina para a região Sul do Brasil. *Iberoamerican Journal of Industrial Engineering*, v. 1, p. 234-248, 2011.

SOUZA, F. M.. **Modelos Box & Jenkins Aplicados a Previsão de Demanda de Leitos Hospitalares**. Monografia (Especialização em Estatística e Modelagem Quantitativa)-Universidade Federal de Santa Maria, Santa Maria, 2006.

WOLFF, L. et al. **Análise do índice BOVESPA sob enfoque de séries temporais**. *Global Manager*, v.1, p.1-10, 2011.