

57  
32

ET-32



UNIVERSIDADE EDUARDO  
UNIVERSIDADE EDUARDO  
MONDLANE



FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE MATEMÁTICA E INFORMÁTICA

TRABALHO DE LICENCIATURA EM ESTATÍSTICA

MODELO DE PREVISÃO DO CONSUMO DE ENERGIA  
ELÉCTRICA FACTURADA, BASEADO NA METODOLOGIA  
DE BOX-JENKINS  
CASO: CIDADE DE MAPUTO

**Autor:**

Hélio Rafael Langa

**Orientadora:**

Dra. Rafica Abdulrazac

Maputo, Abril de 2008

ET-32



## DECLARAÇÃO DE HONRA

Declaro que este trabalho é resultado da minha própria investigação, e que não foi submetido para outro grau que não seja o indicado – Licenciatura em Estatística na Faculdade de Ciências da Universidade Eduardo Mondlane.

As ideias originais nele expressas são da inteira responsabilidade do autor.

Maputo, Abril de 2008

*Hélio Rafael Langa*  
O autor  
.....  
(Hélio Rafael Langa)

A Deus e seus arcanjos  
que me iluminaram  
durante a minha formação  
e que continuam a iluminar-me

A minha querida mãe,  
o amor, dedicação, carinho  
e esforço que envidou  
para a minha formação

Aos meus irmãos, Abel e Nilza,  
A minha amiga prima Dinha  
Ao meu sobrinho Igor  
A minha namorada Stela

E a toda a minha família  
Que a todo o momento me incentivou

DEDICO

## Agradecimentos

A Professora Dra. Rafica Abdulrazac, um especial agradecimento, pela disponibilidade mostrada para supervisionar o presente trabalho, pela paciência e confiança e pelas valiosas contribuições para o desenvolvimento do mesmo.

Ao Prof. Doutor Rafael Conde, pelo incentivo à realização do trabalho e pela dedicação mostrada à turma de Estatística do 4<sup>o</sup> ano/2007 durante o curso.

Aos meus amigos, especialmente ao Madlhaya-nhoka, Júnior “ Anão clarinho”, Zito “Mafikisolo” e aos meus colegas, Dr. Osvaldo Loquia, Dr. Ussimane Dique, Dr. Adelino Juga, Gouveia, Timbe, Tembe e Cardozo cuja realização do trabalho só foi possível graças ao seu apoio emocional, material, logístico, etc.

Aos colegas do curso de Estatística pela amizade e pelo desafio.

A todos professores do curso de Estatística, por tornarem possível a realização do curso.

O meu muito Obrigado.

*"O pensamento estatístico é um modo de pensar que reconhece que as nossas observações sobre o mundo não podem ser completamente exactas: elas estão sempre sujeitas a alguma incerteza; é um processo de pensamento que reconhece que a variação está sempre presente em tudo o que fazemos".*

**Box, George E. P., 1993**

**Resumo:** Neste trabalho foi estimado o Consumo da Energia Eléctrica Facturada para clientes domésticos utilizando a linha de baixa tensão na Cidade de Maputo para o período de Junho de 2007 à Maio de 2008. Para o efeito, foi utilizada a metodologia de Box-Jenkins tendo as estimativas dos parâmetros sido feitas pelo método de máxima verosimilhança. A série do Consumo da Energia Eléctrica Facturada não apresentou movimentos sazonais no período em análise e não se manteve estacionaria, pois identificaram-se três níveis diferentes de consumo. Dos modelos estimados, o que melhor se ajustou aos dados foi o  $ARIMA(0,1,1)$ . As previsões indicam uma tendência decrescente no Consumo da Energia Eléctrica Facturada.

**Palavras chave:** Series temporais, Estacionaridade, Modelos Autoregressivos Integrados Media Móvel, Consumo da Energia Eléctrica Facturada

## LISTA DE TABELAS

Tabela 1. Propriedades da ACF e da PACF para vários modelos ARIMA.....	21
Tabela 2. Propriedades da ACF e da PACF para modelos SARIMA.....	21
Tabela 3. Estatísticas da Regressão de ADF.....	35
Tabela 4. Teste de raiz unitária de ADF.....	36
Tabela 5. Estatísticas da Regressão de ADF.....	36
Tabela 6. Estatísticas do teste de ADF.....	37
Tabela 7. Estatísticas da Estimação do ARIMA(0,1,1) incluindo o parâmetro de deslocamento.....	39
Tabela 8. Estatísticas da Estimação do modelo ARIMA(0,1,1) sem o parâmetro de deslocamento.....	39
Tabela 9. Estatística da Estimação do Modelo ARIMA (1,1,0) incluindo o parâmetro de deslocamento.....	40
Tabela 10. Estatísticas da Estimação do modelo ARIMA(1,1,0) sem o parâmetro de deslocamento.....	40
Tabela 11. Estatísticas da estimação do modelo ARIMA(1,1,1) incluindo o parâmetro de deslocamento.....	41
Tabela 12. Testes estatísticos para os pressupostos de normalidade, media dos resíduos igual a zero e variância constante.....	46
Tabela 13. Estatísticas do desempenho nas previsões dentro da amostra dos modelos seleccionados.....	46
Tabela 14. Previsão do Consumo para o período de Junho de 2007 a Maio de 2008.....	47

## LISTA DE FIGURAS

<b>Figura 1.</b> Evolução do Consumo da Energia Eléctrica Facturada (Janeiro de 2000 à Maio de 2007).....	31
<b>Figura 2.</b> Consumo da Energia Eléctrica Facturada média mensal e os respectivos Intervalos de Confiança.....	33
<b>Figura 3.</b> Correlograma da Função de Autocorrelação da série do Consumo da Energia Eléctrica Facturada em nível.....	33
<b>Figura 3.1.</b> Correlograma da Função de Autocorrelação Parcial da serie do Consumo da Energia Eléctrica Facturada em nível.....	35
<b>Figura 4.</b> Série do Consumo da Energia Eléctrica Facturada em primeira diferença.....	35
<b>Figura 5.</b> Correlograma da Função de Autocorrelação da série do Consumo da Energia Eléctrica Facturada em primeira diferença .....	36
<b>Figura 5.1.</b> Correlograma da Função de Autocorrelação Parcial da série do Consumo da Energia Eléctrica Facturada em primeira.....	37
<b>Figura 6.</b> Série dos resíduos após o ajuste do modelo ARIMA(1,1,0).....	40
<b>Figura 7.</b> Série dos resíduos após o ajuste do modelo ARIMA(0,1,1).....	41
<b>Figura 8.</b> Histograma da série dos resíduos do modelo ARIMA(1,1,0).....	41
<b>Figura 9.</b> Histograma da série dos resíduos do modelo ARIMA(0,1,1).....	42
<b>Figura 10.</b> Correlograma da Função de Autocorrelação da série dos resíduos do modelo ARIMA(1,1,0).....	43
<b>Figura 10.1.</b> Correlograma da Função de Autocorrelação Parcial da série dos resíduos do modelo ARIMA(1,1,0).....	44
<b>Figura 11.</b> Correlograma da Função de Autocorrelação da série dos resíduos do modelo ARIMA(0,1,1).....	45
<b>Figura 11.1.</b> Correlograma da Função de Autocorrelação Parcial da série dos resíduos do modelo ARIMA(0,1,1).....	46

## ÍNDICE

RESUMO.....	iv
LISTA DE TABELAS.....	v
LISTA DE FIGURAS.....	vi
<b>1. INTRODUÇÃO .....</b>	<b>9</b>
1.1. O problema .....	12
1.2. Objectivos.....	12
1.2.1. Geral .....	12
1.2.2. Específicos.....	12
<b>2. REVISÃO BIBLIOGRÁFICA .....</b>	<b>13</b>
2.1. Séries Temporais.....	13
2.2. Teste de Raiz Unitária .....	19
2.3. O método de Box-Jenkins .....	20
<b>3. MATERIAL E MÉTODOS .....</b>	<b>25</b>
3.1. Material .....	25
3.1.1. Dados usados .....	25
3.1.2. Fonte de dados .....	25
3.2. Métodos.....	25
3.2.1. Análise dos dados.....	25
3.2.1.2. Identificação.....	27
3.2.1.3. Estimaco.....	27
3.2.1.4. Diagnostico .....	28
3.2.1.5. Teste LM ao modelo ARCH .....	29
3.2.2. Comparaco dos modelos.....	30

<b>4. RESULTADOS E DISCUSSÃO</b> .....	31
4.1. Séries e representação gráfica.....	31
4.2. Teste de Raiz Unitária.....	34
4.3. Identificação.....	36
4.4. Estimação.....	37
4.4.1. Estimação do modelo Médias Moveis, ARIMA (0,1,1).....	37
4.4.2. Estimação do modelo Autoregressivo, ARIMA (1,1,0).....	38
4.4.3. Estimação do modelo Autoregressivo Integrado de Médias Moveis, ARIMA (1,1,1)...	39
4.3.3. Diagnostico.....	40
4.4. Comparação dos modelos.....	45
4.5. Previsão.....	45
<b>5. CONCLUSÕES</b> .....	47
5.1. Recomendações.....	48
5.2. Limitações do estudo.....	48
<b>6. BIBLIOGRAFIA</b> .....	49
<b>ANEXO A</b> .....	53
<b>ANEXO B</b> .....	61

# CAPITULO I

---

## 1. INTRODUÇÃO

A Energia Eléctrica constitui uma das fontes de energia mais usadas para a humanidade, em particular nas grandes cidades. Ela é amplamente usada em residências, cidades, comércio e indústrias, sendo nas residências uma fonte de iluminação e produção de calor para o funcionamento de ar condicionados, geleiras, congeladores, rádios, televisores, etc. Desta forma, ela pode também ser considerada como um bem ou serviço de primeira necessidade.

Na cidade de Maputo até 2003, cerca de 45.9% dos agregados familiares tinham acesso a energia eléctrica. Porém, até este período a nível nacional, apenas 6.9% tinham acesso a este serviço nas suas residências, IAF (2002/2003).

De acordo com o Plano Estratégico da Electricidade de Moçambique para o período 2006-2009, constituem acções prioritárias para a electrificação:

- I. A extensão das redes de transporte e distribuição de forma a assegurar a comercialização da energia disponível da Cahora Bassa e a redução da produção térmica;
- II. Assegurar a electrificação de 41 Sedes Distritais, 18 Postos Administrativos e 62 Localidades;
- III. Assegurar um nível de acesso de 9%, correspondente a 582.218 clientes domésticos, até finais de 2009;
- IV. Garantir uma taxa média de 70.000 novas ligações/ano;

Para que estes objectivos sejam alcançados, a avaliação e monitoria das actividades devem ser feitas periodicamente pelos gestores. Deste modo, a utilização de métodos estatísticos de previsão torna-se fundamental como ferramenta de auxílio no processo de tomada de decisão, uma vez que, a formulação de um planeamento incorrecto poderá incorrer em prejuízos irrecuperáveis.

A modelagem de séries temporais possibilita descrever um processo estocástico, utilizando valores passados da variável de interesse (estudo), baseando-se na ideia de que observações

passadas contêm informações sobre o nível e comportamento da série temporal em análise, Pessanha *et al* (2004).

Entre as diversas metodologias utilizadas para este fim, destacam-se os modelos de Box & Jenkins e os modelos de Amortecimento Exponencial de Holt e Winters, Pessanha *et al* (2004).

No presente trabalho, será aplicada a metodologia de Box & Jenkins, com o intuito de se obter uma visão prospectiva a médio prazo para o consumo facturado de energia eléctrica na Cidade de Maputo, utilizando os modelos ARIMA ou SARIMA (Auto Regressivo – Integrado – Media Móvel ou Auto Regressivo – Integrado – Media Móvel Sazonal).

O uso desta metodologia, prende-se com a fácil compreensão e implementação da técnica, com a precisão dos resultados e com o sucesso em que muitas vezes se obteve, Miranda (2002).

Por exemplo, Segundo Abdel-Aal & Al-Garni (1997), os modelos Box-Jenkins têm sido largamente utilizados para modelagem e previsão em aplicações médicas, ambientais, financeiras e de engenharia. Os próprios autores aplicaram essa metodologia para prever o consumo mensal de energia eléctrica no leste da Arábia Saudita. Para Slini *et al* (2001), a modelagem de séries temporais é apropriada à previsão da qualidade do ar. Esses autores aplicaram a metodologia de Box-Jenkins para prever a qualidade do ar na cidade de Atenas, na Grécia. Além de aplicações nas áreas citadas, Lim & McAleer (2001) mencionam, com base em diferentes métodos de previsão de séries temporais, que a literatura na área de previsão de demanda turística é numerosa. Os referidos autores aplicaram a metodologia de Box-Jenkins e modelaram a procura de turistas que chegam à Austrália provenientes de três países da Ásia.

Assim, com este estudo pretende-se sugerir um modelo de previsão de Consumo Facturado de Energia Eléctrica.

O presente trabalho está organizado em 5 capítulos, a saber:

No capítulo 1 é feita a apresentação do tema bem como a definição do problema de estudo. São também apresentados os objectivos gerais e específicos a atingir.

A revisão da literatura é apresentada no capítulo 2. Neste capítulo são introduzidos os conceitos de series temporais, estacionaridade que constituem uma base teórica à aplicação do método de Box-Jenkins. De forma a facilitar a compreensão dos resultados empíricos obtidos são também

apresentada as principais etapas da metodologia os principais aspectos teóricos dos modelos ARIMA focalizando as suas propriedades estatísticas.

O capítulo 3 apresenta a metodologia usada para a realização do estudo. É feita uma descrição dos dados utilizados e da fonte dos mesmos. Apresenta-se, também, as principais etapas para a aplicação dos modelos ARIMA nomeadamente, identificação, estimação e diagnóstico.

No capítulo 4 são apresentados os resultados das estimações dos modelos propostos no estudo e sua respectiva discussão. Ainda neste capítulo encontram-se apresentados os resultados sobre a previsão do Consumo de Energia Eléctrica.

Por fim, no capítulo 5, são apresentadas as principais conclusões e recomendações para futuras pesquisas.

## 1.1. O problema

São vários os factores que influenciam o consumo de um determinado bem ou serviço, e no caso da energia eléctrica, factores *Económicos* (inflação, preço, etc.), *Cronológicos* (sazonalidade, fins-de-semana, períodos de férias, etc.), *Demográficos* (numero de consumidores, índices de mobilidade, etc.), *Sócias* (acontecimentos festivos, desportivos, etc.), *Tecnológicos* (novos electrodomésticos, aquisição de novas tecnologias, etc.) e muitos outros imensuráveis podem determinar a variação nos níveis de consumo. Na maioria dos casos não se torna fácil fazer previsões usando modelos exploratórios que possam incluir todos estes factores, pois surgem por vezes erros de especificação dos próprios modelos. Dai que surge o seguinte problema: *Como prever o Consumo da Energia Eléctrica Facturada sem incluir no modelo os factores que a influenciam?*

## 1.2. Objectivos

### 1.2.1. Geral

- Pretende-se como presente trabalho, desenvolver um modelo de previsão do Consumo da Energia Eléctrica Facturada para a Cidade de Maputo para o período de Junho de 2007 à Maio de 2008;

### 1.2.2. Específicos

Constituem objectivos específicos para este trabalho:

- Verificar a ocorrência de comportamentos periódicos no consumo;
- Verificar a evolução no Consumo da Energia Eléctrica Facturada no período em análise;
- Identificar o modelo da classe ARIMA ou SARIMA que melhor se ajusta aos dados da série;
- Prever o Consumo da Energia Eléctrica Facturada para o período de Junho de 2007 à Maio de 2008.

## CAPITULO II

---

### 2. REVISÃO BIBLIOGRÁFICA

Neste capítulo são apresentados alguns conceitos básicos relacionados com as séries temporais que sustentarão a explicação do método de previsão baseado na metodologia de Box-Jenkins.

#### 2.1. Séries Temporais

**Definição 1:** Segundo Morettin & Tolo (1987), uma série temporal é um conjunto de observações ordenadas no tempo e são compostas por quatro elementos:

1. *Tendência:* verifica o sentido de deslocamento da série ao longo de vários anos.
2. *Ciclo:* movimento ondulatório que ao longo de vários anos tende a ser periódico.
3. *Sazonal:* movimento ondulatório de curta duração, em geral, inferior a um ano; associada, na maioria dos casos, a mudanças climáticas.
4. *Ruído aleatório ou erro:* compreende a variabilidade intrínseca aos dados e não pode ser modelado.

Como uma série temporal tem os dados colectados sequencialmente ao longo do tempo, espera-se que ela apresente correlação serial no tempo (Sentana, 1995). Os modelos de Box-Jenkins, genericamente conhecidos por ARIMA (*Auto Regressive Integrated Moving Averages*) e na literatura em português por Auto-regressivos Integrados de Médias Móveis, são modelos matemáticos que visam captar o comportamento da correlação serial ou autocorrelação entre os valores da série temporal, e com base nesse comportamento realizar previsões futuras. Se essa estrutura de correlação for bem modelada, fornecerá boas previsões (Ribeiro & Werner, 2003).

Segundo Fava (2000), os modelos ARIMA resultam da combinação de três componentes denominados “filtros”: o componente auto-regressivo (AR), o filtro de integração (I) e o componente de médias móveis (MA). Uma série pode ser modelada pelos três filtros ou apenas por um subconjunto deles, resultando assim em modelos não-estacionários e estacionários.

**Definição 2:** Segundo Rossetti (2006), *séries não estacionárias* são aquelas que mudam o comportamento ao longo do tempo.

**Definição 3:** Segundo Gujarati (2006), *séries estacionárias* são aqueles que assumem que o processo está em “equilíbrio”. Ainda, segundo o mesmo autor, um processo é considerado fracamente estacionário se sua média e variância se mantêm constantes ao longo do tempo e a função de autocovariância depende apenas da defasagem entre os instantes de tempo  $t$ , isto é:

$$(i) E(X_t) = \mu$$

$$(ii) \gamma_{t,t} = \sigma_x < \infty$$

$$(iii) \gamma_{t,s} = \gamma_{t-s}$$

onde:  $t$  e  $s$  representam o período:  $t, s = 1, 2, 3, \dots, n$  e  $t \neq s$

$X$  representa a série temporal.

Segundo Siklos (2002), um processo é também considerado fortemente estacionário se todos os momentos conjuntos são invariantes a translações no tempo.

**Definição 4:** Quando a série dos resíduos apresenta média e variância constante no tempo, diz-se que o termo de erro é um *Ruído Branco* (Gujarati, 2006).

Quando um processo estacionário é descrito pelos seus valores passados os modelos resultantes são os auto-regressivo (AR). Num modelo auto-regressivo, a série de dados históricos  $X_t$  é descrita por seus valores passados regredidos e pelo *ruído branco*  $\xi_t$ . Assim, um modelo AR(p) é dado por:

$$X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \xi_t = \sum_{j=1}^p \phi_j X_{t-j} + \xi_t \quad (1)$$

onde  $\phi$  é o parâmetro que descreve como  $X_t$  se relaciona com  $X_{t-i}$  para  $i = 1, 2, \dots, p$ .

Segundo o mesmo autor, o modelo AR(p) dado pela equação 1 pode ser reescrito conforme apresentado na equação 2, utilizando o operador de defasagem  $B^1$ .

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) X_t = \phi(B) X_t = c + \xi_t \quad (2)$$

O modelo auto-regressivo de ordem 1 ou AR(1) é a versão mais simples dessa classe de modelos, sua apresentação algébrica é dada pela equação 3:

$$X_t = c + \phi X_{t-1} + \xi_t \quad (3)$$

<sup>1</sup> O operador B é definido como  $B^n X_t = X_{t-n}$

Para o modelo ser estacionário é necessário que  $|\phi| < 1$  (condição de estacionariedade) e que as autocovariâncias ( $\gamma_k$ ) sejam independentes (Mendes *et al*, 2006). No caso do modelo AR(1), as autocovariâncias são dadas por:

$$\gamma_k = \phi_1^k \gamma_0 \quad (4)$$

e as autocorrelações  $\rho_k$  são dadas pela equação 2.5 (Gujarati, 2006):

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi_1^k \quad (5)$$

Segundo Ribeiro e Werner (2003), para um modelo AR, a função de autocorrelação decai exponencialmente quando  $\phi_1$  é positivo; quando  $\phi_1$  é negativo, a função de autocorrelação também decai exponencialmente, mas apresenta alternância de sinais positivos e negativos.

Porem quando o processo estacionário é descrito pelos termos de erro o modelo resultante é o de Medias moveis (MA).

Segundo

(1999), em um modelo de médias móveis (*moving average*), a série  $X_t$  resulta da combinação dos ruídos brancos  $\xi_t$  do período actual com aqueles ocorridos em períodos anteriores. Assim, um modelo de médias móveis de ordem q ou MA(q) é dado por:

$$X_t = \mu + \xi_t - \theta_1 \xi_{t-1} - \theta_2 \xi_{t-2} - \dots - \theta_q \xi_{t-q} = \mu - \sum_{j=0}^q \theta_j \xi_{t-j}, \theta_0 \equiv 0 \quad (6)$$

onde  $\theta_j$  é o parâmetro que descreve como  $X_t$  se relaciona com o valor  $\xi_{t-j}$  para  $j = 1, 2, \dots, q$  e  $\mu$  é a media de  $X_t$ .

O modelo MA(q) dado pela equação 6 pode ser reescrito, como em 7, utilizando o operador de defasagem  $B$ .

$$X_t = (1 + \theta_1 B + \dots + \theta_q B^q) \xi_t = \theta(B) \xi_t \quad (7)$$

O modelo MA(1) é a versão mais simples dessa classe de modelos. Sua apresentação algébrica é dada pela equação 8:

$$X_t = \mu + \theta_1 \xi_t + \xi_t \quad (8)$$

As autocorrelações  $\rho_k$ , que nada mais são do que as autocovariâncias divididas pela variância, são dadas por:

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = -\frac{\theta_1 \sigma_\xi^2}{(1 + \theta_1^2) \sigma_\xi^2} = -\frac{\theta_1}{(1 + \theta_1^2)} \text{ e } \rho_k = 0 \text{ para } k > 1 \quad (9)$$

Segundo Poskitt (1987), em um modelo MA a função de autocorrelação apresenta apenas a primeira autocorrelação não nula e as demais iguais a zero. A primeira autocorrelação será positiva se  $\theta_1$  for menor que zero e negativa se  $\theta_1$  for maior que zero.

Segundo Nelson (1973), uma propriedade importante da MA(1), proveniente da função de autocorrelação, é que sua "memória" é de somente um período. Uma dada observação, por exemplo  $X_{53}$ , está correlacionada apenas a seu antecessor  $X_{52}$  e a seu sucessor  $X_{54}$ , mas não a qualquer outro membro da série.

Para obter a condição de invertibilidade, isto é, para transformar um modelo MA(1) em um processo invertível, é preciso impor a restrição de que  $|\theta_1| < 1$ . Segundo Abraham e Ledolter (1983), para que um processo MA(q) se torne invertível, é necessário que as raízes da equação:

$\theta(B) \equiv 1 + \theta_1 B + \dots + \theta_q B^q = 0$  sejam maiores que um. Contudo, em alguns casos, pode ser necessário utilizar um grande número de parâmetros em modelos puramente AR ou puramente MA. Nesses casos, é vantajoso misturar os componentes de um modelo AR como os componentes de um modelo MA, gerando, assim, um modelo Auto-Regressivo Medias Moveis (ARMA). Segundo Gujarati (2006), o modelo ARMA(p,q) exigirá um número menor de termos para que seja o mais parcimonioso e pode ser expresso conforme a equação 10:

$$X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \xi_t - \theta_1 \xi_{t-1} - \theta_2 \xi_{t-2} - \dots - \theta_q \xi_{t-q} \quad (10)$$

O modelo ARMA mais simples é o ARMA(1,1), dado pela equação 11:

$$X_t = c + \phi_1 X_{t-1} + \theta_1 \xi_{t-1} + \xi_t \quad (11)$$

A função de autocorrelação do modelo ARMA(1,1) é dada por:

$$\rho_1 = \frac{(1 - \phi_1 \theta_1)(\phi_1 - \theta_1)}{1 + \theta_1^2 + 2\phi_1 \theta_1}, \rho_k = \phi_1 \rho_{k-1} \text{ para } k > 1 \quad (12)$$

Segundo Ribeiro e Werner (2003), a função de autocorrelação do modelo ARMA(p,q) apresenta características da função MA(q) para as defasagens  $k < q$ , pelo fato de a “memória” do componente de médias móveis durar apenas q períodos. Para defasagens maiores que  $k + 1$  as características são iguais às de um modelo AR(p).

Pela definição (2), pode-se afirmar que quando uma série temporal apresenta média e variância dependentes do tempo, é porque ela é não estacionária. Quando tal facto ocorre, está-se perante modelos não estacionários.

Segundo Masih (2002), a não-estacionariedade de uma série implica que:

- a) há inclinação nos dados e eles não permanecem ao redor de uma linha horizontal ao longo do tempo;
- b) a variação dos dados não permanece essencialmente constante sobre o tempo, isto é, as flutuações aumentam ou diminuem com o passar do tempo, indicando que a variância está se alterando.

Como a maioria dos procedimentos de análise estatística de séries temporais supõe que estas sejam estacionárias, será necessário transformá-las caso ainda não sejam. Segundo Morretin & Tolo (1987), a transformação mais comum consiste em tomar diferenças sucessivas da série original até obter uma série estacionária.

A primeira diferença de  $X_t$  é definida por 13.

$$\Delta X_t = X_t - X_{t-1} \quad (13)$$

e o operador  $\Delta$  é chamado de operador de diferença. Generalizando 13, a n-esima diferença da serie é dada por :

$$\Delta^n X_t = \Delta^{n-1} X_t - \Delta^{n-1} X_{t-1} = \sum_{r=0}^n (-1)^r \binom{n}{r} X_{t-r}, \text{ onde } \binom{n}{r} = \frac{n!}{r!(n-r)!} \quad (14)$$

Segundo os mesmos autores, será suficiente tomar uma ou duas diferenças para que a série se torne estacionária.

O número  $d$  de diferenças necessárias para tornar a série estacionária é denominado ordem de integração (Grasa, 1989).

Se  $d$  for um número inteiro não negativo, então serie temporal  $X_t$  é dita ser um processo ARIMA(p,d,q) ou um processo integrado de ordem  $d$  se

$$Y_t = (1 - B)^d X_t = \Delta^d X_t \quad (15)$$

for um processo ARMA causal (Brockwell & Davis, 1996).

Assim, um modelo ARMA(p,q) é um modelo ARIMA(p,0,q). Alternativamente,  $\{X_t\}$  deve satisfazer a seguinte condição:

$$\phi(B)[(1 - B)^d X_t - \mu] = \theta(B)\xi_t,$$

onde  $\{\xi_t\}$  é o ruído branco com media zero,  $\phi(B)$  e  $\theta(B)$  são polinómios de  $p$  e  $q$ , respectivamente,  $\phi(B)$  é um operador estacionário,  $\mu$  é a media de  $\Delta^d X_t$  e  $d$  é a ordem de diferença. A ordem de diferença será 0, 1 ou 2 para a maioria dos processos (Granger & Newbold, 1986).

Os modelos ARIMA exploram a autocorrelação entre os valores da série em instantes sucessivos, mas quando os dados são observados em períodos inferiores a um ano, a série também pode apresentar autocorrelação para uma estação de sazonalidade  $s$ . Os modelos que contemplam as séries que apresentam autocorrelação sazonal são conhecidos como SARIMA.

Segundo Brockwell & Davis (1996), os modelos SARIMA contêm uma parte não sazonal, com parâmetros (p,d,q), e uma sazonal, com parâmetros (P,D,Q)<sub>s</sub> e o modelo mais geral é dado pela equação 16 chegando-se assim aos modelos sazonais multiplicativos ARIMA(p,d,q)x(P,D,Q)<sub>s</sub> ou SARIMA :

$$\phi(B)\Phi(B^s)(1 - B)^d(1 - B^s)^D X_t = \theta(B)\Theta(B^s)\xi_t \quad (16)$$

onde :  $\Phi(B^s) = (1 - \Phi_{1s}B^s - \dots - \Phi_{ps}B^{ps})$ , é a parte auto-regressiva sazonal de ordem P e estação sazonal  $s$ ;

$\Theta(B^s) = (1 - \Theta_{1s}B^s - \dots - \Theta_{qs}B^{qs})$ . é a parte sazonal de médias móveis de ordem Q e estação sazonal  $s$ .

$\phi(B) = (1 - \phi_1B - \dots - \phi_pB^p)$ , é a parte auto-regressiva não-sazonal de ordem p;

$\theta(B) = (1 - \theta_1B - \dots - \theta_qB^q)$ , é a parte não-sazonal de médias móveis de ordem q;

$(1 - B^s)^D$ , é a parte de integração sazonal de ordem D e estação sazonal  $s$ ;

$(1 - B)^d$ , é parte de integração não-sazonal de ordem d;

## 2.2. Teste de Raiz Unitária

Para detectar a não-estacionariedade de uma série, pode-se aplicar os testes estatísticos de raiz unitária. O teste de raiz unitária mais usado é o de Dickey- Fuller aumentado (Menezes, 2004).

Segundo o mesmo autor, o teste ADF também conhecido por teste  $\tau$  (Dickey e Fuller, 1979, 1981), baseia-se na seguinte regressão para a variável  $X_t$ :

$$\Delta X_t = \beta_0 + \beta_1 t + (\rho - 1)X_{t-1} + \sum_{k=1}^p \gamma_k \Delta X_{t-k} + \xi_t \quad (17)$$

sendo a hipótese nula  $\rho = 1$  ou  $\delta = (\rho - 1) = 0$ , usando para este efeito os valores críticos de MacKinnon (1991, 1996).

O número de *lags* ou defasamentos do modelo é escolhido de modo a que os resíduos  $\xi_t \sim \text{iid}(0, \sigma^2_\varepsilon)$ . Note-se que  $\beta_0$  é uma constante,  $\beta_1 t$  captura uma eventual tendência determinística existente nos dados e  $(\rho - 1)X_{t-1}$  modela a respectiva tendência estocástica. Fazendo  $\beta_1 = \gamma_k = 0$ , a equação de Dickey-Fuller 17 reduz-se a um processo AR(1) que pode ser visto como uma realização do processo descrito na equação 3.

O procedimento usual nos testes ADF consiste em testar a hipótese nula  $\rho = 1$  (existência de uma raiz unitária) para a variável em níveis  $X_t$  e em primeiras diferenças  $\Delta X_t$ , contra a hipótese alternativa  $|\rho| < 1$  (estacionariedade). Este teste pode considerar a existência de uma constante e uma tendência linear ( $\tau_t$ ), de uma constante ( $\tau_\mu$ ) ou de nenhuma delas ( $\tau$ ), enquanto componentes determinísticas do processo  $(\beta_0 + \beta_1 t)$ . Para além do termo que contém o coeficiente de autoregressão  $\rho$ , a equação 13 incorpora ainda a possibilidade de existirem termos autoregressivos da variável dependente  $\Delta X_t$ . Se a hipótese nula não for rejeitada em níveis mas o for em primeiras diferenças, então a variável  $X_t \sim I(1)$ , o que significa que a série é integrada de ordem 1. Se a hipótese nula for apenas rejeitada em segundas diferenças, a variável  $X_t \sim I(2)$ , o que significa que a série é integrada de ordem 2 e assim sucessivamente.

Dickey e Fuller (1979, 1981) propuseram dois testes para analisar as hipóteses acima descritas. O primeiro baseia-se na distribuição limite da estimativa OLS de  $\rho$ . A estatística do teste ADF utilizada neste contexto obedece à forma tradicional dos testes  $t$ :

$$t_\tau = \frac{\hat{\rho} - 1}{S_{\hat{\rho}-1}} \quad (18)$$

mas não é assintoticamente normal nem simétrica. Deste modo, os valores críticos usuais da distribuição *t*-Student não são válidos neste âmbito. Fuller (1976), Dickey e Fuller (1981) e mais tarde MacKinnon (1991, 1996) fornecem os valores críticos adequados a esta situação. A verdadeira distribuição dos testes de Dickey-Fuller sob a hipótese nula é dada por (ver Dufrénot e Mignon, 2002):

$$t_{\rho} \underset{(H_0)}{\sim} \frac{W^2(1) - 1}{2\sqrt{\int_0^1 W^2(s) ds}} \quad (19)$$

onde  $W(\cdot)$  representa um processo de Wiener no intervalo  $[0, 1]$ . O segundo, baseia-se na distribuição da estatística  $T(\rho-1)$ , onde  $T$  denota a dimensão da amostra.

Segundo Menezes (2006), para a estatística  $\tau$ , nota-se que  $\tau_r < \tau_\mu < \tau$  e os valores críticos do teste ADF pouco diferem dos valores críticos da distribuição Normal standardizada. Deste modo, o uso indevido destes últimos (ou, indiferentemente, dos valores críticos da distribuição *t*-Student) conduziria a uma sobre-rejeição da hipótese nula, aumentando o problema há medida que se introduzem mais componentes determinísticas no modelo.

Um resultado importante relativo aos testes de Dickey-Fuller é que a distribuição assintótica da estatística  $t_\tau$  é independente do número de desfasamentos das primeiras diferenças incluídos na regressão ADF. Estes, contudo, são necessários para eliminar qualquer autocorrelação de ordem superior remanescente na variável residual  $e$ , deste modo, validar o pressuposto de resíduos  $\xi_t \sim iid$  que está na base do uso do método dos mínimos quadrados. Para além disso, a inclusão no modelo de uma componente de médias móveis (do inglês *Moving Average* ou MA) com um número adequado de termos defasados das primeiras diferenças, não altera a validade assintótica do teste ADF (Said e Dickey, 1984).

### 2.3. O método de Box-Jenkins

Segundo Morretin & Toloï. (1987), a construção dos modelos Box-Jenkins é baseada em um ciclo iterativo, no qual a escolha do modelo é feita com base nos próprios dados. Segundo Box & Jenkins (1976), são três as etapas para construção do modelo:

#### 1. Identificação:

Consiste em descobrir qual dentre as várias versões dos modelos de Box-Jenkins, sejam eles sazonais ou não, descreve o comportamento da série. A identificação do modelo a ser estimado

ocorre pelo comportamento das funções de autocorrelações (ACF) e das funções de autocorrelações parciais (PACF) e ou ainda pelo uso de um critério de selecção de modelos.

Para utilizar a primeira estratégia é necessário reconhecer modelos AR, MA e ARMA por meio das características da ACF e da PACF. A Tabela 1 resume as propriedades da ACF e da PACF para diversos modelos ARIMA (Mills, 1990). Para modelos sazonais, deve ser analisado o comportamento da ACF e da PACF próximo da defasagem sazonal, por exemplo, defasagem 12 para dados mensais e defasagem 4 para dados trimestrais. A Tabela 2 resume as propriedades da ACF e da PACF para modelos SARIMA (Bowerman & O'Connell, 1987).

**Tabela 1.** Propriedades da ACF e da PACF para vários modelos ARIMA

MODELO	ACF	PACF
(P,D,0)	Decaimento	Picos nas defasagens s, 2s, ..., Ps e corte após Ps
(0,D,Q)	Picos nas defasagens s, 2s, ..., Qs e corte após Qs	Decaimento
(P,D,0) Ou (0,D,Q)	Picos nas defasagens s, 2s, ..., Qs e corte após Qs	Picos nas defasagens s, 2s, ..., Ps e corte após Ps
(P,D,0) e (0,D,Q)	Decaimento rápido na defasagem sazonal	Decaimento rápido na defasagem sazonal
Nenhum operador sazonal	Valores pequenos em todas as defasagens sazonais (não há picos)	Valores pequenos em todas as defasagens sazonais (não há picos)

**Tabela 2.** Propriedades da ACF e da PACF para modelos SARIMA

MODELO	ACF	PACF
(1,d,0)	Decaimento exponencial ou oscilatório	$\phi_{kk} = 0$ para $k > 1$
(2,d,0)	Decaimento exponencial ou senoidal	$\phi_{kk} = 0$ para $k > 2$
(p,d,0)	Decaimento exponencial e/ou senoidal	$\phi_{kk} = 0$ para $k > p$
(0,d,1)	$\rho_k = 0$ para $k > 1$	Decaimento exponencial
(0,d,2)	$\rho_k = 0$ para $k > 2$	Decaimento exponencial ou senoidal
(1,d,q)	$\rho_k = 0$ para $k > q$	Dominado pela combinação linear de decaimento exponencial e/ou senoidal
(1,d,1)	Decaimento exponencial a partir da defasagem 1	Dominado por decaimento exponencial a partir da defasagem 1
(p,d,q)	Decaimento exponencial e/ou senoidal depois da defasagem q-p	Dominado por decaimento exponencial ou senoidal depois da defasagem q-p

A segunda estratégia para identificar os valores de p, d e q é utilizar um critério de informação que seleccione os modelos por meio de um conjunto de "regras" (Mills, 1990).

Os critérios de selecção para modelos ARIMA mais utilizados são o *AIC* (*Akaike information criterion*), o *AICC* (*Akaike information criterion corrected*) e o *BIC* (*Bayesian information criterion*).

Segundo Hurvich *et al* (1989), na literatura sobre processos autoregressivos integrados media móvel é frequente encontrar referência a estes dois critérios acima referidos que permitem escolher entre modelos alternativos, especificados com um número diferente de desfasamentos da variável endógena. Estes critérios, conhecidos como critérios de informação, visam escolher o número de parâmetros que minimiza o valor do respectivo critério de informação. A função a minimizar deverá, por isso, conter informação sobre os ganhos de precisão do modelo dentro da amostra e sobre as perdas de capacidade predictiva fora da amostra.

Para Dionísio *et al* (2004), a filosofia subjacente aos critérios de informação considera, portanto, dois factores competitivos que influenciam a escolha do modelo: (1) uma função da soma dos quadrados dos resíduos obtidos na estimação do modelo e (2) uma penalização resultante da perda de graus de liberdade devida à adição de parâmetros ao modelo.

Segundo o mesmo autor, por cada parâmetro adicional relativo aos desfasamentos, a soma dos quadrados dos resíduos diminui e o valor da penalização aumenta. A escolha óptima ocorre, então, quando a redução da soma dos quadrados dos resíduos não compensa mais o aumento da penalização introduzida.

Os critérios de informação, em geral, apenas se diferenciam no que respeita ao termo que captura a penalização do processo, sendo os dois mais populares os seguintes:

$$\begin{aligned} AIC(p) &= \ln|\hat{\Omega}(p)| + (pN^2) \frac{2}{T} \\ SBC(p) &= \ln|\hat{\Omega}(p)| + (pN^2) \frac{\ln T}{T} \end{aligned} \quad (20)$$

$AIC(p)$  denota o critério de informação Akaike (1974);  $SBC(p)$  denota o critério de informação Bayesiano proposto por Schwarz (1978). A matriz  $\Omega$  é a matriz de covariâncias,  $p$  é a ordem do processo (ou número de *lags*) e  $T$  o número de observações corrigido pelo número de desfasamentos. Note-se que  $pN^2$  representa o número total de parâmetros a estimar.

Segundo Braumann (2005), o critério de informação de Akaike tem a desvantagem do respectivo estimador não ser assintoticamente consistente ao contrário do critério de informação Bayesiano que, por isso, tende a ser mais parcimonioso na escolha do número óptimo de desfasamentos. Um

outro critério de informação igualmente usado neste contexto é o critério de Hannan-Quinn (1979) o qual, à semelhança do critério SBC, também utiliza um estimador consistente:

$$HQIC(p) = \ln|\hat{\Omega}(p)| + (pN^2) \frac{2}{T} \ln(\ln T) \quad (21)$$

Uma ideia interessante introduzida, neste contexto, por Poskitt e Tremayne (1987) refere-se ao conceito de *carteira de modelos*. A ideia base consiste em comparar o modelo com o processo de ordem  $p_1$  seleccionado, recorrendo à estatística:

$$\mathfrak{R} = \exp\left[-\frac{1}{2}T(SC(p_1) - SC(p))\right], \quad (22)$$

onde SC representa o critério de selecção em análise. Usando  $\sqrt{10}$  como limite superior, a escolha do "critério mais adequado" é a que permite otimizar a carteira de modelos que melhor se ajusta ao processo seleccionado (Poskitt e Tremayne, 1987).

Na prática, a selecção de modelos é feita calculando o valor do critério (o BIC, por exemplo) para todos os modelos ARIMA associados aos valores de  $p$ ,  $d$  e  $q$  de forma que  $p, q = 0, 1, 2, 3, 4, 5$  e  $d = 0, 1$ . Assim, selecciona-se o modelo que apresenta o menor valor do BIC e modelos alternativos cuja diferença para o valor mínimo do BIC seja inferior a 2 (Brockwell & Davis, 1996).

Para a modelagem SARIMA, a quantidade de modelos investigados é maior, pois além dos valores de  $p$ ,  $d$  e  $q$ , deve-se incluir ainda os valores para  $P, Q = 0, 1, 2$  e  $D = 0, 1$ . Porém, os modelos são seleccionados pelos mesmos critérios que os utilizados para os modelos ARIMA.

## 2. Estimação:

Consiste em estimar os parâmetros  $\phi$  e  $\Phi$  do componente auto-regressivo, os parâmetros  $\theta$  e  $\Theta$  do componente de médias móveis e a variância de  $\xi_t$ , utilizando o método de máxima verosimilhança (ML)<sup>2</sup>.

A estimação da ML é difícil e geralmente requer muito tempo de processamento computacional. Desta forma, existem alternativas que aproximam a função de máxima verosimilhança. Duas dessas alternativas são o MQE (mínimos quadrados exactos) e o MQC (mínimos quadrados condicional). Contudo, alguns estudos têm sugerido que o método de máxima verosimilhança é superior aos acima referidos (Ansley & Newbold, 1980).

Para a modelagem SARIMA o procedimento é idêntico ao mostrado para a modelagem ARIMA, com a superioridade da estimação por máxima verosimilhança sendo ainda mais pronunciada para modelos sazonais (Ansley & Newbold, 1980).

### 3. Verificação e Diagnostico:

Consiste em avaliar se o modelo estimado é adequado para descrever o comportamento dos dados.

A correcta especificação de um modelo ARIMA ou SARIMA é verificada no termo  $\xi_T$ , pois ele deve constituir um processo *ruído branco* (Granger & Newbold, 1986). Assim, a verificação da adequabilidade do modelo é efectuada nas autocorrelações amostrais dos erros  $\xi_t$ , as quais seguem

assintoticamente uma distribuição normal, com média zero e desvio padrão  $1/\sqrt{n}$ , se forem provenientes de um *ruído branco*. Como os erros verdadeiros  $\{\xi_T\}$  não são conhecidos, a inferência baseia-se nas estimativas dos erros, os resíduos  $\hat{\xi}_T$ .

Dessa forma, se o modelo estiver correctamente especificado, os resíduos não devem apresentar correlação serial, pois toda a dinâmica dos dados já foi capturada pelo modelo (Cribari-Neto, 2000). Segundo Granger & Newbold (1986), a autocorrelação amostral dos resíduos de ordem  $j$  é calculada da seguinte forma:

$$r_j(\hat{\xi}) = \frac{\sum_{t=j+1}^T \hat{\xi}_t \hat{\xi}_{t-j}}{\sum_{t=1}^T \hat{\xi}_t^2} \quad (23)$$

e os valores das autocorrelações residuais devem estar contidos no intervalo de confiança assintótico de 95%.

$$\left[ -\frac{2}{\sqrt{T}}; \frac{2}{\sqrt{T}} \right] \quad \text{onde } T \text{ indica o número de observações da série.}$$

Em adição ao exame das autocorrelações individuais dos resíduos, pode ser usado um teste conjunto das primeiras  $m$  autocorrelações que é o teste Ljung-Box. Tal teste compara o valor de:

$$Q = T(T+2) \sum_{j=1}^m (T-j)^{-1} r_j^2(\hat{\xi}_j) \quad (24)$$

com os valores da tabela de distribuição do  $X^2$  (qui-quadrado), com  $(m - p - q)$  graus de liberdade.

Segundo Granger *et al* (1986), as hipóteses do teste LB são as seguintes:

H<sub>0</sub>: Os resíduos estimados são puramente aleatórios (*o modelo é adequado*);

H<sub>a</sub>: Os resíduos estimados não são puramente aleatórios (*o modelo não é adequado*);

sendo que a rejeição da hipótese nula é feita para valores de  $Q$  maiores que o valor crítico da tabela  $X^2$ . O valor de  $m$  deve ser pelo menos igual a  $\sqrt{T}$  (Cribari-Neto, 2000).

## CAPITULO III

---

### 3. MATERIAL E MÉTODOS

Neste capítulo é apresentado o material necessário para a realização do presente estudo. Fez-se a ilustração das etapas principais da aplicação do método de Box-Jenkins.

#### 3.1. Material

##### 3.1.1. Dados usados

Para a realização do presente trabalho recolheu-se os dados do Consumo da Energia Eléctrica Facturada da tarifa doméstica dos clientes utilizando a linha de baixa tensão que correspondiam a Cidade de Maputo.

Os dados mensais referem-se ao período de Janeiro de 2000 a Maio de 2007 e estão medidos em kWh. No total a série comporta 89 observações.

Para a análise de dados foram usados dois pacotes estatísticos específicos, que são o Eviws V.4.1 o SPSS V.13.0.

##### 3.1.2. Fonte de dados

A base de dados usada é secundária e foi fornecida pelo Departamento de Estatística e Informática da Electricidade de Moçambique, Área comercial da Cidade de Maputo. Os dados são referentes a Dependências que fazem as respectivas facturações e enviam mensalmente ao departamento acima referido.

#### 3.2. Métodos

##### 3.2.1. Análise dos dados

A análise dos dados consistiu primeiro em análises gráficas para que se pudesse verificar possíveis indícios de não estacionaridade da série e sazonalidade tal como a tendência.

Estimaram-se três Modelos: ARIMA(0,1,1), ARIMA(1,1,0) e ARIMA(1,1,1).

### 3.2.1.1. Teste de Raiz Unitária

Neste trabalho foi usado a estatística de Dickey-Fuller Aumentado (ADF) para testar a presença de ou não de raiz unitária na série.

O teste de Dickey-Fuller baseia-se no seguinte modelo:

$$\Delta X_t = \alpha + \beta t + \eta X_{t-1} + \xi_t \quad (25)$$

$$\text{Sendo que } \eta = \sum_{i=1}^p \rho_i - 1$$

$X$  denota a variável dependente e  $\Delta$  o operador de diferença  $\Delta X_t = X_t - X_{t-1}$ . Os parâmetros a serem estimados são  $\alpha, \beta, \eta$ . As estatísticas  $\tau_r$  e  $\tau_\mu$  e  $\tau$  apresentadas por Dickey & Fuller (1998) corresponde ao teste  $t$  para a estimativa do coeficiente da variável  $X_{t-1}$  da equação (25). Estas estatísticas são especificadas para um modelo que inclui o parâmetro de deslocamento e uma tendência ( $\tau_r$ ), um modelo incluindo apenas constante ( $\tau_\mu$ ) e um modelo sem constante e sem tendência ( $\tau$ ). As hipóteses testadas nesses modelos correspondem a uma hipótese nula de que a série não é estacionária ( $H_0 : X_t$  não é  $I(0)$  ou  $\eta = 0$ ); contra a hipótese alternativa de que a série não é integrada, ou seja, a série é estacionária ( $H_1 : X_t$  é  $I(0)$ ).

Foram também utilizadas as estatísticas  $\phi_3$  e  $\phi_1$ , obtidas por Dickey & Fuller (1976 e 1981), que testam se os coeficientes da variável tendência e o coeficiente da variável  $X_{t-1}$  e se a constante respectivamente são estatisticamente iguais a zero em (25).

Pode-se incorporar na equação (25) valores defasados da variável endógena ( $X_t$ ) a fim de se eliminar a presença de autocorrelação entre os termos de erro, ou seja,

$$\Delta X_t = \alpha + \beta t + \eta X_{t-1} + \sum_{i=1}^{p-1} \lambda_i \Delta X_{t-i} + \xi_t \quad (26)$$

$$\text{sendo que } \lambda_i = - \sum_{j=i+1}^p \rho_j$$

sendo que  $p$  é a ordem do modelo auto-regressivo ou o número de defasagens suficientes para que os resíduos resultantes sejam não correlacionados (*white noise*). Neste caso, tem-se o teste ADF denominado de Dickey-Fuller Aumentado, (Engle, 1997).

### 3.2.1.2. Identificação

Esta fase compreendeu a análise gráfica dos correlogramas para que se pudesse identificar um padrão ou comportamento dos mesmos na tentativa de se verificar o possível modelo a ser estimado. Como passo inicial, identificou-se o modelo ARIMA(0,1,1). Dada a complexidade desta etapa optou-se pela identificação de outros dois modelos que são o ARIMA(1,1,0) e o ARIMA(1,1,1) para que posteriormente se procedesse a comparação dos mesmos.

Para a construção dos correlogramas, usaram-se as funções de autocorrelações (ACF) e autocorrelações parciais (PACF).

$$(ACF) \tau_k = \frac{\sum_{t=k+1}^T (X_t - \bar{X})(X_{t-k} - \bar{X})}{\sum_{t=1}^T (X_t - \bar{X})^2}$$

$$(PACF) \phi_k = \begin{cases} \tau_1, k=1 \\ \frac{\tau_k - \sum_{j=1}^{k-1} \phi_{k-1,j} \tau_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1,j} \tau_{k-1}}, k > 1 \end{cases}$$

Em seguida para se testar se as  $k$  autocorrelações e autocorrelações parciais eram estatisticamente significativas usou-se a estatística  $Q$  de Ljung-Box e os valores de  $p$  para verificar a significância dos mesmos.

A estatística  $Q$ , testa a hipótese nula de que as primeiras  $k$  auto correlações são iguais a zero ( $h_0 : \tau_k = 0$ ) e ela é calculada pela equação 21 apresentada no capítulo 2.

### 3.2.1.3. Estimação

A estimação foi feita pelo método da máxima verosimilhança, sendo necessário recorrer a função de densidade de probabilidade da distribuição de Gumbel (f.d.p), da variável aleatória  $x$ , associada aos valores do cfe e é dada por:

$$f(x, \beta_i, \alpha) = \frac{1}{\alpha} \exp \left\{ - \left( \frac{x - \beta_i}{\alpha} \right) - \exp \left[ - \left( \frac{x - \beta_i}{\alpha} \right) \right] \right\} \quad (27)$$

em que  $x$  é a variável associada aos valores do cfe do período em análise,  $\beta$  é denominado parâmetro de posição e  $\alpha$  é o parâmetro de deslocamento.

A função de distribuição acumulada da variável  $x$ , para os valores da cfe é dada por:

$$F(x, \beta_i, \alpha) = \int_0^x f(x, \beta_i, \alpha) dx = \exp \left\{ - \exp \left[ - \left( \frac{x - \beta_i}{\alpha} \right) \right] \right\} \quad (28)$$

A probabilidade de que ocorra um consumo maior que um certo valor  $x$  é dada por :

$$P = 1 - F(x, \beta_i, \alpha)$$

Os estimadores dos parâmetros  $\beta_i$  e  $\alpha$  pelo método da máxima verosimilhança (EMV) são obtidos a partir da f.d.p conjunta da distribuição de Gumbel e são respectivamente:

$$\hat{\beta}_i = -\hat{\alpha} \ln \left[ \frac{1}{n} \sum_{j=1}^n \exp \left( \frac{-x_j}{\hat{\alpha}} \right) \right] \quad (29)$$

$$\hat{\alpha} = \frac{1}{n} \sum_{j=1}^n x_j - \frac{\sum_{j=1}^n \exp \left( \frac{-x_j}{\hat{\alpha}} \right)}{\sum_{j=1}^n \exp \left( \frac{-x_j}{\hat{\alpha}} \right)} \quad (30)$$

em que  $\ln$  é o operador do logaritmo neperiano.

Mais como os valores de  $\hat{\beta}_i$  e  $\hat{\alpha}$  dos EMV não possuem expressões explícitas, exigindo cálculos numéricos interactivos, primeiro determinou-se  $\hat{\alpha}$  como solução de 4 e substituiu-se esse valor em 3 para obter os  $\hat{\beta}_i$ .

Visto que  $\hat{\alpha}$  não pode ser obtido analiticamente, e conforme Vivanco (1994), o procedimento interactivo de Newton-Raphson é recomendável neste caso, a estimação destes parâmetros foi feita usando o programa elaborado no software Eviws.

### 3.2.1.4. Diagnostico

A verificação das propriedades dos resíduos  $\xi_i \sim IIDN(0, \sigma^2)$ , para diagnosticar se os modelos seleccionados se ajustam bem aos dados da serie do cfe e se os resíduos são puramente aleatória ou se são um ruído branco, foi feita através da análise gráfica da série dos resíduos e os testes foram usados para confirmação das análises.

O teste de K-S foi usado para testar a normalidade dos resíduos sobre a hipótese nula de que ( $h_0 : \xi \sim IIDN$ ). Segundo Boné *et al* (2002) o teste K-S apresenta uma vantagem perante os outros testes (*Jaquerta Berra por exemplo*) porque ele pode ser aplicado sem depender de um número específico de observações para que seja valido.

Contudo para verificar se a media dos resíduos obedecia ao pressuposto, o teste  $t$  foi usado sobre a hipótese nula de que  $h_0 : \mu_\xi = 0$  e a estatística usada foi  $t$ -Student que é dada por:

$$t = \frac{\bar{X}_\xi - \mu_\xi}{\frac{s_\xi}{\sqrt{N_\xi}}} \quad (31)$$

### 3.2.1.5. Teste LM ao modelo ARCH

Para detectar se a variância dos resíduos dos modelos estimados era ou não constante (*homoscedastico*), foi usado o teste LM por forma a detectar a existência ou não de um efeito ARCH nos modelos.

Segundo Masih (2002), o teste ao modelo ARCH baseia-se na equação seguinte:

$$\hat{\xi}_t^2 = \omega + \alpha \hat{\xi}_{t-1}^2 + \gamma \quad (32)$$

onde:  $\hat{\xi}_t^2$  é variância do erro ( $h_t^2$ ) no tempo  $t$

$\hat{\xi}_{t-1}^2$  é a variância do erro ( $h_{t-1}^2$ ) no tempo  $t-1$

$\gamma$  é o erro.

Na verdade o que se faz na pratica é verificar se " $\alpha$ " é estatisticamente significativo, tendo como hipóteses as seguintes alternativas:

$H_0 : \alpha = 0 \Rightarrow$  Hipótese de homoscedasticidade

$H_1 : \alpha \neq 0 \Rightarrow$  Hipótese de heterocedestacidade

Segundo Siklos (2002), para testar as hipóteses acima referidas, o teste baseia-se numa distribuição Qui-Quadrado, com  $P$  graus de liberdade, dado por  $R^2 * n \sim X_p^2$

onde:  $P \Rightarrow$  é o numero de defasagens em  $\hat{\xi}_t^2$

$n \Rightarrow$  é número de observações

$R^2 \Rightarrow$  é o coeficiente de determinação da equação estimada

### 3.2.2. Comparação dos modelos

O desempenho na previsão, para os modelos ARIMA(1,1,0) e ARIMA(0,1,1), foi avaliado a partir das seguintes estatísticas ou medidas, sugeridas por McClain e Humphreys (1996).

$$\text{Raiz do erro quadrático médio (REQM)} = \sqrt{\frac{1}{h+1} \sum_{t=S}^{s+h} (\hat{X}_t - X_t)^2} \quad (33)$$

onde  $h$  é o número de períodos de previsão,  $S$  é o tamanho da amostra,  $\hat{X}_t$  é o consumo previsto e  $X_t$  o consumo actual.

$$\text{Erro absoluto médio (EAM)} = \frac{1}{h+1} \sum_{t=S}^{s+h} |\hat{X}_t - X_t| \quad (34)$$

$$\text{Erro absoluto médio percentual (EAMP)} = 100 \times \frac{1}{h+1} \sum_{t=S}^{s+h} \left| \frac{(\hat{X}_t - X_t)}{X_t} \right| \quad (35)$$

$$\text{Coeficiente de desigualdade de Theil (CDT)} = \frac{REQM}{\sqrt{\frac{1}{h+1} \sum_{t=S}^{s+h} \hat{X}_t + \frac{1}{h+1} \sum_{t=S}^{s+h} X_t}} \quad (36)$$

As duas primeiras estatísticas são usadas como medidas relativas para comparar a qualidade da previsão (ou ajuste) para a mesma série em diferentes modelos; quanto menor o erro, melhor a habilidade preditiva do modelo.

As duas últimas são invariantes em relação a escala de medida da variável dependente. O coeficiente de desigualdade de Theil sempre situa-se entre 0 e 1, onde 0 indica um perfeito ajuste (Braumann, 2005).

Também foram usadas as estatísticas de Akaike e Schwarz obedecendo o critério dos valores mínimos uma vez que estas medidas são de extrema penalidade, isto é, o modelo que apresenta valores mais baixos para esta estatística é que apresenta melhor desempenho (Gujarat, 2006).

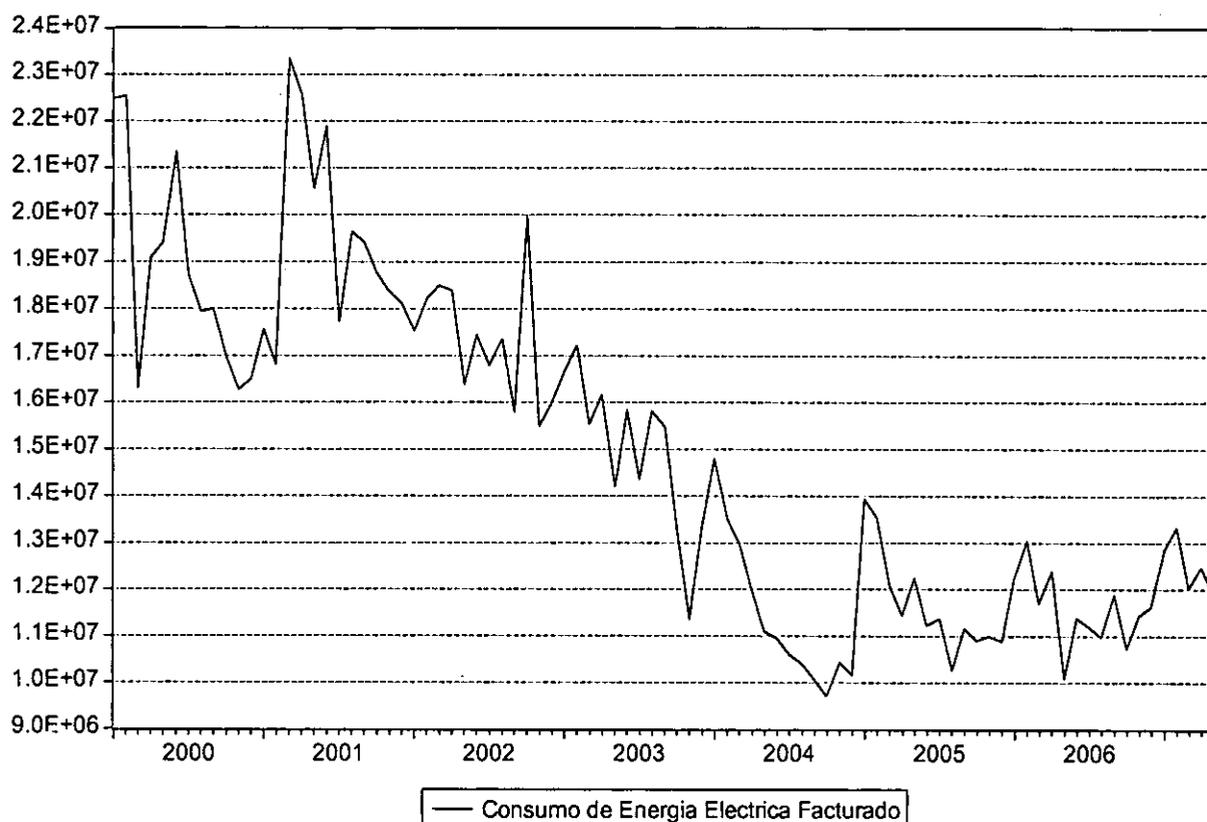
## CAPITULO IV

### 4. RESULTADOS E DISCUSSÃO

#### 4.1. Séries e representação gráfica

A figura 1, mostra a existência de três níveis no consumo durante o período em análise. O primeiro período vai de Janeiro de 2000 à Maio de 2001 em que o consumo aparenta estar em volta de uma valor constante embora haja oscilações no consumo nesse período, com o pico mais baixo em Junho de 2001 e o mais elevado em Fevereiro de 2001. O segundo período vai de Junho de 2001 a Novembro de 2004 onde se observou uma tendência decrescente significativa no consumo. O terceiro e ultimo nível, compreende o período de Dezembro de 2004 a Maio de 2007 onde o consumo parece ter se estabilizado por volta dos 12.000.000,00 kWh ao mês.

Figura 1. Evolução do Consumo da Energia Eléctrica Facturada (Janeiro de 2000 à Maio de 2007).



O decréscimo observado no segundo nível identificado anteriormente, pode estar associado ao facto de os consumidores, principalmente os do sector domestico, estarem preferir a usar o serviço pré pago do que o facturado e consequentemente diminuição dos clientes do serviço facturado.

Uma análise feita empiricamente pode constatar que este decréscimo é acompanhado por uma diminuição no número de clientes usando a energia eléctrica facturada<sup>2</sup>.

Estes níveis levaram a suspeitas de existência, de não estacionaridade na série, uma vez que o consumo não se manteve estável em volta de um certo valor no período em análise.

No entanto, durante o mesmo período o consumo médio mensal foi de 14.879.506,00 kWh ao mês e a dispersão relativamente ao consumo médio foi de 3.626.783,00 kWh. Neste período o consumo mais elevado observou-se no mês de Março de 2001 e o registo foi de 23.340.916,00 kWh e o consumo mínimo foi de 9.694.054,00 kWh no mês de Outubro de 2004.

A figura 2, mostra a existência de igualdade no consumo médio mensal<sup>3</sup>, uma vez que nenhum valor médio esteve fora dos intervalos de confiança dos respectivos meses, evidenciando assim a ausência de factores sazonais.

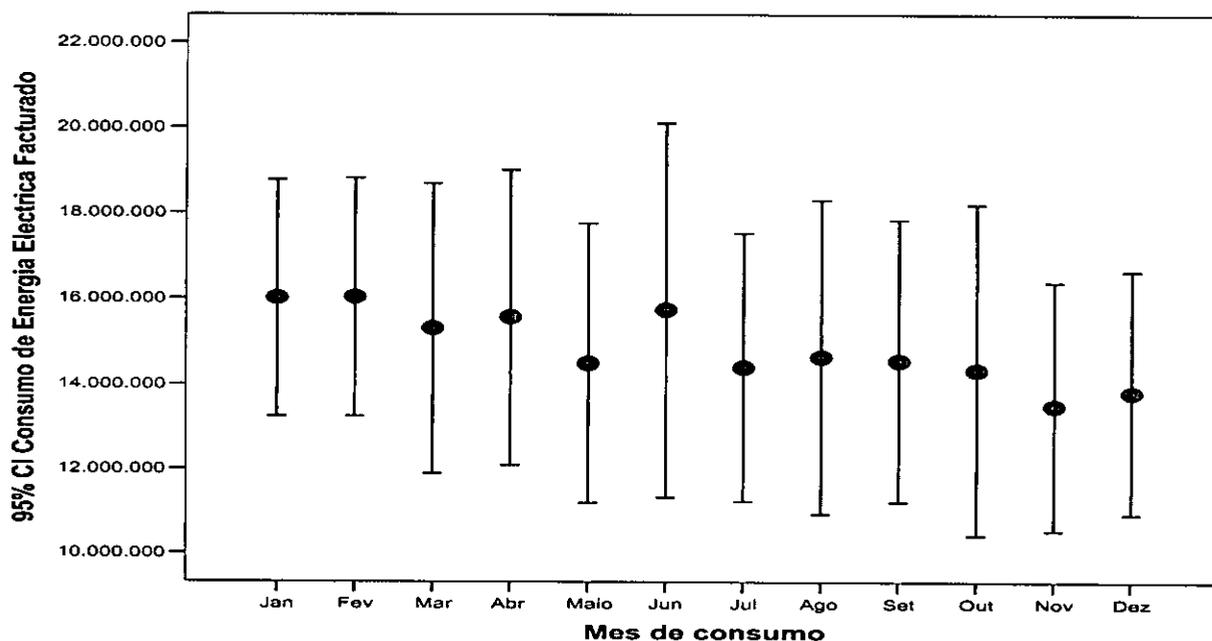
Porém, como ilustra a figura 1, o mês de Junho apresentou uma maior variabilidade no consumo de energia eléctrica uma vez que, para este mês o intervalo de confiança apresentou a maior amplitude relativamente aos restantes meses.

A ausência de factores sazonais pode ser confirmada pela figura 3, pois ela não apresenta o padrão de uma série sazonal.

<sup>2</sup> Para mais informações sobre o número de clientes consultar as tabelas em anexo B.

<sup>3</sup> Procedeu-se ao teste F para a verificação de diferença no consumo médio mensal, que não rejeita a hipótese de igualdade do consumo médio mensal ( $H_0 : \mu_J = \mu_F = \mu_M = \mu_A = \dots = \mu_D$ ) o que comprovar as suspeitas e assim as análises irão ignorar um comportamento sazonal na série. Ver anexo A, tabela 11.

Figura 2. Consumo da Energia Eléctrica Facturada média mensal e os respectivos Intervalos de Confiança



A figura 3, apresenta picos significativos na função de autocorrelação e autocorrelação parcial, evidenciando assim, mais uma vez, a não estacionaridade da série em nível, sugerindo a aplicação de modelos não estacionários com o filtro de integração (I) chamados de ARIMA como se pode constatar anteriormente com a figura 1.

Uma vez que as análises gráficas tenham evidenciado a não estacionaridade da série, procedeu-se ao teste de raiz unitária de Dikey-Fuller Aumentado para a confirmação do mesmo.

Figura 3. Correlograma da Função de Autocorrelação da serie do Consumo da Energia Eléctrica Facturada em nível

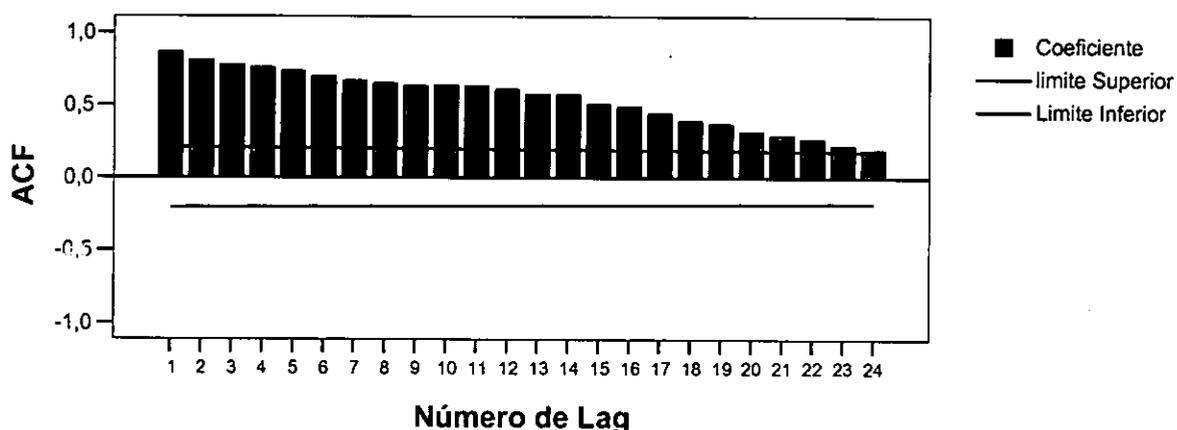
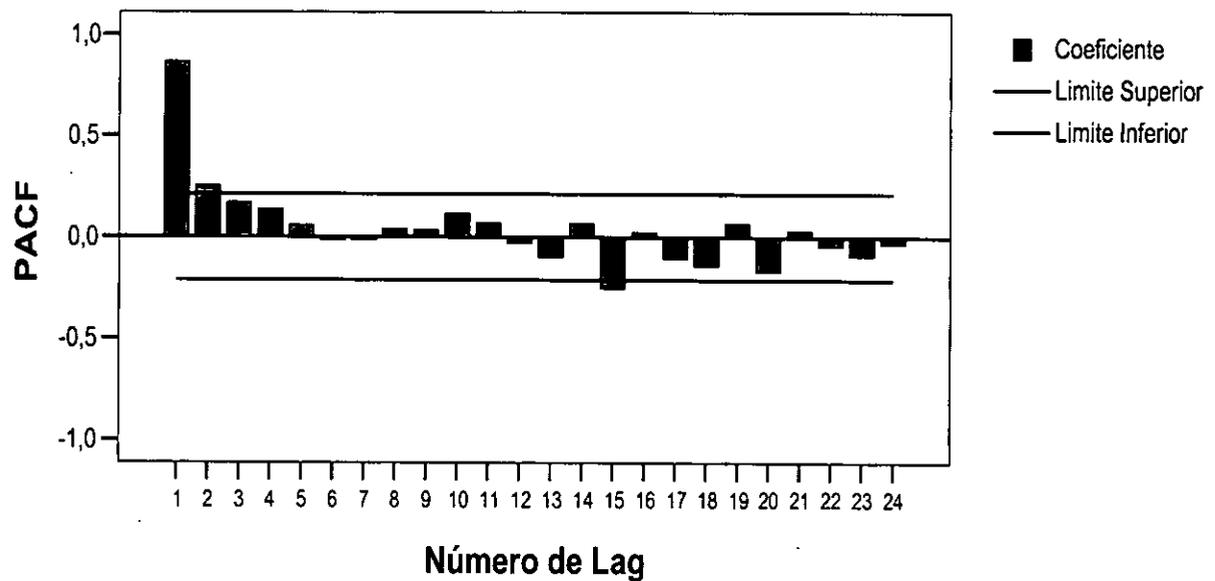


Figura 3.1. Correlograma da Função de Autocorrelação Parcial da serie do Consumo da Energia Eléctrica Facturada em nível



#### 4.2. Teste de Raiz Unitária

As estatísticas da tabela 1, mostraram que pelos valores de  $t$  todos os parâmetros são estatisticamente significativos pois os valores  $p^4$  são todos eles inferiores a 5%.

No entanto como o teste ADF não considera somente a estatística  $t$  e os  $p$  values, mas sim a estatística  $t$  é comparada com as estatísticas sugeridas por Mackinno em 1996, não se rejeitou a 1% e a 5% a hipótese nula ( $H_0 : \delta = 0$ ) de que a regressão estimada na mesma tabela possui uma raiz unitária, isto porque o valor absoluto da estatística  $t$  calculada foi inferior ao valor absoluto da estatística de Mackinno a estes níveis de significancia (Ver tabela 3). Assim, uma vez não rejeitada a hipótese nula, confirmaram-se as suspeitas de que a série do CFE é não estacionária em nível.

Tabela 3. Estatísticas da Regressão de ADF

Teste da Equacao de Dickey-Fuller Aumentado  
Variavel Dependente: D(CFE)

	Coeficientes	Erro Padrao	Estatística t	prob
CFE(-1)	-0,33	0,10	-3,39	0,00
D(CFE(-1))	-0,22	0,11	-2,00	0,05
C	6.389.554,00	2.036.827,00	3,14	0,00
T	-35.327,75	13.664,35	-2,59	0,01

<sup>4</sup> O valor  $p$ , é também conhecido como a probabilidade de rejeição da hipótese nula

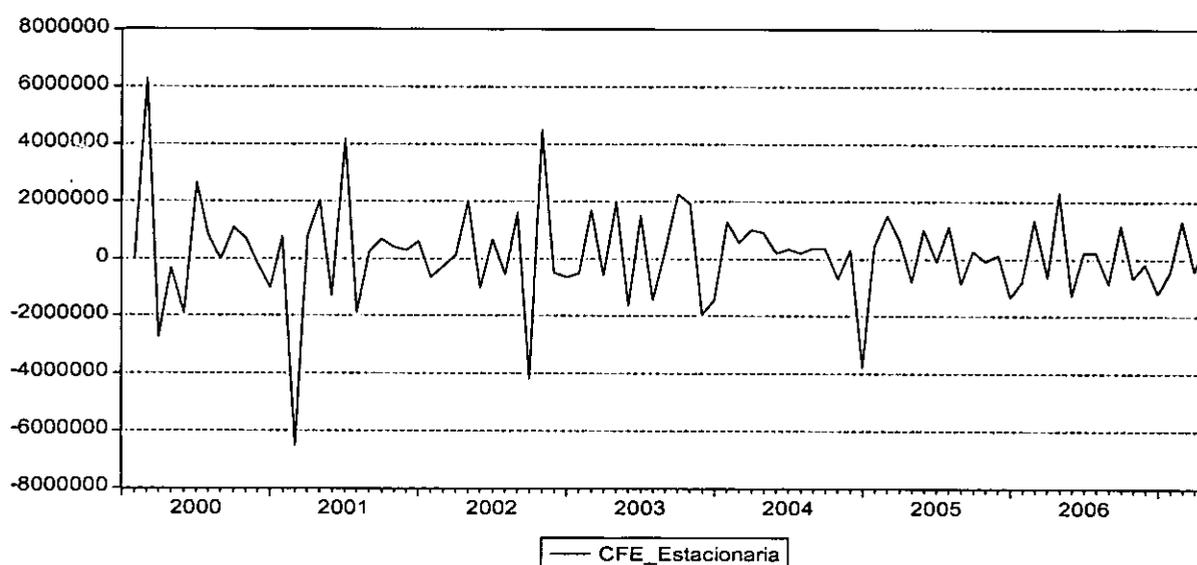
Tabela 4. Teste de raiz unitária de ADF

	t-Statistic	Prob.*
Estatística do teste de Dickey/Fuller Aumentado	-3,38	0,06
Valores críticos do teste:		
1% level	-4,06	
5% level	-3,46	
10% level	-3,16	

\*MacKinnon (1996)

Observando a figura 4 da série em primeira diferença, pode-se verificar facilmente que ela durante o período em análise estabiliza-se em volta de zero. É de referir que ao se transformar a série em primeira diferença perdeu-se uma observação. Assim, a série a passou a comportar 88 observações.

Figura 4 . Série do Consumo da Energia Eléctrica Facturada em primeira diferença



Desta forma, tal como anteriormente estimou-se a equação de regressão de Dickey-Fuller da série da primeira diferença que se encontra na tabela 5. Pela tabela 6, pode-se verificar que o coeficiente (-1.383.869,00), é estatisticamente significativo pois o valor da estatística t para o mesmo coeficiente é maior em termos absolutos. Assim, não se rejeita a hipótese nula, isto é, a série da primeira diferença não tem uma raiz unitária, logo, a série do CFE é estacionária na primeira diferença ( $CFE \rightarrow I(1)$ ).

Tabela 5. Estatísticas da Regressão de ADF

Teste da Equação de Dickey-Fuller Aumentado  
Variável Dependente : D (CFE,2)

Variável	Coefficiente	Erro Padrão	Estatística - t	Prob.
D(CFE(-1))	-1.383.869,00	0,10*	-13,72	0,00
C	-411.697,80	353.517,40	-1,16	0,25
T	5.433,08	684.737,00	0,79	0,43

\* significativa a 1%, 5% e 10%

Tabela 6. Estatísticas do teste de ADF

Hipótese nula: D (CFE) tem raiz unitária		t-Statistic	Prob.*
Estatística do Teste de Dickey-Fuller Aumentado		-13,73	0,00
Valores críticos	1% level	-4,07	
	5% level	-3,46	
	10% level	-3,16	

\*MacKinnon (1996)

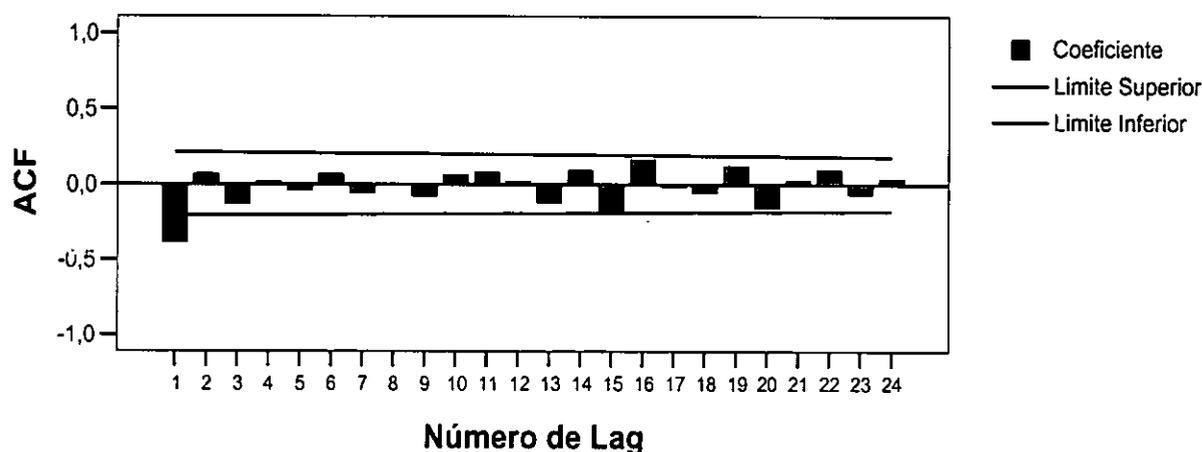
### 4.3. Identificação

Conforme visto no capítulo 2, depois de se verificar a estacionaridade da série torna-se agora necessário identificar os valores  $p, d, q$  e o padrão das funções de autocorrelação e autocorrelação parcial. Contudo o valor de  $d$  já foi identificado anteriormente uma vez que se pode constatar que a série é estacionaria em primeira diferença, logo a série CFE é  $I(1)$  e  $d = 1$ .

Analisando a figura 5, pode-se constatar facilmente que a função de autocorrelação apresenta picos significativos e a função de autocorrelação parcial decaimento exponencial o que muitos autores para este tipo de padrão de ACF e PACF sugerem que o modelo a estimar seja um MA, onde os  $lag^5$  1 até 8 são estatisticamente significante e todos os outros são iguais a zero. Neste contexto o valor do parâmetro  $q=8$ .

No entanto apenas a defasagem 1 é que se mostrou significativa<sup>6</sup> ao se estimar a equação logo  $q=1$ .

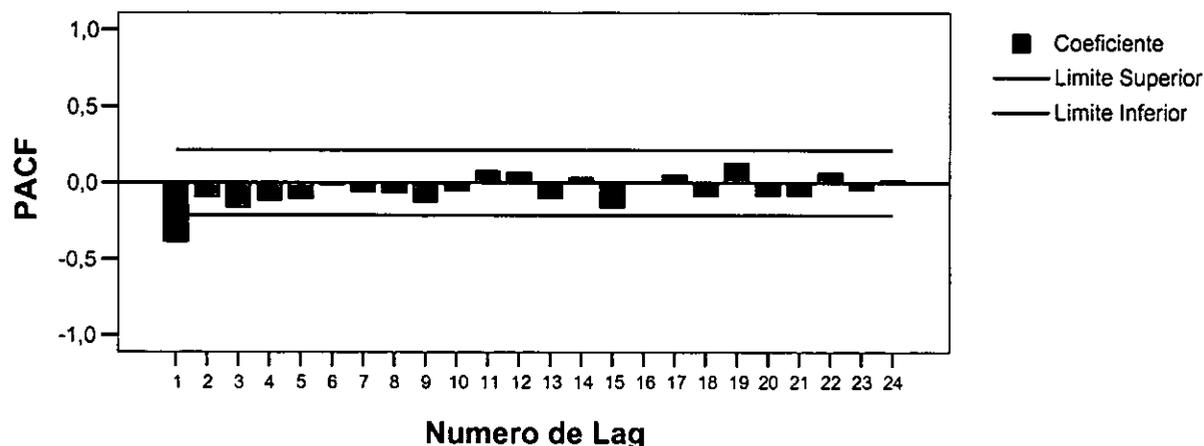
Figura 5. Correlograma da Função de Autocorrelação da série do Consumo da Energia Eléctrica Facturada em primeira diferença



<sup>5</sup> Lag é também chamado de número de defasagem ou período

<sup>6</sup> Ver anexo A, tabela 1 e 2

Figura 5.1. Correlograma da Função de Autocorrelação Parcial da série do Consumo da Energia Eléctrica Facturada em primeira diferença



Contudo, como a estratégia de identificação do modelo a ser estimado é feita de uma forma empírica com base na análise dos correlogramas amostrais e não se baseia em nenhum teste estatístico, então, o modelo ARIMA(0,1,1) identificado pode não ter sido o único modelo da Classe ARIMA que melhor se ajusta aos dados. Assim, as figuras 5 e 5.1 podem sugerir também os modelos ARIMA(1,1,0) e ARIMA(1,1,1).

#### 4.4. Estimação

A estimação consistiu em produzir os três modelos, isto é, depois de se estimar os respectivos parâmetros procedeu-se a verificação da sua significância estatística.

##### 4.4.1. Estimação do modelo Médias Móveis, ARIMA (0,1,1)

A tabela 7 contém os valores das estatísticas da estimação dos parâmetros do modelo ARIMA (0,1,1). Para o modelo estimado, a constante (-100.980,80) não estatisticamente significativa pois a probabilidade para a estatística  $t$  é de 0,18. Deste modo, não se rejeitou a hipótese nula de que o verdadeiro parâmetro populacional seja igual a zero. Contudo, o coeficiente (-0,56) para o processo media móvel é estatisticamente significativo uma vez que para a estatística  $t$  a probabilidade é de 0,00.

Assim rejeita-se a hipótese nula de que o verdadeiro parâmetro médias móveis seja igual a zero. No entanto, uma vez não rejeitada a hipótese da constante ser igual a zero, o modelo a ser estimado deve ser um modelo medias móveis que não apresente deslocamento<sup>7</sup>.

<sup>7</sup> Segundo Gujaratt (2006), o parâmetro beta zero é chamado de deslocamento.

Tabela 7. Estatísticas da Estimação do ARIMA(0,1,1) incluindo o parâmetro de deslocamento

Variável dependente D (CFE)				
Variável	Coefficiente	Erro Padrão	t-Statistic	Prob.
C	-100.980,80	74.293,80	-1,36	0,18
MA(1)	-0,56	0,09	-6,40	0,00
R-quadrado	0,18	Variância media dependente		-120.860,90
R-quadrado ajustado	0,17	S.D. variância dependente		1.706.356,00
S.E. da regressão	1.550.818,00	Crítério de informação Akaike		31,37
Soma dos quadrados dos residos	2.07E+14	Crítério de Schwarz		31,43
Log likelihood	-1.378,23	F- Estatística		19,33
Durbin-Watson stat	1,85	Prob(F-statistic)		0,00

Uma vez verificada a não significância estatística da constante no modelo anterior que sugeriu a eliminação do deslocamento, estimou-se o mesmo modelo não levando em conta o parâmetro deslocamento, os resultados da estatística do modelo encontram-se na tabela 8.

No entanto, o coeficiente do processo Médias Moveis permaneceu estatisticamente significativo. Os valores do Critério de Informação de Akaike e de Schwarz não diferem em unidades, mais sim em valores decimais o que mostrou que a exclusão do deslocamento não criou nenhum efeito no modelo.

Tabela 8. Estatísticas da Estimação do modelo ARIMA(0,1,1) sem o parâmetro de deslocamento

Variável dependente D (CFE)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
MA(1)	-0,52	0,09	-5,79	0,00
R-quadrado	0,18	Variância media dependente		-120.860,90
R-quadrado ajustado	0,18	S.D. variância dependente		1.706.356,00
S.E. da regressão	1.557.260,00	Crítério de informação Akaike		31,37
Soma dos quadrados dos residos	2.11E+14	Crítério de Schwarz		31,40
Logaritimo máximo-verossimil	-1.379,11	F- Estatística		1,88

#### 4.4.2. Estimação do modelo Autoregressivo, ARIMA (1,1,0)

Tal como para a estimação do modelo ARIMA (0,1,1), a tabela 9, contém as estatísticas da estimação do modelo ARIMA (1,1,0). Para o modelo estimado o parâmetro (-120.846,80) não é estatisticamente significativo, pois a estatística  $t$  para este parâmetro apresenta uma probabilidade de 0,33, sendo assim, não se rejeita a hipótese nula de que o verdadeiro parâmetro populacional seja igual a zero. Contudo, o coeficiente do processo auto regressivo é estatisticamente significativo, isto é, rejeita-se a hipótese nula de que o verdadeiro parâmetro populacional seja igual a zero.

Uma vez verificada a não significância estatística do deslocamento, partiu-se para a estimação do mesmo modelo, desta vez, não considerando o parâmetro deslocamento.

Tabela 9. Estatística da Estimação do Modelo ARIMA (1,1,0) incluindo o parâmetro de deslocamento

Variável dependente: D(CFE)

Variável	Coefficiente	Erro padrão	Estatística-t	P
C	-120846.80	124196.10	-0.97	0.33
AR(1)	-0.38	0.10	-3.78	0.00
R-quadrado	0.14	Variancia media dependente		-122955.50
R-quadrado ajustado	0.133	S.D. variancia dependente		1716135.00
S.E. da regressão	1597502.00	Critério de informação Akaike		31.43
S. dos quadrados erros	2.17E+14	Critério de Schwarz		31.49
Log máximo-verossimil	-1365.14	F- Estatística		14.25
Estatística Durbin-Watson	1.89	Prob(F-statistic)		0.00

Do mesmo modo que se procedeu a estimação do ARIMA(0,1,1) sem o referido deslocamento, para o modelo ARIMA(1,1,0) uma vez identificada a não significância estatística do parâmetro de deslocamento, procedeu-se também a estimativa do modelo ARIMA(1,1,0) sem o parâmetro de deslocamento. As estatísticas da estimação encontram-se na tabela 10 e os resultados mostram que o coeficiente do processo auto regressivo permaneceu significativo.

Tabela 10. Estatísticas da Estimação do modelo ARIMA(1,1,0) sem o parâmetro de deslocamento

Variavel dependente: D(CFE)

Variavel	Coefficiente	Erro padrão	Estatística-t	Prob.
AR(1)	-0.37	0.10	-3.72	0.00
R-quadrado	0.13	Variancia media dependente		-122955.50
R-quadrado ajustado	0.13	S.D. variancia dependente		1716135.00
S.E. da regressão	1596967.00	Critério de informação Akaike		31.41
S. dos quadrados erros	2.19E+14	Critério de Schwarz		31.45
Log máximo-verossimil	-1365.62	Estatística Durbin-Watson		1.88

#### 4.4.3. Estimação do modelo Autoregressivo Integrado de Médias Móveis, ARIMA (1,1,1)

Ao se estimar o modelo auto regressivo integrado de médias móveis, as estatísticas que se encontram na tabela 9, mostram que tanto o parâmetro de deslocamento como o coeficiente do processo autoregressivo não se mostram estatisticamente significativos, pois as estatísticas t dos respectivos parâmetros quando comparados com as estatísticas de MacKinnon e as respectivas probabilidades levaram a rejeição das hipóteses nulas de que os verdadeiros parâmetros populacionais seja iguais a zero. Contudo, apenas o coeficiente do processo *médias móveis* é que se mostrou ser estatisticamente significativo.

Desta forma, seria necessário a estimação apenas do modelo *medias moveis* sem a inclusão do parâmetro de deslocamento, para tal, torna-se desnecessário uma vez que já foi estimado no ponto 4.3.2.1 que é o modelo ARIMA(0,1,1).

Tabela 11. Estatísticas da estimação do modelo ARIMA(1,1,1) incluindo o parâmetro de deslocamento

Variável dependente : D(CFE)				
Variável	Coefficiente	Erro padrão	Estatística t	Prob.
C	-80584.01	51526.64	-1.56	0.12
AR(1)	0.26	0.16	1.65	0.10
MA(1)	-0.78	0.11	-7.36	0.00
R-quadrado	0.25	Variância média dependente		-122955.50
R-quadrado ajustado	0.23	S.D. variância dependente		1716135.00
S.E. da regressão	1507676.00	Critério de informação Akaike		31.32
S. dos quadrados erros	1.91E+14	Critério de Schwarz		31.41
Log máximo-verossimil	-1359.59	F- Estatística		13.71
Estatística Durbin-Watson	1.87	Prob(F-statistic)		0.00

### 4.3.3. Diagnostico

Uma vez que o modelo ARIMA(1,1,1) não se mostrou estatisticamente significativo para estimar os valores do consumo de energia eléctrica, apenas os modelos ARIMA(0,1,1) e ARIMA(1,1,0) foram diagnosticados. Analisando as figuras 6 e 7, pode-se verificar que os resíduos estão distribuídos aleatoriamente ao redor da média zero. O histograma desses valores é mostrado nas figuras 8 e 9.

Figura 6. Série dos resíduos após o ajuste do modelo ARIMA(1,1,0)

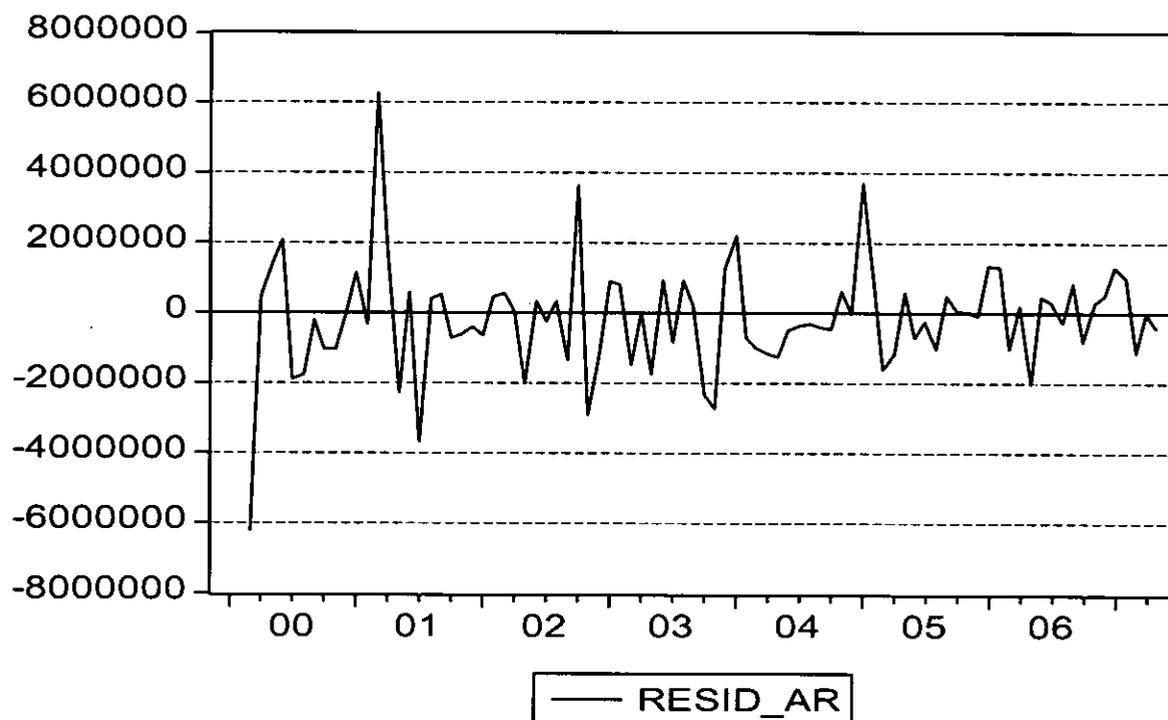
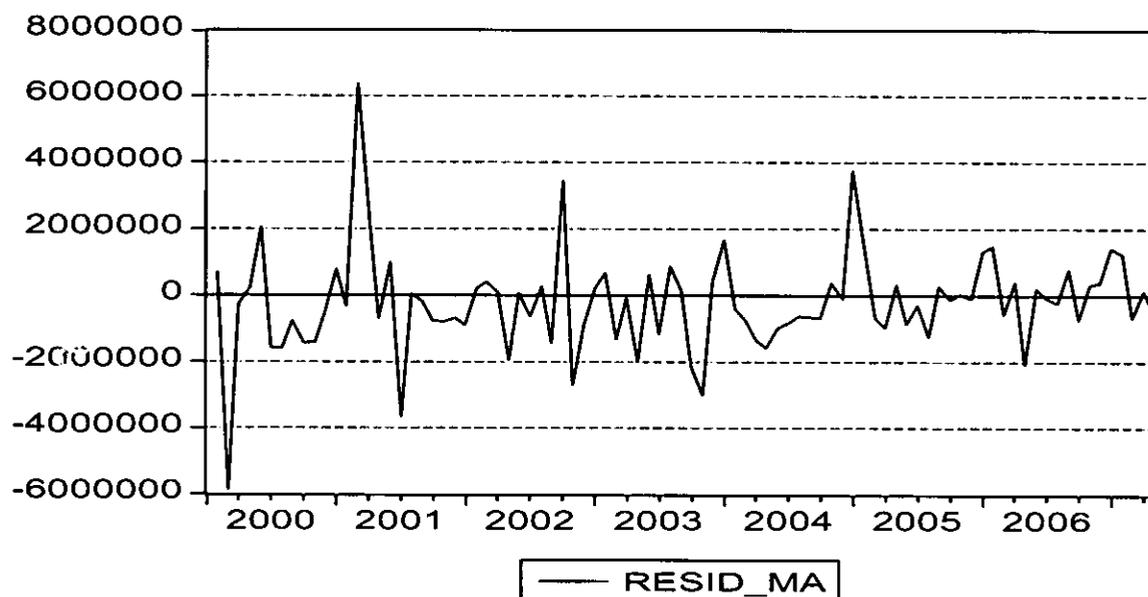


Figura 7. Série dos resíduos após o ajuste do modelo ARIMA(0,1,1)



Pode-se também verificar pelos histogramas que os resíduos dos modelos estimados se distribuem normalmente em torno de zero.

Figura 8. Histograma da série dos resíduos do modelo ARIMA(1,1,0)

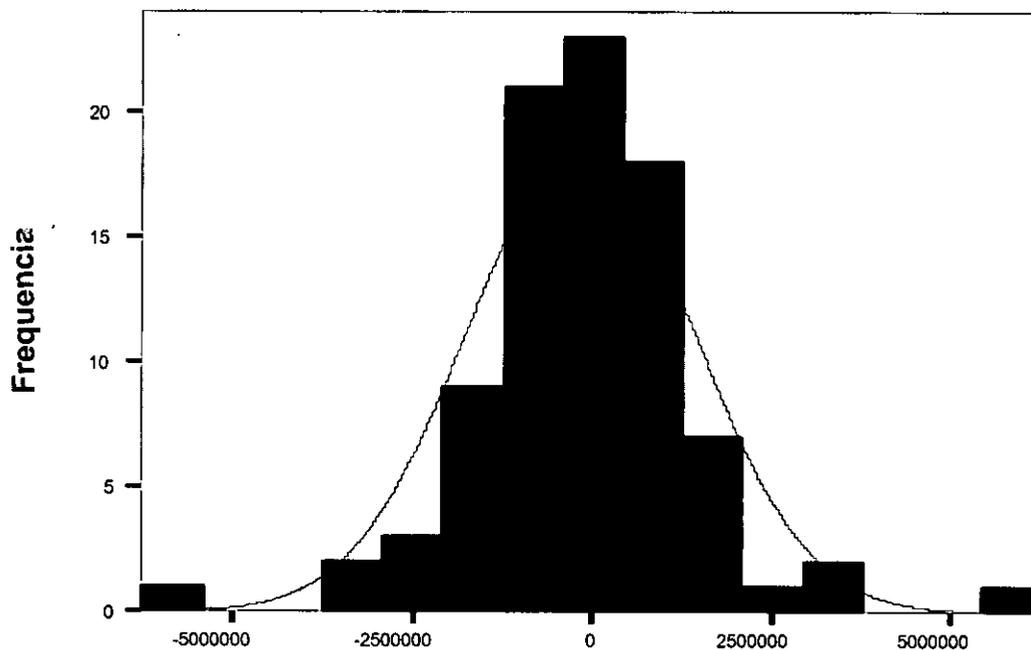
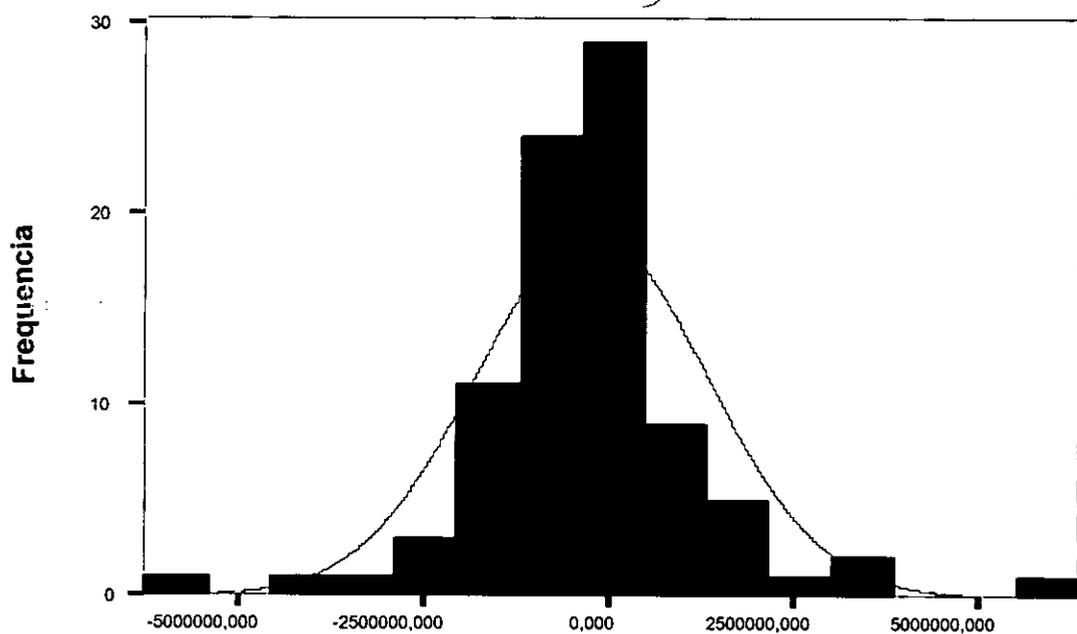


Figura 9. Histograma da série dos resíduos do modelo ARIMA(0,1,1)



Analisando os correlogramas das séries dos resíduos dos modelos, constata-se que após o ajuste não existe mais autocorrelação e autocorrelação parcial dos resíduos. Assim, pode-se afirmar que as séries dos resíduos são estacionárias em nível e todas as autocorrelações e autocorrelações parciais são estatisticamente iguais a zero.

Figura 10. Correlograma da Função de Autocorrelação da série dos resíduos do modelo ARIMA(1,1,0)

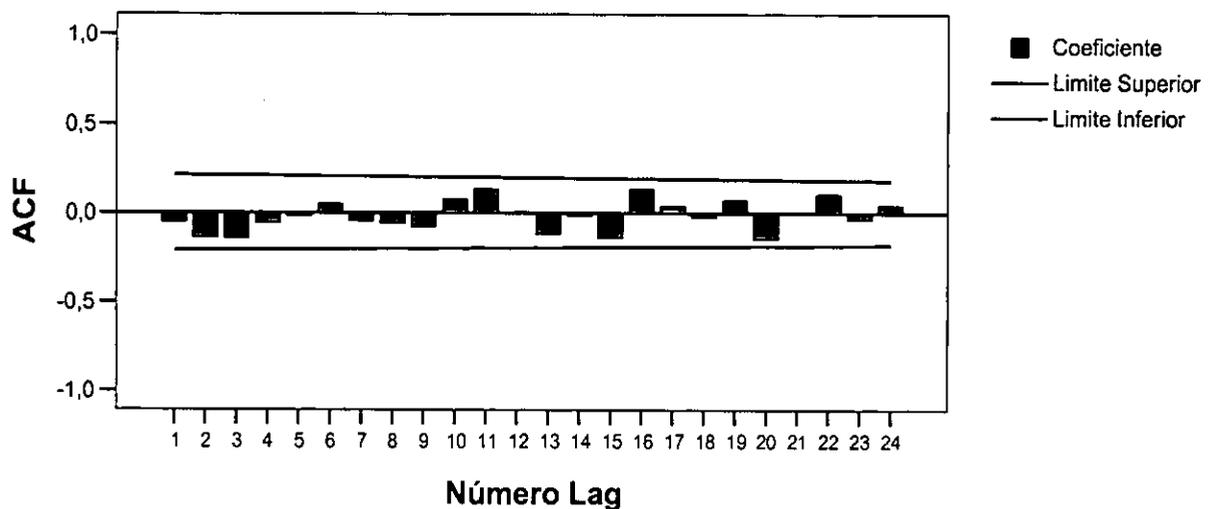


Figura 10.1. Correlograma da Função de Autocorrelação Parcial da série dos resíduos do modelo ARIMA(1,1,0)

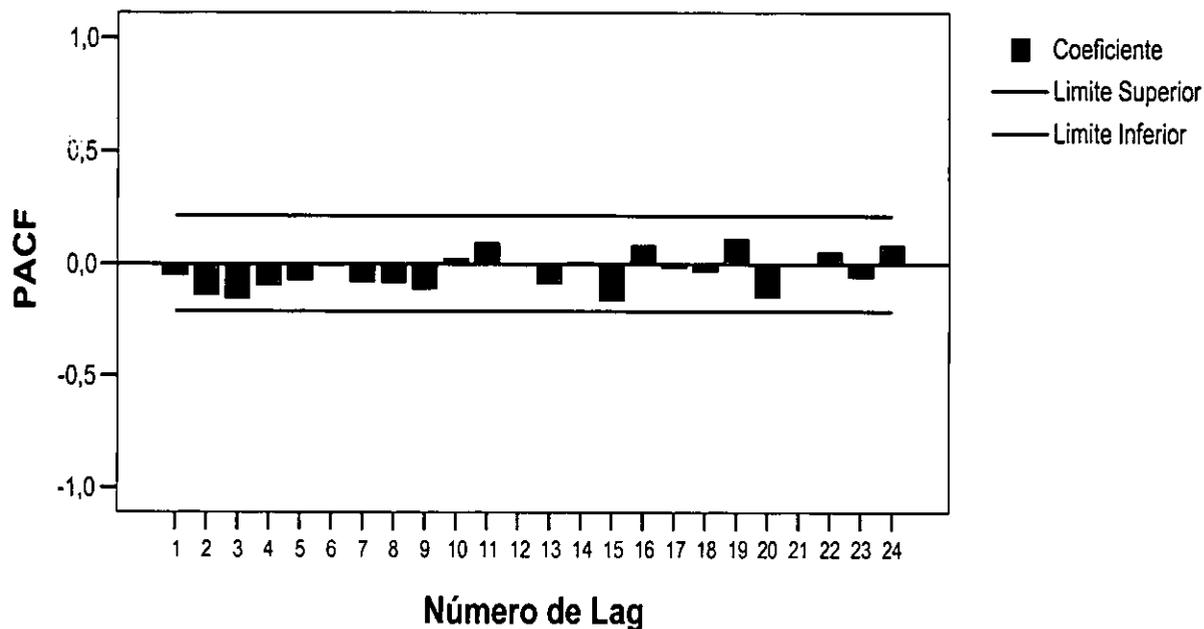


Figura 11. Correlograma da Função de Autocorrelação da série dos resíduos do modelo ARIMA(0,1,1)

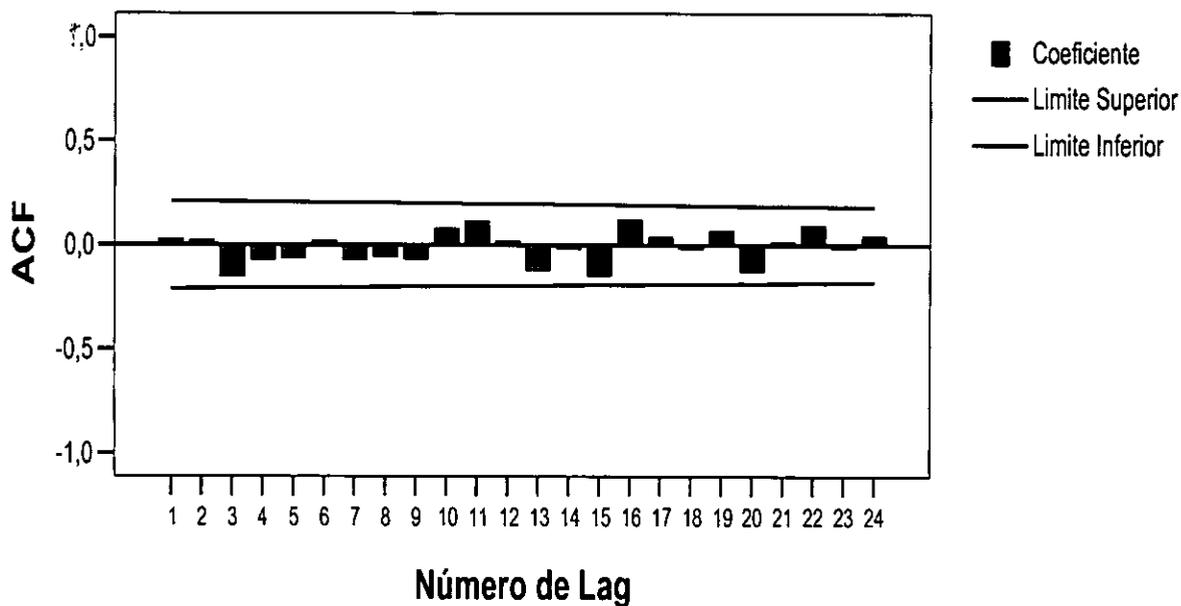
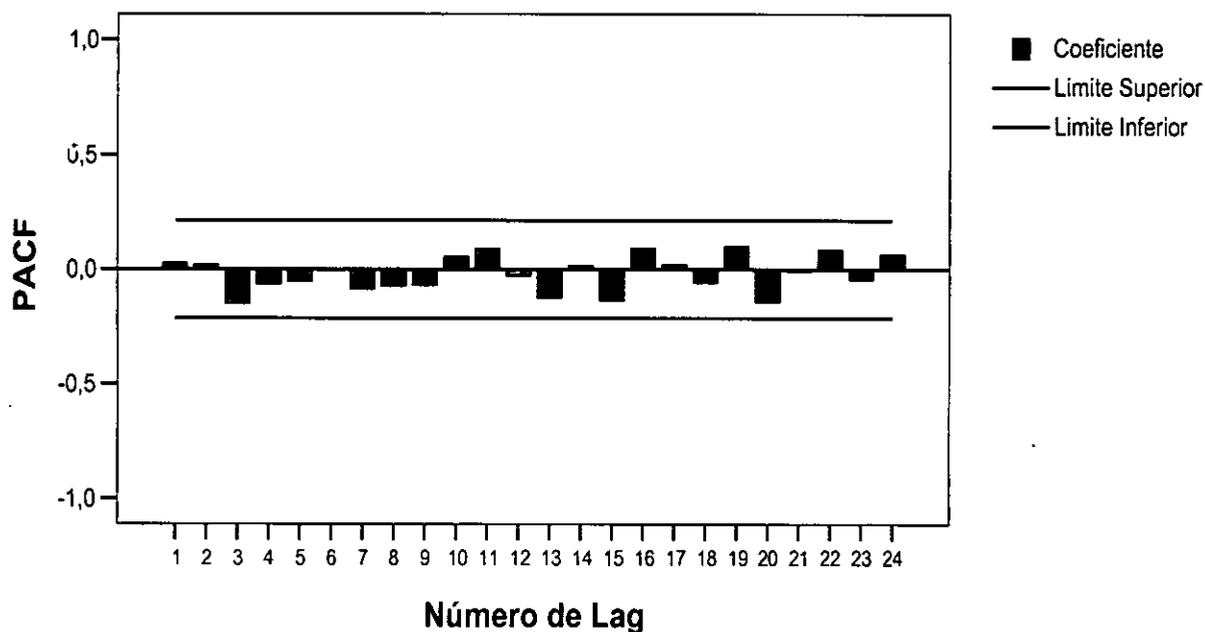


Figura 11.1. Correlograma da Função de Autocorrelação Parcial da série dos resíduos do modelo ARIMA(0,1,1)



A tabela 12. mostra os resultados das estatísticas dos teste e a sua significância estatística para sustentar as análises anteriores. Assim, e de acordo com a mesma tabela, não se rejeita a 5% a hipótese nula de que os resíduos dos modelos estimados seguem distribuição normal pois o valor do sig tanto para os resíduos do modelo ARIMA(1,1,0) como para os do modelo ARIMA(0,1,1) seguem distribuição normal.

O teste *t Student* para os mesmos níveis de significância não rejeita a hipótese nula de que as médias dos resíduos estimados são estatisticamente iguais a zero. O modelo ARCH estimado não é estatisticamente significativo, assim as séries dos resíduos dos modelos não apresentam o efeito ARCH, logo a série apresenta variância constante pois os resíduos ao quadrado não dependem dos seus valores passados.

Deste modo, os dois modelos estimados ajustam-se aos dados da série do CFE e os resíduos são puramente aleatórios, isto é, são um ruído branco.

Neste contexto, os dois modelos poderiam ser usados para estimar ou prever os valores futuros do consumo de Energia Eléctrica Facturada para a Cidade de Maputo.

Tabela 12. Testes estatísticos para os pressupostos de normalidade, media dos resíduos igual a zero e variância constante.

	<i>Kolmogorov-Smirnov (K-S)</i>	<i>t Student</i>	<i>ARCH Teste LM</i>	<i>QLB</i>
<i>Resíduos do modelo ARIMA(1,1,0)</i>	0,93 (0,35)	-0,97 (0,34)	0,02 (0,90)	7,3 (0,84)
<i>Resíduos do modelo ARIMA(0,1,1)</i>	1,28 (0,08)	-1,36 (0,18)	0,00 (0,95)	6,04 (0,91)

#### 4.4. Comparação dos modelos

Feita a verificação e diagnóstico dos modelos estimados, a tabela 13 mostra as estatísticas do desempenho da previsão dentro da amostra usando os modelos estimados.

As estatísticas apresentam valores mais baixos para o modelo ARIMA(0,1,1) sugerindo assim um melhor desempenho do modelo acima referido.

Sendo assim o modelo usado para prever o CFE é o ARIMA(0,1,1).

Tabela 13. Estatísticas do desempenho nas previsões dentro da amostra dos modelos seleccionados.

	<i>REQM</i>	<i>EAM</i>	<i>EAMP</i>	<i>CDT</i>	<i>AIC</i>	<i>SIC</i>
<i>ARIMA(1,1,0)</i>	1.587.763,00	1.101.240,00	7,37	0,05	31,42	31,45
<i>ARIMA(0,1,1)</i>	1.548.387,00	1.056.502,00	7,12	0,05	31,37	31,39

#### 4.5. Previsão

A partir do modelo ARIMA(0,1,1) procedeu-se a previsão do consumo de energia eléctrica facturado na Cidade de Maputo e seus limites de confiança a 95%, para o período 06/2007 à 5/2008, como mostra a Tabela 14.

Neste período, as previsões apontam para um consumo de 11.917.716,54 kWh para o mes de Junho de 2007 sendo que para este período de previsão o consumo mais elevado é 10.749.081,91 kWh para o mes de Maio de 2008 sendo o consumo mais baixo no mesmo período de previsão.

Pela mesma tabela, pode-se constatar que o consumo de Energia Eléctrica Facturado, apresenta uma tendência decrescente como se pode observar pela mesma tabela .

Tabela 14. Previsão do Consumo da Energia Eléctrica Facturada para o período de Junho de 2007 a Maio de 2008

<i>Mês</i>	<i>Limite inferior</i>	<i>Previsão</i>	<i>Limite superior</i>
<i>Junho/2007</i>	8.805.834,13 kWh	11.917.716,54 kWh	15.029.598,95 kWh
<i>Julho/2007</i>	8.341.887,50 kWh	11.811.477,03 kWh	15.281.066,56 kWh
<i>Agosto/2007</i>	7.904.706,98 kWh	11.705.237,52 kWh	15.505.768,06 kWh
<i>Setembro/2007</i>	7.487.823,63 kWh	11.598.998,01 kWh	15.710.172,39 kWh
<i>Outubro/2007</i>	7.086.942,03 kWh	11.492.758,50 kWh	15.898.574,96 kWh
<i>Novembro/2007</i>	6.699.043,74 kWh	11.386.518,99 kWh	16.073.994,22 kWh
<i>Dezembro/2007</i>	6.321.915,72 kWh	11.280.279,47 kWh	16.238.643,22 kWh
<i>Janeiro/2008</i>	5.953.881,00 kWh	11.174.039,96 kWh	16.394.198,92 kWh
<i>Fevereiro/2008</i>	5.593.634,80 kWh	11.0678.00,45 kWh	16.541.966,10 kWh
<i>Março/2008</i>	5.240.139,69 kWh	10.961.560,94 kWh	16.682.982,18 kWh
<i>Abril/2008</i>	4.892.555,78 kWh	10.855.321,43 kWh	16.818.087,06 kWh
<i>Maio/2008</i>	4.550.192,61 kWh	10.749.081,91 kWh	16.947.971,22 kWh

## CAPITULO V

---

### 5. CONCLUSÕES

Neste trabalho analisou-se o Consumo da Energia Eléctrica Facturada na Cidade de Maputo, baseando-se na metodologia de Box-Jenkins. Destacam-se as seguintes conclusões:

- A série do Consumo da Energia Eléctrica Facturada não é sazonal, isto é, o Consumo da Energia Eléctrica Facturado no período em análise não apresentou movimentos ondulatórios de curta duração observadas em períodos idênticos de cada ano. Isto pode ter ocorrido devido ao facto da série em análise não ter englobado todos os consumidores de energia eléctrica, mais sim apenas uma parte deles.
- No período em análise destacaram-se três diferentes níveis de consumo conforme referido no capítulo 4. O primeiro nível foi o que apresentou os consumos mais elevados. O segundo nível o consumo apresentou uma tendência decrescente. O terceiro e ultimo nível o consumo manteve-se relativamente constante por volta dos 120.000.000,00 kWh.
- De acordo com os resultados obtidos, o modelo que melhor se ajustou e produziu as melhores previsões dentro da amostra é o ARIMA(0,1,1) ou médias moveis. Assim, o Consumo da Energia Eléctrica Facturada pode ser estimado pelos erros de estimação do período actual e do período anterior.
- As previsões para o período de Junho de 2007 a Maio de 2008, mostram que o Consumo da Energia Eléctrica Facturada tenderá a decrescer nos próximos meses. Uma das causas que poderá justificar esse baixo consumo, é o uso de meios alternativos como o Petróleo, Gás, Lenha e Carvão que são a segunda fonte de iluminação e energia na Cidade de Maputo.

### **5.1. Recomendações**

Para estudos futuros relacionados com o Consumo da Energia Eléctrica , recomenda-se que a série a ser usada possa englobar todos os consumidores tanto os do consumo facturado como os do consumo pré pago para que possa ter uma visão mais abrangente e precisa da situação real.

### **5.2. Limitações do estudo**

Apesar de existir, hoje, na literatura uma vasta colecção de artigos, livros e pesquisas sobre a metodologia de Box-Jenkins, as principais limitações observadas durante a realização do presente trabalho foram as seguintes:

A impossibilidade de se obter o Consumo Geral da Energia de todos os clientes da Cidade de Maputo, pois apenas os dados dos clientes facturados é que foram usados excluindo-se assim os clientes do pré pago (Credeleec), o que possivelmente poderá ter contribuído para a não identificação de factores sazonais no consumo de energia.

O acesso a fontes bibliográficas primárias, o que levou a que se recorresse constantemente à fontes bibliográficas secundárias como por exemplo a Internet.

## 6. BIBLIOGRAFIA

- [1] Ansley, C.F. & Newbold, P. (1980). Finite sample properties of estimators for autoregressive moving average models. *Journal of Econometrics*. 13, 159-183.
- [2] Abdel-Aal, R. E.; Al-Garni, Z. (1997). Forecasting monthly electric energy consumption in eastern Saudi Arabia using univariate time-series analysis. *Energy*, v. 22, n. 11, p. 1059-1069.
- [3] Asche, F., H. Bremnes e C.R. Wessells (1999), Product aggregation, market integration, and relationships between prices: an application to world salmon markets, *American Journal of Agricultural Economics* 81 (Agosto), 568-581.
- [4] Braumann, C.A. (2005), Introdução às equações diferenciais estocásticas e aplicações. *XIII Congresso Anual da Sociedade Portuguesa de Estatística*, Edições SPE.
- [5] Brockwell, P.J. & Davis, R.A. (1996). *Introduction to Time Series and Forecasting*. NewYork: Springer-Verlag.
- [6] Bowerman, B.L. & O'Connell, R.T. (1987). *Time Series Forecasting, Unified Concepts and Computer Implementation*, 2ª Edição. Boston: Duxbury Press.
- [7] Box, G.E.P., & Jenkins, G.M. (1970). *Time Series Analysis, Forecasting and Control*. San Francisco: Holden Day.
- [8] Cribari-Neto, F. (2000). *Método de Previsão de Arrecadação Tributária*. MIMEO
- [9] Dickey, D.A. e W.A. Fuller (1979), Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association* 74 (Junho), 427-431pp.
- [10] Dickey, D.A. e W.A. Fuller (1981), Likelihood ratio statistics for autoregressive time series with a unit root, *Econometrica* 49, 1057-1072.

- [11] Dufrénot, G. e V. Mignon (2002), *Recent Developments in Nonlinear Cointegration with Applications to Macroeconomics and Finance*. Dordrecht, NL: Kluwer Academic Publishers.
- [12] Dionísio, A., R. Menezes e D.A. Mendes (2004), Mutual information: a measure of dependency for nonlinear time series, *Physica A* 344, 326-329.
- [13] Engle, R. F., YOO, B. Forecasting and testing in cointegrated systems. *Journal of econometrics*, v. 35, p. 143-59, 1987.
- [14] Fava, W.A. (2000), *Introduction to Statistical Time Series*. New York: John Wiley.
- [15] Gujarati, D.N (2006). *Econometria básica*. 4ª edição (tradução). São Paulo. Editora Campus.
- [16] Granger, C.W.J. & Newbold, P. (1986). *Forecasting Economic Time Series*. New York: Academic Press.
- [17] Hurvich, C.M. e C.L. Tsai (1989), Regression and time series model selection in small samples, *Biometrika* 76, 297-307.
- [18] Hannan-Quinn. (1979). *Introdução à economia*. 20ª edição São Paulo. Editora Atlas.
- [19] Instituto Nacional de Estatística de Moçambique (2004); Inquérito aos Agregados Familiares 2002-2003,( [www.ine.gov.mz](http://www.ine.gov.mz) )
- [20] Lim, C., McAleer, M. Time series forecasts of international travel demand for australia. *Tourism Management*. (Artigo aceito em julho de 2001, aguarda impressão.)
- [21] Morettin, P. A. e Toloí, C. M. C. (1987) *Previsão de séries temporais* (2a ed.). Ed. Atual, São Paulo.
- [22] Mills, T.C. (1990). *Time Series Techniques for Economists*. Cambridge: Cambridge University Press.

- [23] Masih, A.M.M. e R. Masih (2002), Propagative causal price transmission among international stock markets: evidence from the pre- and post globalization period, *Global Finance Journal* 13, 63-91.
- [24] MacKinnon, J.G. (1991), Critical values for co-integration tests, em Engle R.F. e C.W.J. Granger (eds.), *Long-Run Economic Relationships*, Oxford: Oxford University Press, 267-276.
- [25] MacKinnon, J.G. (1996), Numerical distribution functions for unit root and cointegration tests, *Journal of Applied Econometrics* 11, 601-618.
- [26] MacKinnon, J.G., A.A. Haug e L. Michelis (1999), Numerical distribution functions of likelihood ratio tests for cointegration, *Journal of Applied Econometrics* 14, 563-577.
- [27] Menezes, R., A. Dionísio e D.A. Mendes (2004), Asymmetric price transmission within the Portuguese stock market, *Physica A* 344, 312-316
- [28] Menezes, R., N.B. Ferreira, e D.A. Mendes (2006), Asymmetric signals in econophysics modeling: dynamic volatility and threshold adjustment models, em Dionísio, A., A.H. Reis e R.N. Rosa (eds.), *Proceedings of the Perspectives on Econophysics Workshop*, Évora: Universidade de Évora, 87-98.
- [29] Marques, C.R. (1998), *Modelos Dinâmicos, Raízes Unitárias e Cointegração*. Lisboa: EDINOVA.
- [30] McClain, K. T. Humphreys, H. B. (1996); Atahunpa. Measuring risk in the mining sector with arch models with important observations on sample size. *Journal of empirical finance*, 369-391.
- [31] Mendes, W.A. (2006). *Introduction to Statistical Time Series*. New York. John Wiley

- [32] Nelson, D.B. (1973). *Conditional Heteroskedasticity in Asset Returns: A New Approach*. *Econometrica*, 59. pp 347-370.
- [33] Pessanha, J.F.M, Leal, L.H.C e Silva, M.A.M (2004); *Uma Aplicação de Modelos de Séries Temporais na Previsão da Demanda por Transporte Aéreo de Passageiros*. Rio de Janeiro. pp 5-11.
- [34] Poskitt, D.S. e A.R. Tremayne (1987), Determining a portfolio of linear time series models, *Biometrika* 74, 125-137.
- [35] Rossetti, J.P. (2006). *Introdução à economia*. 20ª edição São Paulo. Editora Atlas
- [36] Said, S.E. e D.A. Dickey (1984), Testing for unit roots in autoregressive-moving average models of unknown order, *Biometrika* 71, 599-607.
- [37] Error! No index entries found.] Slini, L.; Karatzas, K.; Moussiopoulos, N. Statistical analysis of environmental data as the basis of forecasting: an air quality application. *The Science of the Total Environment*, v. 288, p. 227-237, 2001.
- [38] Siklos, P.L. (2002), Asymmetric adjustment from structural booms and slumps, *Economic Letters* 77, 329-333.
- [39] Sentana, E. (1995). *Quadratic ARCH Models*. *Reviews of Economic Studies*, 60. pp 639.



## ANEXO A

**Tabela 1. Estimação do ARIMA(0,1,8)**

Dependent Variable: D(CFE)  
 Method: Least Squares  
 Date: 01/24/08 Time: 07:46  
 Sample(adjusted): 2000:02 2007:05  
 Included observations: 88 after adjusting endpoints  
 Convergence achieved after 30 iterations  
 Backcast: 1999:06 2000:01

Variable	Coefficient	Std. Error	t-Statistic	Prob.
MA(1)	-0.460026	0.112607	-4.085233	0.0001
MA(2)	-0.012187	0.121447	-0.100349	0.9203
MA(3)	-0.186713	0.119866	-1.557682	0.1233
MA(4)	0.009901	0.125255	0.079051	0.9372
MA(5)	-0.030313	0.124725	-0.243035	0.8086
MA(6)	0.111386	0.122173	0.911710	0.3647
MA(7)	-0.066378	0.123370	-0.538043	0.5920
MA(8)	0.000127	0.112537	0.001129	0.9991

**Tabela 2. Estimação do ARIMA(8,1,0)**

Dependent Variable: D(CFE)  
 Method: Least Squares  
 Date: 04/08/08 Time: 07:10  
 Sample(adjusted): 2000:10 2007:05  
 Included observations: 80 after adjusting endpoints  
 Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-0.447996	0.117677	-3.807004	0.0003
AR(2)	-0.141508	0.127538	-1.109534	0.2709
AR(3)	-0.140745	0.125888	-1.118015	0.2673
AR(4)	-0.188615	0.126665	-1.489083	0.1408
AR(5)	-0.169144	0.128064	-1.320775	0.1908
AR(6)	-0.034396	0.127093	-0.270633	0.7874
AR(7)	-0.064267	0.115521	-0.556319	0.5797
AR(8)	-0.052672	0.104418	-0.504434	0.6155

**Tabela 3. Equação dos resíduos ao quadrado do ARIMA(0,1,1)**

Test Equation:  
 Dependent Variable: RESID^2  
 Method: Least Squares  
 Date: 01/21/08 Time: 22:26  
 Sample(adjusted): 2000:03 2007:05  
 Included observations: 87 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.39E+12	7.11E+11	3.359109	0.0012
RESID^2(-1)	0.013359	0.108467	0.123158	0.9023
R-squared	0.000178	Mean dependent var		2.42E+12
Adjusted R-squared	-0.011584	S.D. dependent var		6.12E+12
S.E. of regression	6.16E+12	Akaike info criterion		61.75852
Sum squared resid	3.22E+27	Schwarz criterion		61.81520
Log likelihood	-2684.495	F-statistic		0.015168
Durbin-Watson stat	1.631684	Prob(F-statistic)		0.902273

**Tabela 4. Teste ARCH dos resíduos do ARIMA(0,1,1)**

ARCH Test:

F-statistic	0.015168	probabilidade	0.902273
Obs*R-squared	0.015522	probabilidade	0.900850

**Tabela 5. Equação dos resíduos do ARIMA(1,1,0)**

Test Equation:  
 Dependent Variable: RESID^2  
 Method: Least Squares  
 Date: 01/21/08 Time: 22:30  
 Sample(adjusted): 2000:04 2007:05  
 Included observations: 86 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.11E+12	5.77E+11	3.658076	0.0004
RESID^2(-1)	-0.005343	0.085064	-0.062806	0.9501
R-squared	0.000047	Mean dependent var	2.10E+12	
Adjusted R-squared	-0.011857	S.D. dependent var	4.93E+12	
S.E. of regression	4.96E+12	Akaike info criterion	61.32617	
Sum squared resid	2.07E+27	Schwarz criterion	61.38325	
Log likelihood	-2635.025	F-statistic	0.003945	
Durbin-Watson stat	1.935360	Prob(F-statistic)	0.950070	

Tabela 6. Teste ARCH dos residuos do ARIMA(1,1,0)

ARCH Test:

F-statistic	0.003945	probability	0.950070
Obs*R-squared	0.004038	Probability	0.949330

Tabela 7. Serie original do Consumo facturado de energia electrica

Meses	Consumo de Energia Electrica da Cidade de Maputo (Consumo Facturado kwh)											
	Ano											
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Jan	22.486.105	17.544.035	17.511.182	16.669.542	14.783.124	13.938.993	12.237.606	12.842.171				
Feb	22.547.471	16.796.370	18.185.904	17.212.171	13.496.089	13.530.461	13.041.378	13.330.572				
Mar	16.286.971	23.340.916	18.477.789	15.527.978	12.968.201	12.047.445	11.684.180	11.995.816				
Apr	19.064.546	22.582.818	18.369.297	16.146.849	11.980.559	11.419.618	12.367.306	12.456.443				
May	19.403.273	20.560.980	16.373.194	14.179.762	11.091.249	12.237.813	10.072.996	11.850.343				
Jun	21.359.089	21.888.691	17.435.511	15.830.113	10.916.636	11.216.683	11.365.180	-				
Jul	18.713.640	17.712.459	16.774.504	14.346.201	10.585.537	11.357.000	11.173.472	-				
Aug	17.931.501	19.635.705	17.352.748	15.812.951	10.394.717	10.249.628	10.967.161	-				
Sep	17.985.438	19.417.315	15.767.966	15.471.986	10.051.914	11.141.741	11.868.855	-				
Oct	16.930.541	18.756.273	19.973.518	13.246.706	9.694.054	10.881.702	10.708.712	-				
Nov	16.258.372	18.358.581	15.466.298	11.337.271	10.425.811	10.983.716	11.408.623	-				
Dec	16.491.813	18.086.739	15.977.733	13.315.226	10.135.930	10.864.986	11.609.548	-				

Fonte: Departamento de Estatística e Informatica, Electricidade de Moçambique, cidade de Maputo

**Tabela 8. Numero de Clientes**

Meses	Numero de clientes											
	Ano											
	2000	2001	2002	2003	2004	2005	2006	2007				
Jan	68.017	71.233	72.962	68.211	59.144	46.358	45.190	47.464				
Feb	68.099	71.808	73.051	67.575	58.729	46.318	45.218	48.404				
Mar	68.155	71.701	73.358	66.853	50.111	46.207	45.549	48.794				
Apr	68.373	71.608	73.689	66.247	48.185	46.116	45.772	49.210				
May	68.581	71.862	73.710	65.754	47.859	45.856	45.937	49.384				
Jun	68.464	71.911	73.908	65.731	48.074	45.684	46.110	-				
Jul	69.396	72.027	73.834	65.533	45.511	45.342	46.116	-				
Aug	70.216	72.314	73.721	65.396	46.468	44.941	46.431	-				
Sep	70.770	72.366	72.785	63.196	46.275	44.807	46.467	-				
Oct	70.789	72.399	73.445	60.788	46.168	44.860	46.544	-				
Nov	70.924	72.404	69.301	61.247	46.125	45.006	46.546	-				
Dec	71.120	72.765	68.979	58.899	46.139	45.098	46.783	-				

Fonte: Departamento de Estatística e Informática, Electricidade de Moçambique, cidade de Maputo

**Tabela 9. Teste de distribuição normal para os resíduos do ARIMA(1,1,0) e ARIMA(0,1,1)**

		One-Sample Kolmogorov-Smirnov Test	
		Resíduos (1,1,0)	Resíduos (0,1,1)
N		88	88
Normal Parameters <sup>a,b</sup>	Mean	-162836.49	-225156.08
	Std. Deviation	1579324.0	1551994.8
Most Extreme Differences	Absolute	.099	.135
	Positive	.099	.135
	Negative	-.089	-.112
Kolmogorov-Smirnov Z		.929	1.270
Asymp. Sig. (2-tailed)		.354	.079

a. Test distribution is Normal.

b. Calculated from data.

**Tabela 10. Teste de hipótese para verificar se a médias dos resíduos Do ARIMA(1,1,0) e ARIMA(0,1,1) são ou não iguais a zero**

One-Sample Test						
Test Value = 0						
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Resíduos (1,1,0)	-.967	87	.336	-162836.49	-497463.28	171790.30
Resíduos (0,1,1)	-1.361	87	.177	-225156.08	-553992.36	103680.20

**Tabela 11. Teste de hipótese para verificação de igualdade no consumo médio mensal**

ANOVA					
Consumo de Energia Eléctrica Facturado					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	5.938E+013	11	5.399E+012	.379	.961
Within Groups	1.098E+015	77	1.426E+013		
Total	1.158E+015	88			

**Tabela 12. Estatísticas das autocorrelações da serie do cfe em nivel**

Series: Consumo de Energia Electrica Facturado

Lag	Autocorrelação	Erro Padrão <sup>a</sup>	estatística de Box-Ljung		
			Valor da Estat	df	Sig. <sup>b</sup>
1	,861	,104	68,237	1	,000
2	,805	,104	128,581	2	,000
3	,775	,103	185,112	3	,000
4	,757	,102	239,686	4	,000
5	,732	,102	291,334	5	,000
6	,698	,101	338,812	6	,000
7	,666	,101	382,570	7	,000
8	,647	,100	424,473	8	,000
9	,631	,099	464,792	9	,000
10	,635	,099	506,070	10	,000
11	,630	,098	547,333	11	,000
12	,608	,098	586,243	12	,000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

**Tabela 13. Estatísticas das autocorrelações da serie do cfe em primeira diferença**

Series: Consumo de Energia Electrica Facturado

Lag	Autocorrelação	Erro Padrão <sup>a</sup>	Estatística de Box-Ljung		
			Valor da Estat	df	Sig. <sup>b</sup>
1	-,379	,105	13,055	1	,000
2	,063	,104	13,424	2	,001
3	-,127	,104	14,925	3	,002
4	,012	,103	14,940	4	,005
5	-,035	,102	15,059	5	,010
6	,062	,102	15,432	6	,017
7	-,051	,101	15,690	7	,028
8	,000	,101	15,690	8	,047
9	-,073	,100	16,219	9	,062
10	,054	,099	16,519	10	,086
11	,075	,099	17,091	11	,105
12	,015	,098	17,114	12	,145

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

**Tabela 14. Estatísticas das autocorrelações da serie dos residuos em nivel do ARIMA(1,1,0).**

Series: Residuos (1,1,0)

Lag	Autocorrelação	Erro Padrão <sup>a</sup>	Estatística de Box-Ljung		
			Valor da Estat	df	Sig. <sup>b</sup>
1	-,046	,105	,195	1	,658
2	-,133	,104	1,830	2	,401
3	-,136	,104	3,561	3	,313
4	-,052	,103	3,815	4	,432
5	-,014	,102	3,834	5	,574
6	,045	,102	4,033	6	,672
7	-,041	,101	4,197	7	,757
8	-,054	,101	4,487	8	,811
9	-,072	,100	5,012	9	,833
10	,073	,099	5,550	10	,852
11	,129	,099	7,267	11	,777
12	,002	,098	7,268	12	,839

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

**Tabela 15. Estatísticas das autocorrelações da serie dos residuos em nivel do ARIMA(0,1,1).**

**Autocorrelations**

Series: Residuos (0,1,1)

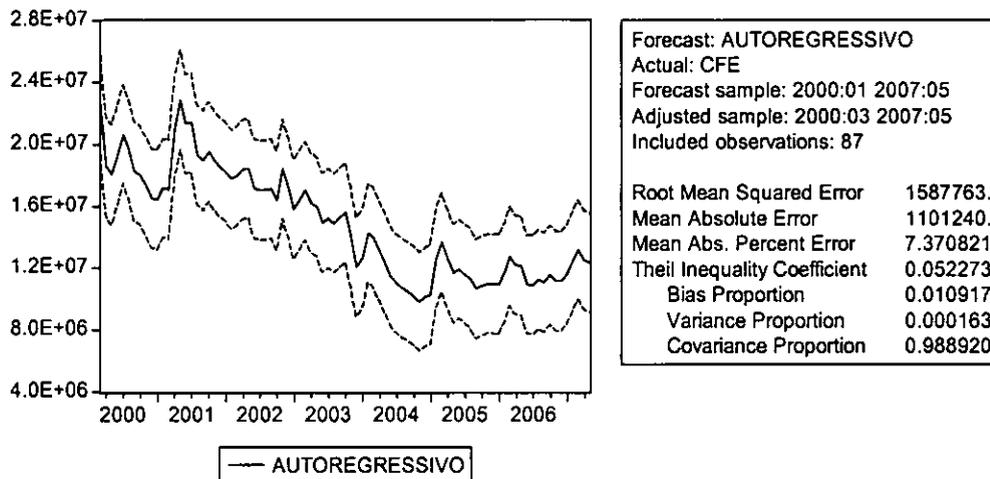
Lag	Autocorrelation	Std.Error <sup>a</sup>	Box-Ljung Statistic		
			Value	df	Sig. <sup>b</sup>
1	,021	,105	,040	1	,842
2	,016	,104	,063	2	,969
3	-,146	,104	2,045	3	,563
4	-,066	,103	2,451	4	,653
5	-,055	,102	2,741	5	,740
6	,014	,102	2,758	6	,839
7	-,064	,101	3,164	7	,869
8	-,050	,101	3,417	8	,906
9	-,062	,100	3,801	9	,924
10	,072	,099	4,325	10	,931
11	,110	,099	5,571	11	,900
12	,013	,098	5,589	12	,935

a. The underlying process assumed is independence (white noise).

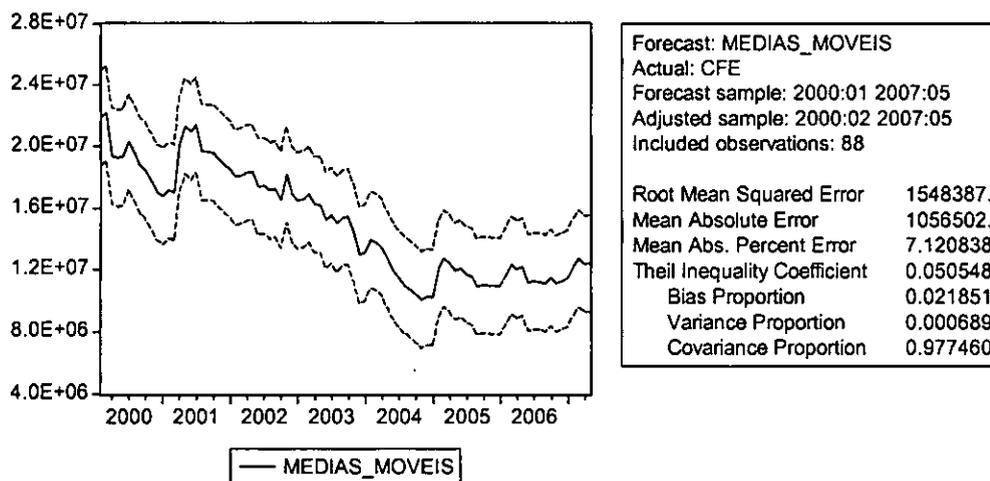
b. Based on the asymptotic chi-square approximation.

ANEXO B

**Figura 1. Previsão dentro da amostra, limites de confidencialidade  
E estatísticas de desempenho do ARIMA(1,1,0)**



**Figura 2. Previsão dentro da amostra, limites de confidencialidade  
E estatísticas de desempenho do ARIMA(1,1,0)**



**Figura 3. Correlograma dos resíduos do ARIMA(1,1,0) e as respectivas estatística QLB**

Included observations: 87

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
. .	. .	1	-0.036	-0.036	0.1179	0.731
* .	* .	2	-0.136	-0.138	1.8046	0.406
* .	* .	3	-0.138	-0.152	3.5588	0.313
. .	* .	4	-0.055	-0.093	3.8414	0.428
. .	* .	5	-0.012	-0.066	3.8546	0.571
. .	. .	6	0.047	-0.002	4.0647	0.668
. .	* .	7	-0.041	-0.075	4.2257	0.753
. .	* .	8	-0.054	-0.077	4.5098	0.808
* .	* .	9	-0.071	-0.105	5.0161	0.833
. .	. .	10	0.073	0.024	5.5502	0.852
* .	* .	11	0.129	0.089	7.2336	0.780
. .	. .	12	0.002	-0.004	7.2340	0.842

**Figura 4. Correlograma dos resíduos do ARIMA(0,1,1) e as respectivas estatística QLB**

Included observations: 88

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
. .	. .	1	0.035	0.035	0.1124	0.737
. .	. .	2	0.028	0.027	0.1843	0.912
* .	* .	3	-0.138	-0.140	1.9504	0.583
* .	. .	4	-0.065	-0.057	2.3500	0.672
* .	. .	5	-0.066	-0.055	2.7715	0.735
. .	. .	6	0.002	-0.010	2.7718	0.837
* .	* .	7	-0.073	-0.088	3.2863	0.857
* .	* .	8	-0.060	-0.079	3.6432	0.888
* .	* .	9	-0.066	-0.071	4.0779	0.906
. .	. .	10	0.074	0.054	4.6267	0.915
* .	* .	11	0.115	0.089	5.9878	0.874
. .	. .	12	0.024	-0.022	6.0455	0.914