



UNIVERSIDADE EDUARDO MONDLANE

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE MATEMÁTICA E INFORMÁTICA

TRABALHO DE LICENCIATURA EM ESTATÍSTICA

APLICAÇÃO DOS MODELOS UNIVARIADOS PARA A
ANÁLISE DAS IMPORTAÇÕES DE AUTOMÓVEIS LIGEIROS, NO
PERÍODO DE JANEIRO DE 2002 À JUNHO DE 2008
CASO: PROVÍNCIA E CIDADE DE MAPUTO

Autor: Cachimo Combo Assane

Maputo, Janeiro de 2009

UNIVERSIDADE EDUARDO MONDLANE

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE MATEMÁTICA E INFORMÁTICA

TRABALHO DE LICENCIATURA EM ESTATÍSTICA

**APLICAÇÃO DOS MODELOS UNIVARIADOS PARA A
ANÁLISE DAS IMPORTAÇÕES DE AUTOMÓVEIS LIGEIROS, NO
PERÍODO DE JANEIRO DE 2002 À JUNHO DE 2008
CASO: PROVÍNCIA E CIDADE DE MAPUTO**

Autor: Cachimo Combo Assane

Supervisor: dr. Tiago Guilherme Devesse

Maputo, Janeiro de 2009

DECLARAÇÃO DE HONRA

Declaro que o presente trabalho é resultado da minha profunda investigação e foi realizado apenas para ser submetido como **Trabalho de Licenciatura em Estatística** na Universidade Eduardo Mondlane.

Maputo, 2008

Cachimo Combo Assane

(Cachimo Combo Assane)

DEDICATÓRIA

Dedico este trabalho à todos naturais e simpatizantes da ilha do Ibo, terra que testemunhou o meu nascimento, e hoje me sinto muito orgulhoso por pertencer a esta maravilhosa ilha.

Ao meu avó Assane Combo que sempre acreditou e esperou que um dia pudesse realizar o seu sonho, de ver um dos seus netos licenciado.

Às minhas sobrinhas Muassite e Tima, que muito cedo perderam o amor e carinho da sua mãe e que sempre podem contar com todo meu apoio

À toda minha família

À minha Djia.

AGRADECIMENTOS

Em primeiro lugar agradeço à ALLAH (o único, rico em clemência, abundante em misericórdia), pela protecção, oportunidade de viver com saúde e forças para que eu pudesse enfrentar todos os obstáculos nesta caminhada que foi concluída com muita satisfação e alegria.

Ao meu orientador, dr. Tiago Guilherme Devesse, agradeço pelo incentivo dado, pela confiança, paciência, amizade, pela dedicação e disponibilidade em estar sempre pronto a me ajudar e pelas suas valiosas contribuições nesta pesquisa.

Aos meus pais, Combo Assane e Nema Buana, aos meus irmãos Bimo, Dede, Cumanda, Binha, agradeço de coração, pela ajuda, carinho, incentivo, paciência e apoio.

Aos meus familiares com os quais não pude conviver muito este último ano por causa deste trabalho.

Aos meus tios Sheik Mussa Tamimo, Ámina, Mussa Imbeve, Amade Assane, pelo apoio moral, material, logísticos.

Aos colegas do piso, Guiliche, Mussalama, Truzão, Chato, Eridja, Massada, Tomo, Culuze, Chane, Sélio, André, Chivale, Lampião, Marcelo e à todos membros do movimento UJOMU que carinhosamente aceitaram a minha maneira de ser e sempre incentivaram e apoiaram as minhas decisões.

Aos colegas de Licenciatura em Estatística e colegas da Universidade, especialmente ao meu amigo e irmão Bilika, pelas horas de estudo, por todas as suas contribuições nos trabalhos, debates e seminários.

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

Muito obrigado a todos que de uma ou de outra forma, contribuíram para a concretização deste trabalho, pois o mérito também é desse conjunto de pessoas.

Resumo: *O presente trabalho desenvolve um modelo de previsão das importações de automóveis ligeiros para província e cidade de Maputo, com base nos dados do período de Janeiro de 2002 à Junho de 2008. Para o efeito, foram aplicados os modelos univariados desenvolvidos por Box e Jenkins. Foi apresentada uma revisão da literatura sobre alguns modelos estatísticos de previsão, principalmente o modelo ARIMA, com suas particularidades e aplicabilidades. Durante o período em estudo, constatou-se a não estacionariedade da série em nível, tornando-a através da transformação em primeira diferença. Dos modelos identificados, o ARIMA(0,1,1) sugeriu melhor ajuste aos dados, por conseguinte, o melhor para a previsão. As previsões das importações para o período de Julho de 2008 à Março de 2009 indicam uma tendência crescente.*

Palavras chaves: *Estacionariedade, Importações de automóveis ligeiros, Modelos ARIMA de Box e Jenkins, Séries temporais.*

Lista das abreviaturas

- ACF— *Autocorrelation Function*
AR - *Auto-Regressive*
ARCH- *Auto-Regressive Conditional Heteroscedasticity*
ARIMA- *Auto-Regressive Integrated Moving Average*
ARMA— *Auto-Regressive Moving Average*
ATM— *Autoridade tributária de Moçambique*
GARCH -*Generalized ARCH*
IAL— *Importacoes de automóveis ligeiros*
INAV— *Instituto Nacional de Viação*
MA— *Moving Average*
PACF— *Partial Autocorrelation Function*
PE— *Processo Estocástico*
SARIMA— *Seasonal Auto Regressive Integrated Moving Average*
USAID— *United States Agency International Development*

Lista de Tabelas

Tabela 1. Padrões teóricos das Funções de autocorrelação e autocorrelação parcial.....	25
Tabela 2. Estatísticas descritiva da série de IAL no período de Janeiro de 2002 a Junho de 2008.....	39
Tabela 3. Teste de homogeneidade das variâncias.....	40
Tabela 4. Teste de normalidade nas importações mensais.....	41
Tabela 5. Teste de igualdade das médias mensais de IAL.....	41
Tabela 6. Estatística de teste de raiz unitária de ADF.....	43
Tabela 6.1. Estatística de teste para Regressão de ADF.....	43
Tabela 7. Estatística de teste de raiz unitária de ADF para a série de IAL em primeira diferença.....	44
Tabela 7.1. Estatística de teste para Regressão de ADF para a série de IAL em primeira diferença.....	44
Tabela 8. Estimação do modelo ARIMA(0,1,26) sem o termo intercepto.....	46
Tabela 9. Estimação do modelo ARIMA(0,1,25) sem intercepto.....	46
Tabela 10. Estimação do modelo ARIMA(0,1,1) sem intercepto.....	47
Tabela 11. Estatísticas de teste de Box-Ljung para resíduos do modelo ARIMA(0,1,26).....	48
Tabela 12. Teste de hipótese da média dos resíduos de ARIMA(0,1,26) ser igual a zero.....	49
Tabela 13. Teste de igualdade de variâncias dos resíduos ARIMA(0,1,26).....	49
Tabela 14. Teste ARCH dos resíduos do ARIMA(0,1,26).....	49
Tabela 15. Teste de normalidade para os resíduos do modelo ARIMA(0,1,26).....	50
Tabela 16. Estatísticas do desempenho nas previsões dos modelos estimados.....	53
Tabela 17. Previsão das importações de automóveis ligeiros para o período de Julho de 2008 à Março de 2009	54
Tabela 1A. Dados mensais de importações de automóveis ligeiros na província e cidade de Maputo.....	60
Tabela 1B. Número médio mensal de IAL e seus respectivos intervalos de confiança.....	61
Tabela 1C. Estimação do modelo ARIMA(0,1,26) com intercepto.....	62
Tabela 2C. Estimação do modelo ARIMA(0,1,25) com intercepto.....	62
Tabela 3C. Estimação do modelo ARIMA(0,1,1) com intercepto.....	62
Tabela 1D. Estatísticas de teste de Box-Ljung para resíduos do modelo ARIMA(0,1,25).....	63
Tabela 2D. Estatísticas de teste de Box-Ljung para resíduos do modelo ARIMA(0,1,1).....	63
Tabela 3D. Teste de hipótese da média dos resíduos de ARIMA(0,1,25) e ARIMA(0,1,1) ser igual a zero.....	64
Tabela 4D. Teste de normalidade para os resíduos dos modelos ARIMA(0,1,25) e ARIMA(0,1,1).....	64
Tabela 5D. Teste de igualdade de variâncias dos resíduos dos modelos ARIMA(0,1,25) e ARIMA(0,1,1).....	64
Tabela 6D. Teste ARCH dos resíduos do ARIMA(0,1,25).....	64
Tabela 7D. Teste ARCH dos resíduos do ARIMA(0,1,1).....	64

Lista de Figuras

Figura 1. Representação gráfica das importações de automóveis ligeiros no período de Janeiro de 2002 a Junho de 2008	38
Figura 2. Gráfico de barra do erro de IAL e os respectivos intervalos de confiança	40
Figura 3. Correlograma da função de autocorrelação da série IAL em nível	42
Figura 4. Correlograma da função de autocorrelação parcial da série IAL em nível	42
Figura 5. Representação gráfica da série de IAL em primeira diferença	44
Figura 6. Correlograma da função de autocorrelação da série IAL em primeira diferença	45
Figura 7. Correlograma da função de autocorrelação parcial da série IAL em primeira diferença	45
Figura 8. Representação gráfica dos resíduos obtidos do modelo ARIMA(0,1,26)	48
Figura 9. Correlograma da função de autocorrelação dos resíduos obtidos do modelo ARIMA(0,1,26)	48
Figura 10. Correlograma da função de autocorrelação dos resíduos obtidos do modelo ARIMA(0,1,26)	48
Figura 11. Histograma dos resíduos obtidos do modelo ARIMA(0,1,26)	50
Figura 12. Representação gráfica dos resíduos obtidos do modelo ARIMA(0,1,25)	51
Figura 13. Correlograma da função de autocorrelação dos resíduos obtidos do modelo ARIMA(0,1,25)	51
Figura 14. Correlograma da função de autocorrelação dos resíduos obtidos do modelo ARIMA(0,1,25)	51
Figura 15. Histograma dos resíduos obtidos do modelo ARIMA(0,1,25)	51
Figura 16. Representação gráfica dos resíduos obtidos do modelo ARIMA(0,1,1)	52
Figura 17. Correlograma da função de autocorrelação dos resíduos obtidos do modelo ARIMA(0,1,1)	52
Figura 18. Correlograma da função de autocorrelação dos resíduos obtidos do modelo ARIMA(0,1,25)	52
Figura 19. Histograma dos resíduos obtidos do modelo ARIMA(0,1,1)	52
Figura 1B. Representação gráfica da série das importações de automóveis ligeiros e a sua tendência	61
Figura 1E. Previsão dentro da amostra, limites de confiança e estatísticas de desempenho do ARIMA(0,1,26)	65
Figura 2E. Previsão dentro da amostra, limites de confiança e estatísticas de desempenho do ARIMA(0,1,25)	65
Figura 3E. Previsão dentro da amostra, limites de confiança e estatísticas de desempenho do ARIMA(0,1,1)	65

ÍNDICE

Capítulo I

1. INTRODUÇÃO.....	1
1.1. DEFINIÇÃO DO PROBLEMA.....	4
1.2. JUSTIFICATIVA.....	4
1.3. OBJECTIVO GERAL.....	5
1.3.1. Objectivos específicos.....	5

Capítulo II

2. REVISÃO BIBLIOGRÁFICA.....	6
2.1. Automóvel.....	6
2.2. Importação.....	6
2.3. Abordagem de Séries Temporais.....	7
2.3.1. Estacionaridade.....	9
2.3.2. Função de autocorrelação.....	11
2.3.3. O Modelo auto-regressivo (AR).....	13
2.3.3.1. Função de autocorrelação parcial.....	15
2.3.4. O Modelo de médias móveis (MA).....	17
2.3.5. O modelo auto-regressivo de média móveis (ARMA).....	18
2.3.6. O modelo auto-regressivo integrado de médias móveis (ARIMA).....	19
2.4. Testes da estacionaridade.....	21
2.4.1. Teste de Análise Gráfica e Correlograma.....	21
2.4.2. O Teste de Raiz Unitária.....	21
2.5. Etapas da metodologia de Box-Jenkins.....	24
2.5.1. Identificação do modelo.....	24
2.5.2. Estimção dos parâmetros do modelo.....	25
2.5.3. verificação de diagnóstico.....	26
2.5.3.1. Análise dos resíduos do modelo.....	27
2.6. Critério de selecção de Modelos.....	27

Capítulo III

3. METODOLOGIA.....	30
3.1. Material e Métodos.....	30
3.1.1. Dados usados.....	30
3.1.2. Fonte de dados.....	30
3.1.3. Análise de dados.....	30
3.1.3.1. Técnicas descritivas.....	30
3.1.3.2. Análise de sazonalidade.....	31

3.1.3.2. Análise de sazonalidade	31
3.1.4. Testes de estacionaridade	32
3.1.4.1. Análise gráfica.....	32
3.1.4.2. Teste de raiz unitária	33
3.1.5. Identificação dos parâmetros.....	34
3.1.6. Estimação	34
3.1.7. Verificação de diagnóstico	35
3.1.8. Critério de selecção do modelo	36

Capítulo IV

4. ANÁLISE E DISCUSSÃO DOS RESULTADOS	38
4.1. Análise exploratória de dados	38
4.2. Testes de estacionaridade	42
4.2.1. Testes de raiz unitária.....	43
4.3. Identificação dos parâmetros do modelo.....	44
4.4. Estimação dos modelos	46
4.5. Verificação de diagnóstico	47
4.5.1. Selecção do modelo.....	53
4.6. Previsão	54

Capítulo V

5. CONCLUSÕES.....	55
5.1. Recomendações.....	56
5.2. Limitações do estudo.....	56
6. REFERÊNCIAS BIBLIOGRÁFICAS	57
ANEXOS.....	60
ANEXO A: Dados de importações de automóveis ligeiros	60
ANEXO B: Análise exploratória de dados.....	61
ANEXO C: Estimação dos modelos com interceptos	62
ANEXO D: Análise residual dos modelos ARIMA(0,1,25) e ARIMA(0,1,1)	63
ANEXO E: Previsão dentro da amostra	65

1. INTRODUÇÃO

O investimento dos empresários na área de importação e venda de viaturas, na sua maioria em segunda mão e provenientes principalmente da Europa, Japão e dos Emiratos Árabes Unidos (Dubai), sobretudo na capital país, influencia no crescimento do parque automóvel, revelando um desconforto tanto com custos económicos como ambientais.

O número de veículos no parque automóvel nacional passou de 240 mil em 2006 para 265 mil em Junho de 2008. Destes, cerca de 70 por cento são importados a partir do sul do país, particularmente através das fronteiras com a África do Sul e Suazilândia, bem como a partir do Porto de Maputo (INAV, 2007)¹.

Para algumas opiniões, este fenómeno é reflexo do crescimento económico. Inicialmente um objecto de luxo, quase exclusivo para elites, rapidamente se fez acessível às massas. Para outros, há que fazer uma leitura cautelosa deste fenómeno que no lugar de ser resultado do almejado desenvolvimento pode vir a criar sérios problemas, sob ponto de vista ambiental e, conseqüentemente, de saúde pública (USAID², 2006).

Martins e Laugeni (1999) consideram a previsão como um processo metodológico para a determinação de dados futuros baseados em modelos estatísticos, matemáticos ou econométricos, ou ainda em modelos subjectivos apoiados em uma metodologia de trabalho clara e previamente definida.

Vários modelos e métodos Univariados de previsões já têm sido experimentados, com graus variados de sucesso. Dentre os quais destacam-se os modelos AR (*Auto-Regressive*), ARMA(*Auto-Regressive Moving Average*), ARIMA(*Auto-Regressive Integrated Moving Average*), SARIMA(*Seasonal Auto Regressive Integrated Moving Average*), ARCH(*Auto-Regressive Conditional Heteroscedasticity*) e GARCH (*Generalized ARCH*).

¹ Afirmou o chefe do Departamento de Veículos no Instituto Nacional de Viação (INAV) -www.zambezia.co.mz- 05/08/2008

² Estudo feito pela Agencia Internacional dos Estados Unidos para Desenvolvimento (USAID) em Dezembro de 2006, sobre "Política dos transportes urbanos para o Grande Maputo" encomendado pela confederação das Associações Empresariais de Moçambique (CTA)

Segundo MAKRIDAKIS *et al.*(1983), as técnicas de previsão auxiliam na tomada de decisões em actividades que necessitam de planeamento e redução da incerteza, tornando assim os possíveis e futuros riscos mais visíveis e conseqüentemente mais controláveis.

Portanto, a utilização de modelos estatísticos Univariados para previsão das importações de viaturas, pode constituir uma ferramenta fundamental no processo de tomada de decisão, uma vez que o aumento rápido do parque automóvel nacional tem, naturalmente, reflexos na circulação. Muito claramente, na cidade de Maputo. Primeiro, devido à falta de espaços de estacionamento em muitas zonas. Segundo, devido a congestionamentos e engarrafamentos em determinadas artérias e a certas horas, o que provavelmente tem originado um número elevado de acidentes.

A modelagem de séries temporais possibilita descrever um processo estocástico, utilizando valores passados da variável de interesse, baseando-se na ideia de que observações passadas contêm informações sobre o padrão de comportamento da série temporal. Entre as diversas metodologias utilizadas para este fim, destacam-se os modelos univariados de Box & Jenkins e os modelos de Amortecimento Exponencial (Morretin e Toloí, 1987).

Silva e Silva (1996) aplicaram o método de Box e Jenkins à série de preço do metro cúbico de carvão vegetal em Minas Gerais, no período de Janeiro de 1980 a Dezembro de 1992. Das várias metodologias estudadas por esses autores, a que forneceu o melhor modelo foi a de Box & Jenkins. Paula (1999) aplicou o método para modelagem e previsão da população brasileira. 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.

Com o presente trabalho pretende-se explorar as diversas técnicas oferecidas pela metodologia de Box e Jenkins (1976) para analisar a evolução temporal nas importações de automóveis ligeiros na cidade e província de Maputo.

O trabalho comporta 5 capítulos, a saber:

No capítulo 1 é feita a apresentação e/ou fundamentação do tema bem como a definição do problema do estudo. São também apresentados os objectivos geral e específicos a alcançar e as principais motivações que conduziram na realização da pesquisa.

A revisão da literatura é apresentada no capítulo 2. Neste capítulo são introduzidos alguns conceitos fundamentais tais como séries temporais, estacionariedade, definição de alguns modelos estatísticos de previsão como AR, MA, ARMA, ARIMA, SARIMA, bem como os aspectos teóricos sobre as principais etapas da metodologia de Box & Jenkins.

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

No capítulo 4 são apresentados os resultados das estimações dos modelos ARIMA propostos neste estudo e sua respectiva discussão, bem como resultados sobre a previsão das importações de automóveis ligeiros para o período de Julho de 2008 à Março de 2009.

Por último são apresentadas, no capítulo 5, as principais conclusões do estudo e algumas recomendações para futuras pesquisas.

1.1. DEFINIÇÃO DO PROBLEMA

Os problemas ligados ao aumento do parque automóvel e do sector dos transportes, impõem que a sociedade encontre soluções alternativas claras e adequadas a um ambiente urbano moderno. Em primeiro lugar, a evolução do sector dos transportes está ligada a um incremento do volume de emissões de gases de combustão. Em segundo, o seu impacto é visível na qualidade de vida das cidades e as consequências nefastas para a população; congestionamentos de tráfego, o aumento do tempo de deslocação dentro das cidades, o estacionamento abusivo, as dificuldades em desenvolver e manter as infra-estruturas necessárias, são exemplos claros desses efeitos.

Segundo a USAID (2006), os líderes empresariais e cívicos preconizam o uso de transportes público de grande dimensão como uma das formas de aliviar o congestionamento e o caos produzido pela grande frota de pequenos autocarros.

Na maioria dos problemas o passado traz informações relevantes sobre o que irá ocorrer no futuro, pois existe “correlação” entre as variáveis em diversos instantes. É claro que o conhecimento do passado não nos diz exactamente como será o futuro, e então sempre existe incerteza associada às nossas previsões, mas podemos ter uma boa ideia de quais serão os valores mais prováveis no futuro, ou seja, podemos especificar previsões futuras e limites de confiança (Pires, 2001).

Considerando que as importações futuras obedecem um processo aleatório, coloca-se a seguinte questão: *Como modelar a série das importações de automóveis ligeiros tendo em conta apenas a informação dos registos passados?*

1.2. JUSTIFICATIVA

Segundo Fava (2000), a metodologia sugerida por Box e Jenkins é uma das mais utilizadas para descrever séries cronológicas. Ela parte do princípio de que cada valor da série (temporal) pode ser explicado por valores prévios, a partir do uso da estrutura de correlação temporal que geralmente há entre os valores da série. Para Pires (2001), esta metodologia constrói modelos que descrevem com precisão, e de forma parcimoniosa, o processo gerador da série temporal, proporcionando, dessa forma, previsões precisas de valores futuros.

Segundo FONSECA *et al*, (1995) a principal vantagem desta metodologia reside no facto de se poder trabalhar com séries temporais, com elevada autocorrelação nos resíduos. Portanto, o que nas outras metodologias mais tradicionais (como, por exemplo, nos modelos de regressão) são problemas, no

método Box-Jenkis acaba sendo uma vantagem, pois o padrão das autocorrelações regular e parcial permite analisar o comportamento intrínseco da variável em estudo, possibilitando, dessa forma, um nível a mais de explicação sobre a própria variável — os Modelos Univariados.

Neste contexto, considerando as vantagens da metodologia desenvolvida por Box e Jenkins, trabalharei com a série mensal de importações de viaturas para estimação dum modelo capaz de fazer melhores previsões.

1.3. OBJECTIVO GERAL

- Desenvolver um modelo de previsão a partir da série das importações de automóveis ligeiros referente a província e cidade de Maputo no período de Janeiro de 2002 à Junho de 2008;

1.3.1. Objectivos específicos

- Fazer uma análise descritiva da série de registos mensais das importações de automóveis ligeiros para verificar o seu comportamento ao longo do período em estudo;
- Identificar e estimar o modelo ARIMA de Box-Jenkis que melhor se ajusta aos dados da série;
- Utilizar o modelo estimado para fazer previsões das importações de automóveis ligeiros para o período de Julho de 2008 à Março de 2009;

2. REVISÃO BIBLIOGRÁFICA

Este capítulo apresenta alguns conceitos fundamentais tais como séries temporais, estacionariedade, definição de alguns modelos estatísticos de previsão, como AR, MA, ARMA, ARIMA, SARIMA e abordagem sobre principais aspectos teóricos da metodologia de Box & Jenkins.

2.1. Automóvel

Segundo DIPLOMA MINISTERIAL N°128/2007, de 26 de Setembro, artigo 1, são consideradas as seguintes definições:

Definição 1: *Automóvel* – veículo de tracção mecânica destinada a transitar pelos seus próprios meios na via pública.

Definição 2: *Automóvel Ligeiro* – veículo automóvel cuja a lotação ou peso bruto não excedam respectivamente, nove (9) lugares ou a 3500 kg.

Definição 3: *Peso bruto* – conjunto da tara e da carga que o veículo pode transportar.

Definição 4: *Tara* – peso de um veículo sem carga.

2.2. Importação

Pelo DECRETO N° 30 / 2002, de 2 de Dezembro, artigo 1, consideram-se as seguintes definições:

Definição 5: *Importação* — a entrada de mercadorias no território aduaneiro;

Definição 6: *Território Aduaneiro* — todo o espaço geográfico em que a República de Moçambique exerce a sua soberania;

A entrada ou saída de mercadorias, pessoas e veículos no ou do território aduaneiro está sujeita ao controlo das Alfândegas e somente poderá realizar-se através dos portos, aeroportos e estâncias aduaneiras devidamente habilitadas para o efeito e em conformidade com as normas deste decreto, normas complementares e de regulamentação da matéria (DECRETO N° 30 / 2002, de 2 de Dezembro, artigo 3).

2.3. Abordagem de Séries Temporais

Definição 7: Uma série temporal (também chamada por alguns autores série cronológica ou série histórica) é um conjunto de observações de uma dada variável, ordenadas no tempo (não necessariamente igualmente espaçadas), e que apresentam como característica central a presença de uma dependência serial entre elas, isto é, dependência entre instantes de tempo (Ehlers, 2005, Ribeiro e Paula, 2000).

Morettin e Toloi (1987) mostram que, quando se analisa uma série temporal os objectivos básicos são:

- i) Modelagem do fenómeno sob consideração;
- ii) Obtenção de conclusões em termos estatísticos;
- iii) Avaliação da adequação do modelo em termos de previsão.

Deste modo, a aplicação de séries temporais utiliza a avaliação dos dados históricos para obter modelos de previsão.

O objecto fundamental da análise de uma série temporal é determinar suas componentes básicas buscando identificar um padrão de comportamento da série que possibilite fazer previsões (Makridakis *et al.*, 1998).

Para a análise de qualquer série temporal, é essencial que, primeiramente, seja feita a representação gráfica dos dados ao longo do tempo, a fim de obter uma visão qualitativa geral do seu comportamento (FONSECA *et al.*, 1995). Entretanto, a inspecção visual do comportamento da série é apenas um indicador, sendo necessárias, as análises estatísticas para uma conclusão mais verdadeira sobre o comportamento da série.

Segundo Spiegel (1985), as curvas obtidas pelas séries temporais podem apresentar quatro características principais:

1. *Tendência*: indica a direcção geral dos valores estudados. A principal característica deste componente é o movimento suave registado num período longo de tempo, direccionando os dados de modo constante, crescente ou decrescente. Makridakis *et al.* (1998), asseguram que um padrão de tendência existe quando há um aumento ou diminuição do valor médio dos dados a longo prazo.

2. *As variações Cíclicas*: referem-se às oscilações a longo prazo ou os desvios em torno da recta ou da curva de tendência. Esses ciclos, como são frequentemente denominados, podem ser ou não periódicos, isto é, podem ou não seguir exactamente padrões análogos, depois de intervalos de tempos iguais. Para Makridakis *et al.* (1998), são alterações resultante de factores excepcionais, como greves ou catástrofes climáticas, que não podem ser previstos e, portanto, inclusos no modelo.

3. *As variações sazonais*: são flutuações que se repetem periodicamente, acompanhando um padrão temporal (relacionadas ao factor tempo), como por exemplo, influências climáticas ou a demanda por algum produto ocorrem somente em determinado período do ano, onde se pode verificar a tendência de picos de vendas. A sazonalidade refere-se às mudanças ou variações cíclicas de curto prazo. São bastante parecidos com os cíclicos, com a diferença de que os fenómenos cíclicos são caracterizados por variações que oscilam em torno da tendência a intervalos aproximadamente regulares de tempo e em longo prazo. Porém, alguns autores preferem não fazer diferença entre os ciclos e a sazonalidade (ex.: Morettin & Tolo, 1995).

4. *Aleatoriedade (ou Erro)*: Também chamado de erro residual possui uma sequência de variáveis aleatórias independentes e identicamente distribuídas (*i.i.d*), frequentemente com média zero e variância constante. São caracterizados por sua duração curta e intensidade variável. Para Mentzer e Bienstock (1998) esta componente da série temporal é denominada “ruído”. Neste caso, os dados passados apresentam variações aleatórias. Estas variações, por sua vez, não podem ser modeladas pelas técnicas de previsão e são, de modo geral, tratadas pelas médias e desvios padrão ou variâncias.

Definição 8: Segundo Parzen (1972), Um fenómeno aleatório que sugere um processo definido por uma lei de probabilidade durante um tempo é chamado *Processo Estocástico*.

De ponto de vista da teoria matemática de probabilidade, define-se *Processo Estocástico* como um conjunto $\{X(t), t \in T\}$ de variáveis aleatórias que descreve a evolução no tempo ou espaço de um processo físico, onde t representa geralmente o tempo, $X(t)$ é chamado de *estado do processo* no instante t , e é uma variável aleatória. O conjunto T é o conjunto de índices, ou *espaço paramétrico* do processo estocástico. Se o conjunto de índice T é um conjunto finito ou contável, por exemplo, $T = \{0, 1, 2, \dots\}$ diz-se que o processo estocástico é um *processo de parâmetro discreto*. Ao contrário, se o conjunto T é um intervalo (finito ou infinito) de números reais, por exemplo $T = \{t: t \geq 0\}$, o processo estocástico é de parâmetros contínuo.

Uma série temporal $\{X_t\}$ é a realização de uma família de variáveis aleatórias $\{X_t\}$. De outra maneira, um modelo de série temporal para dados observados $\{X_t\}$ é uma especificação das distribuições de uma sequência de variáveis aleatórias $\{X_t\}$ da qual $\{X_t\}$ é denominada uma realização (Brockwell & Davis, 1996).

Para a caracterização da sequência de variáveis aleatórias são necessários somente os momentos de primeira e segunda ordem da distribuição conjunta (Granger & Newbold, 1986). O momento de primeira ordem é definido como o valor esperado ou média de X_t :

$$\mu = E(X_t). \quad (1)$$

O momento da segunda ordem é definido como o produto esperado ou covariância entre X_t e X_{t+k} :

$$\gamma_k = \text{cov}(X_t, X_{t+k}) = E[(X_t - \mu)(X_{t+k} - \mu)] \quad (2)$$

Define-se também a variância de X_t como

$$\gamma_{t,t} = \gamma_0 = \text{cov}(X_t, X_t) = E[(X_t - \mu)(X_t - \mu)] = E[(X_t - \mu)^2] = \text{Var}(X_t) = \sigma_x^2 \quad (3)$$

2.3.1. Estacionaridade

Um pressuposto comum em muitas técnicas lidando com uma série temporal é que esta é estacionária.

Definição 9: Segundo Gujarati (2006), diz-se que um processo estocástico é *estacionário* quando a sua média e a sua variância são constante ao longo do tempo e quando o valor da covariância entre dois períodos de tempo depende apenas da distância, do intervalo ou da defasagem entre os dois períodos de tempo, e não do próprio tempo em que a covariância é calculada. Na literatura sobre séries temporais, tal processo estocástico é conhecido como *fracamente estacionário*, ou *estacionário em covariâncias*, ou *estacionário de segunda ordem*.

A série temporal estocástica fracamente estacionária tem as seguintes propriedades:

(i) *Média:* $E(X_t) = \mu$

(ii) *Variância:* $\text{Var}(X_t) = E(X_t - \mu)^2 = \sigma^2$

(iii) *Covariância:* $\gamma_k = E[(X_t - \mu)(X_{t+k} - \mu)]$

Onde γ_k , representa a *covariância* (ou autocovariância) na defasagem k , isto é, a medida de dependência entre dois valores X (X_t e X_{t+k}) separados por k períodos.

Se uma série temporal não for estacionária só poderemos estudar seu comportamento para o período considerado. Cada conjunto de dados da série temporal será, portanto, um episódio específico. Em consequência, não é possível generalizá-lo para outros períodos de tempo. Por conseguinte, para a finalidade de fazer previsões, essas séries temporais (não estacionárias) tem pouco valor prático (Spiegel, 1985). Por tanto, segundo Brockwell & Davis, (1996), Se a série não for estacionária, é habitualmente possível convertê-la numa outra estacionária recorrendo a algumas técnicas apropriadas:

- a) Se os dados revelam uma tendência ("trend") crescente ou decrescente, pode tentar-se ajustar uma curva e subtraí-la aos dados, ficando com uma série de resíduos. Como o propósito do ajustamento é remover a tendência de longo prazo, habitualmente um ajustamento de uma recta é suficiente;
- b) se a variância não é constante, extrair o logaritmo ou a raiz quadrada da série pode ajudar a estabilizá-la.
- c) Em qualquer caso, pode-se diferenciar a série. Isto é, dada uma série com elementos X_t , constrói-se a série das diferenças Y_t tal que:

$$Y_t = X_t - X_{t-1} \quad (4)$$

Segundo Morettin & Toloí (1987), a transformação mais comum consiste em diferenciar sucessivamente a série original até obter uma série estacionária. A primeira diferença de X_t é definida como 5:

$$\Delta X_t = X_t - X_{t-1}, t = 1, 2, 3, \dots \quad (5)$$

A segunda é dada por 6:

$$\Delta^2 X_t = \Delta(\Delta X_t) = \Delta(X_t - X_{t-1}) = X_t - 2X_{t-1} - X_{t-2} \quad (6)$$

O operador Δ é chamado operador de diferença que segundo Pires (2001), o seu uso visa o ajustamento do ruído de uma série, com a conseqüente remoção da tendência da série. Generalizando 6, a n-ésima diferença da série é dada por

$$\Delta^n X_t = \Delta^{n-1} X_t - \Delta^{n-1} X_{t-1} = \sum_{r=0}^n (-1)^r * \binom{n}{r} * X_{t-r}, \text{ onde } \binom{n}{r} = \frac{n!}{r!(n-r)!} \quad (7)$$

Em situações normais, ainda segundo os autores acima citados, será suficiente tomar uma ou duas diferenças para que a série se torne estacionária. Para dados não sazonais, a primeira diferença é em geral suficiente para induzir estacionariedade aproximada.

O número d de diferenças para tornar a série estacionária é denominado ordem de integração. Deve-se notar que n observações são perdidas ao se calcular a n -ésima diferença.

O outro operador B denominado operador de defasagem ou de atraso³ transforma o valor de um processo no instante t no valor anterior $t-1$, ou seja:

$$Bx_t = x_{t-1} \quad (8)$$

Generalizando (8),

$$B^n x_t = x_{t-n}, n = 0,1,2,\dots \quad (9)$$

O operador B pode ser utilizado na forma polinomial, de maneira que

$$x_t + d_1 x_{t-1} + d_2 x_{t-2} + \dots + d_n x_{t-n} = k_t \quad (10)$$

pode ser escrito como

$$(1 + d_1 B + d_2 B^2 + \dots + d_n B^n) x_t = k_t \quad \text{ou} \quad d(B)x_t = k_t, \quad (11)$$

$$\text{onde } d(B)x_t = (1 + d_1 B + \dots + d_n B^n). \quad (12)$$

Para os diferentes tipos de processos estocásticos (ou séries temporais) estacionários, existe um tipo especial de processo *puramente aleatório* denominado *Ruído Branco*.

Definição 10: Segundo Gujarati (2006), dizemos que um processo estocástico é puramente aleatório ou Ruído branco, quando a sua média é zero, a variância σ^2 é constante e ele é serial não correlacionado.

2.3.2. Função de autocorrelação

Definição 11: Segundo Granger & Newbold (1986), *Autocorrelação de um processo (série temporal)* é autocovariância padronizada, “memória” de um processo, extensão para a qual o valor tomado no tempo t , depende daquele tomado no tempo $t+k$. Por outras palavras, a autocorrelação mede a

³ Conhecido como “Backward shift operator” (Fava 2000)

intensidade com que um valor observado no instante $t+k$ é influenciado por aquele observado no tempo t e é definida como:

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad (13)$$

Define-se a função de autocovariância de uma série temporal estacionária como

$$\gamma_x(k) = \text{cov}(X_{t+k}, X_t). \quad (14)$$

Segundo Hamilton (1994), o tamanho de $\gamma_x(k)$ depende da escala em que X_t é medida. Portanto, para efeito de interpretação, é mais útil padronizar a função de autocovariância dando origem a uma função de autocorrelação (*ACF*) que mede a correlação entre X_t e X_{t+k} .

Deste modo, a função de autocorrelação de uma série temporal estacionária $\{X_t\}$ é definida como

$$\rho_x(k) = \frac{\gamma_x(k)}{\gamma_x(0)} = \frac{\gamma_x(k)}{\sigma^2} = \text{cor}(X_{t+k}, X_t). \quad (15)$$

Conforme Ehlers (2005), a equação 15 cumpre as seguintes propriedades:

1. $\rho_x(0) = 1$
2. A correlação entre X_t e X_{t+k} é a mesma que entre X_t e X_{t-k} , ou seja $\rho_x(k) = \rho_x(-k)$.
3. $-1 < \rho_x(k) < 1$

As funções de autocovariância e de autocorrelação fornecem uma medida útil do grau de dependência entre os valores de uma série temporal em diferentes períodos. As autocorrelações medem ainda o tamanho e a força da “memória” do processo.

O gráfico das autocorrelações amostrais *versus* k é chamado de *Correlograma* e constitui uma ferramenta poderosa para identificação das características da série temporal (Enders, 1995). Porém, não é simples examinar um correlograma e extrair dele as correspondentes propriedades populacionais. O que se faz necessário é averiguar alguns modelos plausíveis que provejam correlogramas de formas reconhecidas.

As funções de autocovariância e autocorrelação amostrais podem ser calculadas para qualquer conjunto de dados e não estão restritas a observações de séries temporais estacionárias (Brockwell & Davis, 1996). Para dados contendo tendência, a ACF exibirá um decaimento lento na medida que t aumenta

2.3.3. O Modelo auto-regressivo (AR)

Segundo Ehlers (2005), num processo *Auto-Regressivo*, o valor presente da série temporal X_t é expresso linearmente em termos dos valores passados da série e da perturbação aleatória ξ_t ocorrendo no instante t . A ordem deste processo depende do valor mais antigo sobre o qual a regressão tem lugar. Mais especificamente o modelo auto-regressivo de ordem p $AR(p)$ pode escrever-se como

$$X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \xi_t = c + \sum_{j=1}^p \phi_j X_{t-j} + \xi_t, \quad (16)$$

onde a série $\{\xi_t\}$ é um processo puramente aleatório ou ruído branco com média zero e variância σ_ξ^2 e os vários ϕ_i são constantes reais ou parâmetros que descreve como X_t se relaciona com X_{t-i} do para $i = 1, 2, \dots, p$.

O modelo $AR(p)$ dado pela equação 16 pode ser reescrito conforme apresentado na equação 17 utilizando o polinómio do operador de defasagem

$$\phi(B)X_t = c + \xi_t. \quad (17)$$

O polinómio $\phi(B)$ de ordem p é chamado de polinómio AR e tem-se que

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

Segundo Murteira *et al* (1993), Para que a condição de estacionariedade para modelos $AR(p)$ seja satisfeita é necessário que as raízes do polinómio $\phi(B)$ estejam fora do círculo unitário (no plano complexo).

Para um modelo auto-regressivo de ordem 1 ou $AR(1)$ que é a versão mais simples dessa classe de modelos, cuja sua representação algébrica é dada pela equação 19:

$$X_t = c + \phi_1 X_{t-1} + \xi_t, \quad (19)$$

a condição de estacionariedade é satisfeita quando $|\phi| < 1$ e que as autocovariâncias (γ_k) sejam independentes. No caso do modelo $AR(1)$, as autocovariâncias são dadas por:

$$\gamma_k = \phi_1^k \gamma_0 \quad (20)$$

e as autocorrelações ρ_k são dadas pela equação 21

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi_1^k \quad (21)$$

A função de autocorrelação decai exponencialmente quando ϕ_1 é positivo; quando ϕ_1 é negativo, a função de autocorrelação também decai exponencialmente, mas apresenta alternância de sinais positivos e negativos.

Assumindo que a condição de estacionariedade está satisfeita, a média do processo $AR(p)$ é obtida tomando os valores esperados em (16)

$$\mu = c + \phi_1 \mu + \phi_2 \mu + \dots + \phi_p \mu, \quad (22)$$

ou ainda,

$$\mu = \frac{c}{(1 - \phi_1 - \phi_2 - \dots - \phi_p)}. \quad (23)$$

As autocovariâncias são calculadas multiplicando-se ambos os lados de (22) por $(X_{t-j} - \mu)$ e tomando os valores esperados,

$$\gamma_k = \begin{cases} \phi_1 \gamma_{k-1} + \dots + \phi_p \gamma_{k-p}, & k = 1, 2, \dots \\ \phi_1 \gamma_1 + \dots + \phi_p \gamma_p + \sigma^2, & k = 0 \end{cases} \quad (24)$$

A autocorrelação é obtida dividindo-se a equação acima por γ_0

$$\rho_k = \phi_1 \rho_{k-1} + \dots + \phi_p \rho_{k-p}, \quad k = 1, 2, \dots \quad (25)$$

O comportamento da função de autocorrelação de um processo auto-regressivo é uma mistura de decaimento exponencial e/ou decaimento senoidal. Se as raízes da equação auto-regressiva forem reais,

então as autocorrelações decairão exponencialmente. Caso as raízes sejam complexas, o decaimento será na forma senoidal (Granger & Newbold, 1986).

2.3.3.1. Função de autocorrelação parcial

A autocorrelação entre duas observações seriais X_t e X_{t+k} eliminando a dependência dos termos intermediários $X_{t+1}, X_{t+2}, \dots, X_{t+k-1}$ é chamada de *Autocorrelação Parcial* (Mills, 1990) e é representada por:

$$\phi_{kk} = \text{cor}(X_t, X_{t+k} | X_{t+1}, \dots, X_{t+k-1}) \quad (26)$$

Define-se a *função de autocorrelação parcial (PACF)* como a sequência de correlações entre (X_t, X_{t-1}) , (X_t, X_{t-2}) , (X_t, X_{t-3}) e assim por diante, desde que os efeitos de defasagens anteriores sobre X_t permaneçam constantes (Hill, Griffiths & Judge, 1999).

Segundo Pires (2001), A função de autocorrelação parcial, $\{\phi_{kk}; k \in N_0\}$, de uma série estacionária $\{X_t; t \in Z\}$ de média nula é tal que o seu coeficiente de ordem k , ϕ_{kk} , representa o coeficiente de correlação parcial entre X_t e X_{t+k} uma vez eliminado o efeito das variáveis aleatórias $\{X_{t+1}, X_{t+2}, \dots, X_{t+k-1}\}$, o qual é o valor ϕ_{kk} no modelo de regressão linear com erros normais:

$$X_{t+k} = \phi_{k1}X_{t+k-1} + \phi_{k2}X_{t+k-2} + \dots + \phi_{kk}X_t + \xi_{t+k} \quad (27)$$

onde ξ_t é ruído branco e ξ_{t+k} independente de $\{X_{t+k-j}; j > 1\}$.

Multiplicando ambos os membros da equação de regressão (27) por X_{t+k-j} , com $j=1, 2, \dots, k$, e tomando valores esperados, obtém-se:

$$E[X_{t+k} * X_{t+k-j}] = \phi_{k1}E[X_{t+k-1} * X_{t+k-j}] + \phi_{k2}E[X_{t+k-2} * X_{t+k-j}] + \dots + \phi_{kk}E[X_t * X_{t+k-j}] + E[\xi_{t+k} * X_{t+k-j}] = \gamma_j = \phi_{k1}\gamma_{j-1} + \phi_{k2}\gamma_{j-2} + \dots + \phi_{kk}\gamma_{j-k}, \quad j=1, 2, \dots, k \quad (28)$$

Dividindo cada uma das equações por γ_0 , obtém-se o seguinte sistema de equações lineares:

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \dots \\ \rho_k \end{bmatrix} = \begin{bmatrix} \rho_0 & \rho_1 & \rho_2 & \dots & \rho_{k-1} \\ \rho_1 & \rho_0 & \rho_1 & \dots & \rho_{k-2} \\ \rho_2 & \rho_1 & \rho_0 & \dots & \rho_{k-3} \\ \dots & \dots & \dots & \dots & \dots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_0 \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \phi_{k3} \\ \dots \\ \phi_{kk} \end{bmatrix} \quad (29)$$

Usando a regra de Cramer, conclui-se então que:

$$\phi_{00} = 1, \quad \phi_{11} = \rho_1, \quad \phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

e, em geral:

$$\phi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \dots & \rho_{k-3} & \rho_2 \\ \dots & \dots & \dots & \dots & \dots \\ \rho_{k-1} & \rho_{k-2} & \dots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \dots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \dots & \rho_{k-3} & \rho_{k-2} \\ \dots & \dots & \dots & \dots & \dots \\ \rho_{k-1} & \rho_{k-2} & \dots & \rho_1 & 1 \end{vmatrix}} \quad (30)$$

Assim, para um processo $AR(p)$ não há correlação entre X_t e X_{t+k} para $k > p$ (Mills, 1990). Então, todos os valores de ϕ_{kk} para $k > p$ são zero e a $PACF$ para um processo $AR(p)$ puro apresenta um “corte” para defasagens maiores que p (Ender, 1995).

Assim, pode-se resumir que um processo $AR(p)$ é descrito por:

- Possuir uma função de autocorrelação, ACF , que é uma combinação de decaimentos exponenciais e senoidais e tamanho finito; e
- Possuir uma função de autocorrelação parcial, $PACF$, que é zero para defasagens maiores que P .

2.3.4. O Modelo de médias móveis (MA)

Num processo dito de *média móvel*, o valor presente da série expressa-se em função dos valores presente e passados das perturbações aleatórias, que formam uma série de ruído branco (Ehlers, 2005). A ordem deste processo depende do valor mais antigo da série de ruído branco considerado; para *um processo de média móvel de ordem q*, $MA(q)$ a série exprime-se como

$$X_t = \mu + \xi_t - \theta_1 \xi_{t-1} - \dots - \theta_q \xi_{t-q} = \mu - \sum_{j=0}^q \theta_j \xi_{t-j}, \theta_0 \equiv 1, \quad (31)$$

onde $\{\xi_t\}$ é ruído branco. De um modo alternativo, recorrendo ao operador de defasagem, a equação 31 pode ser reescrita na forma

$$X_t = \theta(B)\xi_t, \text{ onde } \theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q, \quad (32)$$

$\theta(B)$ é o polinómio do operador de defasagem B .

Um processo $MA(q)$ estacionária é considerado *invertível* se as raízes de

$$1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q = 0$$

se encontrarem fora do círculo unitário, ou seja.

$$\{X_t, t \in Z\} \text{ é invertível} \Leftrightarrow [\theta(B) = 0 \Rightarrow |B| > 1].$$

Para um modelo de média móvel de ordem 1 ou $MA(1)$ que é a versão mais simples dessa classe de modelos, cuja sua representação algébrica é dada pela equação 33:

$$X_t = \xi_t - \theta \xi_{t-1}, \quad t \in Z \quad (33)$$

onde $\{\xi_t; t \in Z\} \sim \text{Ruído Branco}(0, \sigma^2)$ então a série é estacionária para todo o θ real é invertível se e só se $-1 < \theta < 1$. Além disso,

$$\gamma_k = \begin{cases} (1 + \theta^2)\sigma^2 & k = 0 \\ -\theta\sigma^2 & k = 1 \\ 0 & k > 1 \end{cases}, \quad \rho_k = \begin{cases} 1, & k = 0 \\ \frac{-\theta}{1 + \theta^2}, & k = 1 \\ 0, & k > 1 \end{cases}, \quad \phi_{kk} = \begin{cases} 1, & k = 0 \\ \frac{-\theta^k(1 - \theta^2)}{1 - \theta^2(k+1)}, & k \geq 1 \end{cases} \quad (34)$$

Dessa maneira, para um processo $MA(q)$ a ACF apresenta um “corte” para zero para defasagens maiores que q ($k > q$) (Mills, 1990).

Pode-se resumir que um processo $MA(q)$ é descrito por:

- Possuir uma ACF que é zero para defasagens maiores que q ; e
- Possuir uma PACF que é uma combinação de decaimentos exponenciais e senoidais e tamanho infinito.

2.3.5. O modelo auto-regressivo de média móvel (ARMA)

Segundo Ehlers (2005), 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 ARMA.

O modelo ARMA (p,q) de ordem (p,q) , que é uma generalização dos modelos AR (p) e MA (q) , exigirá um número menor de termos e pode ser expresso conforme a equação (35).

$$X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \xi_t - \theta_1 \xi_{t-1} - \theta_2 \xi_{t-2} - \dots - \theta_q \xi_{t-q} \quad (35)$$

$$\text{ou } X_t = c + \sum_{j=1}^p \phi_j X_{t-j} + \xi_t - \sum_{j=1}^q \theta_j \xi_{t-j}, t \in Z \quad (36)$$

com $\{\xi_t; t \in Z\} \sim \text{Ruido Branco}(0, \sigma_\xi^2)$, para algum $\sigma_\xi^2 \in (0, \infty)$ e com $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$
e $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$

Os polinómios $\phi(B)$ e $\theta(B)$ são respectivamente, o polinómio auto-regressivo e o polinómio de média móvel da série $\{X_t; t \in Z\}$ e não possuem raízes comuns. Então:

- A série $\{X_t; t \in Z\}$ do modelo ARMA (p,q) é estacionária se e só o polinómio $\phi(B)$ não possuir raízes no círculo unitário; *i.e.*,
 $\{X_t; t \in Z\}$ é estacionária $\Leftrightarrow [\phi(B) = 0 \Rightarrow |B| > 1]$.
- A série $\{X_t; t \in Z\}$ do modelo ARMA (p,q) é invertível se e só o polinómio $\theta(B)$ não possuir raízes no círculo unitário; *i.e.*,
 $\{X_t; t \in Z\}$ é invertível $\Leftrightarrow [\theta(B) = 0 \Rightarrow |B| > 1]$.

As séries ARMA (1,1) são especificações mais simples dentro as séries ARIMA (p,q), e têm a seguinte representação:

$$X_t = c + \phi X_{t-1} + \xi_t - \theta \xi_{t-1}, \quad t \in Z \quad (37)$$

onde $\{\xi_t\}$ é ruído branco, então:

- $\{X_t; t \in Z\}$ é estacionária se e só $-1 < \phi < 1$, sendo nesse caso a representação de médias móveis do modelo dado por:

$$X_t = \xi_t + (\phi - \theta) \sum_{j=1}^{\infty} \phi^{j-1} \xi_{t-j}, \quad t \in Z. \quad (38)$$

- $\{X_t; t \in Z\}$ é invertível se e só $-1 < \theta < 1$, sendo nesse caso a representação auto-regressiva do modelo dado por:

$$X_t = (\phi - \theta) \sum_{j=1}^{\infty} \theta^{j-1} X_{t-j} + \xi_t, \quad t \in Z. \quad (39)$$

Assumindo que $\{X_t; t \in Z\}$ é estacionária e invertível ($-1 < \phi, \theta < 1$), tem-se:

$$\mu = 0$$

$$\gamma_0 = \frac{1 - 2\phi\theta + \theta^2}{1 - \phi^2} \sigma_{\xi}^2 \quad (40)$$

$$\rho_k = \frac{(\phi - \theta)(1 - \phi\theta)}{1 - 2\phi\theta + \theta^2} \phi^{k-1}, \quad k \in N \quad (41)$$

e a função de autocorrelação parcial decai para zero através da combinação de séries geométricas e/ou sinusoidais amortecidas a partir da ordem 0.

2.3.6. O modelo auto-regressivo integrado de médias móveis (ARIMA)

Os processos definidos anteriormente são apropriados para séries temporais estacionárias. Significa isto que a média da série temporal e as covariâncias entre as suas observações não variam com o tempo. Se o processo não for estacionário, como foi referido anteriormente, terá que se proceder à sua estacionarização. A estacionarização de uma série temporal recomendada por Box e Jenkins, consegue obter-se com operações de diferenciação, mesmo que aplicada sucessivas vezes.

Segundo Ehlers (2005), um processo ARMA no qual X_t é substituído pela sua d -ésima diferença $\Delta^d X_t$, é capaz de descrever alguns tipos de séries não estacionárias. Denotando a série diferenciada por

$$W_t = \Delta^d X_t = (1-B)^d X_t \quad (42)$$

A série W_t é denominada um processo auto-regressivo integrado de médias móveis ou seja ARIMA (p,d,q) dado por

$$W_t = \alpha_1 W_{t-1} + \dots + \alpha_p W_{t-p} + \xi_t + \beta_1 \xi_{t-1} + \dots + \beta_q \xi_{t-q} \quad (43)$$

ou, equivalentemente

$$\phi(B)(1-B)^d X_t = \theta(B)\xi_t \quad (44)$$

Da equação 44 acima pode-se notar que o modelo para X_t é claramente não estacionário já que o polinómio auto-regressivo $\phi(B)(1-B)^d$ tem exactamente d raízes sobre o círculo unitário. Um processo que se torna estacionário após d diferenças é dito *não estacionário homogéneo*.

Segundo Morettin & Toloi (1987), Altas correlações em múltiplos do período s podem indicar sazonalidade estocástica. As séries temporais que apresentam esse comportamento podem ser representadas pelo modelo ARIMA sazonal (conhecido por SARIMA) da seguinte forma

$$\Phi(B^s)\Delta_s^d X_t = \Theta(B^s)\xi_t \quad (45)$$

em que:

- $\Phi(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_p B^{ps}$ é o operador auto-regressivo sazonal de ordem P , estacionário;
- $\Theta(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_q B^{qs}$ é o operador de médias móveis sazonal de ordem Q , invertível;
- $\Delta_s = 1 - B^s$ é o operador diferença sazonal, tal que

$$\Delta_s X_t = X_t - X_{t-s} \quad \text{e} \quad \Delta_s^D = (1 - B^s)^D, \text{ em que } D \text{ é o número de diferenças sazonais.}$$

Uma diferença para o modelo ARIMA usual é que os parâmetros $\alpha_t, \alpha_{t+1}, \dots$ não são ruído brancos. Relacionando-os com o ruído branco tem-se

$$\phi(B)\Delta^d \alpha_t = \theta(B)\xi_t \quad (46)$$

substituindo a equação 46 na equação 45 resulta

$$\phi(B)\Phi(B^s)\Delta^d \Delta_s^D X_t = \theta(B)\Theta(B^s)\xi_t, \quad (47)$$

chamado de modelo sazonal multiplicativo de ordem $(p,d,q)x(P,D,Q)_s$, ou de $SARIMA(p,d,q)x(P,D,Q)_s$. Significa que se deve tomar d diferenças simples e D diferenças sazonais da série X_t , para se obter o processo estacionário

$$W_t = \Delta^d \Delta_s^D X_t \quad (48).$$

2.4. Testes da estacionaridade

2.4.1. Teste de Análise Gráfica e Correlograma

Segundo Gujarat (2006), antes de fazer testes formais, é sempre bom representar graficamente as séries temporais em estudo. Os gráficos dão uma ideia inicial da provável natureza da série temporal, sendo percepção intuitiva o ponto de partida de testes de estacionariedade mais formais.

Um outro teste simples de estacionaridade é o de **Correlograma amostral** baseado na *função de autocorrelação*. Para séries temporais estacionárias, o correlograma enfraquece rapidamente, enquanto para séries temporais não estacionárias ele enfraquece gradualmente, ou seja, os coeficientes de autocorrelação nas varias defasagens são muito altos.

O correlograma de um processo puramente de ruído branco apresenta coeficientes de autocorrelação em torno de zero em varias defasagens. Essa é a imagem de um correlograma de uma série temporal estacionária. Assim, se o correlograma de uma série temporal efectiva se assemelha ao correlograma de uma serie temporal de ruído branco, podemos dizer que essa série temporal provavelmente é estacionária

(Gujarat, 2006).

2.4.2. O Teste de Raiz Unitária

No nível formal, a estacionariedade pode ser testada descobrindo se a série tem ou não uma raiz unitária. O primeiro tipo de teste de raiz unitária foi desenvolvido por DICKEY e FULLER (1979 e 1981). Estes autores consideraram um processo auto-regressivo de ordem 1 (*AR (1)*), conforme descrito a seguir:

$$X_t = \rho X_{t-1} + \xi_t, t=1,2,\dots \quad -1 \leq \rho \leq 1 \quad (49)$$

onde X_0 : valor inicial fixo, ξ_t : seqüências de variáveis aleatórias identicamente e independentemente distribuídas, i.e. $\xi_t \sim IID(0, \sigma_\xi^2)$.

A hipótese nula é de que X_t é não estacionária, ou seja, é uma variável aleatória sem um termo constante (ou intercepto, também denominado *drift*), contra a hipótese alternativa de que X_t é um processo *AR (1)*. Portanto, tem-se que: $H_0: \rho = 1$ contra $H_a: |\rho| < 1$. Para a realização deste teste de hipótese utiliza-se no processo de estimação o modelo de mínimos quadrados ordinários. No entanto, os testes para detecção de raiz unitária e/ou de estacionaridade não utilizam a distribuição padrão *t de Student*, mas sim os valores das distribuições denominadas τ (*tau*) e que foram tabeladas por FULLER (1996)

Outra forma de conduzir o teste de raiz unitária é manipular a equação (49) da seguinte forma: deve-se subtrair X_{t-1} de ambos os lados dessa equação para obter a equação (50).

$$\Delta X_t = (\rho - 1)X_{t-1} + \xi_t \quad (50)$$

que pode ser escrita alternativamente como:

$$\Delta X_t = \delta X_{t-1} + \xi_t \quad (51)$$

onde: $\delta = \rho - 1$.

Nesse caso, a hipótese nula é de que há raiz unitária se $\delta = 0$, enquanto que a hipótese alternativa é de que $\delta < 0$. Portanto, o teste de raiz unitária pode ser realizado com as equações (49), (50) ou (51), pois todas são equivalentes, o que muda em cada caso é a hipótese nula.

Segundo Mills (1990), para se iniciar o teste de raiz unitária é necessário tomar a variável em nível e não diferenciada, porque algumas séries podem ser estacionárias em nível. Se fosse aplicada uma diferença sobre uma variável já estacionária isso implicaria sobre diferenciá-la, conseqüentemente, obteriam-se modelos com especificação enviesada.

Segundo Hill *at al.* (1994), na equação 51, pode ser incorporados termos relacionados com a presença de intercepto β_0 e tendência determinística ($\beta_1 t$) quando o termo de resíduo for *IID*. Logo, o modelo com intercepto fica representado da seguinte maneira:

$$\Delta X_t = \beta_0 + \delta X_{t-1} + \xi_t, t = 1, 2, \dots \quad (52)$$

e o modelo com intercepto e tendência assume o seguinte formato:

$$\Delta X_t = \beta_0 + \beta_1 t + \delta X_{t-1} + \xi_t, t = 1, 2, \dots \quad (53)$$

No caso da existência de intercepto a estatística utilizada é denominada τ_μ e para se testar a presença de intercepto e de tendência deve-se utilizar a estatística τ_τ . DICKEY e FULLER (1981), apresentaram valores críticos para testar de maneira conjunta a presença de um termo de intercepto e/ou tendência e de raiz unitária. Esses testes são denominados testes ϕ e correspondem ao teste *F* padrão. No caso do teste denominado ϕ_1 , testa-se a hipótese nula $(\beta_0, \rho) = (0, 1)$ contra a hipótese alternativa $(\beta_0, \rho) \neq (0, 1)$. O teste ϕ_2 tem como hipótese nula $(\beta_0, \beta_1, \rho) = (0, 0, 1)$ contra a hipótese alternativa $(\beta_0, \beta_1, \rho) \neq (0, 0, 1)$. Finalmente, a estatística ϕ_3 , onde a hipótese nula é $(\beta_0, \beta_1, \rho) = (\beta_0, 0, 1)$ versus a hipótese alternativa $(\beta_0, \beta_1, \rho) \neq (\beta_0, 0, 1)$.

Ainda com relação aos valores críticos, ao invés de utilizar nos testes de raiz unitária os valores críticos elaborados por Dickey & Fuller (1979 e 1981) e Fuller (1996), pode-se usar os valores críticos construídos por MACKINNON (1991), pois são obtidos de forma directa (sem a necessidade de cálculos adicionais), eles têm como base simulações e são suficientemente exactos para os propósitos práticos.

Segundo Gujarat (2006), ao fazer o teste de DICKEY-FULLER para 51, 52 e 53, a pressuposição foi de que o termo de erro ξ_t era não correlacionado. Mas no caso de ξ_t apresentar correlação, DICKEY e FULLER desenvolveram um teste conhecido como *Teste de Dickey-Fuller aumentado* que consiste em estimar a seguinte regressão:

$$\Delta X_t = \beta_0 + \beta_1 t + \delta X_{t-1} + \sum_{i=1}^p \gamma_i \Delta X_{t-i} + \xi_t, t = 1, 2, \dots \quad (54)$$

onde ξ_t é um termo de erro de ruído branco puro e $\Delta X_{t-1} = (X_{t-1} - X_{t-2})$, $\Delta X_{t-2} = (X_{t-2} - X_{t-3})$, etc.

O número de termos de diferenças defasados ser incluído é muitas vezes, determinado empiricamente, sendo que a ideia incluir um número suficiente de defasagens no termo ∇y_{t-j} para eliminar a correlação serial dos resíduos e torná-los ruído branco (*white noise*). No teste de Dickey-Fuller aumentado, continua-se testando a hipótese nula $\delta = 0$ e segue a mesma distribuição assintótica que a estatística de Dickey-Fuller, de modo que podem ser usados os mesmos valores críticos.

A principal vantagem desse teste é que ele pode ser aplicado não somente em modelos com erros de Médias Móveis (MA), mas também, em modelos típicos, nos quais as ordens dos polinómios Auto-regressivos (AR) e de Médias Móveis (MA) são desconhecidas.

2.5. Etapas da metodologia de Box-Jenkins

A construção do modelo, segundo a abordagem clássica, segue um ciclo iterativo formado pelas fases de *identificação, estimativa e verificação*. A fase de previsão é considerada um passo adicional, quando o modelo já se encontra ajustado (Tran, 2002).

2.5.1. Identificação do modelo

A fase de identificação consiste em descobrir qual dentre as várias versões dos modelos de Box-Jenkins, sejam eles sazonais ou não, descreve o comportamento da série. Determina-se a estrutura do modelo, o número de observações passadas e ruídos brancos passados que apresentam uma correlação significativa com a observação actual. Esta fase se baseia na análise de autocorrelações e autocorrelações parciais.

Segundo Mills (1990), no processo de identificação de um modelo ARIMA (também chamado de *identificação estrutural*) procura-se encontrar procedimento que permitem identificar a estrutura do modelo, isto é, conhecer os valores “*p*”, “*d*” e “*q*” que caracterizam a estrutura do modelo ARIMA para a série em estudo.

Um ponto importante na identificação do modelo ARIMA é parcimónia, ou seja, deve-se buscar o modelo com melhor ajuste e melhor número possível de parâmetros. Em outras palavras, dados dois modelos com ajustes igualmente bons, escolhe-se aquele com menor número de parâmetros. Existem duas justificativas para isso: i) simplicidade na estrutura identificada, o que permite uma melhor

compreensão do processo subjacente e ii) em geral os modelos super estimados são maus em termos preditivos, pois se ajustam bem ao que acontece dentro da amostra mas não conseguem generalizar o comportamento da série (Brockwell & Davis, 1996)

O primeiro passo na identificação de um modelo ARIMA para representar uma série temporal é a escolha da ordem de diferenciação, ou seja, quantas diferenças são necessárias para transformar a série original numa outra estacionária.

Em geral, o número correcto de diferenças (d) é a menor ordem de diferenciação tal que a série resultante flutue ao redor de um valor “médio” bem definido e cuja função de autocorrelação decai a zero rapidamente. Se, após a diferenciação, o gráfico da ACF da série ainda exhibe valores significantes em defasagens altas, é necessário diferenciar a série mais vezes.

Gujarat (2006) apresentou a Tabela 1 que resume como usar a ACF e PACF para identificação modelos ARIMA.

Tabela 1. Padrões teóricos das Funções de autocorrelação e autocorrelação parcial

Tipos de Modelos	Padrão típico da ACF	Padrão típico da PACF
$AR(p)$	Declina exponencialmente ou com um padrão de onda senóide amortecida, ou ambos	Apresenta picos significativos Até p defasagens.
$MA(q)$	Apresenta picos significativos até q defasagens	Declina exponencialmente
$ARIMA(p,q)$	Diminui exponencialmente	Diminui exponencialmente

Fonte. (Gujarat, 2006)

2.5.2. Estimação dos parâmetros do modelo

Após a identificação do número de diferença (d) e dos graus dos polinómio AR e MA (p e q , respectivamente), o próximo passo é estimar os parâmetros desconhecidos do modelo, isto é, encontrar os ϕ_i 's e θ_j 's.

Segundo Ehlers (2005), na estimação dos parâmetros desconhecidos procuramos minimizar a soma do quadrado dos erros, ou seja, no caso de uma ARIMA (p,d,q) buscamos minimizar:

$$S(\underline{\xi}) = \sum_{t=1}^T \hat{\xi}_t^2 \quad (55)$$

onde $\xi_t = \theta^{-1}(B)\phi(B)\Delta^d X_t$, e T é tamanho da série.

Por causa das defasagens presentes nos polinômios $\phi(B)$ e $\theta(B)$ é frequente impossível calcular a soma do quadrado dos erros como descrita acima, pois ela dependeria de observações inexistentes ($X_0, X_{-1}, X_{-2}, \dots$). Box e Jenkins (1976) sugerem um procedimento conhecido como “backforecasting”, que gera previsões para estes valores iniciais desconhecidos. Assumindo a normalidade dos erros pode-se escrever a função de verosimilhança sem muita dificuldade, e proceder à estimação dos parâmetros por máxima verosimilhança. Os resultados obtidos pela aplicação deste método são equivalentes aos encontrados pela minimização da soma do quadrado dos erros (equação 55). A estimação por máxima verosimilhança possui uma grande vantagem: a distribuição assintótica (quando T tende a infinito) é conhecida e pode-se encontrar facilmente o erro padrão destes estimadores, o que permite estabelecer a significância dos parâmetros.

2.5.3. verificação de diagnóstico

Uma vez identificada a ordem e estimadas eficientemente os parâmetros de um modelo, procede-se a verificação da sua adequação antes de utilizá-lo por exemplo para fazer previsões. Pode-se fazer testes de sobreajustamento, que consistem em incluir parâmetros extras no modelo e verificar sua significância estatística.

Segundo Brockwell & Davis, (1996), Nesta etapa procura-se responder a seguinte questão: O modelo identificado é adequado e capaz de bem representar a série observada? Existem diversos testes para tentar responder a esta pergunta. A grande maioria dos testes baseia-se nos resíduos dos modelos.

Para o mesmo autor, acima citado, considera que geralmente um bom modelo deve:

- produzir previsões suficientemente próximas da realidade (mas não exactamente iguais aos valores observados);
- Ter resíduos independentes com média zero e variância constante (ruído branco), e então o correlograma dos resíduos não deve revelar a existência de dependência serial;

- Ser parcimonioso, isto é, apresentar um modelo relativamente pequeno de parâmetros.

2.5.3.1. Análise dos resíduos do modelo

Conforme já mencionado, num “bom” modelo, os resíduos não devem exibir qualquer estrutura de dependência. Em particular, a média dos resíduos deve ser zero e a variância deve ser constante. Também, a ACF não deve ser significativa em qualquer “lag” (defasagem). A existência de autocorrelações sugere que o modelo não foi correctamente especificado.

Ao invés de olhar para as autocorrelações residuais individualmente pode-se testar se um grupo de autocorrelações é significativamente diferente de zero através das chamadas estatísticas Q. Segundo Box e Jenkins (1976), Um teste muito usado para verificar a existência de autocorrelações nos resíduos é o teste de Ljung e Box ou teste Portmanteau, descrito a seguir.

No teste de Ljung e Box deseja-se testar a hipótese nula de que as k primeiras autocorrelações dos resíduos são nulas, isto é: $H_0 : \rho_1^{(\xi)} = \rho_2^{(\xi)} = \dots = \rho_k^{(\xi)} = 0$ onde os $\rho_i^{(\xi)}$ são as autocorrelações dos resíduos. A estatística do teste é:

$$Q = T(T+2) \cdot \sum_{i=1}^k \frac{\hat{\rho}_i^{(\xi)^2}}{(T-i)} \quad (56)$$

Onde T é o tamanho da série. A regra de decisão deste teste é: se o valor da estatística Q for grande rejeita-se a hipótese de que as k primeiras autocorrelações dos resíduos sejam nulas. O valor da estatística Q, Deve-se compará-lo aos pontos percentuais de uma variável Qui-Quadrado com $k - (p+q+1)$ graus de liberdade, onde k é o numero de autocorrelações sendo testadas e p e q são os graus dos polinómios AR e MA respectivamente.

A escolha usual de quantas autocorrelações devem ser testadas em conjunto é fazer k aproximadamente igual a \sqrt{T} (Makridakis *et al.*, 1998).

2.6. Critério de selecção de Modelos.

Segundo Ehlers (2005), Em muitas aplicações vários modelos podem ser julgados adequados em termos do comportamento dos resíduos. Uma forma de “discriminar” entre estes modelos competidores

é utilizar os chamados **critérios de informação** que levam em conta não apenas a qualidade do ajuste ou seja, que minimizam a variância dos resíduos mas também penalizam a inclusão de parâmetros extras. Assim, um modelo com mais parâmetros pode ter um ajuste melhor mas não necessariamente ser a preferível em termos de critério de informação. A regra básica consiste em seleccionar o modelo cujo critério de informação calculado seja mínimo.

A regra mais utilizada em séries temporais é o chamado **critério de informação de Akaike**, denotado por **AIC**. A definição mais comumente utilizada é:

$$AIC = -2 \log \text{verossimilhança maximizada} + 2m \quad (57)$$

Onde m é o número de parâmetros (em modelos ARIMA (p,q) , $m=p+q+1$). Para dados normalmente distribuídos e usando-se estimativas de máxima verossimilhança para os parâmetros pode se mostrar que

$$AIC = T * \left[\log(\hat{\sigma}_\xi^2) \right] + 2m \quad (58)$$

onde $\hat{\sigma}_\xi^2 = (1/T) * \sum \xi_i^2$.

O critério AIC tende a super estimar a ordem do polinómio AR , o que motivou a criação de outros critérios de informações que é basicamente modificações do AIC na forma de penalizar a inclusão de parâmetros extras. O mais famoso deles é o **critério de informação Bayesiano**, denotado por **BIC** é dado por

$$BIC = -2 * \log \text{verossimilhança maximizada} + m + m * (\log T)$$

Ou ainda,

$$BIC = T * (\log(\hat{\sigma}_\xi^2)) - (T - m) * \log\left(1 - \frac{m}{T}\right) + m * \log(T) + m * \log\left(\frac{\left(\frac{\hat{\sigma}_x^2}{\hat{\sigma}_\xi^2} - 1\right)}{m}\right) \quad (60)$$

De referir que estas medidas não têm nenhum significado quando olhadas individualmente, i.e. considerando-se um único modelo. Assim, tanto o AIC quanto o BIC podem assumir valores quaisquer, inclusive valores negativos, já que eles dependem da forma da função de verossimilhança.

Vale salientar também que ao usar tais critérios para comparar modelos, a estimação precisa ser feita no mesmo período amostral de modo que os modelos sejam comparáveis

Um outro critério denominado por **Critério de Schwartz (SBC Criterion)** foi sugerido por Schwartz em 1978, e é uma simplificação do critério BIC acima, que a mesma eficiência no que diz respeito à super-estimação de p pelo critério AIC. É dado por:

$$SBC = T * \left(\log \left(\hat{\sigma}_\epsilon^2 \right) \right) + m * \log(T) \quad (61)$$

Nem sempre os três critérios irão concordar quanto ao melhor modelo dentro de uma classe de modelos em consideração, mas em geral as diferenças entre os resultados dos três critérios serão pequenas.

3. METODOLOGIA

Neste capítulo é apresentado o material e métodos necessários para efectivação do estudo tendo em conta os seus objectivos. São apresentadas as etapas principais da aplicação da metodologia de Box-Jenkins.

3.1. Material e Métodos

3.1.1. Dados usados

Para a realização do estudo foram utilizados os dados mensais das importações de automóveis ligeiros registados nas terminais das Alfandegas de Moçambique referente a província e cidade de Maputo, no período de Janeiro de 2002 à Junho de 2008. Portanto, a série de dados contém 78 observações.

A análise de dados foi feita com auxílio dos pacotes estatísticos tais como: Eviews Versão 4.0, SPSS V.13.0 e Excel Versão 2003.

3.1.2. Fonte de dados

A base de dados usada é secundária e foi fornecida pela direcção da tecnologia de informação e comunicação da Autoridade Tributária de Moçambique na cidade de Maputo. A instituição é responsável na conservação dos dados registados em todo território Aduaneiro. Portanto, a partir desta, foi adquirida a base de dados das importações registadas na terminal da província de Maputo.

3.1.3. Análise de dados

3.1.3.1. Técnicas descritivas

A representação gráfica da série de IAL constituiu o primeiro passo de todas as análises feitas no estudo. O gráfico temporal serviu como base na revelação de padrões de comportamento importantes, isto é, foi a partir do gráfico que verificou-se indício da não estacionaridade e a sua conseqüente tendência.

3.1.3.2. Análise de sazonalidade

Para verificação da componente sazonal na série das importações, foi testada a significância estatística das importações médias mensais. Para tal, foi usado o teste F de *Snedecor* da análise de variância (ANOVA) cuja sua estatística, segundo Pestana & Gageiro (2005), é:

$$F_{(k-1; n-k)} = \frac{\text{variação entre os } k \text{ grupos}/(k-1)}{\text{variação dentro de cada grupo}/(n-k)} = \frac{\sum_{i=1}^k n_i (\bar{Y}_i - \bar{Y}_{global})^2}{k-1} \cdot \frac{n-k}{\sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2}$$

onde,

n_j = número de elementos do grupo i ;

\bar{Y}_i = média do grupo i ;

\bar{Y}_{global} = média global;

Y_{ij} = a observação j do grupo i ; com $j = 1, 2, 3, \dots, n_i$;

k = número de categorias do factor.

Sendo μ_1 a μ_k as médias das importações em k meses ($k=1, 2, \dots, 12$), as hipóteses a testar são:

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_k$$

$$H_a : \exists(i, j) \mu_i \neq \mu_j, i \neq j$$

Portanto, se a hipótese nula for rejeitada, podemos afirmar que existe pelo menos um mês que difere significativamente dos outros quanto ao número médio das importações, ou seja, a série em estudo apresenta indício da sazonalidade.

A função da autocorrelação também ajuda na identificação da componente sazonal na série. Uma série com sazonalidade, a sua ACF exhibe um comportamento similar nos respectivos períodos.

Segundo os autores acima citados, antes da aplicação do teste F, é necessário verificar previamente os seguintes pressupostos:

1. As observações devem ser independentes entre si.
2. As observações dentro de cada grupo devem ter distribuição normal.
3. As variâncias de cada grupo devem ser iguais entre si, ou seja, deve haver homocedasticidade.

A normalidade é inferida através dos testes Kolmogorov-Smirnov ou Shapiro Wilk, cuja as hipóteses são:

H_0 : A distribuição das importações no grupo i é normal

H_a : A distribuição das importações no grupo i não é normal, $i = (\text{Janeiro, Fevereiro, ..., Dezembro})$

Para se inferir a homogeneidade das variâncias dos grupos, usa-se o teste de Levene, cujas hipóteses são:

$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_{12}^2$

$H_a : \exists(i, j) \sigma_i^2 \neq \sigma_j^2, i \neq j$

Regra de decisão:

Para todos os testes utilizados no trabalho, se o valor do erro tipo I escolhido ($\alpha = 0.05$) for maior ou igual ao erro do tipo I associado ao teste (sig), rejeita-se a hipótese nula (H_0). Se o α for menor que sig , não se rejeita a hipótese nula.

3.1.4. Testes de estacionaridade

3.1.4.1. Análise gráfica

A estacionaridade é primariamente analisada a partir da representação gráfica da série em nível. Com base no gráfico, muitas vezes são facilmente identificados os padrões como, tendência, padrões cíclicas, sazonalidade, etc. Obviamente, não é fácil dizer que uma série é estacionária ou não, apenas pela análise do gráfico. Por isso, é também analisado o correlograma da função de autocorrelação para verificação do comportamento dos seus respectivos coeficientes nas várias defasagens.

Segundo Gujarati (2006), a escolha de número de defasagens é basicamente uma questão empírica. Um método empírico é calcular a função de autocorrelação até um terço ou um quarto do tamanho da série. Portanto, como a série de IAL contém 78 observações mensais, segundo essa regra, consideramos defasagens de 26 meses.

3.1.4.2. Teste de raiz unitária

A análise das estatísticas do teste de raiz unitária é fundamental para uma conclusão mais verdadeira sobre o comportamento da série, no diz respeito a sua estacionariedade.

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

O teste de ADF consistiu em estimar a regressão apresentada na equação (54):

$$\Delta X_t = \beta_0 + \beta_1 t + \delta X_{t-1} + \sum_{i=1}^p \gamma_i \Delta X_{t-i} + \xi_t, t = 1, 2, \dots$$

onde ξ_t é um termo de erro de ruído branco puro e $\Delta X_t = (X_t - X_{t-1})$.

Os parâmetros a serem estimados são β_0, β_1, δ . A estatística utilizada para testar a existência de intercepto é denominada τ_μ . Enquanto que para se testar a presença de intercepto e de tendência foi utilizada a estatística τ_τ .

As hipóteses testadas nos modelos são:

H₀: $\delta = 0$, isto é, X_t não é estacionária

H_a: $\delta < 0$, ou seja, X_t é $I(0)$

Se o valor absoluto calculado da estatística tau ($|\tau|$) exceder o valor crítico nas estatísticas tau de Mackinnon, rejeitamos a hipótese nula de que $\delta = 0$ e, nesse caso a série de IAL é estacionária. Por outro lado, se o valor absoluto calculado $|\tau|$ não exceder o valor crítico de tau, não rejeitamos a hipótese nula, o que significa que a série de IAL não é estacionária ou integrada.

3.1.5. Identificação dos parâmetros

A etapa da identificação consistiu em analisar os correlogramas das funções de autocorrelação e autocorrelação parcial de modo a avaliar o seu comportamento nas diversas defasagens. Segundo as constatações, foi identificado um modelo ARIMA(0,1,26). Devido a complexidade da etapa, sugeriu-se a identificação dos modelos ARIMA(0,1,25) e ARIMA(0,1,1) para comparação dos seus resultados.

As funções de autocorrelações (ACF) e autocorrelações parcial (PACF) são dadas pelas seguintes equações:

$$ACF: \tau_k = \frac{\sum_{t=k+1}^T (X_t - \bar{X})(X_{t-k} - \bar{X})}{\sum_{t=1}^T (X_t - \bar{X})^2} \quad PACF: \phi_k = \begin{cases} \tau_1, k=1 \\ \frac{\tau_k - \sum_{j=1}^{k-1} \phi_{k-1,j} \tau_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1,j} \tau_{k-1}}, k > 1 \end{cases}$$

3.1.6. Estimação

A estimação dos parâmetros dos modelos foi feita pelo método da máxima verosimilhança, tendo se recorrido a função de densidade de probabilidade da distribuição de Gumbel (f.d.p), da variável aleatória x , associada aos valores de IAL e é dada por:

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

onde x é a variável associada aos valores de IAL no período em análise, β é denominado parâmetro de posição e α é o parâmetro de deslocamento.

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

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

A probabilidade de que ocorra um numero de importações maior que um certo valor x é dada por:

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

Os estimadores dos parâmetros β_i e α pelo método da máxima verosimilhança são obtidos a partir da f.d.p conjunta da distribuição de Gumbel e respectivamente representados pelas seguintes equações:

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

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

3.1.7. Verificação de diagnóstico

Para sabermos se os modelos identificados se ajustam razoavelmente aos dados da série de IAL, foram obtidos os seus resíduos com suas respectivas funções de autocorrelação e de autocorrelação parcial. O objectivo é verificar se os resíduos obtidos estão normalmente distribuídos com média igual a zero e variância constante, ou seja, são ruído branco.

Foi usado o teste de **Ljung e Box** para testar a hipótese nula de que as k primeiras autocorrelações dos resíduos são nulas, isto é: $H_0 : \rho_1^{(\xi)} = \rho_2^{(\xi)} = \dots = \rho_k^{(\xi)} = 0$ onde os $\rho_i^{(\xi)}$ são as autocorrelações dos resíduos. A estatística do teste é apresentada na equação (56).

Para testar a normalidade dos resíduos, foi usado o teste de K-S, cujas hipóteses são:

H_0 : Os resíduos têm distribuição normal

H_a : Os resíduos não estão normalmente distribuídos

Para verificar se a média dos resíduos é igual ou não a zero, foi usado o teste *t-Student*, cujas hipóteses são:

$H_0 : \mu_\xi = 0$

$H_a : \mu_\xi \neq 0$

A estatística do teste é:
$$t = \frac{\bar{X}_\xi - \mu_\xi}{\frac{S_\xi}{\sqrt{N_\xi}}}$$

Quanto a verificação da variância constante dos resíduos entre os meses, foi usado o teste Levene, cujas hipóteses são as mesmas apresentadas para análise da sazonalidade.

Ainda para o mesmo propósito, foi também o teste LM de modo a detectar a existência ou não de um efeito ARCH nos modelos. Segundo Masih(2002), o teste baseia-se na equação seguinte:

$$\hat{\xi}_t^2 = c + \gamma \hat{\xi}_{t-1}^2$$

onde: $\hat{\xi}_t^2$ é a 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

γ é o erro

O objectivo é testar a significância estatística do α , tal como:

$H_0 : \alpha = 0$, há homoscedasticidade dos resíduos

$H_1 : \alpha \neq 0$, há heterocedasticidade dos resíduos

O teste das hipóteses acima baseia-se numa distribuição Qui-Quadrado, com P graus de liberdade, dado por $R^2 * n \sim X_p^2$

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

n é número de observações

R^2 é o coeficiente de determinação da equação estimada.

3.1.8. Critério de selecção do modelo

Para a selecção do modelo (entre os três propostos) que se ajusta razoavelmente aos dados sobre importações de automóveis, foram usados critérios de informação de Akaike e Schwartz que levam em conta não apenas a qualidade do ajuste ou seja, que minimizam a variância dos resíduos mas também penalizam a inclusão de parâmetros extras. A regra básica consistiu em seleccionar o modelo cujo critério de informação calculado seja mínimo.

O desempenho na previsão, para os três modelos identificados, pode ser avaliado a partir das seguintes estatísticas ou medidas (Issler, 1999)

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

onde h é o número de períodos de previsão, S é o tamanho da amostra, \hat{X}_t é o número de importações previstas e X_t número de importações actuais.

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

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

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

As duas primeiras estatísticas dependem da escala de medida da variável dependente. 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.

4. ANÁLISE E DISCUSSÃO DOS RESULTADOS

4.1. Análise exploratória de dados

A representação gráfica dos dados constitui o passo fundamental na identificação do comportamento da série ao longo do tempo. A Figura 1 apresenta a análise gráfica da série das importações de automóveis ligeiros no período de Janeiro de 2002 à Junho de 2008. Durante este período, as importações observam um comportamento que pode ser dividido em três estágios diferentes. O primeiro estágio corresponde ao período de Janeiro de 2002 a Julho de 2005, que sugerem uma tendência crescente das importações. O segundo pertence ao período de Agosto de 2005 a Dezembro de 2006, com uma estabilização das importações mensais em torno de um valor médio de aproximadamente 600 automóveis. O último estágio corresponde ao período de Janeiro de 2007 à Junho de 2008, cuja as importações retornam a uma tendência crescente.

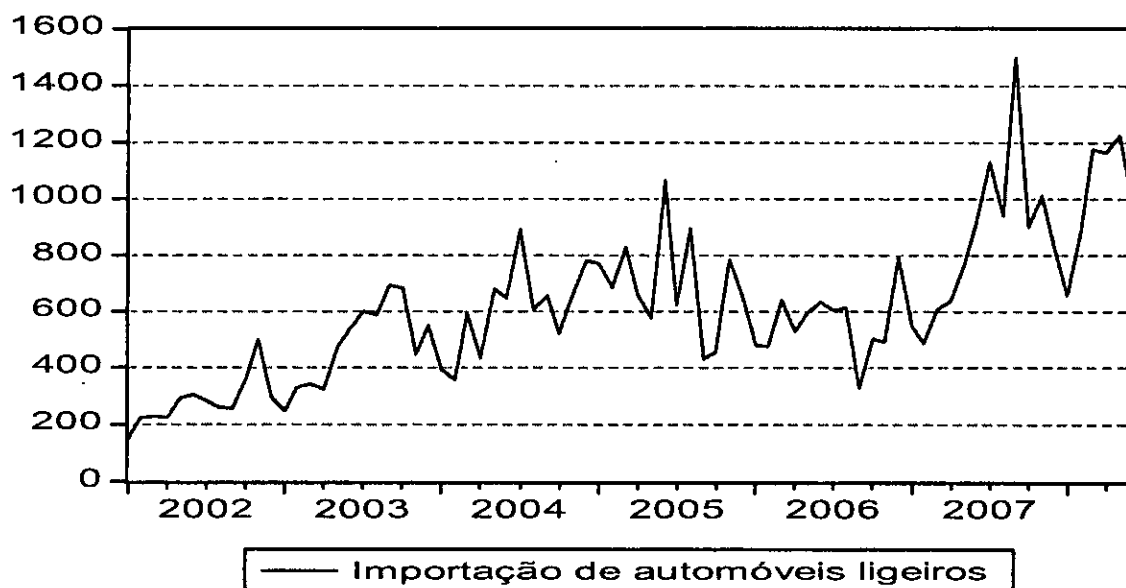


Figura 1. Representação gráfica das importações de automóveis ligeiros no período de Janeiro de 2002 a Junho de 2008

O número médio anual das importações tende a crescer rapidamente, conforme se pode observar a partir da Tabela 1 em Anexo A. Segundo USAID (2006), este crescimento pode ser resultado de uma política governamental, que permite a importação, aparentemente, sem qualquer tipo de controlo de qualidade, de viaturas usadas e recondicionadas. Apesar da mesma política facilitar muitos moçambicanos, em ter acesso a transporte próprio.

As estatísticas da Tabela 2 mostram que o número médio mensal de importações de automóveis ligeiros durante o período em análise foi cerca de 615, com desvio médio em relação a média de aproximadamente 269 viaturas. Em Outubro de 2007 houve maior número de importações com 1499 automóveis, enquanto que Janeiro de 2002 foi o período com menor quantidade das importações, tendo se registado apenas 153 viaturas.

Tabela 2. Estatísticas descritiva da série de IAL no período de Janeiro de 2002 a Junho de 2008

	N	Minimum	Maximum	Mean	Std. Deviation
IAL	78	153	1499	614.31	268.594
Valid N (listwise)	78				

Em média, regista-se maior importação de viaturas no mês de Junho e queda nos primeiros dois meses do ano (Figura 2 e Tabela 1 em anexo B). Entretanto, a autoridade tributária de Moçambique justifica este acontecimento pelo facto da maior parte da população, nos últimos meses do ano, preocupar-se em acumular as suas rendas para épocas das festas, pelo que, no princípio do ano normalmente, regista-se pouca possibilidade da procura deste bem, provavelmente devido a insuficiência de fundos.

A presença da componente sazonal na série de IAL é averiguada empiricamente efectuando-se um teste F, com o objectivo de identificar a significância estatística das médias entre os meses. Na Figura 2 nota-se que as diferenças observadas entre os meses parecem não significativas.

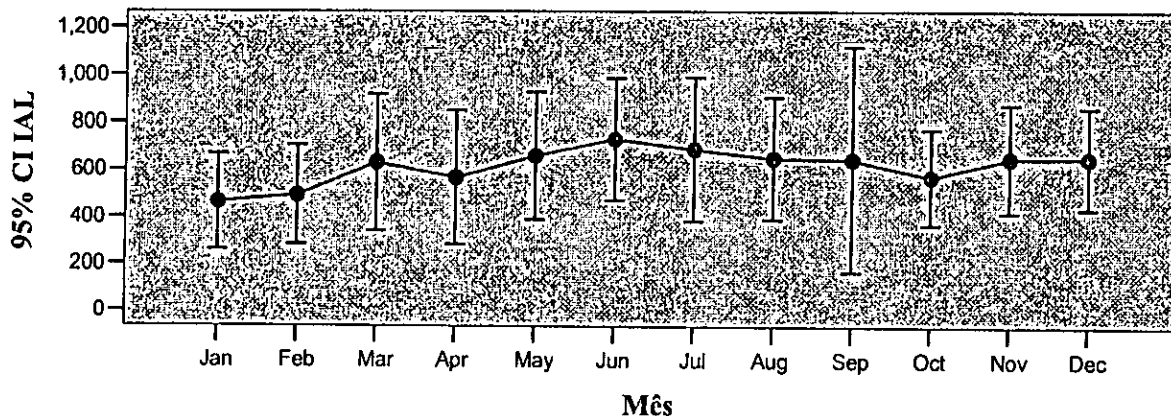


Figura 2. Gráfico de barra do erro de IAL e os respectivos intervalos de confiança

Em relação aos pressupostos para aplicação do teste F, o teste de Levene apresentado na Tabela 3, permite afirmar que não se rejeita a hipótese nula da homocedasticidade, ou seja, as diferenças de dispersões (desvios padrões) observadas nas importações consoante os meses, não são consideradas estatisticamente significativas.

Tabela 3. Teste de homogeneidade das variâncias

IAL			
Levene Statistic	df1	df2	Sig.
.395	11	66	.954

A Tabela 4 infere a normalidade através dos testes Kolmogorov-Smirnov ou Shapiro Wilk, onde não se rejeita a hipótese nula da normalidade, podendo se afirmar que a distribuição das importações dentro de cada mês é normal.

Tabela 4. Teste de normalidade nas importações mensais

Mes	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
IAL						
Jan	.125	7	.200*	.981	7	.966
Feb	.225	7	.200*	.939	7	.629
Mar	.207	7	.200*	.948	7	.712
Apr	.239	7	.200*	.906	7	.366
May	.215	7	.200*	.912	7	.413
Jun	.191	7	.200*	.939	7	.626
Jul	.256	6	.200*	.939	6	.654
Aug	.233	6	.200*	.903	6	.393
Sep	.290	6	.125	.824	6	.096
Oct	.268	6	.200*	.915	6	.469
Nov	.254	6	.200*	.886	6	.297
Dec	.239	6	.200*	.876	6	.252

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

O teste F da Tabela 5 dá a confirmação da não existência de diferenças significativas entre as médias mensais de IAL, dado que não se rejeita a hipótese nula de igualdade das médias entre os meses.

Tabela 5. Teste de igualdade das médias mensais de IAL

IAL					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	455589.6	11	41417.233	.536	.872
Within Groups	5099409	66	77263.773		
Total	5554999	77			

Com base nas análises feitas e tomando em conta a representação da Figura 3, pode-se concluir que a série de IAL não apresenta a componente sazonal, uma vez que não foram detectadas diferenças significativas entre as importações médias mensais e a função de autocorrelação não apresenta padrão de uma série sazonal.

4.2. Testes de estacionaridade

Para analisar a estacionaridade da série de IAL, recorreremos primeiramente a representação gráfica da Figura 1. Duma forma geral, pode-se notar que a série parece ter uma tendência⁴ crescente, o que dá indícios de se estar perante uma série não estacionária em nível. O mesmo se pode afirmar examinando o correlograma das funções de autocorrelação e autocorrelação parcial onde o coeficiente de autocorrelação começa com um valor muito alto e tende muito lentamente a zero à medida que as defasagens aumentam (Figura 3). Mas como não é suficiente identificar se a série é ou não estacionária apenas na análise da figura 1 e as funções de Autocorrelação (ACF) e Autocorrelação parcial (PACF), ela é submetida a um teste de raiz unitária para a devida confirmação.

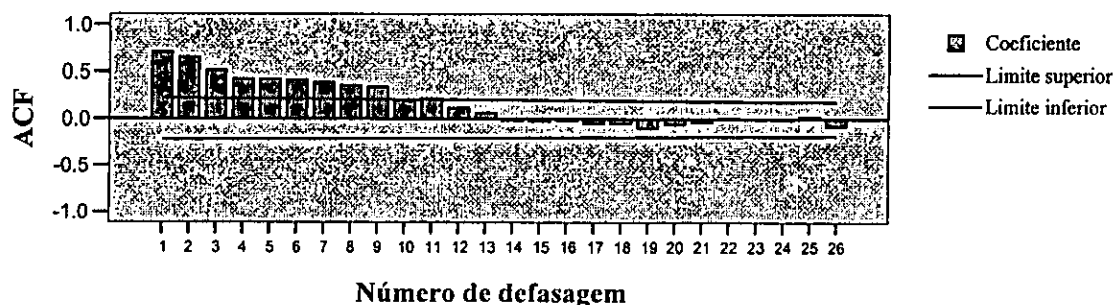


Figura 3. Correlograma da função de autocorrelação da série IAL em nível

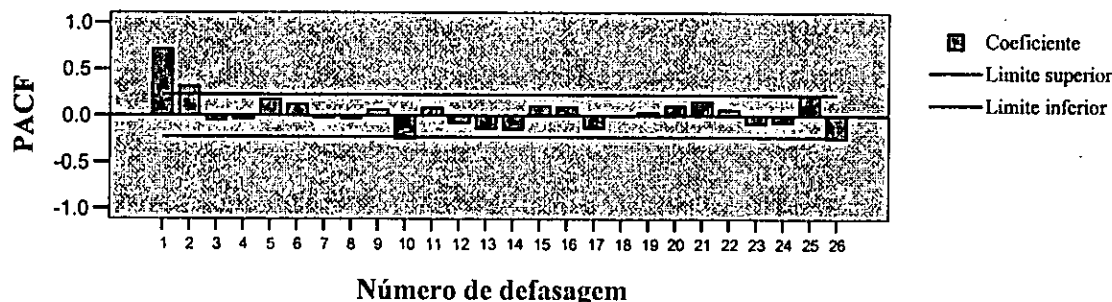


Figura 4. Correlograma da função de autocorrelação parcial da série IAL em nível

⁴ A linha da tendência (crescente) pode ser observada na figura 1 em anexo A

4.2.1. Testes de raiz unitária

Tendo se constatado que a série parece não estacionária em nível, utilizou-se um teste de raiz unitária de Dickey-Fuller Aumentado para a confrontação da hipótese.

A Tabela 6 apresenta a estatística de teste de raiz unitária para regressão estimada na Tabela 6.1. O valor da estatística de ADF é (-3.283628) e associada a probabilidade de 0.0768. Comparando-o com os valores críticos da estatística sugerida por MacKinnon (1996), a níveis de 1% e 5% de significância, este em termo absoluto é menor, pelo que não se rejeita a hipótese nula de que a regressão estimada possui raiz unitária, podendo se concluir que, de facto a série de IAL não é estacionária em nível.

Tabela 6. Estatística de teste de raiz unitária de ADF

	Estatística t	Prob.*
Estatística do teste de Dickey-Fuller Aumentado	-3.283628	0.0768
1% level	-4.083355	
5% level	-3.470032	
10% level	-3.161982	

*MacKinnon (1996)

Tabela 6.1. Estatística de teste para Regressão de ADF

Variável Dependente: D (IAL)

Variável	Coefficiente	Erro Padrão	Estatística t	Prob.
IAL(-1)	-0.394545	0.120155	-3.283628	0.0016
D(IAL(-1))	-0.293757	0.114511	-2.565320	0.0124
C	123.5311	50.72549	2.435286	0.0174
TREND	3.384611	1.316571	2.570778	0.0122

Confirmado a não estacionaridade da série, efectuou-se a transformação da mesma em primeira diferença. Conforme mostra a representação gráfica da Figura 5, durante o período em análise a série se estabiliza em torno de um valor constante zero e com base nos testes de raiz unitária apresentados nas Tabelas 7 e 7.1, uma vez que o valor absoluto calculado da estatística de ADF é maior que os absolutos críticos de MacKinnon para todos níveis de significância, rejeita-se a hipótese nula de que a regressão estimada possui raiz unitária e conclui-se que a série de IAL é estacionária em primeira diferença, ou seja, é integrada de ordem 1, I(1).

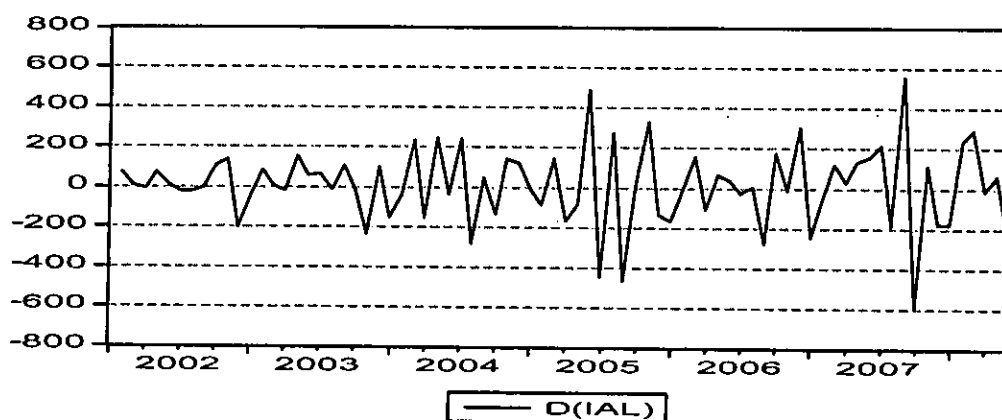


Figura 5. Representação gráfica da série de IAL em primeira diferença

Tabela 7. Estatística de teste de raiz unitária de ADF para a série de IAL em primeira diferença

	Estatística t	Prob.*
Estatística do teste de Dickey-Fuller Aumentado	-14.55207	0.0001
Teste de valores críticos:		
1% level	-4.083355	
5% level	-3.470032	
10% level	-3.161982	

*MacKinnon (1996)

Tabela 7.1. Estatística de teste para Regressão de ADF para a série de IAL em primeira diferença
Variável Dependente: D(IAL,2)

Variável	Coefficiente	Erro Padrão	Estatística t	Prob.
D(IAL(-1))	-1.495993	0.102803	-14.55207	0.0000
C	14.33152	40.78883	0.351359	0.7263
TREND	0.076950	0.902757	0.085239	0.9323

4.3. Identificação dos parâmetros do modelo

Após ser garantida a condição básica para aplicação da metodologia de Box-Jenkins, segue-se o passo de identificação dos valores de p , d e q . Para tornar a série de IAL estacionária, foi necessário processar uma transformação de primeira diferença, ou seja, a série foi diferenciada uma vez, portanto, $d=1$. Da Figura 6 permite visualizar que as autocorrelações com picos nas defasagens 1, 25, e 26 parecem estatisticamente significativas, uma vez que os seus respectivos valores estão fora de intervalo de 95% de confiança.

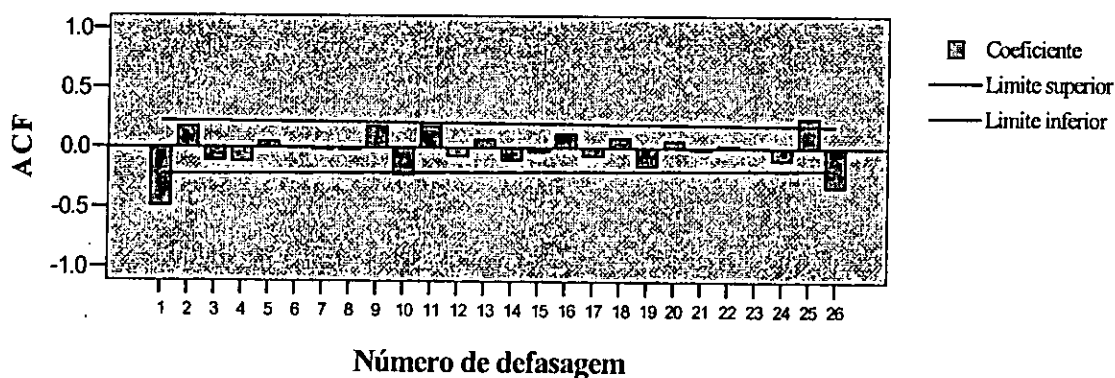


Figura 6. Correlograma da função de autocorrelação da série IAL em primeira diferença

A função autocorrelação parcial (PACF) decai exponencialmente com padrão senoidal amortecido. Portanto, pode se propor um processo que seja no máximo MA(26), mas não sendo necessário a inclusão de todos os termos MA até 26, pois a partir do correlograma de ACF sabemos que apenas os termos MA das defasagens 1, 25, 26 são significativas.

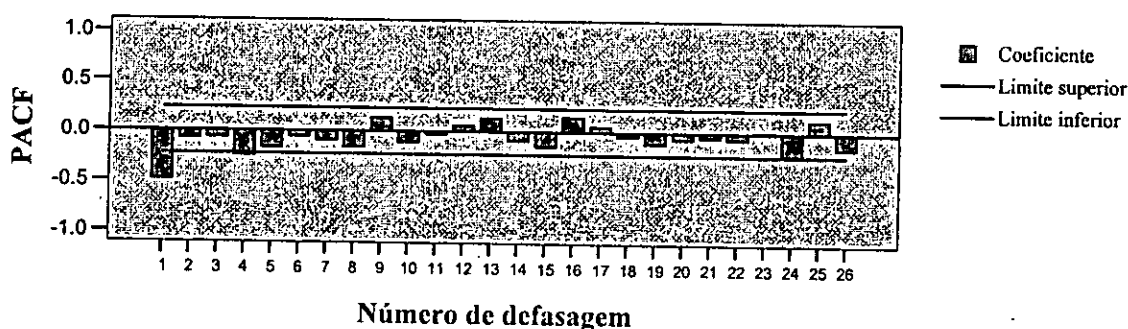


Figura 7. Correlograma da função de autocorrelação parcial da série IAL em primeira diferença

Como já foi referido no capítulo 2, o ponto importante na identificação do modelo ARIMA é parcimônia, ou seja, deve-se buscar o modelo com melhor ajuste e melhor número possível de parâmetros. Sendo assim, para além do modelo contendo os termos MA(1), MA(25) e MA(26), será estimado um modelo que contenha apenas o termo MA(1) e outro com os termos MA(1) e MA(25) para comparação dos seus resultados. Uma vez que não existem regras fixas estabelecidas para escolha dos parâmetro, sendo apenas feita com auxílio das funções de ACF e PACF.

4.4. Estimação dos modelos

As Tabelas 1C, 2C e 3C em anexo, mostram a estimação dos parâmetros dos modelos ARIMA(0,1,26), ARIMA(0,1,25) e ARIMA(0,1,1) incluindo os seus respectivos interceptos. Nota-se que a um nível de 5% de significância não se rejeitam as hipóteses nulas de que os verdadeiros interceptos são iguais a zero, o que justificaria a estimação dos mesmos sem os termos constante.

A seguir são apresentadas as estimções dos modelos acima referidos sem inclusão dos interceptos. Como se pode ver a partir das Tabelas 8, 9 e 10, os coeficientes de médias móveis estimados em todos modelos são estatisticamente significativos a qualquer nível de significância. Portanto, os três modelos serão submetidos na etapa de verificação de diagnóstico.

Tabela 8. Estimação do modelo ARIMA(0,1,26) sem o termo intercepto

Variável Dependente D(IAL)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
MA(1)	-0.404713	0.074874	-5.405240	0.0000
MA(25)	0.374328	0.065129	5.747514	0.0000
MA(26)	-0.850376	0.028858	-29.46770	0.0000
R-quadrado	0.245611	Variância média dependente		11.14286
R-quadrado ajustado	0.245611	S.D. Variância dependente		194.4036
S.E. da regressão	168.8503	Critério de informação de Akaike		13.10880
S. dos quadrados dos erros	2166792.	Critério de Schwarz		13.13924
Log máxima-verossimilhança	-503.6890	Estatística de Durbin-Watson		2.023751

Tabela 9. Estimação do modelo ARIMA(0,1,25) sem intercepto

Variável Dependente: D(IAL)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
MA(1)	-0.555450	0.090679	-6.125486	0.0000
MA(25)	0.406295	0.082430	4.928950	0.0000
R-quadrado	0.334692	Variância média dependente		11.14286
R-quadrado ajustado	0.325821	S.D. Variância dependente		194.4036
S.E. da regressão	159.6217	Critério de informação de Akaike		13.00912
S. dos quadrados dos erros	1910932.	Critério de Schwarz		13.07000
Log máxima-verossimilhança	-498.8512	Estatística de Durbin-Watson		2.099704

Tabela 10. Estimação do modelo ARIMA(0,1,1) sem intercepto

Variável Dependente: D(IAL)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
MA(1)	-0.581461	0.093119	-6.244274	0.0000
R-quadrado	0.535166	Variância média dependente		11.14286
R-quadrado ajustado	0.522603	S.D. Variância dependente		194.4036
S.E. da regressão	134.3211	Critério de informação de Akaike		12.67652
S. dos quadrados dos erros	1335119.	Critério de Schwarz		12.76784
Log máxima-verossimilhança	-485.0462	Estatística de Durbin-Watson		1.979963

4.5. Verificação de diagnóstico

Depois de identificados e estimados os parâmetros dos três modelos, procede-se a verificação da sua adequação antes de utilizá-los para fazer previsões. Nesta etapa procurar-se-á verificar se os resíduos dos modelos estimados são ou não independentes com média igual a zero e variância constante, ou seja, são ruídos brancos.

Analisando as Figuras 8, 9 e 10, pode se afirmar que a série dos resíduos gerada por modelo ARIMA(0,1,26) se estabiliza em torno de um valor médio constante zero e que as funções de autocorrelação e autocorrelação parcial dos resíduos não revelam a existência de dependência serial, ou seja, elas não são significante em qualquer defasagem

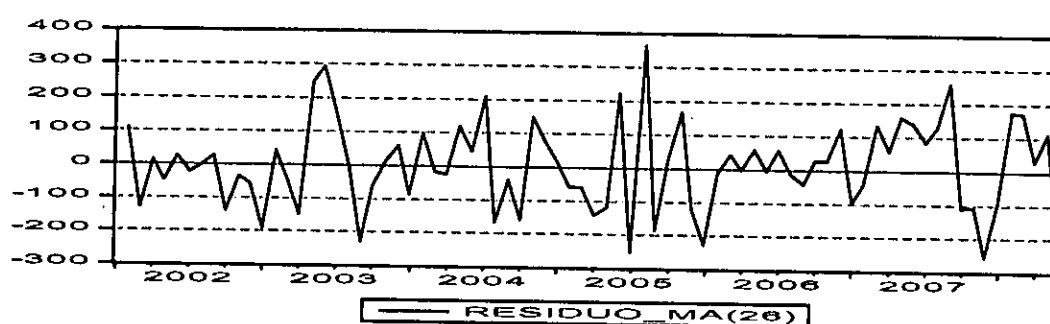


Figura 8. Representação gráfica dos resíduos obtidos do modelo ARIMA(0,1,26)

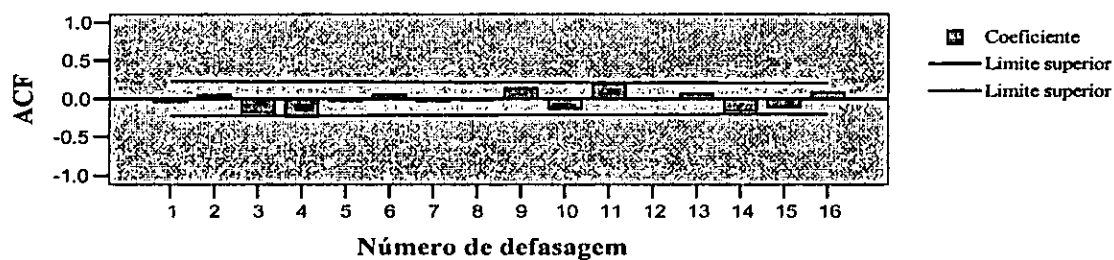


Figura 9. Correlograma da função de autocorrelação dos resíduos obtidos do modelo ARIMA(0,1,26)

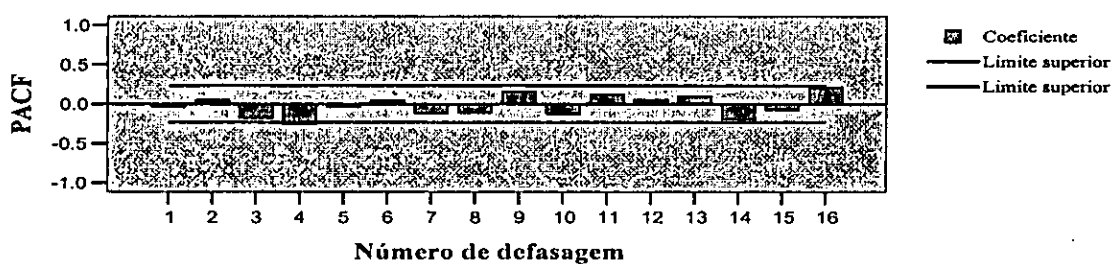


Figura 10. Correlograma da função de autocorrelação parcial dos resíduos obtidos do modelo ARIMA(0,1,26)

O teste de Ljung e Box apresentado na Tabela 11 não rejeita a hipótese nula das 16 primeiras autocorrelações dos resíduos não diferem estatisticamente de zero.

Tabela 11. Estatísticas de teste de Box-Ljung para resíduos do modelo ARIMA(0,1,26)

Series: RESIDUO MA(26)

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-.033	.114	.087	1	.768
2	.049	.114	.279	2	.870
3	-.175	.114	2.798	3	.424
4	-.230	.118	7.213	4	.125
5	-.023	.123	7.258	5	.202
6	.052	.124	7.493	6	.278
7	-.027	.124	7.556	7	.373
8	-.015	.124	7.576	8	.476
9	.132	.124	9.137	9	.425
10	-.130	.128	10.672	10	.384
11	.187	.127	13.888	11	.239
12	-.007	.131	13.892	12	.308
13	.066	.131	14.300	13	.353
14	-.160	.131	16.783	14	.268
15	-.108	.134	17.920	15	.267
16	.082	.135	18.586	16	.291

a. The underlying process assumed is MA with the order equal to the lag number minus one. The Bartlett approximation is used.

b. Based on the asymptotic chi-square approximation.

Os resultados dos testes mostrados pelas Tabelas 12,13 e 14 abaixo confirmam que a média dos resíduos não difere estatisticamente de zero e a variância dos mesmos não varia consoante o mês, isto é, é constante. O teste de ARCH também não rejeita a hipótese nula da insignificância do modelo estimado, ou seja, os resíduos não apresentam o efeito ARCH.

Tabela 12. Teste de hipótese da média dos resíduos de ARIMA(0,1,26) ser igual a zero

Teste de Hipotese Media = 0						
	t	df	Sig. (2-tailed)	Diferença média	Intervalo de 95% de confiança da diferença	
					Lower	Upper
RESIDUO_MA(26)	.426	76	.671	6.4239501	-23.6235	36.471449

Tabela 13. Teste de igualdade de variâncias dos resíduos ARIMA(0,1,26)

	Levene Statistic	df1	df2	Sig.
RESIDUO_MA(26)	.732	11	65	.704

Tabela 14. Teste ARCH dos resíduos do ARIMA(0,1,26)

ARCH Test:

F-statistic	0.776930	Probability	0.735596
Obs*R-squared	23.30781	Probability	0.615481

A normalidade dos resíduos é testada a partir da Figura 11 e Tabela 15 abaixo. O valor da estatística de Kolmogorov-Smirnov é de 0.058 associado a uma probabilidade de 0.2, pelo que não se rejeita a hipótese nula da normalidade e conclui-se que a série dos resíduos está normalmente distribuída.

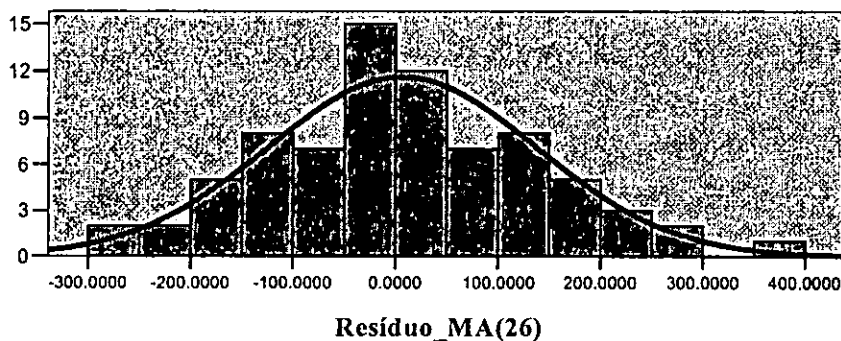


Figura 11. Histograma dos resíduos obtidos do modelo ARIMA(0,1,26)

Tabela 15. Teste de normalidade para os resíduos do modelo ARIMA(0,1,26)

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
RESIDUO_MA(26)	.058	77	.200*	.991	77	.860

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Portanto, as análises feitas revelam que os resíduos estimados a partir do modelo ARIMA(0,1,26), estão normalmente distribuídos com média zero e variância constante, ou seja, são ruído branco.

A seguir foram também verificados os diagnósticos dos resíduos obtidos a partir dos modelos ARIMA(0,1,25) e ARIMA(0,1,1) tendo se chegado as mesmas conclusões obtidas para o modelo ARIMA(0,1,26), isto quer dizer que os resíduos destes modelos também são normalmente distribuídos com média zero e variância constante, conforme se pode confirmar a partir dos resultados abaixo apresentados.

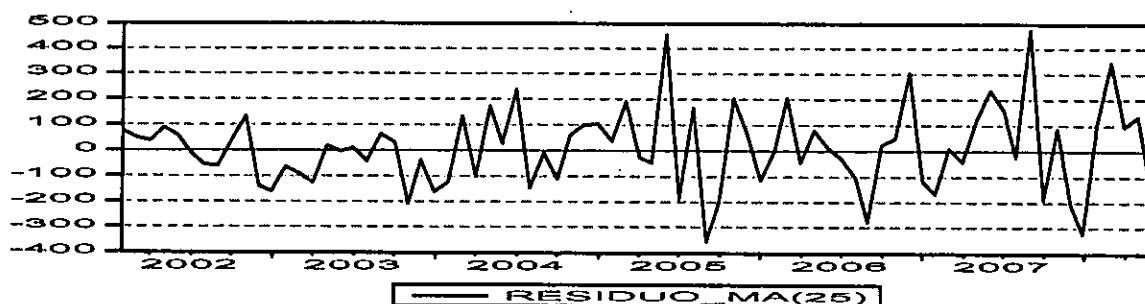


Figura 12. Representação gráfica dos resíduos obtidos do modelo ARIMA(0,1,25)

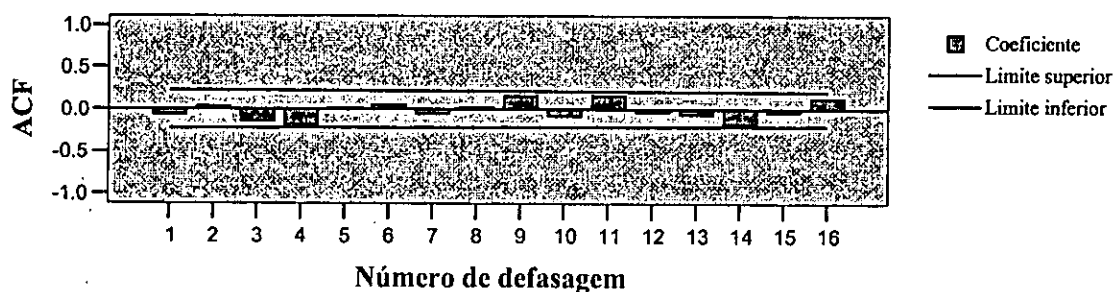


Figura 13. Correlograma da função de autocorrelação dos resíduos obtidos do modelo ARIMA(0,1,25)

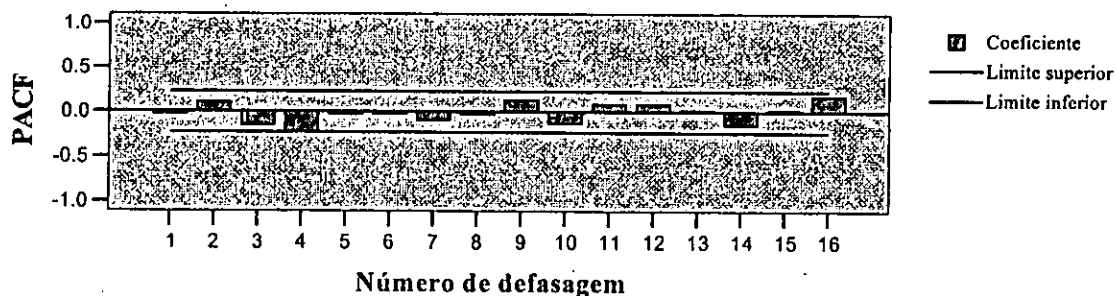


Figura 14. Correlograma da função de autocorrelação dos resíduos obtidos do modelo ARIMA(0,1,25)

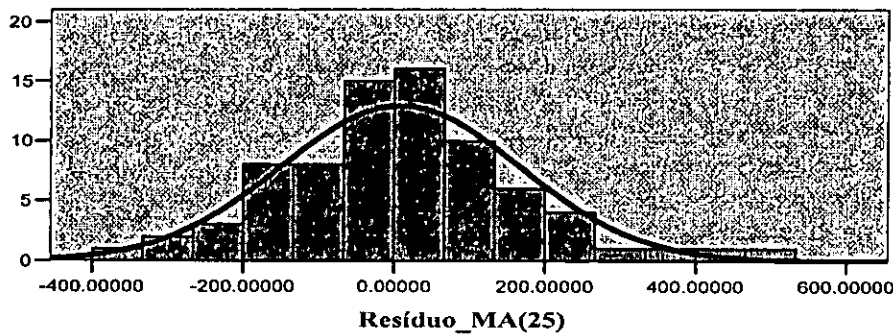


Figura 15. Histograma dos resíduos obtidos do modelo ARIMA(0,1,25)

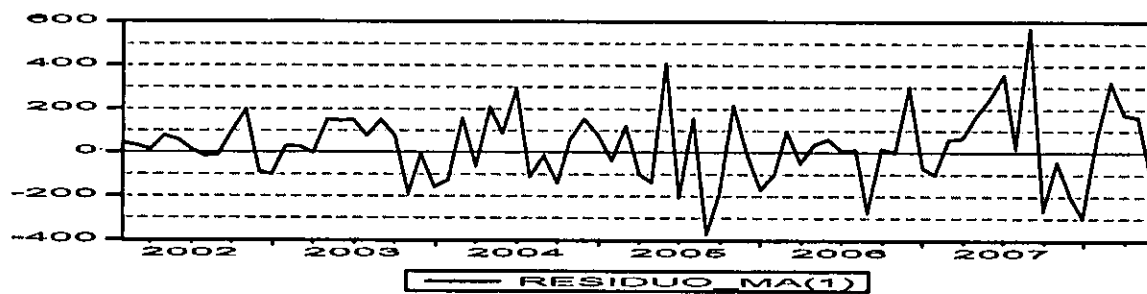


Figura 16. Representação gráfica dos resíduos obtidos do modelo ARIMA(0,1,1)

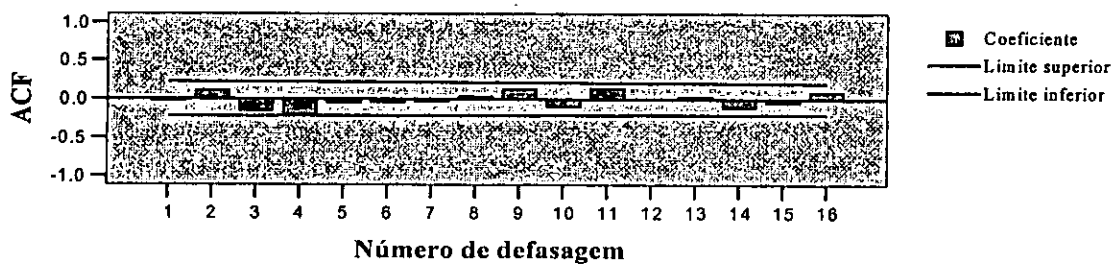


Figura 17. Correlograma da função de autocorrelação dos resíduos obtidos do modelo ARIMA(0,1,1)

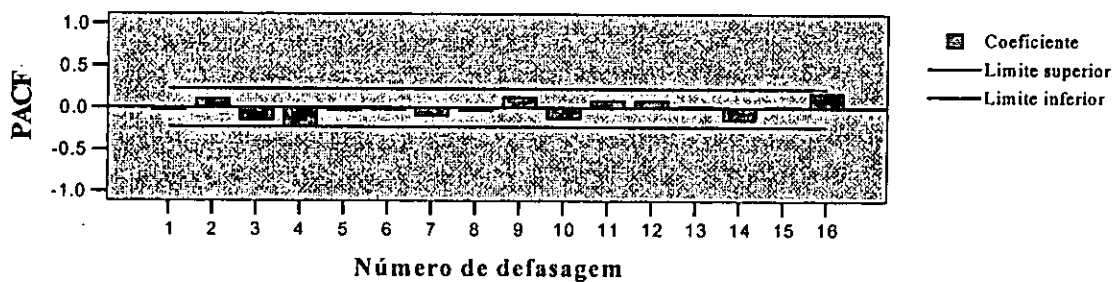


Figura 18. Correlograma da função de autocorrelação parcial dos resíduos obtidos do modelo ARIMA(0,1,1)

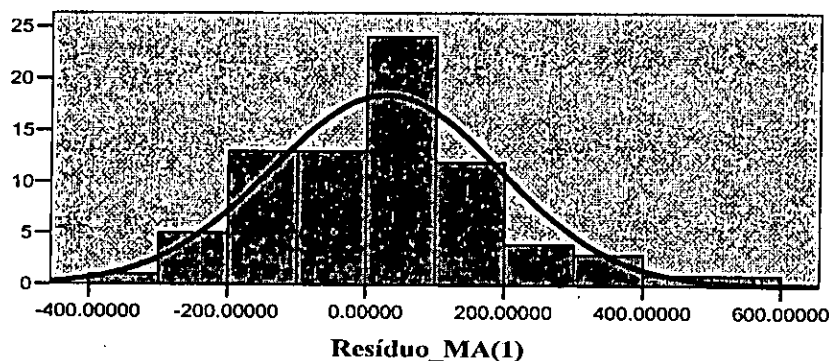


Figura 19. Histograma dos resíduos obtidos do modelo ARIMA(0,1,1)

Com estes resultados conclui-se que os três modelos obtidos podem ser adequados para descrever a série de IAL.

4.5.1. Seleção do modelo

O diagnóstico feito revela a adequação dos três modelos estimados à série das importações. Contudo, para fins de previsão, apenas um modelo será escolhido com base nos critérios de seleção. A Tabela 16 apresenta as estatísticas do desempenho da previsão dentro da amostra para cada um dos modelos.

Tabela 16. Estatísticas do desempenho nas previsões dos modelos estimados

Modelo	Estatística de desempenho					
	REQM	EAM	EAMP	CDT	AIC	SIC
ARIMA(0,1,1)	131.6784	103.9263	18.0880	0.0985	12.6765	12.7678
ARIMA(0,1,25)	157.5351	120.9157	21.01927	0.1190	13.0091	13.0700
ARIMA(0,1,26)	167.7503	127.2951	20.6438	0.1282	13.1088	13.1392

As estatísticas do modelo ARIMA(0,1,1) mostram valores menores em relação às dos outros, sugerindo assim, melhor desempenho. Por isso, este modelo foi usado para a previsão das importações de automóveis ligeiros nos próximos meses.

4.6. Previsão

Uma das razões para popularidade da metodologia de Box & Jenkins é o sucesso na previsão (Gujarati, 2006).

Com base no modelo ARIMA(0,1,1) procedeu-se a previsão das importações de automóveis ligeiros para o período de Julho de 2008 à Março de 2009, como mostra a Tabela 17, onde são apresentados valores das importações com seus respectivos limites inferiores e superiores de intervalos de 95% de confiança.

Tabela 17. Previsão das importações de automóveis ligeiros para o período de Julho de 2008 à Março de 2009

Mês	Limite inferior	Previsão	Limite superior
Julho/2008	744.91799	1010	1417.92629
Agosto/2008	715.01311	1038	1447.83117
Setembro/2008	687.37125	1066	1475.47303
Outubro/2008	661.54521	1081	1501.29907
Novembro/2008	637.21817	1081	1525.62611
Dezembro/2008	614.15595	1081	1548.68833
Janeiro/2009	592.17964	1177	1570.66464
Fevereiro/2009	571.14892	1135	1591.69536
Marco/2009	550.95132	1182	1611.89296

5. CONCLUSÕES

O trabalho teve como objectivo desenvolver um modelo de previsão sobre a série das importações de automóveis ligeiros ao nível da cidade e província de Maputo. A utilização da metodologia Box & Jenkins possibilitou a identificação da estrutura do modelo em causa.

Ao longo do período, as importações de automóveis registaram um crescimento médio anual significativo. Em média, o número das importações é mais alto em Junho e mais baixo em Janeiro.

Dentre os modelos identificados, o ARIMA(0,1,1) evidenciou melhor desempenho estatístico, apresentando melhores resultados de previsão, tanto dentro como fora da amostra. Por isso, a metodologia de Box & Jenkins pode ser aplicada para descrever a serie a das importações de viaturas nos próximos períodos.

As previsões revelam um crescimento gradual das importações de viaturas para o período de Julho de 2008 à Março de 2009, com estabilidade nos meses de Outubro, Novembro e Dezembro de 2008 em torno de 1081 viaturas. A tendência do aumento de viaturas foi também constatada pela USAID(2006).

5.1. Recomendações

A metodologia sugerida pelo Box & Jenkins revela uma adequação à série das importações de automóveis ligeiros. Portanto, é recomendada a sua aplicação para análise das importações de automóveis pesados, principalmente para os mini-buses. Segundo a USAID(2006), estes autocarros são considerados os transportes públicos que presta serviço à maioria da população ao nível das duas cidades.

Para além dos modelos univariados analisados, propõe-se que sejam feitas as análises multivariadas para inclusão das variáveis que influenciam na importação de viatura. Por exemplo, taxas aduaneiras, preços de combustíveis.

Algumas das formas para fazer face ao aumento de viaturas, recomenda-se a construção de mais parques de estacionamento comuns e a implementação de medidas que incentivam o uso do transporte público como alternativa do transporte individual

5.2. Limitações do estudo

As principais limitações encontradas durante a realização do estudo foram:

- Dificuldades no acesso as fontes bibliográficas primárias o que levou a que se recorresse constantemente à fontes bibliográficas secundárias;
- A insuficiência da informação dos dados sobre as importações de automóveis pesados, fez com que o estudo analisasse apenas a serie de importações de automóveis ligeiros.
- Dificuldade na obtenção de informação adicional de interesse, por parte das ATM.

6. REFERÊNCIAS BIBLIOGRÁFICAS

Box, G.E.P., Jenkins, G (1976); *Time Series Analysis: Forecasting and Control*, Holden-Day.

Brockwell, P.J. & Davis, R.A. (1996). *Introduction to Time Series and Forecasting*. New York. Springer-Verlag.

Bruno, M.(2001). *Modelo de previsão para arrecadação tributária*. Escola de Administração fazendária-ESAF.Brasil

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. pp 427-431.

Dickey, D.A. e Fuller W.A. (1981). *Likelihood ratio statistics for autoregressive time series with a unit root*. Econometrica, 49. pp 1057-1072.

Ehlers, S.R.(2005), *Análise de séries temporais*. 3ª edição. Departamento de Estatística. UFPR. Curitiba. Brasil

Enders, W. (1995). *Applied Econometric Time Series*. New York: John Wiley & Sons.

Fava, V. L (2000). *Manual de econometria*. IN: VASCONCELOS, M. A. S.; ALVES, D. São Paulo: Editora Atlas.

Fonseca, J.S.; Martins, G.A.; Toledo, G. L (1995). *Estatística Aplicada*. São Paulo:Atlas.

Fuller, W.A. (1976). *Introduction to Statistical Time Series*. New York. John Wiley.

Granger, C.W.J. & Newbold, P. (1974). *Experience with forecasting univariate time series and the combination of forecasts*. Journal of the Royal Statistics Society A, 137, 131-146.

Granger, C.W.J. & Newbold, P. (1986). *Forecasting Economic Time Series*. New York: Academic Press.

Gujarati, D.N (2006); *Econometria básica*, Editora Campus, 4ª edição (tradução); São Paulo.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press.

- Hill, R.C., Griffiths, W.E. & Judge, G.G. (1999). *Econometria*. São Paulo: Editora Saraiva.
- Issler, J.V (1999). *Estimating and Forecasting the Volatility of Brazilian Finance series using ARCH models*. Rio de Janeiro. 49 pp. Graduate School of Economics – EPGE.
- Martins, P. G.; Laugeni, F. P (1999); *Administração da Produção*. Saraiva. São Paulo.
- MacKinnon, J.G. (1991). *Critical values for co-integration tests* em Engle, R.F. e C.W.J. Granger (eds.), *Long-Run Economic Relationships*. pp 267-276. Oxford. Oxford University Press.
- MacKinnon, J.G. (1996). *Numerical distribution functions for unit root and cointegration tests*. Journal of Applied Econometrics, 11. pp 601-618.
- Makridakis, S.; Wheelwright, S.; Hyndman, R. J (1983). *Forecasting methods and applications*.3. ed. New York: John Wiley & Sons.
- Masih, A.M.M & R.Masih (2002), *propagative causal price transmission among international stock markets: evidence from the pre-end post globalization*.Global Finance Journal 13, 63-91.
- Mills, T.C. (1990). *Time Series Techniques for Economists*. Cambridge:Cambridge University Press.
- Mentzer, J.T.; Bienstock, C.C. (1998). *Sales Forecasting Management*. Thousand Oaks: Sage.
- Moretlin, P.A; Toloí, C.M (1987). *Séries Temporais*. 2.ed. (Actual).São Paulo.
- Morretin, P.A; Toloí, C.M (1995). *Previsão de Séries Temporais*.: Actual Editora. São Paulo
- Murteira, B.J.F. MAuller, D.A. e Turkman K.F (1993) *Análise de Sucessões Cronológicas*. McGraw Hill, Lisboa.
- Parzen, E.(1972), *Procesos Estocásticos*. Departamento de Estatística. Universidade de Stanford. Madrid. Espanha

Paula, A.V.(1999) *Previsão da População Brasileira através dos modelos ARIMA de Box e Jenkins*. Monografia final do curso de Especialização em Métodos estatísticos Computacionais da Universidade Federal de Juiz de Fora. Brasil.

Pestana, M. H; Gageiro, J. N (2005). *Análise de Dados para Ciências Sociais: A complementaridade do SPSS*. 4ª Edição. Edições Sílabo. Lisboa. Portugal.

Pires, A.P (2001), *Notas de Séries Temporais*. Instituto superior Técnico LMAC. Lisboa.

Ribeiro, L.C; Paula, A.V (2000), *Previsão de população através dos modelos ARIMA de Box e Jenkins: Um exercício para Brasil*. Universidade Federal de Minas Gerais. Caxambu. Brasil
Disponível em:

www.abep.nepo.unicamp.br/anais_2000.htm

Silva, M. L.; Silva, J. M. A (1996); *Análise do comportamento temporal dos preços do carvão vegetal: aplicação do modelo e a avaliação da metodologia "Box and Jenkins"*. Revista Arvore, Vol.20.

Slini, L.; Karatzas, K.; Moussiopoulos, N. (2001). *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.

Spiegel, M.R (1985), *Estatística*. McGraw-Hill. Brasil.

Tran, N. N. (2002). *Automatic arima time series modeling and forecasting for adaptive input/output prefetching*. University of Illinois.

UEM. (1994); *Regulamento do trabalho de licenciatura para cursos da faculdade de ciências*; Maputo.

ANEXOS

ANEXO A: Dados de importações de automóveis ligeiros

Tabela 1A. Dados mensais de importações de automóveis ligeiros na província e cidade de Maputo

Meses	2002	2003	2004	2005	2006	2007	2008
Jan	153	248	396	772	482	548	655
Feb	223	332	359	686	479	487	887
Mar	229	343	592	829	636	604	1177
Apr	224	324	435	659	529	633	1164
May	293	478	680	576	592	765	1226
Jun	305	536	646	1066	630	917	1011
Jul	282	600	889	622	603	1129	
Aug	260	587	605	897	610	940	
Sep	257	693	653	432	329	1499	
Oct	362	683	521	454	503	901	
Nov	500	449	660	783	493	1011	
Dec	295	549	781	650	794	834	
Médias	281.9167	485.1667	601.4167	702.1667	556.6667	855.6667	1020

Fonte: Autoridade Tributária de Moçambique

ANEXO B: Análise exploratória de dados

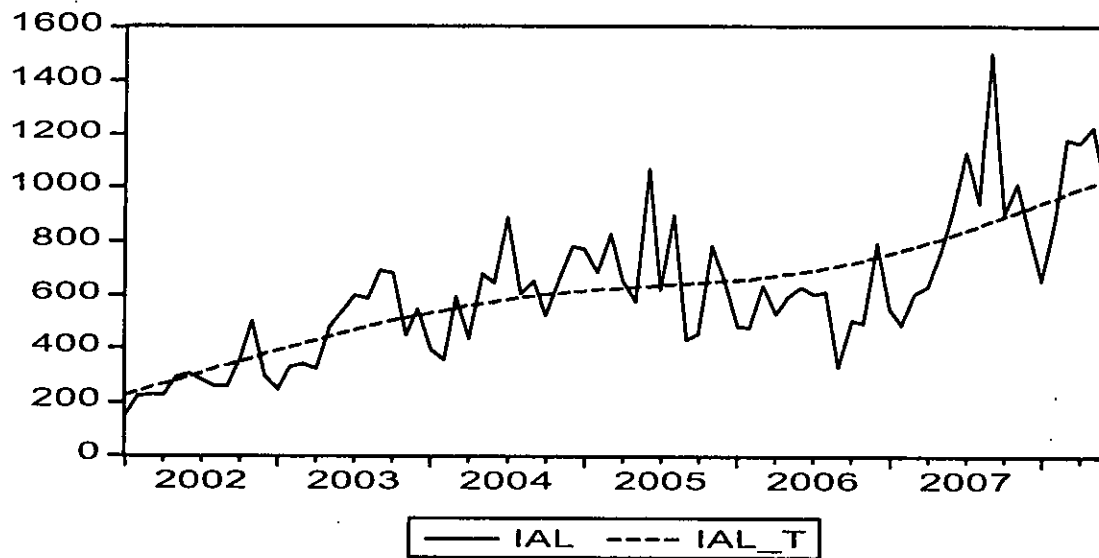


Figura 1B. Representação gráfica da série das importações de automóveis ligeiros e a sua tendência

Tabela 1B. Número médio mensal de IAL e seus respectivos intervalos de confiança

IAL		95% CI para media das Importacoes						
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
Jan	7	464.86	218.602	82.624	262.68	667.03	153	772
Feb	7	493.29	227.117	85.842	283.24	703.33	223	887
Mar	7	630.00	312.019	117.932	341.43	918.57	229	1177
Apr	7	566.86	306.810	115.963	283.11	850.61	224	1164
May	7	658.57	291.907	110.331	388.60	928.54	293	1226
Jun	7	730.14	277.532	104.897	473.47	986.82	305	1066
Jul	6	687.50	289.497	118.187	383.69	991.31	282	1129
Aug	6	649.83	246.902	100.797	390.73	908.94	260	940
Sep	6	643.83	453.309	185.063	168.11	1119.55	257	1499
Oct	6	570.67	192.841	78.727	368.29	773.04	362	901
Nov	6	649.33	217.098	88.630	421.50	877.16	449	1011
Dec	6	650.50	203.950	83.262	436.47	864.53	295	834
Total	78	614.31	268.594	30.412	553.75	674.87	153	1499

ANEXO C: Estimação dos modelos com interceptos

Tabela 1C. Estimação do modelo ARIMA(0,1,26) com intercepto

Variável Dependente: D(IAL)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	7.272517	9.509091	0.764796	0.4469
MA(1)	-0.414178	0.073775	-5.614082	0.0000
MA(25)	0.385632	0.064096	6.016497	0.0000
MA(26)	-0.847487	0.029641	-28.59155	0.0000

Tabela 2C. Estimação do modelo ARIMA(0,1,25) com intercepto

Variável Dependente: D(IAL)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	11.66764	16.03455	0.727657	0.4691
MA(1)	-0.566065	0.089492	-6.325316	0.0000
MA(25)	0.403328	0.083037	4.857184	0.0000

Tabela 3C. Estimação do modelo ARIMA(0,1,1) com intercepto

Variável Dependente; D(IAL)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	11.60770	6.852465	1.693945	0.0944
MA(1)	-0.650453	0.087650	-7.421043	0.0000

ANEXO D: Análise residual dos modelos ARIMA(0,1,25) e ARIMA(0,1,1)

Tabela 1D. Estatísticas de teste de Box-Ljung para resíduos do modelo ARIMA(0,1,25)

Series: RESIDUO MA(25)

Lag	Autocorrelation	Std.Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-.063	.112	.313	1	.576
2	.034	.111	.409	2	.815
3	-.140	.110	2.018	3	.569
4	-.185	.110	4.854	4	.303
5	.012	.109	4.866	5	.432
6	.052	.108	5.095	6	.532
7	-.054	.107	5.352	7	.617
8	.019	.107	5.383	8	.716
9	.153	.106	7.472	9	.588
10	-.079	.105	8.037	10	.625
11	.157	.104	10.314	11	.502
12	-.030	.103	10.396	12	.581
13	-.061	.103	10.750	13	.632
14	-.172	.102	13.613	14	.479
15	-.034	.101	13.729	15	.546
16	.126	.100	15.314	16	.502

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

b. Based on the asymptotic chi-square approximation.

Tabela 2D. Estatísticas de teste de Box-Ljung para resíduos do modelo ARIMA(0,1,1)

Series: RESIDUO MA(1)

Lag	Autocorrelation	Std.Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-.023	.112	.043	1	.835
2	.104	.111	.916	2	.633
3	-.151	.110	2.788	3	.425
4	-.196	.110	6.002	4	.199
5	-.048	.109	6.194	5	.288
6	-.039	.108	6.323	6	.388
7	-.028	.107	6.389	7	.495
8	.028	.107	6.456	8	.596
9	.128	.106	7.923	9	.542
10	-.093	.105	8.705	10	.560
11	.136	.104	10.410	11	.494
12	.003	.103	10.411	12	.580
13	.024	.103	10.464	13	.656
14	-.109	.102	11.616	14	.637
15	-.044	.101	11.808	15	.694
16	.085	.100	12.525	16	.707

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

b. Based on the asymptotic chi-square approximation.

Tabela 3D. Teste de hipótese da média dos resíduos de ARIMA(0,1,25) e ARIMA(0,1,1) ser igual a zero

Teste de Hipotese: Media = 0						
	t	df	Sig. (2-tailed)	Diferença media	Intervalo de 95% de confiança da media	
					Lower	Upper
RESIDUO_MA(25)	.473	76	.638	8.5280292	-27.4097	44.46579
RESIDUO_MA(1)	1.469	76	.146	27.880	-9.91	65.67

Tabela 4D. Teste de normalidade para os resíduos dos modelos ARIMA(0,1,25) e ARIMA(0,1,1)

	Kolmogorov-Smimov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
RESIDUO_MA(25)	.065	77	.200*	.981	77	.297
RESIDUO_MA(1)	.067	77	.200*	.985	77	.514

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Tabela 5D. Teste de igualdade de variâncias dos resíduos dos modelos ARIMA(0,1,25) e ARIMA(0,1,1)

	Levene			
	Statistic	df1	df2	Sig.
RESIDUO_MA(25)	1.358	11	65	.214
RESIDUO_MA(1)	1.805	11	65	.071

Tabela 6D. Teste ARCH dos resíduos do ARIMA(0,1,25)

ARCH Test:			
F-statistic	0.776930	Probability	0.735596
Obs*R-squared	23.30781	Probability	0.615481

Tabela 7D. Teste ARCH dos resíduos do ARIMA(0,1,1)

ARCH Test:			
F-statistic	0.234814	Probability	0.629409
Obs*R-squared	0.240397	Probability	0.623919

ANEXO E: Previsão dentro da amostra

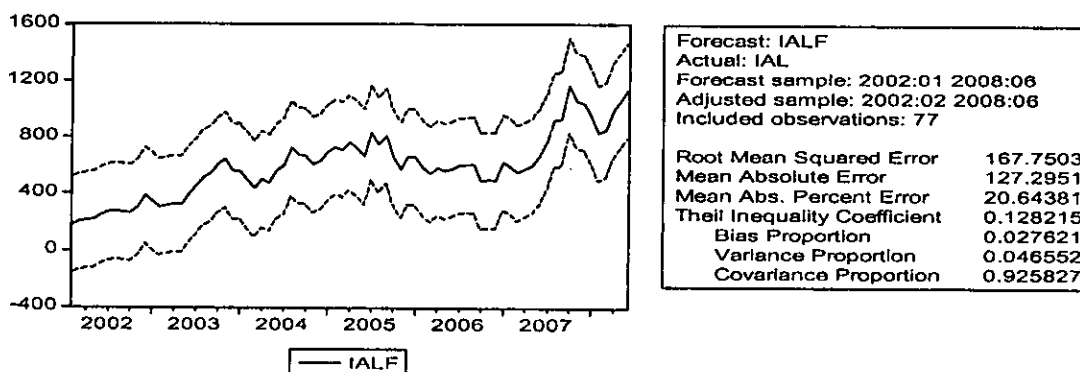


Figura 1E. Previsão dentro da amostra, limites de confiança e estatísticas de desempenho do ARIMA(0,1,26)

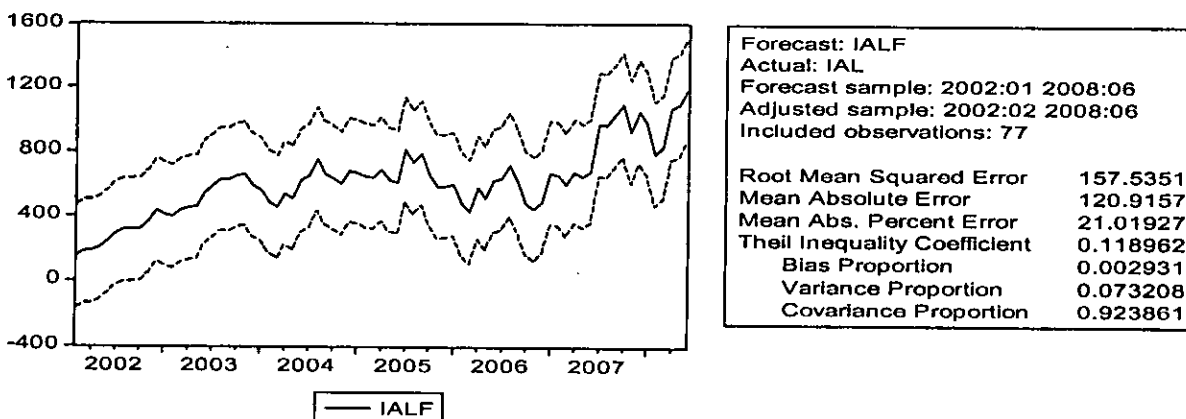


Figura 2E. Previsão dentro da amostra, limites de confiança e estatísticas de desempenho do ARIMA(0,1,25)

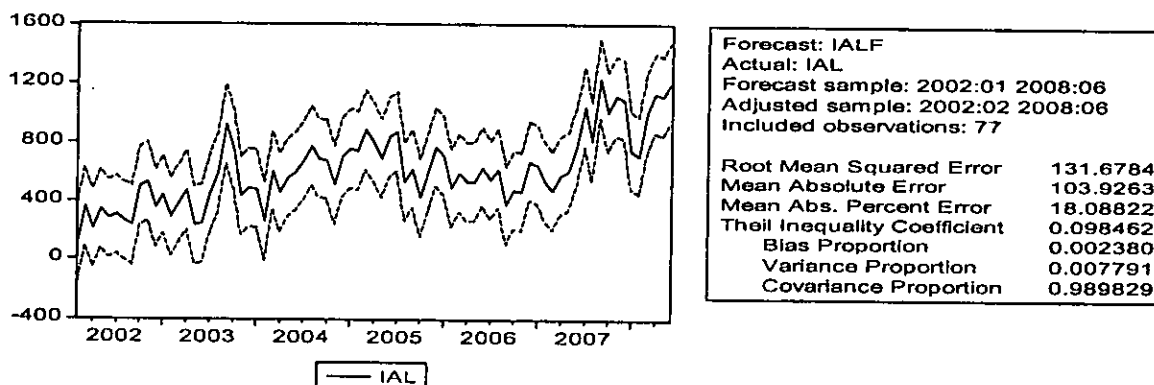


Figura 3E. Previsão dentro da amostra, limites de confiança e estatísticas de desempenho do ARIMA(0,1,1)