



UNIVERSIDADE EDUARDO MONDLANE

Faculdade de Ciências

Departamento de Matemática e Informática

Trabalho de Licenciatura em Estatística

**MODELOS DE PREVISÃO PARA ARRECADAÇÃO DO
IRPS NA PROVÍNCIA DO NIASSA**

Autor: Momade Zacarias Momade

Maputo, Abril de 2010



UNIVERSIDADE EDUARDO MONDLANE

Faculdade de Ciências

Departamento de Matemática e Informática

Trabalho de Licenciatura em Estatística

**MODELOS DE PREVISÃO PARA ARRECADAÇÃO DO
IRPS NA PROVÍNCIA DO NIASSA**

Autor: Momade Zacarias Momade

Supervisor: Dr. Alberto Mulenga

Maputo, Abril de 2010

DECLARAÇÃO DE HONRA

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

Maputo, aos ____ de _____ de _____

O Estudante

Momade Zacarias Momade

DEDICATÓRIA

Aos meus pais Zacarias Momade e Fátima Momade, pelo carinho e educação que me deram.

Aos meus irmãos Manjuma, Madania, Abacar, Muaziza e Momade, pela alegria de tê-los como irmãos.

AGRADECIMENTOS

À Allah, pela vida.

Ao meu supervisor, Dr. Alberto Mulenga, pela disponibilidade, orientação, compreensão e paciência que teve com minhas deficiências.

Ao meu primo Tapú, pelo apoio moral e incentivo.

À Associação Moçambicana para o Desenvolvimento e Ensino (ADEF), pelas aulas de Teologia que foram fundamentais ao meu bem-estar moral e espiritual durante este percurso de formação.

À Autoridade Tributária de Moçambique, pelo estágio que me concedeu.

RESUMO

A previsão do Imposto sobre o Rendimento das Pessoas Singulares – IRPS é fundamental para o planeamento das actividades do governo uma vez que este constitui o segundo principal imposto do Estado Moçambicano. Este trabalho desenvolve dois modelos de previsão para a arrecadação do IRPS nas Direcções de Área Fiscal de Cuamba e Lichinga cujo somatório corresponde ao IRPS da Província do Niassa. Para a análise dos dados foram utilizados os modelos Auto-Regressivos Integrados de Média Móvel – ARIMA baseados na metodologia de Box-Jenkins e, modelos de Ajustamento Exponencial Simples de Holt e de Holt-Winters. Os resultados mostram que a série de arrecadação do IRPS na Direcção de Área Fiscal de Cuamba não cumpriu com os pressupostos para a aplicação dos métodos de Ajustamento Exponencial. E o modelo que melhor se ajustou à série foi $ARIMA(1,1,27)$. Enquanto que a série de arrecadação do IRPS na Direcção de Área Fiscal de Lichinga, mostrou-se adequada para aplicação dos métodos de Box-Jenkins e de Ajustamento Exponencial de Holt e foi-lhe estimada os modelos $ARIMA(1,1,21)$ e $Holt(0.4,0.02)$ respectivamente. Comparando as estatísticas de desempenho na previsão, conclui-se que o modelo $ARIMA(1,1,21)$ é o mais adequado e o método de Box-Jenkins produziu melhores resultados em relação ao método de Ajustamento Exponencial. Finalmente recomenda-se a sua aplicação para a previsão do IRPS das restantes Províncias de Moçambique.

Palavras-chave: IRPS, Modelos ARIMA, Ajustamento Exponencial, Box-Jenkins, Holt e Holt-Winters.

ÍNDICE

I.	INTRODUÇÃO	1
1.1.	Definição do Problema	2
1.2.	Justificação do tema	2
1.3.	Objectivos	3
II.	REVISÃO DA LITERATURA	4
2.1.	O conceito de imposto e sua classificação	4
2.2.	Previsão de Receitas Públicas	5
2.2.1.	O Processo de Previsão	6
2.2.2.	Métodos de Previsão	7
2.3.	Séries Temporais	9
2.3.1.	Representação gráfica de Séries Temporais	10
2.3.2.	Decomposição de Séries Temporais	10
2.3.3.	Estacionaridade	11
2.4.	Modelos Auto-Regressivos Integrados de Média Móvel	13
2.4.1.	Método de Box-Jenkins	14
2.4.2.	CrITÉrios de selecção de Modelos ARIMA	17
2.5.	Modelos de Ajustamento Exponencial	18
2.5.1.	Ajustamento Exponencial Simples	18
2.5.2.	Ajustamento Exponencial de Holt	19
2.5.3.	Ajustamento Exponencial de Holt-Winters	20
2.6.	Métodos de avaliação de desempenho na previsão	21
III.	MATERIAL E MÉTODOS	23
3.1.	Material	23
3.2.	Métodos	24
3.2.1.	Análise descritiva	24
3.2.2.	Análise de estacionaridade	24

3.2.3.	Análise da sazonalidade	25
3.2.4.	Método de Box-Jenkins para modelos ARIMA	26
3.2.5.	Critérios de selecção de Modelos ARIMA	27
3.2.6.	Método de Ajustamento Exponencial	27
3.2.7.	Métodos de avaliação de desempenho na previsão	28
IV.	ANÁLISE E DISCUSSÃO DOS RESULTADOS	29
4.1.	Série de arrecadação do IRPS na Direcção de Área Fiscal de Cuamba	29
4.1.1.	Análise exploratória da série de IRPSCuamba	29
4.1.2.	Teste de estacionaridade	33
4.1.3.	Identificação dos Parâmetros	34
4.1.4.	Estimação dos Modelos ARIMA	35
4.1.5.	Diagnóstico de verificação do Modelo	36
4.1.6.	Selecção do Modelo ARIMA da série de IRPSCuamba	40
4.1.7.	Estimação do modelo de Ajustamento exponencial para a série de IRPSCuamba	41
4.1.8.	Previsão do IRPSCuamba	42
4.2.	Série de arrecadação do IRPS na Direcção de Área Fiscal de Lichinga	43
4.2.1.	Análise exploratória da série de IRPSLichinga	43
4.2.2.	Estimação do modelo $ARIMA(1,1,21)$	45
4.2.3.	Estimação do modelo de Ajustamento exponencial para a série de IRPSLichinga	45
4.2.4.	Selecção do modelo de previsão das arrecadações de IRPS na DAF de Lichinga	46
4.2.5.	Previsão do IRPSLichinga	47
V.	CONCLUSÕES E RECOMENDAÇÕES	48
5.1.	Conclusões	48
5.2.	Recomendações	49
	REFERÊNCIAS BIBLIOGRÁFICAS	50
ANEXO A:	Série de arrecadação do IRPS na Direcção de Área Fiscal de Cuamba	52
ANEXO B:	Série de arrecadação do IRPS na Direcção de Área Fiscal de Lichinga	55

LISTA DE FIGURAS

Figura 1:	Boxplots da série de IRPSCuamba antes e depois da correcção do outlier	29
Figura 2:	Representação da Série de IRPSCuamba	30
Figura 3:	Representação da série de IRPSCuamba em primeiras diferenças	34
Figura 4:	Correlogramas da série de IRPSCuamba em primeiras diferenças	35
Figura 5:	Correlogramas dos resíduos obtidos do modelo $ARIMA(1,1,27)$	37
Figura 6:	Tendência da série de IRPSCuamba	41
Figura 7:	Representação da série de IRPSLichinga	43
Figura 8:	Representação da série de IRPSLichinga em primeiras diferenças	44
Figura 9:	Tendência da série das arrecadações de IRPS na DAF de Lichinga	45
Figura 1A:	Previsão dentro da amostra através do modelo $ARIMA(0,1,27)$ e estatísticas de desempenho na previsão	53
Figura 2A:	Previsão dentro da amostra através do modelo $ARIMA(1,1,27)$ e estatísticas de desempenho na previsão	54
Figura 1B:	Correlogramas da série de IRPSLichinga em primeiras diferenças	55
Figura 2B:	Previsão de IRPSLichinga através do modelo $ARIMA(1,1,21)$ e estatísticas de desempenho na previsão	56

LISTA DE TABELAS

Tabela 1:	Taxas do Imposto sobre o Rendimento das Pessoas Singulares	5
Tabela 2:	Padrões teóricos da ACF e da PACF para modelos ARIMA	15
Tabela 3:	Relação das empresas que cessaram suas actividades a partir de 2007	30
Tabela 4:	Estatísticas da série de arrecadação do IRPS na DAF de Cuamba	31
Tabela 5:	Variação percentual e receitas anuais do IRPS na DAF de Lichinga	31
Tabela 6:	Teste de normalidade da série do IRPSCuamba dentro de cada mês	32
Tabela 7:	Teste de homogeneidade das variâncias	32
Tabela 8:	Teste de igualdade das arrecadações médias mensais da série IRPSCuamba	33
Tabela 9:	Teste de raiz unitária de ADF para a série de IRPSCuamba em nível	33
Tabela 10:	Teste de raiz unitária de ADF para a série de IRPSCuamba em primeiras diferenças	34
Tabela 11:	Estimação do Modelo $ARIMA(1,1,27)$ da série de IRPSCuamba	36
Tabela 12:	Estimação do Modelo $ARIMA(0,1,27)$ da série de IRPSCuamba	36
Tabela 13:	Teste de normalidade aos resíduos do modelo $ARIMA(1,1,27)$	37
Tabela 14:	Teste t da média dos resíduos do modelo $ARIMA(1,1,27)$	38
Tabela 15:	Teste de variância dos resíduos do modelo $ARIMA(1,1,27)$	38
Tabela 16:	Teste de Ljung-Box aos resíduos do modelo $ARIMA(0,1,27)$	39
Tabela 17:	Teste de variância dos resíduos do modelo $ARIMA(0,1,27)$	39
Tabela 18:	Teste de normalidade aos resíduos do modelo $ARIMA(0,1,27)$	40
Tabela 19:	Teste t da média dos resíduos do modelo $ARIMA(0,1,27)$	40
Tabela 20:	Seleção do Modelo ARIMA da série de IRPSCuamba	41
Tabela 21:	Valores previstos e realizados do IRPS mensal para a DAF de Cuamba	42
Tabela 22:	Estatísticas de desempenho na previsão do modelo $ARIMA(1,1,27)$	42
Tabela 23:	Estatísticas da série de arrecadação do IRPS na DAF de Lichinga	43
Tabela 24:	Taxa de crescimento anual e total das arrecadações de IRPS na DAF de Lichinga	44
Tabela 25:	Teste de raiz unitária de ADF para a série de IRPSLichinga em nível	44
Tabela 26:	Teste de raiz unitária de ADF para a série de IRPSLichinga em primeiras diferenças	45
Tabela 27:	Estimação do Modelo $ARIMA(1,1,21)$ da série de IRPSLichinga	45

Tabela 28:	Constantes de suavização do modelo de $Holt(\alpha, \beta)$	46
Tabela 29:	Estatísticas de desempenho na previsão dos Modelos $ARIMA(1,1,21)$ e $Holt(0.4,0.02)$	46
Tabela 30:	Valores previstos e realizados do IRPS mensal para a DAF de Lichinga	47
Tabela 1A:	Cobranças do IRPS na Direcção de Área Fiscal de Cuamba	52
Tabela 2A:	Estatísticas da Série IRPSCuamba de acordo com os grupos de meses	52
Tabela 3A:	Teste de Ljung-Box aos resíduos do modelo $ARIMA(1,1,27)$	53
Tabela 1B:	Cobranças do IRPS na Direcção de Área Fiscal de Lichinga	55
Tabela 2B:	Teste de igualdade das arrecadações médias mensais da série IRPSLichinga	56

LISTA DAS ABREVIATURAS

ACF	Função de autocorrelação
ACIS	Associação Comercial e Industrial de Sofala
ADF	Dickey-Fuller Aumentado
AIC	Critério de Informação de Akaike
AR	Auto-Regressivo
ARIMA	Auto-Regressivos Integrados de Média Móvel
ARMA	Auto-Regressivos de Média Móvel
AT	Autoridade Tributária de Moçambique
BIC	Critério de informação Bayesiano
CIRPS	Código do Imposto sobre o Rendimento das Pessoas Singulares
DAF	Direcção de Área Fiscal
ICE	Imposto sobre Consumos Específicos
IRPC	Imposto sobre o Rendimento das Pessoas Colectivas
IRPS	Imposto sobre o Rendimento das Pessoas Singulares
IRPSCuamba	Série de arrecadação do IRPS na Direcção de Área Fiscal de Cuamba
IRPSLichinga	Série de arrecadação do IRPS na Direcção de Área Fiscal de Lichinga
ISPC	Imposto Simplificado para Pequenos Contribuintes
MA	Média Móvel
PACF	Função de autocorrelação parcial
SBCC	Critério de informação Bayesiano de Schwartz corrigido
SQR_{ε}	Soma dos Quadrados dos Erros

I. INTRODUÇÃO

O Imposto sobre o Rendimento das Pessoas Singulares – IRPS, é um imposto directo que incide sobre o valor global anual dos rendimentos mesmo quando provenientes de actos ilícitos (artigo 1 do CIRPS¹, 2002).

Segundo Guajardo e Miranda (2000), a actividade de previsão de receitas ganha importância do ponto de vista de gestão, pois permite que o administrador tenha à sua disposição um instrumento seguro de apoio à tomada de decisão, que lhe permitirá um melhor planeamento das finanças do Estado, proporcionando a optimização das políticas públicas.

Uma das finalidades da Autoridade Tributária de Moçambique, adiante designada AT, órgão do aparelho do Estado Moçambicano, com autonomia administrativa, exercendo a sua actividade em todo o território nacional, é a de efectuar a Previsão e realizar a análise estatística de Cobrança de Receitas Tributárias e Aduaneiras sob sua administração, através da sua Divisão de Análise, Previsão e Controlo de Receitas, órgão responsável, no âmbito das funções a ela atribuídas, pelas alíneas g) e h), do n° 3, do artigo 22/2008, do Estatuto Orgânico da Autoridade Tributária de Moçambique.

Prever recursos futuros com base na informação passada é uma tarefa muito sensível daí que exige muito cuidado, habilidade e conhecimento científico a partir da escolha dos métodos, da organização dos dados, para além do conhecimento das variáveis envolvidas na previsão. Guajardo e Miranda (2000) afirmam que uma boa previsão pode evitar desequilíbrios orçamentais, objectivo de preocupação constante das Autoridades governamentais.

Segundo Makridakis et al. (1984), as técnicas de previsão auxiliam na tomada de decisões em actividades que necessitam de planeamento e da redução de incerteza, tornando assim os possíveis e futuros riscos mais visíveis e consequentemente mais controláveis.

Martins e Laugeni (1999), afirmam que a previsão é 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.

¹ Código do Imposto sobre o Rendimento das Pessoas Singulares.

Vários são os modelos utilizados para previsão de receitas tributárias, neste trabalho, usa-se os métodos quantitativos univariados, onde analisam-se as séries de arrecadação do IRPS da Província do Niassa a fim de se conhecer os seus devidos modelos de previsão através da modelagem ARIMA e/ou Ajustamento Exponencial.

1.1. Definição do Problema

Segundo ACIS² (2008), para a realização dos fins que se propõe, e que se traduzem na realização de despesas orientadas fundamentalmente à satisfação das necessidades públicas, o Estado necessita de obter receitas. Tais receitas provêm prioritariamente da cobrança de tributos (prestações compulsivas pecuniárias ou em espécie) que podem revestir a natureza de impostos ou de taxas ou de venda e exploração de bens do património público (receitas patrimoniais).

Para um bom planeamento dos programas governamentais, uma vez que as receitas seguem processos aleatórios e que tudo é previsto no orçamento do Estado a partir da receita possível de ser obtida, é importante prever as receitas seja de forma agrupada (exemplo receitas tributárias compostas por impostos e taxas) ou individual (exemplo Imposto sobre o Rendimento das Pessoas Singulares – IRPS e Imposto sobre o Valor Acrescentado – IVA) construindo modelos com uma base científica que minimizem o erro de previsão. Diante deste contexto surge a seguinte questão:

- Como conhecer um modelo de previsão para a arrecadação mensal do IRPS, tendo em conta a sua informação passada?

1.2. Justificação do tema

A razão da escolha do tema está ligada ao facto de o mesmo possibilitar a aplicação dos conhecimentos adquiridos ao longo do curso e à sua importância no planeamento das actividades de um Estado, pois a previsão de receitas provenientes de impostos é um instrumento bastante valioso para os administradores públicos. Sua finalidade é a de contribuir para a tomada correcta de decisões e providências e permitir que sejam evitados riscos e aproveitadas as oportunidades.

² Associação Comercial e Industrial de Sofala.

1.3. Objectivos

Geral

- Desenvolver dois modelos de previsão para a arrecadação do IRPS nas Direcções de Área Fiscal de Cuamba e de Lichinga cujo somatório corresponde ao IRPS da Província do Niassa.

Específicos

- Fazer análise descritiva das séries de arrecadação mensal do IRPS no período de Janeiro de 2005 à Novembro de 2009.
- Identificar e estimar os modelos que melhor se adequam aos dados das séries em análise.
- Fazer previsões para o primeiro trimestre de 2010 a partir dos modelos estimados.

Mais adiante, o trabalho conta com mais 4 capítulos: Revisão da Literatura onde são descritos os modelos ARIMA baseados na metodologia de Box-Jenkins e, modelos de Ajustamento Exponencial Simples de Holt e de Holt-Winters. No capítulo de Material e Métodos são indicados passos necessários para a aplicação dos modelos anteriormente mencionados e as séries em estudo. No capítulo da Análise e Discussão dos Resultados, neste capítulo são analisadas as séries em estudo e discutidos os seus respectivos resultados e finalmente no capítulo de Conclusões e Recomendações são apresentadas as principais conclusões do estudo e deixadas algumas recomendações para futuras investigações.

II. REVISÃO DA LITERATURA

2.1. O conceito de imposto e sua classificação

Segundo ACIS (2008), imposto é uma prestação obrigatória, avaliável em dinheiro, exigida por uma entidade pública para a prossecução de fins públicos, sem contraprestação individualizada, e cujo facto tributário assenta em manifestações de capacidade contributiva.

Montoro e Franco (2002), afirmam que os impostos podem ser classificados de diversas formas, dependendo do enfoque que se queira ressaltar. Do ponto de vista da incidência económica, eles classificam-se em impostos directos e indirectos.

O imposto directo é aquele que incide directamente no rendimento ou na riqueza (Imposto sobre o Rendimento das Pessoas Colectivas – IRPC, Imposto sobre o Rendimento das Pessoas Singulares – IRPS e Imposto Especial sobre o Jogo). A pessoa que recolhe o imposto (contribuinte direito) é quem aguenta com o ónus tributário (contribuinte de facto), ou seja, o contribuinte de direito é a mesma pessoa do contribuinte de facto.

Por outro lado, o imposto indirecto, que incide sobre as operações de circulação de mercadorias e sobre as prestações de serviços (Imposto sobre o Valor Acrescentado – IVA, Imposto sobre Consumos Específicos – ICE, direitos de importação, etc.), tem como característica a dicotomia entre as duas espécies de contribuintes. Nesta espécie tributária quem tem o dever legal de recolher o imposto não é aquele que aguenta com o ónus do tributo, portanto o contribuinte de direito e de facto são pessoas distintas.

Imposto sobre o rendimento das pessoas singulares

Natureza do IRPS

Os rendimentos referidos provêm das seguintes categorias:

Primeira categoria: Rendimentos do trabalho dependente;

Segunda categoria: Rendimentos empresariais e profissionais;

Terceira categoria: Rendimentos de capitais e das mais-valias;

Quarta categoria: Rendimentos prediais;

Quinta categoria: Outros rendimentos.

Incidência pessoal do IRPS

Ficam sujeitas a IRPS as pessoas singulares que residam em território moçambicano e as que, nele não residindo, aqui obtenham rendimentos. Existindo agregado familiar, o imposto é devido pelo conjunto dos rendimentos das pessoas que o constituem, considerando-se como sujeitos passivos aquelas a quem incumbe a sua direcção (artigo 17 do CIRPS, 2002).

Âmbito da sujeição do IRPS

O IRPS pago pelas pessoas residentes em território moçambicano, incide sobre a totalidade dos seus rendimentos, incluindo os obtidos fora desse território. Tratando-se de não residentes, o IRPS incide unicamente sobre os rendimentos obtidos em território moçambicano (artigo 19 do CIRPS, 2002).

Tabela 1: Taxas do Imposto sobre o Rendimento das Pessoas Singulares

Rendimentos Colectáveis Anuais em Meticais (R)	Taxas
$R \leq 42.000$	10%
$42.000 < R \leq 168.000$	15%
$168.000 < R \leq 504.000$	20%
$504.000 < R \leq 1.512.000$	25%
$R > 1.512.000$	32%

Fonte: Artigo 64 do CIRPS (2002).

2.2. Previsão de Receitas Públicas

A previsão de receitas públicas envolve o uso de técnicas analíticas para projectar a quantidade de recursos financeiros disponíveis num determinado tempo futuro. No sector público, as receitas provêm prioritariamente da cobrança de tributos (prestações compulsivas pecuniárias ou em espécie) que podem revestir a natureza de impostos ou de taxas ou de venda e exploração de bens do património público (receitas patrimoniais). A previsão procura identificar o relacionamento entre os factores que afectam as receitas (alíquotas tributárias, vendas macroeconómicas, etc.) e a arrecadação governamental (tributos e outras receitas).

As previsões das receitas podem ser aplicadas tanto à receita total quanto às suas fontes individuais. Obviamente, não há um método único para se projectar receitas. Assim sendo, os

diferentes métodos investigados tendem a trabalhar em função do tipo de receita que se pretende prever.

Segundo Guajardo e Miranda (2000), a previsão de receita pública está intimamente ligada ao processo da política e, portanto, sujeita ao escrutínio e até mesmo à pressão política, factores que influenciarão decisivamente os resultados da arrecadação.

2.2.1. O Processo de Previsão

A política fiscal do governo Moçambicano é afectada pelo contexto em que ela é formada. Lida tanto com interesses políticos, quanto económicos. Portanto, é essencial estabelecer as hipóteses e os procedimentos com os quais as partes envolvidas coadunam, assim como um mecanismo para avaliar a validade da receita que se prevê. Dessa forma, é fundamental que se estabeleça um procedimento bastante disciplinado.

Guajardo e Miranda (2000), sugerem um procedimento de sete etapas que devem ser aplicadas a cada tipo de receita a ser prevista:

1ª Etapa: A primeira etapa envolve a selecção de um período de tempo sobre o qual os dados serão examinados. O tamanho do período de tempo depende da disponibilidade e da qualidade dos dados, do tipo de receita a ser prevista e do grau de exactidão desejado.

2ª Etapa: Na segunda etapa, são examinados os dados a fim de determinar todos os padrões, taxas de variação, ou tendências que estejam evidentes. Uma vez identificados tais componentes, o responsável pela previsão precisará decidir o grau que a receita é previsível. Isto é feito examinando-se as características subjacentes à receita, tais como as estruturas de alíquotas usadas, as mudanças na demanda, ou as variações sazonais ou cíclicas.

3ª Etapa: Os responsáveis pela previsão, em seguida, precisam compreender as hipóteses subjacentes associadas às fontes de receitas. Necessitam de considerar a que grau a receita é afectada por circunstâncias económicas, mudanças na demanda do cidadão e nas políticas do governo, entre outros efeitos. Estas suposições ajudam a determinar o método de previsão que será adequado a cada caso.

4ª Etapa: Esta etapa consiste em projectar os valores dos grupos de receitas para o horizonte de tempo desejado. O método seleccionado para executar a projecção depende da natureza e do tipo de receita.

5ª Etapa: Depois das projecções forem feitas, as estimativas são avaliadas, para que se verifique a fiabilidade e a validade dos procedimentos. Para avaliar a validade das estimativas, as hipóteses associadas às fontes de receitas devem ser reexaminadas. A fiabilidade é avaliada através de uma análise de sensibilidade, que envolve a verificação dos elementos chaves usados para criar as estimativas.

6ª Etapa: Na sexta etapa, os grupos de receitas são acompanhados e comparados às estimativas. Tal acompanhamento serve para avaliar a exactidão das projecções e determinar a probabilidade de ocorrência de quedas ou aumentos no orçamento.

7ª Etapa: Finalmente, como as circunstâncias que afectam a geração de receitas governamentais mudam, a previsão precisa ser actualizada.

2.2.2. Métodos de Previsão

Existe uma grande quantidade de técnicas de previsão disponíveis (Frank, 1993). As técnicas dividem-se em qualitativas informais à quantitativas formais. Em geral, os responsáveis pela previsão usam uma variedade de técnicas, reconhecendo que algumas apresentam melhores resultados do que outras, dependendo da natureza da fonte de receita.

Métodos Qualitativos de Previsão

Os métodos qualitativos de previsão baseiam-se em conjecturas sobre a futura arrecadação de determinadas receitas. Estas técnicas são frequentemente de aproximações conjecturais ou não extrapolativos. Além de não apresentar uma formalidade matemática adequada, tais técnicas normalmente não fornecem especificações rigorosas das hipóteses subjacentes ao modelo.

Entre os métodos qualitativos de previsão, o mais usado é o da previsão conjectural. Esta técnica baseia-se na utilização de um indivíduo ou um grupo pequeno de pessoas que fazem avaliações de futuras circunstâncias prováveis e analisam os efeitos destas sobre a receita a ser prevista.

Embora seja *ad hoc*, a técnica pode produzir estimativas muito boas, especialmente quando pessoas experientes estão envolvidas no processo.

As aproximações conjecturais tendem a trabalhar quando o ambiente económico está mudando rapidamente. Quando as circunstâncias económicas, políticas ou administrativas estão em fluxo intenso, métodos quantitativos não podem capturar informações importantes sobre os factores que provavelmente estão alterando os padrões históricos, sendo pouco recomendados para o caso.

Uma variação da aproximação conjectural é a chamada previsão de consenso. Aqui, os peritos, familiarizados com os factores que afectam um tipo particular de receita, encontram-se para discutir algumas circunstâncias gerais, a fim de definir o que provavelmente acontecerá à arrecadação de determinada receita. A previsão de consenso tende a funcionar quando há pouca informação histórica para subsidiar a elaboração de um método quantitativo de previsão, por exemplo quando se deseja prever a arrecadação de um novo imposto como é o caso do Imposto Simplificado para Pequenos Contribuintes – ISPC.

Segundo Guajardo e Miranda (2000), as aproximações conjecturais de previsão têm certamente seu lugar entre os métodos de previsão. Com alguma extensão, uma perspectiva conjectural pode complementar outras técnicas de previsão, mesmo aquelas baseadas em métodos quantitativos extremamente rigorosos. Como pode-se suspeitar, entretanto, as aproximações conjecturais podem estar sujeitas à polarização e a outras fontes de erro.

Métodos Quantitativos de Previsão

Métodos quantitativos baseiam-se em dados numéricos relevantes para uma determinada fonte de receita em análise. Tais métodos procuram, explicar as hipóteses e os procedimentos usados para gerar as previsões, além de atribuírem, uma margem de erro às previsões, fornecendo uma indicação do grau de incerteza associado às estimativas.

Por outro lado, os métodos quantitativos de previsão podem ser classificados em métodos de previsão univariados e multivariados:

Os métodos de previsão univariados são baseados na análise de uma única série temporal ou seja uma única variável. Aqui, o comportamento passado da variável ou da série temporal em estudo explica os seus futuros valores.

Os métodos de previsão multivariados, por sua vez, utilizam a relação de correlação estatística existente entre várias variáveis independentes para estimar o futuro comportamento da variável a ser estimada, através da construção de modelos de regressão.

Segundo Cirincione et al. (1999), em geral, os métodos quantitativos são melhores do que os métodos qualitativos na previsão das futuras receitas. Métodos quantitativos mais simples também executam previsões tão boas quanto os métodos mais complexos (Makridakis et al., 1984).

Segundo Frank (1993), os modelos univariados normalmente apresentam melhor desempenho que os modelos multivariados, ao menos a curto prazo, dado a incerteza associada à captura de todos os factores económicos relevantes que influenciam a geração da receita.

2.3. Séries Temporais

Segundo Ehlers (2005), uma série temporal 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.

De acordo com Brockwell e Davis (1996), uma série temporal X_t é a realização de uma família de variáveis aleatórias, ou seja, um modelo de série temporal para dados observados é uma especificação das distribuições de uma sequência de variáveis aleatórias X_t da qual x_t é denominada uma realização.

São necessários para a caracterização da sequência de variáveis aleatórias somente os momentos de primeira e segunda ordem da distribuição conjunta (Newbold e Granger, 1974). O momento de primeira ordem é definido como o valor esperado ou média de X_t :

$$\mu_t = E(X_t) \tag{1}$$

O momento de segunda ordem é definido como a covariância entre X_t e X_s :

$$\gamma_{t,s} = cov(X_t, X_s) = E[(X_t - \mu_t)(X_s - \mu_s)] \quad (2)$$

Define-se também a variância de X_t como:

$$\gamma_{t,t} = cov(X_t, X_t) = var(X_t) = E[(X_t - \mu_t)(X_t - \mu_t)] = E[(X_t - \mu_t)^2] \quad (3)$$

2.3.1. Representação gráfica de Séries Temporais

Segundo Murteira e Black (1983), o primeiro e fundamental passo em qualquer estudo que envolve séries temporais, é a representação gráfica das mesmas, visto que, revelam padrões de comportamento importantes, como a tendência, padrões cíclicos, alterações estruturais, observações aberrantes (outlier), etc. A representação gráfica mostra a visão geral do comportamento da série temporal, porém, é necessário efectuar análises estatísticas formais.

2.3.2. Decomposição de Séries Temporais

Segundo Spiegel (1985), as curvas obtidas pelas séries temporais podem apresentar quatro características principais: tendência (T), variação sazonal (S), variação cíclica (C) e a aleatoriedade (A). Estas características juntas determinam a função: $X_t = f(T_t, S_t, C_t, A_t)$.

Tendência: indica a direcção geral dos valores estudados. A principal característica desta componente é um movimento suave registado num período longo de tempo, direccionando os dados de modo constante, crescente ou decrescente. Makridakis e Wheelwright (1989), asseguram que um padrão de tendência existe quando há um aumento ou diminuição do valor médio dos dados a longo prazo.

Variações sazonais: de acordo com Murteira e Black (1983), são movimentos oscilatórios de curta duração (períodos não superiores a um ano), que traduzem a influência de factores cuja actuação é periódica, no sentido de aumentar ou diminuir a intensidade do fenómeno. Segundo Wallis e Thomas (1971), a sazonalidade pode ser definida como o conjunto dos movimentos ou flutuações com período igual ou inferior a um ano, sistemáticos, mas não necessariamente regulares que ocorrem numa série temporal.

Morettin e Tolo (2004), afirmam ser difícil definir, tanto do ponto de vista conceitual como estatístico, a sazonalidade e, empiricamente, consideram como sazonais os fenômenos que ocorrem regularmente de ano para ano.

Variações cíclicas: referem-se às oscilações a longo prazo em torno da curva de tendência, estas podem ser ou não periódicas. Para Murteira e Black (1983), as variações cíclicas exprimem a influência de factores excepcionais, como greves, catástrofes climáticas que não podem ser previstos e, portanto, incluídos no modelo.

Aleatoriedade: também chamado de erro residual possui uma sequência de variáveis aleatórias independentes e identicamente distribuídas, frequentemente com média zero e variância constante, são caracterizados por sua duração curta e intensidade variável.

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

2.3.3. Estacionaridade

Segundo Gujarati (2006), um processo estocástico é estacionário quando a sua média é constante ($E(X_t) = \mu$) e a sua variância também é constante ($var(X_t) = E(X_t - \mu)^2 = \sigma^2$) 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 dois períodos de tempo e não do próprio tempo em que a covariância é calculada.

Conforme Spiegel (1985), se uma série temporal não for estacionária, só pode-se estudar seu comportamento para o período considerado. Para a finalidade de fazer previsões, essas séries temporais não estacionárias não têm valor prático. Segundo Brockwell e Davis (1996), se a série não for estacionária, é habitualmente possível convertê-la numa outra estacionária recorrendo a algumas técnicas apropriadas, tais como:

1. Se os dados revelam uma tendência crescente ou decrescente, pode-se ajustar uma curva e subtraí-la aos dados, ficando com uma série de resíduos. Como o propósito de ajustamento é remover a tendência de longo prazo, um ajustamento de uma recta é suficiente;

2. Se a variância não é constante, transformar a série X_t em $Y_t = \ln X_t$ ou $Y_t = \sqrt{X_t}$ pode ajudar a estabilizá-la.
3. Em qualquer caso, pode-se diferenciar a série, isto é, dada uma série com elementos X_t , constrói-se a série das diferenças, tal que: $Y_t = \Delta X_t = X_t - X_{t-1}$.

Função de autocorrelação

Segundo Newbold e Granger (1974), a autocorrelação mede a intensidade com que um valor observado no instante $t + k$ é influenciado por aquele observado no tempo t e é definido como:

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad (4)$$

Segundo Hamilton (1994), o tamanho da autocovariância de uma série temporal estacionária $\rho_x(k) = cov(X_{t+k}, X_t)$ depende da escala em que X_t é medida. Sendo assim, para efeitos de interpretação, é 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 é:

$$\rho_x(k) = \frac{\gamma_k(k)}{\gamma_x(0)} = \frac{\gamma_k(k)}{\sigma^2} = cor(X_{t+k}, X_t) \quad (5)$$

Segundo Brockwell e Davis (1996), 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.

Função de autocorrelação parcial

Segundo Mills (1990), 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 e é representada por:

$$\varphi_{kk} = cor(X_t, X_{t+k} | X_{t+1}, X_{t+2}, \dots, X_{t+k-1}) \quad (6)$$

2.4. Modelos Auto-Regressivos Integrados de Média Móvel

Os Modelos Auto-Regressivos Integrados de Média Móvel – ARIMA são modelos univariados, isto é, nestes modelos a variável dependente é explicada a partir dela própria. Devido a sua popularidade, não serão apresentados com detalhe neste trabalho aconselhando-se a quem queira se actualizar mais acerca destes modelos a consultar os seguintes livros: Box e Jenkins (1976), Newbold e Granger (1974) e Enders (1995).

Dependendo da série com a qual se esteja trabalhando, se é estacionária ou não, o processo estocástico gerador da série pode ser descrito por um modelo específico. Dessa forma, para séries estacionárias os modelos apropriados são, o auto-regressivo de ordem p : $AR(p)$, o de média móvel de ordem q : $MA(q)$ e o auto-regressivo de média móvel de ordem p e q : $ARMA(p, q)$.

O modelo auto-regressivo de ordem p : $AR(p)$ é representado pela formula:

$$X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t = \sum_{j=1}^p \phi_j X_{t-j} + \varepsilon_t \quad (7)$$

onde ε_t é um processo de ruído branco com média zero e variância constante, ϕ_i ($i = 1, 2, \dots, p$) são parâmetros que descrevem como X_t se relaciona com X_{t-i} .

O modelo de médias móveis é dado pela forma:

$$X_t = \mu + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t = \sum_{j=0}^q \theta_j \varepsilon_{t-j} + \varepsilon_t \quad (8)$$

onde θ_i ($i = 1, 2, \dots, q$) são parâmetros que descrevem como X_t se relaciona com ε_{t-i} .

Um processo $ARMA(p, q)$ é uma generalização dos modelos $AR(p)$ e $MA(q)$, sendo definido como:

$$X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (9)$$

Ou mesmo por, $X_t = \sum_{j=1}^p \phi_j X_{t-j} + \sum_{j=0}^q \theta_j \varepsilon_{t-j} + \varepsilon_t$

Quando a série for não estacionária, o modelo apropriado é o auto-regressivo integrado de média móvel de ordem p , d e q : $ARIMA(p, d, q)$. Onde d indica a ordem de integração a partir da qual a série torna-se estacionária. Denotando a série diferenciada por $Y_t = \Delta^d X_t$ o seu modelo, isto é, o $ARIMA(p, d, q)$, é dado por:

$$Y_t = \alpha_1 Y_{t-1} + \dots + \alpha_p Y_{t-p} + \beta_1 \varepsilon_{t-1} + \dots + \beta_q \varepsilon_{t-p} + \varepsilon_t \quad (10)$$

2.4.1. Método de Box-Jenkins

Segundo Newbold e Granger (1974), o objectivo da metodologia de Box-Jenkins é encontrar um modelo estocástico linear da classe $ARIMA(p, d, q)$ que possa ser utilizado para fornecer previsões de futuros valores da série.

Um conceito importante nessa metodologia é o princípio da parcimónia (Enders, 1995). Tal princípio sugere que modelos mais simples, com poucos parâmetros, produzem melhores previsões que modelos super parametrizados. Um modelo parcimonioso ajusta bem os dados sem incorporar coeficientes inúteis.

Segundo Box e Jenkins (1976), a fim de se obter melhores resultados na utilização da metodologia de Box-Jenkins, três hipóteses básicas precisam ser observadas.

A primeira é relativa ao tamanho inicial da amostra geralmente aceite, que deve ser de, no mínimo, 50 observações.

A segunda suposição é a de que a série de dados seja estacionária, isto é, que a série varie em torno de uma média constante e com uma variância constante. Se a série não for estacionária, ela deve ser diferenciada ou deve-se proceder à adição de uma tendência no tempo.

A terceira hipótese é a de que a série seja homoscedástica, isto é, tenha uma variância constante ao longo do tempo.

A estratégia de modelagem é baseada em um ciclo de três etapas iterativas: Identificação dos parâmetros do modelo, estimação do modelo e diagnóstico de verificação do modelo.

Identificação dos parâmetros do modelo

Segundo Newbold e Granger (1974), esta etapa consiste em seleccionar valores para p , d , q e é considerada a mais difícil e delicada por envolver subjectivismo e julgamento pessoal e não há consenso sobre qual a melhor estratégia a ser seguida.

Além de identificar os valores para p e q , o grau de diferenciação da série (d) precisa de ser conhecido. Para tal, faz-se o teste de raiz unitária através da estatística de Dickey-Fuller Aumentado (ADF) ou uma outra estatística específica depois de uma inspecção da ACF e da PACF e do gráfico 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 baseia-se na análise de ACF e PACF.

Tabela 2: Padrões teóricos da ACF e da PACF para modelos ARIMA

Tipo de modelo	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 desfasagens
MA(q)	Apresenta picos significativos até q desfasagens	Declina exponencialmente
ARIMA(p, q)	Diminui exponencialmente	Diminui exponencialmente

Fonte: (Gujarati, 2006)

Estimação do modelo

Nesta etapa de estimação, os coeficientes identificados na etapa anterior são estimados usando técnicas estatísticas.

Assumindo que um modelo ARIMA da forma

$(1 - \Phi_1 B - \dots - \Phi_p B^p)(1 - B)^d X_t = (1 - \theta_1 + \dots + \theta_q B^q) \varepsilon_t$, seja escolhido conforme a etapa anterior, o objectivo seguinte é estimar, utilizando o método de máxima verossimilhança, os parâmetros $\Phi = (\Phi_1, \dots, \Phi_p)'$, $\theta = (\theta_1, \dots, \theta_q)'$ e σ_ε^2 , a variância de ε_t .

Mesmo sabendo-se que o método de máxima verossimilhança é o mais sugerido por vários autores, existem outros métodos de estimação entre eles destacam-se o método dos mínimos quadrados exactos e o método dos mínimos quadrados condicional.

Diagnóstico de verificação do modelo

Esta etapa indica se o modelo identificado e estimado descreve adequadamente o comportamento dos dados da série X_t . Caso o modelo não seja adequado, o ciclo deve começar novamente (Cribari-Neto, 2000).

A correcta especificação de um modelo ARIMA é verificada no termo do erro ε_t , pois ele deve constituir um processo de ruído branco (Newbold e Granger, 1974). Assim, a verificação da adequabilidade do modelo é efectuada nas autocorrelações amostrais dos erros ε_t , as quais seguem assintoticamente uma distribuição normal, com média zero e variância constante, se

forem provenientes de um ruído branco. Como os verdadeiros erros ε_t não são conhecidos, a inferência baseia-se nas estimativas dos erros, os resíduos $\hat{\varepsilon}_t$.

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

No teste de Ljung-Box deseja-se testar a hipótese nula de que as k primeiras autocorrelações dos resíduos são nulas, isto é:

$H_0: \rho_1^{(\varepsilon)} = \rho_2^{(\varepsilon)} = \dots = \rho_k^{(\varepsilon)} = 0$, onde os $\rho_i^{(\varepsilon)}$ são as autocorrelações dos resíduos. A estatística do teste é dada pela equação (11).

$$Q = T(T + 2) \sum_{i=1}^k \frac{\rho_1^{(\hat{\varepsilon})^2}}{(T - i)} \quad (11)$$

onde T é o tamanho da série. O valor da estatística Q é comparado com o valor tabelado da distribuição Qui-Quadrado (χ^2) com $k - (p + q + 1)$ graus de liberdade e a hipótese nula de que as k primeiras autocorrelações dos resíduos sejam nulas é rejeitada se o valor de Q for superior ao valor de χ^2 .

A escolha usual de quantas autocorrelações devem ser testadas em conjunto é feita por aproximação, tomando $k = \sqrt{T}$ (Makridakis et al., 1998).

2.4.2. Critérios de selecção de Modelos ARIMA

Segundo Ehlers (2005), em muitas aplicações vários modelos podem ser julgados adequados em termos de comportamento dos resíduos. O critério de informação é um dos métodos usados para seleccionar um modelo entre vários concorrentes.

Os critérios de selecção mais utilizados de modelos ARIMA são o Critério de Informação de Akaike - *AIC*, Critério de Informação Bayesiano - *BIC* e o Critério de Informação Bayesiano de Schwartz Corrigido - *SBCC*. Esses critérios incorporam um termo de penalidade para o aumento do número de parâmetros (p e q) no modelo, de forma que modelos mais “parcimoniosos”, ou seja, com o menor número de parâmetros, sejam escolhidos. As equações para esses critérios, sendo T o número de observações, são (Cribari-Neto, 2000):

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

Para dados normalmente distribuídos e usando estimadores de máxima verossimilhança para os parâmetros pode-se mostrar que:

$$AIC = T * [\log(\sigma_{\hat{\epsilon}}^2)] + 2m \quad (12)$$

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

$$\text{ou mesmo, } BIC = T \log(\sigma_{\hat{\epsilon}}^2) - (T - m) \log\left(1 - \frac{m}{T}\right) + m \log(T) + m \log\left[\frac{\left(\frac{\sigma_{\hat{\epsilon}}^2}{\sigma_{\hat{\epsilon}}^2} - 1\right)}{m}\right] \quad (13)$$

$$SBCC = T \log(\sigma_{\hat{\epsilon}}^2) + m \log(T) \quad (14)$$

onde m é o número de parâmetros em modelos ARIMA, isto é, $m = p + q + 1$.

Na prática o *AIC* super estima assintoticamente a ordem verdadeira do modelo (Granger e Newbold, 1986) apresentando uma tendência de escolher modelos super parametrizados (Cribari-Neto, 2000).

O *BIC* é um critério consistente, de forma que ele fornece estimativas de p e q que convergem em probabilidade para os valores verdadeiros à medida que T tende ao infinito (Brockwell e Davis, 1996). Já o *AIC* não é consistente, no entanto ele é assintoticamente eficiente para

modelos puramente auto-regressivos. A regra consiste em seleccionar o modelo cujo seu valor de critério de informação é mínimo.

2.5. Modelos de Ajustamento Exponencial

O método de Ajustamento exponencial para descrever as técnicas de previsão de séries temporais, isola na série até 4 factores: nível, tendência linear, factor sazonal e um elemento residual não previsível, às vezes chamado erro aleatório.

Segundo Soares (1991), o método de Ajustamento exponencial, também chamado “suavização exponencial”. O nome “suavização” provém do facto de a série, após reduzida a seus componentes estruturais, ter menor número de variações bruscas, mostrando um comportamento mais suave. De acordo com o mesmo autor, o termo “exponencial” aparece porque os processos de “suavização” envolvem médias aritméticas ponderadas, onde os pesos decrescem exponencialmente na medida em que se avança no passado.

2.5.1. Ajustamento Exponencial Simples

A forma mais simples de ajustamento exponencial é aquela que é feita para séries temporais que não possuem sazonalidade nem tendência crescente ou decrescente.

O método consiste em estimar o “nível” (ou a “média”) da série e usar esse nível como previsão de futuros valores. O nível da série no tempo t é estimado pela equação (15):

$$\bar{x}_t = \alpha x_t + \alpha(1 - \alpha)x_{t-1} + \alpha(1 - \alpha)^2 x_{t-2} + \dots, \quad (15)$$

onde: α é a constante de suavização e $\alpha \in (0,1)$.

Uma forma mais simples de escrever a equação (15) é obtida substituindo t por $(t - 1)$ em seguida multiplicar os dois membros da equação (15) por $(1 - \alpha)$ e adicionar os dois membros o termo αx_t , o que leva à equação (16):

$$\bar{x}_t = \alpha x_t + (1 - \alpha)\bar{x}_{t-1} \quad (16)$$

Assim os futuros valores ($f_{n,h}$) são estimados a partir da seguinte equação:

$$f_{n,h} = \bar{x}_n; \quad h = 1, 2, 3, \dots, n. \quad (17)$$

Para iniciar-se com as previsões é necessário especificar um valor inicial, que usualmente é: $\bar{x}_1 = x_1$.

O peso de cada termo é determinado pelo valor de α . A escolha desta constante é feita de maneira que seu valor minimize a soma dos quadrados dos erros (Janacek, 2001). O erro de previsão é definido pela equação:

$$e_t = x_t - f_{t-1,1}; \quad t = 3, 4, \dots, n. \quad (18)$$

a soma dos quadrados dos erros é dada por:

$$SQR_\varepsilon = \sum_{t=3}^n e_t^2 = \sum_{t=3}^n (x_t - f_{t-1,1})^2 \quad (19)$$

2.5.2. Ajustamento Exponencial de Holt

Caso a série temporal apresente tendência crescente ou decrescente, a equação (17) não é válida para fazer previsões. O método de ajustamento exponencial de $Holt(\alpha, \beta)$ leva em consideração esses movimentos e permite estimar também o actual coeficiente angular da série. Se o gráfico apresentar tendências crescentes e decrescentes numa só série, é aconselhável dividi-la de acordo com as tendências formando novas séries.

Assim, o nível e o coeficiente angular da série são dados, respectivamente, por:

$$\bar{x}_t = \alpha x_t + (1 - \alpha)(\bar{x}_{t-1} + T_{t-1}) \quad (20)$$

$$T_t = \beta(\bar{x}_t - \bar{x}_{t-1}) + (1 - \beta)T_{t-1}$$

com $0 < \alpha < 1$ e $0 < \beta < 1$, constantes de suavização.

As previsões são obtidas supondo um acréscimo ou decréscimo contínuo dado pela última estimativa de inclinação, assim:

$$f_{n,h} = \bar{x}_n + hT_n; \quad h = 3, 4, 5, \dots, n. \quad (21)$$

Os possíveis valores iniciais do algoritmo são:

$$\bar{x}_2 = x_2$$

$$T_2 = x_2 - x_1$$

O erro de previsão e a soma dos quadrados dos erros são dados pelas equações 18 e 19 respectivamente.

2.5.3. Ajustamento Exponencial de Holt-Winters

Segundo Cribari-Neto (2000), se a série temporal apresentar movimentos sazonais de período s , o método adequado é o de Holt-Winters que integra a componente sazonal. Assim, o método de Holt-Winters é definido de modo que para cada período seja necessário estimar um factor de sazonalidade, F_t . No instante t , a última estimativa do factor de sazonalidade para o período é F_{t-s} , (obtido do mesmo período do ano anterior). As equações para o nível, a inclinação e o factor de sazonalidade, considerando que a sazonalidade é aditiva, são:

$$\begin{aligned}\bar{x}_t &= \alpha(x_t - F_{t-s}) + (1 - \alpha)(\bar{x}_{t-1} + T_{t-1}) \\ T_t &= \beta(\bar{x}_t - \bar{x}_{t-1}) + (1 - \beta)T_{t-1} \\ F_t &= \gamma(x_t - \bar{x}_t) + (1 - \gamma)F_{t-s}\end{aligned}\tag{22}$$

com $0 < \alpha < 1$, $0 < \beta < 1$ e $0 < \gamma < 1$, constantes de suavização.

As previsões são dadas por:

$$\begin{aligned}f_{n,h} &= \bar{x}_n + hT_n + F_{t+h-s}; \quad h = 1, 2, 3, \dots, s. \\ f_{n,h} &= \bar{x}_n + hT_n + F_{t+h-2s}; \quad h = s + 1, s + 2, \dots, 2s. \\ &\vdots\end{aligned}\tag{23}$$

Considerando a sazonalidade multiplicativa, as equações para o nível, a inclinação e o factor de sazonalidade são, respectivamente:

$$\begin{aligned}\bar{x}_t &= \alpha\left(\frac{x_t}{F_{t-s}}\right) + (1 - \alpha)(\bar{x}_{t-1} + T_{t-1}) \\ T_t &= \beta(\bar{x}_t - \bar{x}_{t-1}) + (1 - \beta)T_{t-1} \\ F_t &= \gamma\left(\frac{x_t}{\bar{x}_t}\right) + (1 - \gamma)F_{t-s}\end{aligned}\tag{24}$$

onde: $0 < \alpha < 1$, $0 < \beta < 1$ e $0 < \gamma < 1$, constantes de suavização.

As previsões são dadas por:

$$f_{n,h} = (\bar{x}_n + hT_n)F_{t+h-s}; h = 1,2,3, \dots, s. \quad (25)$$

$$f_{n,h} = (\bar{x}_n + hT_n)F_{t+h-2s}; h = s + 1, s + 2, \dots, 2s.$$

⋮

Segundo Brockwell e Davis (1996), Os possíveis valores iniciais do algoritmo podem ser os seguintes:

$$\bar{x}_{s+1} = x_{s+1}$$

$$T_{s+1} = \frac{(x_{s+1} - x_1)}{s}$$

$$F_i = x_i - (x_i + T_{s+1}(i - 1)); i = 1,2, \dots, s.$$

Tanto para a sazonalidade aditiva como para a sazonalidade multiplicativa, os valores das constantes de suavização são calculados de forma a minimizar a soma dos quadrados dos erros de previsão.

As formas aditivas e multiplicativas do método de Holt-Winters fornecem previsões diferentes. Se a série apresentar oscilações sazonais aproximadamente constantes, o modelo aditivo é mais adequado, porém, se as oscilações sazonais forem proporcionais ao nível da série, o modelo multiplicativo é mais indicado. Alternativamente, pode-se utilizar os dois procedimentos e escolher aquele que fornece a menor soma dos quadrados dos erros de previsão (Cribari-Neto, 2000).

2.6. Métodos de avaliação de desempenho na previsão

Quando o objectivo da análise de séries temporais é de realizar previsão, depois da estimação do modelo é necessário avaliar o seu desempenho na previsão. Os métodos de avaliação de desempenho sugeridos na literatura utilizam os erros de previsão em seus cálculos, entre eles destacam-se os seguintes:

Raiz do erro quadrático médio (REQM)

$$REQM = \sqrt{\frac{1}{(h+1)} \sum_{t=s}^{s+h} (\hat{x}_t - x_t)^2} \quad (26)$$

Erro absoluto médio (EAM)

$$EAM = \frac{1}{(h+1)} \sum_{t=s}^{s+h} |\hat{x}_t - x_t| \quad (27)$$

Erro absoluto médio percentual (EAMP)

$$EAMP = 100 \frac{1}{h+1} \sum_{t=s}^{s+h} \frac{|\hat{x}_t - x_t|}{x_t} \quad (28)$$

Coefficiente de desigualdade de Theil (CDT).

$$CDT = \frac{REQM}{\sqrt{\frac{1}{(h+1)} \sum_{t=s}^{s+h} \hat{x}_t^2} + \sqrt{\frac{1}{(h+1)} \sum_{t=s}^{s+h} x_t^2}} \quad (29)$$

onde:

h é o número de períodos de previsão;

s é o tamanho da amostra;

\hat{x}_t é o IRPS previsto e x_t é o IRPS realizado.

As estatísticas $REQM$ e EAM apresentadas pelas fórmulas (26) e (27) respectivamente, dependem da escala de medida da variável dependente. Elas são usadas como medidas relativas para comparar a qualidade de previsão para a mesma série em diferentes modelos e quanto menor o erro melhor é a habilidade preditiva do modelo.

As estatísticas (28) e (29) são invariantes em relação a escala de medida da variável dependente. O CDT sempre situa-se entre 0 e 1, onde 0 indica um perfeito ajuste.

Os dois modelos (ARIMA e Ajustamento Exponencial) aqui apresentados, pelo facto de serem os mais aplicados para fazer previsão, serão usados para a previsão do IRPS, apurando-se aquele que verificar todos os pressupostos para a sua aplicação e apresentar menor erro de previsão.

III. MATERIAL E MÉTODOS

3.1. Material

Dados usados

Segundo AT, na Província do Niassa a cobrança do IRPS é efectuada em duas Direcções de Áreas Fiscais, uma situada na Cidade de Lichinga e a outra no Distrito de Cuamba.

Para a concretização dos objectivos do trabalho e de acordo com a modalidade de cobrança anteriormente expressa, foram usadas duas séries mensais correspondentes ao período que vai de Janeiro de 2005 à Novembro de 2009 o que corresponde a 59 observações para cada uma delas.

- Série de arrecadação do IRPS na Direcção de Área Fiscal de Cuamba que abreviadamente designamos por IRPSCuamba.
- Série de arrecadação do IRPS na Direcção de Área Fiscal de Lichinga que abreviadamente designamos por IRPSLichinga.

De notar que as arrecadações estão expressas em metical (unidade monetária de Moçambique).

Tamanho da amostra

O tamanho da amostra é considerado aceitável pois segundo Box e Jenkins (1976), para modelos ARIMA o tamanho mínimo deve ser de 50 observações. E segundo Newbold e Granger (1974), para modelos de Ajustamento Exponencial o tamanho deve ser superior a 30 observações.

Mesmo sabendo que se tem duas séries de dados, para evitar-se a repetição de procedimentos, somente a série do IRPSCuamba foi analisada detalhadamente, enquanto para a série do IRPSLichinga foram apresentados os resultados.

A base de dados foi obtida da AT, através da sua Divisão de Análise, Previsão e Controlo de Receitas, depois de um estágio de 3 meses.

Os dados foram processados com auxílio dos Softwares Estatísticos: Eviews versão 4.1, SPSS versão 17.0 e Microsoft Office Excel 2007.

3.2. Métodos

3.2.1. Análise descritiva

Outliers

A detecção de outlier é feita com auxílio de Boxplot e caso for considerado influente nos resultados é removido através do método sugerido por Venables e Ripley (1999), que consiste em substituí-lo pela média dos demais valores observados nos mesmos meses da sua ocorrência. O outlier é considerado influente nos resultados se o seu desvio em relação a média for superior a média da série.

Para sintetizar a informação contida nos dados, são calculados alguns indicadores estatísticos como, medidas de tendência central e de dispersão.

3.2.2. Análise de estacionaridade

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

Para séries temporais estacionárias, o correlograma da função de autocorrelação enfraquece rapidamente, enquanto para séries não estacionárias ele enfraquece gradualmente, ou seja, os coeficientes de autocorrelação nas várias defasagens são muito altos.

Ao nível formal, a estacionaridade pode ser testada descobrindo se a série tem ou não uma raiz unitária através da estatística de Dickey-Fuller Aumentado (ADF).

O teste da raiz unitária consiste em estimar a equação (30):

$$\Delta X_t = \beta_0 + \beta_1 t + \delta X_{t-1} + \sum_{i=1}^p \gamma_i \Delta X_{t-i} + \varepsilon_t \quad t = 1, 2, \dots \quad (30)$$

onde ε_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 para testar a existência de intercepto é denominada τ_μ . Enquanto para testar a presença de intercepto e de tendência é denominada τ_τ .

$$H_0: \delta = 0 \quad (31)$$

$$H_a: \delta < 0$$

A hipótese nula é rejeitada se o valor absoluto calculado da estatística tau ($\tau = \frac{\hat{\delta}}{s_{\hat{\delta}}}$) exceder o valor crítico nas estatísticas tau de Mackinnon, caso contrário, ela não é rejeita.

Todos os testes são realizados a um nível de significância de 5% ($\alpha = 0.05$). Para os outputs provenientes dos softwares usados neste trabalho, a hipótese nula é rejeitada se a probabilidade (sig.) associada ao teste for menor do que $\alpha = 0.05$, caso contrário não se rejeita a hipótese nula.

3.2.3. Análise da sazonalidade

A presença da componente sazonal é verificada pelo meio de análise gráfica e de seguida, para o mesmo propósito, é testada a significância estatística das arrecadações médias mensais do IRPS com o auxílio do teste F de Fisher da análise de variância cuja sua estatística, segundo Pestana e Gageiro (2005), é:

$$F_{(k-1; n-k)} = \frac{\frac{\sum_{i=1}^k n_i (\bar{Y}_i - \bar{Y}_{global})^2}{(k-1)}}{\frac{\sum_{i=1}^k \sum_{j=1}^n (Y_{ij} - \bar{Y}_i)^2}{(n-k)}} \quad (32)$$

onde, n_i é o número de elementos do grupo i ; \bar{Y}_i é a média do grupo i ; \bar{Y}_{global} é a média global; Y_{ij} é a observação j do grupo i , com $j = 1, 2, \dots, n$; k é o número de categorias do factor.

As hipóteses a testar são:

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k \quad (33)$$

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

sendo μ_1, \dots, μ_k as médias das arrecadações em k meses ($k = 1, 2, \dots, 12$).

Se a hipótese nula for rejeitada, podemos afirmar que existe pelo menos um mês que difere significativamente dos outros no valor médio arrecadado, ou seja, a série em estudo apresenta indício da sazonalidade.

Segundo Pestana e Gageiro (2005), antes de aplicar o teste F é necessário verificar os seguintes pressupostos:

1. As observações dentro de cada grupo têm distribuição normal;
2. As observações são independentes entre si;
3. As variâncias de cada grupo são iguais entre si, ou seja, há homocedasticidade.

Para a normalidade de uma distribuição recorre-se ao teste não paramétrico de aderência à normal de Kolmogorov-Smirnov (K-S), com a correção Lilliefors ou ao teste Shapiro Wilks, cujas hipóteses são:

$$H_0: X \sim N(\mu, \sigma^2) \quad (34)$$

$$H_a: X \not\sim N(\mu, \sigma^2)$$

A independência das observações é avaliada mediante o conhecimento da variável em estudo e da forma que os dados são observados (exemplo: dada a série de arrecadação diária de um dado imposto Moçambicano, as observações serão independentes se o valor arrecadado hoje não depender de nenhum outro anterior).

A homogeneidade das variâncias é testada através do teste de Levene com as seguintes hipóteses:

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_{12}^2 \quad (35)$$

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

3.2.4. Método de Box-Jenkins para modelos ARIMA

Identificação do modelo

A etapa da identificação consiste em analisar os correlogramas das funções de autocorrelação e autocorrelação parcial de modo a avaliar os seus comportamentos nas várias defasagens.

Estimação dos parâmetros

A estimação dos parâmetros dos modelos é feita pelo método da máxima verossimilhança, mediante a função de densidade de probabilidade da distribuição de Gumbel dada por:

$$f(x, \beta, \alpha) = \frac{1}{\alpha} \exp\left(-\frac{x-\beta_i}{\alpha}\right) * \exp\left(-\exp\left(-\frac{x-\beta_i}{\alpha}\right)\right) \quad (36)$$

Diagnóstico de verificação do modelo

Para testar a hipótese nula de que as primeiras k autocorrelações dos resíduos são nulas é usado o teste de Ljung-Box, cuja estatística do teste é dada pela equação (11).

Para a normalidade dos resíduos recorre-se ao teste não paramétrico de aderência à normal de Kolmogorov-Smirnov (K-S), ver a fórmula (34).

O teste t-Student é usado para verificar se a média dos resíduos é igual a zero e tem como hipóteses:

$$H_0: \mu_\varepsilon = 0 \quad (37)$$

$$H_a: \mu_\varepsilon \neq 0$$

A estatística do teste é:

$$t = \frac{\bar{X}_\varepsilon - \mu_\varepsilon}{\frac{S_\varepsilon}{\sqrt{n_\varepsilon}}} \quad (38)$$

O teste de Levene é usado para verificar se a variância dos resíduos é ou não constante, para a formulação das hipóteses ver a fórmula (35).

3.2.5. Critérios de selecção de Modelos ARIMA

Os Modelos Auto-Regressivos Integrados de Média Móvel são seleccionados pelo meio das equações (12) e (14).

3.2.6. Método de Ajustamento Exponencial

O método de ajustamento exponencial é considerado como um método ingénua na sua aplicação, na medida em que não exige muitas estratégias de modelagem, excepto a verificação da

existência das componentes de tendência e sazonalidade. O funcionamento deste método encontra-se descrito no capítulo da revisão da literatura.

3.2.7. Métodos de avaliação de desempenho na previsão

A avaliação do desempenho na previsão dos modelos de Ajustamento Exponencial e Auto-Regressivos Integrados de Média Móvel é feita através das fórmulas (26), (27), (28) e (29).

IV. ANÁLISE E DISCUSSÃO DOS RESULTADOS

4.1. Série de arrecadação do IRPS na Direcção de Área Fiscal de Cuamba

Estudo de outliers da série de IRPSCuamba

Ao se analisar o Boxplot da série de arrecadação do IRPS na Direcção de Área Fiscal (DAF) de Cuamba no período de Janeiro de 2005 à Novembro de 2009, apresentado à esquerda na Figura 1, constata-se que a observação 36 correspondente ao valor arrecadado no mês de Dezembro de 2007 é atípica.

O valor arrecadado naquele mês foi de 1,113,682.73 meticais e é superior a média em 2.36 vezes e tem um desvio maior do que a média da série. Assim, por seu desvio ser superior a média da série este outlier é considerado influente nos resultados e, será substituído pela média dos meses de Dezembro na série. A Figura 1 à direita mostra o Boxplot da série de IRPSCuamba depois da correcção do outlier.

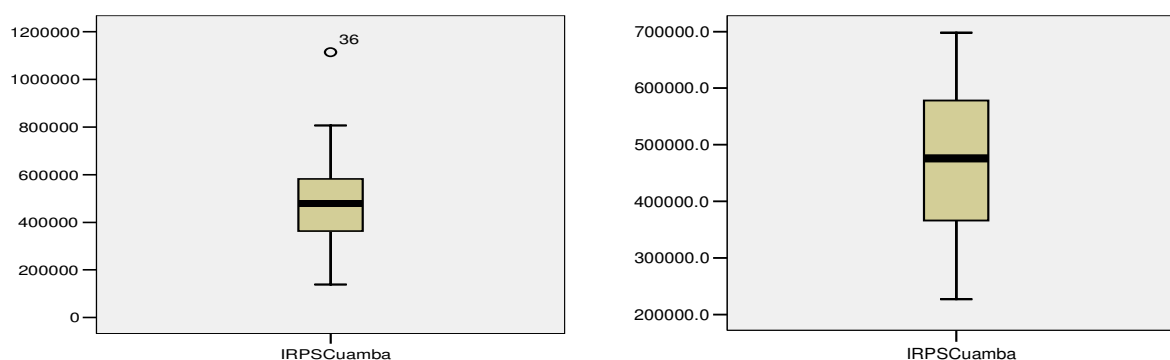


Figura 1: Boxplots da série de IRPSCuamba antes e depois da correcção do outlier

4.1.1. Análise exploratória da série de IRPSCuamba

As arrecadações do IRPS na Direcção de Área Fiscal de Cuamba no período de Janeiro de 2005 a Novembro de 2009, mostradas pela Figura 2, apresentam um comportamento que pode ser dividido em dois estágios. O primeiro estágio sugere uma tendência crescente das arrecadações e corresponde ao período de Janeiro de 2005 à Junho de 2007 e o segundo sugere uma tendência decrescente e compreende ao período de Julho de 2007 à Novembro de 2009. Este

comportamento da série também pode ser verificado a partir do gráfico de tendência da série (Figura 6).

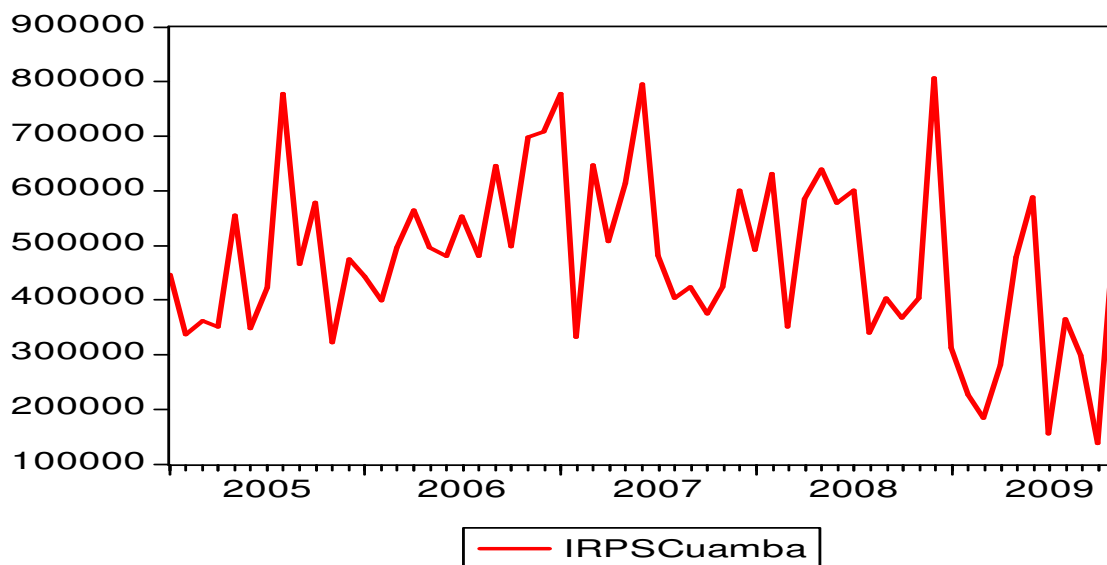


Figura 2: Representação da série de IRPSCuamba

Segundo Direcção de Área Fiscal de Cuamba, a tendência decrescente de arrecadação verificada a partir de Julho de 2007 é justificada pelo facto de algumas empresas terem cessado as suas actividades a partir desse ano. A Tabela 3, ilustra as empresas e as suas respectivas contribuições anuais.

Tabela 3: Relação das empresas que cessaram suas actividades a partir de 2007

Ano	Nome da empresa	Contribuição anual em meticais	Observação
2007	CLUSA	225,399.84	CESSOU
2007	IBIS	217,800.00	CESSOU
2007	INTERMON OXFAM	101,439,48	CESSOU
2007	UCASN	437,515.32	CESSOU
2008	TML	1,078,272.00	CESSOU
2008	PRATICAL ACTION PAMA	227,859.72	CESSOU
2009	OXFAM GB	862,416.00	CESSOU

Fonte: Direcção de Área Fiscal de Cuamba

De acordo com a Tabela 4, o valor médio mensal das arrecadações durante o período de Janeiro de 2005 à Novembro de 2009 foi de 474,826.54 meticais, com um desvio padrão de 155,153.13 meticais. As arrecadações máximas e mínimas foram verificadas em Dezembro de 2008 e Outubro de 2009 respectivamente.

Tabela 4: Estatísticas da série de arrecadação do IRPS na DAF de Cuamba

	N	Minimum	Maximum	Mean	Std. Deviation
IRPSCuamba	59	138132.3	806698.2	474826.54	155153.1259
Valid N (listwise)	59				

A Tabela 5 apresenta a receita do IRPS arrecadada anualmente, na Direcção de Área Fiscal de Cuamba, durante o período em estudo e as suas respectivas taxas de crescimento, sendo 2009 ano em que se registou o maior decréscimo de 32.66% relativamente ao ano anterior.

Tabela 5: Variação percentual e receitas anuais do IRPS na DAF de Cuamba

Ano	Receita	Varição%
2005	5,442,619.3	-----
2006	6,461,714.1	18.72%
2007	6,386,104.8	-1.17%
2008	6,198,352.4	-2.94%
2009	4,174,092.2	-32.66%

As probabilidades (sig.) associadas aos testes de aderência à normalidade, de Kolmogorov-Smirnov com a correcção de Lilliefors ou o Shapiro-Wilk apresentados na Tabela 6, são superiores a $\alpha = 0.05$ o que permite concluir que as distribuições das arrecadações da série IRPSCuamba dentro de cada mês são normais.

Tabela 6: Teste de normalidade da série do IRPSCuamba dentro de cada mês

	Meses	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
IRPSCuamba	Janeiro	.307	5	.139	.870	5	.265
	Fevereiro	.263	5	.200*	.893	5	.370
	Marco	.204	5	.200*	.976	5	.911
	Abril	.243	5	.200*	.885	5	.332
	Maiο	.203	5	.200*	.921	5	.536
	Junho	.225	5	.200*	.971	5	.882
	Julho	.254	5	.200*	.882	5	.318
	Agosto	.286	5	.200*	.800	5	.081
	Setembro	.243	5	.200*	.939	5	.660
	Outubro	.243	5	.200*	.945	5	.703
	Novembro	.224	5	.200*	.917	5	.508
	Dezembro	.164	4	.	.991	4	.964

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

a. Lilliefors Significance Correction

As dispersões dos valores arrecadados conforme o mês não são consideradas significativamente diferentes entre si, para $\alpha = 0.05$, uma vez que o teste de Levene tem associado a probabilidade de 0.9926 (Tabela 7). As dispersões conforme o mês podem ser vistas através da Tabela 2A em anexo A.

Tabela 7: Teste de homogeneidade das variâncias

IRPSCuamba			
Levene Statistic	df1	df2	Sig.
0.2427	11	47	0.9926

Com a verificação da normalidade e da homocedasticidade e pelo facto de as observações serem independentes umas das outras pode se prosseguir com o teste F.

O teste F com probabilidade ($sig. = 0.341$) maior do que $\alpha = 0.05$, permite concluir que não existe diferenças significativas entre as médias mensais da série IRPSCuamba (Tabela 8).

Tabela 8: Teste de igualdade das arrecadações médias mensais da série IRPSCuamba

IRPSCuamba					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	297574207585.50	11	27052200690	1.157	0.341
Within Groups	1098630356719.8	47	23375113973		
Total	1396204564305.3	58			

Deste modo, pode se concluir que apesar da série de IRPSCuamba mostrar oscilações em torno da média, não apresenta variações sazonais, uma vez que o teste F da Tabela 8 permite afirmar que não há diferenças significativas entre as médias mensais da série IRPSCuamba.

4.1.2. Teste de estacionaridade

A partir da Figura 6 constata-se que a série de IRPSCuamba apresenta duas tendências, uma crescente e a outra decrescente, que dão indícios da mesma não ser estacionária. Para uma conclusão fiável a série é submetida a um teste de raiz unitária.

De acordo com a Tabela 9, em termos absolutos, o valor da estatística de ADF (-1.754678) associado a probabilidade de 0.0650 é menor do que os valores críticos de Mackinnon, pelo que não se rejeita a hipótese nula, podendo se concluir que a série de IRPSCuamba não é estacionária em nível.

Tabela 9: Teste de raiz unitária de ADF para a série de IRPSCuamba em nível

	Estatística t	Prob.*
Estatística de teste do Dickey-Fuller Aumentado	-1.754678	0.0650
Valores críticos:	1% nível	-3.548007
	5% nível	-2.712478
	10% nível	-2.586380

*MacKinnon (1996)

Comprovada a não estacionaridade da série, efectuou-se a transformação da mesma em primeiras diferenças. Como ilustra a Figura 3, a série em primeiras diferenças estabiliza-se em torno de zero dando um indício de estacionaridade da mesma.

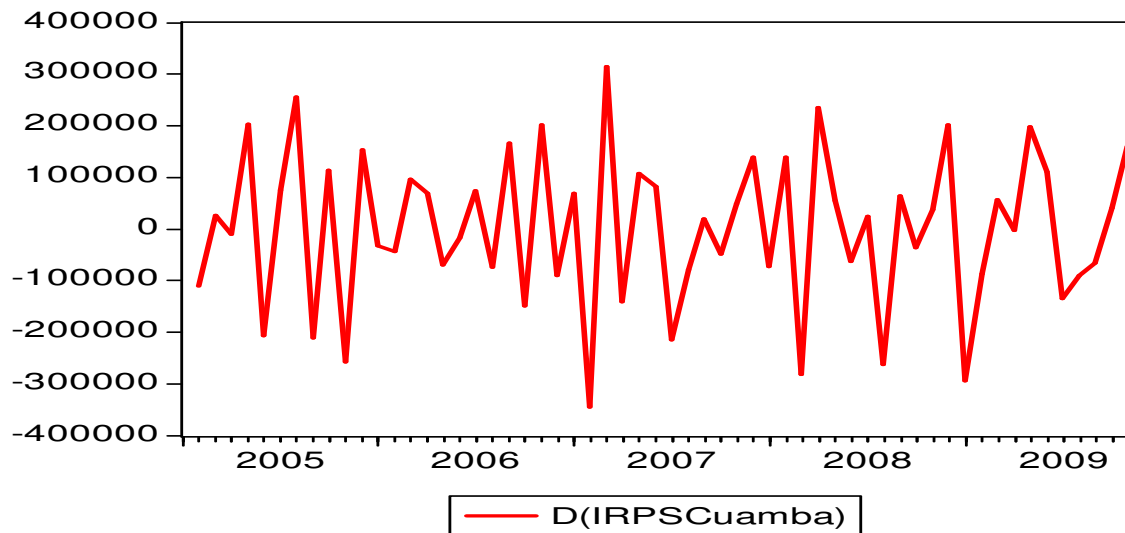


Figura 3: Representação da série de IRPSCuamba em primeiras diferenças

Conforme a Tabela 10, em termos absolutos, o valor da estatística de Dickey-Fuller Aumentado (-5.750324) associado a probabilidade de 0.0000 é maior do que os valores críticos de Mackinnon, pelo que se rejeita a hipótese nula, podendo se concluir que a série de IRPSCuamba é estacionária em primeiras diferenças ou integrada de ordem 1.

Tabela 10: Teste de raiz unitária de ADF para a série de IRPSCuamba em primeiras diferenças

	Estadística t	Prob.*
Estadística de teste de Dickey-Fuller Aumentado	-5.750324	0.0000
Valores críticos:		
1% nível	-3.548208	
5% nível	-2.912631	
10% nível	-2.594027	

*MacKinnon (1996)

4.1.3. Identificação dos Parâmetros

Ao se analisar os correlogramas das funções de autocorrelação e autocorrelação parcial da série de IRPSCuamba em primeiras diferenças, apresentados pela Figura 4, constata-se que a função de autocorrelação é significativa nas defasagens 1 e 27, assim o parâmetro q toma os valores 1 e 27. Por sua vez a função de autocorrelação parcial decai com padrão de onda senóide e apresenta picos significativos nas defasagens 1, 5 e 11, sendo estes os valores do parâmetro p .

Importa realçar que o parâmetro d é igual a 1 uma vez que a série é estacionária em primeiras diferenças.

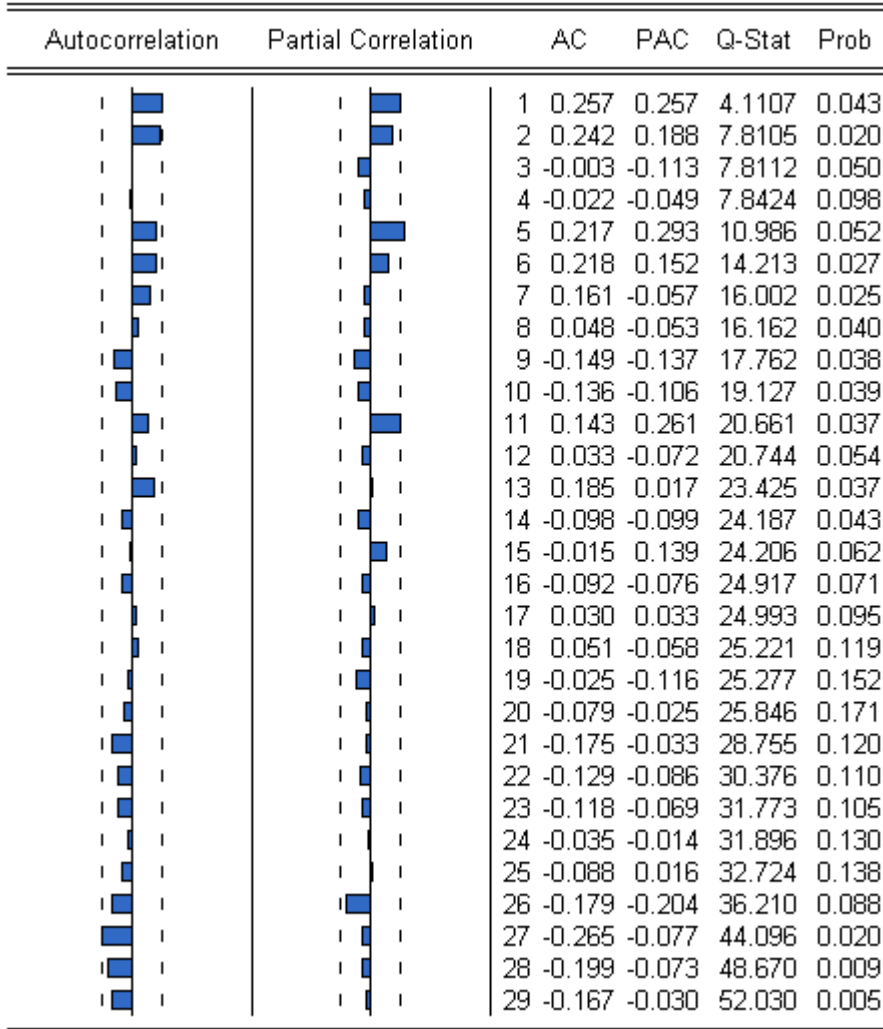


Figura 4: Correlogramas da série de IRPSCuamba em primeiras diferenças

4.1.4. Estimação dos Modelos ARIMA

A partir dos parâmetros p, d e q identificados anteriormente e dos padrões teóricos das funções de autocorrelação e autocorrelação parcial para modelos ARIMA apresentados na Tabela 2, foram estimados os modelos $ARIMA(1,1,27)$ e $ARIMA(0,1,27)$, todos estes com interceptos, conforme ilustram as Tabelas 11 e Tabela 12 respectivamente.

A apresentação de somente dois modelos, já que a partir dos parâmetros identificados vários modelos poderiam ser estimados, deveu-se ao facto de estes serem os únicos que apresentaram todos os seus coeficientes significativos a um nível de significância de 5%, valores de Durbin-Watson próximos de 2 indicando a não presença de auto correlação serial e coeficientes de determinação (R-quadrado) superiores a 50%.

Tabela 11: Estimação do Modelo $ARIMA(1,1,27)$ da série de IRPSCuamba

Variável	Coefficiente	Desv. Padrão	Estatística t	Prob.
Constante	454051.5	23961.73	18.94903	0.0000
AR(1)	0.360872	0.125202	2.882316	0.0056
MA(27)	-0.896509	0.028882	-31.04017	0.0000
R-quadrado	0.670347	V. Média dependente		474666.2
R-quadrado ajustado	0.658359	S.D. Variância dependente		122635.3
S.E. da regressão	71680.36	C. Informação de Akaike		25.24816
S. Q. dos Erros	2.83E+11	Critério de Schwarz		25.35473
Log M. verossimilhança	-729.1966	Estat. de Durbin-Watson		1.994844

Tabela 12: Estimação do Modelo $ARIMA(0,1,27)$ da série de IRPSCuamba

Variável	Coefficiente	Desv. Padrão	Estatística t	Prob.
Constante	458868.5	15578.65	29.45497	0.0000
MA(27)	-0.885922	0.032549	-27.21842	0.0000
R-quadrado	0.619536	V. Média dependente		474191.2
R-quadrado ajustado	0.612861	S.D. Variância dependente		121628.2
S.E. da regressão	75677.65	C. Informação de Akaike		25.33966
S. Q. dos Erros	3.26E+11	Critério de Schwarz		25.41009
Log M. verossimilhança	-745.5201	Estat. de Durbin-Watson		2.273422

4.1.5. Diagnóstico de verificação do Modelo

Modelo $ARIMA(1, 1, 27)$

Ao se examinarem as funções de autocorrelação e autocorrelação parcial da série dos resíduos do modelo $ARIMA(1,1,27)$ apresentados pela Figura 5, a um nível de significância de 5%, verifica-se que nenhuma das autocorrelações geral e parcial é, individualmente, significativa. As probabilidades (sig.) associadas ao teste de Ljung-Box, como mostra a Tabela 3A em anexo A,

são superiores a $\alpha = 0.05$ pelo que não se rejeita a hipótese nula de que as primeiras 16 autocorrelações dos resíduos do modelo $ARIMA(1,1,27)$ são nulas. Em outras palavras, os correlogramas das autocorrelações geral e parcial dão a impressão de que os resíduos estimados a partir do modelo $ARIMA(1,1,27)$ são puramente aleatórios.

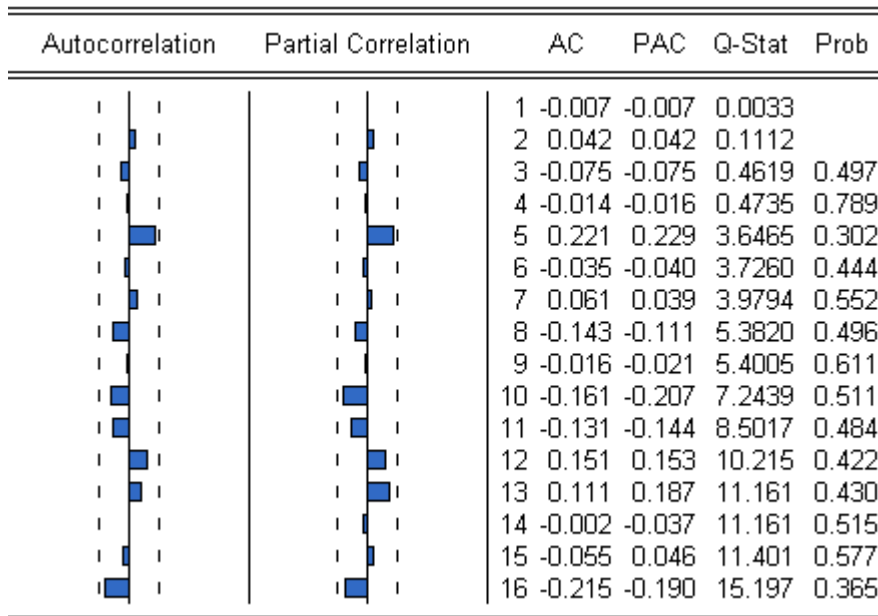


Figura 5: Correlogramas dos resíduos obtidos do modelo $ARIMA(1,1,27)$

A probabilidade ($sig. = 0.64$) associada ao teste de aderência à normalidade, Kolmogorov-Smirnov com a correcção de Lilliefors apresentado na Tabela 13, é superior a $\alpha = 0.05$ pelo que não se rejeita a hipótese nula de que a distribuição dos resíduos do modelo $ARIMA(1,1,27)$ é normal.

Tabela 13: Teste de normalidade aos resíduos do modelo $ARIMA(1,1,27)$

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
resíduo $ARIMA(1,1, 27)$.114	57	.064	.966	57	.107

a. Lilliefors Significance Correction

A estatística $t = -0.429$ com a probabilidade de 0.670, mostra que para $\alpha = 0.05$, não se rejeita a hipótese nula, inferindo-se que a média dos resíduos do modelo $ARIMA(1,1,27)$ é igual a zero

(Tabela 14). De facto, o zero está incluído no intervalo de confiança a 95%, o que confirma a não rejeição da hipótese nula.

Tabela 14: Teste t da média dos resíduos do modelo $ARIMA(1,1,27)$

Test Value = 0						
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
residuo ARIMA(1,1, 27)	-.429	56	.670	-333.9145	-1564.599	1281.770

As variâncias dos resíduos do modelo $ARIMA(1,1,27)$ de acordo com o mês não são consideradas significativamente diferentes entre si, para $\alpha = 0.05$, uma vez que o teste de Levene tem associado a probabilidade de 0.332 (Tabela 15).

Tabela 15: Teste de variância dos resíduos do modelo $ARIMA(1,1,27)$

residuo ARIMA(1,1, 27)			
Levene Statistic	df1	df2	Sig.
1.168	11	45	0.332

Modelo $ARIMA(0, 1, 27)$

As probabilidades (sig.) associadas ao teste de Ljung-Box, como mostra a Tabela 16, são superiores a $\alpha = 0.05$ pelo que não se rejeita a hipótese nula de que as primeiras 16 autocorrelações dos resíduos do modelo $ARIMA(0,1,27)$ não diferem de zero, ou seja, os resíduos estimados a partir do modelo $ARIMA(0,1,27)$ são puramente aleatórios.

Tabela 16: Teste de Ljung-Box aos resíduos do modelo $ARIMA(0,1,27)$

Series: residuo ARIMA(0,1, 27)

Lag	Autocorrelation	Std.Error ^a	Box-Ljung Statistic	
			df	Sig. ^b
1	-.032	.142	1	.851
2	.052	.142	2	.846
3	-.063	.143	3	.867
4	-.014	.143	4	.906
5	.231	.143	5	.711
6	-.036	.152	6	.715
7	.071	.160	7	.762
8	-.163	.164	8	.776
9	-.016	.163	9	.709
10	-.181	.163	10	.612
11	-.141	.166	11	.688
12	.161	.168	12	.797
13	.131	.171	13	.597
14	-.004	.172	14	.573
15	-.065	.172	15	.600
16	-.245	.173	16	.540

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.

As variâncias dos resíduos do modelo $ARIMA(0,1,27)$ de acordo com os grupos formados a partir dos meses não são consideradas significativamente diferentes entre si, para $\alpha = 0.05$, uma vez que o teste de Levene tem associado a probabilidade de 0.336 (Tabela 17).

Tabela 17: Teste de variância dos resíduos do modelo $ARIMA(0,1,27)$

residuo ARIMA(0,1, 27)

Levene Statistic	df1	df2	Sig.
1.032	11	45	.336

A probabilidade ($sig. = 0.055$) associada ao teste de aderência à normalidade, Kolmogorov-Smirnov com a correção de Lilliefors apresentado na Tabela 18, é superior a $\alpha = 0.05$ pelo que não se rejeita a hipótese nula de que a distribuição dos resíduos do modelo $ARIMA(0,1,27)$ é normal.

Tabela 18: Teste de normalidade aos resíduos do modelo $ARIMA(0,1,27)$

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
residuo $ARIMA(0,1, 27)$.103	57	.055	.902	57	.114

a. Lilliefors Significance Correction

A estatística $t = -0.326$ com a probabilidade de 0.502, mostra que para $\alpha = 0.05$, não se rejeita a hipótese nula, inferindo-se que a média dos resíduos do modelo $ARIMA(0,1,27)$ é igual a zero (Tabela 19). De facto, o zero está incluído no intervalo de confiança a 95%, o que confirma a não rejeição da hipótese nula.

Tabela 19: Teste t da média dos resíduos do modelo $ARIMA(0,1,27)$

	Test Value = 0					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
residuo $ARIMA(0,1, 27)$	-.326	56	.502	-383.7651	-2436.473	1381.410

4.1.6. Selecção do Modelo $ARIMA$ da série de IRPSCuamba

As análises feitas Anteriormente aos resíduos dos modelos $ARIMA(1,1,27)$ e $ARIMA(0,1,27)$, demonstram de que os dois modelos ajustam-se aos dados. Mas para efeitos de previsão somente um modelo tem que ser escolhido pelo meio dos critérios de selecção de $SBCC$ e AIC .

Como mostra a Tabela 20, o modelo seleccionado foi o $ARIMA(1,1,27)$ pois apresenta os seus valores de $SBCC$ e AIC inferiores aos do modelo $ARIMA(0,1,27)$. Este modelo será comparado com o modelo estimado pelo método de Ajustamento exponencial, através das suas estatísticas de desempenho, a fim de se apurar o modelo final que servirá para fazer previsões das arrecadações de IRPS na Direcção de Área Fiscal de Cuamba.

Tabela 20: Selecção do Modelo ARIMA da série de IRPSCuamba

Modelo	Critério de selecção	
	<i>SBCC</i>	<i>AIC</i>
<i>ARIMA(1,1,27)</i>	25.35473	25.24816
<i>ARIMA(0,1,27)</i>	25.41009	25.33966

4.1.7. Estimação do modelo de Ajustamento Exponencial para a série de IRPSCuamba

A série de arrecadação do IRPS na Direcção de Área Fiscal de Cuamba no período de Janeiro de 2005 à Novembro de 2009, como mostra a Figura 6, apresenta duas tendências, a primeira crescente que corresponde ao período de Janeiro de 2005 a Junho de 2007 (30 observações) e a segunda decrescente que compreende ao período de Julho de 2007 à Novembro de 2009 (29 observações).

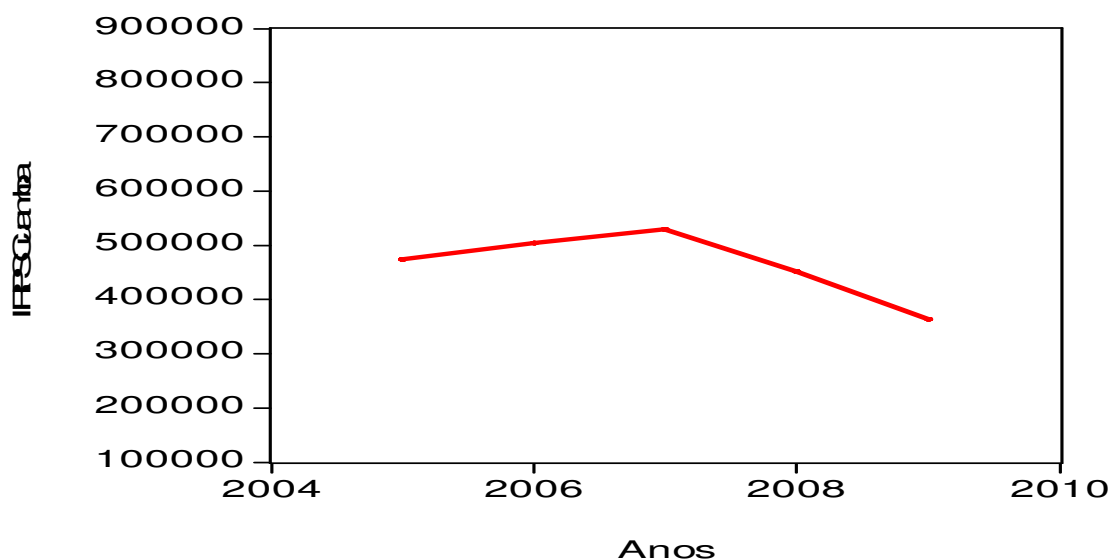


Figura 6: Tendência da série de IRPSCuamba

Para que o modelo a ser estimado se ajustasse aos dados, a série de IRPSCuamba foi dividida em duas de acordo com as tendências. Feita a divisão se constatou que as duas séries resultantes não cumpriam com o pressuposto de tamanho de amostra para a aplicação do método de Ajustamento exponencial pelo que, este método, não foi adequado para descrever a série de IRPSCuamba.

Finalmente e para todos os efeitos o modelo apurado para fazer previsões das arrecadações de IRPS na Direcção de Área Fiscal de Cuamba é *ARIMA(1,1,27)*.

4.1.8. Previsão do IRPSCuamba

Para se atingir o objectivo proposto neste trabalho, a partir do modelo *ARIMA(1,1,27)* apurado, realizaram-se previsões de IRPS a ser arrecadado pela Direcção de área Fiscal de Cuamba, no primeiro trimestre de 2010 incluindo o mês de Dezembro de 2009, cujos valores previstos e os respectivos intervalos de confiança estão dispostos na Tabela 21 incluindo os valores realizados e os respectivos erros de previsão. As estatísticas de desempenho na previsão dentro da amostra encontram-se na Tabela 22.

Tabela 21: Valores previstos e realizados do IRPS mensal para a DAF de Cuamba

Mês/Ano	Limite inferior (Mt)	Previstos	Limite superior (Mt)	Realizados	Erro de previsão
Dezembro/09	258,232.29	513,903.87	769,575.45	510,383.62	-3,520.25
Janeiro/10	268,196.67	530,965.90	793,735.13	521,674.45	-9,291.45
Fevereiro/10	160,835.99	448,920.95	737,005.92	456,393.76	7,472.81
Março/10	67,835.57	365,075.97	662,316.37	400,432.34	35,356.37

Tabela 22: Estatísticas de desempenho na previsão do modelo *ARIMA(1,1,27)*

Estatística	Valor
REQM	69801.94
EAM	55551.74
EAMP	11.96470
CDT	0.072076

4.2. Série de arrecadação do IRPS na Direcção de Área Fiscal de Lichinga

4.2.1. Análise exploratória da série de IRPSLichinga

As arrecadações do IRPS na Direcção de Área Fiscal de Lichinga no período de Janeiro de 2005 à Novembro de 2009, mostradas pela Figura 7, sugerem uma tendência crescente ao longo dos cinco anos em análise. Esta tendência da série também pode ser verificada a partir da Figura 9.

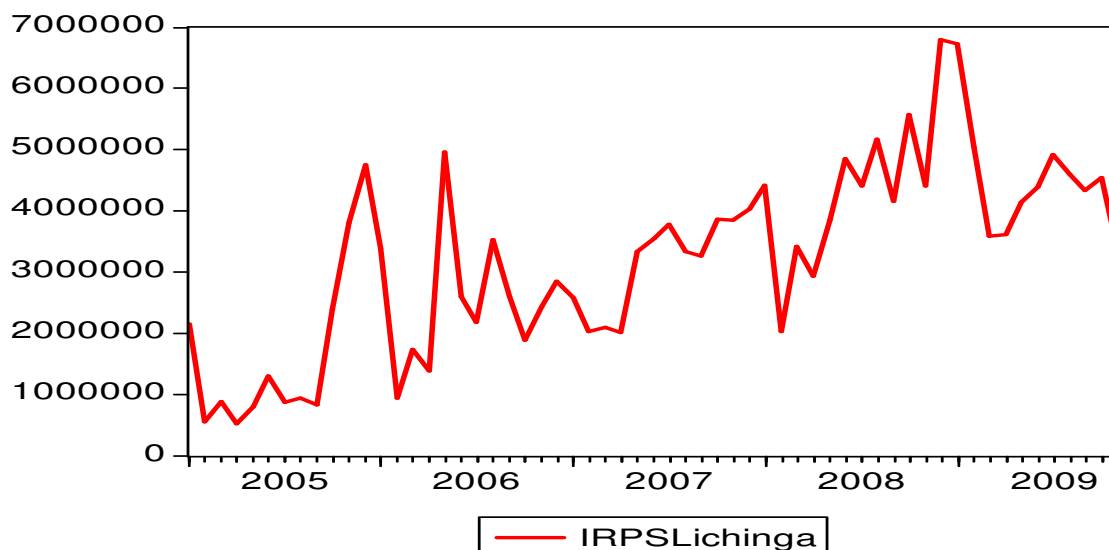


Figura 7: Representação da série de IRPSLichinga

De acordo com a Tabela 23, o valor médio mensal das arrecadações durante o período de Janeiro de 2005 à Novembro de 2009 foi de 3,207,544.3 meticais, com um desvio padrão de 1,514,741.59 meticais. As arrecadações mínimas e máximas foram verificadas em Abril de 2005 e Dezembro de 2008 respectivamente.

Tabela 23: Estatísticas da série de arrecadação do IRPS na DAF de Lichinga

	N	Minimum	Maximum	Mean	Std. Deviation
IRPSLichinga	59	528743.6	6799587.0	3207544.3	1514741.5926
Valid N (listwise)	59				

A Tabela 24 apresenta a receita do IRPS arrecadada anualmente, na DAF de Lichinga, e as suas respectivas taxas de crescimento, sendo 2006 ano em que se registou o maior crescimento durante os cinco anos em estudo de 54.13% relativamente ao ano anterior.

Tabela 24: Taxa de crescimento anual e total das arrecadações de IRPS na DAF de Lichinga

Ano	Arrecadação	Crescimento%
2005	19,763,547.3	-----
2006	30,462,203.5	54.13%
2007	37,713,162.7	23.80%
2008	52,008,692.3	37.91%
2009	53,904,628.3	3.65%

A série de IRPSLichinga não é estacionária em nível mas sim em primeiras diferenças como mostram as Tabelas 25 e 26 e a Figura 8.

Tabela 25: Teste de raiz unitária de ADF para a série de IRPSLichinga em nível

	Estatística t	Prob.*
Estatística de teste de Dickey-Fuller Aumentado	-2.436873	0.0735
Valores críticos:	1% level	-3.548208
	5% level	-2.912631
	10% level	-2.594027

*MacKinnon (1996)

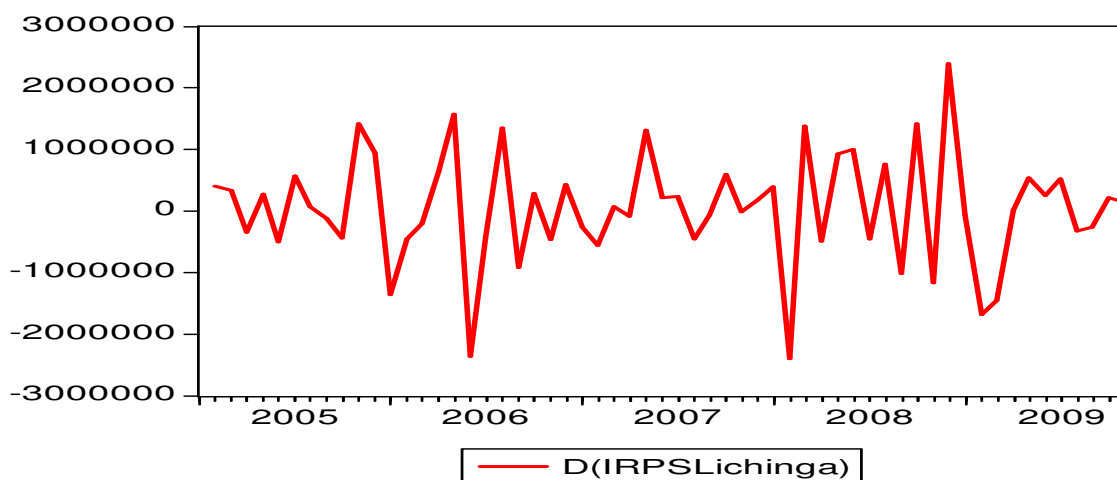


Figura 8: Representação da série de IRPSLichinga em primeiras diferenças

Tabela 26: Teste de raiz unitária de ADF para a série de IRPSLichinga em primeiras diferenças

	Estatística t	Prob.*
Estatística de teste de Dickey-Fuller Aumentado	-10.03638	0.0000
Valores críticos:	1% nível	-3.550396
	5% nível	-2.913549
	10% nível	-2.594521

*MacKinnon (1996)

4.2.2. Estimação do modelo *ARIMA(1, 1, 21)*

Tabela 27: Estimação do Modelo *ARIMA(1,1,21)* da série de IRPSLichinga

Variável	Coefficiente	Desv. Padrão	Estatística t	Prob.
AR(1)	-0.408876	0.123638	-3.307034	0.0017
MA(21)	-0.853608	0.034952	-24.42242	0.0000
R-quadrado	0.564548	V. Média dependente		37641.41
R-quadrado ajustado	0.556631	S.D. Variância dependente		919009.6
S.E. da regressão	611931.6	C. Informação de Akaike		29.52109
S. Q. dos Erros	2.06E+13	Critério de Schwarz		29.59277
Log M. verossimilhança	-839.3509	Estat. de Durbin-Watson		2.003018

4.2.3. Estimação do modelo de Ajustamento Exponencial para a série de IRPSLichinga

A série de IRPSLichinga apresenta uma tendência crescente (Figura 9) e, não apresenta variações sazonais (Tabela 2B em anexo B).

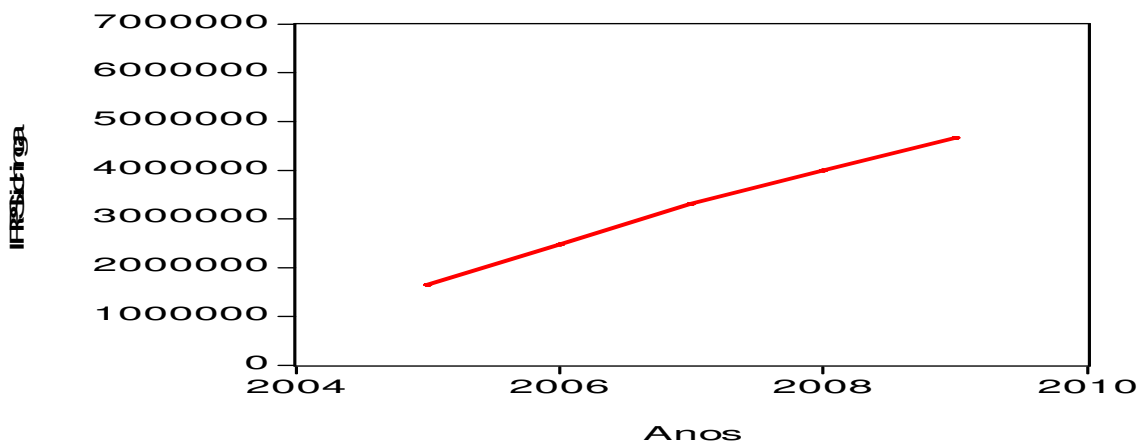


Figura 9: Tendência da série das arrecadações de IRPS na DAF de Lichinga

O método adequado para séries que apresentam estas características é o método de Ajustamento exponencial de $Holt(\alpha, \beta)$ que leva em consideração essa tendência e permite estimar também a inclinação actual da série.

Tabela 28: Constantes de suavização do modelo de $Holt(\alpha, \beta)$

Série	α	β	SQR_{ε}	gl dos erros
IRPSLichinga	0.4000	0.0200	4.1E+13	57

Assim, o nível e a inclinação da série são, respectivamente:

$$\bar{x}_t = 0.4x_t + 0.6(\bar{x}_{t-1} + T_{t-1})$$

$$T_t = 0.02(\bar{x}_t - \bar{x}_{t-1}) + 0.98T_{t-1}$$

Os valores iniciais são:

$$\bar{x}_2 = 2550397.6$$

$$T_2 = 403891.9$$

Os futuros valores do modelo $Holt(0.4,0.02)$ são estimados a partir da seguinte equação:

$$f_{n,h} = \bar{x}_n + hT_n; h = 3,4,5, \dots, n.$$

4.2.4. Selecção do modelo de previsão das arrecadações de IRPS na DAF de Lichinga

De acordo com a Tabela 29, as estatísticas de desempenho na previsão do modelo $ARIMA(1,1,21)$ apresentam todos os seus valores menores do que os do modelo $Holt(0.4,0.02)$ pelo que foi o modelo seleccionado para fazer previsões das arrecadações de IRPS na Direcção de Área Fiscal de Lichinga.

Tabela 29: Estatísticas de desempenho na previsão dos Modelos $ARIMA(1,1,21)$ e $Holt(0.4,0.02)$

Estatística	$ARIMA(1,1,21)$	$Holt(0.4,0.02)$
REQM	601100.1	822593.7
EAM	472403.8	616155.5
EAMP	14.21984	18.20225
CDT	0.079688	0.111135

4.2.5. Previsão do IRPSLichinga

A partir do modelo $ARIMA(1,1,21)$ apurado, realizaram-se previsões para o primeiro trimestre de 2010 incluindo o mês Dezembro de 2009, cujos valores estimados e os respectivos intervalos de confiança estão dispostos na Tabela 30 incluindo os valores realizados e os correspondentes erros de previsão.

Tabela 30: Valores previstos e realizados do IRPS mensal para a DAF de Lichinga

Mês/Ano	Limite inferior (Mt)	Previstos	Limite superior (Mt)	Realizados	Erro de previsão
Dezembro/09	3,529,016.83	5,049,942.85	6,570,868.88	4,904,325.18	-145,617.67
Janeiro/10	4,177,142.00	5,919,145.70	7,661,149.43	6,027,438.34	108,292.64
Fevereiro/10	3,462,616.04	5,314,064.05	7,165,512.05	5,332,167.92	18,103.87
Março/10	3,061,491.54	4,994,505.14	6,927,518.73	5,045,318.37	50,813.23

V. CONCLUSÕES E RECOMENDAÇÕES

5.1. Conclusões

Neste trabalho foram desenvolvidos dois modelos de previsão para a arrecadação do Imposto sobre o Rendimento das Pessoas Singulares – IRPS nas Direcções de Área Fiscal de Cuamba e Lichinga cujo somatório corresponde ao IRPS da Província do Niassa. As duas séries em estudo correspondem ao período de Janeiro de 2005 a Novembro de 2009.

- A série de arrecadação do IRPS na Direcção de Área Fiscal de Cuamba, apresentou duas tendências. A primeira tendência de arrecadação é crescente e corresponde ao período de Janeiro de 2005 a Junho de 2007 e a segunda é decrescente e compreende ao período de Julho de 2007 a Novembro de 2009. Segundo a Direcção de Área Fiscal de Cuamba, a tendência decrescente de arrecadação verificada a partir de Julho de 2007 é justificada pelo facto de algumas empresas terem cessado as suas actividades a partir desse ano.
- A série de arrecadação do IRPS na Direcção de Área Fiscal de Cuamba não apresentou variações sazonais, não foi estacionária em nível, mas sim em primeiras diferenças e nem cumpriu com os pressupostos para a aplicação dos métodos de Ajustamento Exponencial. O método adequado para esta série foi o de Box-Jenkins e, o modelo apurado para fazer previsão da arrecadação do IRPS na Direcção de Área Fiscal de Cuamba para o primeiro trimestre de 2010 foi $ARIMA(1,1,27)$ por ter este apresentado melhor desempenho na previsão de ponto de vista estatístico.
- A série de arrecadação do IRPS na Direcção de Área Fiscal de Lichinga apresentou uma tendência crescente ao longo dos 5 anos em estudo. A série não apresentou variações sazonais e nem foi estacionária em nível, mas sim em primeiras diferenças. Esta Cumpriu com os pressupostos para a aplicação tanto dos métodos de Ajustamento Exponencial como de Box-Jenkins e, o modelo apurado, para fazer previsão da arrecadação do IRPS na Direcção de Área Fiscal de Lichinga para o primeiro trimestre de 2010 foi $ARIMA(1,1,21)$ por apresentar as mesmas razões do modelo anterior.

5.2. Recomendações

- O método de Box-Jenkins produziu melhores resultados em relação ao método de Ajustamento exponencial, recomendando-se a sua aplicação para a previsão do IRPS das restantes Províncias de Moçambique.
- Recomenda-se para futuros trabalhos a inclusão de variáveis explicativas independentes que influenciam na arrecadação do IRPS e/ou a combinação do método de Box-Jenkins com o método de Ajustamento Exponencial. Essa combinação pode ser pensada como uma combinação linear entre os dois métodos, de forma que os pesos associados aos métodos possam ser estimados.

REFERÊNCIAS BIBLIOGRÁFICAS

- Associação Comercial e Industrial de Sofala – ACIS (2008). Quadro Legal para Impostos em Moçambique, 1ª edição. 58 pp. Sofala.
- Autoridade Tributária de Moçambique (2008). Estatuto Orgânico da Autoridade Tributária de Moçambique. Revisto.
- Box, G. E. P. e G. M. Jenkins (1976). Time Series Analysis. Forecasting and Control. 1ª edição. New Jersey, Prentice Hall.
- Brockwell, P.J. e R.A. Davis (1996). Introduction to Time Series and Forecasting. New York, Springer-Verlag.
- Cirincione, C., G.A. Gurrieri e B. Van de Sande (1999). Municipal Government Revenue Forecasting. Issues of Method and Data. pp 26-46, Public Budgeting and Finance.
- Código do Imposto sobre o Rendimento das Pessoas Singulares – CIRPS. (2002). 60 pp. Moçambique.
- Cribari-Neto, F. (2000). Método de Previsão de Arrecadação Tributária. MIMEO.
- 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 and Sons.
- Frank, H.A. (1993). Budgetary Forecasting in Local Government. New Tools and Techniques. Westport, CT. Quorum Books.
- Guajardo, S.A. e R. Miranda (2000). An Elected Official's Guide to Revenue Forecasting. Chicago, Government Finance Officers Association.
- Gujarati, D.N. (2006). Econometria Básica, 4ª edição. São Paulo, Editora Campus.
- Hamilton, J. D. (1994). Time Series Analysis. Princeton University Press.
- Janacek, G. (2001). Practical Time Series. London, Arnold Publishers.

- Makridakis, S., A. Andersen, R. Carbone, R. Fildes, M. Hibon, R. Lewandowski, J. Newton e R. Winkler (1984) The Forecasting Accuracy of Major Time Series Methods. New York, John Wiley and Sons.
- Makridakis, S. e S.C. Wheelwright (1989). Forecasting Methods for Management. 5ª edição. New York, John Wiley and Sons.
- Martins, P. G. e F. P. Laugeni (1999). Administração da Produção. São Paulo, Editora Saraiva.
- Mills, T.C. (1990). Time Series Techniques for Economists. Cambridge, Cambridge University Press.
- Montoro, F. e A. Franco (2002). Manual de Economia, 3ª edição. São Paulo. Editora Saraiva.
- Morettin, P. A. e C.M.C. Toloi (2004). Análise de Séries Temporais. São Paulo, Edgar Blucher.
- Murteira, B.J.F. e G.h.J. Black (1983). Estatísticas Descritiva. Portugal, Editora McGraw-Hill.
- Newbold, P. e C.W.J. Granger (1974). Experience with Forecasting Univariate Time Series and the Combination of Forecasts. Journal of the Royal Statistics Society A, pp 131-146.
- Pestana, M.H. e J.N. Gageiro (2005). Análise de Dados para Ciências Sociais. A Complementaridade do SPSS, 4ª Edição. Lisboa, Portugal, Edições Sílabo.
- Parzen, E. (1972). Processos Estocásticos. Departamento de Estatística. Universidade de Stanford. Madrid, Espanha.
- Soares, J. F. (1991). Introdução à Estatística. Rio de Janeiro, Guanaba.
- Spiegel, M.R. (1985). Estatística. Brasil, Editora McGraw-Hill.
- Venables, W.N. e B.D. Ripley (1999). Modern Applied Statistics with S-PLUS, 3ª Edição. New York, Springer -Verlag.
- Wallis, K. F. e J. J. Thomas (1971). Seasonal Variation in Regression Analysis. Journal of the Royal Statistical Society, Volume 134. pp 57-72.

ANEXO A: Série de arrecadação do IRPS na Direcção de Área Fiscal de Cuamba

Tabela 1A: Cobranças do IRPS na Direcção de Área Fiscal de Cuamba

Mês	2005	2006	2007	2008	2009
Janeiro	446,639.9	443,662.5	777,521.4	491,574.8	313,390.6
Fevereiro	336,577.7	399,476.6	332,565.5	631,300.8	226,900.1
Março	362,199.1	495,738.1	647,639.8	350,641.1	183,480.3
Abril	351,167.8	564,976.3	507,031.5	584,914.1	280,339.5
Mai	554,388.1	496,191.3	614,476.7	639,924.4	478,270.3
Junho	347,498.5	479,398.9	796,188.5	577,081.6	589,002.9
Julho	422,627.4	553,166.7	481,956.1	601,408.5	154,370.6
Agosto	777,718.7	479,433.1	403,685.4	339,265.5	365,113.1
Setembro	465,785.2	646,171.8	423,863.7	403,822.5	298,186.3
Outubro	579,613.1	497,031.1	375,435.5	367,389.2	138,132.3
Novembro	322,352.4	697,981.2	424,510.8	404,332.1	498,789.8
Dezembro	476,051.8	708,486.8	1,113,683	806,698.2	-----

Fonte: Autoridade Tributária de Moçambique

Tabela 2A: Estatísticas da Série IRPSCuamba de acordo com os grupos de meses

IRPSCuamba							
	N	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower Bound	Upper Bound		
Janeiro	5	494557.84	171599.3306	281489.127	707626.549	313390.6	777521.4
Fevereiro	5	385364.15	150800.0844	198121.117	572607.189	226900.1	631300.8
Marco	5	407939.54	173869.3577	192052.218	623826.862	183480.0	647639.8
Abril	5	457685.83	135028.7691	290025.472	625346.195	280339.5	584914.0
Mai	5	556650.08	70856.1058	468670.602	644629.554	478270.3	639924.4
Junho	5	557834.07	164733.0765	353290.938	762377.199	347498.5	796188.5
Julho	5	442705.83	174977.9007	225442.068	659969.588	154370.6	601408.5
Agosto	5	473043.17	178342.2948	251601.965	694484.379	339265.5	777718.7
Setembro	5	447565.85	127048.9685	289813.719	605317.980	298186.3	646171.8
Outubro	5	391520.18	167001.4529	184160.485	598879.869	138132.3	579613.0
Novembro	5	469593.24	142289.1311	292917.946	646268.528	322352.4	697981.2
Dezembro	4	648116.69	142123.5840	421966.348	874267.023	476051.8	806698.2
Total	59	474826.54	155153.1259	434393.411	515259.661	138132.3	806698.2

Tabela 3A: Teste de Ljung-Box aos resíduos do modelo $ARIMA(1,1,27)$

Series: residuo ARIMA(1,1, 27)

Lag	Autocorrelation	Std.Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	-.007	.132	.003	1	.954
2	.042	.132	.111	2	.946
3	-.075	.133	.462	3	.927
4	-.014	.133	.473	4	.976
5	.221	.133	3.646	5	.601
6	-.035	.140	3.726	6	.714
7	.061	.140	3.979	7	.782
8	-.143	.140	5.382	8	.716
9	-.016	.143	5.401	9	.798
10	-.161	.143	7.244	10	.702
11	-.131	.146	8.502	11	.668
12	.151	.148	10.215	12	.597
13	.111	.151	11.161	13	.597
14	-.002	.152	11.161	14	.673
15	-.055	.152	11.401	15	.724
16	-.215	.153	15.197	16	.510

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.

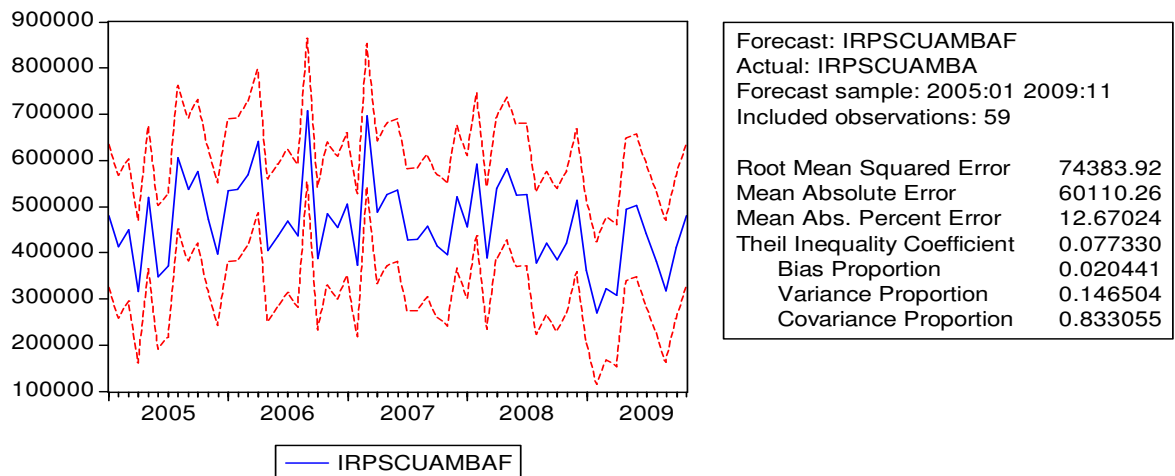


Figura 1A: Previsão dentro da amostra através do modelo $ARIMA(0,1,27)$ e estatísticas de desempenho na previsão

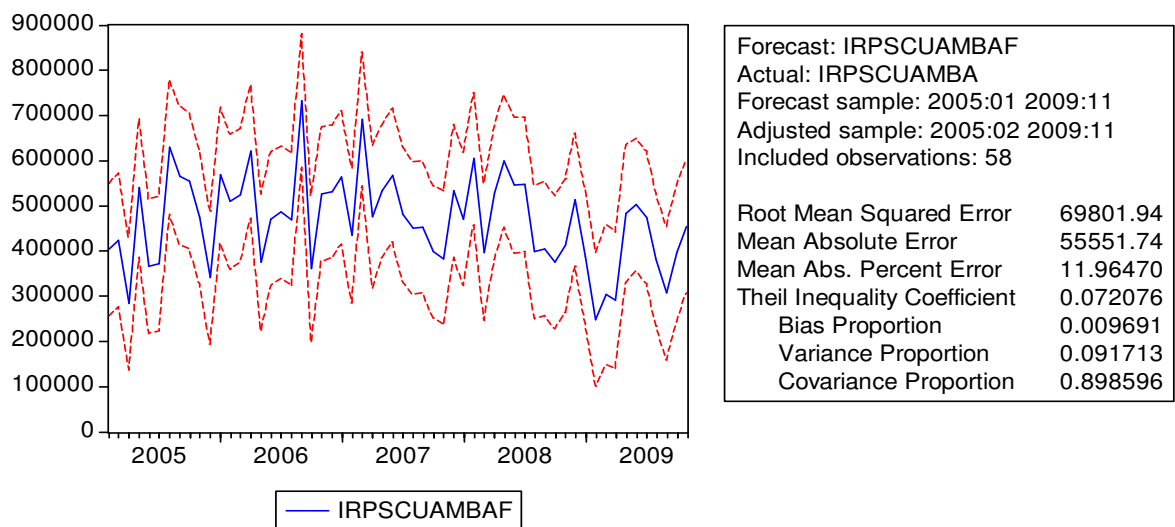


Figura 2A: Previsão dentro da amostra através do modelo $ARIMA(1,1,27)$ e estatísticas de desempenho na previsão

ANEXO B: Série de arrecadação do IRPS na Direcção de Área Fiscal de Lichinga

Tabela 1B: Cobranças do IRPS na Direcção de Área Fiscal de Lichinga

Mês	2005	2006	2007	2008	2009
Janeiro	2146505.7	3394070.8	2579876	4425453	6727983
Fevereiro	2550397.6	942156.70	2021029	2031766	5046982
Marco	879829.10	1734393.0	2095506	3415178	3587787
Abril	528743.60	1379066.9	2015255	2922636	3605366
Maio	797731.20	4956886.8	3328626	3854087	4145337
Junho	1295648.3	2590313.6	3546400	4851775	4392984
Julho	865645.30	2177227.5	3784574	4401676	4913294
Agosto	940067.40	3523505.6	3336981	5166250	4596926
Setembro	825887.10	2601785.5	3263471	4155587	4337806
Outubro	2383709.5	1891104.6	3859410	5578413	4547088
Novembro	3801317.1	2423924.3	3848973	4406285	3395958
Dezembro	4748065.4	2847768.2	4033063	6799587	-----

Fonte: Autoridade Tributária de Moçambique

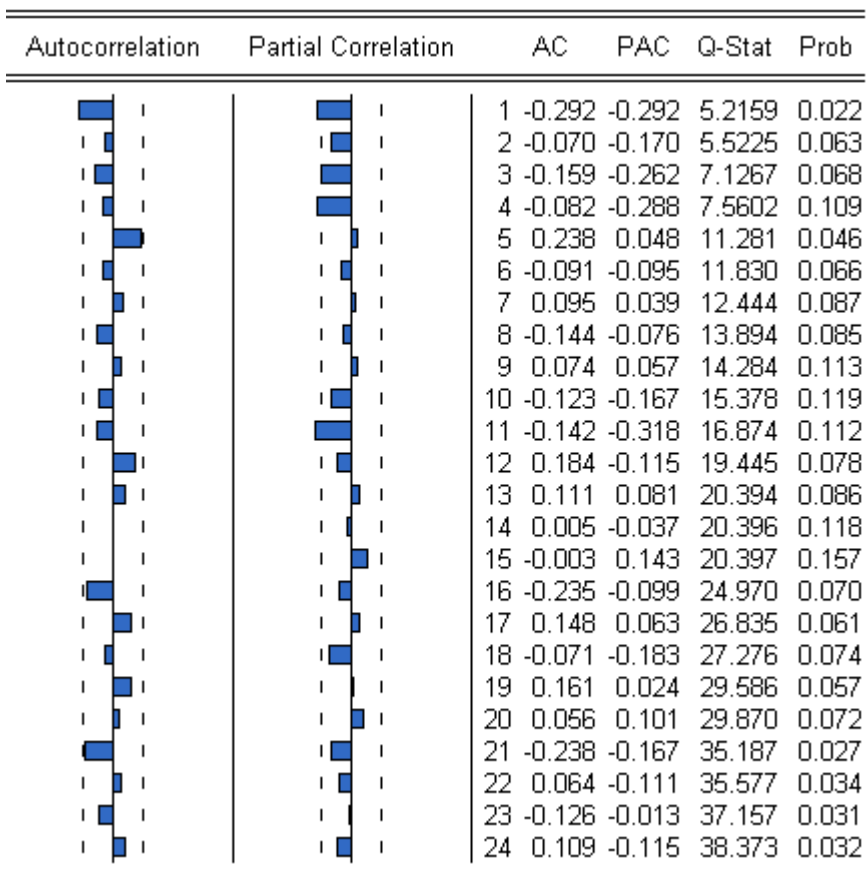


Figura 1B: Correlogramas da série de IRPS Lichinga em primeiras diferenças

Tabela 2B: Teste de igualdade das arrecadações médias mensais da série IRPSLichinga

IRPSLichinga					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	28420899729074.9	11	583718157188.621	1.160	0.339
Within Groups	104656741622658	47	226739183460.809		
Total	133077641351733	58			

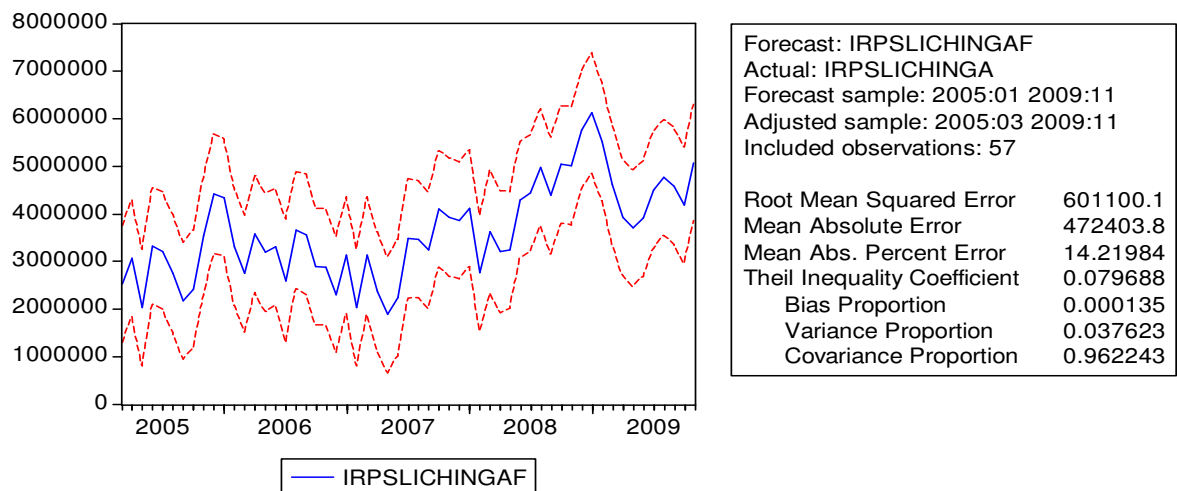


Figura 2B: Previsão de IRPSLichinga através do modelo $ARIMA(1,1,21)$ e estatísticas de desempenho na previsão

