

**UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE CIÊNCIAS NATURAIS E EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA E
MODELGAEM QUANTITATIVA**

**COMPARAÇÃO DE TÉCNICAS DE PREVISÃO EM ANÁLISE DE
SÉRIES TEMPORAIS: uma proposta via *wavelets* e modelos tradicionais
aplicadas à série de arrecadação do ICMS no RS**

MONOGRAFIA DE ESPECIALIZAÇÃO

Fábio Mariano Bayer

**Santa Maria, RS, Brasil
2009**

**COMPARAÇÃO DE TÉCNICAS DE PREVISÃO EM ANÁLISE
DE SÉRIES TEMPORAIS: uma proposta via *wavelets* e modelos
tradicionais aplicadas à série de arrecadação do ICMS no RS**

por

Fábio Mariano Bayer

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

Orientador: Prof. Dr. Adriano Mendonça Souza

**Santa Maria, RS, Brasil
2009**

**Universidade Federal de Santa Maria
Centro de Ciências Naturais e Exatas
Programa de Pós-Graduação em Estatística e Modelagem Quantitativa**

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

**COMPARAÇÃO DE TÉCNICAS DE PREVISÃO EM ANÁLISE DE
SÉRIES TEMPORAIS: uma proposta via *wavelets* e modelos tradicionais
aplicadas à série de arrecadação do ICMS no RS**

elaborada por
Fábio Mariano Bayer

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

COMISSÃO EXAMINADORA:

**Adriano Mendonça Souza, Dr. (UFSM)
(Presidente/Orientador)**

Roselaine Ruviaro Zanini, Dra. (UFSM)

Luis Felipe Dias Lopes, Dr. (UFSM)

Santa Maria, 9 de Janeiro de 2009.

AGRADECIMENTO

À Débora, pelo carinho, amor, dedicação e, sobretudo, paciência dedicados a mim durante esta etapa.

ao Adriano, mais do que um orientador, um grande amigo, que teve participação fundamental na conclusão desta pesquisa;

ao Programa de Pós-Graduação em Estatística e Modelagem Quantitativa e, principalmente, aos membros da banca examinadora por viabilizarem a defesa da presente monografia e por muito contribuírem para o resultado final da mesma;

à Analice, grande amiga que me ajudou nos momentos mais difíceis;

à Sandra, secretária do curso de Pós-Graduação em Estatística e Modelagem Quantitativa, sempre prestativa e disposta a nos ajudar;

à Universidade Federal de Santa Maria;

aos demais familiares, colegas e amigos, que deram força e entenderam meus momentos de ausência e ansiedade.

RESUMO

Monografia de Especialização
Programa de Pós-Graduação em Estatística e Modelagem Quantitativa
Universidade Federal de Santa Maria

COMPARAÇÃO DE TÉCNICAS DE PREVISÃO EM ANÁLISE DE SÉRIES TEMPORAIS: uma proposta via *wavelets* e modelos tradicionais aplicadas à série de arrecadação do ICMS no RS

AUTOR: Fábio Mariano Bayer

ORIENTADOR: Adriano Mendonça Souza

Data e Local da Defesa: Santa Maria, 9 de Janeiro de 2009.

Esta pesquisa tem o objetivo de explorar e comparar modelos tradicionais de previsão da classe ARIMA e de alisamento exponencial, assim como propor o uso conjunto destas técnicas tradicionais de previsão com a decomposição via *wavelets*. A comparação é feita por meio de um estudo empírico aplicado à série de arrecadação mensal do ICMS no estado do Rio Grande do Sul. A série em estudo compreende o período de janeiro de 1998 a outubro de 2008 deflacionados a preços constantes de outubro de 2008 utilizando o IGP-DI. Dentre os modelos de alisamento exponencial foi utilizado o algoritmo de Holt-Winters aditivo e, dentre os modelos da classe ARIMA, foram utilizados os modelos SARIMA, pois a série em estudo apresenta um comportamento sazonal. Estes mesmos modelos foram utilizados para a modelagem e previsão de cada uma das subséries da decomposição *wavelet*. As técnicas de previsão foram comparadas mediante diferentes horizontes de previsão e pôde-se verificar que a proposta do uso conjunto dos algoritmos de Holt-Winters com a decomposição *wavelet* originou uma melhora significativa às previsões, principalmente para os horizontes de previsão $h = 2$ e $h = 12$.

Palavras Chave: Séries Temporais, *Wavelets*, Comparação, Previsão, ICMS.

ABSTRACT

Monograph of Specialization
Post Graduation Program in Statistics and Quantitative Modeling
Federal University of Santa Maria

COMPARISON OF FORECASTING TECHNIQUES IN TIMES SERIES ANALYSIS: a proposal using wavelets and traditional models applied to the tax revenue of ICM in the RS

AUTHOR: Fábio Mariano Bayer
ADVISOR: Adriano Mendonça Souza
Date and Place of defence: Santa Maria, 9 de Janeiro de 2009.

The main purpose of this research is to compare and explore traditional ARIMA and exponential smoothing models, likewise to propose a combined use of these forecasting techniques with the wavelets decomposition. The comparison is made by mean a empirical study applied to the monthly tax revenue of ICM of the Rio Grande do Sul state. The series period is from January, 1998 till October, 2008 using the date deflated at constant price of October, 2008 using the IGP-DI. Among the exponential smoothing models the additive Holt-Winter model was used, the class of ARIMA and the SARIMA models, because the series shows a seasonal behavior. These same models were used to modeled and forecast each of sub-series of the decomposition in wavelets. The forecasting techniques were compared using different forecasting horizons and was possible to observe that the proposed use of combine algorithmic of Holt-Winters and wavelets decomposition gave a better results to the forecasted values, principally to the horizons with steps of forecasting equal $h = 2$ e $h = 12$.

Key words: Time series, *Wavelets*, Comparison, Forecasting, ICMS.

LISTA DE FIGURAS

FIGURA 1 - EXEMPLO DE ANÁLISE MULTIRESSOLUÇÃO <i>WAVELET</i> COM $J = 4$	33
FIGURA 2 – PREVISÃO COM DECOMPOSIÇÃO <i>WAVELETS</i>	34
FIGURA 3 - SÉRIE DA ARRECADAÇÃO MENSAL DO ICMS-RS NO PERÍODO ANALISADO.....	41
FIGURA 4 - DECOMPOSIÇÃO EM SAZONALIDADE, TENDÊNCIA E IRREGULAR DA SÉRIE DO ICMS-RS	42
FIGURA 5 - ANÁLISE GRÁFICA DA SAZONALIDADE DA SÉRIE ICMS-RS	43
FIGURA 6 - CORRELOGRAMA E CORRELOGRAMA PARCIAL DA SÉRIE ICMS	44
FIGURA 7 - CORRELOGRAMA E CORRELOGRAMA PARCIAL DA SÉRIE COM UMA DIFERENÇA	44
FIGURA 8 – ANÁLISE DE DIAGNÓSTICO DO MODELO SARIMA PARA $H = 2$	47
FIGURA 9 – ANÁLISE DE DIAGNÓSTICO DO MODELO SARIMA PARA $H = 6$	48
FIGURA 10 – ANÁLISE DE DIAGNÓSTICO DO MODELO SARIMA PARA $H = 12$	48
FIGURA 11 – VALORES REAIS E PREVISTOS UTILIZANDO O ALGORITMO DE HOLT-WINTERS PARA AS DUAS SUBSÉRIES DA DECOMPOSIÇÃO <i>WAVELET</i> ($H = 6$)	51
FIGURA 12 – VALOR REAL E PREVISTO DA SÉRIE ICMS-RS UTILIZANDO ALGORITMO DE HOLT-WINTERS COM DECOMPOSIÇÃO <i>WAVELET</i> ($H = 6$)	51
FIGURA 13 - VALORES REAIS E PREVISTOS UTILIZANDO MODELOS ARIMA PARA AS DUAS SUBSÉRIES DA DECOMPOSIÇÃO <i>WAVELET</i> ($H = 6$)	52
FIGURA 14 - VALOR REAL E PREVISTO DA SÉRIE ICMS-RS UTILIZANDO MODELOS SARIMA E DECOMPOSIÇÃO <i>WAVELET</i> ($H = 6$)	52
FIGURA 15 – VALORES REAIS E PREVISTOS PARA $H = 2$	55
FIGURA 16 - VALORES REAIS E PREVISTOS PARA $H = 6$	55
FIGURA 17 - VALORES REAIS E PREVISTOS PARA $H = 12$	56
FIGURA 18 – PREVISÃO PARA OS PRÓXIMOS 12 MESES DE ARRECADAÇÃO DO ICMS-RS	57

LISTA DE TABELAS

TABELA 1 - CRITÉRIOS DE AVALIAÇÃO SOBRE A QUALIDADE DAS PREVISÕES.....	35
TABELA 2 - MEDIDAS DESCRITIVAS DA SÉRIE DO ICMS-RS.....	40
TABELA 3 – TESTE ADF PARA A SERIE DA ARRECADAÇÃO DO ICMS-RS.....	45
TABELA 4 – MELHORES MODELOS SARIMA PARA A SÉRIE DO ICMS-RS	46
TABELA 5 – CONSTANTES DE ALISAMENTO PARA O ALGORITMO DE HOLT-WINTERS PARA A SÉRIE ICMS-RS	49
TABELA 6 – COMPARAÇÃO DAS DIFERENTES TÉCNICAS DE PREVISÃO	54
TABELA 7 - PREVISÃO PARA OS PRÓXIMOS 12 MESES DE ARRECADAÇÃO DO ICMS-RS	57

LISTA DE REDUÇÕES

- AIC – critério de informação de Akaike (*Akaike Information Criterion*)
- AICc – critério de informação de Akaike corrigido
- AR – modelo auto-regressivo
- ARIMA – modelo auto-regressivo integrado e de médias móveis
- ARMA - modelo auto-regressivo e de médias móveis
- BIC – Critério de informação Bayesiano (*Bayesian Information Criterion*)
- Cor - correlação
- Cov - covariância
- DB(P) - wavelets de Daubechies com P momentos nulos
- EMV - estimadores de máxima verossimilhança
- EQM – Erro quadrático médio
- FAC – função de autocorrelação
- FAV – função de autocovariância
- FGV – Fundação Getúlio Vargas
- GARCH – modelo auto-regressivo de heteroscedasticidade condicional generalizado
(*generalized autoregressive conditional heteroskedasticity*)
- H_0 – Hipótese nula
- H_1 – Hipótese alternativa
- i.i.d.* – independente e identicamente distribuída
- ICMS - Imposto sobre Circulação de Mercadorias e Prestação de Serviços
- IGP-DI - Índice Geral de Preços – Disponibilidade Interna
- MA – modelo de médias móveis (*moving average*)
- MAPE – erro absoluto percentual médio
- RB – ruído branco
- RS – Rio Grande do Sul
- SARIMA – modelo auto-regressivo integrado e de médias móveis sazonal
- SEBRAE – Serviço Brasileiro de Apoio às Micro e Pequenas Empresas
- teste ADF - teste de Dickey e Fuller aumentado
- teste DF – teste de Dickey e Fuller
- TW – transformada wavelet direta
- TWI – transformada wavelet inversa
- V.A. – variável aleatória
- Var - variância

SUMÁRIO

1	INTRODUÇÃO	11
1.1	TEMA DA PESQUISA	12
1.2	JUSTIFICATIVA E IMPORTÂNCIA DA PESQUISA	12
1.3	PROBLEMA DE PESQUISA	12
1.4	OBJETIVOS	13
1.4.1	<i>Objetivo Geral</i>	13
1.4.2	<i>Objetivos específicos</i>	13
1.5	DELIMITAÇÃO DA PESQUISA	14
1.6	ORGANIZAÇÃO DO TRABALHO	14
2	REFERENCIAL TEÓRICO-EMPÍRICO.....	15
2.1	ANÁLISE DE SÉRIES TEMPORAIS	15
2.1.1	<i>Conceitos Básicos</i>	15
2.1.2	<i>Modelos de Decomposição</i>	18
2.1.3	<i>Algoritmos de Alisamento Exponencial</i>	19
2.1.4	<i>Modelagem ARMA e extensões</i>	22
2.1.5	<i>Metodologia Box-Jenkins</i>	25
2.1.6	<i>Testes de Raiz Unitária</i>	28
2.2	<i>WAVELETS</i>	30
2.2.1	<i>Previsões utilizando decomposição wavelet</i>	33
2.3	CRITÉRIOS DE AVALIAÇÃO E COMPARAÇÃO DE MODELOS	35
3	METODOLOGIA	37
3.1	BANCO DE DADOS	37
3.2	ASPECTOS COMPUTACIONAIS	38
4	APRESENTAÇÃO E ANÁLISE DOS RESULTADOS.....	40
4.1	ANÁLISE EXPLORATÓRIA	40
4.2	MODELAGEM BOX-JENKINS.....	43
4.2.1	<i>Identificação da ordem de integração</i>	43
4.2.2	<i>Estimação dos parâmetros e escolha do melhor modelo</i>	45
4.2.3	<i>Análise de diagnóstico</i>	46
4.3	MODELOS DE ALISAMENTO EXPONENCIAL	49
4.4	MODELOS DE PREVISÃO COM DECOMPOSIÇÃO <i>WAVELET</i>	50
4.5	COMPARAÇÃO DAS DIFERENTES TÉCNICAS DE PREVISÃO	53
4.6	PREVISÃO PARA OS PRÓXIMOS MESES 12 MESES.....	56
5	CONSIDERAÇÕES FINAIS E RECOMENDAÇÕES	58
	APÊNDICE	64

1 INTRODUÇÃO

A área de análise e previsão de séries temporais tem recebido considerável atenção nas últimas décadas. Um número crescente de métodos e técnicas estatísticas tem sido desenvolvidas com o objetivo de obter previsões cada vez mais acuradas, sejam em séries temporais macroeconômicas, como séries de consumo ou produção (como em DIEBOLD, 1998), ou em outras áreas como recursos naturais, oceanografia, agropecuária (exemplos, ROCHE, 1995; LABYS, 1999; MORANA, 2001; DOOLEY e LENIHAN, 2005; FERNANDEZ, 2007).

Com o avanço da tecnologia e da capacidade de armazenagem e processamento dos sistemas computacionais, diversos modelos e técnicas quantitativas de previsão têm sido pesquisados, complementando e aprimorando as análises qualitativas por uma série de fatores, principalmente por uma maior precisão (WINKLHOFER et al., 1996).

Modelos quantitativos de previsão utilizam basicamente dados históricos para detectar padrões de comportamento ao decorrer do tempo e então estimá-los no futuro. Diversas técnicas estatísticas têm sido usadas para este fim, entre as mais populares estão os modelos da classe ARIMA, ou modelos Box-Jenkins, e os algoritmos de alisamento exponencial.

Nos últimos anos, a literatura de modelos de previsão vem dando destaque à técnicas como combinação de múltiplas previsões individuais de diferentes modelos estatísticos (exemplo, CHEN e YANG, 2007; CLEMEN, 1989; CORDEIRO e CORDEIRO, 2004), assim como a utilização da decomposição de séries via *wavelets*. Esta última vem ganhando popularidade principalmente nas áreas de economia e finanças, por sua capacidade de decompor uma série em subséries de baixa e alta frequência sem perder sua localização no tempo (LIN e STEVENSON, 2001; GENÇAY, WHITCHER e SELÇUK, 2005). No entanto, segundo Fernandez (2007), o uso de *wavelets* com o propósito de previsão não é muito explorado na literatura. Alguns poucos exemplos da utilização da decomposição *wavelet* para tal fim são Arinõ (1995), Wong et al. (2003), Lima (2004); Homsy, Portugal e Araújo (2000).

Com isso, esta pesquisa visa explorar os modelos tradicionais de previsão da classe ARIMA e de alisamento exponencial aplicados aos dados de arrecadação do Imposto sobre Circulação de Mercadorias e Prestação de Serviços (ICMS) no Rio Grande do Sul (RS), assim como apresentar uma proposta que visa melhorar a qualidade das previsões com o uso da

decomposição da série temporal via *wavelet* juntamente com os modelos tradicionais para prever cada uma das subséries da decomposição *wavelet*.

1.1 Tema da pesquisa

A presente pesquisa apresenta uma comparação do desempenho dos modelos da classe ARIMA, modelos de alisamento exponencial e uma proposta híbrida de previsão, que une decomposição *wavelet* e modelos tradicionais, quando aplicados a série da arrecadação do ICMS no estado do Rio Grande do Sul.

1.2 Justificativa e importância da pesquisa

No dia 4 de maio de 2000 o Congresso Nacional aprovou a Lei Complementar 101/00, que ficou conhecida como Lei de Responsabilidade Fiscal. A partir do momento que a Lei de Responsabilidade Fiscal entrou em vigor os governantes passaram a se preocupar mais fortemente com a arrecadação de suas unidades administrativas.

Para alcançar o almejado equilíbrio fiscal, torna-se fundamental um bom planejamento orçamentário, onde a base desse planejamento é a receita orçamentária. Com base no rigor imposto pela Lei de Responsabilidade Fiscal, os governantes passam a exigir previsões mais acuradas das receitas futuras de suas instituições fazendárias. A própria Lei determina a execução da previsão da receita e seu rigor técnico, em seus artigos 11 e 12.

Com isso, pode-se destacar a importância do emprego de técnicas de previsão nas decisões estratégicas governamentais. Sendo o ICMS o principal imposto estatal se faz de suma importância contribuir e desenvolver técnicas mais acuradas para a previsão do mesmo.

Tem-se como um importante diferencial da presente pesquisa a utilização de modelos de alisamento exponencial conjuntamente com a decomposição via *wavelet*, uma vez que não encontrou-se trabalhos na literatura que utilize conjuntamente estas duas técnicas.

1.3 Problema de Pesquisa

Com base nas conjecturas realizadas na seção 1.2, pode-se definir o problema de pesquisa da seguinte forma:

Qual é a melhor técnica a ser considerada para a previsão da arrecadação do ICMS no Rio Grande do Sul? E ainda, a utilização de *wavelets* juntamente com os modelos tradicionais melhora essas previsões?

1.4 Objetivos

Neste item são descritos os objetivos que nortearão a presente pesquisa.

1.4.1 Objetivo Geral

Explorar os modelos tradicionais de previsão da classe ARIMA e de alisamento exponencial quando aplicados aos dados de arrecadação do ICMS no Rio Grande do Sul, assim como propor o uso conjunto destas técnicas tradicionais com decomposição via *wavelets*.

1.4.2 Objetivos específicos

Na busca de alcançar o objetivo geral da pesquisa, traçam-se os seguintes objetivos específicos:

- utilizar modelos tradicionais da classe ARIMA e algoritmos de alisamento exponencial pra realizar previsões da série em estudo;
- introduzir a idéia da utilização da decomposição *wavelet* na análise de séries temporais;
- realizar previsões por meio do uso conjunto de *wavelets* e modelos tradicionais de previsão;
- comparar a qualidade das previsões das diferentes técnicas em diferentes horizontes de previsão;
- por fim, traçar previsões da arrecadação do ICMS do Rio Grande do Sul para os próximos 12 meses com a melhor técnica encontrada para este horizonte de previsão.

1.5 Delimitação da Pesquisa

A variável em estudo se restringe à arrecadação mensal do ICMS no estado do Rio Grande do Sul no período de janeiro de 1998 a outubro de 2008, sem contemplar qualquer possível interação com outras variáveis exógenas, em uma abordagem exclusivamente univariada.

As técnicas de previsão são os modelos da classe ARIMA e algoritmos de alisamento exponencial. São utilizadas *wavelets* de Daubechies, que são *wavelets* ortogonais e de suporte compacto, não fazendo qualquer referência a outros tipos de *wavelets* como de Haar, Morlet, Chapéu Mexicano, Symlets, Coiflets, Splines entre outras. A previsão nos diferentes níveis da decomposição *wavelet* também fica restrita ao uso dos modelos ARIMA e de alisamento exponencial.

1.6 Organização do Trabalho

O presente trabalho está organizado em 5 capítulos. No capítulo 1 aborda-se uma introdução ao trabalho de pesquisa, contendo a importância do trabalho, justificativa e delimitação do tema. O segundo capítulo apresenta a revisão de literatura, fornecendo o aporte teórico-estatístico para o desenvolvimento da pesquisa. No capítulo 3 é exposta a metodologia, bem como a apresentação dos dados em estudo e os aspectos computacionais da pesquisa. O capítulo 4 apresenta a realização do estudo empírico, com a aplicação das técnicas e os resultados obtidos com a mesma e, no quinto e último capítulo, são apresentadas as conclusões dos resultados obtidos.

No Apêndice são encontradas as implementações computacionais desenvolvidas e utilizadas na pesquisa.

2 REFERENCIAL TEÓRICO-EMPÍRICO

Os itens abordados neste capítulo servem de embasamento teórico para o desenvolvimento desta pesquisa. Inicialmente, são apresentados os modelos de séries temporais mais tradicionais, posteriormente uma breve introdução à *wavelets* e, por último, encontra-se a proposta do uso conjunto das técnicas.

2.1 Análise de Séries Temporais

Uma série temporal é um conjunto de observações de qualquer fenômeno aleatório ordenadas no tempo. A análise de séries temporais consiste em procurar alguma relação de dependência existente temporalmente nos dados, identificando o mecanismo gerador da série com o objetivo de extrair periodicidades relevantes nas observações, descrever o seu comportamento e fazer previsões.

Em diversas áreas do conhecimento, observações de interesse são obtidas em instantes sucessivos de tempo e isto pode ocorrer de forma discreta, por exemplo, a cada hora, ou então de forma contínua, como um traçado eletrocardiográfico. Com isso, tem-se séries temporais discretas e/ou séries temporais contínuas.

Nas seções imediatamente seguintes são introduzidos conceitos básicos¹ para a compreensão e estudo de séries temporais, assim como os modelos utilizados nesta pesquisa. As definições aqui apresentadas, assim como um estudo mais completo e detalhado pode ser encontrado em Morettin (2006), Morettin e Toloí (2006), Brockwell e Davis (2002), Hamilton (1994), Box et al. (1994).

2.1.1 Conceitos Básicos

Na análise de séries temporais é natural supor que cada observação y_t é um valor realizado de certa variável aleatória Y_t . Desta forma, a série temporal $\{y_t, t \in \mathbf{T}_0\}$ é uma realização da coleção de variáveis aleatórias $\{Y_t, t \in \mathbf{T}_0\}$. Portanto, em análise de séries

¹ Conceitos básicos não no sentido de serem triviais, mas sim no sentido de serem a base de conhecimentos para o estudo de séries temporais.

temporais, modelam-se os dados de uma realização (ou parte dela) de um processo estocástico $\{Y_t, t \in \mathbf{T}\}$, onde $\mathbf{T} \supseteq \mathbf{T}_0$.

Formalmente, define-se um processo estocástico como segue:

Definição 1. (Processo Estocástico). Seja \mathbf{T} um conjunto arbitrário. Um processo estocástico é uma coleção $\{Y_t, t \in \mathbf{T}\}$ tal que, para cada $t \in \mathbf{T}$, Y_t é uma variável aleatória.

O conjunto \mathbf{T} é normalmente tomado como o conjunto dos inteiros, podendo ser contínuo em alguns casos. Portanto, a uma determinada realização desse processo dá-se o nome de série temporal.

Um conceito bastante importante em séries temporais diz respeito à estacionariedade. De modo geral, um processo é dito ser estacionário se ele oscila ao redor de uma média constante e com uma variância também constante, ou seja, o primeiro e o segundo momentos mantêm-se constantes durante todo o período analisado. Os dois primeiros momentos são definidos tal como se encontra em (1) e (2).

$$E[Y_t] = \mu_t, \quad (1)$$

$$E[(Y_t - \mu_t)^2] = E[Y_t^2] - E^2[Y_t] = \sigma_t^2. \quad (2)$$

A covariância e a correlação entre Y_t e Y_{t+h} são dadas em (3) e (4).

$$\gamma(h) = Cov(Y_t, Y_{t+h}) = E[(Y_t - \mu_t)(Y_{t+h} - \mu_{t+h})], \text{ para } h \in Z, \quad (3)$$

$$\rho(h) = Cor(Y_t, Y_{t+h}) = \frac{Cov(Y_t, Y_{t+h})}{\sqrt{\sigma_t^2 \sigma_{t+h}^2}}, \text{ para } h \in Z. \quad (4)$$

Como se tratam de medidas entre instantes de tempo de uma mesma série temporal, tendo o interesse na dependência de y_t e seus valores defasados y_{t-h} , os conceitos de correlação e covariância são chamados de autocorrelação e autocovariância, ou ainda, função de autocorrelação (FAC) e função de autocovariância (FAV), respectivamente, as quais são muito utilizadas na identificação de modelos de séries temporais. Uma nomenclatura também utilizada diz respeito aos correlogramas, que são as FAC estimadas expostas de forma gráfica.

Definição 2. A função distribuição conjunta de um conjunto finito de variáveis aleatórias $\{Y_t, t \in \mathbf{T}_0\}$ da coleção $\{Y_t, t \in \mathbf{T}\}$ é

$$F_{Y_{t_1}, \dots, Y_{t_n}}(y_{t_1}, \dots, y_{t_n}) = P\{w : Y(t_1, w) \leq y_{t_1}, \dots, Y(t_n, w) \leq y_{t_n}\},$$

onde n é o número de elementos de \mathbf{T}_0 .

Definição 3. (Estacionaridade forte). Um processo estocástico $\{Y_t, t \in \mathbf{T}\}$ é dito fortemente ou estritamente estacionário se as funções de distribuição conjuntas de Y_{t_1}, \dots, Y_{t_m} e $\{Y_{t_1+h}, \dots, Y_{t_m+h}\}$ são idênticas para todo $n \geq 1$ tal que $t_1, \dots, t_n, t_{n+1}, \dots, t_{n+h} \in \mathbf{T}$.

Com isso, os momentos, se existirem, são constantes ao longo do tempo, sendo esse comportamento raramente observado em dados reais. Contudo, pode-se utilizar uma condição de estacionaridade mais fraca, que se encontra na Definição 4.

Definição 4. (Estacionaridade fraca). Um processo estocástico $\{Y_t, t \in \mathbf{T}\}$ é dito ser fracamente estacionário ou simplesmente estacionário se:

- (i) $E[Y_t]$ é constante para todo t ;
- (ii) $Var(Y_t)$ é constante para t ;
- (iii) $Cov(Y_t, Y_{t+h})$ depende apenas de h para todo t .

É fato que, desde que os primeiros momentos existam, estacionaridade fraca implica estacionaridade forte, mas estacionaridade forte não necessariamente implica estacionaridade fraca. De agora em diante no texto, quando for referido a estacionaridade, na verdade, estará sendo referido a estacionaridade fraca.

A introdução do conceito de estacionaridade é essencial ao estudo de séries temporais, pois esta é a suposição mais comumente feita em tais modelos. No entanto, muitas vezes, em estudos empíricos, as séries são não-estacionárias, não satisfazendo uma ou todas as condições de estacionaridade apresentadas na Definição 4. A violação desta suposição pode ser decorrente de causas como mudanças na estrutura da série, sazonalidade, períodos de volatilidade, entre outras. Em geral, tomam-se diferenças da série original buscando torná-la estacionária.

A d -ésima diferença é dada pelo operador diferença, encontrado em (5).

$$(1 - B)^d = 1 - dB + \frac{d(d-1)}{2!} B^2 - \frac{d(d-1)(d-2)}{3!} B^3 + \dots \quad (5)$$

onde B é o operador retroativo, sendo $B^h y_t = y_{t-h}$.

No caso particular de d inteiro positivo, então a d -ésima diferença de y_t será dada por (6):

$$\Delta^d y_t = (1 - B)^d y_t. \quad (6)$$

Considerando a primeira diferença, ou diferença de ordem 1 ($d=1$) de y_t , torna-se assim mais compreensível o processo de diferenciação inteira, descrito em (7):

$$\Delta y_t = (1 - B)y_t = y_t - y_{t-1}. \quad (7)$$

A segunda diferença de y_t é dada por (8):

$$\Delta_{y_t}^2 = (1 - B)^2 y_t = (1 - 2B + B^2)y_t = y_t - 2y_{t-1} + y_{t-2}. \quad (8)$$

Nos modelos mais tradicionais, como os abordados nesta pesquisa, geralmente utilizam-se $d=0$, $d=1$ ou $d=2$. No entanto, existem modelos que admitem que a ordem de integração d receba valores reais entre $(-1,1)$. Tais modelos são chamados de modelos de integração fracionária, ou modelos de memória longa, os quais fogem do escopo desta pesquisa.

Outro conceito bastante comum e importante, em séries temporais, é o conceito de ruído branco, dado a seguir:

Definição 5. Diz-se que $\{\varepsilon_t, t \in Z\}$ é um ruído branco discreto se as variáveis aleatórias ε_t são não-correlacionadas, isto é, $Cov\{\varepsilon_t, \varepsilon_s\} = 0$, $t \neq s$.

Admitindo que ε_t seja um ruído branco, com média μ e variância σ^2 , denota-se $\varepsilon_t \sim RB(\mu, \sigma^2)$.

Se, ainda, todas as variáveis aleatórias do processo estocástico forem independentes e tiverem a mesma distribuição de probabilidade, tem-se então uma seqüência de variáveis aleatórias independentes e identicamente distribuídas (*i.i.d.*). Obviamente, se as variáveis aleatórias ε_t são independentes elas também serão não-correlacionadas.

A seguir, nas próximas subseções são introduzidos alguns modelos e algoritmos de previsão. Uma seqüência de variáveis aleatórias *i.i.d.*, como definida acima, é chamada de um processo puramente aleatório.

2.1.2 Modelos de Decomposição

Considerando as observações $\{y_t, t = 1, \dots, T\}$ de uma série temporal, um modelo de decomposição consiste em escrever y_t como uma soma de três componentes não-observáveis,

$$y_t = T_t + S_t + a_t, \quad (9)$$

onde T_t e S_t representam a tendência e sazonalidade, respectivamente, enquanto que a_t é uma componente aleatória, cujo processo gerador é um processo estacionário.

O principal interesse em considerar um modelo desse tipo é o de se estimar S_t , assim

como T_t , e então construir uma série livre de sazonalidade e/ou tendência.

A decomposição de uma série temporal também se faz importante em uma primeira análise dos dados, buscando identificar se tal série possui tendência ou sazonalidade. Nesta pesquisa, é utilizada a função `decompose` do pacote `stats` do programa R para tal procedimento. Esta função decompõe a série em suas componentes T_t , S_t e a_t usando médias móveis.

2.1.3 Algoritmos de Alisamento Exponencial

Uma grande classe de métodos de previsão são as suavizações ou alisamentos. Estas técnicas assumem que os valores extremos da série representam a aleatoriedade e, assim, por meio da suavização desses extremos, pode-se identificar o padrão básico da série que é o objeto de interesse.

Esses métodos ou algoritmos de suavização possuem grande popularidade devido à simplicidade, à eficiência computacional e à razoável precisão.

Na Seção 2.1.3.1 será apresentado o alisamento exponencial simples, o qual é adequado para séries localmente constantes, na Seção 2.1.3.2 o algoritmo de Holt, que é adequado para séries que apresentam tendência, e na Seção 2.1.3.3 será visto o algoritmo de Holt-Winters, o qual é indicado para séries sazonais.

2.1.3.1 Alisamento Exponencial Simples

Considere o caso em que a série y_t é localmente composta de seu nível mais um ruído aleatório, isto é,

$$y_t = \mu_t + \varepsilon_t, \quad t=1, \dots, T, \quad (10)$$

onde $E(\varepsilon_t) = 0$, $Var(\varepsilon_t) = \sigma^2$ e μ_t é uma componente desconhecida que pode variar com o tempo.

No alisamento exponencial simples, o nível é estimado pela média ponderada dos dados, sendo que os pesos decaem exponencialmente quando regredidos no tempo. Desse modo, a estimativa do nível é:

$$N_t = \alpha y_t + \alpha(1-\alpha)y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} + \dots, \quad (11)$$

onde α é a constante de alisamento ($0 \leq \alpha \leq 1$) e N_t é o valor exponencial suavizado.

O alisamento exponencial simples pode ser escrito na forma de recorrência, como

$$N_t = \alpha y_t + (1 - \alpha)N_{t-1}. \quad (12)$$

A equação (12) é importante, pois torna o custo computacional muito mais baixo do que utilizando a equação (11).

A previsão de todos os valores futuros é dada pelo último valor exponencialmente alisado, isto é,

$$\hat{y}_t(h) = N_t, \quad h = 1, 2, 3, \dots \quad (13)$$

A escolha da constante de alisamento α pode ser feita de diversas formas, inclusive subjetivas, no entanto, o mais usual é escolher α de modo a minimizar a soma dos quadrados dos erros. O erro de previsão um passo à frente é dado por $e_t = y_t - \hat{y}_{t-1}(1)$; logo, escolhe-se α tal que minimize $S(\alpha) = \sum_{t=3}^T e_t^2$. Esta mesma forma também é usada para determinar as constantes de alisamento dos algoritmos de Holt e Holt-Winters apresentados a seguir.

2.1.3.2 Algoritmo de Holt

O algoritmo de Holt é utilizado em séries não-sazonais decompostas localmente como soma de nível, tendência e um resíduo aleatório com média zero e variância constante, isto é,

$$y_t = \mu_t + T_t + \varepsilon_t, \quad t = 1, \dots, T. \quad (14)$$

O método de alisamento exponencial simples, quando aplicado a uma série que apresenta tendência linear, fornece previsões que subestimam ou superestimam continuamente os valores reais. Para evitar esse erro sistemático utiliza-se o algoritmo de Holt. Esse método é bastante similar ao alisamento exponencial simples, entretanto, a diferença sendo que em vez de suavizar só o nível ele utiliza uma nova constante de alisamento para captar a tendência da série.

Os valores do nível e da tendência da série, no instante t , serão estimados por:

$$\begin{aligned} N_t &= \alpha y_t + (1 - \alpha)(N_{t-1} + \hat{T}_{t-1}), \quad 0 \leq \alpha \leq 1, \\ \hat{T}_t &= \beta(N_t - N_{t-1}) + (1 - \beta)\hat{T}_{t-1}, \quad 0 \leq \beta \leq 1. \end{aligned} \quad (15)$$

Aqui, α e β são constantes de alisamento, as quais são determinadas da mesma maneira que para o alisamento exponencial simples.

As previsões h passos a frente são dadas por,

$$\hat{y}_t(h) = N_t + h\hat{T}_t, \quad h = 1, 2, 3, \dots \quad (16)$$

2.1.3.3 Algoritmo de Holt-Winters

O algoritmo de Holt-Winters é o método de alisamento exponencial usado quando a série possui sazonalidade. A sazonalidade da série pode ser aditiva ou multiplicativa. Para a série aditiva, tem-se,

$$y_t = \mu_t + T_t + S_t + \varepsilon_t, \quad t=1, \dots, T. \quad (17)$$

Para a série com sazonalidade multiplicativa,

$$y_t = \mu_t S_t + T_t + \varepsilon_t, \quad t=1, \dots, T. \quad (18)$$

O algoritmo de Holt-Winters possui duas abordagens, uma para série sazonal aditiva e outra para série sazonal multiplicativa, sendo que o algoritmo de Holt-Winters aditivo é dado por:

$$\begin{aligned} N_t &= \alpha(y_t - \hat{S}_{t-s}) + (1-\alpha)(N_{t-1} + \hat{T}_{t-1}), \quad 0 \leq \alpha \leq 1, \\ \hat{T}_t &= \beta(N_t - N_{t-1}) + (1-\beta)\hat{T}_{t-1}, \quad 0 \leq \beta \leq 1, \\ \hat{S}_t &= \gamma(y_t - N_t) + (1-\gamma)\hat{S}_{t-s}, \quad 0 \leq \gamma \leq 1, \end{aligned} \quad (19)$$

onde o índice s é o número de vezes que a série é observada por ano (geralmente $s = 12$).

As previsões do algoritmo Holt-Winters aditivo são dadas por

$$\begin{aligned} \hat{y}_t(h) &= N_t + h\hat{T}_t + \hat{S}_{t+h-s}, \quad h=1, \dots, s \\ \hat{y}_t(h) &= N_t + h\hat{T}_t + \hat{S}_{t+h-2s}, \quad h=s+1, \dots, 2s. \end{aligned} \quad (20)$$

...

De maneira semelhante, o algoritmo de Holt-Winters multiplicativo é dado por

$$\begin{aligned} N_t &= \alpha \left(\frac{y_t}{\hat{S}_{t-s}} \right) + (1-\alpha)(N_{t-1} + \hat{T}_{t-1}), \quad 0 \leq \alpha \leq 1, \\ \hat{T}_t &= \beta(N_t - N_{t-1}) + (1-\beta)\hat{T}_{t-1}, \quad 0 \leq \beta \leq 1, \\ \hat{S}_t &= \gamma \left(\frac{y_t}{N_t} \right) + (1-\gamma)\hat{S}_{t-s}, \quad 0 \leq \gamma \leq 1. \end{aligned} \quad (21)$$

onde s é o número de vezes que a série é observada por ano. As previsões do algoritmo Holt-Winters multiplicativo são dadas por

$$\begin{aligned} \hat{y}_t(h) &= (N_t + h\hat{T}_t)\hat{S}_{t+h-s}, \quad h=1, \dots, s \\ \hat{y}_t(h) &= (N_t + h\hat{T}_t)\hat{S}_{t+h-2s}, \quad h=s+1, \dots, 2s. \end{aligned} \quad (22)$$

...

2.1.4 Modelagem ARMA e extensões

Nesta seção são apresentadas as principais extensões dos modelos auto-regressivos de médias móveis (ARMA). A classe de modelos mais comumente utilizada em análise de séries temporais é a dos modelos auto-regressivos integrados e de médias móveis, abreviadamente, ARIMA, ou ainda modelos Box-Jenkins, que são modelos ARMA para processos não-estacionários. Estes são modelos estatísticos lineares, propostos originalmente por Box e Jenkins (1970), sendo que a idéia básica é que a série temporal em estudo seja gerada por um processo estocástico, cuja natureza pode ser representada a partir de um modelo matemático.

O modelo ARIMA pode ser identificado em sua totalidade, ou parcialmente, tais como os modelos AR, onde o processo estocástico modelado apresenta apenas a parte auto-regressiva, ou modelos MA, onde apenas as médias móveis são de interesse. Os modelos ainda podem ser ampliados para os conhecidos modelos SARIMA, na presença de sazonalidade, ou para os modelos mais gerais SARIMAX que, além da sazonalidade, consideram também variáveis explicativas ao modelo. Este último não será abordado nesta pesquisa, uma vez que são considerados unicamente modelos univariados.

Com exceção dos modelos SARIMAX, todas as outras extensões dos modelos ARMA são apresentadas em maiores detalhes nas subseções seguintes.

2.1.4.1 Modelos AR

Diz-se que Y_t segue um processo auto-regressivo de ordem p , ou, abreviadamente, $AR(p)$, se:

$$Y_t = c + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t, \quad (23)$$

onde $\varepsilon_t \sim RB(0, \sigma^2)$. O processo $AR(p)$ pode ser escrito utilizando o operador retroativo B , da seguinte maneira:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) Y_t = c + \varepsilon_t. \quad (24)$$

Ainda, o processo $AR(p)$ pode ser escrito como

$$\phi(B) Y_t = c + \varepsilon_t, \quad (25)$$

Onde $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ é o polinômio auto-regressivo.

Como um resultado geral, tem-se que um processo $AR(p)$ será estacionário se todas as raízes de $\phi(B)$ estiverem fora de círculo unitário complexo.

2.1.4.2 Modelo MA

Um processo de médias móveis de ordem q , $MA(q)$, é definido por

$$Y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}, \quad (26)$$

onde $\varepsilon_t \sim RB(0, \sigma^2)$.

Em termos de operadores retroativos, o processo $MA(q)$ pode ser escrito como

$$Y_t = \mu + (1 + \theta_1 B + \dots + \theta_q B^q) \varepsilon_t, \quad (27)$$

ou ainda,

$$Y_t = \mu + \theta(B) \varepsilon_t. \quad (28)$$

Vale salientar que um processo $MA(q)$ é sempre estacionário, independente dos valores do parâmetro. No entanto, tem-se que Y_t é dito invertível se todas as raízes de $\theta(B)$ estiverem fora do círculo unitário.

2.1.4.3 Modelos ARMA

Um processo auto-regressivo e de médias móveis de ordem (p, q) , denotado por $ARMA(p, q)$, é definido por:

$$Y_t = c + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}, \quad (29)$$

onde $\varepsilon_t \sim RB(0, \sigma^2)$.

Uma notação mais interessante é dada na equação (30):

$$\phi(B)Y_t = c + \theta(B)\varepsilon_t, \quad (30)$$

onde $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ e $\theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q$.

As condições de estacionaridade e invertibilidade dos modelos ARMA são as mesmas dos modelos AR e MA.

2.1.4.4 Modelos ARIMA

Na análise de séries temporais, procura-se por um conjunto de observações que sejam estáveis ao longo do tempo, garantindo as propriedades de um processo estacionário. Contudo, geralmente as realizações encontradas na natureza não apresentam esta

característica de estacionaridade, sendo necessária a utilização de transformações para torná-la estacionária. O procedimento comumente utilizado é o processo de diferenciação da série. O ato de tomar diferenças na série é denominado de integração, representado pela letra I incluída ao modelo ARMA. Logo, o modelo ARMA integrado passa ser denominado de modelo ARIMA. Isto posto, um processo estocástico segue um modelo ARIMA(p, d, q) se $\Delta^d Y_t$ (como definido na equação (7)) seguir um modelo ARMA(p, q), ou seja, tem-se o modelo dado em (31):

$$\phi(B)\Delta^d Y_t = \theta(B)\varepsilon_t, \quad (31)$$

onde d é inteiro não-negativo e $\varepsilon_t \sim RB(0, \sigma^2)$.

A ordem de integração (d) é o menor número de diferenças necessárias para se alcançar a estacionaridade da série temporal. Com isso, um processo estacionário é dito integrado de ordem zero, ou seja, I(0).

2.1.4.5 Modelos SARIMA

Na Seção 2.1.3.3 foi introduzido o algoritmo de Holt-Winters que é uma generalização do algoritmo de Holt para o caso em que há sazonalidade na série. Nesta seção serão introduzidos os modelos SARIMA, que são uma generalização da classe de modelos ARIMA para processos estocásticos com sazonalidade.

Para introduzir os modelos SARIMA é preciso definir o operador de diferenciação sazonal. Seja s o período da sazonalidade², a diferenciação sazonal é dada por:

$$(1 - B^s)Y_t = Y_t - Y_{t-s}. \quad (32)$$

Com isso, o modelo SARIMA(p, d, q)(P, D, Q) é dado pela equação (33):

$$\begin{aligned} (1 - \phi_1 B - \dots - \phi_p B^p)(1 - \Phi_1 B^s - \dots - \Phi_P B^{sP})(1 - B)^d (1 - B^s)^D Y_t = \\ (1 + \theta_1 B + \dots + \theta_q B^q)(1 + \Theta_1 B^s + \dots + \Theta_Q B^{sQ})\varepsilon_t, \end{aligned} \quad (33)$$

onde $\varepsilon_t \sim RB(0, \sigma^2)$ e (p, d, q) são as ordens do modelo referentes à dinâmica ordinária, enquanto que (P, D, Q) são as ordens do modelo da parte sazonal. Em geral, p, q, P e Q assumem valores 0, 1 ou 2, ao passo que d e D são usualmente 0 ou 1.

Utilizando-se a notação de polinômios auto-regressivos, como já utilizado nos modelos anteriores, tem-se que o modelo SARIMA(p, d, q)(P, D, Q) pode ser escrito como em (34):

² Usualmente, s é o número de vezes por ano em que a série é observada

$$\phi(B)\Phi(B)(1-B)^d(1-B^s)^D Y_t = \theta(B)\Theta(B)\varepsilon_t. \quad (34)$$

2.1.5 Metodologia Box-Jenkins

Esta seção trata de uma metodologia para modelar as séries temporais da classe ARIMA, ou seja, modelos ARMA, ARIMA e extensões, onde a extensão mais geral, aqui proposta, é o modelo SARIMA.

No estabelecimento de um modelo da classe ARIMA para uma série temporal há basicamente três estágios a considerar:

- (i) identificação;
- (ii) estimação;
- (iii) diagnóstico.

Esta metodologia é conhecida como metodologia de Box-Jenkins. Nesta seção serão vistos com poucos detalhes cada um dos estágios da metodologia de Box-Jenkins. Informações mais detalhadas podem ser vistos em Moretin e Tolo (2006) e Box et al. (1994).

Serão apresentadas as etapas da metodologia considerando um particular modelo ARIMA(p, d, q), no entanto, entende-se que a metodologia é a mesma, guardadas as devidas particularidades, para os modelos mais gerais como SARIMA, SARIMAX e outros.

2.1.5.1 Identificação

A identificação do modelo ARIMA a ser ajustado aos dados é, talvez, a fase mais crítica do processo iterativo da metodologia de Box-Jenkins. Este estágio subdivide-se basicamente em outras duas etapas, que são a identificação da ordem de integração (d) e a identificação da ordem (p, q) do modelo.

Para a identificação da ordem de integração existem duas maneiras: a primeira seria analisando o correlograma, onde, caso ele decaia muito lentamente há indicação de que a série é não-estacionária, ou seja, $d > 0$; a outra maneira de decidir se a série é integrada, ou não é por meio dos testes de raiz unitária, os quais estão descritos na Subseção 2.1.6. Se houver evidências de que o processo gerador dos dados é integrado de ordem d , então toma-se d diferenças na série e passa-se para a etapa de identificação da ordem (p, q).

A partir da série estacionarizada, por meio de diferenciação, a identificação da ordem do modelo ARMA também pode ser feita de duas maneiras. A primeira seria analisando os

correlogramas, onde segundo Morettin (2006), os processos AR, MA e ARMA apresentam FAC com as seguintes características especiais:

- (i) um processo $AR(p)$ tem FAC que decai de acordo com exponenciais e/ou senóides amortecidas, infinitas em extensão;
- (ii) um processo $MA(q)$ tem FAC finita, no sentido que ela apresenta um corte após o lag q ;
- (iii) um processo $ARMA(p, q)$ tem FAC infinita em extensão, que decai de acordo com exponenciais e/ou senóides amortecidas após o lag $q-p$.

Estas observações podem ser úteis no procedimento de identificação do modelo a ser ajustado aos dados observados, calculando-se as estimativas das FAC (correlogramas), que se acredita reproduzir adequadamente as verdadeiras FAC desconhecidas. Também existe outra abordagem semelhante, utilizando-se as funções de autocorrelações parciais (FACP), para isso ver Box et al. (1994).

A segunda maneira de identificar a ordem (p, q) do processo é por meio dos critérios de informação ou critérios de seleção de modelos. A idéia básica dos critérios de seleção de modelos consiste em incluir um termo penalizador pela inclusão de novos parâmetros, pois cada parâmetro a mais a ser estimado representa um grau de liberdade a menos.

Um critério de seleção bastante utilizado é o AIC (*Akaike Information Criterion*), proposto em Akaike (1974), que foi desenvolvido a partir dos estimadores de máxima verossimilhança (EMV), para decidir qual o modelo mais adequado quando utilizados muitos modelos com quantidades diferentes de parâmetros, isto é, selecionar um modelo que seja parcimonioso, ou, em outras palavras, que esteja bem ajustado com um número reduzido de parâmetros. Este critério é assintoticamente eficiente, no entanto não é consistente e, em geral, é dado por:

$$AIC = -2l(\hat{\xi}) + 2(k), \quad (35)$$

onde k é o número de parâmetros do modelo e $l(\hat{\mu}, \hat{\phi})$ é o logaritmo da verossimilhança maximizada. No caso do modelo $ARMA(p, q)$, $k = p+q$.

O AIC mostra-se pouco adequado para pequenas amostras. Com o objetivo de melhorar o seu comportamento, Hurvich e Tsai (1989) derivaram o AICc e mostraram que este critério, de fato, tem melhor desempenho que o AIC em pequenas amostras. Ele é assintoticamente equivalente ao AIC e, portanto, é assintoticamente eficiente. Temos que

$$AICc = -2l(\hat{\xi}) + 2(k) \left(\frac{T}{T - k - 1} \right). \quad (36)$$

Baseados em uma perspectiva bayesiana, Schwarz (1978) e Akaike (1978) introduziram critérios equivalentes para seleção de modelos. Esse critério de seleção de modelos, conhecido como BIC (*Bayesian Information Criterion*), é consistente e dado por

$$BIC = -2l(\hat{\xi}) + (k)\log(T). \quad (37)$$

Com base em algum dos critérios de informação, estimam-se diversos modelos concorrentes e escolhe-se p e q de acordo com o modelo que fornece um menor valor para o critério de informação.

2.1.5.2 Estimação

Tendo-se identificado um modelo candidato para a série temporal, o próximo passo é estimar seus parâmetros. Uma vez que é utilizado o programa R para desenvolver a análise dos dados, será descrito brevemente o método de estimação utilizado pelas funções do R responsáveis pelo ajuste de modelos ARIMA.

Considere o modelo $ARIMA(p, d, q)$ e o vetor $\xi = (\phi, \theta, \sigma_a^2)$ de seus $p+q+1$ parâmetros, onde $\phi = (\phi_1, \dots, \phi_p)$, $\theta = (\theta_1, \dots, \theta_q)$ e no caso em que $\phi_0 = 0$. Caso isso não seja verdadeiro teremos $p+q+2$ parâmetros.

Para estimar ξ , um dos métodos empregados será o de máxima verossimilhança: dadas as T variáveis aleatórias Y_1, \dots, Y_T , considerando a função de verossimilhança $L(\xi | Y_1, \dots, Y_T)$ encarada como função de ξ . Os estimadores de máxima verossimilhança (EMV) de ξ serão os valores que maximizam L ou a função log-verossimilhança $l = \log L$. Para se determinar os EMV trabalha-se com a suposição que o processo ε_t (erro) é normalmente distribuído.

Como o estimador de máxima verossimilhança, neste caso, não possui forma fechada, precisa-se de métodos de otimização para maximizar a função L ou l . Os métodos de otimização necessitam de valores iniciais, sendo que no R esses valores iniciais são dados por meio de uma variação do método de mínimos quadrados.

2.1.5.3 Diagnóstico

Após identificar o modelo e estimar seus parâmetros, faz-se necessário verificar se o mesmo representa adequadamente os dados. Se a análise de diagnóstico revelar qualquer insuficiência é preciso considerar outro modelo alternativo, caso contrário, o modelo está apto para fazer previsões.

Nesta seção é apresentado o teste de Ljung-Box, que é um teste útil no diagnóstico de um modelo ajustado, uma vez que ele torna possível a identificação de existência de autocorrelação dos erros por meio das autocorrelações estimadas dos resíduos.

Box e Pierce (1970) sugeriram um procedimento para testar a hipótese de não-autocorrelação dos erros, que apesar de não detectar quebras específicas no comportamento de ruído branco, pode indicar se esses valores são muito grandes em valores absolutos. Esses autores formulam um teste global envolvendo as m primeiras autocorrelações dos resíduos (r_i), na qual deseja-se testar a seguinte hipótese nula

$$H_0 : \rho_1 = \rho_2 = \dots = \rho_m = 0 \quad (38)$$

Se o modelo for apropriado, ou seja, sob H_0 , a estatística de teste dada por

$$Q_{BP} = T \sum_{i=1}^m r_i^2 \quad (39)$$

possui distribuição assintótica χ^2 com $m-p-q$ graus de liberdade. Este fato nos permite concluir pela má qualidade do ajuste quando o valor da estatística de teste Q_{BP} exceder o valor crítico associado ao nível de significância adotado. A rejeição da hipótese nula sugere que o modelo estimado não é apropriado para descrever a série temporal em questão, indicando uma estrutura de correlação serial significativa nos resíduos e, portanto, nos erros, já que os resíduos devem ser boas estimativas para os erros.

Uma modificação deste teste foi proposta por Ljung e Box (1978), o qual possui a seguinte estatística de teste:

$$Q_{LB} = T(T+2) \sum_{i=1}^k \frac{r_i^2}{T-i}. \quad (40)$$

A estatística Q_{LB} também converge em distribuição para a distribuição χ^2 com $m-p-q$ graus de liberdade. Neste caso, também tem-se a hipótese nula de ruído branco rejeitada para valores suficientemente grandes de Q_{LB} .

2.1.6 Testes de Raiz Unitária

No processo de modelagem Box-Jenkins existem pressupostos, já mencionados nas seções anteriores, que devem ser satisfeitos e, portanto, testados. Dentre as suposições mais importantes está a de estacionaridade.

Nesta seção é apresentado o teste ADF, o qual é utilizado para inferir sobre a estacionaridade da série, admitindo como hipótese nula a não-estacionaridade da mesma. A idéia intuitiva do teste vem do fato de que uma raiz unitária, encontrada no polinômio auto-regressivo, sugere que a série segue um processo não-estacionário, ou seja, integrado de ordem um.

Considere o processo gerador dos dados da seguinte forma:

$$Y_t = \alpha Y_{t-1} + u_t. \quad (41)$$

Dickey e Fuller em 1979 e 1981 propuseram um teste de raiz unitária, o qual possui a forte hipótese de que $u_t \sim RB(0, \sigma^2)$, que se tornou conhecido como teste DF, ou teste de Dickey e Fuller. Relaxando a hipótese de ruído branco no termo u_t da equação (41), e assumindo que $u_t \sim ARMA(p, q)$, Said e Dickey (1984) propuseram um teste mais amplo conhecido como teste ADF, ou teste de Dickey e Fuller aumentado.

Para o teste ADF, considere os seguintes possíveis modelos estimados:

$$\Delta Y_t = \rho Y_{t-1} + \sum_{i=1}^k \gamma_i \Delta Y_{t-i} - i + \varepsilon_t, \quad (42)$$

$$\Delta Y_t = \mu + \rho Y_{t-1} + \sum_{i=1}^k \gamma_i \Delta Y_{t-i} - i + \varepsilon_t, \quad (43)$$

$$\Delta Y_t = \mu + \beta T_t + \rho Y_{t-1} + \sum_{i=1}^k \gamma_i \Delta Y_{t-i} - i + \varepsilon_t, \quad (44)$$

onde $\varepsilon_t \sim RB(0, \sigma^2)$, μ é uma constante e T_t é a componente de tendência. De acordo com o comportamento dos dados escolhe-se um dos três modelos de regressão anterior e estimam-se os parâmetros pelo método de mínimos quadrados. Com isso, as hipóteses a serem testadas são

$$\begin{cases} H_0 : \rho = 0 \\ H_1 : \rho = 1 \end{cases}$$

A estatística do teste é dada por

$$\tau = \frac{\hat{\rho}}{e.p.(\hat{\rho})} \quad (45)$$

em que $e.p.(\hat{\rho})$ é o erro padrão de $\hat{\rho}$.

A distribuição da estatística τ sob H_0 não segue distribuição *t-Student*, sendo que os valores críticos para o teste ADF foram obtidos via simulação. Com isso, rejeitando-se H_0 se τ for menor que o valor crítico ao nível de significância adotado. Ou seja, se τ é menor que o valor crítico, conclui-se que há evidências, ao nível α de significância, contra a hipótese nula de não estacionaridade.

Uma importante decisão a ser tomada no teste ADF é a escolha do número de truncagens auto-regressivas k , pois, em geral, o teste ADF torna-se sensível a diferentes valores de k . Em suma, a bibliografia da área sugere os seguintes procedimentos para a seleção ótima de k :

- (i) fixar um valor máximo de k , seguindo a regra $k_{\max} = \text{Int}\left\{12\left(T/100\right)^{1/4}\right\}$;
- (ii) fixado o k_{\max} seleciona-se o k que apresentar um menor critério de informação no intervalo entre zero e k_{\max} .

Para o passo (ii) acima, alguns autores sugerem a utilização do critério AIC, sendo que o critério BIC tende a escolher um número muito pequeno de truncagens auto-regressivas.

2.2 Wavelets

Segundo Lima (2004), a análise de séries temporais possui diversas ferramentas, onde talvez, a mais conhecida e utilizada nas diferentes áreas do conhecimento seja a análise de Fourier. A análise de Fourier quebra a série em seus componentes senoidais, cada um com dada frequência, e então descreve uma função em um intervalo de tempo (janela) como a combinação linear dessas funções senoidais.

Alguns autores, como Chiann (1997, *apud* Lima 2004), afirmam que a análise de Fourier é fundamental em áreas onde o interesse básico é a procura de periodicidades nos dados. Em muitos estudos a análise de Fourier é extremamente usada porque a frequência contém grande importância no estudo e é apropriada para processos estacionários.

No entanto, a análise de Fourier possui uma séria desvantagem, pois quando transforma a série para o domínio da frequência, a informação do tempo é perdida. Com isso, ao observar a transformada de Fourier de uma série é impossível saber quando determinado evento ocorreu (GENÇAY et al., 2002).

Dada esta característica da análise de Fourier, as *wavelets* tornam-se competidoras ou mesmo alternativas para a análise de Fourier, tornando-se uma forte ferramenta principalmente para dados não-estacionários.

As *wavelets* são funções matemáticas que ampliam intervalos de dados, separando-os em diferentes componentes de frequência e permitindo a análise de cada componente em sua escala correspondente. Para Misiti et al. (1997, apud Lima, 2004), este processo de decomposição envolve a passagem de uma série temporal por um filtro de uma *wavelet*, dando origem a duas novas séries: (i) uma série é chamada de aproximada, associada a baixas frequências e (ii) a outra série é denominada de detalhada, envolvendo altas frequências.

Quando as *wavelets* consideradas são ortogonais, pode-se tomar a transformada *wavelet* inversa das duas subséries (aproximação e detalhe) e então reconstruir o sinal original.

Em termos matemáticos, seja $L^2(\mathfrak{R})$ o espaço das funções de quadrado integrável e seja a função escala $\phi(x)$ que possui energia finita, é oscilatória, possui média zero e um decaimento rápido em ambos os lados. Além disso, a família $\{\phi(x-k), k \in \mathbb{Z}\}$ forma uma base ortonormal para um subespaço de referência V_0 . Assim, $\{0\} \subset \dots \subset V_{-1} \subset V_0 \subset V_1 \subset \dots \subset L^2(\mathfrak{R})$.

Ainda, $f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1}$, onde $j \in \mathbb{Z}$. Ou seja, os espaços são relacionados pelo fato de que uma função f para ir de V_j para V_{j+1} é escalonada por 2.

Considere a projeção de uma função f (sinal de interesse) no espaço V_j , no qual é gerado por todas as translações diádicas de uma função escala $\phi_{j,l}$ para a escala j , dada pela seguinte expressão:

$$(Pv_j f)(x) = \sum_{l=-\infty}^{+\infty} c_{j,l} \phi_{j,l}(x), \quad (46)$$

onde os $c_{j,l}$ seriam os pesos da combinação linear.

De Nielsen (1998), temos que, se Pv_j e Pw_j são os operadores de projeção ortogonal de f em V_j e W_j , respectivamente, onde W_j é complemento ortogonal de V_j , ou seja, $V_j = V_{j-1} \oplus W_{j-1}$, então $Pv_j f = Pv_{j-1} f + Pw_{j-1} f$. Desta maneira, esta função f pode ser decomposta em vários níveis de detalhes de acordo com as condições anteriores. Portanto, a projeção (46) também possui sua formulação em termos de funções escalas $\phi_{j,l}$ e funções *wavelets* ψ_{j+1} , dada em (47).

$$(Pv_j f)(x) = \sum_{l=-\infty}^{+\infty} c_{j-1,l} \phi_{j-1,l}(x) + \sum d_{j-1,l} \psi_{j-1,l}(x) \quad (47)$$

A ortogonalidade de funções escalas (ϕ) e funções *wavelets* (ψ) junto com a expansão da análise multiresolução conduzem à relação entre os coeficientes da função escala ($c_{j,l}$) e os coeficientes *wavelets* ($d_{j,l}$) nos diferentes níveis de resolução. Isto produz um algoritmo rápido e acurado, denotado algoritmo piramidal ou transformada *wavelet* rápida (MALLAT, 1989).

Há um mapeamento entre a seqüência de coeficientes $\{c_{j,l}, l \in Z\}$ do nível mais fino j com a seqüência de coeficientes $\{c_{j+1,l}, l \in Z\}$ e $\{d_{j+1,l}, l \in Z\}$ do nível $j+1$. As principais derivações desse mapeamento são a equação de dilatação dada por

$$\phi(x) = \sqrt{2} \sum_{k=0}^{2P-1} a_k \phi(2x - k) \quad (48)$$

e a equação *wavelet*

$$\psi(x) = \sqrt{2} \sum_{k=0}^{2P-1} b_k \phi(2x - k), \quad (49)$$

onde P é o número de momentos nulos e a_k e b_k são os filtros.

Considerando um sinal $c_{j,l}$ com 2^j valores discretos da função f na escala j , obtém-se a relação:

$$c_{j-1,l} = \sum_{k=0}^{2P-1} a_k c_{j,2l+k}, \quad d_{j-1,l} = \sum_{k=0}^{2P-1} b_k c_{j,2l+k}. \quad (50)$$

Aplicando (50) recursivamente para $j = J, J-1, \dots, J_0 + 1$, onde $J - J_0$ é a profundidade da expansão, tem-se a transformada *wavelet* rápida da função f , denotada por $TW(f) = \bar{f}$. Pela ortogonalidade das funções *wavelet* e escala, pode-se tomar a transformada inversa de \bar{f} , denotada por $TWI(\bar{f})$ e recompor o sinal original f , ou seja $TWI(\bar{f}) = f$. A transformada *wavelet* inversa parcial é dada por:

$$c_{j,l} = \sum_{n=n_1(l)}^{n_2(l)} c_{j-1,n} a_{l-2n} + d_{j-1,n} b_{l-2n}. \quad (51)$$

Consecutivamente, a transformada *wavelet* inversa $TWI(\bar{f})$, é obtida aplicando-se recursivamente a equação (51) para $j = J_0 + 1, \dots, J - 1, \dots, J$.

Um exemplo da análise multiresolução pode ser visto na Figura 1.

Fonte: Nielsen (1998)

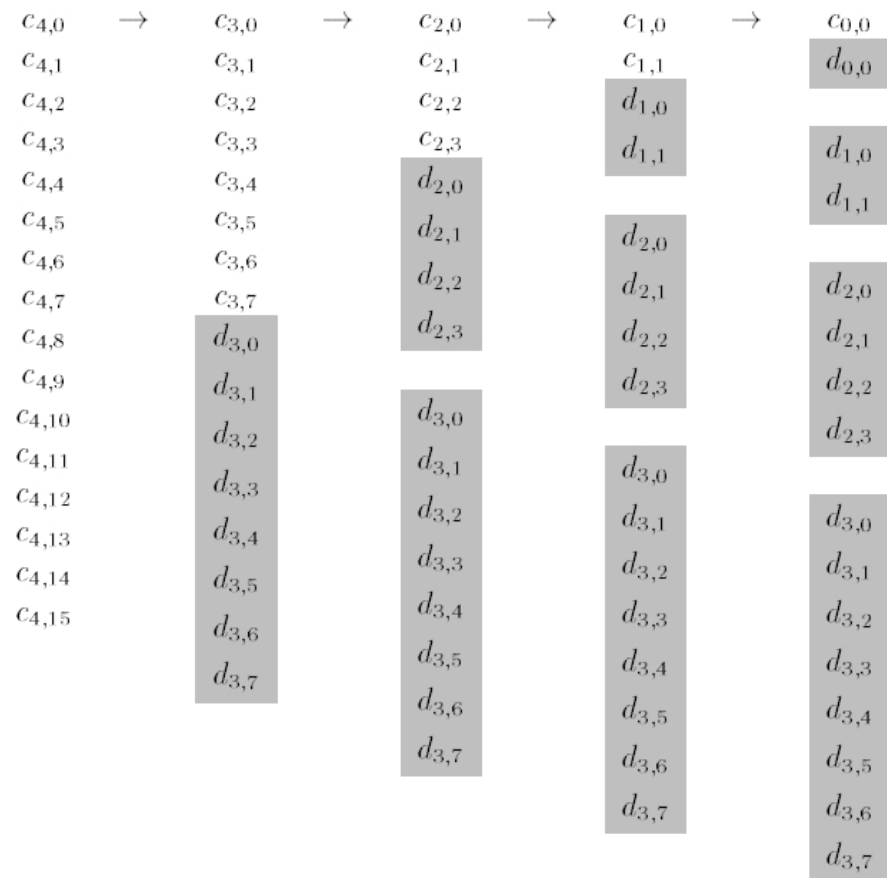


Figura 1 - Exemplo de análise multiresolução *wavelet* com $J = 4$

Na Figura 1, o primeiro vetor $c_{4,l}$, para $l = 0, 1, \dots, 15$ seria o vetor de dados originais de entrada. Com a decomposição em um nível, tem-se duas série $c_{3,l}$ e $d_{3,l}$, com $l = 0, 1, \dots, 7$; com a decomposição *wavelet* em dois níveis, ter-se-ia a terceira coluna, e assim sucessivamente.

2.2.1 Previsões utilizando decomposição *wavelet*

Seja uma função f de interesse, na qual na prática utiliza-se uma amostra discreta dessa função f representada por um vetor S com T observações. Como visto anteriormente, dado o sinal S de entrada, a transformada *wavelet* em um nível dá origem a duas novas subséries: (i) uma série é chamada de aproximada (C), associada a baixas frequências de tamanho $T/2$ e (ii) a outra série é denominada de detalhada (D), envolvendo altas frequências e também com $T/2$ observações.

Portanto, em análise *wavelet*, um sinal é dividido em aproximação e detalhe. Esta aproximação pode ser novamente dividida em aproximação e detalhe de segundo nível, e o processo se repete. Este processo é chamado de análise multiresolução.

Percebe-se com isso que, para a análise multiresolução são necessárias $T = 2^J$ observações no vetor S inicial. No entanto, nesta pesquisa o interesse é somente na TW com um nível, ou seja, dado o vetor de entrada S de tamanho T , tem-se a saída \bar{S} , que na verdade são dois vetores de tamanho $T/2$ cada um, onde um será o vetor de aproximação (C) e outro de detalhes (D) e, com isso, os dados necessitam apenas de um número par de observações.

Segundo a seção anterior, foi visto que, aplicando a transformada *wavelet* inversa $TWI(\bar{S})$, tem-se a série original S novamente, ou seja, $TWI(\bar{S}) = S$. Contudo, a idéia é utilizar métodos de previsão, como modelos ARIMA e de alisamento exponencial, para prever cada uma dessas subséries e então fazer a transformada inversa com esses valores previstos, tendo então, o sinal original previsto. Ou seja, dado que \hat{S} é o vetor de valores previstos de \bar{S} , obtido utilizando algum método tradicional de previsão, o interesse é obter a previsão \hat{S} de S , sendo que isso é dado por $TWI(\hat{S}) = \hat{S}$. Este processo pode ser visto na Figura 2.

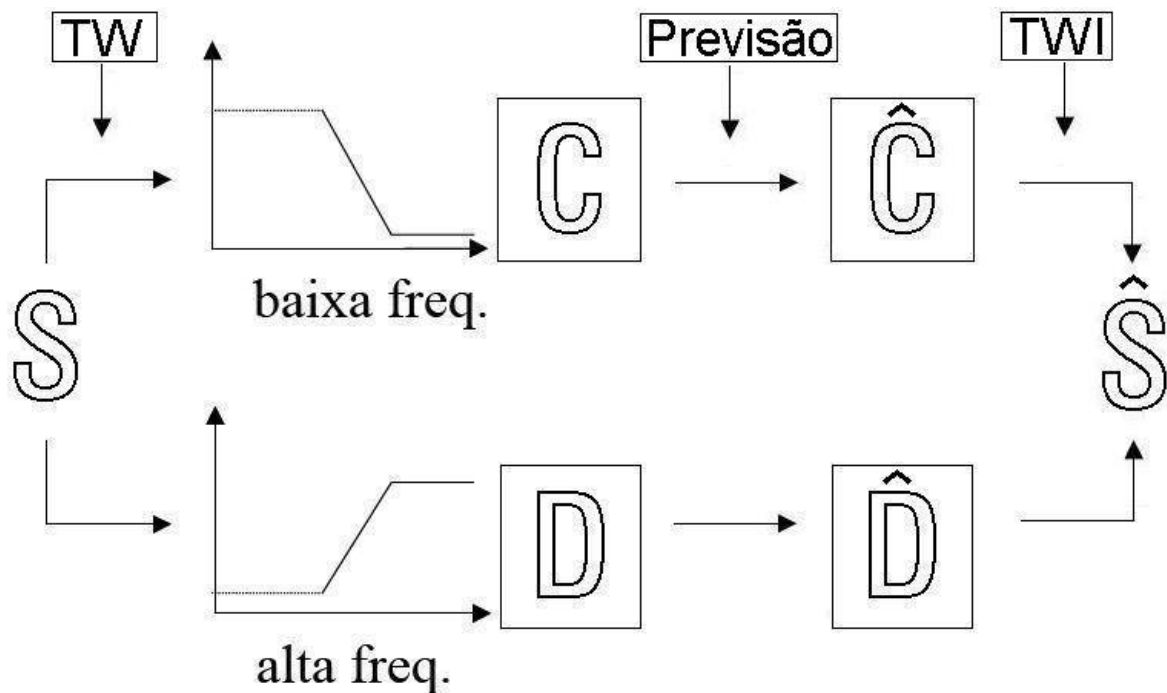


Figura 2 – Previsão com decomposição *wavelets*

Com isso, o objetivo é verificar se a modelagem dos dados separadamente em alta frequência e baixa frequência traz um melhor desempenho para as previsões do que quando feitos diretamente na série original.

Fernandez (2007) e Lima (2004) afirmam que existem poucos estudos neste sentido na literatura internacional. Os autores citam alguns trabalhos como Tak (1995), Arino (1995), Homsy et al. (2000) e Wong et al. (2003). Em seu trabalho, Lima (2004), utiliza a decomposição da série temporal via *wavelets* conjuntamente com modelos econométricos da classe ARIMA e GARCH e também redes neurais. Os resultados deste trabalho, assim como os demais trabalhos citados anteriormente, apontam para bons resultados quando utilizada a decomposição *wavelet* em métodos de previsão.

Por estes relatos, percebe-se que o uso da decomposição via *wavelet* modelando separadamente as subséries de baixas e altas frequências contribui positivamente para a qualidade das previsões. Isto ressalta o objetivo da presente pesquisa, que é utilizar a decomposição via *wavelet* conjuntamente com modelos ARIMA e modelos de alisamento exponencial.

2.3 Critérios de avaliação e comparação de modelos

O principal objetivo do trabalho é avaliar e comparar diferentes modelos de previsão em séries temporais. Para isso, faz-se necessário definir alguns critérios de avaliação de previsões.

Os critérios de avaliação de previsões utilizados neste trabalho encontram-se evidenciados na Tabela 1. Estes critérios serão calculados para cada um dos modelos em estudo, considerando também diferentes horizontes de previsão (h). São considerados $h = 2, 6, 12$, onde y_i é o valor real observado no instante i e \hat{y}_i é o valor previsto no mesmo instante i .

Tabela 1 - Critérios de avaliação sobre a qualidade das previsões

Critério	Notação	Expressão
Erro quadrático médio	EQM	$\frac{1}{h} \sum_{i=1}^h (y_{n+1} - \hat{y}_{n+1})^2$
Erro absoluto percentual médio	MAPE	$\frac{1}{h} \sum_{i=1}^h \frac{ y_{n+1} - \hat{y}_{n+1} }{ y_{n+1} }$

Para a avaliação das diferentes técnicas de previsão inicialmente são retiradas as h últimas observações amostrais. Posteriormente, a partir da série com $T-h$ observações aplicam-se as técnicas de previsão, obtendo-se os valores previstos para estas h últimas observações. Com isso, utilizando os critérios de avaliação para estes valores reais e previstos, determina-se qual é a melhor ou pior técnica de previsão.

3 METODOLOGIA

Neste capítulo é apresentada a abordagem metodológica utilizada para o desenvolvimento da pesquisa.

Este trabalho, de acordo com Gil (2006), caracteriza-se como uma pesquisa aplicada, uma vez que integra pesquisa bibliográfica e estudo de caso, desenvolvida de forma a quantificar as informações obtidas, caracterizando-se também como uma abordagem quantitativa. Quanto aos objetivos, classifica-se como descritiva, pois se conhece a direção de causalidade, utilizando-se testes de hipóteses. Entende-se o processo como um método científico indutivo, pois trata-se de um estudo teórico aplicado a uma determinada situação, em que generalizações poderão ser frutos de constatações particulares da realidade, gerando possíveis hipóteses à estudos futuros.

3.1 Banco de dados

Os dados em análise referem-se à arrecadação mensal do ICMS no Rio Grande do Sul (RS) no período de janeiro de 1998 a outubro de 2008, disponíveis on-line no site da Fundação de Economia e Estatística no endereço <http://www.fee.tche.br>.

O ICMS é o principal imposto de competência estadual e, segundo o SEBRAE, ele incide sobre a circulação de mercadorias, prestações de serviços de transporte interestadual, ou intermunicipal, de comunicações, de energia elétrica. Também incide sobre a entrada de mercadorias importadas e serviços prestados no exterior. Cada Estado da Federação tem liberdade para adotar regras próprias relativas à cobrança desse imposto, respeitados os requisitos mínimos fixados na Constituição Federal e pelo Código Tributário Nacional.

Contudo, a série encontra-se sobre o efeito da inflação acumulada no período em análise, fazendo-se necessário o deflacionamento desta, ou seja, a conversão de valores correntes (ou nominais) em moeda de poder aquisitivo constante (valor real).

A fim de expressar os dados a preços constantes de outubro de 2008, utiliza-se o Índice Geral de Preços – Disponibilidade Interna (IGP-DI). O IGP-DI é calculado mensalmente pela Fundação Getúlio Vargas (FGV) e foi instituído em 1944 com a finalidade de medir o comportamento de preços em geral da economia brasileira.

3.2 Aspectos computacionais

Para o cumprimento dos objetivos propostos na presente pesquisa utiliza-se o programa computacional R. O R é um *software* livre que pode ser obtido gratuitamente em <http://www.r-project.org> e utiliza a linguagem S, que é uma linguagem interpretada de programação de alto nível, desenvolvida para análise de dados e gráficos.

Para a modelagem e análise dos dados foram utilizados, dentre outros, os pacotes `forecast` e `urca`. O primeiro, utilizado para o ajuste de modelos ARIMA, Holt-Winters e também para a geração de previsões dos mesmos; o segundo, para a realização dos testes de raiz unitária.

Também foram criadas algumas funções úteis para o desenvolvimento do trabalho. Dentre elas estão as funções definidas como: `auto.sarima`, `bestk.adf`, `wave.hw` e `wave.arima`.

A função `auto.sarima` recebe de entrada um vetor com a série de interesse e retorna o modelo SARIMA que minimiza o critério AIC. Esta função também apresenta um resultado gráfico do teste de diagnóstico do melhor modelo encontrado por meio da função `tsgdiag(stats)`.

A função `bestk.adf` foi desenvolvida a partir de uma modificação da função `ur.df` do pacote `urca`. Assim como a função `ur.df`, a função `bestk.adf` também realiza o teste ADF de raiz unitária para uma dada série de entrada, no entanto, ela não necessita da informação do valor k (número de truncagens auto-regressivas), pelo contrário, ela encontra o valor de k ótimo segundo o critério de AIC, ou seja, utiliza o k que minimiza o AIC do modelo de regressão estimado no teste ADF.

As funções `wave.hw` e `wave.arima` são as funções mais importantes propostas nesta pesquisa. A primeira recebe uma série temporal como dado de entrada, assim como o número desejado de previsões passos a frente, e retorna estas h previsões utilizando a decomposição *wavelet* em um nível na qual cada subsérie, de alta e baixa frequência, são previstas utilizando o algoritmo de Holt-Winters e, então, toma a transformada inversa desses valores previstos. Já a função `wave.arima`, possui um procedimento semelhante, no entanto utiliza o modelo SARIMA que minimiza o critério AIC em cada uma das subséries para prever as mesmas.

Também foram criadas outras funções menos importantes, assim como utilizadas outras funções já prontas, principalmente do pacote `stats`, que é um pacote padrão do R que

já encontra-se instalado e disponível. Todas as funções criadas e modificadas apresentam-se no Apêndice.

4 APRESENTAÇÃO E ANÁLISE DOS RESULTADOS

Neste capítulo são apresentados e discutidos os resultados da análise de séries temporais da série do ICMS, assim como a comparação do desempenho dos diferentes modelos de previsão considerados e propostos.

4.1 Análise exploratória

Para um melhor entendimento da variável em estudo no período analisado, a Tabela 2 apresenta algumas medidas descritivas da série da arrecadação mensal do ICMS no RS, assim como evidencia-se na Figura 3 o gráfico temporal da mesma. Vale salientar que a série encontra-se em milhões de reais e deflacionadas a preços constantes de outubro de 2008.

Tabela 2 - Medidas descritivas da série do ICMS-RS

Estatística	Série do ICMS (milhões de R\$)
Média	1063,094
Mediana	1043,299
Desvio padrão	130,826
Coefficiente de Variação (%)	12,306
Mínimo	749,258
Máximo	1428,656

Nota-se que a arrecadação média do ICMS no período de janeiro de 1998 a outubro de 2008 foi de R\$ 1063,094 milhões, com desvio padrão de R\$ 130,826 milhões e arrecadação mediana de R\$ 1043,299. O coeficiente de variação abaixo de 30% indica que a média é bastante representativa, não havendo muita variabilidade em torno da mesma.

Uma importante ferramenta na análise de séries temporais é a análise gráfica da série ao longo do tempo. Trata-se de uma análise simples, mas bastante informativa, que se encontra na Figura 3.

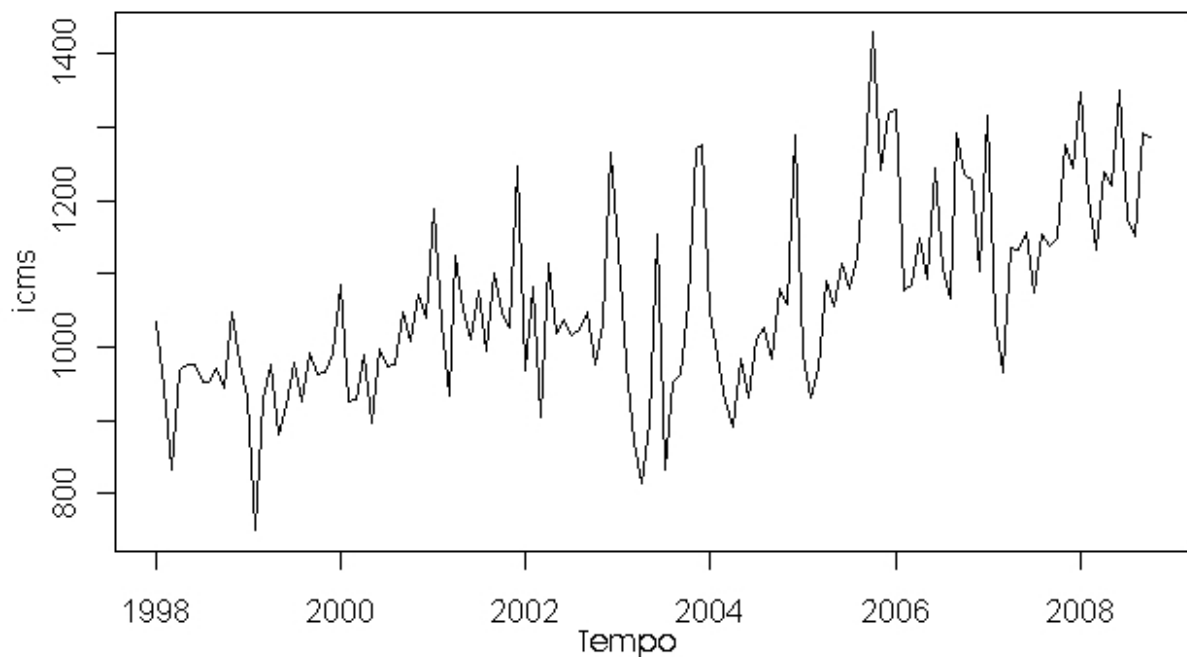


Figura 3 - Série da arrecadação mensal do ICMS-RS no período analisado

Pela simples inspeção visual da Figura 3, percebe-se uma tendência crescente ao longo do tempo, assim como um fator sazonal. Para melhor identificar tais componentes da série temporal do ICMS, são apresentadas as Figura 4 e Figura 5.

A Figura 4 apresenta a decomposição da série nas componentes de sazonalidade, tendência e irregular. Percebe-se claramente a sazonalidade da série, assim como a tendência crescente.

Outro gráfico bastante informativo a respeito da sazonalidade da série encontra-se na Figura 5. Nesta figura é fácil identificar que os meses de janeiro e dezembro detêm a maior arrecadação e o mês de março, em geral, possui a menor arrecadação anual. A tendência crescente na série também fica bastante evidenciada na Figura 5 (b), uma vez que observa-se um crescimento ao longo dos anos na maioria dos meses.

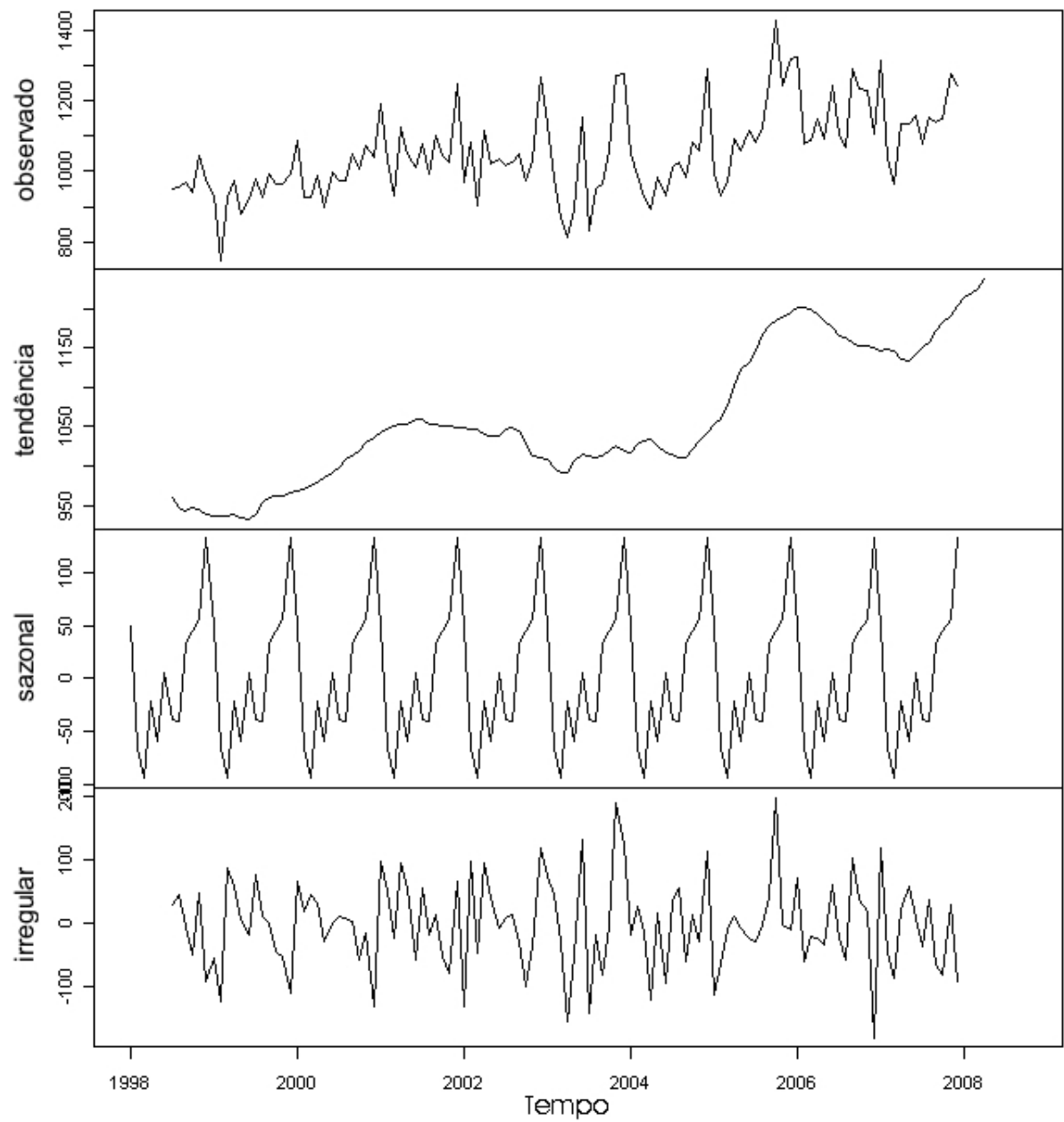


Figura 4 - Decomposição em sazonalidade, tendência e irregular da série do ICMS-RS

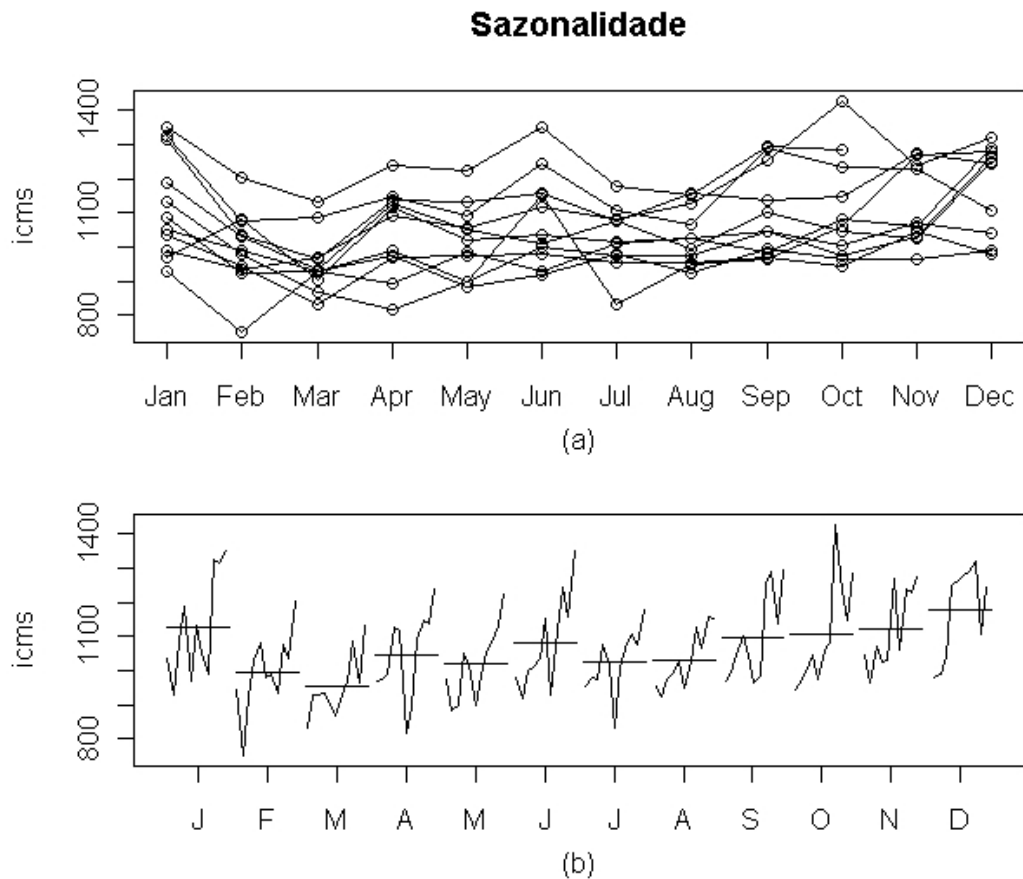


Figura 5 - Análise gráfica da sazonalidade da série ICMS-RS

4.2 Modelagem Box-Jenkins

Nesta seção serão explorados os modelos da classe ARIMA mediante a metodologia de Box-Jenkins.

4.2.1 Identificação da ordem de integração

As funções de autocorrelação amostral e autocorrelação parcial amostral da série do ICMS em nível e diferenciada de ordem 1 encontram-se nas Figura 6 e Figura 7.

Pela análise desses correlogramas percebe-se que a série parece ser integrada de ordem um, pelo seu decaimento lento no correlograma da série em nível e seu comportamento constante no correlograma da série diferenciada, sendo, assim, não-estacionária.

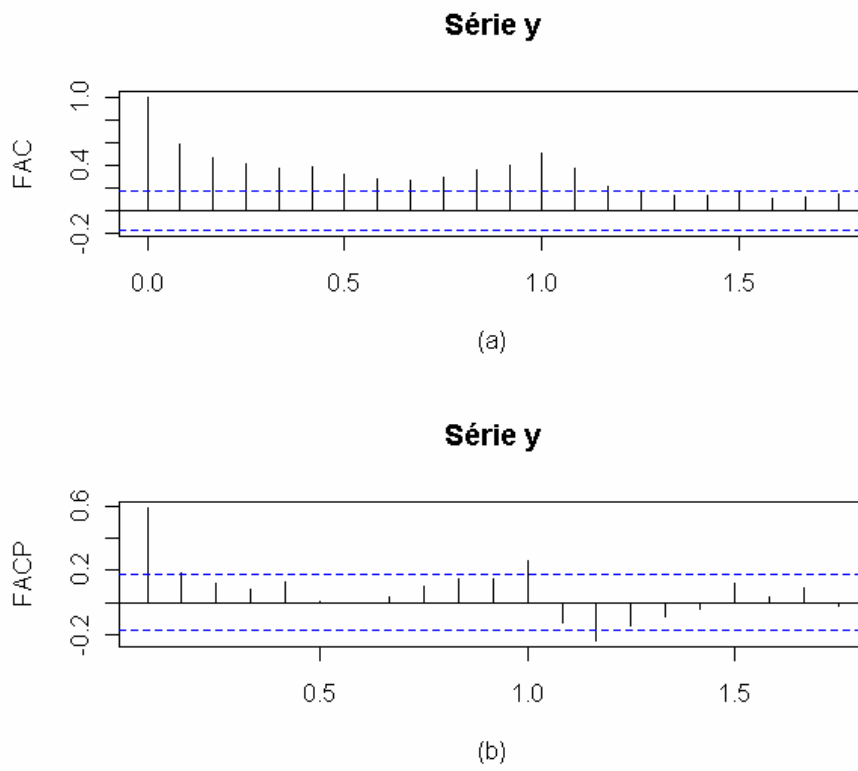


Figura 6 - Correlograma e correlograma parcial da série ICMS

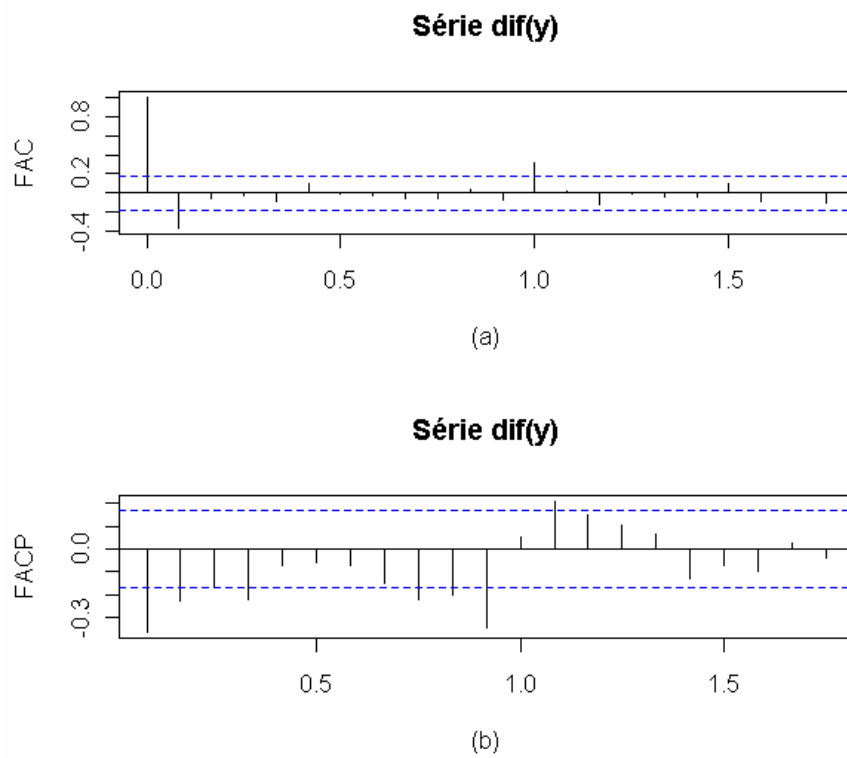


Figura 7 - Correlograma e correlograma parcial da série com uma diferença

Pelos correlogramas também se pode analisar a presença de sazonalidade, revelada pelo pico de correlação significativa na defasagem sazonal (1.0). Isso indica que um modelo ARIMA com sazonalidade seja o mais indicado, ou seja, um modelo SARIMA.

No entanto, para ter uma decisão mais robusta a respeito da não-estacionariedade da série, são apresentados na Tabela 3 os resultados do teste ADF para as séries em nível e diferenciada.

Tabela 3 – Teste ADF para a serie da arrecadação do ICMS-RS

Série	Estatística de Teste	Valor crítico (5%)	k_{max}	$k_{ótimo}$	Decisão
ICMS	-1,923	-3,430	13	13	Não rejeita-se H_0
dif(ICMS)	-3,043	-2,580	13	13	Rejeita-se H_0

Uma vez que a hipótese nula do teste ADF é de presença de raiz unitária, ou seja, de não-estacionariedade, pode-se afirmar pelos resultados da Tabela 3 que, ao nível de 5% de significância, a série é integrada de ordem 1, como já haviam sugeridos os correlogramas das Figuras 6 e 7.

Vale destacar, na Tabela 3, que o número de truncagens auto-regressivas ótimo escolhido pelo critério de minimização do AIC para o teste ADF foi o mesmo valor máximo calculado de acordo com a regra $k_{max} = \text{Int}\{12(T/100)^{1/4}\}$. Também foram feitos testes utilizando-se o critério BIC, sendo que os valores $k_{ótimo}$ em alguns casos foram menores que os encontrados utilizando-se o AIC, confirmando o conhecimento a priori de que o BIC penaliza mais fortemente a inclusão de novos parâmetros no modelo. No entanto, os resultados dos testes ADF com esse número de truncagens menor não foi alterado.

4.2.2 Estimação dos parâmetros e escolha do melhor modelo

Nesta seção apresentam-se os resultados da modelagem dos dados utilizando a função `auto.sarima`, que trata-se de uma generalização da função `auto.arima` do pacote `forecast`. Esta função estima e calcula o AIC de todos os modelos SARIMA até o modelo de maior ordem dado por SARIMA(6,2,6)(2,2,2). No entanto, como já foi verificada na seção anterior que a série é integrada de ordem 1, esta busca restringe-se pelo modelo que minimiza o AIC apenas para o caso em que $d=1$.

Foram considerados os modelos para os dados sem as h últimas observações, onde $h=2, 6, 12$. Esses modelos encontram-se na Tabela 4 com seus respectivos valores calculados de AIC, AICc e BIC.

Tabela 4 – Melhores modelos SARIMA para a série do ICMS-RS

Modelos	Parâmetros estimados	AIC	AICc	BIC
$h = 2$				
SARIMA(0,1,2)(1,2,1) ₁₂	$\theta_1 = -0,7733$	1026,64	1027,25	1039,81
	$\theta_2 = 0,0387$			
	$\Phi_1 = -0,3322$			
	$\Theta_1 = -0,9846$			
$h = 6$				
SARIMA(0,1,1)(0,1,2) ₁₂	$\theta_1 = -0,6736$	1201,34	1201,71	1212,18
	$\Theta_1 = -0,6047$			
	$\Theta_1 = -0,2761$			
$h = 12$				
SARIMA(0,1,1)(0,1,2) ₁₂	$\theta_1 = -0,6959$	1131,37	1131,77	1141,99
	$\Theta_1 = -0,6021$			
	$\Theta_1 = -0,2941$			

Os resultados da Tabela 4 evidenciam que os modelos para $h = 6$ e $h = 12$ são os mesmos e com pequenas diferenças nos parâmetros estimados. No entanto, o modelo encontrado para $h = 2$ é bastante diferente dos demais, considerando um parâmetro a mais e também duas diferenças sazonais.

Os erros-padrão, assim como os p -valores, para os estimadores dos parâmetros da Tabela 4 não são apresentados, pois, segundo Ansley e Newbold (1980), os estimadores pontuais de máxima verossimilhança para modelos ARIMA possuem bom desempenho, no entanto, a estimação intervalar não funciona bem em amostras finitas, uma vez que a distribuição dos estimadores não está próxima, em geral, da distribuição limite (Normal).

4.2.3 Análise de diagnóstico

Os diagnósticos dos modelos foram feitos utilizando-se a função `tsdiag(stats)` do R. Esta função traça os resíduos padronizados, o correlograma dos resíduos do modelo, assim

como uma interpretação gráfica do teste de Ljung-Box. O resultado da análise de diagnóstico do modelo SARIMA para $h = 2$ apresenta-se na Figura 8, para $h = 6$ na Figura 9 e, finalmente, para $h = 12$ a análise de diagnóstico evidencia-se na Figura 10. A análise destas figuras valida os modelos para traçar previsões, uma vez que não apresentam nenhuma autocorrelação significativamente diferente de zero no correlograma residual, os p -valores dos testes de Ljung-Box apresentam-se todos acima do limite tracejado indicando que não rejeita-se a hipótese nula de que as m primeiras autocorrelações dos erros são iguais a zero e os resíduos padronizados também encontram-se razoavelmente dentro dos limites de -2 a 2.

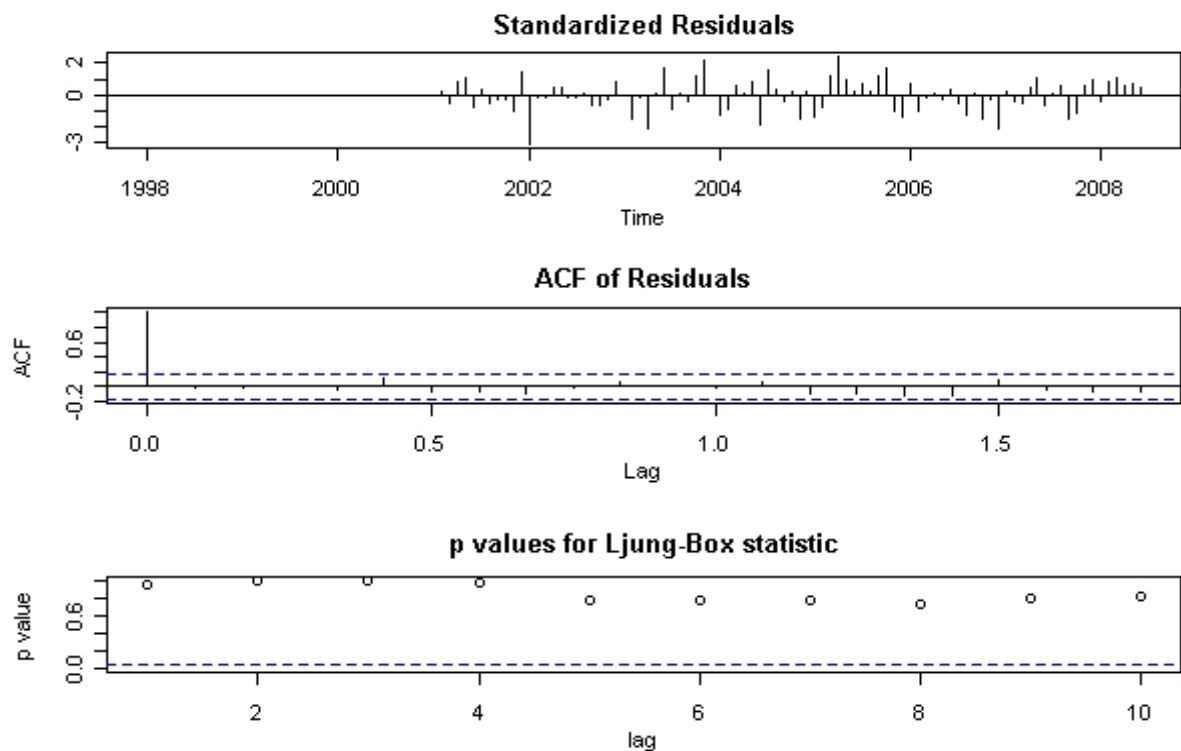


Figura 8 – Análise de diagnóstico do modelo SARIMA para $h = 2$

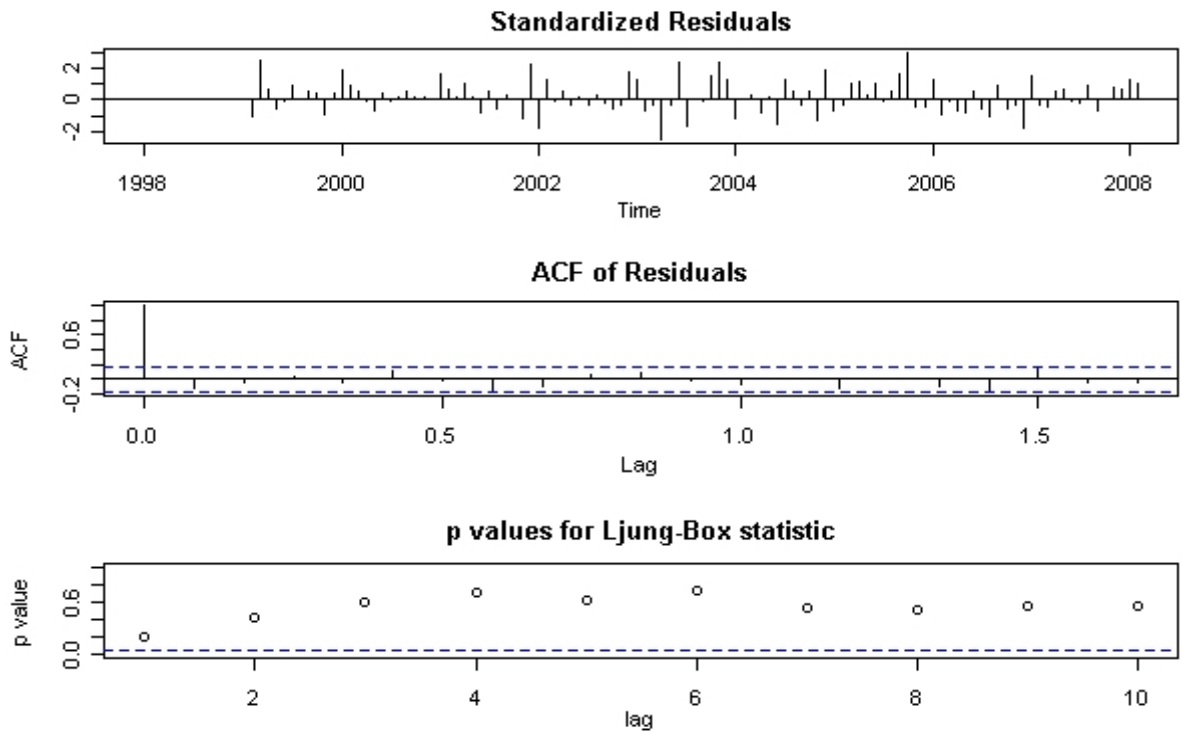


Figura 9 – Análise de diagnóstico do modelo SARIMA para $h = 6$

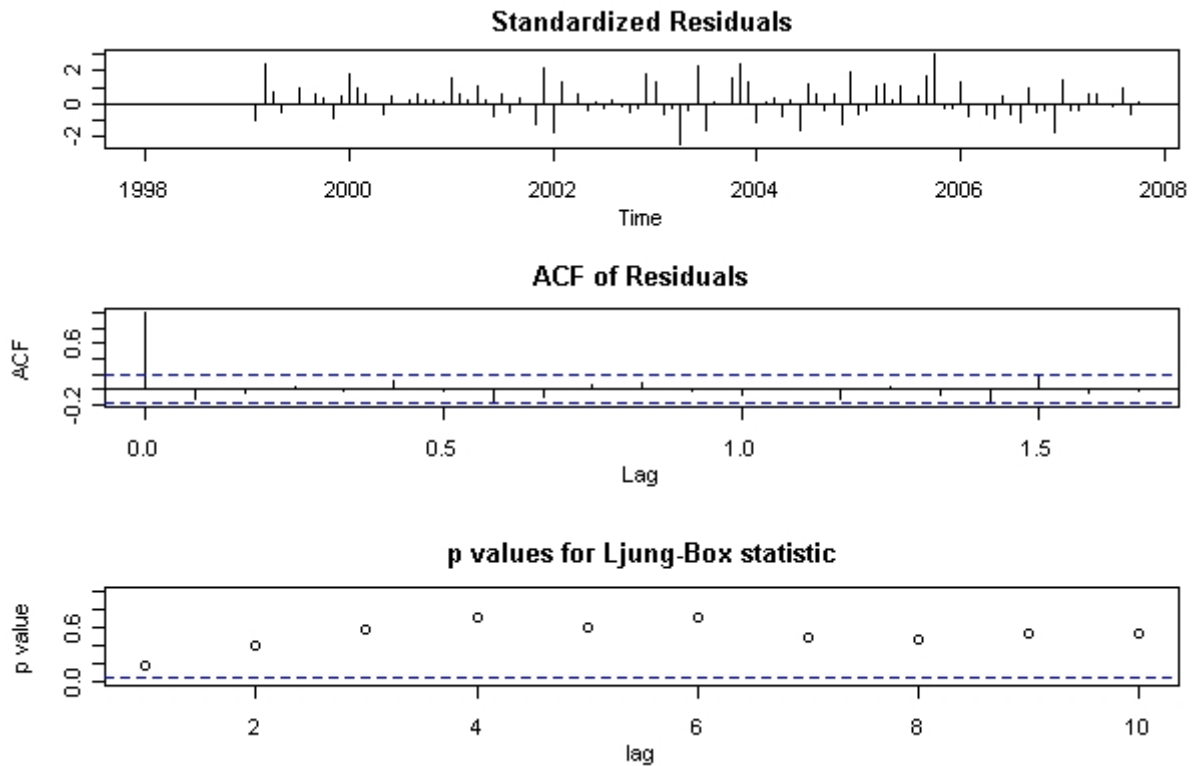


Figura 10 – Análise de diagnóstico do modelo SARIMA para $h = 12$

Como os modelos SARIMA foram validados pela análise de diagnóstico, considera-se que estes sejam os melhores modelos da classe ARIMA que representam os dados e aptos a realizar previsões. Com isso, os mesmos serão utilizados para a comparação com as previsões dos demais modelos considerados na pesquisa. Estas comparações encontram-se na seção 4.5.

4.3 Modelos de Alisamento Exponencial

Por meio da análise exploratória dos dados na seção 4.1, conclui-se que a série possui sazonalidade e tendência. Uma série com tais características pode ser modelada factivelmente por meio de um alisamento exponencial sazonal de Holt-Winters, ou algoritmo de Holt-Winters.

O método de Holt-Winters oferece duas modelagens distintas que dependem do comportamento da série original. Quando a amplitude sazonal da série aumenta juntamente com o tempo utiliza-se um modelo multiplicativo e, quando a amplitude sazonal é constante, ou seja, independente da variação temporal, utiliza-se o modelo aditivo (WINTERS, 1960). Nesta pesquisa percebe-se, pelas análises até aqui apresentadas, que a componente de sazonalidade é constante, ou seja, sugere a utilização do algoritmo de Holt-Winters aditivo.

Na Tabela 5 encontram-se as constantes de alisamento encontradas para o algoritmo de Holt-Winters considerando os diferentes valores de h .

Tabela 5 – Constantes de alisamento para o algoritmo de Holt-Winters para a série ICMS-RS

Modelo	Constantes de alisamento		
	α	β	γ
Holt-Winters para $h = 2$	0,2487	0,0048	0,0001
Holt-Winters para $h = 6$	0,2372	0,0034	0,0038
Holt-Winters para $h = 12$	0,2105	0,0001	0,0006

Apesar de os dados terem visivelmente forte sazonalidade, as constantes de alisamento γ , referentes ao fator de sazonalidade do algoritmo de Holt-Winters, são bem próximas de zero.

Estes modelos de alisamento exponencial serão utilizados na seção 4.5 para a comparação com os demais métodos.

4.4 Modelos de previsão com decomposição *wavelet*

Nesta seção são apresentados os resultados da modelagem da série utilizando a decomposição *wavelet* juntamente com os modelos SARIMA e modelos de alisamento exponencial de Holt-Winters. Vale salientar que existem diversas escolhas para as funções *wavelets*, como já citado, porém este trabalho se deterá às *wavelets* de Daubechies com $P = 8$, denominadas DB(8), onde os filtros a_k e b_k são tabelados.

A apresentação dos resultados se dará somente para o caso $h = 6$, uma vez que os resultados para $h = 2$ e $h = 12$ são muito semelhantes a estes.

O primeiro passo desta técnica é decompor a série temporal em duas outras subséries, de baixa e de alta frequência, por meio da transformada *wavelet*. Essa decomposição para a série com $h = 6$, encontra-se nas Figura 11 e Figura 13, com seus respectivos valores previstos utilizando-se o algoritmo de Holt-Winters e modelos ARIMA, respectivamente. Como já mencionado, a série de baixa frequência possui um comportamento similar ao da série original, isto pode ser visto comparando-se a aproximação da Figura 11 com a série original na Figura 3.

Como visto nas seções anteriores, as *wavelets* de Daubechies são ortogonais, podendo tomar a transformada *wavelet* inversa destas subséries e então recompor o sinal original. No entanto, o objetivo do trabalho é utilizar os valores previstos destas subséries para fazer a transformada inversa e com isso obter os valores previstos da série original.

Fazendo a transformada inversa com os valores previstos de cada subsérie utilizando o algoritmo de Holt-Winters obtêm-se os resultados evidenciados na Figura 12. Da mesma forma, utilizando-se os valores previstos das subséries de baixa e alta frequência com modelos ARIMA, vistos na Figura 13, tem-se o resultado apresentado na Figura 14.

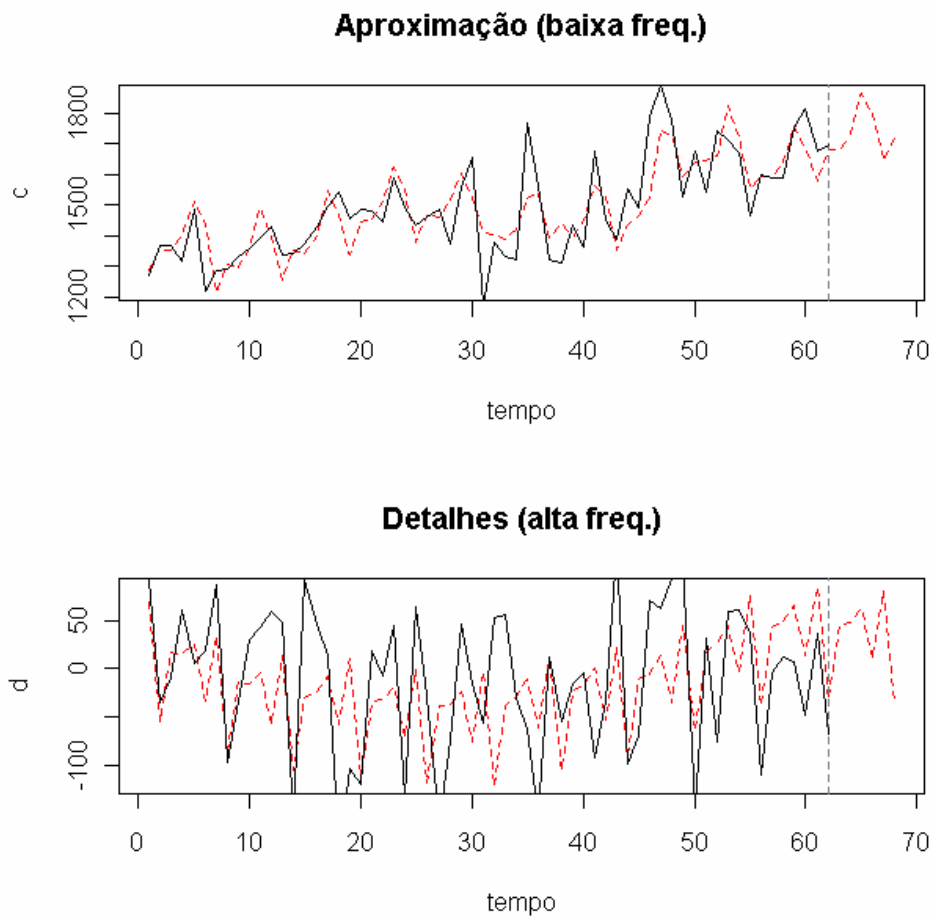


Figura 11 – Valores reais e previstos utilizando o algoritmo de Holt-Winters para as duas subséries da decomposição *wavelet* ($h=6$)

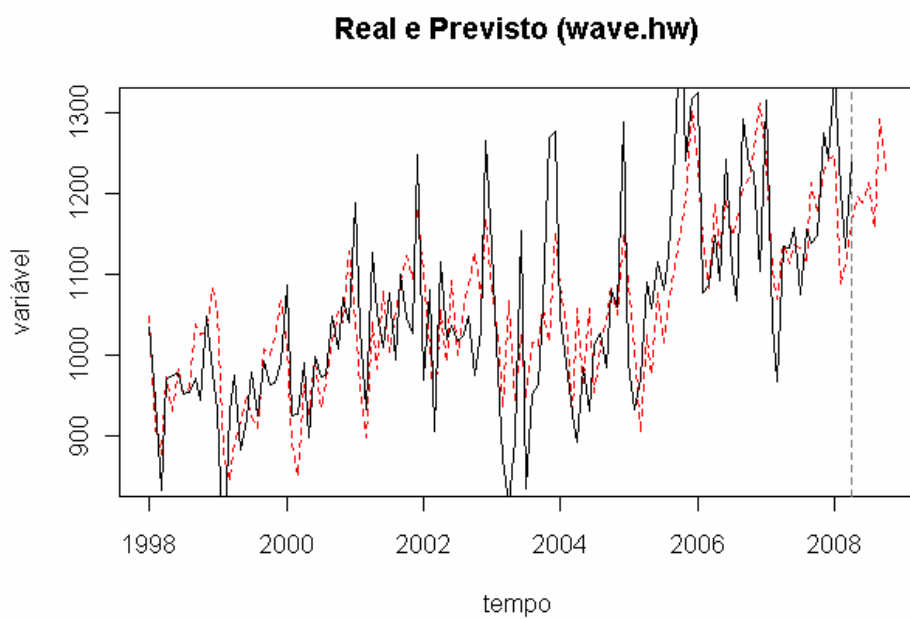


Figura 12 – Valor real e previsto da série ICMS-RS utilizando algoritmo de Holt-Winters com decomposição *wavelet* ($h=6$)

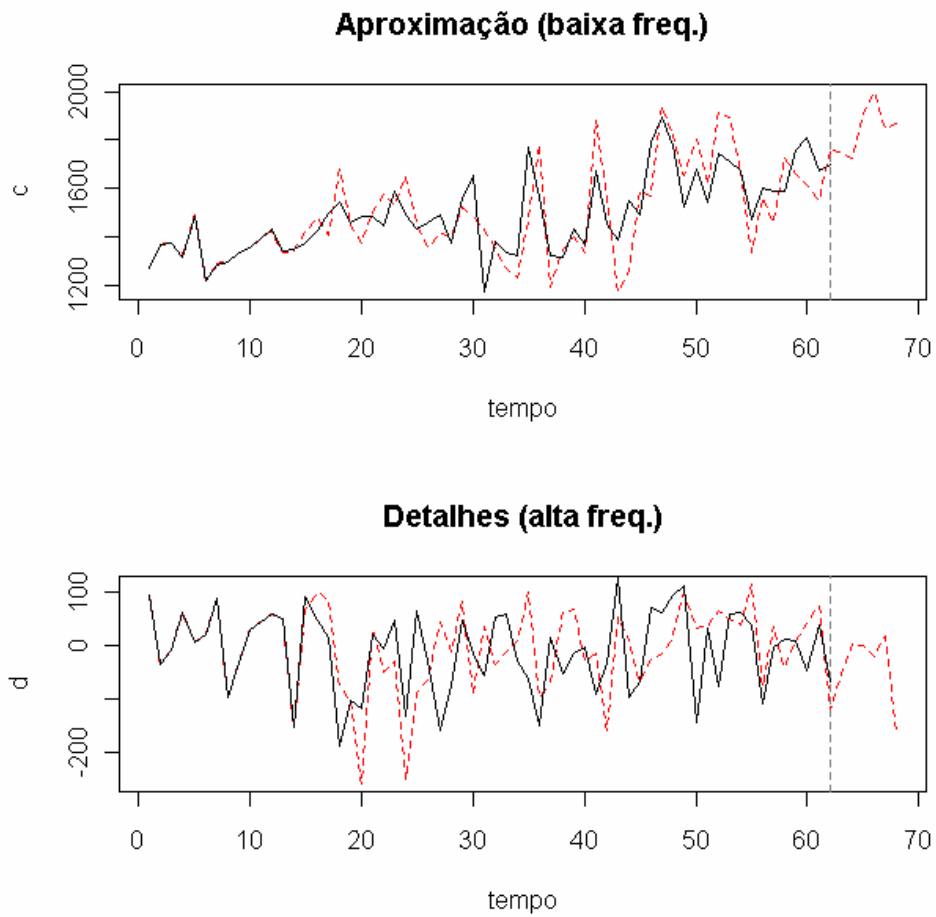


Figura 13 - Valores reais e previstos utilizando modelos ARIMA para as duas subséries da decomposição *wavelet* ($h=6$)

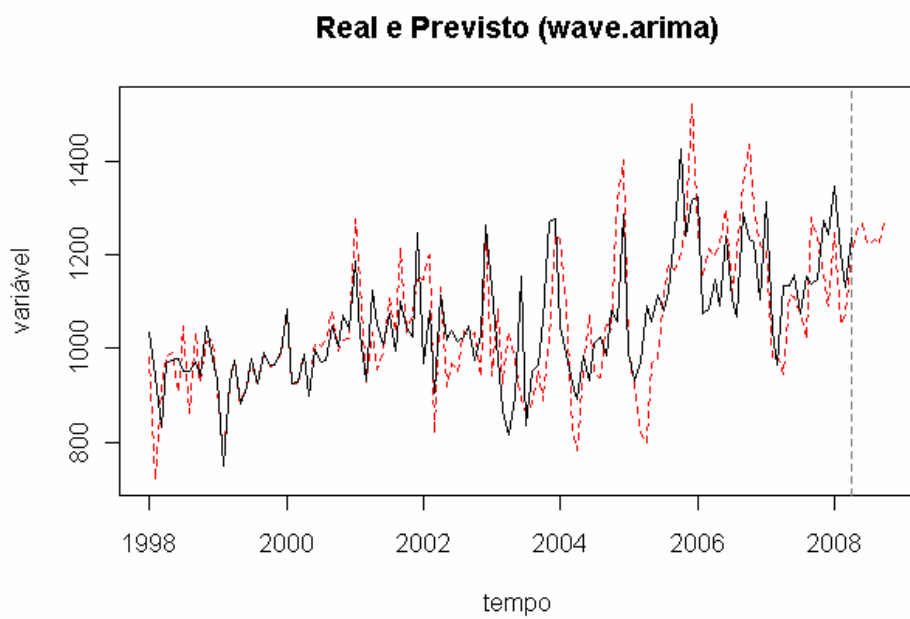


Figura 14 - Valor real e previsto da série ICMS-RS utilizando modelos SARIMA e decomposição *wavelet* ($h=6$)

A partir dos resultados das técnicas apresentadas acima obtém-se os critérios de avaliação, que serão apresentados na seção 4.5 juntamente com as demais técnicas de previsão e diferentes valores de h .

4.5 Comparação das diferentes técnicas de previsão

Nesta seção serão comparadas as diferentes técnicas de previsão exploradas nesta pesquisa. Esta comparação é feita utilizando-se os critérios de avaliação EQM e MAPE para diferentes valores de h . Para facilitar a exposição dos resultados são utilizados os rótulos *wave.hw* e *wave.arima* para as técnicas de previsão com decomposição *wavelet* que utilizam, respectivamente, o algoritmo de Holt-Winters e modelos da classe ARIMA para prever as subséries de alta e baixa frequência. Os resultados apresentam-se na Tabela 6.

Os resultados da comparação dos critérios de avaliação mostram que para $h = 2$ e $h = 12$ o método *wave.hw* foi o que obteve melhor resultado, alcançando um erro absoluto percentual médio abaixo de 1% no caso de $h = 2$. Já para $h = 6$ o modelo SARIMA obteve ligeira superioridade em termos de MAPE, seguido de perto pelo *wave.hw*. Porém, em termos de EQM o método *wave.hw* foi o que obteve pior resultado para $h = 6$.

Um resultado importante da comparação das técnicas quando aplicadas à série de arrecadação do ICMS-RS é que os modelos SARIMA, quando utilizados diretamente na série original, obtiveram melhores resultados do que quando utilizados juntamente com a decomposição *wavelet*. Por outro lado, o algoritmo de Holt-Winters obteve melhoras significativas em todos os diferentes horizontes de previsão quando utilizado conjuntamente com a decomposição *wavelet*. Este é um resultado positivo para a presente pesquisa, pois o diferencial da mesma se dá justamente na proposta conjunta de *wavelets* com algoritmos de alisamento exponencial, já que trabalhos utilizando decomposição *wavelet* com modelos ARIMA já são encontrados em periódicos internacionais.

Outro fato importante a favor da técnica *wave.hw* é o fato dela ser muito mais simples computacionalmente, já que, para a estimação dos parâmetros dos modelos ARIMA são utilizados métodos numéricos para maximizar as funções de verossimilhança que, em geral, são funções complicadas para tal fim. Além disso, outro fator positivo para a técnica *wave.hw* é o fato de que os algoritmos de Holt-Winters não possuem tantos pressupostos teóricos, que precisam ser testados, como na modelagem Box-Jenkins.

Tabela 6 – Comparação das diferentes técnicas de previsão

Técnica	EQM	MAPE
<i>h = 2</i>		
SARIMA	286,936	0,013
Holt-Winters	187,521	0,010
wave.hw	122,656	0,008
wave.arima	530,221	0,014
<i>h = 6</i>		
SARIMA	3281,957	0,034
Holt-Winters	3862,000	0,038
wave.hw	5323,000	0,037
wave.arima	3719,000	0,044
<i>h = 12</i>		
SARIMA	16380,000	0,095
Holt-Winters	12912,000	0,083
wave.hw	10956,000	0,076
wave.arima	28353,000	0,130

Como forma de melhor avaliar as previsões das diferentes técnicas, apresentam-se os valores reais e previstos nas Figura 15, Figura 16 e Figura 17, para $h = 2$, 6 e 12, respectivamente.

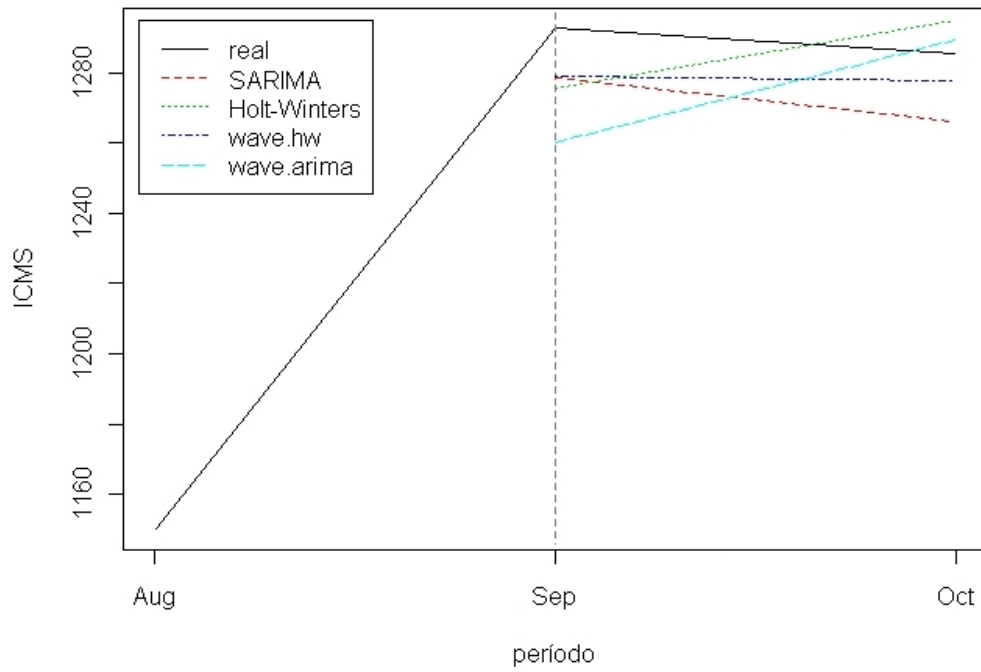


Figura 15 – Valores reais e previstos para $h = 2$

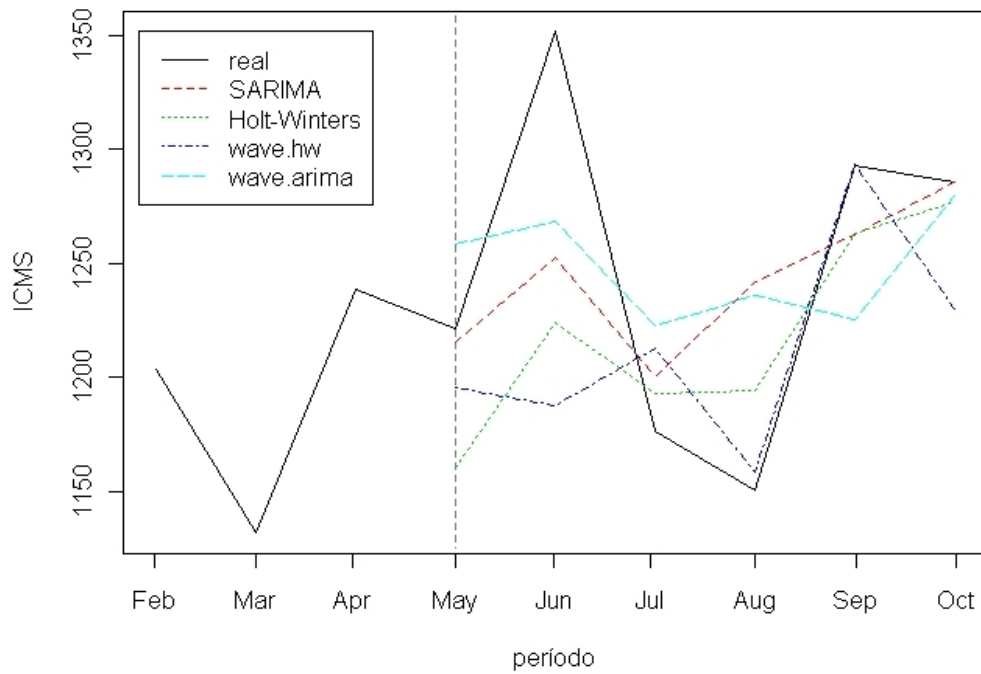


Figura 16 - Valores reais e previstos para $h = 6$

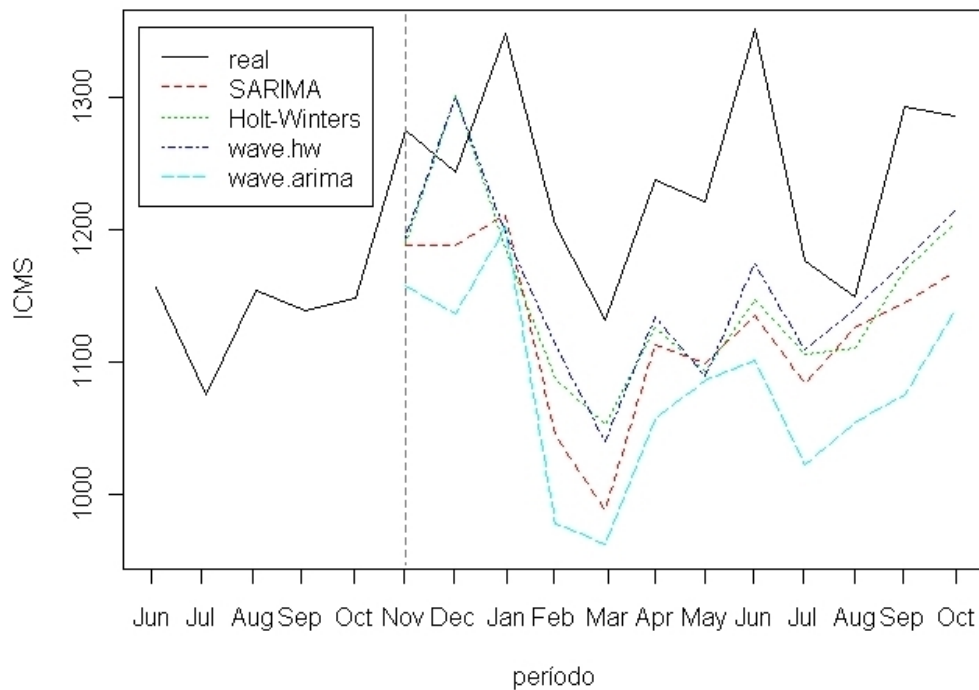


Figura 17 - Valores reais e previstos para $h = 12$

Como já discutido nos resultados da Tabela 6, pela análise gráfica das previsões conclui-se que a técnica *wave.hw* obteve bons resultados, principalmente para os horizontes de previsão de tamanhos 2 e 12, obtendo para estes horizontes as melhores previsões.

Tem-se como um dos objetivos do trabalho traçar previsões para os próximos 12 meses de arrecadação do ICMS utilizando a melhor técnica encontrada e, pelo fato da técnica *wave.hw* ter obtido os melhores resultados para este horizonte de interesse ($h = 12$), utiliza-se esta técnica na próxima seção para prever os próximos 12 meses de arrecadação de ICMS no estado do Rio Grande do Sul.

4.6 Previsão para os próximos meses 12 meses

Considerando os resultados da comparação dos modelos de previsão para a série do ICMS-RS da seção anterior, nesta seção são apresentadas as previsões para os próximos doze meses de arrecadação do imposto, utilizando-se a decomposição *wavelet* juntamente com o

algoritmo de Holt-Winters, a qual foi denominada wave.hw, pois este é apontado como o melhor método de previsão para o horizonte $h = 12$.

Os resultados numéricos das previsões encontram-se na Tabela 7. As mesmas previsões são apresentadas de modo gráfico, juntamente com o gráfico da série no período amostrado, na Figura 18.

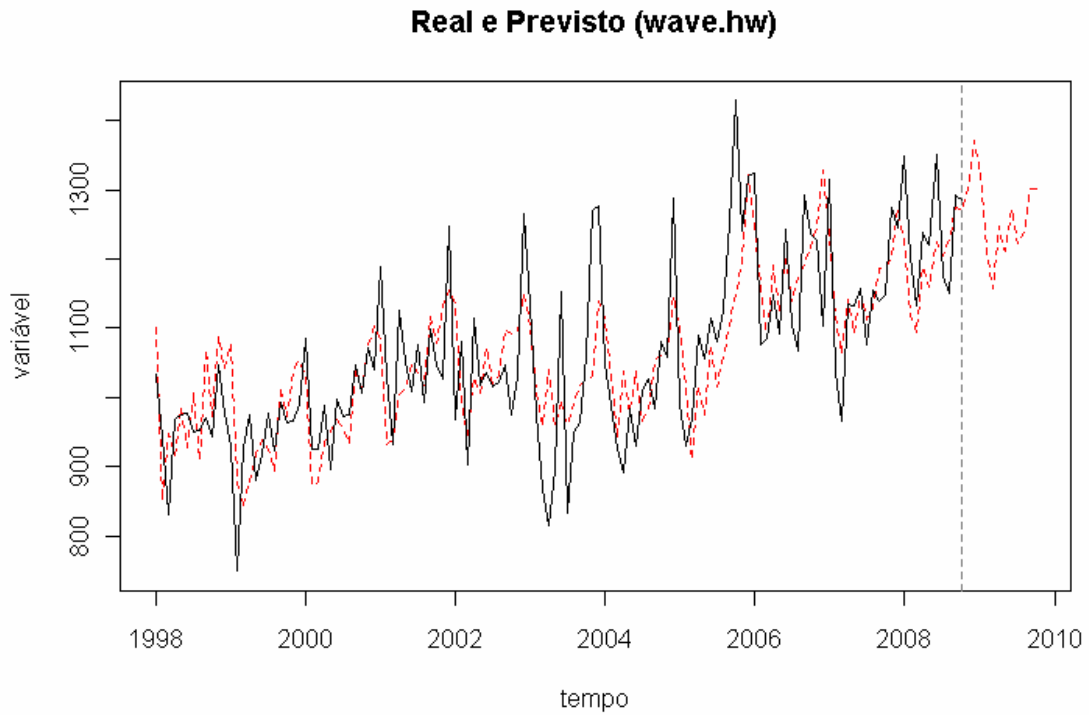


Figura 18 – Previsão para os próximos 12 meses de arrecadação do ICMS-RS

Tabela 7 - Previsão para os próximos 12 meses de arrecadação do ICMS-RS

Período	Previsão
Nov/2008	1299,454
Dez/2008	1370,998
Jan/2009	1324,259
Fev/2009	1212,397
Mar/2009	1157,171
Abr/2009	1246,996
Mai/2009	1210,724
Jun/2009	1271,241
Jul/2009	1223,013
Ago/2009	1238,016
Set/2009	1300,648
Out/2009	1301,927

5 CONSIDERAÇÕES FINAIS E RECOMENDAÇÕES

Neste trabalho efetuou-se uma comparação de métodos tradicionais de previsão com uma proposta de uso conjunto destes métodos com a decomposição *wavelet*. A comparação foi feita por meio de um estudo empírico aplicado à série de arrecadação mensal do ICMS no estado do Rio Grande do Sul. A série em estudo compreende o período de janeiro de 1998 a outubro de 2008 a preços constantes de outubro de 2008 utilizando o IGP-DI.

A série do ICMS-RS possui padrão sazonal visível em seu comportamento no decorrer do período. Com isso, foram utilizados modelos que procuram captar o comportamento sazonal da série temporal. Dentre os modelos de alisamento exponencial foi utilizado o algoritmo de Holt-Winters aditivo e, dentre os modelos da classe ARIMA, foram utilizados os modelos SARIMA. Estes mesmos modelos foram utilizados para a modelagem e previsão de cada uma das subséries da decomposição *wavelet*.

Um ponto forte do trabalho foi a utilização de um *software* livre para o desenvolvimento da pesquisa e, principalmente, a criação de funções para o programa R, nas quais podem ser reproduzidas por quem tiver interesse, podendo ser modificadas de acordo com a necessidade do pesquisador. Todas as funções criadas e modificadas podem ser consultadas no Apêndice.

As técnicas de previsão foram comparadas mediante diferentes horizontes de previsão e pôde-se verificar que a proposta do uso conjunto dos algoritmos de Holt-Winters com a decomposição *wavelet* originou uma melhora significativa às previsões, principalmente para os horizontes de previsão $h = 2$ e $h = 12$. Já o uso de modelos da classe ARIMA juntamente com a decomposição *wavelet* não obteve resultados superiores aos modelos ARIMA utilizados diretamente para a série original.

Esta pesquisa tornou-se importante, pois tratou de uma abordagem exploratória de diferentes modelos de previsão, juntamente com uma proposta de uso conjunto de modelos tradicionais de previsão com decomposição *wavelet*, proposta esta que resultou em previsões mais acuradas. Com isso, uma vez que o ICMS é o principal imposto estadual, o uso desta técnica viabiliza uma melhor tomada de decisão no âmbito das estratégias governamentais no que diz respeito ao planejamento orçamentário do estado.

Deixam-se como sugestões para trabalhos futuros a utilização de Redes Neurais, assim como um estudo exploratório e comparativo do uso de diferentes funções *wavelets* além da *wavelet* de Daubechies com 8 momentos nulos (DB(8)).

REFERÊNCIAS

AKAIKE, H.. A new look at the statistical model identification. **IEEE Transaction on Automatic Control**, AC-19, 1974. p. 716-723.

AKAIKE, H.. A Bayesian analysis of the minimum AIC procedure. **Annals of the Institute of Statistical Mathematics A**, 30, 1978. p.9-14.

ANSLEY, C. F.; NEWBOLD, P.. Finite sample properties of estimators for autoregressive moving average models. **Journal of Econometrics**, 13, 1980. p. 159-183.

ARINO, M. A.. Time series forecasts via wavelets: an application to cars Sales in the Spanish market. Discussion Paper 95-30, ISDS, Duke University.
(<http://www.isds.duke.edu/research/papers/>)

BOX, G. E. P.; PIERCE, D. A.. Distribution of autocorrelations in autoregressive moving average models. **Journal of the American Statistical Association**, 65, 1970. p. 1509-1526.

BOX, G. E. P.; JENKINS, G. M. **Time Series Analysis: Forecasting and Control**. San Francisco: Holden-Day, 1970.

BOX, G. E. P.; JENKINS, G. M.; REINSEL, G.. **Time Series Analysis: Forecasting and Control**. Third Edition. Englewood Cliffs: Prentice Hall, 1994.

BRASIL. Lei n. 101 de 04 de Maio de 2000. Estabelece normas de finanças públicas voltadas para a responsabilidade na gestão fiscal e dá outras providências. **Diário Oficial [da] República Federativa do Brasil**. Brasília, DF, 2000. Disponível em:
<http://www.planalto.gov.br/ccivil_03/Leis/LCP/lcp101.htm> Acesso em 28 de dez. de 2008.

BROCKWELL, P.J.; Davis, R.A.. **Introduction to Time Series and Forecasting**, Second Edition. Springer-Verlag, New York, 2002.

CHEN, Z.; YANG, Y.. Time Series Models for Forecasting: Testing or Combining?. **Studies in Nonlinear Dynamics & Econometrics**, Berkeley Electronic Press, vol. 11(1), 2007, p. 1385-1385.

CLEMEN, R.. Combining forecasts: a review and annotated bibliography. **International Journal of Forecasting**, 5, 1989, p. 559–583.

CORDEIRO, D. M.; CORDEIRO, G. M. Modelo de Previsão Combinada: Uma Aplicação Série Mensal das Notificações de Dengue no Estado de Pernambuco. **Rev. Matemática e Estatística**, São Paulo, v.22, n.3, 2004, p.57-80.

DICKEY, D.; Fuller, W.. Distribution of the estimators for auto-regressive time series with unit root. **Journal of the American Statistical Association**, v. 74, n. 366, Jun. 1979. p. 427-431.

DICKEY, D. and Fuller, W.. Likelihood ratio statistics for auto-regressive time series with a unit root. **Econometrica**, v. 49, n. 4, Jun. 1981. p. 1057-1072.

DIEBOLD, F.. The past, present, and future of macroeconomic forecasting. **Journal of Economic Perspectives**, 12, 1998, p. 175–192.

DOOLEY, G., LENIHAN, H.. An assessment of time series methods in metal price forecasting. **Resources Policy**, 30, 2005, p. 208–217.

FERNADEZ, V. Wavelet- and SVM-based forecasts: An analysis of the U.S. metal and materials manufacturing industry. **Resources Policy**, 32, 2007, p. 80–89

GENÇAY, R.; SELÇUK, F.; WHITCHER, B. **An introduction to wavelets and other filtering methods in finance and economics**. New York: Academic Press, 2002. 359 p.

GENÇAY, R., SELÇUK, F., WHITCHER, B.. Multiscale systematic risk. **Journal of International Money and Finance**, 24 (1), 2005, p. 55–70.

GIL, A. C. **Como Elaborar Projetos de Pesquisa**. 4 ed. Atlas, São Paulo, 2006. 176 p.

HAMILTON, J.D. **Time Series analysis**. Princeton University Press, Princeton – New Jersey, N.J. 1994.

HOMSY, G. V.; PORTUGAL, M. S.; ARAÚJO, J. P. **Wavelets e previsões de séries de tempo: uma análise empírica**. Porto Alegre, UFRGS, 2000.

HURVICH, C. M.; Tsai, C. L..Regression and time series model selection in small samples. **Biometrika**, v. 76, n. 2, 1989. p. 297-307.

LABYS, W., 1999. **Modeling Mineral and Energy Markets**. Kluwer Academic Press, USA.

LIMA, F. G. **Um método de análise e previsão de sucessões cronológicas unidimensionais lineares e não-lineares**. 2004. 221f. Tese de Doutorado (Doutorado em Administração) – Universidade de São Paulo, São Paulo, 2004.

LIN, S.-J.; STEVENSON, M.. Wavelet Analysis of the Cost-of-Carry Model. **Studies in Nonlinear Dynamics & Econometrics**: Vol. 5: No. 1, Article 7, 2004.

LJUNG, G.; BOX, G. E. P.. On a measure of lack of fit in time series models. **Biometrika**, 65, 297-303, 1978.

MALLAT, S.G. A Theory for Multiresolution Signal Decomposition, the Wavelet Representation. **IEEE, Transactions on Pattern Analysis and Machine Intelligence**. v.11, n.7, jul. 1989. p. 574-693, 1989.

MORANA, C.. A semiparametric approach to short-term oil price forecasting. **Energy Economics**. 23 (3), 2001, p. 325–338.

MORETTIN, P. A. **Econometria Financeira: Um curso em séries temporais financeiras**. 17º Simpósio Nacional de Probabilidade e Estatística, ABE, Caxambu, 2006.

MORETTIN, P. A.; TOLOI, C. M. C.. **Análise de Séries Temporais**. Segunda Edição. São Paulo: Editora E. Blücher-Associação Brasileira de Estatística, 2006. 564p.

NIELSEN, O.M. **Wavelets in Scientific Computing**. 1998. 232f. Ph.D. Thesis, Technical University of Denmark, 1998.

ROCHE, J.. **Forecasting Commodity Markets**. Probus Publishing Company, London, 1995.

SAID, D.; DICKEY, D.. Testing for unit roots in autoregressive-moving average models of unknown order. **Biometrika**, v. 71, n. 3, 1984. p. 599-607.

SCHWARZ, G.. Estimating the dimension of a model. **The Annals of Statistics**, v.6, n.2, 1978. p. 461-464.

SERVIÇO BRASILEIRO DE APOIO À MICRO E PEQUENAS EMPRESAS – SEBRAE. **O que é ICMS?**. São Paulo, 2008. Disponível em :
<http://www.sebraesp.com.br/faq/contabilidade/legislacao_enquadramento_legalizacao/icms> Acesso em 28 de dez. de 2008.

TAK, B. **A new method for forecasting stock prices using artificial neural network and wavelet theory**. Ph.D. Thesis (Doutorado em Economia). Universidade da Pensilvânia, Estados Unidos da América, 1995.

WINTERS, P. R.. Forecasting Sales By Exponentially Weighted Moving Average. **Management Science**. v. 6, 1960. p. 324-342.

WINKLHOFER, H.; WITT, S. F.; DIAMATOPOULOS, A. Forecasting practice: a review of the empirical literature and an agenda for the future research. **International Journal of Forecasting**, Elsevier Science B, v. 1, n. 12, p. 193-221, 1996.

WONG, H.; IP, W.; XIE, Z; LUI, X. Modelling and forecasting by wavelets, and the application to exchange rates. **Journal of Applied Statistics**, v. 30, n. 5, 2003. p. 537-553.

ZOU, H.; YANG, Y. Combining time series models for forecast. **International Journal of Forecasting**. n. 20, 2004. p. 69-84.

Apêndice A – Funções implementadas em R

```

# Este arquivo possui as funções wave.hw e wave.arima. Estas decompõem a série
# temporal em duas subséries de alta e baixa frequência utilizando wavelet de
# Daubechies, faz a previsão destas subséries utilizando o algoritmo de
# Holt-Winters (wave.hw) ou modelos SARIMA (wave.arima) e então faz a transformada
# inversa obtendo uma previsão para a série original. As funções recebem de
# entrada uma série temporal e o número de previsões passos a frente.

wave.hw<-function(c,previsao=12,N=8)
{
  source("ffiltros.r") # ANTES DE RODAR TEM QUE RODAR O ARQUIVO Função Filtros.r

  ### Funções para pegar os filtros
  a<-f.a(N)
  b<-f.b(a)
  ### Agora já tenho os filtros a e b

  D<-2*N # número de momentos nulos
  niveis<-1 # número de níveis para análise multiresolução
  frequencia<-frequency(c) # número de observações dentro de um ano
  data.inicio<-start(c)

  fim<-length(c) # número de elementos da série de entrada

  inicio<-1
  if(fim%%2 == 1)
  { inicio<-2
    print("MENSAGEM:")
    print("Foi retirada a primeira observação, pois o número de elementos da
      série era impar!")
  }

  c.aux<-t(as.matrix(c)) #transforma o arquivo lido em matriz realmente
  c0<-cbind(c.aux[inicio:fim])

  cl<-c0 #cl é para utilizar no meio do caminho. Ele vai recebendo os detalhes
  tamanho1<-length(c0) #conta o número de observações

  cj<-c0
  for(j in (niveis-1):0) #começa o laço da MRA
  {
    tamanho2<-length(cj) #número de observações no nível
    hwa.aux1<-hw(ts(as.vector(cj),frequency=frequencia), h=D,
      seasonal="additive") # ajuste holt-winters (hw) auxiliar
    cj<-c(cj,hwa.aux1$mean) # usa as previsões hw para os D valores além dos dados
    #cj<-c(cj,cj[1:D]) #periodicidade circular

    ### *****
    ### Inicio da transformada direta ortonormal (Daubechies)

    for(l in 0:((tamanho2/2)-1))
    { c1<-0 # inicializa vetor aproximações
      for(k in 0:(D-1))
      { c1<-c1+a[k+1]*cj[(2*k)+k+1] #1ª formula para TWD
      }
      if(l==0)
        cj_1<-c1
      else
        cj_1<-c(cj_1,c1)

    ### agora são os detalhes

    d1<-0 #inicializa vetor detalhes
    for(k in 0:(D-1))
    { d1<-d1+b[k+1]*cj[(2*k)+k+1] #2ª formula para TWDD
    }
    if(l==0)

```

```

    dj_1<-d1
  else
    dj_1<-c(dj_1,d1)
}

### fim da transformada direta ortonormal
### *****

for(m in 1:(tamanho2/2)) # coloca vetor aproximação/vetor detalhes dentro
  # do vetor principal
  {
    cl[m]<-cj_1[m] #vetor aproximação no inicio
    cl[(tamanho2/2)+m]<-dj_1[m] #vetor detalhes no fim
  }
  cj<-cj_1 #atualização para a nova interação
}

cl_geral<-cl #para guardar o cl para todos os thresholds

### *****
### Inicio da previsão nas diferentes subséries

tamanho3<-length(cl)

aprox<-cl[1:(tamanho3/2)]
detalhes<-cl[((tamanho3/2)+1):tamanho3]

# previsão para frente e para trás da parte grosseira
hwa.aux2<-hw(ts(as.vector(aprox),frequency=(frequencia/2)), h=(previsao+D),
  seasonal="additive") # previsão para frente
hwa.aux3<-hw(ts(as.vector(rev(aprox)),frequency=(frequencia/2)), h=D,
  seasonal="additive") # previsão para trás

# previsão para frente e para trás dos detalhes
hwa.aux4<-hw(ts(as.vector(detalhes),frequency=(frequencia/2)), h=(previsao+D),
  seasonal="additive") # previsão para frente
hwa.aux5<-hw(ts(as.vector(rev(detalhes)),frequency=(frequencia/2)), h=D,
  seasonal="additive") # previsão para trás

aprox.prev<-c(rev(hwa.aux3$mean),hwa.aux2$fitted,hwa.aux2$mean)
detalhes.prev<-c(rev(hwa.aux5$mean),hwa.aux4$fitted,hwa.aux4$mean)

#aprox.prev<-c(rev(hwa.aux3$mean),aprox,hwa.aux2$mean)
#detalhes.prev<-c(rev(hwa.aux5$mean),detalhes,hwa.aux4$mean)

### Fim da previsão nas diferentes subséries
### *****

### *****
### Inicio da transformada Inversa (Daubechies)

cj_1<-aprox.prev

dj_1<-detalhes.prev
tamanho4<-(tamanho1/2)+(previsao/2)

for(l in 0:(2*tamanho4-1))
{
  n1<-ceiling((l-(D)+1)/2)
  n2<-floor(l/2)
  ci<-0
  for(n in n1:n2)
  {
    ci<-ci+cj_1[n+1+D]*a[(l-2*n)+1]+dj_1[n+1+D]*b[(l-2*n)+1]
  }
}

```

```

    if(l==0)
      sinal_previsao<-ci
    else
      sinal_previsao<-c(sinal_previsao,ci)
  }
  ts.previsao<-ts(sinal_previsao,start=c(data.inicio[1],data.inicio[2]),
    frequency=frequencia)

### Fim da transformada Inversa (Daubechies)
### *****

par(mfrow=c(2,1))
plot(aprox,main="Aproximação (baixa freq.)",ylab="c",xlab="tempo",type="l")
plot(detalhes,main="Detalhes (alta freq.)",ylab="d",xlab="tempo",type="l")

plot(aprox.prev[(D+1):(length(aprox.prev)-D)],main="Aproximação (baixa freq.)",
  ylab="c",xlab="tempo",type="l",lty=2,col="red")
abline(v = (tamanho1/2), col = gray(.50),lty=2)
lines(aprox,main="Aproximação (baixa freq.)",lty=2,col="red")

plot(detalhes.prev[(D+1):(length(detalhes.prev)-D)],main="Detalhes (alta freq.)",
  ylab="d",xlab="tempo",type="l",lty=2,col="red")
abline(v = (tamanho1/2), col = gray(.50),lty=2)
lines(detalhes,main="Detalhes (alta freq.)",lty=2,col="red")

data.fim<-end(c)
v.abline<-data.fim[1]+((data.fim[2]-1)/12)
v.abline
par(mfrow=c(1,1))
plot(ts.previsao,main="Real e Previsto (wave.hw)",lty=2,col="red",xlab="tempo",
  ylab="variável",ylim = c(min(c),max(c)))
abline(v = v.abline, col = gray(.50),lty=2)
lines(c)

return(ts.previsao[(tamanho1+1):(tamanho1+previsao)])
} #fim da função wave.hw

#####
#####

wave.arima<-function(c,previsao=12,N=8)
{
  source("ffiltros.r") # ANTES DE RODAR TEM QUE RODAR O ARQUIVO Função Filtros.r

  ### Funções para pegar os filtros
  a<-f.a(N)
  b<-f.b(a)
  ### Agora já tenho os filtros a e b

  D<-2*N # número de momentos nulos
  niveis<-1 # número de níveis para análise multiresolução
  frequencia<-frequency(c) # número de observações dentro de um ano
  data.inicio<-start(c)

  fim<-length(c) # número de elementos da série de entrada

  inicio<-1
  if(fim%%2 == 1)
  { inicio<-2
    print("MENSAGEM:")
    print("Foi retirada a primeira observação, pois o número de elementos da
      série era impar!")
  }
}

```

```

c.aux<-t(as.matrix(c)) #transforma o arquivo lido em matriz realmente
c0<-cbind(c.aux[inicio:fim])

cl<-c0 #cl é para utilizar no meio do caminho. Ele vai recebendo os detalhes
tamanho1<-length(c0) #conta o número de observações

cj<-c0
for(j in (niveis-1):0) #começa o laço da MRA
{
  tamanho2<-length(cj) #número de observações no nível
  hwa.aux1<-hw(ts(as.vector(cj),frequency=frequencia), h=D,
    seasonal="additive") # ajuste holt-winters (hw) auxiliar
  cj<-c(cj,hwa.aux1$mean) # usa as previsões hw para os D valores além dos dados
  #cj<-c(cj,cj[1:D]) #periodicidade circular

  ### *****
  ### Inicio da transformada direta ortonormal (Daubechies)

  for(l in 0:((tamanho2/2)-1))
  { c1<-0 # inicializa vetor aproximações
    for(k in 0:(D-1))
    { c1<-c1+a[k+1]*cj[(2*k)+k+1] #1ª formula para TWD
    }
    if(l==0)
      cj_1<-c1
    else
      cj_1<-c(cj_1,c1)

    ### agora são os detalhes

    d1<-0 #inicializa vetor detalhes
    for(k in 0:(D-1))
    { d1<-d1+b[k+1]*cj[(2*k)+k+1] #2ª formula para TWDD
    }
    if(l==0)
      dj_1<-d1
    else
      dj_1<-c(dj_1,d1)
  }

  ### fim da transformada direta ortonormal
  ### *****

  for(m in 1:(tamanho2/2)) # coloca vetor aproximação/vetor detalhes dentro
    # do vetor principal
  { cl[m]<-cj_1[m] #vetor aproximação no inicio
    cl[(tamanho2/2)+m]<-dj_1[m] #vetor detalhes no fim
  }
  cj<-cj_1 #atualização para a nova interação
}

cl_geral<-cl #para guardar o cl para todos os thresholds

### *****
### Inicio da previsão nas diferentes subséries

tamanho3<-length(cl)

aprox<-cl[1:(tamanho3/2)]
detalhes<-cl[((tamanho3/2)+1):tamanho3]

# previsão para frente e para trás da parte grosseira
hwa.aux2<-auto.sarima(ts(as.vector(aprox),frequency=(frequencia/2)),
  f=(frequencia/2)) # previsão para frente
hwa.aux3<-auto.sarima(ts(as.vector(rev(aprox)),frequency=(frequencia/2)),

```

```

f=(frequencia/2)) # previsão para trás

# previsão para frente e para trás dos detalhes
hwa.aux4<-auto.sarima(ts(as.vector(detalhes),frequency=(frequencia/2)),
  f=(frequencia/2)) # previsão para frente
hwa.aux5<-auto.sarima(ts(as.vector(rev(detalhes)),frequency=(frequencia/2)),
  f=(frequencia/2)) # previsão para trás

fit.sarima2<-fitted.Arima(hwa.aux2)
fit.sarima4<-fitted.Arima(hwa.aux4)

pred.sarima2<-forecast(hwa.aux2, h=((previsao+D)))
pred.sarima3<-forecast(hwa.aux3, h=(D))
pred.sarima4<-forecast(hwa.aux4, h=((previsao+D)))
pred.sarima5<-forecast(hwa.aux5, h=(D))

aprox.prev<-c(rev(pred.sarima3$mean),fit.sarima2,pred.sarima2$mean)
detalhes.prev<-c(rev(pred.sarima5$mean),fit.sarima4,pred.sarima4$mean)

### Fim da previsão nas diferentes subséries
### *****

### *****
### Inicio da transformada Inversa (Daubechies) *****

cj_1<-aprox.prev

dj_1<-detalhes.prev
tamanho4<-(tamanho1/2)+(previsao/2)

for(l in 0:(2*tamanho4-1))
{ n1<-ceiling((l-(D)+1)/2)
  n2<-floor(l/2)
  ci<-0
  for(n in n1:n2)
  {
    ci<-ci+cj_1[n+1+D]*a[(l-2*n)+1]+dj_1[n+1+D]*b[(l-2*n)+1]
  }
  if(l==0)
    sinal_previsao<-ci
  else
    sinal_previsao<-c(sinal_previsao,ci)
}
ts.previsao<-ts(sinal_previsao,start=c(data.inicio[1],data.inicio[2]),
  frequency=frequencia)

### Fim da transformada Inversa (Daubechies) *****
### *****

par(mfrow=c(2,1))
plot(aprox,main="Aproximação (baixa freq.)",ylab="c",xlab="tempo",type="l")
plot(detalhes,main="Detalhes (alta freq.)",ylab="d",xlab="tempo",type="l")

plot(aprox.prev[(D+1):(length(aprox.prev)-D)],main="Aproximação (baixa freq.)",
  ylab="c",xlab="tempo",type="l",lty=2,col="red")
abline(v = (tamanho1/2), col = gray(.50),lty=2)
lines(aprox,main="Aproximação (baixa freq.)",lty=2,col="red")

plot(detalhes.prev[(D+1):(length(detalhes.prev)-D)],main="Detalhes (alta freq.)",
  ylab="d",xlab="tempo",type="l",lty=2,col="red")
abline(v = (tamanho1/2), col = gray(.50),lty=2)
lines(detalhes,main="Detalhes (alta freq.)",lty=2,col="red")

data.fim<-end(c)
v.abline<-data.fim[1]+((data.fim[2]-1)/12)
v.abline

```

```

par(mfrow=c(1,1))
plot(ts.previsao,main="Real e Previsto (wave.arima)",lty=2,col="red",
     xlab="tempo",ylab="variável",ylim = c(min(c),max(c)))
abline(v = v.abline, col = gray(.50),lty=2)
lines(c)

return(ts.previsao[(tamanho1+1):(tamanho1+previsao)])
} #fim da função wave.arima

#####
#####

# Essa função escolhe o melhor valor de k para o teste adf e retorna
# o teste feito!
# Os valores de entrada da função são idênticos aos da função original ur.df(urca),
# porém sem precisar informar o lag=k e informando em crit se deseja usar "AIC" ou "BIC".
# Se não informar qual critério usar ele usará o AIC, mas dará warnings.
# Esta função precisa da função ur.df.aic que é a função ur.df(urca) na qual eu
# modifiquei para imprimir o aic (ou BIC) da regressão em um arquivo.

bestk.adf<-function(y, type1 = c("none", "drift", "trend"),crit1=c("AIC","BIC"))
{
  source("ur.df.aic.r") # precisa da função que eu modifiquei
  print(crit1)
  if(crit1 == "BIC") # se optar por BIC precisa do pacote stats4
    library(stats4)

  T<-length(y)
  k.max<-round((12*(T/100)^(1/4)),0) # dá o valor máximo que k pode assumir

  write.table(0, file="aic_ur_df.fab",quote = FALSE,row.names = FALSE,
             col.names = FALSE) # zera o arquivo que vão ser impressos os aic's

  for(i in 1:k.max) # faz o teste com todos os tamanhos de k <=k.max
  { ur.df.aic(y,type=type1,lag=i,crit=crit1)
  }

  aic.ur.df<-read.table("aic_ur_df.fab")
  vec<-as.vector(aic.ur.df[,1])

  min.aic<-vec[2]
  otimo.k<-vec[3]

  for(i in 2:k.max) # vê qual é o valor de k que gerou o menor aic
  { if(vec[2*i] < min.aic)
    { min.aic<-vec[2*i]
      otimo.k<-vec[(2*i)+1]
    }
  }

  print(c("O k máximo é:",k.max))
  print(c("O valor de k que minimiza o critério de escolha é:",otimo.k))
  print("H0: Existe raiz unitária (d=1)")
  print("Interpretação do resultado")
  print("test-statistic < critical values (nível alfa) => Rejeite a hipótese nula")

  resultado<-ur.df(y,type=type1,otimo.k)
  return(summary(resultado))
}

```

```
#####
#####
```

Essa função é a mesma ur.df(urca) só que eu modifiquei para retornar o valor
do aic do modelo considerado

```
ur.df.aic<-function (y, type = c("none", "drift", "trend"), lags = 1,crit=c("AIC","BIC"))
{
  if (ncol(as.matrix(y)) > 1)
    stop("\ny is not a vector or univariate time series.\n")
  if (any(is.na(y)))
    stop("\nNAs in y.\n")
  y <- as.vector(y)
  lag <- as.integer(lags)
  if (lag < 0)
    stop("\nLags must be set to an non negative integer value.\n")
  CALL <- match.call()
  DNAME <- deparse(substitute(y))
  type <- type[1]
  x.name <- deparse(substitute(y))
  lags <- lags + 1
  z <- diff(y)
  n <- length(z)
  x <- embed(z, lags)
  z.diff <- x[, 1]
  z.lag.1 <- y[lags:n]
  tt <- lags:n
  if (lags > 1) {
    z.diff.lag = x[, 2:lags]
    if (type == "none") {
      result <- lm(z.diff ~ z.lag.1 - 1 + z.diff.lag)
      tau <- coef(summary(result))[1, 3]
      teststat <- as.matrix(tau)
      colnames(teststat) <- "tau1"
    }
    if (type == "drift") {
      result <- lm(z.diff ~ z.lag.1 + 1 + z.diff.lag)
      tau <- coef(summary(result))[2, 3]
      phi1.reg <- lm(z.diff ~ -1 + z.diff.lag)
      phi1 <- anova(phi1.reg, result)$F[2]
      teststat <- as.matrix(t(c(tau, phi1)))
      colnames(teststat) <- c("tau2", "phi1")
    }
    if (type == "trend") {
      result <- lm(z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
      tau <- coef(summary(result))[2, 3]
      phi2.reg <- lm(z.diff ~ -1 + z.diff.lag)
      phi3.reg <- lm(z.diff ~ z.diff.lag)
      phi2 <- anova(phi2.reg, result)$F[2]
      phi3 <- anova(phi3.reg, result)$F[2]
      teststat <- as.matrix(t(c(tau, phi2, phi3)))
      colnames(teststat) <- c("tau3", "phi2", "phi3")
    }
  }
  else {
    if (type == "none") {
      result <- lm(z.diff ~ z.lag.1 - 1)
      tau <- coef(summary(result))[1, 3]
      teststat <- as.matrix(tau)
      colnames(teststat) <- "tau1"
    }
    if (type == "drift") {
      result <- lm(z.diff ~ z.lag.1 + 1)
      phi1.reg <- lm(z.diff ~ -1)
      phi1 <- anova(phi1.reg, result)$F[2]
    }
  }
}
```

```

    tau <- coef(summary(result))[2, 3]
    teststat <- as.matrix(t(c(tau, phi1)))
    colnames(teststat) <- c("tau2", "phi1")
  }
  if (type == "trend") {
    result <- lm(z.diff ~ z.lag.1 + 1 + tt)
    phi2.reg <- lm(z.diff ~ -1)
    phi3.reg <- lm(z.diff ~ 1)
    phi2 <- anova(phi2.reg, result)$F[2]
    phi3 <- anova(phi3.reg, result)$F[2]
    tau <- coef(summary(result))[2, 3]
    teststat <- as.matrix(t(c(tau, phi2, phi3)))
    colnames(teststat) <- c("tau3", "phi2", "phi3")
  }
}
rownames(teststat) <- "statistic"
testreg <- summary(result)
res <- residuals(testreg)
if (n < 25)
  rowselec <- 1
if (25 <= n & n < 50)
  rowselec <- 2
if (50 <= n & n < 100)
  rowselec <- 3
if (100 <= n & n < 250)
  rowselec <- 4
if (250 <= n & n < 500)
  rowselec <- 5
if (n >= 500)
  rowselec <- 6
if (type == "none") {
  cval.tau1 <- rbind(c(-2.66, -1.95, -1.6), c(-2.62, -1.95,
    -1.61), c(-2.6, -1.95, -1.61), c(-2.58, -1.95, -1.62),
    c(-2.58, -1.95, -1.62), c(-2.58, -1.95, -1.62))
  cvals <- t(cval.tau1[rowselec, ])
  testnames <- "tau1"
}
if (type == "drift") {
  cval.tau2 <- rbind(c(-3.75, -3, -2.63), c(-3.58, -2.93,
    -2.6), c(-3.51, -2.89, -2.58), c(-3.46, -2.88, -2.57),
    c(-3.44, -2.87, -2.57), c(-3.43, -2.86, -2.57))
  cval.phi1 <- rbind(c(7.88, 5.18, 4.12), c(7.06, 4.86,
    3.94), c(6.7, 4.71, 3.86), c(6.52, 4.63, 3.81), c(6.47,
    4.61, 3.79), c(6.43, 4.59, 3.78))
  cvals <- rbind(cval.tau2[rowselec, ], cval.phi1[rowselec,
    ])
  testnames <- c("tau2", "phi1")
}
if (type == "trend") {
  cval.tau3 <- rbind(c(-4.38, -3.6, -3.24), c(-4.15, -3.5,
    -3.18), c(-4.04, -3.45, -3.15), c(-3.99, -3.43, -3.13),
    c(-3.98, -3.42, -3.13), c(-3.96, -3.41, -3.12))
  cval.phi2 <- rbind(c(8.21, 5.68, 4.67), c(7.02, 5.13,
    4.31), c(6.5, 4.88, 4.16), c(6.22, 4.75, 4.07), c(6.15,
    4.71, 4.05), c(6.09, 4.68, 4.03))
  cval.phi3 <- rbind(c(10.61, 7.24, 5.91), c(9.31, 6.73,
    5.61), c(8.73, 6.49, 5.47), c(8.43, 6.49, 5.47),
    c(8.34, 6.3, 5.36), c(8.27, 6.25, 5.34))
  cvals <- rbind(cval.tau3[rowselec, ], cval.phi2[rowselec,
    ], cval.phi3[rowselec, ])
  testnames <- c("tau3", "phi2", "phi3")
}
}
colnames(cvals) <- c("1pct", "5pct", "10pct")
rownames(cvals) <- testnames

```

Minhas inclusões


```

if(crit=="BIC")
{ crit.df<-BIC(result) # calcula o aic da regressão
  #print("BIC")
  #print(crit.df)
}

if(crit=="AIC")
{ crit.df<-AIC(result) # calcula o aic da regressão
  #print("AIC")
  #print(crit.df)
}

#print(c(aic.df,lag))
write.table(c(crit.df,lag), file="aic_ur_df.fab",append=TRUE,
quote = FALSE,row.names = FALSE,col.names = FALSE) # imprime
#em um arquivo o aic e o valor de k
### Fim das minhas inclusões

new("ur.df", y = y, model = type, cval = cvals, lags = lag,
teststat = teststat, testreg = testreg, res = res,
test.name = "Augmented Dickey-Fuller Test")
}

#####
#####

# Esta função utiliza a função auto.arima(forecast) para encontrar o modelo
# SARIMA que minimiza o critério AIC dentre todos os modelos até o modelo de
# ordem maior igual SARIMA(6,2,6)(2,2,2)

auto.sarima<-function(x,f=12,d.ini=0,d.fim=2)
{ sarima<-auto.arima(x,d=0,D=0,max.p = 6, max.q = 6,
max.P = 2, max.Q = 2, start.p=0, start.q=0, start.P=0, start.Q=0,
stationary = FALSE, ic = "aic", stepwise=FALSE)

aicmin<-sarima$aic
for(d1 in d.ini:d.fim)
{ for(D1 in 0:2)
{ sarima.aux<-auto.arima(x,d=d1,D=D1,max.p = 6, max.q = 6,
max.P = 2, max.Q = 2, start.p=0, start.q=0, start.P=0, start.Q=0,
stationary = FALSE, ic = "aic", stepwise=FALSE,
approximation=length(x)>100 | frequency(x)>f)
if(sarima.aux$aic<aicmin)
{ sarima<-sarima.aux
aicmin<-sarima.aux$aic
}
}
}
}
tsdiag(sarima) #diagnostico dos resíduos do modelo encontrado
return(sarima) # retorna o melhor modelo encontrado
}

#####
#####

## Função que plota as FAC e FACP estimadaas num mesmo gráfico
## Primeiro plota da série e depois da série diferenciada

correlog<-function(y)
{
par(mfrow=c(2,1))

```

```

acf(y,ylab="FAC",main="Série y",xlab="(a)")
pacf(y,ylab="FACP",main="Série y",xlab="(b)")
acf(diff(y),ylab="FAC",main="Série dif(y)",xlab="(a)")
pacf(diff(y),ylab="FACP",main="Série dif(y)",xlab="(b)")

par(mfrow=c(1,1))
}

#####
#####

# Função que retorna algumas medidas de avaliação das previsões.
# Entra-se com o vetor de observações reais e com o vetor de valores previstos
# e ela retorna uma matriz com os critérios de avaliação que podem ser vistos abaixo.

crit<-function(real,prev)
{
  n<-length(real)
  if(n!=length(prev))
  { return(print("ERRO: Os dois vetores devem ter a mesma dimensão!!!"))
  }

  #agreg<-sum(real)-sum(prev) # erro agregado
  #agregp<-agreg/sum(real) # erro agregado percentual
  eqm<-sum((real-prev)^2)/n # erro quadrático médio
  mape<-sum(abs(real-prev)/real)/n # erro absoluto médio percentual

  resultado<-rbind(c("EQM",round(eqm,3)),
                  c("MAPE",round(mape,3)))

  return(resultado)
}

#####
#####

# Essa função recebe de entrada uma matriz com k colunas (variáveis) e
# T linhas (observações) e imprime no prompt do R a matriz com as principais
# estatísticas descritivas.
# Mais no futuro ela também imprimirá um arquivo com a tabela em código tex

descritiva<-function(X)
{ X<-as.matrix(X) # transforma em matriz realmente, assim pode-se entrar com
  # um vetor linha também
  d<-dim(X) #dimensão da matriz
  T<-d[1] # número de linhas (total de observações)
  k<-d[2] # número de colunas (total de variáveis)

  m<-7 #número de medidas descritivas utilizadas

  mat<-c("Variáveis","Média","Mediana","DP","Var","CV","Min","Max")

  for(i in 1:k)
  { linha<-rep(0,(m+1))
    linha[1]<-i
    linha[2]<-round(mean(X[,i]),3)
    linha[3]<-round(median(X[,i]),3)
    linha[4]<-round(sd(X[,i]),3)
    linha[5]<-round(var(X[,i]),3)
    linha[6]<-round(((linha[4]/linha[2])*100),3)
    linha[7]<-round(min(X[,i]),3)
  }
}

```

```

linha[8]<-round(max(X[,i]),3)

mat<-rbind(mat,linha)

}

print(mat)

}

#####

#####

#### Filtros da TW Daubechies conforme N
#### N é o número de momentos nulos da Wavelet de Daubechies ## D=2*N
#### *****

f.a<-function(N)
{
  if (N==1) #filtros db1
    a<-c(0.7071067811865475,0.7071067811865475)
  if (N==2) #filtros db2
    a<-c(0.4829629131445341,0.8365163037378077,0.2241438680420134,-0.1294095225512603)
  if (N==3) #filtros db3
    a<-c(0.3326705529500825,0.8068915093110924,0.4598775021184914,-
0.1350110200102546,-0.0854412738820267,0.0352262918857095)
  if (N==4) #filtros db4
    a<-c(0.2303778133088964,0.7148465705529154,0.6308807679298587,-
0.0279837694168599,-0.1870348117190931,0.0308413818355607,0.0328830116668852,-
0.0105974017850690)
  if (N==5) #filtros db5
    a<-c(0.1601023979741929,0.6038292697971895,0.7243085284377726,0.1384281459013203,-
0.2422948870663823,-0.0322448695846381,0.0775714938400459,-0.0062414902127983,-
0.0125807519990820,0.0033357252854738)
  if (N==6) #filtros db6
    a<-c(0.1115407433501095,0.4946238903984533,0.7511339080210959,0.3152503517091982,-
0.2262646939654400,-0.1297668675672625,0.0975016055873225,0.0275228655303053,-
0.0315820393174862,0.0005538422011614,0.0047772575109455,-0.0010773010853085)
  if (N==7) #filtros db7
    a<-c(0.0778520540850037,0.3965393194818912,0.7291320908461957,0.4697822874051889,-
0.1439060039285212,-0.2240361849938412,0.0713092192668272,0.0806126091510774,-
0.0380299369350104,-0.0165745416306655,0.0125509985560986,0.0004295779729214,-
0.0018016407040473,0.0003537137999745)
  if (N==8) #filtros db8
    a<-c(0.0544158422431072,0.3128715909143166,0.6756307362973195,0.5853546836542159,-
0.0158291052563823,-0.2840155429615824,0.000472484573912,0.1287474266204893,-
0.0173693010018090,-0.0440882539307971,0.0139810279174001,0.0087460940474065,-
0.0048703529934520,-0.0003917403733770,0.0006754494064506,-0.0001174767841248)
  if (N==9) #filtros db9
    a<-c(0.0380779473638778,0.2438346746125858,0.6048231236900955,0.6572880780512736,
0.1331973858249883,-0.2932737832791663,-0.0968407832229492,0.1485407493381256,
0.0307256814793385,-0.0676328290613279,0.0002509471148340,0.0223616621236798,-
0.0047232047577518,-0.0042815036824635,0.0018476468830563,0.0002303857635232,-
0.0002519631889427,0.0000393473203163)
  if (N==10) #filtros db10
    a<-c(0.0266700579005473,0.1881768000776347,0.5272011889315757,0.6884590394534363,
0.2811723436605715,-0.2498464243271598,-0.1959462743772862,0.1273693403357541,
0.0930573646035547,-0.0713941471663501,-0.0294575368218399,0.0332126740593612,
0.0036065535669870,-0.0107331754833007,0.0013953517470688,0.0019924052951925,-
0.0006858566949564,-0.0001164668551285,0.0000935886703202,-0.0000132642028945)

  return(a)
}

```

```
f.b<-function(a)
{ N<-length(a)/2
  b<-((-1)^(0))*a[(2*N)-0]
  for(k in 2:(2*N))
    b<-c(b,((-1)^(k-1))*a[(2*N)-(k-1)])
  return(b)
}
```