

# Utilizando Séries Temporais na Previsão da Arrecadação do Imposto de Renda

Olívio Armando Cordeiro Júnior

Monografia apresentada ao Programa de Pós-  
graduação do Isc/Cefor como parte das exigências  
do curso de Especialização em Orçamento Público.

Brasília  
Distrito Federal - Brasil  
Julho - 2007

ISC / CENTRO DE FORMAÇÃO, TREINAMENTO E APERFEIÇOAMENTO -  
CEFOR

# Utilizando Séries Temporais na Previsão da Arrecadação do Imposto de Renda

**Olívio Armando Cordeiro Júnior**

Orientador: Prof. Jorge Abrahão de Castro

Monografia apresentada ao Programa de Pós-graduação do Isc/Cefor como parte das exigências do curso de Especialização em Orçamento Público.

Brasília  
Distrito Federal - Brasil  
Julho - 2007

ISC / CENTRO DE FORMAÇÃO, TREINAMENTO E APERFEIÇOAMENTO -  
CEFOR

# Agradecimentos

A finalização de um trabalho, mesmo que simples, como uma monografia de um curso de especialização, é sempre uma ocasião repleta de alegria e de satisfação. Nesse momento é apropriado que sejam feitos os agradecimentos àqueles que, direta, ou indiretamente, estiveram associados à etapa que se encerra.

Antes de tudo, aos meus pais, que sempre me incentivaram a trilhar o caminho do estudo e da dedicação. Sem os bons exemplos e o esforço que tiveram em me fornecer a melhor educação que lhes foi possível dar, certamente eu não estaria aqui e minha vida seria infinitamente mais difícil.

Ao Tribunal de Contas da União, que tem me proporcionado o melhor dos ambientes, bem como a possibilidade de estar sempre sempre em contínua evolução, tanto profissionalmente, como em termos pessoais.

Ao professor e orientador Jorge Abrahão, que teve a boa vontade e disposição de ler o texto, fazendo os aconselhamentos adequados para que algumas falhas não fossem cometidas e os excessos fossem retirados.

Ao professor Alexandre Ywata, que ao ministrar com grande competência o curso de *Econometria e Séries Temporais* no ISC, me despertou para o interesse com relação à matéria.

Ao Marcelo de Sousa Eira e ao Charles Mathusalém Soares Evangelista, dirigentes, respectivamente, da SEMAG e de sua 2<sup>a</sup> Diretoria, os quais apoiaram a realização, tanto do curso de séries temporais, como o de Especialização em Orçamento Público.

A todos os colegas da SEMAG, que fazem da unidade um lugar ímpar, dos mais agradáveis em que já tive a oportunidade de trabalhar.

À minha namorada, Shara, com quem tenho compartilhado momentos maravilhosos, repletos de felicidade.

À minha filha, Mariana, sem cuja existência minha vida seria infinitamente mais vazia e desprovida de significado.



# SUMÁRIO

	Página
Agradecimentos	ii
Lista de Quadros . . . . .	vii
Lista de Figuras . . . . .	viii
<b>Introdução</b>	<b>1</b>
Metodologia . . . . .	3
<b>1 Revisão Bibliográfica</b>	<b>5</b>
<b>2 A Previsão no Processo Orçamentário</b>	<b>7</b>
2.1 Métodos Qualitativos e Quantitativos de Previsão . . . . .	8
2.2 O Processo de Previsão do Poder Executivo Federal . . . . .	10
2.3 O Modelo da Secretaria da Receita Federal . . . . .	12
2.3.1 A Acurácia das Previsões da SRF . . . . .	14
<b>I Análise da Série Reduzida Deflacionada</b>	<b>17</b>
<b>3 Suavização Exponencial</b>	<b>18</b>
3.1 Índices de Deflacionamento . . . . .	19
3.2 Métodos Baseados na Média . . . . .	24
3.2.1 Média Simples . . . . .	25
3.2.2 Média Móvel . . . . .	26
3.3 Métodos de Suavização Exponencial . . . . .	27
3.3.1 Modelos sem tendência e sem sazonalidade (N-N) . . . . .	30
3.3.2 Modelos com Tendência Aditiva e Sazonalidade Multiplicativa (A-M) . . . . .	30
3.3.3 Modelos com Tendência Aditiva e Sazonalidade Aditiva (A-A) . . . . .	33

3.3.4	Modelos com Tendência Multiplicativa e Sazonalidade Multiplicativa (M-M) . . . . .	35
3.3.5	Modelos com Tendência Amortecida Aditiva e Sazonalidade Aditiva (DA-A) . . . . .	35
3.4	Conclusão . . . . .	36
<b>4</b>	<b>Metodologia Box&amp;Jenkins</b>	<b>41</b>
4.1	Análise da Série . . . . .	43
4.2	Conclusão . . . . .	49
 <b>II Análise da Série Reduzida não Deflacionada</b>		<b>55</b>
<b>5</b>	<b>Suavização Exponencial</b>	<b>56</b>
5.1	Modelo Log - Tendência Aditiva e sem Sazonalidade (A-N) . . . . .	57
5.2	Modelo Log - Tendência Aditiva e Sazonalidade Multiplicativa (A-M) . . . . .	58
5.3	Modelo Log - Tendência Aditiva e Sazonalidade Aditiva (A-A) . . . . .	59
5.4	Modelo Log - Tendência Multiplicativa e Sazonalidade Multiplicativa (M-M)	60
5.5	Modelo /s Log - Tendência Aditiva e Sazonalidade Multiplicativa (A-M) . . . . .	61
5.6	Modelo s/ Log - Tendência Multiplicativa e Sazonalidade Multiplicativa (M-M)	62
5.7	Conclusão . . . . .	63
<b>6</b>	<b>Metodologia Box&amp;Jenkins</b>	<b>65</b>
6.1	Conclusão . . . . .	70
<b>7</b>	<b>Conclusão</b>	<b>73</b>
	<b>Referências Bibliográficas</b> . . . . .	<b>75</b>

# Lista de Quadros

	<b>Página</b>
2.1 Valores Previstos e Arrecadados de Imposto de Renda - 2004/2006 . . . . .	16
3.1 Inflação (IGP-DI) apurada pela FGV e Prevista pela SPE . . . . .	22
3.2 Suavização Exponencial - IGP-DI/SPE . . . . .	39
3.3 Suavização Exponencial - IGP-DI/FGV . . . . .	40
4.1 Parâmetros das Equações de Regressão . . . . .	47
4.2 Melhores Modelos - AIC/BIC . . . . .	48
4.3 Metodologia de Box-Jenkins - Série Inflacionada com Previsões da SPE . . .	53
4.4 Metodologia de Box-Jenkins - Série Inflacionada com IGP-DI Mensurado pela FGV . . . . .	54
5.1 Suavização Exponencial s/ Deflacionamento - Valores Previstos e Arrecada- dos 2006 . . . . .	64
6.1 Parâmetros das Equações de Regressão . . . . .	66
6.2 Melhores Modelos - AIC/BIC . . . . .	68
6.3 Previsões para Arrecadação do Imposto de Renda Sem Deflacionamento - Box&Jenkins . . . . .	72



# Lista de Figuras

	<b>Página</b>
2.1	Histograma dos Erros de Arrecadação da SRF . . . . . 15
3.1	Evolução dos Índices de Inflação . . . . . 20
3.2	Evolução do AIC - IGP-DI/IPCA Ponderados . . . . . 21
3.3	IR - Evolução da Arrecadação - Série Deflacionada . . . . . 24
3.4	IR - Previsão pela Média Simples . . . . . 26
3.5	IR - Média Móvel . . . . . 28
3.6	IR - Média Móvel - FAC Erros . . . . . 28
3.7	IR - Suavização Exponencial Simples . . . . . 31
3.8	IR - Alisamento Exponencial A-M . . . . . 32
3.9	IR - Suavização Exponencial Modelo A-A . . . . . 34
3.10	IR - Suavização Exponencial Modelo M-M . . . . . 36
3.11	Suavização Exponencial Modelo DA-A . . . . . 37
4.1	Série do Imposto de Renda Normal e em Log . . . . . 45
4.2	IR - Evolução da Arrecadação Média - Série Deflacionada . . . . . 45
4.3	IR - FAC e FACP - Série Deflacionada . . . . . 46
4.4	IR - Correlogramas das FAC e FACP das Séries sem Tendência . . . . . 48
4.5	Modelo SARIMA (0,0,0)(1,0,0) c/ Tendência Linear . . . . . 50
4.6	Modelo SARIMA (0,0,0)(1,0,1) c/ Tendência Linear . . . . . 50
4.7	Modelo SARIMA (0,0,0)(1,0,0) c/ Tendência Quadrática . . . . . 51
4.8	Modelo SARIMA (0,0,0)(1,0,1) c/ Tendência Quadrática . . . . . 51
5.1	IR - Evolução da Arrecadação não Deflacionada . . . . . 57
5.2	Suavização Exponencial Modelo A-N Log . . . . . 58
5.3	Suavização Exponencial Modelo A-M Log . . . . . 59
5.4	Suavização Exponencial Modelo A-A Log . . . . . 60

5.5	Suavização Exponencial Modelo M-M Log . . . . .	61
5.6	Suavização Exponencial Modelo A-M Sem Log . . . . .	62
5.7	Suavização Exponencial Modelo M-M Sem Log . . . . .	63
6.1	IR - Evolução da Série não Deflacionada . . . . .	66
6.2	IR - FAC e FACP da Série não Deflacionada . . . . .	67
6.3	IR - Série não Deflacionada sem Tendência . . . . .	68
6.4	Box&Jenkins - SARIMA(0,0,0)(1,0,1) . . . . .	69
6.5	Box&Jenkins - SARIMA(0,0,0)(1,0,0) . . . . .	69
6.6	Box&Jenkins - SARIMA(0,0,0)(1,0,2) . . . . .	70

# Introdução

A Secretaria de Macroavaliação Governamental-SEMAG do Tribunal de Contas da União possui, dentre suas atribuições regimentais, a de avaliar a previsão de arrecadação de receitas realizada anualmente pelo Poder Executivo e inserida nos anexos à Lei Orçamentária Anual. Atualmente esse trabalho é feito de forma qualitativa, apenas verificando se os parâmetros adotados pela Secretaria de Política Econômica do Ministério da Fazenda, e repassados aos órgãos arrecadadores para que cada um faça suas previsões, estão coerentes com as expectativas do Mercado e de instituições de pesquisa (IPEA, Banco Central).

Considerando, porém, a necessidade do aprimoramento desse trabalho, aventou-se a possibilidade de se iniciar a utilização de métodos estatísticos e econométricos nas análises das previsões realizadas, buscando uma fundamentação de forma mais técnica e confiável das conclusões obtidas. Poderia ser avaliado, dessa forma, se os métodos utilizados pela Secretaria da Receita Federal, ou pelo INSS, seriam os mais adequados para realizar as previsões anuais de arrecadação de receita tributária. Ao mesmo tempo, poderia ser verificado até que ponto as previsões realizadas têm conseguido espelhar a arrecadação observada.

A primeira etapa dessa nova estratégia de trabalho envolveu a contratação de um pesquisador do IPEA, professor Alexandre Ywata, para ministrar um curso de *Econometria e Séries Temporais* no ISC, voltado para um grupo de servidores interessados em modelagem estatística. A segunda etapa consistiu na alocação de um servidor para começar a estudar o comportamento das séries de alguns tributos (Imposto de Renda, IPI e COFINS) de forma a realizar alguns trabalhos seminais de previsão no âmbito da própria Secretaria. Contudo, dada a necessidade de alocação do servidor para atividades de auditoria, a continuidade do trabalho ficou prejudicada, sendo paralisado.

Posteriormente, com o surgimento do Curso de Pós-graduação em Orçamento Público ministrado pelo ISC/CEFOP, entendeu-se que seria o momento adequado para o reatamento e finalização dos trabalhos iniciados. Finalmente, deve-se mencionar que esse processo está inserido na expectativa de criação de uma massa crítica no âmbito da Secre-

taria, de servidores com conhecimento avançado na área de métodos quantitativos. Esses servidores seriam alocados para trabalhos futuros, tanto de auditoria, como de auxílio à elaboração do Relatório e Parecer Prévio das Contas do Governo da República, buscando enriquecer tecnicamente essas atividades. Além disso, há um consenso de que as atividades do Tribunal estão a cada dia se tornando mais complexas, saindo da estrita análise legal e enveredando por questões de eficiência, eficácia e efetividade. Esses novos desafios exigem servidores mais qualificados e instrumentos mais sofisticados, dentre eles a análise estatística.

A Constituição Federal estabelece, em seu artigo 165, parágrafo 8º, que a Lei Orçamentária Anual conterà a previsão de receita a ser arrecadada, bem como a fixação da despesa a ser executada. Considerando que um dos princípios orçamentários básicos é o do equilíbrio, ou seja, a receita deve ser igual à despesa, torna-se bastante importante que se faça uma previsão a mais acurada possível do montante a ser arrecadado, de molde a não se fixarem despesas em montante superior às receitas.

Essa questão é particularmente importante, uma vez que há um eterno conflito entre os setores governamentais responsáveis pelos gastos e aqueles responsáveis pela estabilização fiscal. Os que gastam têm todo o interesse em maximizar os dispêndios governamentais, seja para aumentar a quantidade de bens adquiridos/nível de serviços que prestam à população, seja para aumentar o próprio nível de poder ou influência (maximização do *bureau*). Aqui no Brasil esse tipo de comportamento tem sido visto nos últimos anos quando da apreciação, pelo Congresso Nacional, da Lei Orçamentária Anual enviada pelo Poder Executivo. Frequentemente o Poder Legislativo tem alterado, para maior, a previsão de arrecadação, visando dessa forma ao aumento das possibilidades de gasto, acomodando as inúmeras demandas dos grupos de pressão existentes.

Já os responsáveis pela estabilização fiscal têm consciência de que as demandas são infinitas e devem necessariamente ser limitadas à situação real da arrecadação, sob pena de haver um descontrole da dívida pública (caso a opção para fazer face ao excesso de despesa seja o endividamento) ou da inflação (caso a opção seja pela emissão de moeda).

A solução encontrada no Brasil nos últimos anos para esse embate tem sido o contingenciamento orçamentário. Ao longo de cada ano o Poder Executivo baixa decretos de contingenciamento, procurando adequar os dispêndios previstos às possibilidades econômicas existentes. Esse tipo de política, no entanto, não é eficiente, gerando frequentemente discontinuidades em ações governamentais, que não raro trazem prejuízos no médio e longo prazos. Exemplos que podem ser citados são a interrupção de obras para cujo reinício se faz necessária nova implantação de infra-estrutura, gerando gastos dobrados ou triplicados

nesse item, ou então a interrupção de ações de controle fitossanitário, possibilitando a ocorrência de patologias que acabam por prejudicar o consumidor ou as exportações nacionais.

É nesse contexto que entra em cena a necessidade de realizar da forma mais fidedigna possível a previsão das receitas a serem arrecadadas. A utilização de procedimentos robustos de previsão de arrecadação, aceitos consensualmente pela comunidade técnica, além de potencialmente inibir a tentação de superestimativa de receita por parte de determinados setores, pode ainda assegurar um melhor fluxo de gastos por parte do governo, evitando o contingenciamento e tornando a gestão dos programas mais racional, minimizando as possíveis discontinuidades derivadas de interrupções.

## Metodologia

A metodologia de trabalho a ser desenvolvida para a consecução do presente trabalho é relativamente simples. O ponto primordial envolve a aquisição de conhecimento na área de séries temporais, o que se iniciou com o já citado curso ministrado pelo Professor Ywata, tendo sido complementado por estudos posteriores.

A seguir, pretende-se utilizar esses procedimentos econométricos para a modelagem do comportamento da série bruta (sem quaisquer restituições ou acertos) da arrecadação do **Imposto sobre a Renda e Proventos de Qualquer Natureza**<sup>1</sup> no período de janeiro de 1995 a junho de 2005. Os valores mensais dessa arrecadação são obtidos mediante consulta à listagem L&88, gerada mensalmente pelo SERPRO, com informações desde janeiro de 1985 para os tributos administrados pela Secretaria da Receita Federal. A não utilização de dados anteriores a 1995 se deve à elevada taxa de inflação que permeava o período, o que poderia causar distorções nas análises. Os valores de arrecadação real entre janeiro e dezembro de 2006 serão utilizados para a verificação da acurácia das previsões realizadas.

O trabalho será dividido em duas partes. Na primeira, os valores da série serão deflacionados para uma mesma data, agosto de 2004, estudando-se a sua evolução, com base na suavização exponencial e metodologia Box&Jenkins, com avaliação de estacionariedade e de ocorrência de quebras estruturais, tanto com a utilização de métodos gráficos como numéricos. A escolha do índice é explicada na Seção 3.1 do trabalho, onde foi feita uma

---

<sup>1</sup>Note-se que essa série pode ser decomposta em três grandes rubricas: Imposto de Renda Pessoa Física-IRPF, Imposto de Renda Pessoa Jurídica-IRPJ e Imposto de Renda Retido na Fonte-IRRF. Cada uma dessas rubricas pode ser decomposta, ainda, em numerosos subcomponentes: IRPF-Declaração de Ajuste Anual, IRPF-Carnê Leão etc.

rápida análise na qual se procurou identificar qual dos índices de inflação seria o mais apropriado para o deflacionamento da série, mediante a minimização de um critério de informação, Akaike Information Criteria-AIC. Na segunda parte do trabalho são realizadas as mesmas análises, mas com a série com valores não deflacionados.

Existe uma grande quantidade de sistemas computacionais voltados para cálculos estatísticos: *SAS*, *E-views*, *R*, *Stata*, *SACS etc.*, devendo cada pesquisador escolher aquele que melhor lhe convier para trabalhar. Critérios ligados a preço, portabilidade, capacidade computacional, facilidade de uso, dentre outros, são comumente utilizados para a escolha adequada. O programa adotado no presente trabalho foi o  $R^2$  uma vez que é um *software* livre<sup>3</sup>, gratuito, é de fácil instalação e possui todas as facilidades instaladas necessárias para a execução dos procedimentos de análise das séries. Como existe uma grande comunidade de pesquisadores em todo o mundo utilizando o sistema, há um numeroso conjunto de manuais e textos técnicos disponíveis para *download* na *Internet*, o que facilita sobremaneira a sua utilização.

Como já foi dito anteriormente, além de realizar as análises das séries de arrecadação tributária, outro dos objetivos do presente trabalho é sedimentar conhecimentos e aprimorar a utilização de técnicas estatísticas para uso posterior nos demais trabalhos da SEMAG. Procurar-se-á, dessa forma, explicar e exemplificar muitos dos resultados teóricos obtidos na literatura. Embora isso possa vir a tornar o texto um pouco mais pesado do que o desejável, a necessidade de gerar modelos, fazer sua simulação computacional, analisar e descrever os resultados fará com que se adquira um domínio maior sobre o assunto.

---

<sup>2</sup>Versão 2.4.1, liberada para uso em fevereiro de 2007.

<sup>3</sup>Disponível para *download* em [www.cran.org](http://www.cran.org).

# Capítulo 1

## Revisão Bibliográfica

A literatura sobre séries temporais é vasta e abrangente, se estendendo desde livros com teor profundamente teórico-matemático (e.g. Hamilton, 1994), até livros mais leves, voltados para a aplicação da metodologia Box & Jenkins sem maiores considerações teóricas (e.g. Bowerman & O'Connell, 1987). A escolha de qual caminho trilhar foi difícil, mas optou-se por um meio termo; nem o rigor matemático que arriscaria a transformar o trabalho uma tese de doutorado (sem que o autor tivesse as qualificações teóricas condizentes com o desafio), nem o simplismo que, além de pouco agregar em termos de conhecimento, ainda poderia resultar em uma compreensão distorcida dos modelos utilizados, bem como dos resultados obtidos.

O embasamento teórico para a realização do trabalho se lastreou, então, fundamentalmente nas notas de aula do professor Alexandre Ywata, no curso de Estatística e Econometria ministrado, bem como no livros *Forecasting Methods and Applications*, de Spyro Makridakis, Steven C. Wheelwright e Rob J. Hyndman e *Applied Econometric Time Series* de Walter Enders.

Uma outra valiosa fonte de informações, explicações e exercícios bastante práticos foi o *Study Book* da cadeira *Statistics for Climate Research*, elaborado pelo professor Peter Dunn, da *University of Southern Queensland*. Os numerosos exemplos e as explicações simples e diretas foram de grande utilidade na compreensão do assunto.

Com relação propriamente à utilização de séries temporais para a previsão de receitas tributárias federais, há dois trabalhos que sobressaem, ambos ganhadores do prêmio de monografias da STN.

O primeiro trabalho é de Bruno Stephan Veras de Melo, servidor da Secretaria da Receita Federal, que fez estudos do comportamento da arrecadação do Imposto de Renda.

Em seu trabalho foram utilizadas as metodologias de alisamento exponencial e SARIMA, baseada na metodologia desenvolvida por Box & Jenkins. Além disso, foi explicada de forma bastante didática a metodologia desenvolvida pela Secretaria da Receita Federal (Método dos Indicadores) para a realização de suas previsões. O autor realizou trabalhos de previsão utilizando a série do Imposto de Renda, comparando seus resultados com aqueles obtidos pela SRF. Em média os resultados obtidos pelo autor se aproximaram mais da realidade do que os obtidos pelo órgão arrecadatório.

O segundo trabalho foi apresentado um ano após o primeiro, em 2002, por Marcelo Siqueira, também servidor da Secretaria da Receita Federal, tendo sido agraciado com o primeiro lugar do Prêmio da STN no tema Tributação, Orçamento e Sistemas de Informação sobre Administração Financeira Pública. Pode-se dizer que é uma evolução em relação ao primeiro, uma vez que além de modelar a arrecadação do Imposto de Renda, o autor estudou também o comportamento de outros tributos federais: Imposto de Importação, Imposto sobre Produtos Industrializados e CPMF. Em suas conclusões o autor também se mostrou bastante otimista com relação à utilização de séries temporais para a modelagem e previsão da arrecadação de tributos.

Além destes textos, será utilizado como fundamentação teórica o livro *An Introduction to State and Local Public Finance* de autoria dos professores Thomas Garrett e John Leatherman da *West Virginia University*. Em seu capítulo IV, (*Revenue Forecasting*), faz-se toda uma apresentação teórica da aplicação de séries temporais à previsão de receita, incluindo-se uma excelente revisão bibliográfica. A expectativa é que esse texto, somado aos trabalhos do Bruno e de Siqueira, bem como aos livros citados, formem o arcabouço teórico do trabalho.



## Capítulo 2

# A Previsão no Processo Orçamentário

Este capítulo tem por objetivo apresentar as principais justificativas, legais e teóricas, para a realização do trabalho de previsão de receitas no âmbito do processo de elaboração da Lei Orçamentária Anual (LOA).

Primeiramente, a questão legal. O artigo 165 da Constituição Federal, em seu §8º, estabelece que a Lei Orçamentária Anual não deverá conter dispositivo estranho à *previsão de receita* e fixação de despesa. Assim, a determinação para que seja feita a previsão das receitas é um *mandamus* constitucional do qual o elaborador do orçamento não pode se esquivar.

Adicionalmente, a Lei de Responsabilidade Fiscal, Lei Complementar nº 101/2000, estabelece no seu artigo 11 que a instituição, *previsão* e efetiva arrecadação de *todos os tributos* da competência constitucional de cada ente da Federação são requisitos essenciais para a responsabilidade da gestão fiscal. O seu parágrafo único veda que sejam feitas transferências voluntárias para o ente que não observar esse *mandamus* com relação aos impostos.

O artigo 12 do mesmo ato legal determina, ainda, que as previsões de receita deverão observar as normas técnicas e legais, considerando no seu cálculo os efeitos das alterações na legislação, da variação do índice de preços, do crescimento econômico ou de qualquer outro fator relevante. As previsões deverão ser acompanhadas de demonstrativo de sua evolução nos últimos três anos, da projeção para os dois seguintes àquele a que se referirem, bem como da metodologia de cálculo e premissas utilizadas.

Com relação a este artigo, vale a pena observar que *não existem* normas técnicas que determinem de forma inequívoca como deve ser feito um trabalho de previsão (Garrett e Leatherman, 2000, página 1). Existem, na verdade, diversos métodos que podem ser

utilizados, qualitativos e quantitativos, conforme pode ser visto em Makridakis et al. (1998). Outrossim, ao estabelecer que no cálculo das previsões devem ser levados em conta "os efeitos das alterações na legislação, da variação do índice de preços, do crescimento econômico ou de qualquer outro fator relevante", o legislador está limitando demasiadamente a margem de discricionariedade da equipe encarregada das previsões. A se levar ao pé da letra o disposto na Lei, a utilização de séries temporais univariadas para realizar as previsões estaria descartada, uma vez que nesta metodologia são levados em consideração apenas os próprios valores pretéritos da série para o cálculo dos seguintes.

Ainda que não estivesse positivada na norma, a previsão de receita deveria ser realizada a cada ciclo orçamentário, uma vez que compreender o processo de previsão é uma etapa essencial no planejamento orçamentário e nos processos de gestão conforme exposto em Kyobe e Danninger (2005). Considerando que um dos princípios orçamentários é justamente o do equilíbrio,  $\text{Receitas} \equiv \text{Despesas}$ , há que se fazerem previsões razoavelmente acuradas para as receitas, de forma a se equilibrarem com as despesas fixadas, sob pena de haver excesso de gastos (com reflexos negativos na política fiscal, uma vez que os recursos que estão faltando deverão ser obtidos de alguma forma, mediante a obtenção de empréstimos ou emissão monetária) ou excesso de receita (embora seja uma possibilidade teórica, há muito que isso não é visto no Brasil).

Deve-se ressaltar que quando se fala em previsão de receita não se está referindo apenas às receitas tributárias, embora essas de fato sejam as mais importantes. O poder público dispõe de uma série de outras fontes de recursos (e.g. cobrança de taxas pela utilização de serviços, rendas provenientes da exploração de atividades econômicas e recebimento de dividendos) que também devem ser objeto de previsão.

## 2.1 Métodos Qualitativos e Quantitativos de Previsão

A metodologia de previsão de receitas a ser utilizada pode ser de natureza qualitativa ou quantitativa. Nada obsta, no entanto, que as duas sejam utilizadas em conjunto. Após a realização de uma análise numérica, os resultados podem ser apresentados a um grupo de *experts* para validação.

Os processos de natureza qualitativa usualmente se baseiam na opinião de *experts* no setor a ser considerado. Dada a sua experiência de vários anos trabalhando no assunto, testemunhando as mais diversas situações que redundaram em diferentes efeitos sobre a

evolução da variável cujo comportamento se deseja prever, estes profissionais acabaram por adquirir uma alentada capacidade de compreensão dos fatores que afetam a evolução desta variável. Embora não esteja lastreada em estudos numéricos, esta experiência lhes permite prever o que irá acontecer com uma certa dose de acerto, Garrett e Leatherman (2000).

Previsões baseadas em métodos qualitativos são usualmente utilizadas quando não existem estatísticas confiáveis relativas aos períodos passados e que possam ser utilizadas em métodos quantitativos, quando são esperadas modificações significativas no ambiente em que se dá o processo (e.g. uma grande elevação na alíquota de um determinado tributo, um cataclismo que afetou severamente uma determinada região ou setor econômico) ou quando não há equipes competentes para a realização de análises econométricas.

Apesar do contínuo avanço das técnicas estatísticas, os métodos qualitativos de previsão ainda são bastante utilizados. Em estudo publicado pelo FMI, envolvendo uma amostra de trinta e quatro países (da África Subsaariana, da Ásia, América Latina e Oriente Médio), aproximadamente 85% utilizam técnicas de previsão de receitas de caráter subjetivo e de simples extrapolação (Kyobe e Danninger, 2005, página 14).

É claro que para que haja a criação de um grupo capaz de realizar previsão de forma qualitativa é necessário que haja uma certa estabilidade na estrutura administrativa. Se houver uma grande rotatividade nos funcionários encarregados da administração de um determinado tributo, por exemplo, dificilmente será criada uma memória institucional que permita gerar o conhecimento necessário para a realização desse tipo de trabalho.

Embora extensivamente utilizados, os métodos qualitativos, uma vez que envolvem julgamentos subjetivos, podem dar margem a problemas. Guajardo e Miranda (2000), citados por (Garrett e Leatherman, 2000, página 3), apresentam oito das principais fraquezas que uma análise qualitativa pode apresentar<sup>1</sup>:

- eventos âncora - eventos recentes de grande impacto podem influenciar a percepção sobre eventos futuros. Por exemplo, ao ser escolhido para sediar um grande evento esportivo (uma copa do mundo, olimpíadas ou jogos panamericanos) toda a economia de um país, ou de uma região, pode ser afetada de forma positiva durante o período de preparação e realização deste evento. Este efeito, no entanto, tem data certa para acabar, o que deve ser levado em conta quando da realização de previsões;
- disponibilidade de informações - excesso de peso dado ao uso de informações prontamente disponíveis;

---

<sup>1</sup>Os itens a seguir foram livremente traduzidos, com as adaptações que se consideraram adequadas.

- correlação falsa - os encarregados da realização das previsões podem incorporar informações sobre fatores que erroneamente são considerados como influenciadores da arrecadação de receitas;
- inconsistências nos métodos e nos julgamentos - os encarregados da realização das previsões podem utilizar diferentes estratégias de previsão ao longo do tempo, tornando-as menos confiáveis;
- percepções seletivas - ocorrem quando os encarregados das previsões ignoram informações importantes que conflitam com a sua visão sobre relações causais;
- *wishful thinking*<sup>2</sup> - ao darem um peso maior àquilo que os encarregados de previsão ou servidores públicos desejam, os resultados das previsões podem ser irrealistamente otimistas;
- comportamento de grupo - ocorre quando a dinâmica da formação de consenso tende a levar indivíduos a reforçarem os pontos de vista uns dos outros, em detrimento de seus julgamentos independentes;
- pressões políticas - ocorrem quando os encarregados das previsões ajustam as estimativas de forma a ir ao encontro das necessidades de um orçamento equilibrado ou das limitações orçamentárias.

Os métodos quantitativos de previsão de receitas se baseiam na análise numérica de dados pré-existentes, buscando estabelecer um modelo (ou modelos) que os represente(m) da forma mais fiel possível, utilizando-os em seguida para a realização das previsões desejadas. Entre as vantagens que esse tipo de metodologia apresenta podem ser citadas a transparência dos parâmetros utilizados nos cálculos<sup>3</sup> e a possibilidade de estabelecimento de um grau de incerteza associada às estimativas realizadas (Garrett e Leatherman, 2000, página 3).

## 2.2 O Processo de Previsão do Poder Executivo Federal

Tendo em vista as diretrizes constitucionais e legais citadas, o Poder Executivo Federal, por intermédio de seus órgãos da administração direta e indireta, realiza o trabalho

---

<sup>2</sup>Mantido no inglês original, significa tomar os desejos por realidades e tomar decisões, ou seguir raciocínios, baseados nesses desejos em vez de nos fatos ou na racionalidade (Wikipedia, verbete *Wishful thinking*).

<sup>3</sup>Que necessariamente deverão constar nas equações utilizadas.

de previsão das receitas tributárias e não tributárias a serem arrecadadas no final do período (julho de cada exercício até o final do ano) e para os doze meses do ano seguinte.

Essas previsões, no geral, são calculadas tomando por base a evolução de determinados parâmetros que, por pressuposto, afetam a evolução da arrecadação das receitas a serem estimadas. De forma a uniformizar as projeções desses parâmetros, foi decidido que a Secretaria de Política Econômica do Ministério da Fazenda - SPE/MF iria centralizar os cálculos e a coleta dos indicadores a serem utilizados, repassando-os aos demais órgãos a fim de que estes os utilizem na previsão das receitas de cuja arrecadação são responsáveis.

Este tipo de procedimento evita que os órgãos utilizem diferentes valores para a projeção do crescimento real do PIB, inflação, massa salarial etc., o que certamente causaria enorme confusão. Finalizado o trabalho de previsão de cada órgão, os valores calculados são enviados à Secretaria de Orçamento Federal-SOF que os consolida e insere no Projeto de Lei Orçamentária Anual a ser enviado ao Congresso Nacional para deliberação.

Na lista a seguir são elencados os principais parâmetros macroeconômicos consolidados/calculados pela SPE e enviados aos demais órgãos<sup>4</sup>:

- PIB: as projeções para evolução do PIB levam em conta a evolução dessazonalizada nos últimos trimestres do ano, sendo incorporados os valores estimados para as taxas de juros, a produção agrícola e a produção industrial;
- Indicadores industriais: a produção da Indústria de Transformação e de Bebidas e as vendas da Indústria de Fumo e de Veículos são utilizados como parâmetros para as estimativas de receitas orçamentárias. De acordo com o órgão, as projeções iniciais são ajustadas levando-se em conta: a trajetória de longo prazo da série estatística e a relação entre as expectativas de evolução da atividade até o final do período e sua trajetória de longo prazo;
- Importações: as projeções das importações, excluindo-se o item combustíveis, são estimadas tomando por base as expectativas de mercado para o valor anual das importações totais, divulgadas semanalmente pelo Banco Central. Subtraem-se dos valores anuais as estimativas de importação de petróleo fornecidas pela Petrobrás;
- Mercado de Trabalho: com relação a este item, são utilizados os seguintes parâmetros, constantes da Pesquisa Mensal de Emprego (PME) executada pelo IBGE: População Economicamente Ativa (PEA), População Ocupada (PO) com carteira assinada e o Rendimento Médio Nominal (RMN) dos trabalhadores com carteira assinada. Cada um desses parâmetros tem metodologia diferente de previsão;

---

<sup>4</sup>Informações obtidas nas Informações Complementares ao Projeto de Lei Orçamentária Anual.

- Indicadores de Preços: são elaboradas projeções para os seguintes índices de preços: IGP-DI (Índice Geral de Preços - Disponibilidade Interna, da FGV), IPCA (Índice de Preços ao Consumidor Amplo, do IBGE), INPC (Índice Nacional de Preços ao Consumidor, do IBGE) e IPA (Índice de Preços ao Atacado, da FGV) para a Indústria de Transformação, de Bebidas, Fumo e Veículos. As projeções destes parâmetros são revistas periodicamente, em função do comportamento da inflação, bem como das previsões atualizadas do Banco Central (atas das reuniões do Copom e relatórios trimestrais de inflação) e do mercado (expectativas mensais, anuais e para os doze próximos meses para os principais índices de preços, divulgadas semanalmente pelo Banco Central);
- Taxas de câmbio: projeções tomam como base as expectativas do mercado, ajustadas pelas informações referentes à demanda por divisas estrangeiras determinadas pelos fluxos de pagamentos financeiros dos setores público e privado;
- Taxas de juros: projeções tomam por base as expectativas de mercado para as decisões do Copom, levando em conta também a trajetória de juros reais de longo prazo consistente com a taxa projetada de crescimento do PIB;
- Aplicações financeiras: calculadas de acordo com a seguinte equação:  $(M4 - M1 - \text{Dep. Poupança})$ , onde M4 é o maior agregado monetário, M1 é a quantidade de moeda em poder do público somada aos depósitos à vista nos bancos comerciais. As projeções são realizadas individualmente para cada parâmetro da equação, considerando a Teoria Quantitativa da Moeda;
- Taxa de Juros de Longo Prazo/TJLP: projetada considerando o risco-país de longo prazo e as metas de inflação pro-rata 12 meses adiante;
- Taxa Referencial/TR: projeções realizadas com base na taxa over-selic, utilizando-se ainda os parâmetros da Resolução 3.354 do Conselho Monetário Nacional.

## 2.3 O Modelo da Secretaria da Receita Federal

Tomando por base os parâmetros enviados pela SPE, a Secretaria da Receita Federal do Brasil-RFB realiza então os cálculos da previsão de arrecadação dos tributos sob sua responsabilidade. A seguir será apresentada a metodologia desenvolvida pela Secretaria, conforme exposto nas notas metodológicas apresentadas anualmente pelo órgão e em Melo (2001).

O processo de previsão de arrecadação de receitas no âmbito da SRFB pode ser dividido em duas etapas. Na primeira, é analisada a situação de cada tributo, procurando verificar se a arrecadação base no período  $t$  (que será utilizada para a previsão de arrecadação no período  $t + 1$ ) poderá ser utilizada sem alterações, ou haverá necessidade de um ajuste por conta de fatores que ocorreram no período anterior, ou supõe-se que ocorrerão no seguinte. Uma vez detectada a existência de uma fator, ou fatores, que impactarão a nova base de cálculo, os mesmos serão incorporados à nova base de cálculo, que será utilizada nas previsões.

Há numerosos exemplos de fatores que podem ser citados. O mais comum é o ajuste relativo ao número de dias úteis, que varia de período para período, e que impacta a arrecadação de IPI, COFINS e PIS/PASEP (conforme pode ser visto em nota metodológica da SRF). Um outro exemplo, desta vez hipotético, é a eventual exclusão da base de cálculo do ano seguinte dos efeitos relativos ao elevado consumo de cervejas e refrigerantes (IPI-Bebidas) em um ano em que as temperaturas elevadas do verão se estenderam por um período superior ao normal.

Uma vez ajustada, a base de cálculo resultante é multiplicada pelos índices relativos a alterações de preço (representativo do efeito que a inflação projetada para o período seguinte terá sobre o fato econômico objeto do tributo), quantidade (representativo da variação real da quantidade desse fato econômico) e legislação (representativo da variação que alterações na legislação vão causar na arrecadação).

Sinteticamente,

$$A_{i,t} = A_{i,t-1}^a * (1 + \Delta P) * (1 + \Delta Q) * (1 + \Delta L) \quad (2.1)$$

onde,

- $A_{i,t}$  - arrecadação prevista no mês  $i$  do ano  $t$ ;
- $A_{i,t-1}^a$  - arrecadação efetiva no mês  $i$  do ano  $t - 1$ , *ajustada*. Como as previsões no geral são feitas nos meses de junho ou julho, só existem dados de arrecadação efetiva até estes meses. Para os meses seguintes (usualmente julho a dezembro) são utilizados os valores de julho a dezembro do ano anterior ;
- $(1 + \Delta P)$  - índice de preço (inflação no período);
- $(1 + \Delta Q)$  - índice de quantidade;
- $(1 + \Delta L)$  - índice de legislação.

Note-se que todo este trabalho exige da(s) equipe(s) que o realizam um profundo conhecimento empírico do comportamento da arrecadação dos tributos frente às alterações no panorama econômico do Brasil e do mundo. Conseguir quantificar em um parâmetro o efeito que um determinado índice (previsto) de inflação vai ter sobre a arrecadação de um tributo não é tarefa trivial. A experiência e a cultura institucional acumuladas durante anos de trabalho são essenciais para a acurácia das previsões.

### 2.3.1 A Acurácia das Previsões da SRF

No Quadro 2.1 são apresentados os valores referentes às previsões, bem como à arrecadação real do Imposto de Renda agregado ao longo do período 2004/2006<sup>5</sup>, mensalmente, por ano. Esses valores previstos fazem parte do Projeto de Lei Orçamentária Anual enviado para apreciação pelo Congresso Nacional.

Pode ser constatado nos três anos considerados que as previsões da Secretaria da Receita Federal foram sistematicamente inferiores às que efetivamente foram realizadas. Essa diferença foi particularmente elevada em 2005, quando atingiu 12,83% do total. Essa subestimativa sistemática da arrecadação está muito provavelmente ligada a um comportamento conservador por parte da autoridade fazendária. Ao prever a menor uma expectativa de arrecadação, a SRF permite que haja a criação de uma margem de segurança com relação à alocação de gastos orçamentários.

Os erros das previsões mensais apresentam valores diversos, se estendendo desde +0.09% em novembro de 2004 a -21.45% em agosto de 2005. Na Figura 2.1 é apresentado o histograma dos trinta e seis erros mensais incorridos no período observado. Note-se que em apenas dois meses a previsão da SRF superou a arrecadação; nos restantes trinta e quatro meses a previsão foi inferior à arrecadação, evidenciando o caráter conservador dos cálculos.

---

<sup>5</sup>Obtidos junto à listagem L&88 emitida pelo SERPRO mensalmente.



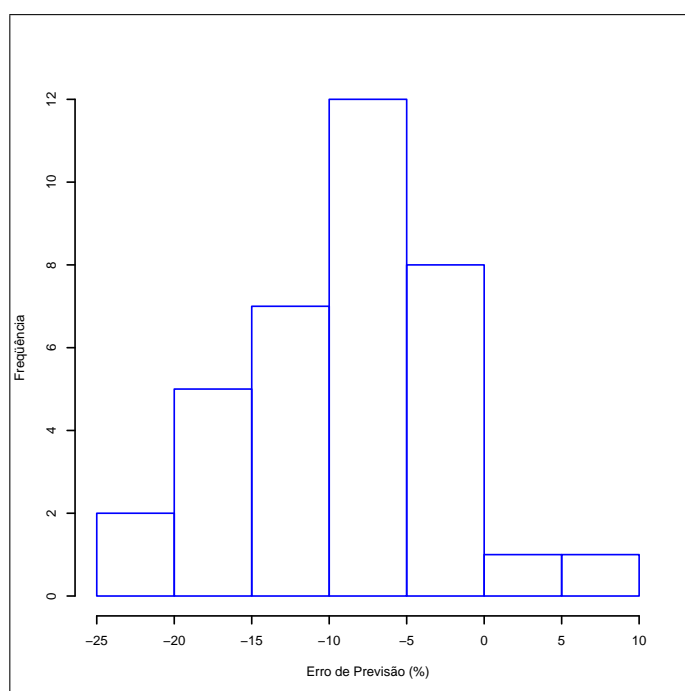


Figura 2.1: Histograma dos Erros de Arrecadação da SRF

Quadro 2.1: Valores Previstos e Arrecadados de Imposto de Renda - 2004/2006

Mês	2004			2005			2006		
	Arrec.*	Prev.*	Δ (%)	Arrec.**	Prev.**	Δ (%)	Arrec.***	Prev.***	Δ (%)
jan	10.105.940,02	8.983.185	-11,11	10.081,97	9.089	-9,85	10.643.316.250,33	9.507.397.313	-10,67
fev	7.464.527,59	7.105.618	-4,81	7.858,25	7.542	-4,02	9.259.073.435,62	8.606.013.260	-7,05
mar	7.889.709,18	7.458.519	-5,47	9.539,31	8.472	-11,19	10.189.952.436,27	9.592.790.814	-5,86
abr	9.198.601,70	8.952.123	-2,68	10.927,58	9.365	-14,30	12.588.188.574,93	11.667.203.521	-7,32
mai	6.817.773,31	7.185.383	5,39	8.501,77	6.956	-18,18	9.180.874.262,48	8.886.017.518	-3,21
jun	7.360.569,67	7.319.956	-0,55	12.477,37	11.083	-11,18	13.956.711.467,90	13.237.302.483	-5,15
jul	8.504.515,08	8.124.361	-4,47	9.250,10	8.593	-7,10	10.397.486.714,18	10.376.072.978	-0,21
ago	7.200.058,18	6.545.924	-9,09	8.990,94	7.062	-21,45	9.485.565.072,44	8.633.888.599	-8,98
set	7.855.788,11	7.031.913	-10,49	8.422,67	6.814	-19,10	10.312.573.908,60	8.492.821.645	-17,65
out	8.676.989,71	7.982.289	-8,01	10.108,37	7.961	-21,24	11.703.111.572,74	9.772.641.683	-16,50
nov	6.530.677,05	6.536.763	0,09	9.549,99	8.424	-11,79	9.492.285.018,92	8.840.966.029	-6,86
dez	12.013.341,08	9.839.264	-18,10	15.505,50	14.299	-7,78	16.606.948.120,64	17.433.981.751	4,98
Total	99.618.490,68	93.065.298	-6,58	121.213,81	105.660	-12,83	133.816.086.835,05	125.047.097.594	-6,55
EPMA		6,69			13,10			7,87	

Fonte: Informações complementares aos projetos de lei orçamentária dos anos 2004 a 2006.

\* (R\$ mil)

\*\* (R\$ milhões)

\*\*\* (R\$)

# Parte I

## Análise da Série Reduzida Deflacionada

## Capítulo 3

# Suavização Exponencial

Nesse capítulo são apresentados os métodos de previsão baseados na média e na suavização exponencial, aplicando-os à série de arrecadação do Imposto de Renda, cujos valores foram deflacionados pelo índice a ser escolhido na forma apresentada na Seção 3.1. Para cada um desses métodos, é feita uma breve introdução, apresentando-se a seguir as equações que são utilizadas para a modelagem da série. Feitas as análises, são apresentados os resultados para as previsões, os gráficos com a série modelada, os correlogramas das funções de autocorrelação e de autocorrelação parcial dos resíduos, bem como os resultados para o Teste Box-Pierce, (Garrett e Leatherman, 2000, página 318), dos resíduos. Após o cálculo dos valores previstos deflacionados, é feito o seu inflacionamento para os meses de 2006, utilizando-se o índice de inflação adotado.

Ao final, na conclusão do capítulo, são apresentados os Quadros 3.2 e 3.3, onde é feita uma comparação entre os resultados obtidos com os métodos de suavização exponencial considerados mais adequados, com base na análise dos resíduos, e os valores efetivamente recolhidos no período de janeiro a dezembro de 2006. No primeiro quadro os valores previstos foram inflacionados com base no índice previsto pela SPE, enquanto no segundo foi utilizado o índice adotado efetivo, calculado pela entidade responsável pela sua elaboração.

Ressalte-se que, com essa metodologia de análise, procurou-se simular da forma mais fidedigna possível a situação vivida pelos técnicos do Governo encarregados de realizar as previsões de arrecadação tributária para o ano seguinte. O projeto de lei orçamentária precisa estar pronto para ser enviado ao Congresso Nacional para apreciação e aprovação até agosto de cada ano. Portanto, as informações disponíveis para as análises são aquelas usualmente existentes até junho de cada exercício. Os órgãos arrecadadores precisam então prever, na verdade, dezoito meses de arrecadação; os seis últimos meses do ano e os doze meses do ano seguinte.

Sob esse aspecto metodológico, o presente trabalho se distancia daqueles realizados por Siqueira (2002) e Melo (2001). Em ambos, o período dos dados utilizados foi até dezembro, com as previsões sendo realizadas para os doze meses do ano seguinte.

A utilização da série deflacionada faz sentido em um ambiente institucional em que a previsão dos índices inflacionários seja realizada por um órgão distinto daquele que realiza a previsão de arrecadação da receita, o qual deve levar em consideração aqueles índices em seus cálculos. Assim, faz-se o deflacionamento dos valores da série para uma data base, constrói-se o modelo, fazem-se as previsões e, posteriormente, inflacionam-se os valores calculados para a data desejada, utilizando-se os índices oficiais.

Basicamente, essa é uma situação similar à que acontece com a RFB, e demais órgãos arrecadadores, quando da elaboração das previsões de arrecadação, uma vez que estes órgãos precisam se ater aos parâmetros de índices inflacionários e crescimento do PIB calculados pela Secretaria de Política Econômica-SPE.

O trabalho de análise da série deflacionada do Imposto de Renda é complementado no Capítulo 4, onde é feita uma rápida apresentação da metodologia de Box&Jenkins, utilizando-a em seguida para a realização das previsões relativas à arrecadação do tributo.

### 3.1 Índices de Deflacionamento

Uma vez definido que a série a ser estudada deve ser deflacionada, há que se escolher o índice a ser utilizado. Essa decisão é importante, considerando que a adoção de um ou outro índice ou, mesmo, uma combinação dentre eles, como inflacionador/deflacionador, gerará uma série temporal diferente, que poderá, ou não, gerar modelos mais capazes de moldar a série original e realizar previsões mais, ou menos, precisas. Os índices que mensuram a inflação comumente utilizados no Brasil exibem comportamentos distintos, de acordo com a metodologia de apuração e abrangência de captação de dados de cada um. A evolução do comportamento temporal dos principais índices de inflação (IGP, IGP-DI, IPCA e INPC) pode ser melhor visualizada na Figura 3.1, para o período jan/1995 a dez/2006.

A Secretaria da Receita Federal do Brasil, em seu trabalho de previsão de arrecadação, utiliza diferentes índices para compor o Efeito Preço a ser utilizado em sua metodologia. De acordo com o tributo a ser estudado, os índices a serem adotados podem ser o INPC, IPA ou o IPCA, ou, mesmo, uma combinação dentre eles. Assim, o cálculo da previsão da arrecadação do Imposto sobre Produtos Industrializados sobre a produção de automóveis, leva em consideração o índice de preços específico do setor. Com relação

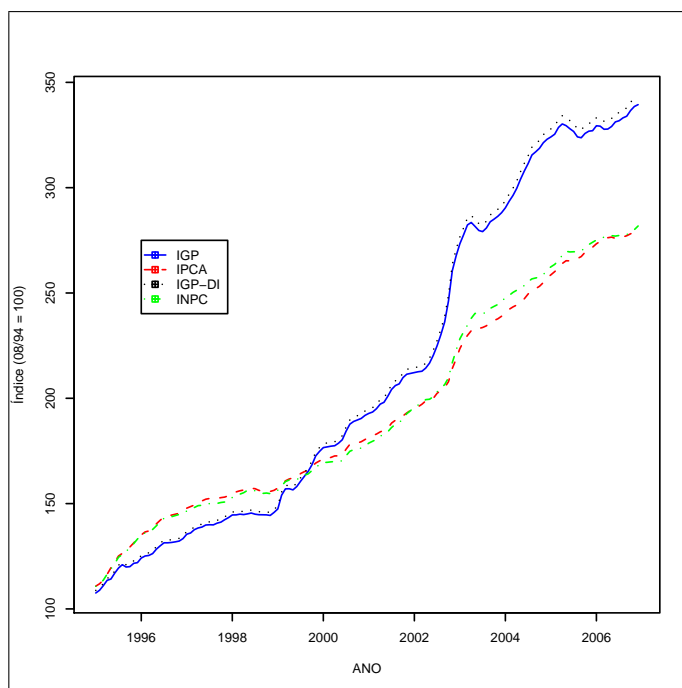


Figura 3.1: IR - Média Móvel

ao Imposto de Importação, as previsões relativas à variação da taxa média de câmbio são levadas em consideração. No que concerne às previsões relativas à arrecadação do Imposto de Renda, em suas duas modalidades, Imposto de Renda Pessoa Física-IRPF e Imposto de Renda Pessoa Jurídica-IRPJ, utiliza-se no geral um índice ponderado (55% IPCA e 45% IGP)<sup>1</sup>.

Já nos trabalhos apresentados por Melo (2001) e Siqueira (2002), os autores utilizaram o IGP-DI como índice para o deflacionamento das séries a serem estudadas, com resultados satisfatórios, segundo as conclusões apresentadas por ambos. O IGP-DI, Índice Geral de Preços - Disponibilidade Interna, é calculado mensalmente pela Fundação Getúlio Vargas do Rio de Janeiro. É um índice composto, refletindo a evolução de preços captada pelo Índice de Preços por Atacado (IPA), Índice de Preços ao Consumidor (IPC-FGV) e Índice Nacional de Preços da Construção Civil (INCC).

Note-se que dos gráficos apresentados na Figura 3.1, pode-se ver que a evolução do IGP e do IGP-DI é semelhante, diferenciando-se ambos da evolução do IPCA e do INPC. Essa diferença no comportamento dos índices é decorrente do fato de o IGP e IGP-DI serem

<sup>1</sup>Informações obtidas junto às notas metodológicas usualmente apresentadas no Projeto de Lei Orçamentária Anual, enviado pelo Poder Executivo ao Congresso Nacional para apreciação.

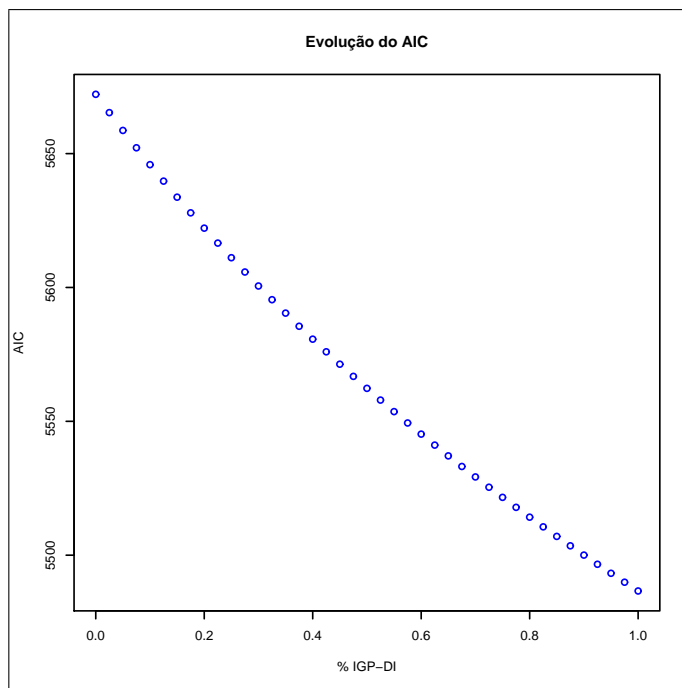


Figura 3.2: Evolução do AIC - IGP-DI/IPCA Ponderados

mais sensíveis ao câmbio, uma vez que captam os preços dos produtos de exportação e de importação com mais fidedignidade. Essa influência do câmbio nos índices fica bastante clara no gráfico. Após a desvalorização cambial do real no final de 1999, os dois índices gerais de preço começam a se descolar do IPCA e do INPC em uma trajetória de crescimento mais elevado. Essa trajetória sofre nova alteração, desta vez em sentido contrário, quando do início da apreciação do real frente às demais moedas, principalmente o dólar americano, a partir de 2005.

De forma a verificar qual o índice, ou ponderação de índices, seria o mais apropriado para moldar a série, iremos utilizar as propriedades do Critério de Informação Akaike (Akaike Information Criteria), ou AIC, (Garrett e Leatherman, 2000, página 360). Considerando que quanto melhor um modelo se adapta a uma série, menor o seu AIC, foi feito um estudo da evolução desse critério de informação, criando-se um índice ponderado a partir do IGP-DI e do IPCA, variando-se a participação de cada um entre 0% a 100%. Para cada uma das séries resultantes do deflacionamento ponderado, utilizou-se o comando `ets` do R para a geração de séries por suavização exponencial, calculando-se então o AIC respectivo. O resultado pode ser visto na Figura 3.2. O IGP-DI, puro, sem qualquer ponderação com o IPCA, se revelou o melhor deflacionador para série do Imposto de Renda, quando a mesma é modelada por suavização exponencial.

Quadro 3.1: Inflação (IGP-DI) apurada pela FGV e Prevista pela SPE

Mês	IGP-DI Apurado	Previsão (SPE)	$\Delta$ (%)
jul/05	-0,40	0,26	165,00
ago/05	-0,79	0,39	149,37
set/05	-0,13	0,44	438,46
out/05	0,63	0,43	31,75
nov/05	0,33	0,45	36,36
dez/05	0,07	0,49	600,00
jan/06	0,72	0,37	48,61
fev/06	-0,06	0,50	933,33
mar/06	-0,45	0,43	195,56
abr/06	0,02	0,53	2550,00
mai/06	0,38	0,68	78,95
jun/06	0,67	0,60	10,45
jul/06	0,17	0,53	211,76
ago/06	0,41	0,61	48,78
set/06	0,24	0,22	8,33
out/06	0,81	0,25	69,14
nov/06	0,57	0,38	33,33
dez/06	0,26	0,24	7,69
EPMA	-	-	468,07



A partir do resultado obtido, bem como dos trabalhos de Melo (2001) e Siqueira (2002), optou-se por utilizar no presente trabalho o IGP-DI, puro, como deflacionador da série sob análise. Ressalte-se que, para uma série de um tributo diferente (e.g. IPI ou Cofins), um outro índice poderia ser utilizado.

É importante entender que, uma vez que seja utilizado um processo de deflacionamento, modelagem e previsão, com posterior inflacionamento dos resultados até a data cuja arrecadação se deseja prever, as incertezas associadas à previsão se avolumam. A primeira delas concerne ao próprio índice a ser utilizado. Como estamos usando a série do Imposto de Renda total, estamos agregando arrecadações tributárias provenientes de diversos fenômenos econômicos, os quais reagem a variações inflacionárias de forma diferente, que poderiam ser captadas por diferentes índices. O IGP-DI adotado foi o melhor índice *possível*.

A segunda se refere ao próprio modelo adotado representar a série sob estudo. Nenhum modelo é perfeito, nem consegue captar a integralidade das características de uma série composta por valores reais, mensurados. Então, quando há a utilização de um modelo, há uma perda de parte da informação contida na série.

Por último, as previsões relativas aos índices de inflação futura. Quão acuradas são essas previsões? No Quadro 3.1 são apresentadas as previsões do IGP-DI para os período de jul/2005 a Dez/2006, realizadas pela SPE e utilizadas nos parâmetros macroeconômicos para elaboração do orçamento de 2006. Não se pode afirmar que as previsões tenham sido muito precisas. Para o mês de abril, por exemplo, houve uma divergência de 2550% entre a previsão elaborada pela SPE e o valor apurado pela FGV. Uma possível forma de tornar essa previsão mais acurada seria mediante a utilização, também, de métodos econométricos na previsão dos índices de inflação futuros, conforme realizado em Neves Cordeiro (2002).

Na Figura 3.3 são apresentados dois gráficos da série de arrecadação do Imposto de Renda, com os valores já convertidos em reais, deflacionados pelo IGP-DI para agosto de 1994. O primeiro gráfico apresenta a série desde janeiro de 1985 a junho de 2005, enquanto o segundo apresenta apenas o período com o qual se vai trabalhar, janeiro de 1995 a junho de 2005. Percebe-se de forma clara a ocorrência de uma alteração no comportamento da série a partir de meados de 1994, período em que houve a implementação do Plano Real.

Como na figura já ficou bastante claro o período em que ocorreu a alteração no comportamento da série, não há necessidade de realização de um teste estatístico para averiguação da ocorrência de quebra estrutural. Com base nesse resultado, optou-se por

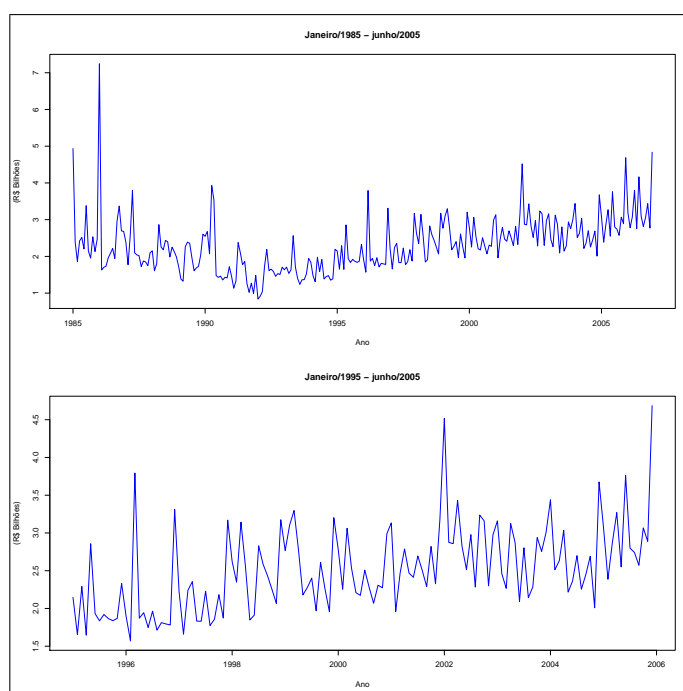


Figura 3.3: Evolução da Arrecadação do Imposto de Renda - Série Deflacionada

trabalhar somente com a segunda parte da série, relativa ao período de estabilidade monetária proporcionado pelo Plano Real.

Note-se que a quebra estrutural da série se dá exatamente no ano do plano de estabilização. A arrecadação vinha em uma trajetória descendente, em termos reais, até meados de 1994. Após o início do processo de estabilização monetária, a arrecadação real inicia uma trajetória crescente, que continua até o final do período. A perda de arrecadação tributária que vinha ocorrendo, provavelmente, tinha relação com as altas taxas de inflação prevalecentes. Como havia sempre um intervalo de tempo para o recolhimento dos tributos e nenhuma indexação é perfeita, estava ocorrendo uma perda para o Tesouro Nacional. É o chamado Efeito Tanzi, em homenagem ao economista Vito Tanzi.

## 3.2 Métodos Baseados na Média

São dois os principais métodos de previsão baseados na média: o de Média Simples e o de Média Móvel. Ambos têm como característica dar a mesma ponderação para todas as observações que serão utilizadas na modelagem da série original.

### 3.2.1 Média Simples

A utilização desse método pressupõe que a série sob estudo seja estacionária e que os seus valores se distribuem ao redor de uma constante (a média), sendo a variância em torno da média também constante Makridakis et al. (1998).

A estacionariedade de uma série pode ser *fraca* ou *forte*. Ela é considerada fortemente estacionária, ou estritamente estacionária, quando a distribuição do conjunto de variáveis aleatórias  $\{y_{t_1}, y_{t_2}, y_{t_3}, \dots, y_{t_s}\}$  é a mesma que a distribuição do conjunto  $\{y_{t_1+k}, y_{t_2+k}, y_{t_3+k}, \dots, y_{t_s+k}\}$  para todos as  $s$ -uplas  $\{y_{t_1}, y_{t_2}, y_{t_3}, \dots, y_{t_s}\}$  e todos os valores de  $k$ , onde  $k$  é inteiro. Assim, a distribuição do modelo é independente do tempo, sendo dependente apenas do intervalo entre cada ponto no tempo (Baillie, 2003).

Porém, não há como testar, a partir de uma realização empírica de uma série temporal, se ela é fortemente estacionária. Assim, normalmente é utilizado um outro conceito, o de estacionariedade fraca. Uma série é dita ser fracamente estacionária se apresenta as seguintes características:

- valor esperado constante, independente do tempo;
- variância constante;
- autocovariâncias constantes, dependentes apenas da diferença entre os períodos considerados.

Neste trabalho, se trabalhará sempre com o conceito de estacionariedade fraca. Assim, quando se disser que uma série é estacionária, implicitamente se estará dizendo que ela é *fracamente* estacionária.

O método da Média Simples modela a série considerando que todos os seus valores podem ser aproximados pela média dos seus valores. Assim,

$$\bar{y}_t = \frac{\sum_{i=1}^n y_t}{n} \quad (3.1)$$

onde  $n$  é o número de pontos observados da série e  $y_t$  representa o valor de cada um. Os valores previstos para  $t > n$ ,  $\hat{y}_t$ , são todos iguais a  $\bar{y}_t$ , multiplicados pelo índice de atualização monetária mensal utilizado (no presente caso o IGP-DI).

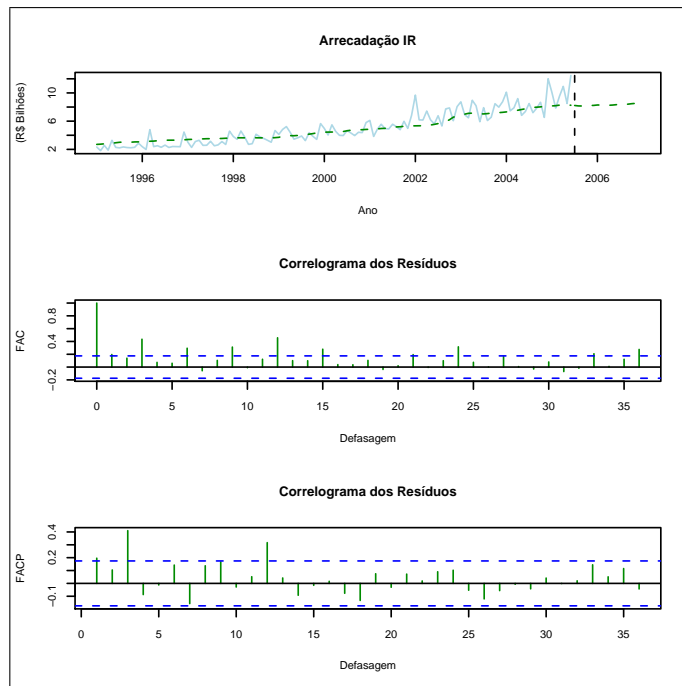


Figura 3.4: IR - Previsão pela Média Simples

No primeiro gráfico da Figura 3.4 são apresentadas as evoluções das séries original (linha azul contínua) e modelada pela média simples (linha verde serrilhada), com a evolução dos valores previstos sendo apresentada após a linha vertical. É visível que esse tipo de modelo não consegue simular um comportamento tão variável como o da arrecadação tributária. Os correlogramas das funções de autocorrelação e de autocorrelação parcial da série dos resíduos indicam que há correlação significativa entre eles, com caráter nitidamente sazonal, de período de doze meses. A Estatística Q do Teste Box-Pierce apresentou o valor de 99.05, extremamente elevado para uma distribuição de  $\chi^2$  com vinte graus de liberdade.

Com base nos resultados encontrados, verifica-se que este modelo não pode ser considerado adequado para a modelagem e previsão da série de arrecadação do Imposto de Renda.

### 3.2.2 Média Móvel

Neste método, cada ponto da nova série modelada é calculado como sendo a média dos valores das  $m$  observações anteriores. Todas as observações são ponderadas pelo mesmo valor ( $\frac{1}{m}$ ). Apesar de já ser um avanço com relação ao método baseado unicamente na média total, ainda assim ele não é adequado para modelar séries que apresentem tendência ou

sazonalidade (Makridakis et al., 1998, pág. 143). A equação 3.2 é utilizada para os cálculos do modelo. O número de pontos conhecidos da série é representado pelo parâmetro  $n$ , enquanto  $m$  representa a ordem do modelo de média móvel.

$$\hat{y}_{n+1} = \frac{1}{m} \sum_{i=n-m+1}^n y_i \quad (3.2)$$

A série deflacionada do Imposto de Renda foi modelada utilizando-se o Método da Média Móvel com  $m = 3, 5, 7$  e  $9$ , e os gráficos da série original e da resultante da modelagem podem ser vistos na Figura 3.5 para cada uma das ordens. Realmente o método parece simular melhor a série que o da média simples. Ainda assim, os correlogramas das funções de autocorrelação dos resíduos, Figura 3.6, demonstram que a modelagem realizada não conseguiu extrair toda a informação presente na série original, dada a existência de sazonalidade nas séries.

A Estatística Q do Teste Box-Pierce também demonstra a não adequação do modelo. Fazendo-se os cálculos da Estatística Q para os resíduos de cada uma das séries, chega-se aos seguintes valores: 55.52, 63.28, 63.75, 62.64, para  $m = 3, 5, 7$  e  $9$ , respectivamente. Dado que, consultando-se a tabela com a distribuição do  $\chi^2$ , verifica-se que há um probabilidade de apenas 0.5% de um valor para a estatística Q superar 39.99, quando o número de graus de liberdade for 20, pode-se rejeitar, com quase 100% de certeza, a hipótese nula de que os resíduos são ruído branco.

Ambos as verificações indicam portanto que também o Método da Média Móvel não é adequado para a modelagem da série do Imposto de Renda.

### 3.3 Métodos de Suavização Exponencial

Diferentemente dos métodos baseados na média, nos métodos de suavização exponencial há uma ponderação dos valores observados, de sorte que às observações mais remotas são consignados pesos menores que às mais recentes, pesos estes que vão decrescendo suavemente (daí o nome do método). De acordo com Gardner (2005), os métodos de suavização exponencial tiveram a sua origem nos Estados Unidos, nas décadas de 1940/50, com os trabalhos seminais de Robert Brown e Charles Holt, para a Marinha Americana, bem como Winters.

São diversos os métodos de suavização possíveis, adotados de acordo com as características da série temporal cujo comportamento se pretende prever. Assim, podem ser utilizados métodos de suavização exponencial para séries com ou sem tendência, com ou

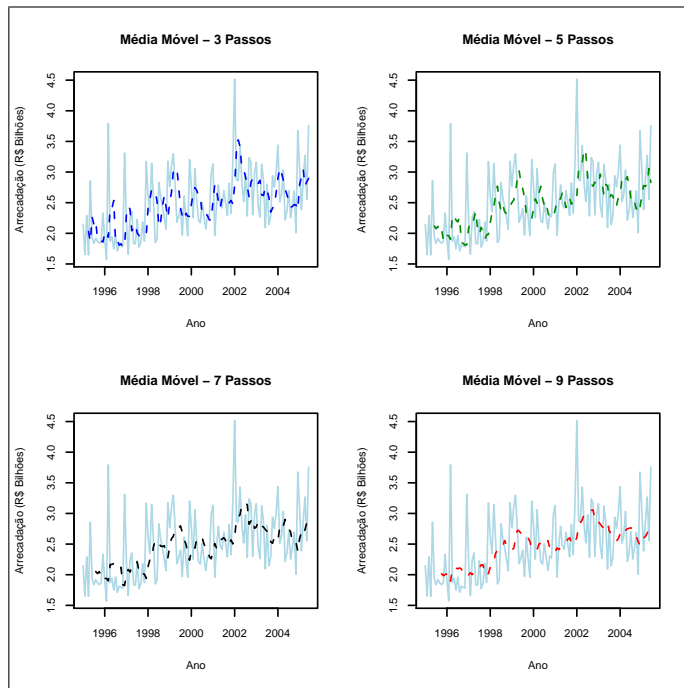


Figura 3.5: IR - Média Móvel

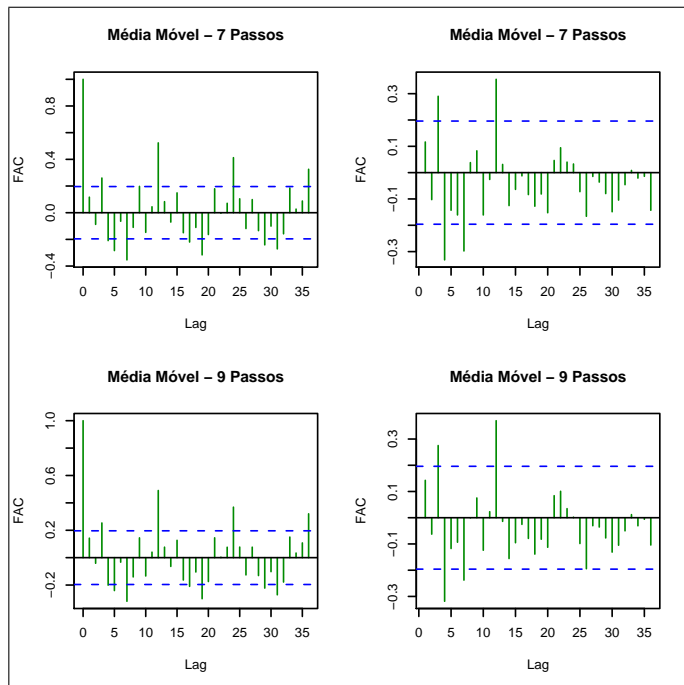


Figura 3.6: IR - Média Móvel - FAC Erros

sem sazonalidade, com tendência amortecida ou não etc. Um quadro completo contendo as diversas possibilidades pode ser encontrado em Gardner (2005, pág. 5).

Nesta seção, inicialmente, será utilizado o método de suavização mais simples para modelar a série do Imposto de Renda: série sem tendência e sem sazonalidade, N-N<sup>2</sup>. Em um segundo momento, e com base nos resultados obtidos, serão utilizados modelos mais complexos e mais próximos do comportamento da série do Imposto de Renda: tendência aditiva e sazonalidade aditiva, A-A, tendência aditiva amortecida e sazonalidade aditiva, DA-A<sup>3</sup>.

Os métodos de suavização exigem o cálculo de parâmetros ( $\alpha$ ,  $\beta$  e  $\gamma$ ) que são utilizados nas equações para a modelagem das séries. Em Makridakis et al. (1998) são sugeridas duas formas para o cálculo desses parâmetros: a utilização de algoritmos de otimização não linear e o uso de simulações computacionais, por intermédio das quais seriam escolhidos os parâmetros que minimizassem algum critério de escolha (e.g. erro quadrático médio, Estatística U de Theil ou Akaike Information Criteria-AIC). Neste trabalho foi utilizado o comando `ets` (*exponential smoothing state space modes*) do R, que permite escolher em que tipo de modelo de suavização exponencial que se quer que os dados sejam amoldados. O comando realiza automaticamente um processo de minimização não linear do critério de informação adotado (AIC, AICC ou BIC), calculando os parâmetros que serão utilizados nas equações de previsão.

Com relação ao critério para a escolha da série modelada que melhor adere à série original, não há uma opinião unânime. Foi adotado o Critério de Informação de Akaike-AIC, tendo em vista sua facilidade de implementação e os bons resultados obtidos em testes de previsão utilizando suavização exponencial reportados por Billah et al. (2005) e Hyndman et al. (2000).

Para avaliar a acurácia da previsão realizada, foram utilizados os índices EPMA (Erro Percentual Médio Absoluto) e EQM (Erro Quadrático Médio), usualmente adotados pela literatura, cujos cálculos são realizados de acordo com as Equações 3.3 e 3.4.

$$EPMA = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} * 100 \right| \quad (3.3)$$

$$EQM = \frac{1}{n} \sum_{i=1}^n (y_t - \hat{y}_t)^2 \quad (3.4)$$

---

<sup>2</sup>*No trend e No seasonality.*

<sup>3</sup>*Damped additive e Additive.*

onde  $n$  é o número de pontos da série a ser ajustada,  $y_t$  é o valor no ponto e  $\hat{y}_t$  é o valor ajustado no ponto.

### 3.3.1 Modelos sem tendência e sem sazonalidade (N-N)

Também conhecido como Modelo de Suavização Exponencial Simples, tem como objetivo prever o comportamento futuro da série com base nas observações e nos erros de previsão anteriores observados, de acordo com a equação 3.5.

$$\hat{y}_{t+1} = y_t + \alpha(y_t - \hat{y}_t) \quad (3.5)$$

O parâmetro  $\alpha$  é conhecido como a constante de suavização, se encontrando no intervalo entre 0.0 e 1.0 e  $(y_t - \hat{y}_t)$  é o termo de erro associado à previsão anterior. A série  $y_t$  é a de valores observados e a série  $\hat{y}_t$  é a de valores ajustados ao modelo com base na equação anterior. Na Figura 3.7 são apresentados os gráficos da série original, do modelo ajustado com as previsões realizadas, bem como os correlogramas das funções de autocorrelação e de autocorrelação parcial dos resíduos. Embora a série ajustada pareça acompanhar razoavelmente a arrecadação, os correlogramas indicam a existência de sazonalidade, com período de doze meses. O Teste Box-Pierce da série dos resíduos forneceu uma estatística no valor de 73.59, bastante elevada para uma distribuição  $\chi^2$  com vinte graus de liberdade, conforme pode ser visto pelo p-valor calculado:  $4.66 \cdot 10^{-8}$ . Verifica-se, portanto, que o modelo não é adequado para representar a série sob análise.

### 3.3.2 Modelos com Tendência Aditiva e Sazonalidade Multiplicativa (A-M)

Esse método permite a captura de características ligadas à tendência e à sazonalidade da série sob análise. Com este objetivo, são utilizados três parâmetros de suavização, um para o nível da série,  $\alpha$ , outro para a tendência,  $\beta$  e o terceiro,  $\gamma$ , para a sazonalidade, cujo comportamento nesse modelo é considerado como multiplicativo. Todos os três parâmetros se encontram no intervalo entre 0.0 e 1.0. As equações que regem os procedimentos de previsão são as seguintes, Makridakis et al. (1998):

$$\begin{aligned} l_t &= \alpha(y_t/s_{t-s}) + (1 - \alpha)(l_{t-1} + b_{t-1}); \\ b_t &= \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}; \\ s_t &= \gamma(y_t/l_t) + (1 - \gamma)s_{t-s}; \\ \hat{y}_{t+m} &= (l_t + mb_t)s_{t-s+m}. \end{aligned} \quad (3.6)$$



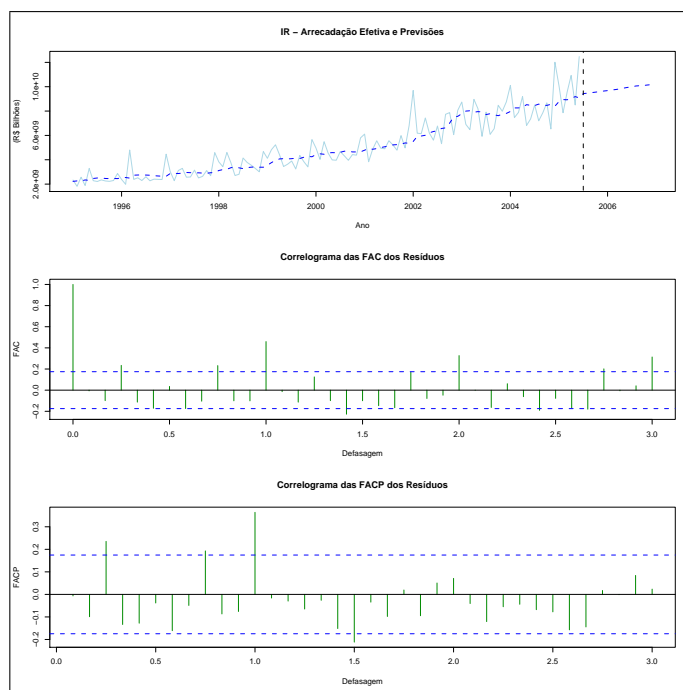


Figura 3.7: IR - Suavização Exponencial Simples

Onde,

- $\alpha$  é o parâmetro de suavização do nível da série;
- $\beta$  é o parâmetro de suavização da tendência da série;
- $\gamma$  é o parâmetro de suavização da sazonalidade da série;
- $l_t$  é a estimativa do nível da série em  $t$ ;
- $b_t$  é a estimativa do coeficiente angular da série em  $t$ ;
- $s_t$  é a estimativa do componente sazonal da série em  $t$ ;

O que diferencia uma *sazonalidade aditiva* de uma *sazonalidade multiplicativa* é a forma como o seu comportamento é considerado no modelo.

Utilizando-se o comando `ets` do pacote `Forecast` do `R`, obteve-se o modelo que melhor se ajustou à série, plotando-se a seguir os gráficos e correlogramas. Diferentemente dos métodos anteriores, os correlogramas das FAC e das FACP indicam que há pouca correlação entre os resíduos, o que permite inferir que o sinal foi em boa parte capturado pelo modelo, um indício de que este pode ser considerado elegível para previsão da arrecadação do Imposto de Renda.

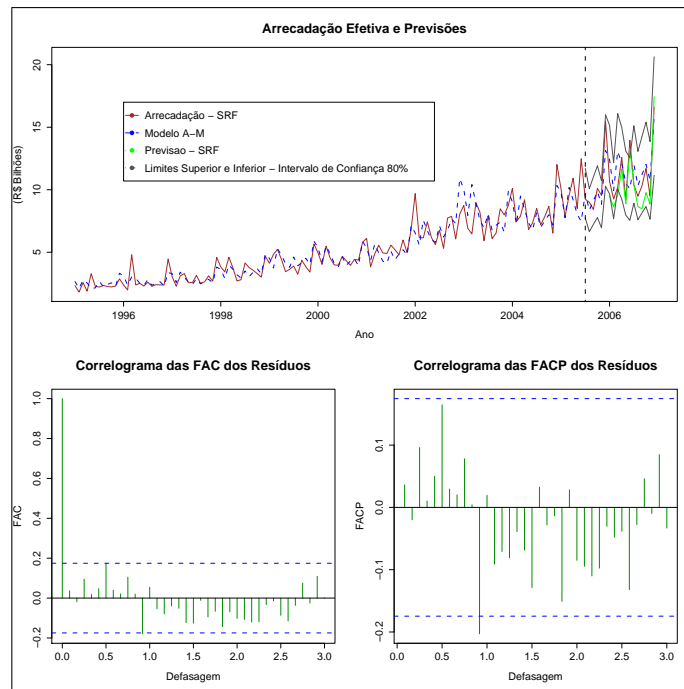


Figura 3.8: Alisamento Exponencial A-M

Calculando-se a estatística de Box-Pierce para este modelo, chega-se ao valor de 13.32, com p-valor de 0.86. Já é um avanço considerável, tendo em vista os resultados obtidos para o modelo anterior.

Considerando que o modelo é aceitável, foram feitas previsões para a arrecadação do Imposto de Renda no período de janeiro a dezembro de 2006, a qual foi inflacionada tanto pelos valores do IGP-DI previstos pela SPE, como por aqueles mensurados pela FGV. Os resultados obtidos com os dois processos de inflacionamento podem ser vistos nos Quadros 3.2 e 3.3, respectivamente, na conclusão do capítulo, Seção 3.4, onde são feitos os comentários pertinentes a cada um dos modelos elegíveis para a realização das previsões.

Além das estimativas pontuais de arrecadação, há que se calcular também os intervalos de confiança em torno dos quais essas estimativas se distribuem. No presente trabalho optou-se por estimar essa distribuição em torno dos intervalos de 90% e 95% de confiança. No R, o procedimento de cálculo é realizado automaticamente utilizando-se o comando `forecast`.

Os valores para os intervalos de confiança não podem, no entanto, ser simplesmente inflacionados pelo IGP-DI (previsto pela SPE, ou mensurado pela FGV), como os valores pontuais, uma vez que esses intervalos dependem do cálculo da variância da dis-

tribuição em cada ponto. Fazer o inflacionamento puro e simples seria um procedimento incorreto. Assim, são realizados os seguintes procedimentos para cada ponto previsto da série deflacionada:

- calculam-se os valores pontuais da previsão,  $x_i$ , bem como os valores máximos,  $x_{i2}$ , e mínimos,  $x_{i1}$ , para o intervalo de confiança (e.g. 95%);
- calcula-se o erro-padrão no ponto para o intervalo de confiança desejado: ( $ep(x_i) = (x_{i2} - x_{i1}) / (2.0 * 1.95)$ );
- calcula-se a variância:  $var(x_i) = (ep(x_i))^2$ ;
- o novo valor pontual ( $x_{in}$ ) será ( $x_i * Ind_i$ ), onde  $Ind_i$  é o índice do IGP-DI em  $i$ ;
- o valor da nova variância no ponto é  $var_n(x_i) = var(Ind_i * x_i) = Ind_i^2 * var(x_i)$ ;
- o valor do novo erro-padrão no ponto é  $ep_n(x_i) = \sqrt{var_n(x_i)}$ ;
- os novos limites para o intervalo de confiança de 95% são  $x_{in} = 1.95 * ep_n(x_{in})$ .

Todas as análises com séries deflacionadas seguirão exatamente essa metodologia de inflacionamento, doravante, motivo pelo qual essa explicação não será repetida nas próximas seções, ou no Capítulo 4, referente à análise utilizando os modelos de Box&Jenkins.

### 3.3.3 Modelos com Tendência Aditiva e Sazonalidade Aditiva (A-A)

Esse método permite a captura de características ligadas à tendência e à sazonalidade da série, ambas consideradas com possuidoras de comportamento aditivo. Assim como no método anterior, são três os parâmetros de suavização a serem calculados, um para o nível da série,  $\alpha$ , outro para a tendência,  $\beta$  e o terceiro,  $\gamma$ , para a sazonalidade, cujos valores se encontram entre 0.0 e 1.0. As equações que regem os procedimentos de previsão são as seguintes, Makridakis et al. (1998):

$$\begin{aligned}
 l_t &= \alpha(y_t - s_{t-s}) + (1 - \alpha)(l_{t-1} + b_{t-1}); \\
 b_t &= \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}; \\
 s_t &= \gamma(y_t - l_t) + (1 - \gamma)s_{t-s}; \\
 \hat{y}_{t+m} &= l_t + mb_t + s_{t-s+m}.
 \end{aligned}
 \tag{3.7}$$

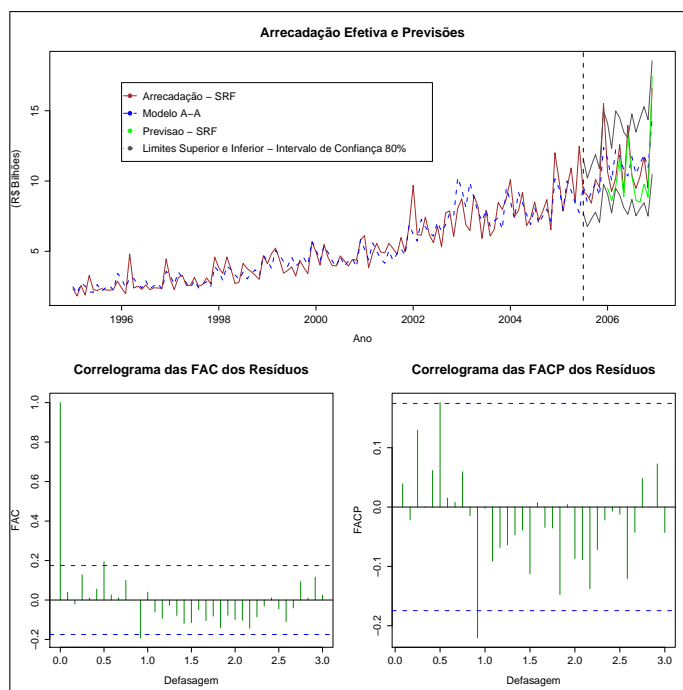


Figura 3.9: IR - Suavização Exponencial Modelo A-A

Onde,

- $\alpha$  é o parâmetro de suavização do nível da série;
- $\beta$  é o parâmetro de suavização da tendência da série;
- $\gamma$  é o parâmetro de suavização da sazonalidade da série;
- $l_t$  é a estimativa do nível da série em  $t$ ;
- $b_t$  é a estimativa do coeficiente angular da série em  $t$ ;
- $s_t$  é a estimativa do componente sazonal da série em  $t$ ;

Na Figura 3.9 são apresentados o gráfico da série modelada, bem como os correlogramas das funções de autocorrelação e de autocorrelação parcial dos resíduos da série A-A que apresentou o menor valor de AIC. A sazonalidade da série foi bem capturada pelo modelo e o resultado da Estatística Q para o Teste de Box-Pierce foi de 21.05, com um p-valor de 0.39. Com base nos critérios adotados, vê-se que não é um modelo tão adequado quanto o sazonal multiplicativo; ainda assim será utilizado para os cálculos das previsões, até como uma forma de comparação entre os diversos modelos elegíveis para os cálculos.

Os resultados para a previsão de arrecadação entre janeiro e dezembro de 2006 podem ser vistos nos Quadros 3.2 e 3.3 na Seção 3.4.

### 3.3.4 Modelos com Tendência Multiplicativa e Sazonalidade Multiplicativa (M-M)

Esse método permite a captura de características ligadas à tendência e à sazonalidade da série, ambas consideradas como possuidoras de comportamento multiplicativo. As equações que regem os procedimentos de previsão são as seguintes, Makridakis et al. (1998):

$$\begin{aligned} l_t &= \alpha(y_t - s_{t-s}) + (1 - \alpha)(l_{t-1} + b_{t-1}); \\ b_t &= \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}; \\ s_t &= \gamma(y_t - l_t) + (1 - \gamma)s_{t-s}; \\ \hat{y}_{t+m} &= l_t + mb_t + s_{t-s+m}. \end{aligned} \tag{3.8}$$

Onde,

- $\alpha$  é o parâmetro de suavização do nível da série;
- $\beta$  é o parâmetro de suavização da tendência da série;
- $\gamma$  é o parâmetro de suavização da sazonalidade da série;
- $l_t$  é a estimativa do nível da série em  $t$ ;
- $b_t$  é a estimativa do coeficiente angular da série em  $t$ ;
- $s_t$  é a estimativa do componente sazonal da série em  $t$ ;

Na Figura 3.10 são apresentados o gráfico da série modelada, bem como os correlogramas das funções de autocorrelação e de autocorrelação parcial dos resíduos do modelo que apresentou o menor valor de AIC. O resultado da Estatística Q para o Teste de Box-Pierce foi de 19.05, com um p-valor de 0.52. Assim como o anterior, também será utilizado para comparação e os resultados obtidos podem ser vistos nos Quadros 3.2 e 3.3 na Seção 3.4.

### 3.3.5 Modelos com Tendência Amortecida Aditiva e Sazonalidade Aditiva (DA-A)

A utilização de um modelo como esse se basearia na hipóteses de que a tendência de crescimento da série não se daria de forma indefinida, mas em algum momento iniciaria

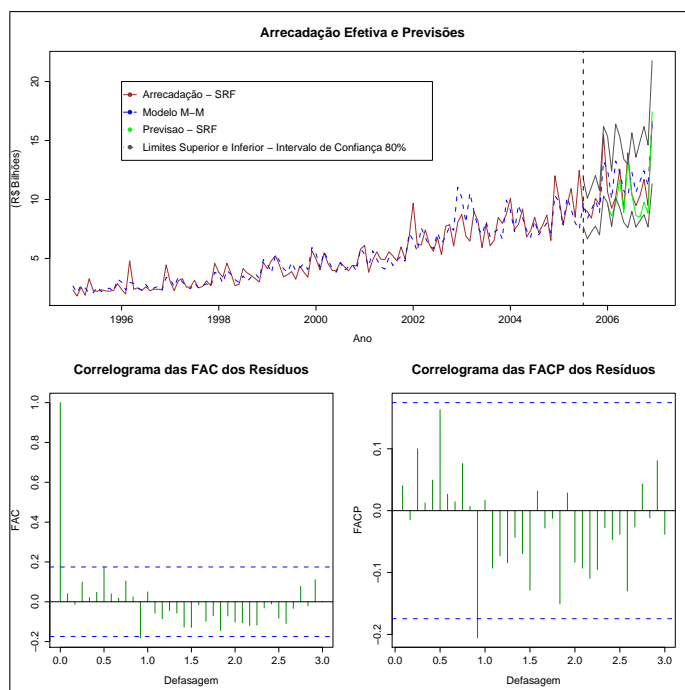


Figura 3.10: IR - Suavização Exponencial Modelo M-M

um processo de estabilização (amortecimento). Em se tratando de arrecadação de tributos, essa idéia não é desprezível, uma vez que em algum momento a trajetória de crescimento da carga tributária necessariamente terá que desacelerar, estabilizando a tendência da série.

O gráfico com as séries de arrecadação efetiva e prevista, juntamente com os correlogramas dos resíduos, podem ser vistos na Figura 3.11. Ambos os correlogramas indicam uma reduzida correlação entre os resíduos. A Estatística Q de Box-Pierce, com valor de 21.05 e p-valor de 0.39, também é um bom indicador. O modelo será utilizado para se verificar se o processo de crescimento da arrecadação do Imposto de Renda já se encontra em fase de estabilização. Os resultados, juntamente com as análises, se encontram na Seção 3.4.

## 3.4 Conclusão

A análise necessariamente terá que se dividir em duas, conforme os índices de inflacionamento, previsto pela SPE e mensurado pela FGV, utilizados.

Quando o inflacionamento dos valores calculados é feito com os índices previstos pela SPE, os resultados podem ser vistos no Quadro 3.2. Dois dos quatro modelos estudados

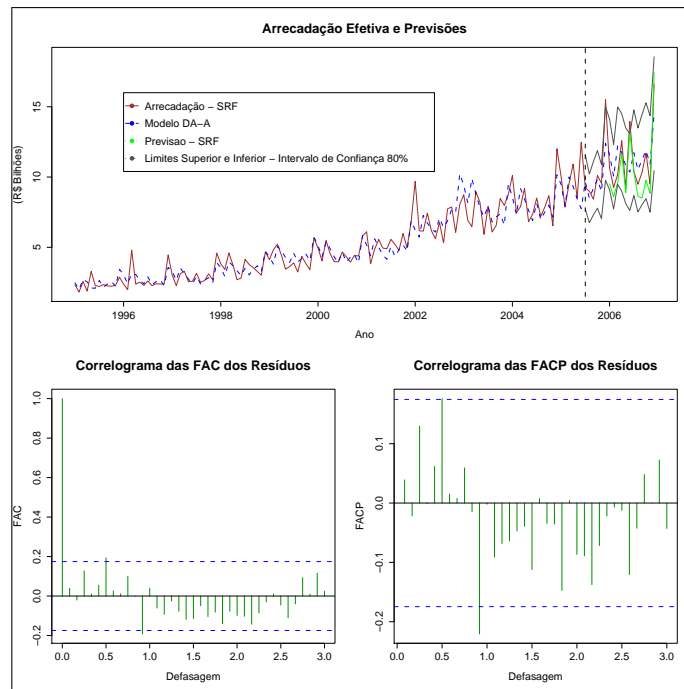


Figura 3.11: Suavização Exponencial Modelo DA-A

sobressaem: o A-M e o A-A, com vantagem para o segundo, uma vez que apresentou o menor erro absoluto para o somatório da arrecadação. O modelo M-M, embora tenha apresentado o menor Erro Quadrático Médio-EQM, foi o que forneceu o maior erro absoluto do somatório, bem como Erro Percentual Médio Absoluto-EPMA.

No geral, pôde ser observado nos gráficos de evolução das séries que a arrecadação do Imposto de Renda realizada pela Secretaria da Receita Federal se manteve dentro do intervalo de previsão de 80%, o que pode ser considerado um indício de uma boa aderência do modelo à realidade.

Apesar disso, diferentemente das previsões da SRF, todos os quatro modelos superestimaram a arrecadação. O corpo técnico que adotasse essa metodologia para as previsões poderia se ver na difícil situação de prever arrecadações superiores àquelas que seriam realizadas, com possíveis conseqüências negativas para o equilíbrio fiscal. Cotejando-se o erro absoluto do somatório e o EPMA dos modelos A-M e A-A com os da previsão da SRF, no Quadro 2.1, na página 16, vê-se que ambos são menores que os da previsão da SRF.

Analisando agora os resultados obtidos com o inflacionamento feito pelos índices mensurados (reais), apresentados no Quadro 3.3, vê-se que os erros absolutos nos quatro modelos para o somatório da arrecadação são sensivelmente menores que aqueles observados

no Quadro 3.2. O modelo A-M, em especial, conseguiu prever o total de arrecadação em 2006 com uma diferença de apenas 0.5%. Esse resultado para 2006 parece indicar que, no limite, se fosse possível acertar perfeitamente os índices de inflação mensais, os modelos de suavização exponencial lograriam prever com razoável precisão a arrecadação total do Imposto de Renda.

Apesar disso, o EPMA de todos os modelos é superior àqueles das séries inflacionadas pelo índice previsto pela SPE. Em outras palavras, os erros mensais são maiores.

Uma última observação, que vale para todos os resultados a serem apresentados no presente trabalho, é que todos os cálculos e previsões foram feitos para apenas um ano, 2006. Para se afirmar taxativamente que este ou aquele modelo é superior aos demais, seria necessário realizar previsões para diversos anos, cotejá-las com a arrecadação obtida e verificar qual o que apresentou os melhores resultados. Se os resultados convergirem para um mesmo modelo, aí sim se poderia concluir pela sua superioridade.

Um outro ponto que deve ser observado é a forma como essas informações geradas a partir de cálculos econométricos seriam inseridas na proposta orçamentária. Ao informar que a previsão para a arrecadação em determinado mês do ano seguinte é de  $x \pm \Delta x$ , com um intervalo de confiança de, e.g., 95%, pode-se incorrer em um risco de serem feitas alterações (para maior) nos gastos previstos, dado que poderia ser alegado que há uma chance razoável de que a arrecadação venha a ser superior à previsão pontual. Dado que a arrecadação pode ser maior, mas também pode ser menor, há um potencial problema de estabilidade fiscal.



Quadro 3.2: Suavização Exponencial - IGP-DI/SPE

Mês (2006)	Valores Arrecadados	Modelo A-M	$\Delta$ (%)	Modelo A-A	$\Delta$ (%)	Modelo M-M	$\Delta$ (%)	Modelo DA-A	$\Delta$ (%)
jan	10.643.316.250,33	12.398.007.739	16.49	11.596.705.093	8.96	12.521.789.265	17.65	11.596.705.093	8.96
fev	9.259.073.435,62	9.929.551.222	7.24	10.022.654.535	8.25	10.051.298.056	8.55	10.022.654.535	8.25
mar	10.189.952.436,27	13.074.915.759	28.31	12.231.605.426	20.04	13.272.041.536	30.25	12.231.605.426	20.04
abr	12.588.188.574,93	12.115.380.127	-3.76	11.744.632.159	-6.70	12.322.880.935	-2.11	11.744.632.159	-6.70
mai	9.180.874.262,48	10.529.583.902	14.69	10.793.840.377	17.57	10.741.579.293	17.00	10.793.840.377	17.57
jun	13.956.711.467,90	10.063.542.670	-27.89	10.353.035.336	-25.82	10.291.599.217	-26.26	10.353.035.336	-25.82
jul	10.397.486.714,18	12.024.510.067	15.65	11.733.692.122	12.85	12.327.362.155	18.56	11.733.692.122	12.85
ago	9.485.565.072,44	10.328.373.713	8.88	10.504.175.834	10.74	10.620.106.356	11.96	10.504.175.834	10.74
set	10.312.573.908,60	11.207.543.028	8.68	11.229.992.289	8.89	11.556.233.445	12.06	11.229.992.289	8.89
out	11.703.111.572,74	12.029.226.562	2.79	11.872.785.682	1.45	12.440.403.797	6.30	11.872.785.682	1.45
nov	9.492.285.018,92	10.745.785.615	13.21	10.927.288.903	15.12	11.148.273.247	17.45	10.927.288.903	15.12
dez	16.606.948.120,64	15.898.791.444	-4.26	14.504.894.547	-12.66	16.545.183.388	-0.37	14.504.894.547	-12.66
Total	133.816.086.835,05	140.345.211.854	4.88	137.515.302.309	2.76	143.838.750.696	7.49	137.515.302.309	2.76
EQM	-	4.18 10 <sup>16</sup>		3.68 10 <sup>17</sup>		3.18 10 <sup>14</sup>		3.68 10 <sup>17</sup>	
EPMA	-	6.67		4.89		9.25		4.89	

Quadro 3.3: Suavização Exponencial - IGP-DI/FGV

Mês (2006)	Valores Arrecadados	Modelo A-M	$\Delta$ (%)	Modelo A-A	$\Delta$ (%)	Modelo M-M	$\Delta$ (%)	Modelo DA-A	$\Delta$ (%)
jan	10.643.316.251,32	12.103.585.759	13,72	11.321.312.065	6,37	12.224.427.780	14,86	11.321.312.065	6,37
fev	9.259.073.435,42	9.639.963.688	4,11	9.730.351.716	5,09	9.758.159.871	5,39	9.730.351.716	5,09
mar	10.189.952.437,74	12.582.356.917	23,48	11.770.815.811	15,51	12.772.056.563	25,34	11.770.815.811	15,51
abr	12.588.188.576,04	11.600.160.801	-7,85	11.245.179.281	-10,67	11.798.837.418	-6,27	11.245.179.281	-10,67
mai	9.180.874.264,09	10.051.274.728	9,48	10.303.527.281	12,23	10.253.640.171	11,68	10.303.527.281	12,23
jun	13.956.711.468,44	9.612.683.358	-31,13	9.889.206.390	-29,14	9.830.522.686	-29,56	9.889.206.390	-29,14
jul	10.397.486.714,18	11.444.679.251	10,07	11.167.884.764	7,41	11.732.927.585	12,84	11.167.884.764	7,41
ago	9.485.565.072,44	9.810.728.918	3,43	9.977.720.063	5,19	10.087.840.297	6,35	9.977.720.063	5,19
set	10.312.573.908,60	10.647.870.805	3,25	10.669.199.015	3,46	10.979.148.633	6,46	10.669.199.015	3,46
out	11.703.111.572,74	11.491.946.291	-1,80	11.342.492.777	-3,08	11.884.758.471	1,55	11.342.492.777	-3,08
nov	9.492.285.018,92	10.285.258.410	8,35	10.458.983.096	10,18	10.670.496.815	12,41	10.458.983.096	10,18
dez	16.606.948.120,64	15.220.972.258	-8,35	13.886.501.894	-16,38	15.839.806.330	-4,62	13.886.501.894	-16,38
Total	133.816.086.840,56	134.491.481.184	0,50	131.763.174.153	-1,53	137.832.622.620	3,00	131.763.174.153	-1,53
EQM	-	2.71 10 <sup>18</sup>		1.22 10 <sup>20</sup>		1.35 10 <sup>20</sup>		1.22 10 <sup>20</sup>	
EPMA	-	10,42		10,39		11,45		10,39	

# Capítulo 4

## Metodologia Box&Jenkins

Uma série temporal, cujos valores podem ser derivados de algum experimento ou da observação de algum fenômeno físico ou social, pode ser dividida em duas partes. A primeira é o *signal*, o componente dos dados que contém a informação e que pode ser modelado mediante a utilização de algum instrumental matemático, e o segundo é o ruído, componente imprevisível, aleatório, oriundo da impossibilidade de incluir no modelo *tudo* o que acontece no universo. Por mais acuradas que sejam as observações e melhores os modelos de previsão construídos, sempre haverá espaço para um acontecimento fortuito (uma revolução, um período chuvoso ou de estiagem muito acima do normal, a quebra inesperada de algum equipamento etc.), gerando incertezas no processo de previsão.

O propósito desse trabalho é buscar modelar da melhor forma possível o sinal existente na série de arrecadação do Imposto de Renda, utilizando então esse modelo para a realização de previsões de arrecadação. No Capítulo 3 foram utilizados métodos baseados na média e na suavização exponencial, cujos resultados foram sintetizados na Seção 3.4. Neste capítulo será utilizada a metodologia desenvolvida Box&Jenkins para a modelagem de dados e previsão com base em séries temporais.

O método, desenvolvido seminalmente nos anos 1970 pelos estatísticos George Box e Gwilym Jenkins é baseado na utilização de funções auto-regressivas (*Autoregressive*, AR) e de média móvel, (*Moving Average*, MA), que, sozinhas ou, mais freqüentemente, combinadas, permitem a reprodução da melhor forma possível da *informação*, sinal, contida nos dados.

Os modelos AR implicam que o sinal da série de tempo pode ser expresso como uma função linear de valores pretéritos da série de tempo. Exemplos:

$$y_t = \alpha_0 + \phi_1 y_{t-1} + \xi_t$$

$$y_t = \alpha_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \xi_t$$

$$y_t = \alpha_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \xi_t$$

De forma a indicar o grau  $n$  de autoregressividade da série, coloca-se o número entre parênteses. Assim, AR(1) é um modelo auto-regressivo de primeira ordem, AR(2), de segunda ordem e assim sucessivamente. Como o próprio nome dá a entender, no caso dos modelos auto-regressivos, o que se faz é uma regressão de  $y_t$  em  $y_{t-1}$ ,  $y_{t-2}$ ,  $y_{t-3}$  ..., calculando-se os valores do coeficientes.

Já os modelos MA são baseados na pressuposição de que o sinal pode ser expresso como uma função dos erros de previsão pretéritos. Ao se fazerem previsões baseadas nos erros ocorridos, busca-se melhorar as novas previsões. Exemplos de modelos de média móvel são:

$$y_t = \alpha_0 + \theta_1 \xi_{t-1} + \xi_t$$

$$y_t = \alpha_0 + \theta_1 \xi_{t-1} + \theta_2 \xi_{t-2} + \xi_t$$

$$y_t = \alpha_0 + \theta_1 \xi_{t-1} + \theta_2 \xi_{t-2} + \theta_3 \xi_{t-3} + \xi_t$$

Uma terceira possibilidade, mais comum, é que seja feita uma conjugação de efeitos auto-regressivos e de média móvel, gerando um modelo misto, ARMA(p,q), onde p é a ordem auto-regressiva e q é a ordem de média móvel. Alguns exemplos de modelos ARMA podem ser vistos a seguir:

$$ARMA(1, 2) : y_t = 2.0 + 0.18y_{t-1} + 0.35\xi_{t-1} + 0.15\xi_{t-2} + \xi_t;$$

$$ARMA(2, 2) : y_t = 2.0 + 0.18y_{t-1} + 0.43y_{t-2} - 0.55\xi_{t-1} + 0.42\xi_{t-2} + \xi_t.$$

Cada um dos modelos pode ser representado por uma equação geral:

$$AR(p) : y_t = \alpha_0 + \sum_{i=1}^p \theta_i y_{t-i} + \xi_t;$$

$$MA(q) : y_t = \alpha_0 + \sum_{k=1}^q \phi_k \xi_{t-k} + \xi_t;$$

$$ARMA(p, q) : y_t = \alpha_0 + \sum_{i=1}^p \theta_i y_{t-i} + \sum_{k=1}^q \phi_k \xi_{t-k} + \xi_t;$$

Os modelos, no entanto, podem se tornar ainda mais complexos quando a série a ser analisada apresenta características não estacionárias na média e/ou sazonais. No intuito de lidar com essa multiplicidade de características, Box&Jenkins propuseram uma equação geral, Equação 4.1, a partir da qual, com a escolha de parâmetros adequados, quaisquer séries temporais poderiam ser modeladas.

$$\phi_p(B)\Phi_p(B^L)\Delta^d\Delta_s^D y_t = \theta_q(B)\Theta_q(B^L)\xi_t \quad (4.1)$$

onde,

- B - operador *backshift*  $\rightarrow By_t = y_{t-1}$ ;
- $\Delta^d$  - operador de diferença :  $(1 - B)^d$ ;
- $\Delta_s^D$  - operador de diferença sazonal:  $(1 - B^s)^D$ ;
- $\xi_t$  - var. aleatória indicativa de choques, apresentando a distribuição  $\xi_t \sim N(0, \sigma^2)$ ;
- s - período da sazonalidade da série;
- d - ordem de diferenciação da série para torná-la estacionária em nível;
- D - ordem de diferenciação sazonal da série para torná-la sazonalmente estacionária;
  
- p - ordem do modelo auto-regressivo associado à série;
- P - ordem do modelo auto-regressivo sazonal associado à série;
- q - ordem do modelo de média móvel associado à série;
- Q - ordem do modelo de média móvel sazonal associado à série;
- $\phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$  - operador auto-regressivo não-sazonal de ordem p;
- $\Phi_P(B^s) = (1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_P B^{Ps})$  - operador auto-regressivo sazonal de ordem P;
- $\theta_q(B) = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q)$  - operador de média móvel não-sazonal de ordem q;
- $\Theta_Q(B^s) = (1 + \Theta_1 B^s + \Theta_2 B^{2s} + \dots + \Theta_Q B^{Qs})$  - operador de média móvel sazonal de ordem Q.

## 4.1 Análise da Série

O objetivo desta parte do trabalho é fazer uma análise da série do Imposto de Renda deflacionada pelo IGP-DI<sup>1</sup>, no período de janeiro de 1995 a junho de 2005, de forma

---

<sup>1</sup>Será considerado aqui que o IGP-DI puro é o índice mais adequado para o deflacionamento/inflacionamento dos dados para sua modelagem com o método de Box&Jenkins, assim como o foi para a suavização exponencial.

a verificar em qual(is) modelo(s) Box&Jenkins previstos pela Equação 4.1 ela melhor adere. Uma vez definido(s) o(s) modelo(s), serão feitas previsões para a arrecadação do tributo no período de janeiro a dezembro de 2006, verificando-se a seguir a acurácia da modelagem realizada.

A opção por utilizar a série de valores conhecidos até junho de 2005, apenas, se deve a que, um dos objetivos do presente trabalho é justamente procurar analisar se a utilização de modelos econométricos de séries temporais traria vantagens no processo de previsão de receitas no âmbito da elaboração da proposta orçamentária.

Os cálculos de previsão de receita que constam do Projeto de Lei Orçamentária Anual (PLOA) são elaborados com dados de arrecadação até junho de cada ano. Portanto, qualquer simulação do processo só poderá utilizar dados efetivos de arrecadação até esse mês. Por esse motivo, cada modelo adotado será utilizado para prever a arrecadação de 18 (dezoito) meses (o período de julho a dezembro de 2005 e o ano inteiro de 2006).

São seis as etapas a serem seguidas para a realização do trabalho no âmbito da metodologia desenvolvida por Box&Jenkins:

1. verificação da estacionariedade da série;
2. utilização dos procedimentos adequados para tornar a série estacionária, caso se identifique a sua não estacionariedade;
3. identificação do(s) modelo(s) Box-Jenkins mais adequado(s) para a sua modelagem;
4. realização do processo de previsão;
5. cotejamento dos resultados obtidos a partir das previsões com a arrecadação efetiva;
6. cálculo dos erros de previsão.

Primeiramente, iremos analisar a série que será utilizada nas previsões. Como visto no Capítulo 3, apenas os valores de arrecadação no período de janeiro de 1995 a junho de 2004 serão utilizados nas previsões. No primeiro gráfico da Figura 4.1 a evolução dessa série (já deflacionada) é mostrada. No segundo gráfico é mostrada a série após a aplicação do operador logaritmo, utilizado para tornar a variância da série estacionária.

A análise de ambos os gráficos permite inferir que a série não é estacionária na média, uma vez que a mesma apresenta uma trajetória de crescimento ao longo do período. Essa primeira impressão é corroborada com o cálculo das médias anuais, cuja evolução pode ser vista no gráfico da Figura 4.2.

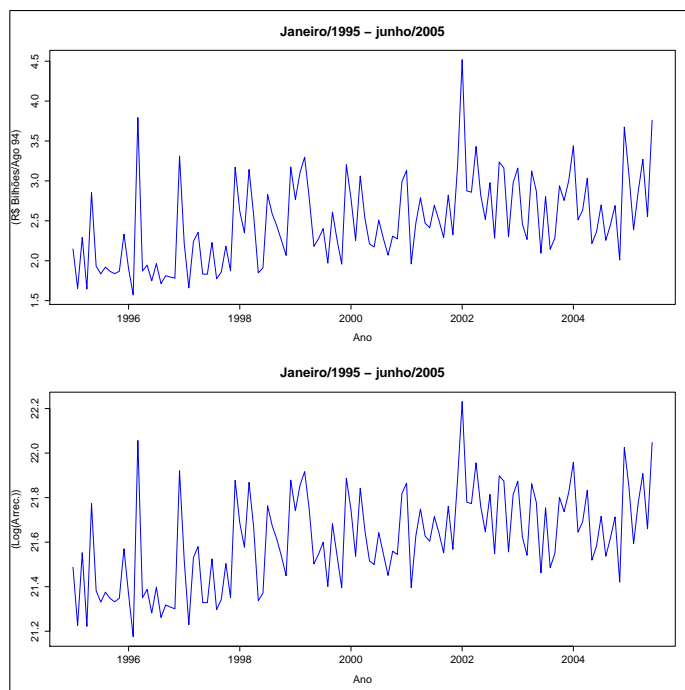


Figura 4.1: Série do Imposto de Renda Normal e em Log

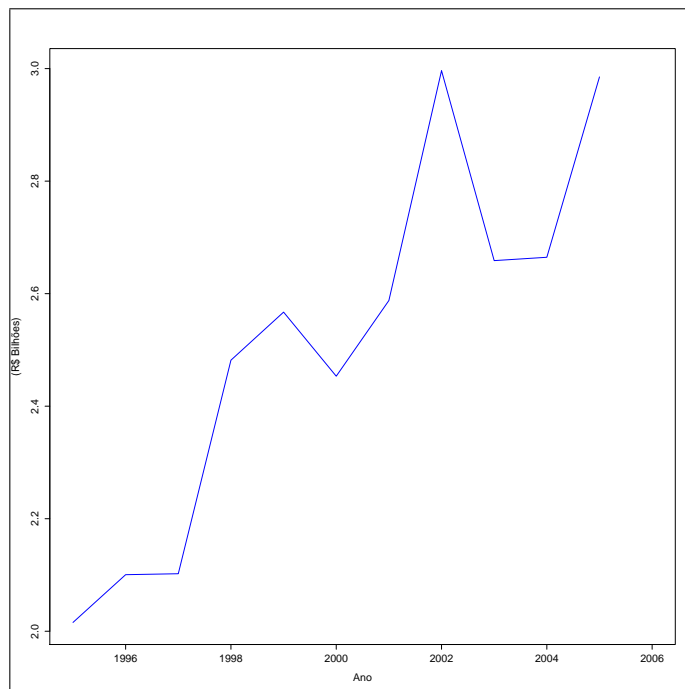


Figura 4.2: Evolução da Média de Arrecadação do Imposto de Renda

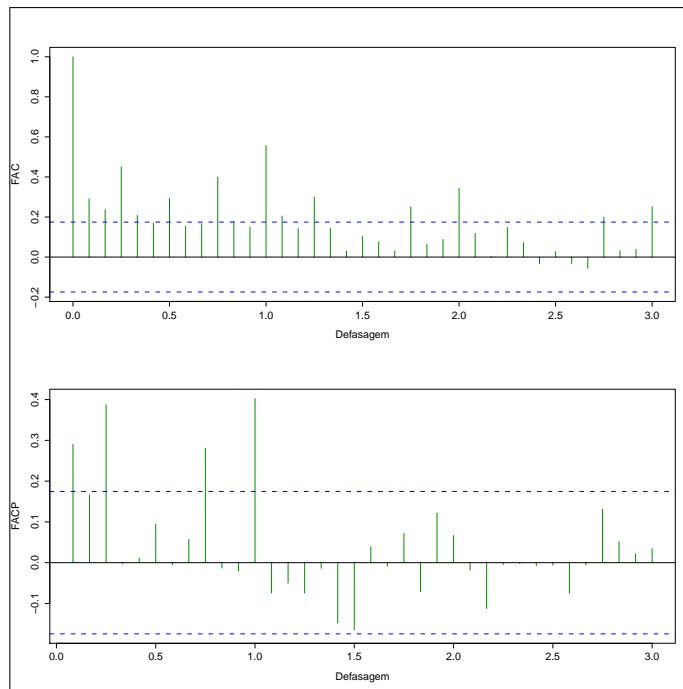


Figura 4.3: FAC e FACP da Série Deflacionada

Calculando-se as funções de autocorrelação e de autocorrelação parcial da série, são obtidos os correlogramas apresentados na Figura 4.3. Há fortes evidências a respeito de seu comportamento sazonal, com picos nas defasagens múltiplas de 12 (12, 24 e 36) do correlograma das FAC, bem como na defasagem 12 do correlograma das FACP, conforme é esperado de uma série sazonal. Uma primeira opinião seria no sentido de uma sazonalidade estacionária, uma vez que os valores dos picos decaem rapidamente.

Nessa parte do trabalho, deve-se decidir como lidar com a não estacionariedade da média, verificando-se se ela decorre da existência de uma raiz unitária ou de uma tendência inerente à série. A verificação de ocorrência de raiz unitária é feita com o Teste de Dickey-Fuller, utilizando o comando `adf.test` (Augmented Dickey-Fuller Test) do **R** cujos resultados, estatística no valor de  $-4.6365$  e  $p$ -valor de  $0.01$ , apontam para a ausência dessa característica. Não haveria, em princípio, necessidade de diferenciação da série. O crescimento do nível médio da série pode então ser aproximado por uma equação simulando uma tendência, a qual teria o seguinte formato:

$$y_t = tend_t + saz_t + \xi_t \quad (4.2)$$

onde  $tend_t$  é a parte da equação que representa a tendência da série, enquanto  $saz_t$  representa o comportamento sazonal. O termo  $\xi_t$  representa o erro. A equação que rege o



Quadro 4.1: Parâmetros das Equações de Regressão

Ordem da Regressão	Parâmetro	Valor	Estat. t	Prob. (t)
linear	$\alpha$	21.415	657.43	$< 2 \cdot 10^{-16}$
	$\beta$	0.003	6.95	$1.83 \cdot 10^{-10}$
Quadrática	$\alpha$	21.343	437.44	$< 2 \cdot 10^{-16}$
	$\beta_1$	$6.473 \cdot 10^{-3}$	3.65	0.000387
	$\beta_2$	$-2.661 \cdot 10^{-5}$	-1.97	0.051474

comportamento da tendência pode ser linear, quadrática etc. Neste trabalho serão considerados dois modelos para a tendência: linear e quadrático. As equações 4.3 a seguir ilustram cada um deles.

$$\begin{aligned} tend_t^{lin} &= \alpha + \beta t \\ tend_t^{quad} &= \alpha + \beta_1 t + \beta_2 t^2 \end{aligned} \quad (4.3)$$

Fazendo-se as regressões em relação a  $t$ , no caso linear, e a  $t$  e  $t^2$ , para a tendência quadrática, obtêm-se os valores para os parâmetros no Quadro 4.1. Em ambos os casos a estatística  $t$  e o p-valor indicam que os parâmetros calculados podem ser considerados estatisticamente significativos.

Lembrando a Equação 4.2, pode-se retirar o efeito da tendência, obtendo-se a Equação 4.4, onde a tendência é representada por um dos termos da Equação 4.3.

$$z_t = y_t - tend_t = saz_t + \xi_t \quad (4.4)$$

Na Figura 4.4 são mostrados os correlogramas das FAC e das FACP para cada uma das séries cuja tendência, linear ou quadrática, foi retirada. O que se pode observar, em ambos os casos é que as séries resultantes da retirada da tendência se mantêm sazonalmente estacionárias. Existem alguns picos não sazonais na terceira e na nona defasagens, tanto no correlograma da FAC, como no da FACP.

De forma a se verificar qual o modelo Box&Jenkins que melhor se adapta à série, será realizado um processo iterativo, variando-se os parâmetros  $p$ ,  $d$ ,  $q$ ,  $P$ ,  $D$  e  $Q$  da Equação 4.1 no comando `arima` do R, apresentado a seguir, obtendo-se diversos modelos diferentes de séries temporais. Para cada modelo são calculados os valores do AIC (Akaike Information Criteria) e BIC (Bayesian Information Criteria) e escolhidos aqueles que apresentam os menores valores.

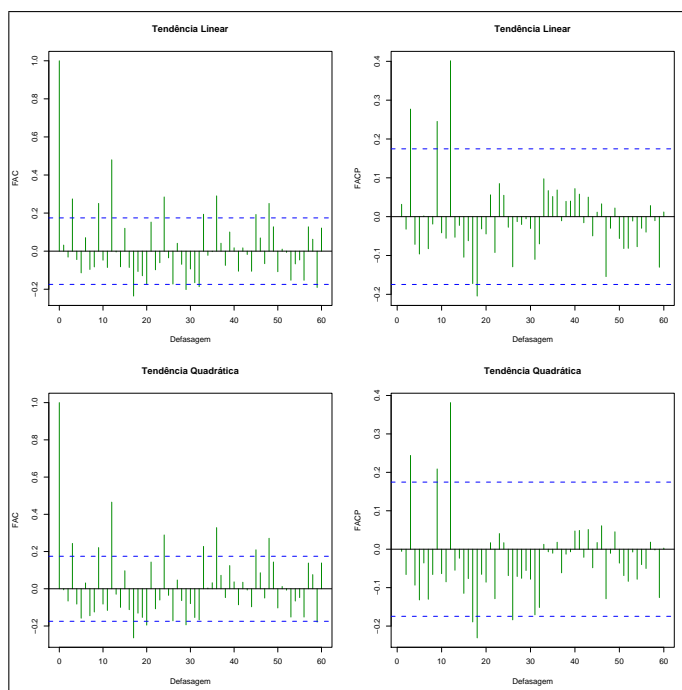


Figura 4.4: Correlogramas das Séries sem Tendência

```
Teste <- arima(Serie, order=c(p,d,q), seasonal=list(order=c(P,D,Q), period
= 12), xreg=NULL, include.mean=TRUE)
```

Os oito modelos que minimizaram os critérios de informação adotados podem ser vistos no Quadro 4.2, sendo todos estacionários na média e na sazonalidade ( $d = 0$  e  $D = 0$ ). Os modelos adotados pelo BIC são mais parcimoniosos que os previstos pelo AIC, conforme a teoria prediz, sendo escolhidos para a realização das previsões de arrecadação.

Quadro 4.2: Melhores Modelos - AIC/BIC

Ordem da Regressão	Hierarquia	AIC	BIC
linear	1º	SARIMA(1,0,3)(1,0,1)	SARIMA(0,0,0)(1,0,0)
	2º	SARIMA(0,0,1)(1,0,1)	SARIMA(0,0,0)(1,0,1)
Quadrática	1º	SARIMA(0,0,0)(1,0,1)	SARIMA(0,0,0)(1,0,1)
	2º	SARIMA(0,0,0)(1,0,2)	SARIMA(0,0,0)(1,0,0)

O resultado `Teste` é um objeto com os parâmetros do modelo arima associados à série temporal. O comando `forecast`, com a sintaxe apresentada a seguir, com base no objeto `Teste`, faz a previsão dos valores da série `n` períodos de tempo à frente (no nosso caso dezoito meses), bem como os intervalos de predição desejados, 80% e 95%.

```
y <- forecast(Teste, h=18, conf=c(80,95))
```

Os resultados de todos esses procedimentos são apresentados graficamente nas Figuras 4.5, 4.7, 4.8 e 4.8. No primeiro gráfico inferior de cada figura, à direita, é mostrada a série tratada, com os valores previstos em azul e os intervalos de predição, em preto e vermelho. Tanto as séries modeladas com tendência linear, como aquelas com tendência quadrática, aparentam ser estacionárias. Os correlogramas das séries dos resíduos corroboram essa impressão, não apresentando picos fora do intervalo de confiança. Note-se que as características de sazonalidade somem das séries dos resíduos, evidenciando que o modelo capturou bem esse comportamento.

No gráfico maior das figuras são mostradas as evoluções das séries de arrecadação, das séries ajustadas pelo modelos<sup>2</sup> e dos valores de arrecadação previstos pela SRF. Regra geral, todos os modelos lograram acompanhar a evolução da série de arrecadação. No período de previsão, o ano de 2006, as três séries, para os quatro modelos, se embaralham, o que pode ser uma evidência de que os modelos, desenvolvidos no presente trabalho e da SRF, apresentam uma certa aderência à realidade.

Nos Quadros 4.4 e 4.3 são apresentados os resultados numéricos das previsões com os modelos, inflacionados pelo IGP-DI previsto pela SPE e o mensurado pela FGV, respectivamente.

## 4.2 Conclusão

Nos quadros 4.3 e 4.4 são apresentados os resultados das previsões para os modelos escolhidos, inflacionados para os meses de 2006 com base no IGP-DI previsto pela SPE e no IGP-DI apurado pela FGV, respectivamente.

Tendo em vista os valores apurados para o Erro Quadrático Médio (EQM) e o Erro Percentual Médio Absoluto (EPMA), os modelos que tiveram a tendência aproximada por

---

<sup>2</sup>Já inflacionadas pelo IGP-DI previsto pela SPE, após a aplicação do operador logaritmo e da inserção das tendências, nessa ordem.

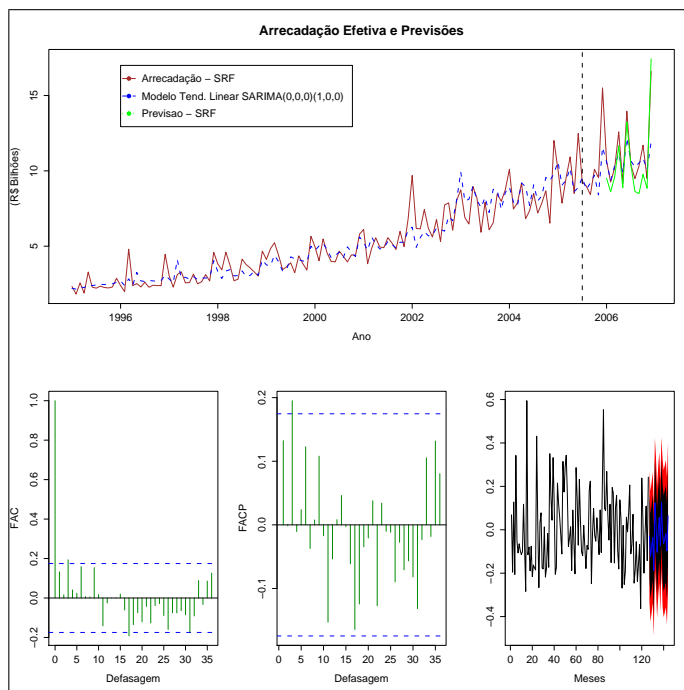


Figura 4.5: Modelo SARIMA  $(0,0,0)(1,0,0)$  c/ Tendência Linear

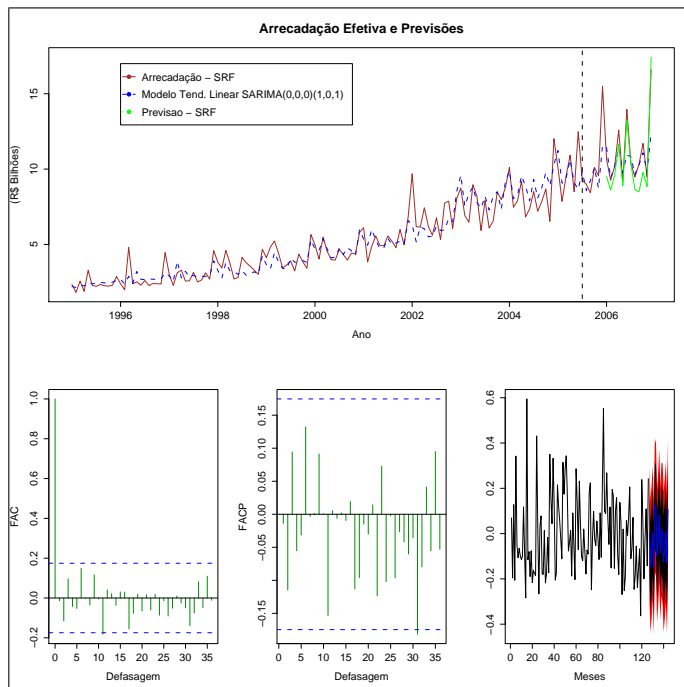


Figura 4.6: Modelo SARIMA  $(0,0,0)(1,0,1)$  c/ Tendência Linear

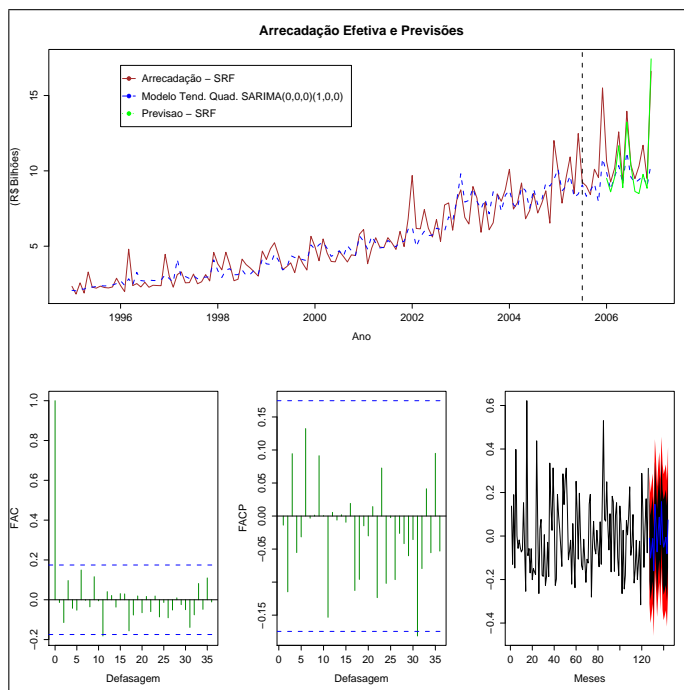


Figura 4.7: Modelo SARIMA  $(0,0,0)(1,0,0)$  c/ Tendência Quadrática

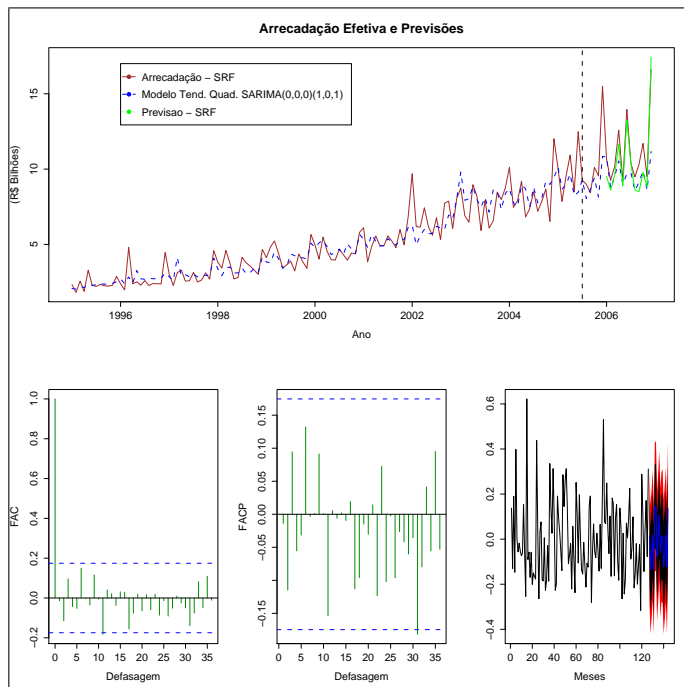


Figura 4.8: Modelo SARIMA  $(0,0,0)(1,0,1)$  c/ Tendência Quadrática

uma curva quadrática se revelaram menos adequados (com erros maiores) que os de tendência linear. Também apresentaram uma diferença maior com relação à soma da arrecadação anual. Como já previsto quando dos cálculos do AIC e BIC, o modelo cujas previsões se mostraram mais acertadas foi o SARIMA(0,0,0)(1,0,0). Com um erro no somatório da arrecadação de -4.51% e um Erro Percentual Médio Absoluto de 7,44%, o resultado obtido com esse modelo é um pouco melhor do que da previsão realizada pela SPE.

Uma segunda observação é que quando as séries com tendência linear são inflacionadas com o índice previsto pela SPE, os resultados se mostram melhores que aqueles obtidos quando o índice utilizado é o mensurado pela FGV. Como os índices previstos pela SPE são maiores que os da FGV, uma conclusão possível é que o método subestimou a arrecadação de 2006, subestimação essa que foi parcialmente corrigida com a aplicação de um índice maior para o inflacionamento. Em outras palavras, um erro corrigindo o outro.

Ao contrário do método de suavização exponencial, Capítulo 3, cuja utilização redundou em uma sobre-estimativa da arrecadação do Imposto de Renda, como pode ser visto nos Quadros 3.2 e 3.3, os modelos Box&Jenkins utilizados nesse capítulo subestimaram a arrecadação real, apresentando um caráter conservador. Considerando que os modelos de suavização exponencial superestimaram a arrecadação, uma possível experiência seria fazer uma combinação dos dois modelos, de molde a verificar se as previsões se tornam mais acuradas. Pode ser uma boa idéia para trabalhos futuros.

Quadro 4.3: Metodologia de Box-Jenkins - Série Inflacionada com Previsões da SPE

Mês	Valores Arrecadados	Tendência Linear			Tendência Quadrática				
		SARIMA(0,0,0)(1,0,0)		SARIMA(0,0,0)(1,0,1)		SARIMA(0,0,0)(1,0,1)			
		Valores Previstos	$\Delta(\%)$	Valores Previstos	$\Delta(\%)$	Valores Previstos	$\Delta(\%)$		
jan/06	10.643.316.250	10.592.796.016	-0.47	11.425.074.866	7.34	9.907.930.760	-6.91	10.808.155.142	1.55
fev/06	9.259.073.435	9.349.197.302	0.97	9.412.143.259	1.65	8.754.566.319	-5.45	8.726.557.456	-5.75
mar/06	10.189.952.436	10.345.321.467	1.53	10.200.484.437	0.10	9.643.804.840	-5.36	9.504.667.859	-6.72
abr/06	12.588.188.574	11.147.014.826	-11.45	11.400.284.431	-9.44	10.351.223.640	-17.77	10.539.307.036	-16.28
maí/06	9.180.874.262	9.880.284.992	7.62	9.754.850.429	6.25	9.183.957.547	0.03	8.983.666.995	-2.15
jun/06	13.956.711.467	12.174.654.209	-12.77	10.935.305.937	-21.65	11.235.475.315	-19.50	9.694.772.202	-30.54
jul/06	10.397.486.714	10.666.369.000	2.59	10.826.554.322	4.13	9.592.074.481	-7.75	9.766.707.791	-6.07
ago/06	9.485.565.072	10.248.422.589	8.04	9.654.964.528	1.79	9.207.608.963	-2.93	8.566.032.576	-9.69
set/06	10.312.573.908	10.524.807.176	2.06	10.266.576.224	-0.45	9.414.299.327	-8.71	9.138.050.398	-11.39
out/06	11.703.111.572	10.846.235.646	-7.32	11.081.710.512	-5.31	9.657.228.483	-17.48	9.872.948.645	-15.64
nov/06	9.492.285.018	10.086.468.722	6.26	9.779.347.460	3.02	8.984.586.762	-5.34	8.620.679.207	-9.18
dez/06	16.606.948.120	11.923.035.932	-28.20	12.446.346.505	-25.05	10.500.325.223	-36.77	11.169.948.503	-32.74
Total	133.816.086.835	127.784.607.885	-4.51	127.183.642.915	-4.96	116.433.081.666	-12.99	115.391.493.816	-13.77
EQM	-	5,49 10 <sup>18</sup>		6,12 10 <sup>18</sup>		2,99 10 <sup>19</sup>		33,25 10 <sup>19</sup>	
EPMA	-	7,44		7,18		11,16		12,31	

Quadro 4.4: Metodologia de Box-Jenkins - Série Inflacionada com IGP-DI Mensurado pela FGV

Mês	Valores Arrecadados	Tendência Linear			Tendência Quadrática				
		SARIMA(0,0,0)(1,0,0)		SARIMA(0,0,0)(1,0,1)		SARIMA(0,0,0)(1,0,1)			
		Valores Previstos	$\Delta(\%)$	Valores Previstos	$\Delta(\%)$	Valores Previstos	$\Delta(\%)$		
jan/06	10.643.316.250	10.341.243.343	-2.84	11.153.757.632	4.79	9.672.641.941	-9.12	10.551.488.223	-0.86
fev/06	9.259.073.435	9.076.535.331	-1.97	9.137.645.519	-1.31	8.499.246.292	-8.21	8.472.054.286	-8.50
mar/06	10.189.952.436	9.955.592.029	-2.30	9.816.211.307	-3.67	9.280.502.970	-8.93	9.146.607.564	-10.24
abr/06	12.588.188.574	10.672.976.256	-15.21	10.915.475.303	-13.29	9.911.026.930	-21.27	10.091.111.881	-19.84
maí/06	9.180.874.262	9.431.470.394	2.73	9.311.733.730	1.43	8.766.773.810	-4.51	8.575.581.510	-6.59
jun/06	13.956.711.467	11.629.214.456	-16.68	10.445.390.539	-25.16	10.732.111.953	-23.10	9.260.434.269	-33.65
jul/06	10.397.486.714	10.152.028.755	-2.36	10.304.489.822	-0.89	9.129.537.517	-12.19	9.295.749.878	-10.60
ago/06	9.485.565.072	9.734.784.842	2.63	9.171.070.134	-3.32	8.746.135.455	-7.79	8.136.714.050	-14.22
set/06	10.312.573.908	9.999.228.803	-3.04	9.753.893.156	-5.42	8.944.176.498	-13.27	8.681.722.639	-15.81
out/06	11.703.111.572	10.361.793.160	-11.46	10.586.750.641	-9.54	9.225.892.494	-21.17	9.431.977.607	-19.41
nov/06	9.492.285.018	9.654.197.559	1.716	9.360.238.451	-1.39	8.599.538.449	-9.41	8.251.226.713	-13.07
dez/06	16.606.948.120	11.414.716.634	-31.27	11.915.716.709	-28.25	10.052.660.888	-39.47	10.693.735.866	-35.61
Total	133.816.086.835	122.423.781.569	-8.51	121.872.372.948	-8.93	111.560.245.202	-16.63	110.588.404.492	-17.36
EQM	-	1,40 10 <sup>19</sup>		1,15 10 <sup>19</sup>		4,74 10 <sup>19</sup>		5,14 10 <sup>19</sup>	
EPMA	-	7,85		8,21		14,87		15,70	



## Parte II

# Análise da Série Reduzida não Deflacionada

# Capítulo 5

## Suavização Exponencial

Nesta segunda parte do trabalho, englobando este capítulo e o seguinte, será feita a análise da série composta pelos valores originais de arrecadação do Imposto de Renda, sem deflacionamento/inflacionamento, para o período de janeiro de 1995 a junho de 2005. Os procedimentos que serão adotados nos dois capítulos não diferem muito daqueles utilizados nos Capítulos 3 e 4, nos quais a série *deflacionada/inflacionada* foi objeto de análise, primeiramente utilizando-se suavização exponencial e, em seguida, a metodologia desenvolvida por Box&Jenkins.

No primeiro gráfico da Figura 6.1 é apresentada a evolução da arrecadação em valores originais. No segundo gráfico da figura é apresentado o comportamento da série após a sua transformação utilizando-se o operador logaritmo. Note-se que o comportamento da série transformada aparenta ser mais regular, o que poderia em um primeiro momento fazer o analista optar apenas pela sua análise. De forma a verificar se isso é verdade, ambas as séries, em valor original e em logaritmo, serão analisadas.

Primeiramente serão utilizados modelos baseados na suavização exponencial. Os resultados das análises são apresentados nas seções que se seguem. No Quadro 5.1, na conclusão do capítulo, são apresentados os resultados das previsões (jan/2006 a dez/2006) para os modelos que melhor se adaptaram ao conjunto de dados existentes.

Foi utilizado o comando `ets` (*exponential smoothing state space modes*) do **R**, que permite escolher em que tipo de modelo de suavização exponencial que se quer que os dados sejam amoldados. O comando realiza automaticamente um processo de minimização não linear do critério de informação adotado (AIC, AICC ou BIC), escolhendo os parâmetros ( $\alpha$ ,  $\beta$  e  $\phi$ ) que serão utilizados nas equações de previsão, conforme já visto no Capítulo 3. Não serão estudados modelos com amortecimento, uma vez que os gráficos das séries não apontam para qualquer comportamento amortecido da arrecadação do tributo.

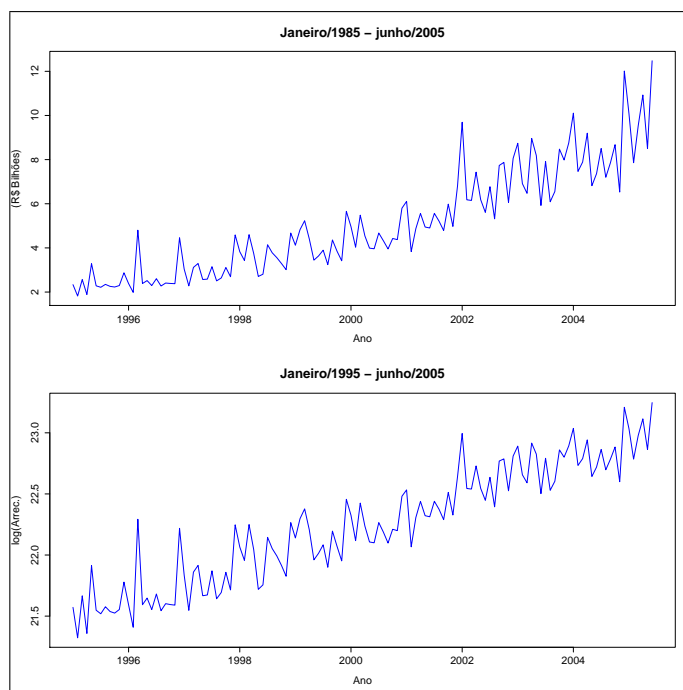


Figura 5.1: Evolução da Arrecadação não Deflacionada

## 5.1 Modelo Log - Tendência Aditiva e sem Sazonalidade (A-N)

No gráfico maior da Figura 5.2 são apresentadas as evoluções das séries de arrecadação realizada, do modelo ajustado e da previsão da SRF. Pode-se ver que, embora a série ajustada acompanhe, em média, a de arrecadação realizada, ela não consegue simular as subidas e descidas da série original. Está faltando algo para o modelo.

Os dois gráficos dos correlogramas das Funções de Autocorrelação e de Autocorrelação Parcial dos resíduos, apresentados na parte inferior da figura, apontam para o elemento que está faltando. Há um claro comportamento sazonal da série, com período de doze meses, que não está sendo captado pelo modelo, o que se reflete nos resíduos, que ainda apresentam informação que não foi utilizada. O Modelo A-N, portanto, não é adequado para o ajuste da série.

É óbvio que já sabíamos que a série tem comportamento sazonal, uma vez que já fizemos esses testes nos capítulos anteriores. O objetivo, porém, é estudar cada método por inteiro, procurando não herdar informações dos demais capítulos. O Teste de Box-Pierce para a verificação do grau de adequação dos resíduos forneceu o valor do  $\chi^2 = 82.62$ , para

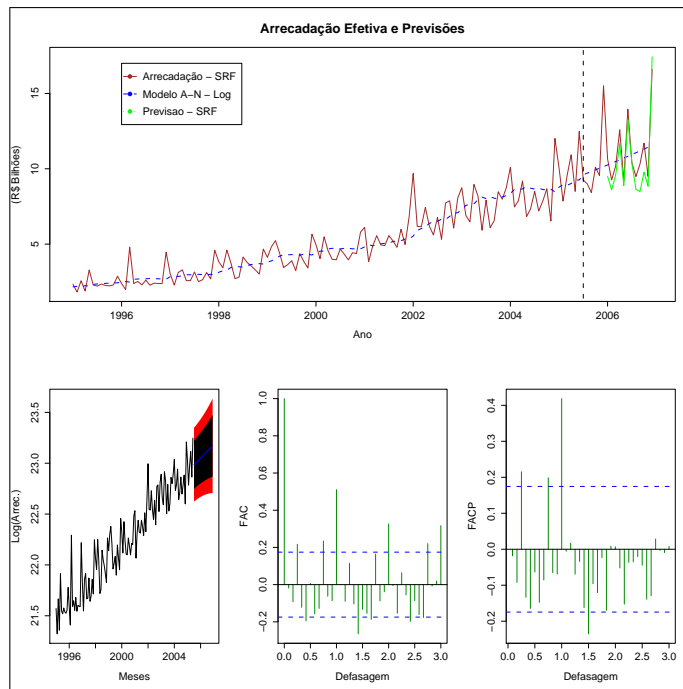


Figura 5.2: Suavização Exponencial Modelo A-N Log

vinete graus de liberdade, com p-valor de  $1.404 \cdot 10^{-9}$ . O valor da estatística do teste está muito alto, outro indicador de que o modelo poderia ser descartado.

Como foi identificado um comportamento sazonal da série, apenas modelos de suavização exponencial que possam simular esse comportamento serão estudados doravante.

## 5.2 Modelo Log - Tendência Aditiva e Sazonalidade Multiplicativa (A-M)

As séries e os correlogramas das FAC e das FACP dos resíduos podem ser vistos na Figura 5.3. Os correlogramas das funções de autocorrelação dos resíduos indicam que o comportamento sazonal foi bem modelado. O valor calculado para o Teste Box-Pierce foi de  $\chi^2 = 18.61$ , com vinte graus de liberdade e p-valor de 0.55, o que também é um resultado bastante superior ao modelo da Seção 5.1.

Comparando-se visualmente os intervalos de previsão de 80% e 95% para a série logaritmizada, apresentada no primeiro gráfico na parte inferior à esquerda de cada figura, vê-se que os intervalos para esse modelo são menores que para o modelo anterior, o que é um outro indicador da superioridade desse modelo ajustado.

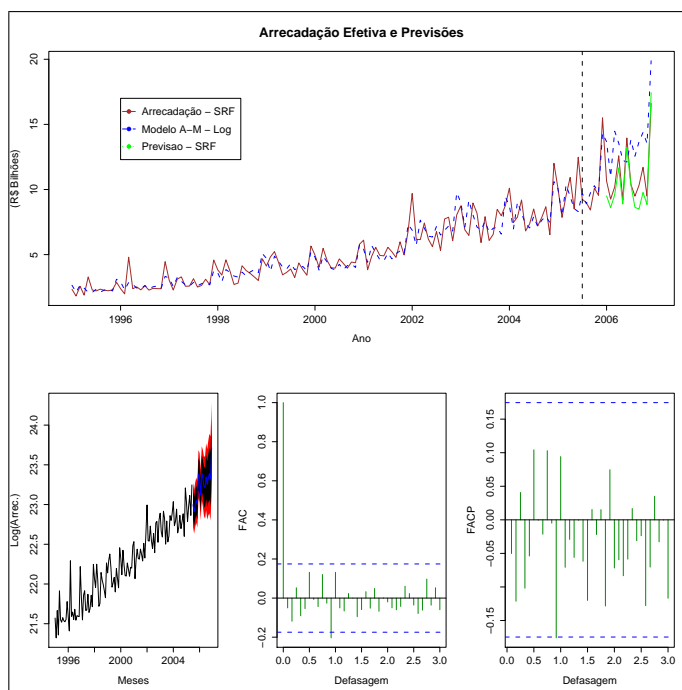


Figura 5.3: Suavização Exponencial Modelo A-M Log

Apesar de um ajuste melhor, pode-se ver, a partir dos gráficos das evoluções das séries, que esse modelo claramente superestima a arrecadação para 2006. A soma dos valores previstos para o ano atinge R\$ 164,85 bilhões, um valor 23,19% superior ao efetivamente arrecadado, R\$ 133,82 bilhões.

### 5.3 Modelo Log - Tendência Aditiva e Sazonalidade Aditiva (A-A)

Os correlogramas das funções de autocorrelação da série dos resíduos indicam que o modelo ajustado captou bem as informações da série original, conforme pode ser visto nos gráficos da Figura 5.4. O resultado para a estatística do Teste Box-Pierce é semelhante ao do modelo anterior, com valor do  $\chi^2$  de 18.15, e p-valor de 0.58.

É outro modelo que superestima consideravelmente a arrecadação (em 29,35%), como pode ser visto no gráfico. É possível que ao considerarmos a tendência como aditiva, estejamos levando a essas superestimativas.

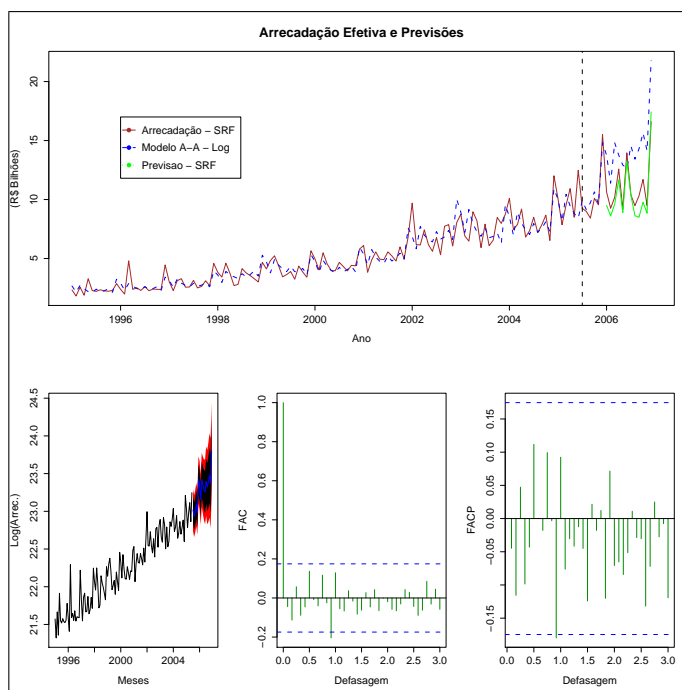


Figura 5.4: Suavização Exponencial Modelo A-A Log

## 5.4 Modelo Log - Tendência Multiplicativa e Sazonalidade Multiplicativa (M-M)

Quarto e último dos modelos de suavização exponencial estudados para a série em logaritmo, foi também o que melhor resultados obteve. O valor calculado para o Teste Box-Pierce foi de 16.73, com o p-valor de 0.67. No Quadro 5.1 são apresentados os valores para as previsões de arrecadação do Imposto de Renda para 2006 utilizando-se este modelo. Os resultados obtidos estão muito distantes dos realizados, com uma superestimativa de 24,37% com relação ao valor arrecadado.

Aparentemente, aplicar o operador logaritmo à série não deflacionada, antes de utilizar o método de suavização exponencial, não pode ser considerada uma boa estratégia para o cálculo de previsões. Esse fato será corroborado a seguir, onde resultados obtidos com os mesmos modelos, mas para séries sem a aplicação do logaritmo, apresentaram resultados superiores para as previsões.

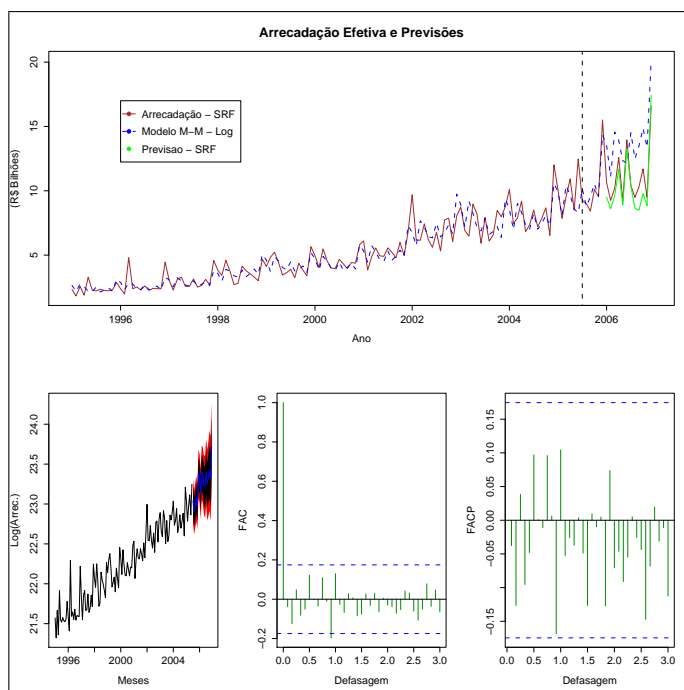


Figura 5.5: Suavização Exponencial Modelo M-M Log

## 5.5 Modelo /s Log - Tendência Aditiva e Sazonalidade Multiplicativa (A-M)

Vamos estudar agora o comportamento da série quando modelada sem a aplicação prévia do operador logarítmico. Como já se sabe que ela apresenta um comportamento sazonal, só serão utilizados modelos que incorporem essa característica às equações. O primeiro modelo tentado é aquele com tendência aditiva e sazonalidade multiplicativa, Modelo A-M.

A Figura 5.6 apresenta o gráfico da série, bem como os correlogramas das FAC e das FACP dos resíduos. Ambos os correlogramas indicam que o comportamento dos resíduos se encontra próximo daquele do ruído branco. Realizado o Teste Box-Pierce, encontra-se o valor para o  $\chi^2$  de 14.12, com p-valor de 0.82. Esse resultado foi melhor que para todos os modelos de suavização exponencial quando aplicado o operador logaritmo natural.

Analisando a evolução das séries no gráfico maior da figura, vê-se que a série do modelo ajustado se adapta bem aos dados reais e as previsões também. Os valores calculados para a previsão estão no Quadro 5.1 e indicam uma superestimativa de 6,33% frente aos valores efetivamente arrecadados.

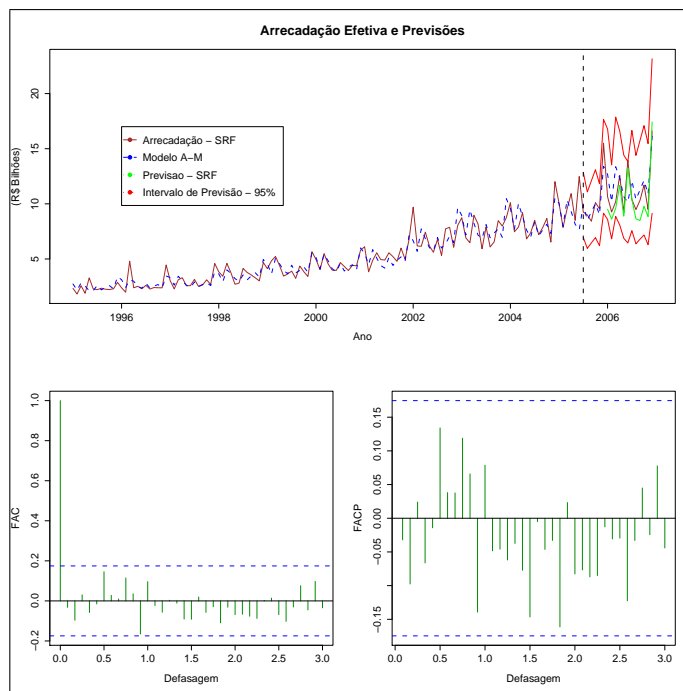


Figura 5.6: Suavização Exponencial Modelo A-M Sem Log

Como não houve nenhuma transformação nos valores da série, podemos utilizar diretamente os intervalos de previsão calculados pelo comando `forecast` para o modelo, representados pelas linhas vermelhas no gráfico. Os valores previstos pela SRF e os valores arrecadados estão dentro do intervalo de 95%, não se encontrando nenhum *outlier*.

## 5.6 Modelo $s/$ Log - Tendência Multiplicativa e Sazonalidade Multiplicativa (M-M)

Esse é o último modelo analisado nesse capítulo. A Figura 5.7 apresenta o gráfico das séries, bem como os correlogramas das FAC e das FACP dos resíduos. Os resultados encontrados se assemelham bastante ao do modelo A-M, com o valor obtido para a estatística do Teste Box-Pierce,  $\chi^2 = 14.42$ , com p-valor de 0.81. Os correlogramas indicam uma boa captura das características de sazonalidade da série.

Apesar desses bons resultados, quando se calculam as previsões, verifica-se que o erro do somatório da arrecadação anual chega a 10,51%, com um EPMA de 16,39. O erro quadrático médio foi de  $4,01 \cdot 10^{18}$ . Esses resultados são bem piores que os obtidos no modelo A-M, apresentados na Seção 5.5 e por esse motivo não serão incluídos no Quadro 5.1. Apesar disso, ainda são melhores que os do modelo M-M  $c/$  Log.



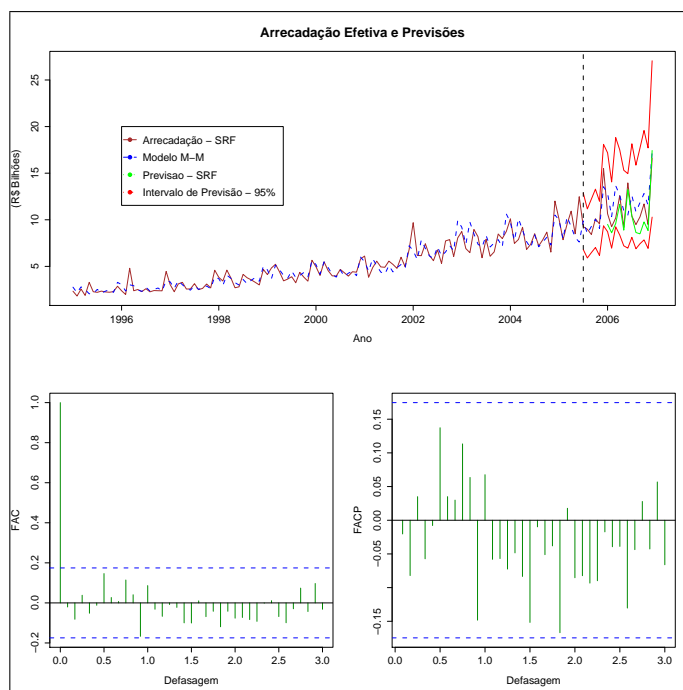


Figura 5.7: Suavização Exponencial Modelo M-M Sem Log

## 5.7 Conclusão

Com base nas análises realizadas nesse capítulo, verificou-se que a aplicação do operador logaritmo à série, em que pese tê-la tornado mais regular, não foi uma boa estratégia. Todos os modelos que tiveram por base a série logaritmizada apresentaram uma capacidade de ajuste à série real e de previsão inferiores aos que utilizaram a série original.

Dos modelos sem log, destacou-se o A-M, com o menor valor para a estatística do Teste Box-Pierce. Dos modelos com log, o que apresentou os melhores resultados foi o M-M. No Quadro 5.1 são apresentados os valores da arrecadação do Imposto de Renda no ano de 2006, bem como os resultados das previsões para os dois modelos citados. Com base nos dois critérios de erro utilizados (EPMA e EQM), vê-se que o Modelo A-M sem a aplicação do operador logaritmo é bastante superior ao melhor dos modelos com aplicação do logaritmo.

Apesar disso, a qualidade das previsões realizadas com esse modelo se demonstrou, no geral, inferior à obtida com os modelos de suavização exponencial inflacionados/deflacionados pelas previsões do IGP-DI, conforme apresentado no Capítulo 3. Também a qualidade das previsões realizadas com a metodologia de Box&Jenkins aplicada à série inflacionada se revelou superior.

Quadro 5.1: Suavização Exponencial s/ Deflacionamento - Valores Previstos e Arrecadados 2006

Mês	Valores Arrecadados	Valores Previstos (A-M s/ Log)	$\Delta(\%)$	Valores Previstos (M-M Log)	$\Delta(\%)$
jan/06	10.643.316.251	12.684.825.143	19,18	13.534.453.327	27,16
fev/06	9.259.073.435	10.173.009.141	9,87	11.104.867.593	19,93
mar/06	10.189.952.438	13.367.882.022	31,19	14.577.922.433	43,06
abr/06	12.588.188.576	12.338.280.642	-1,99	13.897.461.533	10,40
mai/06	9.180.874.264	10.649.541.456	16,00	12.338.409.026	34,39
jun/06	13.956.711.468	10.145.830.796	-27,31	12.151.997.646	-12,93
jul/06	10.397.486.714	12.113.375.768	16,50	14.514.766.460	39,59
ago/06	9.485.565.072	10.382.357.773	9,45	12.483.622.333	31,60
set/06	10.312.573.909	11.269.675.953	9,28	13.480.928.354	30,72
out/06	11.703.111.573	12.132.547.147	3,67	14.865.782.582	27,02
nov/06	9.492.285.019	10.880.130.886	14,62	13.375.973.208	40,91
dez/06	16.606.948.121	16.147.593.175	-2,77	20.105.812.401	21,06
Total	133.816.086.841	142.285.049.902	6,33	166.432.000.000	24,37
EQM	-	$3,24 \cdot 10^{18}$		$1,37 \cdot 10^{19}$	
EPMA	-	13,49		32,57	

Uma outra característica dos resultados obtidos que vale a pena mencionar é que houve uma superestimativa dos valores arrecadados em praticamente todos os meses, para os três modelos. Esse comportamento é semelhante ao encontrado no Capítulo 3 para a série deflacionada.

# Capítulo 6

## Metodologia Box&Jenkins

Nesse capítulo será feita a análise da série de arrecadação do Imposto de Renda sem deflacionamento utilizando-se a metodologia desenvolvida por Box&Jenkins. Os procedimentos utilizados não diferem substantivamente daqueles adotados no Capítulo 4. Ao final do capítulo, após serem deduzidos os modelos que melhor lograram simular o comportamento da série original, são feitas previsões para o montante mensal de arrecadação do tributo para 2006.

No primeiro gráfico da Figura 6.1 é mostrada a evolução da série original. Não há dúvidas acerca de sua não estacionariedade, nem em relação à média, nem à variância. De forma a tentar estacionarizar a variância, é aplicado o operador logaritmo natural a todos os valores da série. O resultado obtido da operação pode ser visto no segundo gráfico da figura, tendo aparentemente conseguido tornar a variância estacionária. Como a aplicação da metodologia pressupõe a estacionariedade da variância, *apenas a série à qual foi aplicada o operador logaritmo será objeto de análise doravante.*

Com relação à média, procurar-se-á analisar qual o procedimento mais apropriado para torná-la estacionária: a aplicação do operador de diferença ( $L = 1 - B$ ), ou o cálculo da equação da curva de tendência, com a posterior retirada de seus efeitos para a realização das previsões. Em um primeiro momento, ao analisar-se a forma do gráfico, já se pensaria em optar pela utilização da linha de tendência, mas alguns testes estatísticos devem ser feitos antes de se adotar qualquer um dos procedimentos.

Na Figura 6.2 são apresentados os correlogramas das FAC e das FACP da série em logaritmo. Como esperado, o comportamento de ambos os correlogramas aponta para a sua não estacionariedade. Realizando-se o Teste de Dickey-Fuller, verifica-se que a mesma **não** apresenta raiz unitária<sup>1</sup>, favorecendo então a primeira opinião de que o comportamento não

---

<sup>1</sup>O resultado obtido para o teste foi de  $-5.93$ , com p-valor de  $0.01$ .

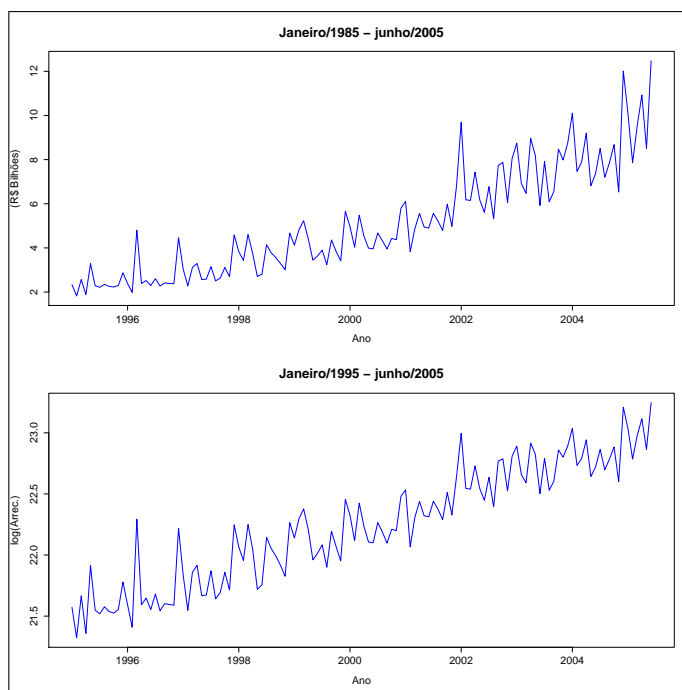


Figura 6.1: Evolução da Série Original e em Logaritmo

Quadro 6.1: Parâmetros das Equações de Regressão

Parâmetro	Valor	Estat. t	Prob. (t)
$\alpha$	21.47	683.92	$< 2 \cdot 10^{-16}$
$\beta$	$1.22 \cdot 10^{-02}$	28.41	$< 2 \cdot 10^{-16}$

estacionário da média da série se deve a um efeito de tendência.

Há que se identificar, então, qual a melhor equação que possa representar essa tendência. Da observação visual do gráfico da série, entendeu-se que a equação da reta<sup>2</sup> seria a mais apropriada. Faz-se, então, uma regressão linear em relação ao tempo, obtendo-se os parâmetros da equação apresentados no Quadro 6.1. O valor calculado para  $R^2$  é 0.8668 e a estatística F é 807.3 com p-valor  $< 2.2 \cdot 10^{-16}$ . Todos esses valores indicam que a reta calculada ajusta de forma bastante satisfatória o comportamento linear dos pontos da série.

Subtraindo-se da série o efeito da tendência e calculando-se os valores das FAC e das FACP, obtêm-se o seu novo gráfico, bem como os correlogramas, os quais são apresentados

<sup>2</sup>Equação da reta:  $y = \alpha + \beta t$ .

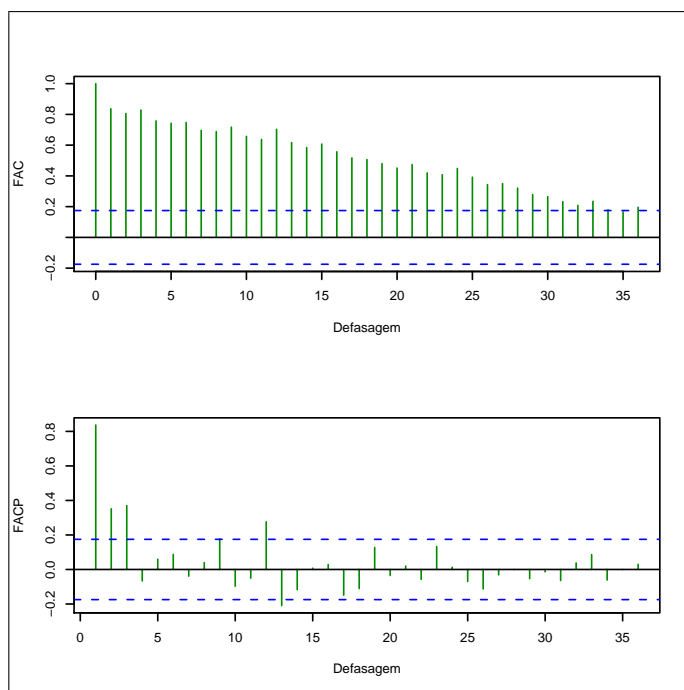


Figura 6.2: FAC e FACP da Série não Deflacionada

no Figura 6.3. Todos os elementos da figura apontam para um comportamento sazonal estacionário da série, com período de doze meses.

Essa série transformada será utilizada então para a escolha dos coeficientes dos modelos SARIMA mais adequados para a previsão. Será adotado o mesmo procedimento utilizado no Capítulo 4: variam-se os parâmetros da equação  $(p,d,q,P,D,Q)$  em um processo iterativo, calculam-se os valores dos critérios de informação, AIC e BIC, adotam-se então os modelos que minimizaram estes critérios. É possível que técnicos com grande experiência em análise de séries temporais, com a simples observação do comportamento dos correlogramas dos resíduos, já consigam inferir quais os modelos são possíveis. Dado que não temos toda essa experiência, preferimos verificar numericamente quais os modelos mais adequados.

Os resultados obtidos para as análises são apresentados no Quadro 6.2. Ambos os critérios de informação são concordantes no sentido de que o modelo SARIMA $(0,0,0)(1,0,1)$  é o mais adequado para representar a série. Como segundo melhor, cada um aponta para um modelo diverso. Como esperado, o AIC é menos parcimonioso que o BIC, apontando para um modelo<sup>3</sup> com três parâmetros a serem determinados. Para as previsões serão utilizados os modelos SARIMA  $(0,0,0)(1,0,1)$ ,  $(0,0,0)(1,0,2)$  e  $(0,0,0)(1,0,0)$ .

<sup>3</sup>SARIMA $(0,0,0)(1,0,2)$ .

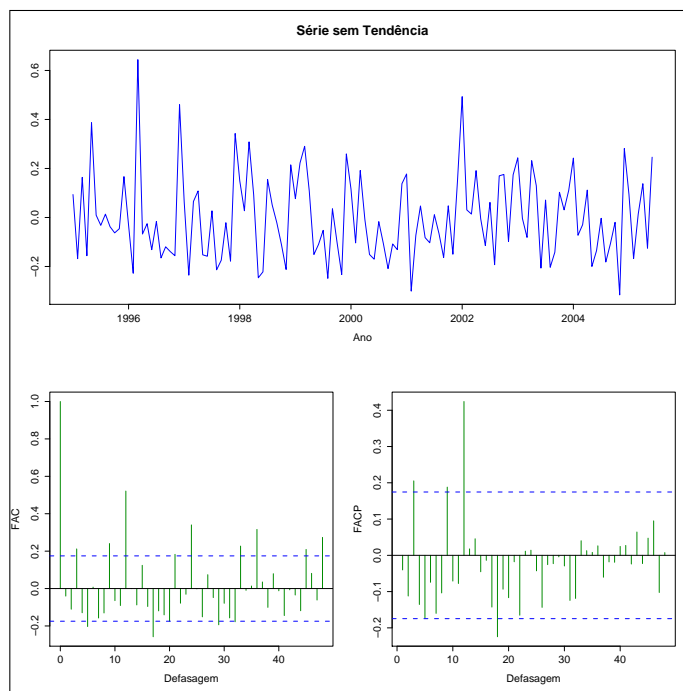


Figura 6.3: Série não Deflacionada sem Tendência

Quadro 6.2: Melhores Modelos - AIC/BIC

AIC	Hierarquia	Modelos
-124,94	1 <sup>o</sup>	SARIMA(0,0,0)(1,0,1)
-124,33	2 <sup>o</sup>	SARIMA(0,0,0)(1,0,2)
-123,92	3 <sup>o</sup>	SARIMA(0,0,0)(2,0,1)
BIC	Hierarquia	Modelos
-123,28	1 <sup>o</sup>	SARIMA(0,0,0)(1,0,1)
-123,17	2 <sup>o</sup>	SARIMA(0,0,0)(1,0,0)
-120,23	3 <sup>o</sup>	SARIMA(0,0,0)(2,0,0)

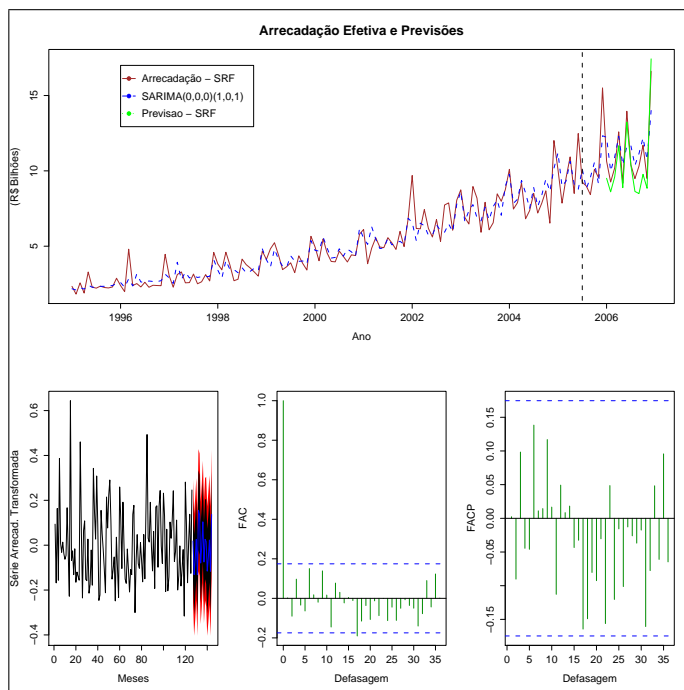


Figura 6.4: Box&Jenkins - SARIMA(0,0,0)(1,0,1)

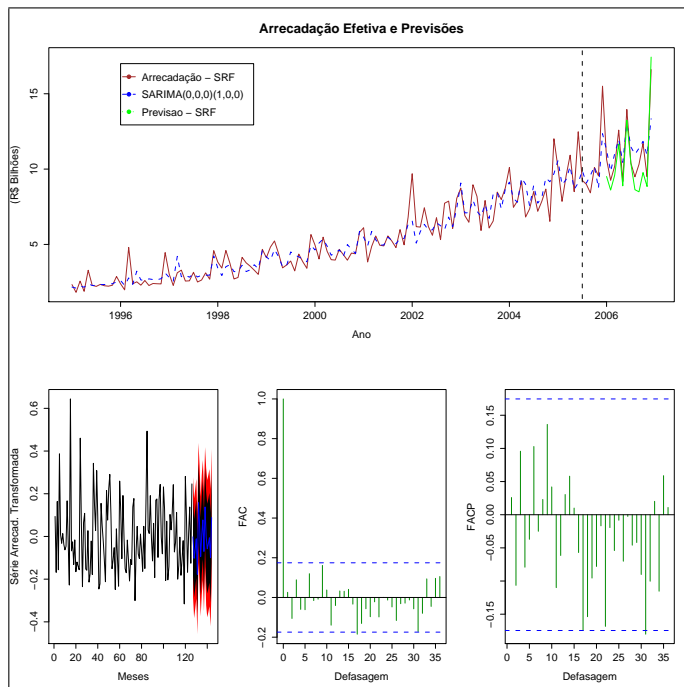


Figura 6.5: Box&Jenkins - SARIMA(0,0,0)(1,0,0)

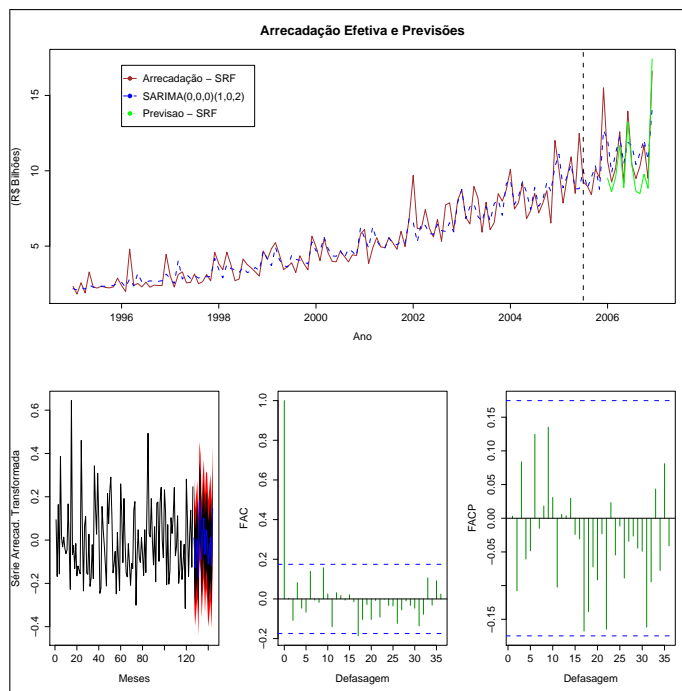


Figura 6.6: Box&Jenkins - SARIMA(0,0,0)(1,0,2)

As figuras 6.4, 6.5 e 6.6 apresentam as séries de arrecadação, as séries com as previsões, bem como os correlogramas das funções de autocorrelação e de autocorrelação parcial das séries dos resíduos. Todos os três modelos aparentam se ajustar à série de arrecadação muito bem. Os resultados para a estatística do Teste de Box-Pierce para as três séries de resíduos são, para o primeiro modelo, 19,72, com p-valor de 0,48, para o segundo, 20,07, com p-valor de 0,45 e para o terceiro, 18,68, com p-valor de 0,54.

No Quadro 6.3 são apresentados os resultados para as previsões com os três modelos utilizados. O que apresentou o melhor resultado, adotando os critérios EQM e EPMA, foi o SARIMA(0,0,0)(1,0,1), que foi exatamente o modelo que minimizou ambos os critérios de informação (AIC e BIC). O segundo modelo, SARIMA(0,0,0)(1,0,0), embora um pouco pior<sup>4</sup>, ainda apresenta erros de ordem de grandeza semelhante ao do primeiro. Já os resultados para o terceiro modelo se revelaram bem piores que para os dois primeiros.

## 6.1 Conclusão

Das quatro metodologias apresentadas em cada um dos capítulos do presente trabalho, a de Box&Jenkins, aplicada à série sem deflacionamento, foi a que melhor conseguiu

<sup>4</sup>Segundo os critérios adotados.



realizar as previsões para a arrecadação de Imposto de Renda de 2006. Os três modelos apresentados neste capítulo lograram calcular o montante de arrecadação anual com um erro, em média, de 2,87%; um erro absoluto inferior a todos os demais modelos, sendo inferior também àquele apresentado pela RFB em sua previsão na Lei Orçamentária Anual. Assim como nos capítulos 3, para a série deflacionada, e 5, para a série não deflacionada, houve uma superestimativa dos valores arrecadados em praticamente todos os meses, para os três modelos.

Quadro 6.3: Previsões para Arrecadação do Imposto de Renda Sem Deflacionamento - Box&amp;Jenkins

Mês	Valores Arrecadados	SARIMA(0,0,0)(1,0,1)		SARIMA(0,0,0)(1,0,0)		SARIMA(0,0,0)(1,0,2)	
		Valores Previsitos	$\Delta(\%)$	Valores Previsitos	$\Delta(\%)$	Valores Previsitos	$\Delta(\%)$
jan/06	10.643.316.251	12.227.847.819	14,89	11.253.507.012	5,73	12.070.611.383	13,41
fev/06	9.259.073.435	9.980.289.534	7,79	9.836.698.821	6,24	9.870.040.462	6,60
mar/06	10.189.952.438	10.914.842.356	7,11	11.026.609.919	8,21	11.188.799.426	9,80
abr/06	12.588.188.576	12.288.447.250	-2,38	11.964.504.447	-4,95	12.259.650.586	-2,61
mai/06	9.180.874.264	10.399.142.777	13,27	10.447.443.519	13,80	10.488.597.589	14,24
jun/06	13.956.711.468	11.648.902.374	-16,54	13.028.055.043	-6,65	11.892.746.375	-14,79
jul/06	10.397.486.714	11.677.424.997	12,31	11.474.533.630	10,36	11.669.382.660	12,23
ago/06	9.485.565.072	10.359.056.865	9,21	10.979.052.347	15,74	10.415.571.956	9,80
set/06	10.312.573.909	11.145.304.322	8,07	11.379.507.772	10,35	11.165.792.138	8,27
out/06	11.703.111.573	12.190.824.881	4,17	11.840.251.117	1,17	12.052.923.228	2,99
nov/06	9.492.285.019	10.737.579.270	13,12	10.916.488.450	15,00	10.720.122.224	12,94
dez/06	16.606.948.121	14.006.380.014	-15,66	13.339.933.692	-19,67	14.135.341.427	-14,88
Soma	133.816.086.841	137.576.042.458	2,81	137.486.585.768	2,74	137.929.579.453	3,07
EQM	-	1,40 10 <sup>19</sup>		1,52 10 <sup>19</sup>		4,74 10 <sup>19</sup>	
EPMA	-	7,85		8,20		14,87	

# Capítulo 7

## Conclusão

O objetivo do presente trabalho foi estudar diferentes formas de realizar a previsão de arrecadação do Imposto de Renda agregado, em valores brutos, sem retificações, utilizando duas metodologias ligadas a séries temporais, já consagradas na literatura técnica: a Suavização Exponencial e os modelos Box&Jenkins. Para cada uma dessas metodologias, foi estudado o comportamento das previsões com base nas séries em valores originais e deflacionados pelo IGP-DI, calculado pela Fundação Getúlio Vargas, do Rio de Janeiro.

Buscou-se avaliar, também, qual a performance desses modelos, quando cotejados com a previsão de arrecadação realizada pela Secretaria da Receita Federal do Brasil-RFB para o ano de 2006.

Das metodologias estudadas, e com base nos parâmetros utilizados para comparação, verificou-se que aquela que apresentou modelos com melhor capacidade de previsão da arrecadação em 2006 foi a de Box&Jenkins, para a série original, sem deflacionamento. Deve-se acrescentar, também, que as quatro metodologias apresentaram modelos cuja capacidade de previsão se revelou superior à apresentada pelo modelo da RFB para o ano considerado.

Esses resultados, no entanto, devem ser vistos com cautela, uma vez que os cálculos foram feitos para apenas um ano. Uma afirmação mais assertiva acerca da pertinência da troca da metodologia atualmente utilizada pela Secretaria só seria possível após um estudo envolvendo diversos anos de arrecadação e previsão. Se, após um estudo como esse, fosse verificado que os métodos aqui apresentados se comportam melhor que o da RFB de forma contumaz, aí sim, alguma proposta nesse sentido poderia ser realizada.

Um outro aspecto que deve ser levado em conta, como foi comentado na Introdução, é que o órgão arrecadador tem por obrigação natural ser conservador em suas previsões.

Considerando as pressões eternas por mais e mais gastos, é sempre melhor prever menos do que realmente será arrecadado. Dos métodos estudados, porém, apenas um, Box&Jenkins para a série deflacionada, subestimou a arrecadação, enquanto os demais a superestimaram. Talvez a RFB não possa correr esse risco.

Por último, deve-se acrescentar que há inúmeros outros estudos que podem ser realizados com base nas séries de arrecadação disponíveis. Ao invés de estudar a série do Imposto de Renda agregado, seria possível estudar a evolução de cada um de seus componentes e ver se as previsões iriam melhorar. Uma outra possibilidade seria estudar a combinação dos diversos modelos estudados e ver como as previsões se comportam. São sugestões que ficam para próximos trabalhos.

# Referências Bibliográficas

- Billah, Baki, King, Maxell L., Snyder, Ralph D., Koehler, Anne B. *Exponential Smoothing Model Selection for Forecasting*. Department of Econometrics and Business Statistics, Monash University, Australia, 2005.
- Bowerman, B. L. e O'Connell, R. T. *Time Series Forecasting Unified Concepts and Computer Implementation*. 2 ed. USA: Duxbury Press. 1987. 540 p.
- Cordeiro, Nilton José Neves. *Modelos Univariados da Previsão da Dinâmica Inflacionária Brasileira*. Dissertação de Mestrado em Estatística, Universidade Federal de Pernambuco, 2002, 183 p.
- Dunn, Peter. *Statistics for Climate Research - Study Book*. Faculty of Sciences - The University of Southern Queensland, 2005. Disponível no sítio [www.usq.edu.au](http://www.usq.edu.au), acessado em 10/12/2006.
- Enders, Walter. *Applied Econometric Time Series* 1 ed. USA: John Wiley & Sons, Inc. 1995. 433 p.
- Gardner, E. S., Jr. *Exponential smoothing: The state of the art - Part II*. Bauer College of Business, University of Houston, 2005. Disponível no sítio [www.bauer.uh.edu/gardner](http://www.bauer.uh.edu/gardner), acessado em 03/01/2007.
- Garrett, Thomas A. e Leatherman, John C. *An Introduction to State and Local Public Finance*. Regional Research Institute, West Virginia University, 2000. Disponível no sítio [www.rr.i.wvu.edu/WebBook/Garrett/chapterfour.htm](http://www.rr.i.wvu.edu/WebBook/Garrett/chapterfour.htm), acessado em 31/08/2006.
- Guaajardo, S.A. e Miranda, R. *An Elected Official's Guide to Revenue Forecasting*. Chicago: Government Finance Officers Association, 2000.
- Hyndman, Rob J., Koehler, Anne B., Snyder, Ralph D., Grose, Simone *A State Space Framework for Automatic Forecasting Using Exponential Smoothing Methods*. Department of Econometrics and Business Statistics, Monash University, Australia, 2000.

Kyobe, Annette e Danninger, Stephan. *Revenue Forecasting - How is it done? Results from a Survey of Low-Income Countries*. IMF Working Paper 05/24. International Monetary Fund. Washington, 2005.

Makridakis, Spyros, Wheelwright, Steven C., Hyndman, Rob J. *FORECASTING Methods and Applications* 3 ed. USA: John Wiley & Sons, Inc. 1998. 642 p.

Melo, Bruno Stephan Veras de. *Modelo de Previsão para Arrecadação Tributária* Brasília: ESAF, 2001 97 p. Monografia vencedora em primeiro lugar do VII Prêmio do Tesouro Nacional - 2001.

Penner, Rudolph G. *Errors in Budget Forecasting*. The Urban Institute, 2001.

Siqueira, Marcelo. *Melhorando a Previsão da Arrecadação Tributária Federal Através da Utilização de Modelos de Séries Temporais*. Brasília: ESAF, 2002 79 p. Monografia vencedora em primeiro lugar do VI Prêmio do Tesouro Nacional - 2002.