



Manuel Guillermo Trespalacios Pérez

CONTRIBUIÇÃO COMPUTACIONAL PARA O MÉTODO DE  
MÍNIMOS QUADRADOS RECURSIVO E APLICAÇÕES  
A MÚLTIPLAS SÉRIES TEMPORAIS.

Campinas

2012





Universidade Estadual de Campinas  
Faculdade de Engenharia Elétrica e de Computação.

Manuel Guillermo Trespalacios Pérez

Contribuição Computacional para o Método de Mínimos  
Quadrados Recursivo e Aplicações a Múltiplas Séries Temporais.

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Engenharia Elétrica. Área de concentração: Automação.

Orientador: Prof. Dr. Gilmar Barreto

Este exemplar corresponde à versão final da dissertação defendida pelo aluno Manuel Guillermo Trespalacios Pérez, e orientada pelo Prof. Dr. Gilmar Barreto

---

Campinas

2012

Ficha catalográfica  
Universidade Estadual de Campinas  
Biblioteca da Área de Engenharia e Arquitetura  
Rose Meire da Silva - CRB 8/5974

T725c Trespacios Pérez, Manuel Guillermo, 1983-  
Contribuição computacional para o método de mínimos quadrados recursivo e aplicações a múltiplas séries temporais. / Manuel Guillermo Trespacios Pérez. – Campinas, SP : [s.n.], 2012.

Orientador: Gilmar Barreto.  
Dissertação (mestrado) – Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Mínimos quadrados. 2. Análise de séries temporais. 3. Processamento dados. I. Barreto, Gilmar, 1958-. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

Informações para Biblioteca Digital

**Título em inglês:** Computational contribution to the recursive least squares method and applications to multiple time series.

**Palavras-chave em inglês:**

Least-squares

Time series analysis

Data processing

**Área de concentração:** Automação

**Titulação:** Mestre em Engenharia Elétrica

**Banca examinadora:**

Gilmar Barreto [Orientador]

Annabell Del Real Tamariz

João Bosco Ribeiro do Val

**Data de defesa:** 22-11-2012

**Programa de Pós-Graduação:** Engenharia Elétrica

## COMISSÃO JULGADORA - TESE DE MESTRADO

**Candidato:** Manuel Guillermo Trespacios Pérez

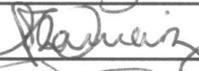
**Data da Defesa:** 22 de novembro de 2012

**Título da Tese:** "Contribuição Computacional para o Método de Mínimos Quadrados Recursivo e Aplicações a Múltiplas Séries Temporais"

Prof. Dr. Gilmar Barreto (Presidente):



Profa. Dra. Annabell Del Real Tamariz:



Prof. Dr. João Bosco Ribeiro do Val:





DEDICO MEU ESTUDO E TRABA-  
LHO FEITO NESTA TESE, A  
MINHA FAMILIA MEU ORIEN-  
TADOR E AMIGOS.



# Agradecimentos

Agradeço,

ao Prof. Gilmar Barreto por sua valorosa orientação.

à agência CAPES o apoio financeiro nos dois anos de mestrado

aos colegas do LCSI.

à minha família o apoio dia a dia concedido.

aos membros da banca examinadora.

ao Prof. Celso Bottura por me transmitir seu interesse pelo conhecimento.

à nosso amigo Murilo Trindade .

à meus amigos da Colômbia e Brasil.

a todos que de alguma forma contribuíram com o meu progresso como aluno e como Ser.



A maravilhosa disposição e harmonia do universo só pode ter tido origem segundo o plano de um Ser que tudo sabe e tudo pode. Isto fica sendo a minha última e mais elevada descoberta.

Isaac Newton



# Trabalhos publicados pelo autor

1. Trespalacios, M., Barreto, G., “Mínimos cuadrados recursivo en las exportaciones de cafe.” ISBN, vol. 1, p. 7, 2012.
2. Trespalacios, M., Barreto, G., “Mínimos cuadrados recursivo aplicado a las exportaciones de café en Colombia.” Symposium ISIT 2012, vol. 1, p. 6, 2012.



# Resumo

Esta tese apresenta uma contribuição computacional para o método de mínimos quadrados recursivo (MQR) e a sua aplicação à múltiplas séries temporais. O objetivo principal é obter estimativas com menor erro médio final, em comparação com as estimativas resultantes da aplicação do método MQR usual. Os resultados das análises foram obtidos aplicando diferentes modificações e ajustes no modelo de dados de entrada e saída. Estes ajustes procuram ter a quantidade suficiente de dados nas séries temporais de entrada que forneçam a melhor relação com a série temporal de saída. O método desenvolvido usa inicialmente, e de forma independente, o método MQR em duas matrizes geradas pelo modelo de dados de entrada e saída, obtendo assim dois estimadores do mesmo modelo; a primeira matriz é composta pela totalidade dos dados do modelo final e a segunda matriz é gerada por um filtro que utiliza a primeira matriz e a nova observação, filtro especificado no presente trabalho. Os dois estimadores calculados para cada nova observação são comparados aplicando-lhes critérios estabelecidos neste trabalho, gerando assim um terceiro estimador que é o melhor estimador do modelo de dados de entrada e saída. O método proposto é chamado de método de mínimos quadrados recursivo 3 (MQR3). Finalmente o método MQR3 é testado com um modelamento de dados de múltiplas séries temporais e seus resultados comparados com os resultados do método MQR. As estatísticas encontradas para diferentes dados de validação mostram estimativas com menor erro médio final para o método MQR3.

Palavras-chave: Mínimos quadrados recursivo, séries temporais, filtro de dados, previsões de dados.



# Abstract

This thesis presents a computational contribution to the recursive least squares method (RLS) and the application to multiple time series. The main objective is to obtain estimates with lower mean final error, when compared to the estimates resulting from the application of the normal MQR method. The analysis results were obtained by applying different modifications and adjustments to the input and output data model. These adjustments seek to have a sufficient amount of data in the incoming time series to provide the highest ratio with the output time series. The method initially uses, independently, the RLS method in two arrays generated by the input and output data model, thus obtaining two estimators of the same model, the first matrix is composed of all of the data for the final model and the second matrix is generated by a filter that uses the first array and the new observation, filter specified in the present work. The two estimators calculated for each new observation are compared by applying them with criteria established in this work, thereby generating a third which is the best estimator of the input and output model. The proposed method is called recursive least squares method 3 (MQR3). Finally, the method is tested with a MQR3 multiple time series data modeling and the results compared with the results of the RLS method. The statistics found for different validation data show estimates with a lower final average error for the MQR3 method.

Key-words: Recursive least-squares, time series, data filter, data predictions.



# Lista de Figuras

2.1	Série temporal de Reservas Internacionais da Colômbia em milhões de dolares. . .	4
2.2	Sistema Dinâmico. . . . .	5
2.3	Série 1 (Produção de café verde Colombiano, medida em milhares de sacos de café) e teste de Tendência. . . . .	10
2.4	Série 2 (Câmbio de Dólar a peso Colombiano.) e teste de Tendência. . . . .	11
2.5	Série 3 (Exportações de petróleo e derivados em milhões de dólares (FOB).) e teste de Tendência. . . . .	12
2.6	Série 4 (Exportações de café em milhões de dólares (FOB). ) e teste de Tendência. . . . .	13
2.7	Série 5 (Reservas Internacionais Líquidas em milhões de dólares. ) e teste de Tendência. . . . .	15
2.8	Série 6 (Volume de exportações de café verde Colombiano, em milhares de sacos de 60 Kg.) e teste de Tendência. . . . .	16
2.9	Medição de temperatura para um processo químico (comportamento estacionário). . . . .	19
2.10	Medição de temperatura para um processo químico (comportamento não-estacionário). . . . .	19
2.11	A população dos EUA, Censo população x Tempo de 1900 a 2000. . . . .	22
5.1	Séries temporais de entrada (janeiro do ano 1970 até maio do ano 2011) . . . . .	38
5.2	Série temporal de entrada (janeiro do ano 1970 até maio do ano 2011) . . . . .	39
5.3	Série temporal de saída (janeiro do ano 1970 até maio do ano 2011). . . . .	39
5.4	Estrutura do Modelo de dados n=0 passos para frente. . . . .	40
5.5	Estrutura do Modelo de dados n=1 passos para frente. . . . .	40
5.6	Erro de saída com modelo de dados n=0 passos para frente. . . . .	41
5.7	Erro de saída com modelo de dados n=1 passos para frente. . . . .	41
5.8	Série temporal de Saída x Série temporal Estimada. (janeiro do ano 1970 até fevereiro do ano 2010). . . . .	46
5.9	Convergência de $\theta_1 MQR$ a $\theta MQ$ . . . . .	47
6.1	Diagrama para o método MQR. . . . .	52
6.2	Diagrama geral do método MQR3. . . . .	53
6.3	Diagrama para o método MQR3. . . . .	54
7.1	Blocos calculados do diagrama para o método MQR3. . . . .	57
7.2	Blocos calculados do diagrama para o método MQR3. . . . .	58

7.3	Blocos calculados do diagrama para o método MQR3. . . . .	59
7.4	Blocos calculados do diagrama para o método MQR3. . . . .	60
8.1	Dados de Validação. . . . .	64
8.2	Saída real (dados de validação), Método MQR3 e Método MQR. . . . .	64
8.3	Percentagem de Erro de saída entre as duas previsões. . . . .	66
8.4	Saída real (dados de validação), Método MQR3, MQR e MQR3 parcial . . . . .	67
8.5	Teste para tendência em dados validação série ICMS -Maranhão (X=observações)	70
8.6	Teste para tendência em dados validação série “Volume de exportações de café verde Colombiano” (X=observações) . . . . .	70
9.1	Série 1 Saída do Sistema x Saída estimada. . . . .	74
9.2	Série 1. Convergência de $\theta MQR$ a $\theta MQ$ . . . . .	75
9.3	Série 2 Saída do Sistema x Saída estimada. . . . .	76
9.4	Série 2. Convergência de $\theta MQR$ a $\theta MQ$ . . . . .	77
9.5	Série 3 Saída do Sistema x Saída estimada. . . . .	78
9.6	Série 3. Convergência de $\theta MQR$ a $\theta MQ$ . . . . .	79
9.7	Série 4 Saída do Sistema x Saída estimada. . . . .	80
9.8	Série 4. Convergência de $\theta MQR$ a $\theta MQ$ . . . . .	81
9.9	Série 5 Saída do Sistema x Saída estimada. . . . .	82
9.10	Série 5. Convergência de $\theta MQR$ a $\theta MQ$ . . . . .	83
9.11	Série 6 Saída do Sistema x Saída estimada. . . . .	84
9.12	Série 6. Convergência de $\theta MQR$ a $\theta MQ$ . . . . .	85

# Lista de Tabelas

3.1	Valores reais e previstos para exemplo modelo simples. . . . .	27
3.2	Valores reais e previstos para exemplo modelo linear. . . . .	28
5.1	Modelos de dados com diferentes séries temporais de entrada . . . . .	42
5.2	Modelos de dados com diferentes séries temporais de entrada (2). . . . .	43
5.3	Critério: % Erro Médio multiplicado pelos Dados de entrada. . . . .	43
5.4	Modelo de dados. . . . .	44
5.5	Dados de validação. . . . .	45
5.6	Primeiros dados de validação. . . . .	47
5.7	Resultados de previsão com MQR (Tabela 1). . . . .	48
5.8	Resultados de previsão com MQR (Tabela 2). . . . .	48
5.9	Erro de previsão com MQR (Tabela 1). . . . .	49
5.10	Erro de previsão com MQR (Tabela 2). . . . .	49
5.11	Porcentagem do erro médio de previsão com MQR. . . . .	49
7.1	Resultados de previsão final com MQR3 (1). . . . .	61
7.2	Resultados de previsão final com MQR3 (2). . . . .	61
7.3	% Erro de previsão final com MQR3 (1). . . . .	61
7.4	% Erro de previsão final com MQR3 (2). . . . .	62
7.5	Porcentagem do erro médio de previsão com MQR3. . . . .	62
8.1	%Diferença entre as duas previsões . . . . .	65
8.2	Porcentagem de erro médio de previsão com os métodos: MQR3, MQR e MQR3 parcial. . . . .	66
8.3	Comparação dos EPAM para a escolha do modelo de alisamento exponencial. . .	69
8.4	Quadro comparativo entre os exemplos e resultados da previsão para o ICMS e Exportações de café. . . . .	69
9.1	Série 1. Porcentagem do erro médio de previsão com os métodos de : MQR3, MQR e MQR parcial. . . . .	75
9.2	Série 2. Porcentagem do erro médio de previsão com os métodos de : MQR3, MQR e MQR parcial. . . . .	77

9.3	Série 3. Percentagem do erro médio de previsão com os métodos de : MQR3, MQR e MQR parcial. . . . .	79
9.4	Série 4. Percentagem do erro médio de previsão com os métodos de : MQR3, MQR e MQR da matriz filtrada. . . . .	81
9.5	Série 5. Percentagem do erro médio de previsão com os métodos de : MQR3, MQR e MQR da matriz filtrada. . . . .	83
9.6	Série 6. Percentagem do erro médio de previsão com os métodos de : MQR3, MQR e MQR da matriz filtrada. . . . .	85

# Lista de Acrônimos e Notação

MQ	Mínimos Quadrados.
MQR	Mínimos Quadrados Recursivo.
MQR3	Mínimos Quadrados Recursivo com Contribuição.

<b>I</b>	matriz identidade de dimensão apropriada
$L > 0$	indica que a matriz $L$ é simétrica definida positiva
$L \geq 0$	indica que a matriz $L$ é simétrica semi-definida positiva
$A$	notação para matrizes (letras maiúsculas do alfabeto latino)
$A'$	( $'$ ), pós-posto a um vetor ou matriz, indica a operação de transposição
$\mathbb{R}$	Conjunto dos números reais
$\mathbb{Z}$	conjunto dos números inteiros
$\mathbb{Z}_+$	conjunto dos números inteiros não negativos
$\mathbb{N}$	conjunto dos números naturais (incluindo o zero)
<b>O</b>	matriz de zeros de dimensão apropriada
$g!$	símbolo ( $!$ ), denota fatorial, isto é, $g! = g(g - 1) \cdots (2)(1)$ para $g \in \mathbb{N}$
$N$	especialmente utilizada para denotar o número de vértices de um polítopo
$n$	especialmente utilizada para representar a ordem uma matriz quadrada
$\Delta_n$	simplex unitário de $N$ variáveis
$\alpha$	especialmente utilizada para representar as incertezas de um sistema



# Sumário

<b>1</b>	<b>Introdução Geral.</b>	<b>1</b>
	<b>Introdução Geral</b>	<b>1</b>
<b>2</b>	<b>Séries Temporais.</b>	<b>3</b>
2.1	Introdução. . . . .	3
2.2	Análise de séries temporais. . . . .	3
2.3	Componentes de uma série temporal. . . . .	5
2.4	Métodos de estimação da tendência. . . . .	6
2.4.1	Tendência polinomial (Procedimento paramétrico). . . . .	6
2.5	Testes para tendência . . . . .	7
2.5.1	Testes antes da estimação de $T_t$ . . . . .	8
2.5.1.1	Exemplos . . . . .	9
2.6	Modelos de séries temporais. . . . .	17
2.6.1	Séries temporais estacionárias e não-estacionárias. . . . .	18
2.6.2	Função de Autocorrelação . . . . .	20
2.6.3	Seleção do Modelo para múltiplas séries temporais de entrada. . . . .	21
<b>3</b>	<b>Mínimos quadrados</b>	<b>23</b>
3.1	Introdução. . . . .	23
3.2	Método de mínimos quadrados - MQ . . . . .	23
3.3	Exemplos numéricos com MQ. . . . .	26
3.3.1	Exemplo com modelo simples . . . . .	26
3.3.2	Exemplo de modelo linear. . . . .	27
3.3.3	Exemplo de modelo não linear. . . . .	29
<b>4</b>	<b>Método de mínimos quadrados recursivo e múltiplas séries temporais.</b>	<b>31</b>
4.1	Introdução. . . . .	31
4.2	Análise multivariada de dados . . . . .	31
4.3	Método de mínimos quadrados recursivo - MQR. . . . .	33

<b>5</b>	<b>Aplicação do método MQR a múltiplas séries temporais.</b>	<b>37</b>
5.1	Introdução . . . . .	37
5.1.1	MQR aplicado nas exportações de café da Colômbia. . . . .	38
5.2	Validação do método MQR. . . . .	44
<b>6</b>	<b>Contribuição computacional para o método de mínimos quadrados recursivo - MQR3</b>	<b>51</b>
6.1	Introdução . . . . .	51
6.2	Diagrama implementado para o método de mínimos quadrados recursivo - MQR	51
6.3	Diagrama implementado para o método de mínimos quadrados recursivo 3 - MQR3 . . . . .	52
<b>7</b>	<b>Aplicação do método MQR3 a múltiplas séries temporais.</b>	<b>57</b>
7.1	Introdução . . . . .	57
7.2	MQR3 aplicado nas exportações de café na Colômbia. . . . .	58
7.3	Validação do método MQR3 . . . . .	60
<b>8</b>	<b>Análise e resultados das aplicações dos métodos MQR e MQR3 à múltiplas séries temporais.</b>	<b>63</b>
8.1	Introdução . . . . .	63
8.2	Análise e resultados dos métodos MQR e MQR3. . . . .	63
8.3	Análise do método MQR3. . . . .	67
8.3.1	Previsão mensal do ICMS do estado de Maranhão para o ano de 2008. . . . .	68
<b>9</b>	<b>Exemplos: aplicações dos métodos MQR e MQR3 a séries temporais com previsão de um passo.</b>	<b>73</b>
9.1	Introdução . . . . .	73
9.2	Resultados com os métodos MQR e MQR3. . . . .	73
9.2.1	Série 1: Produção de café verde Colombiano, medida em milhares de sacos de café. . . . .	73
9.2.2	Série 2: Câmbio de Dólar a peso Colombiano. . . . .	75
9.2.3	Série 3: Exportações de petróleo e derivados em milhões de dólares (FOB). . . . .	77
9.2.4	Série 4: Exportações de café em milhões de dólares (FOB). . . . .	79
9.2.5	Série 5: Reservas Internacionais Líquidas em milhões de dólares. . . . .	81
9.2.6	Série 6: Volume de exportações de café verde Colombiano, em milhares de sacos de 60 Kg. . . . .	83
<b>10</b>	<b>Conclusões.</b>	<b>87</b>
	<b>Bibliografia</b>	<b>91</b>

## Introdução Geral.

A história pode mostrar-nos o poder de cálculo do método de mínimos quadrados e o conceito de previsões de séries temporais. Legendre foi o primeiro a publicar um relato da teoria no ano de 1805, no entanto o conceito está fortemente ligado a Carl Friedrich Gauss, que desenvolveu o método com um tratamento probabilístico de erros de observação em “Theoria Motus Coporum Coelestum” [1].

Esta teoria foi feita para resolver um problema prático, baseada na estimativa de seis coeficientes paramétricos constantes que determinam a órbita elíptica de um corpo planetário, com base em  $n=6$  observações [2]. Nos anos de 1821, 1823 e 1826 foram desenvolvidos outros trabalhos no assunto com o nome de “Combinationis Theoria Obnoxiae mínima Erroribus”, por Gauss. Nestes, Gauss abandonou o tratamento do primeiro artigo, em que ele restringe a atenção para as estimativas que são uma função linear das observações [3], e apresentou seu teorema mais famoso que é: “Entre todas as estimativas lineares de erro consistentes, as estimativas dos mínimos quadrados têm o mínimo erro quadrático médio”. Neste trabalho Gauss desenvolveu um algoritmo sequencial ou de atualização recursiva, para a estimação de mínimos quadrados quando uma nova observação é recebida.

Além de sua importância em termos históricos, a análise de Gauss é interessante porque demonstra a elegância de sua abordagem e poder de sua mente, sendo Gauss capaz de obter com os Mínimos Quadrados Recursivo um algoritmo de baixo consumo computacional. A análise de Gauss representa o nascimento da teoria dos mínimos quadrados recursivo; uma teoria muito à frente de seu tempo que permaneceria latente por quase um século e meio antes de ser redescoberta em duas ocasiões: primeiramente pelo estatístico Plackett no ano de 1950, e mais tarde, de uma forma mais sofisticada por Rudolf Kalman no ano de 1960 [4].

Neste trabalho objetivamos apresentar uma metodologia para modelar séries temporais complexas e reais, usando como base o método de mínimos quadrados recursivo. A capacidade que tem o método MQR para processar múltiplas séries temporais é usada neste trabalho para modelar séries temporais econômicas e de produção. [5]

Normalmente quando o MQR é aplicado para estimar dados, enfatizamos as equações matemáticas dadas pela teoria, deixando de lado o contexto e informação valiosa que tem cada dado de uma série temporal. Procuramos nesta tese gerar um modelo de dados que realmente represente a saída, explorando informações intrínsecas nos dados das séries temporais, melhorando

nos casos em que seja possível o erro médio nas estimativas. Um caso de muita importância, onde será visível e útil aplicar nossa contribuição e metodologia, ocorre quando os dados de entrada em nosso modelo são atípicos, mas reais, porém o uso do MQR usual faz com que as estatísticas incrementem seu erro médio. O conhecimento das séries temporais, a modelagem adequada e a informação adicional do por que dos dados, permitem atingir nosso objetivo.

Estudos dos conceitos, métodos, algoritmos e comparações foram feitos para fundamentar a teoria e os resultados apresentados, com os enfoques principais de aplicar o método MQR e de apresentar uma contribuição computacional para melhorar o erro médio nas estimativas, para múltiplas séries temporais.

Inicialmente conceitos e a teoria básica, clássica das séries temporais são a apresentados. A seguir, introduzimos a teoria do método de mínimos quadrados, fundamental para nosso trabalho e aplicamos o conceito de recursividade ao mesmo. Posteriormente o MQR é aplicado a múltiplas séries temporais e validado.

Com a teoria básica apresentada, aplicamos o método de mínimos quadrados recursivo (MQR) à modelos de dados que construímos, para assim obter estimadores e as previsões correspondentes. Com base em critérios que estabelecemos, escolhemos o melhor modelo de dados e a ele aplicamos o MQR. Para obter melhores estimadores e previsões propomos o método MQR3 que constitui uma modificação do MQR usual. Ao mesmo modelo de dados que foi escolhido aplicamos o MQR3, obtendo o melhor estimador e as previsões correspondentes. Finalmente, aplicamos o método MQR3 a cada uma das séries temporais que apresentamos neste trabalho, testando a efetividade da previsão do mesmo em modelos simples de dados compostos por uma série temporal de entrada e uma de saída.

# Séries Temporais.

## 2.1 Introdução.

Uma série temporal é um conjunto de observações ordenadas no tempo [6]. Para introduzir uma ideia geral do que é uma série temporal, deve-se pensar em observações feitas para o campo de interesse, as quais são obtidas em instantes sucessivos de tempo. As observações podem ser feitas em pequenos intervalos de tempo ou em longos intervalos de tempo, isso depende do que se precisa e da disponibilidade dos dados. Pode-se encontrar ou construir séries temporais em diversas áreas de aplicação. Estas são muito conhecidas e usadas na área de meteorologia, finanças, marketing, entre outras. As séries temporais podem ser discretas ou contínuas. Um exemplo de série temporal contínua é o registro da atividade vulcânica, obtido através de um sismógrafo. Como exemplo de uma série temporal discreta pode-se mencionar os valores mensais das exportações de café de um país, na Bolsa de valores. Um uso possível das séries temporais é entender a sua dinâmica e assim poder realizar previsões, sendo este nosso objetivo principal.

## 2.2 Análise de séries temporais.

Em uma primeira análise do comportamento dos dados no tempo de uma série temporal, podem-se apresentar diversas variações; isto depende das observações e natureza da série. Então quando se sabe quais são os resultados que se procura obter ou aproximar da série, pode-se definir a análise ou método para aplicar, tendo assim um ponto de partida para conseguir o objetivo. A análise da série pode chegar a ser tão complexa quanto for necessário, usando simultaneamente diversas ferramentas para obter os melhores resultados. Um exemplo de série temporal encontra-se na figura 2.1, que representa as reservas internacionais da Colômbia em milhões de dólares. Esta apresenta diferentes comportamentos, e uma análise por intervalos de tempo da série é feita a seguir [7].

Observa-se no gráfico 2.1, que mostra mais de 450 dados na abscissa, um comportamento crescente dos dados o qual chama-se geralmente de tendência, sendo esta positiva para o exemplo. Se no mesmo gráfico só analisam-se os dados entre a observação cento e cinquenta e a observação cento e setenta, encontra-se uma tendência negativa e um resultado da análise totalmente diferente do anterior. Um caso misturando tendências negativas e positivas é visto no

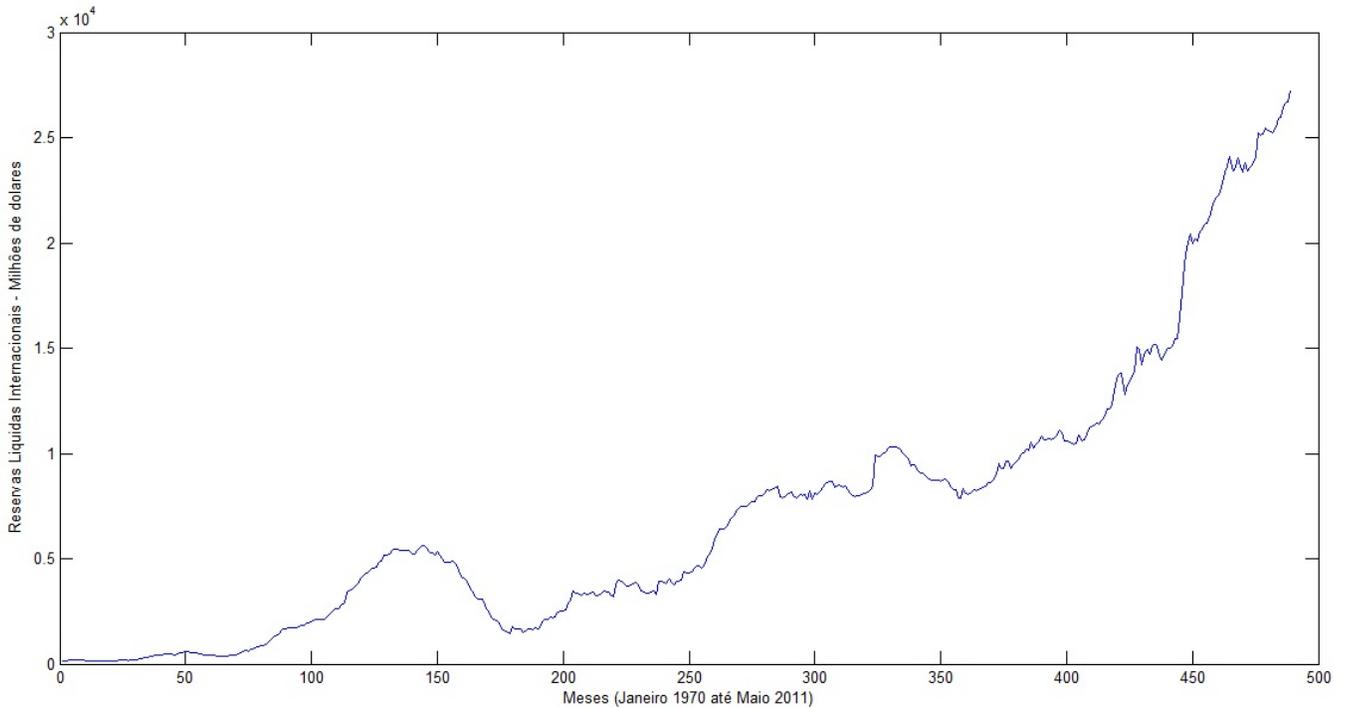


Figura 2.1: Série temporal de Reservas Internacionais da Colômbia em milhões de dólares.

intervalo de dados 300 ao 350 do mesmo gráfico. Evidenciar um conjunto das possíveis análises de uma série temporal e usar as informações intrínsecas dos dados da melhor forma possível é fundamental.

A primeira análise a ser feita nos dados de qualquer série temporal para obter-se uma percepção global da mesma é a gráfica. Com ela podemos classificar o comportamento dos dados dentro de um grupo geral, com o objetivo final de aplicar um algoritmo para o processamento correto dos dados.

Pode-se falar de diferentes tipos de séries temporais e assim entrar nas diversas áreas de estudo. Mas o essencial da análise de séries temporais é a observação, que tem como objetivo básico a inferência estatística.

Para à análise geral de uma série temporal  $Z(t_1), Z(t_2), \dots, Z(t_n)$  observada nos instantes  $t_1, t_2, \dots, t_n$ , apresentam-se três objetivos principais, que vão ajudar consideravelmente na compreensão deste trabalho [8]:

- Modelagem do fenômeno sob consideração.
- Obtenção de conclusões em termos estatísticos.
- Avaliação da adequação do modelo em termos de previsão.

Outros objetivos também podem ser importantes na hora da análise, dentre eles destacamos [6, 9]:

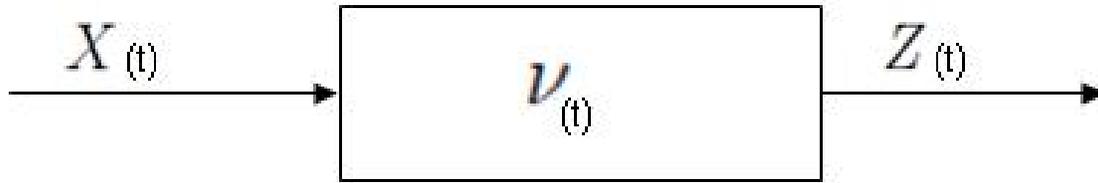


Figura 2.2: Sistema Dinâmico.

- Sumarizar os dados, descrevendo o comportamento da série com algumas técnicas úteis: construção do gráfico, verificação da existência de tendências, ciclos e variações sazonais, construção de histogramas e diagramas de dispersão, etc.
- Análise do sistema dinâmico: muitas situações em diversas áreas das ciências envolvem o conceito de sistema dinâmico, caracterizado por uma série de entrada  $X(t)$ , uma série de saída  $Z(t)$  e uma função de transferência  $\nu(t)$ , mostradas na figura 2.2:

De particular importância são os sistemas dinâmicos, onde a saída é relacionada com a entrada através de uma forma linear envolvendo  $\nu(t)$ ; um exemplo típico é mostrado na equação 2.1.

$$Z(t) = \sum_{j=0}^{\infty} \nu(j)X(t-j) \quad (2.1)$$

Alguns problemas de interesse são: estimar a função de transferência  $\nu(t)$ , conhecendo-se as séries de entrada e saída; fazer previsões da série  $Z(t)$ , com o conhecimento das duas séries; controlar a série de saída  $Z(t)$  [10].

## 2.3 Componentes de uma série temporal.

Um modelo clássico para uma série temporal  $Z(t_1), Z(t_2), \dots, Z(t_n)$  observada nos instantes  $t_1, t_2, \dots, t_n$  supõe que ela pode ser descrita como a soma de três componentes, equação 2.2:

- Tendência  $T_t$ : é um movimento que em geral pode ser representado por uma função suave do tempo  $t$ . Em séries econômicas a tendência é gerada por fatores que são medidos durante períodos longos de tempo.
- Componente sazonal  $S_t$ : é um movimento que pode ser representado por uma função periódica do tempo.
- Termo aleatório ou residual  $a_t$ : é um movimento que pode ser representado de forma puramente aleatória ou como um ruído branco.

$$Z_t = T_t + S_t + a_t \quad (2.2)$$

O modelo clássico apresentado na equação 2.2 é chamado de aditivo, útil quando as variáveis  $T_t$  e  $S_t$  são independentes. Para outros casos podemos ter modelos multiplicativos como é

mostrado na equação 2.3, modelos mistos como na equação 2.4 ou então outros modelos mais complexos [9].

$$Z_t = T_t * S_t * a_t, \quad t = 1, 2, \dots, N \quad (2.3)$$

$$Z_t = T_t * S_t + a_t, \quad t = 1, 2, \dots, N \quad (2.4)$$

## 2.4 Métodos de estimação da tendência.

Querendo modelar convenientemente a componente de tendência  $T_t$  de uma série, vamos considerar novamente os modelos de decomposição já citados:

Modelo Aditivo:

$$Z_t = T_t + S_t + a_t, \quad t = 1, 2, \dots, N \quad (2.5)$$

Agora se suponha que a componente sazonal  $S_t$ , não está presente, isto é:

$$Z_t = T_t + a_t \quad (2.6)$$

onde  $a_t$  é ruído branco, isto é, uma sequência de variáveis aleatórias independentes com média zero e variância  $\sigma_a^2$ .

Menciona-se nesta seção dois tipos de método de se estimar  $T_t$ .

- Procedimento Paramétrico: Consiste em ajustar uma função do tempo, como um polinômio, uma exponencial ou outra função suave de  $t$ .
- Procedimento não paramétrico: Consiste em suavizar ou filtrar os valores da série ao redor de um ponto para estimar a tendência naquele ponto.

Estimando-se a tendência através de  $\hat{T}_t$ , podemos obter a série ajustada para tendência  $\Upsilon_t$  ou livre de tendência:

$$\Upsilon_t = Z_t - \hat{T}_t = \hat{a}_t \quad (2.7)$$

devido à validade de 2.6.

### 2.4.1 Tendência polinomial (Procedimento paramétrico).

O problema mais sério que se tem no momento de estimar  $T_t$  através de um polinômio é que, embora ele possa se ajustar bem ao conjunto de valores observados, extrapolações futuras podem ser bastante ruins.

Suponha então que

$$T_t = \theta_0 + \theta_1 t + \dots + \theta_m t^m \quad (2.8)$$

onde o grau  $m$  do polinômio é bem menor que o número de observações  $N$ .

Para estimar os parâmetros  $\theta_j$  utilizamos o método de mínimos quadrados, ou seja, minimizamos

$$f(\theta_0, \dots, \theta_m) = \sum_{t=1}^N (T_t - \theta_0 - \theta_1 t - \dots - \theta_m t^m)^2 \quad (2.9)$$

para obter os estimadores de mínimos quadrados usuais  $\hat{\theta}_0, \hat{\theta}_1, \dots, \hat{\theta}_m$ . Maiores detalhes do método de mínimos quadrados estão amplamente desenvolvidos nos próximos capítulos.

Utilizando a equação 2.8 temos que a tendência estimada é dada por:

$$\hat{T}_t = \hat{\theta}_0 + \hat{\theta}_1 t + \dots + \hat{\theta}_m t^m \quad (2.10)$$

Então:

1. Se  $m = 1$ , a equação 2.8 reduz-se a uma tendência linear.
2. Se a série apresentar uma tendência exponencial.

$$T_t = a \exp(bt) \quad (2.11)$$

podemos aplicar o logaritmo a  $T_t$ , obtendo:

$$T_t^* = \log(T_t) = \log(a) + bt \quad (2.12)$$

Assim, para estimar  $a$  e  $b$ , ajustamos a  $T_t^*$ , um polinômio de grau 1 em  $t$  com  $\theta_0 = \log a$  e  $\theta_1 = b$ . Os valores estimados para  $a$  e  $b$  serão  $\hat{a} = \exp \hat{\theta}_0$  e  $\hat{b} = \hat{\theta}_1$  [6].

Exemplos que mostram a tendência polinomial são apresentados nas equações 3.17, 3.18 e 3.19 do capítulo 3.

## 2.5 Testes para tendência

Mostra-se que além da inspeção gráfica, podemos utilizar testes de hipóteses estatísticas para verificar se realmente existe tendência na série temporal. Isto pode ser feito de duas maneiras, antes ou depois de se obter a estimativa de  $T_t$  com base nas observações  $Z_t$ ,  $t = 1, 2, \dots, N$ . Se existe outra componente (como  $S_t$ ) na série temporal, é preciso eliminar a componente  $S_t$  antes de testar a presença de  $T_t$ .

As hipóteses a serem testadas serão:

- $H_0$ : Não existe tendência.
- $H_1$ : Existe tendência.

### 2.5.1 Testes antes da estimação de $T_t$ .

Apresentam-se três testes não paramétricos, que em geral se baseiam em hipóteses que podem não estar verificadas para o caso de uma série temporal; uma suposição comum é que as observações constituem uma amostra de uma população.

- Teste de sequências (Wald-Wolfowitz).

Suponha  $m$  como a mediana da série; atribuímos a cada valor de  $Z_t$  os símbolos:  $A$  se  $Z_t \geq m$  e  $B$  se  $Z_t < m$ . Teremos então  $N = (n_1 \text{ pontos } A) + (n_2 \text{ pontos } B)$ .

A estatística do teste é:

$T_1$  = número total de sequências.

Então, rejeitamos  $H_0$  se existem poucas sequências, ou seja se  $T_1$  é pequeno. Para um dado nível de significância  $\alpha$ , rejeitamos  $H_0$  se  $T_1 < w_\alpha$  onde  $w_\alpha$  é o  $\alpha$  - quantil da distribuição de  $T_1$ , [11].

Para  $n_1$  ou  $n_2$  maior que 20 podemos usar uma aproximação normal, isto é,  $T_1 \sim N(\mu, \sigma^2)$  onde:

$$\mu = \frac{2n_1n_2}{N} + 1 \quad (2.13)$$

$$\sigma^2 = \frac{2n_1n_2(2n_1n_2 - N)}{N^2(N - 1)} \quad (2.14)$$

- Teste do sinal (Cox-Stuart)

Agrupam-se as observações em pares  $(Z_1, Z_{1+c}), (Z_2, Z_{2+c}), \dots, (Z_{N-c}, Z_N)$  onde  $c = \frac{N}{2}$ , se  $N$  é par e  $c = \frac{N+1}{2}$  se  $N$  é ímpar. A cada par  $P(Z_i, Z_{i+c})$  associamos o sinal  $+$  se  $Z_i < Z_{i+c}$  e o sinal  $-$  se  $Z_i > Z_{i+c}$ , eliminando os empates. Seja  $n =$  número de pares onde  $Z_i \neq Z_{i+c}$  [6].

Neste teste as hipóteses são:

$$H_0 : P(Z_i < Z_{i+c}) = P(Z_i > Z_{i+c}) \quad \forall i \quad (2.15)$$

$$H_1 : P(Z_i < Z_{i+c}) \neq P(Z_i > Z_{i+c}) \quad \forall i \quad (2.16)$$

A estatística do teste é:

$T_2$  = número de pares com sinal  $+$ .

- Teste baseado no coeficiente de correlação de Spearman

Seja  $R_t$  as posições de cada  $Z_t$ , dentre as  $N$  observações; se houver empates utilize a posição média.

A estatística do teste é

$$T_3 = \sum_{t=1}^N (R_t - t)^2 \quad (2.17)$$

onde  $t = 1, 2, \dots, N$  são as posições “naturais” dos instantes de tempo.

Rejeitamos  $H_0$  se  $T_3 \leq c_1$  ou  $T_3 \geq c_2$ ; os valores de  $c_1$  e  $c_2$  são definidos em [11].

O teste também pode ser feito em função do coeficiente de correlação de Spearman dado por:

$$\rho = 1 - \frac{6T_3}{N(N^2 - 1)} \quad (2.18)$$

### 2.5.1.1 Exemplos

Querendo considerar os testes não paramétricos, vamos apresentar as seguintes seis séries temporais:

1. Produção de café verde Colombiano, medida em milhares de sacos de café.
2. Câmbio de Dólar a peso Colombiano.
3. Exportações de petróleo e derivados em milhões de dólares (FOB).
4. Exportações de café em milhões de dólares (FOB).
5. Reservas Internacionais Líquidas em milhões de dólares.
6. Volume de exportações de café verde Colombiano, em milhares de sacos de 60 Kg.

As séries temporais apresentadas têm 497 dados cada uma, de janeiro 1970 até maio 2011, onde cada observação representa o resultado de um mês [12]. Para cada série temporal vamos aplicar os testes para tendências estudados anteriormente, mostrando os resultados e gráficos obtidos em cada processo.

- A primeira série temporal “produção de café verde Colombiano” é mostrada no gráfico 2.3, junto com os gráficos obtidos depois de aplicar os testes.

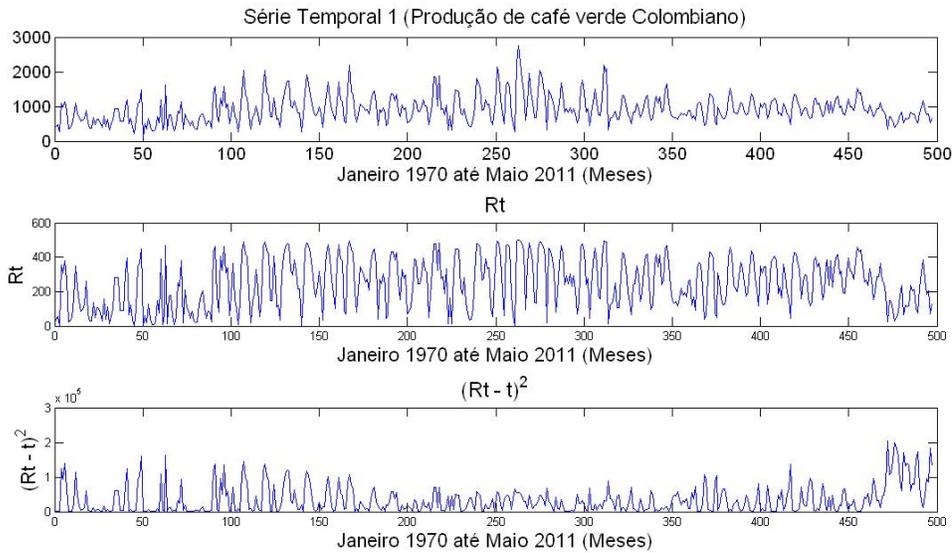


Figura 2.3: Série 1 (Produção de café verde Colombiano, medida em milhares de sacos de café) e teste de Tendência.

Resultado do teste de seqüências (Wald-Wolfowitz):

$$m = 857$$

$$n_1 = 249$$

$$m_1 = 248$$

$$N = 497$$

$$T_1 = 1$$

Rejeita-se  $H_0$  dado que existem poucas seqüências, ou seja se  $T_1$  é pequeno.

Dado que  $n_1$  é maior que 20 pode-se usar a aproximação normal  $T_1 \sim N(\mu, \sigma^2)$  com:

$$\mu = \frac{2n_1n_2}{N} + 1 = 249,4990$$

$$\sigma^2 = \frac{2n_1n_2(2n_1n_2 - N)}{N^2(N - 1)} = 123,9985$$

Resultados do teste baseado no coeficiente de correlação de Spearman:

$$T_3 = \sum_{t=1}^N (R_t - t)^2 = 16701706$$

Como definimos anteriormente na equação 2.17:  $R_t$  como as posições de cada  $Z_t$ , dentre as  $N$  observações e  $t = 1, 2, \dots, N$  as posições “naturais” dos instantes de tempo. Sendo então  $T_3$  a somatória da diferença da posição da série e a posição natural ao quadrado, para a série trabalhada. No caso de uma tendência pura positiva temos  $T_3 = 0$  e para uma tendência pura negativa temos  $T_3 = 40920992$ . O resultado para a série trabalhada neste caso representa um  $\%T_3 = 40,81\%$  do máximo.

Rejeita-se  $H_0$  dado que  $T_3 \leq c_1$  ou  $T_3 \geq c_2$ ; os valores de  $c_1$  e  $c_2$  são definidos em [11].

Calcula-se o coeficiente de correlação de Spearman, obtendo:

$$\rho = 1 - \frac{6T_3}{N(N^2 - 1)} = 0.1837$$

- A segunda série temporal “Câmbio de Dólar a peso Colombiano” é mostrada no gráfico 2.4, junto com os gráficos obtidos depois de aplicar os testes.

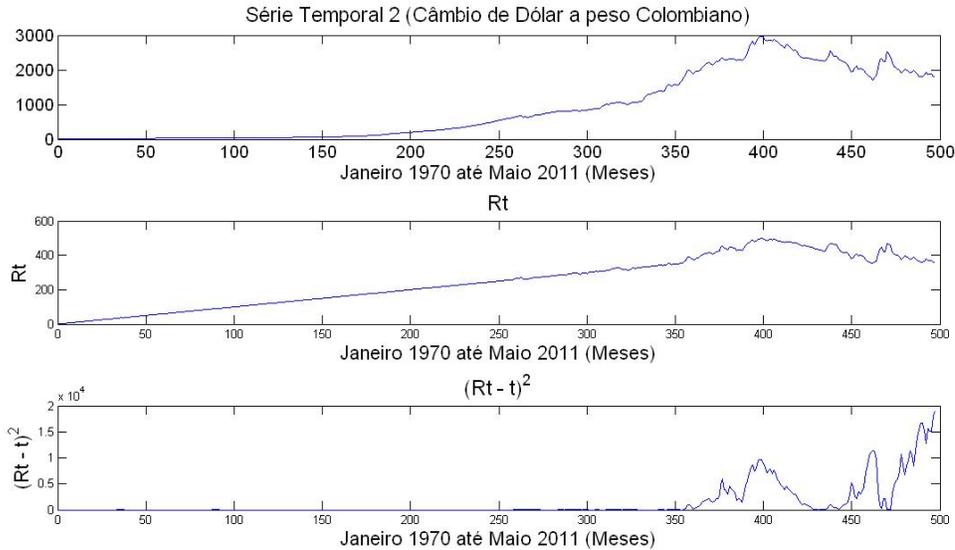


Figura 2.4: Série 2 (Câmbio de Dólar a peso Colombiano.) e teste de Tendência.

Resultado do teste de seqüências (Wald-Wolfowitz):

$$m = 530, 5400$$

$$n_1 = 249$$

$$m_1 = 248$$

$$N = 497$$

$$T_1 = 1$$

Rejeitamos  $H_0$  dado que existem poucas seqüências, ou seja se  $T_1$  é pequeno.

Dado que  $n_1$  é maior que 20 podemos usar a aproximação normal  $T_1 \sim N(\mu, \sigma^2)$  com:

$$\mu = \frac{2n_1n_2}{N} + 1 = 249,4990$$

$$\sigma^2 = \frac{2n_1n_2(2n_1n_2 - N)}{N^2(N - 1)} = 123,9985$$

Resultados do teste baseado no coeficiente de correlação de Spearman:

$$T_3 = \sum_{t=1}^N (R_t - t)^2 = 674000$$

Como definimos anteriormente na equação 2.17:  $R_t$  como as posições de cada  $Z_t$ , dentre as  $N$  observações e  $t = 1, 2, \dots, N$  as posições “naturais” dos instantes de tempo. Sendo então  $T_3$  a somatória da diferença da posição da série e a posição natural ao quadrado, para a série trabalhada. No caso de uma tendência pura positiva temos  $T_3 = 0$  e para uma tendência pura negativa temos  $T_3 = 40920992$ . O resultado para a série trabalhada neste caso representa um  $\%T_3 = 1,65\%$  do máximo.

Rejeitamos  $H_0$  dado que  $T_3 \leq c_1$  ou  $T_3 \geq c_2$ ; os valores de  $c_1$  e  $c_2$  são definidos em [11].

Calculamos o coeficiente de correlação de Spearman, obtendo:

$$\rho = 1 - \frac{6T_3}{N(N^2 - 1)} = 0.9671$$

- A terceira série temporal “Exportações de petróleo e derivados” é mostrada no gráfico 2.5, junto com os gráficos obtidos depois de aplicar os testes.

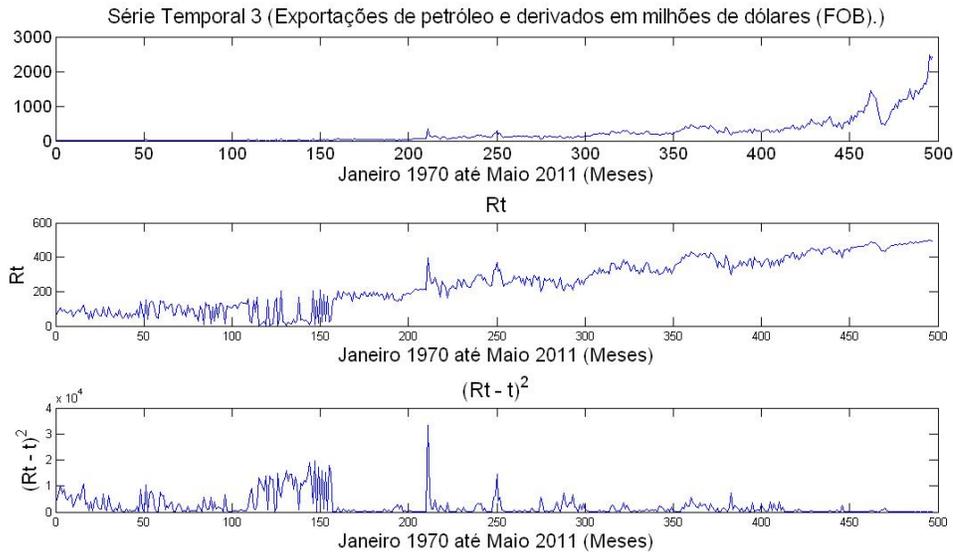


Figura 2.5: Série 3 (Exportações de petróleo e derivados em milhões de dólares (FOB).) e teste de Tendência.

Resultado do teste de sequências (Wald-Wolfowitz):

$$m = 112,0600$$

$$n_1 = 249$$

$$m_1 = 248$$

$$N = 497$$

$$T_1 = 11$$

Rejeitamos  $H_0$  dado que existem poucas sequências, ou seja se  $T_1$  é pequeno.

Dado que  $n_1$  é maior que 20 podemos usar a aproximação normal  $T_1 \sim N(\mu, \sigma^2)$  com:

$$\mu = \frac{2n_1n_2}{N} + 1 = 249,4990$$

$$\sigma^2 = \frac{2n_1n_2(2n_1n_2 - N)}{N^2(N - 1)} = 123,9985$$

Resultados do teste baseado no coeficiente de correlação de Spearman:

$$T_3 = \sum_{t=1}^N (R_t - t)^2 = 1109886$$

Como definimos anteriormente na equação 2.17:  $R_t$  como as posições de cada  $Z_t$ , dentre as  $N$  observações e  $t = 1, 2, \dots, N$  as posições “naturais” dos instantes de tempo. Sendo então  $T_3$  a somatória da diferença da posição da série e a posição natural ao quadrado, para a série trabalhada. No caso de uma tendência pura positiva temos  $T_3 = 0$  e para uma tendência pura negativa temos  $T_3 = 40920992$ . O resultado para a série trabalhada neste caso representa um  $\%T_3 = 2,71\%$  do máximo.

Rejeitamos  $H_0$  dado que  $T_3 \leq c_1$  ou  $T_3 \geq c_2$ ; os valores de  $c_1$  e  $c_2$  são definidos em [11]. Calculamos o coeficiente de correlação de Spearman, obtendo:

$$\rho = 1 - \frac{6T_3}{N(N^2 - 1)} = 0,9458$$

- A quarta série temporal “Exportações de café” é mostrada no gráfico 2.6, junto com os gráficos obtidos depois de aplicar os testes.

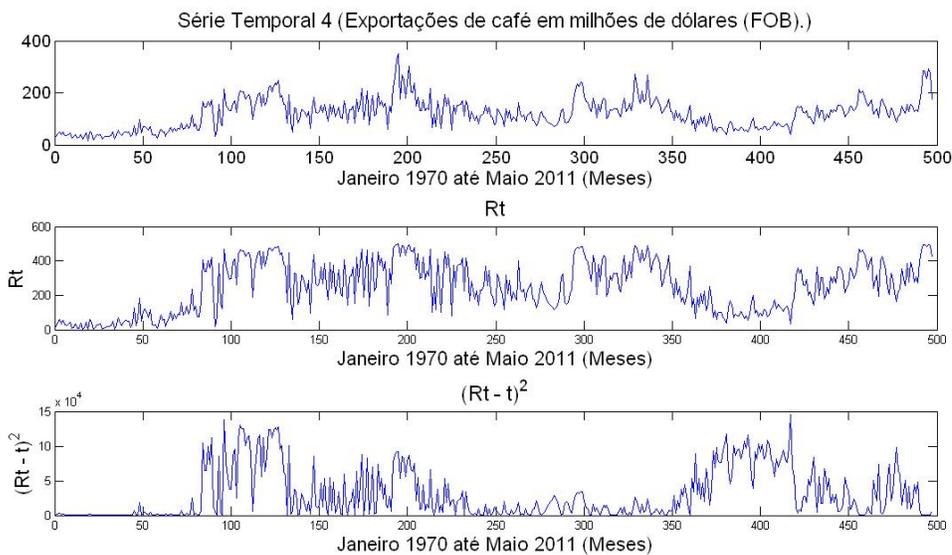


Figura 2.6: Série 4 (Exportações de café em milhões de dólares (FOB).) e teste de Tendência.

Resultado do teste de sequências (Wald-Wolfowitz):

$$m = 119,3200$$

$$n_1 = 249$$

$$m_1 = 248$$

$$N = 497$$

$$T_1 = 1$$

Rejeitamos  $H_0$  dado que existem poucas sequências, ou seja se  $T_1$  é pequeno.

Dado que  $n_1$  é maior que 20 podemos usar a aproximação normal  $T_1 \sim N(\mu, \sigma^2)$  com:

$$\mu = \frac{2n_1n_2}{N} + 1 = 249,4990$$

$$\sigma^2 = \frac{2n_1n_2(2n_1n_2 - N)}{N^2(N - 1)} = 123,9985$$

Resultados do teste baseado no coeficiente de correlação de Spearman:

$$T_3 = \sum_{t=1}^N (R_t - t)^2 = 15070114$$

Como definimos anteriormente na equação 2.17:  $R_t$  como as posições de cada  $Z_t$ , dentre as  $N$  observações e  $t = 1, 2, \dots, N$  as posições “naturais” dos instantes de tempo. Sendo então  $T_3$  a somatória da diferença da posição da série e a posição natural ao quadrado, para a série trabalhada. No caso de uma tendência pura positiva temos  $T_3 = 0$  e para uma tendência pura negativa temos  $T_3 = 40920992$ . O resultado para a série trabalhada neste caso representa um  $\%T_3 = 36,83\%$  do máximo.

Rejeitamos  $H_0$  dado que  $T_3 \leq c_1$  ou  $T_3 \geq c_2$ ; os valores de  $c_1$  e  $c_2$  são definidos em [11].

Calculamos o coeficiente de correlação de Spearman, obtendo:

$$\rho = 1 - \frac{6T_3}{N(N^2 - 1)} = 0,2635$$

- A quinta série temporal “Reservas Internacionais Líquidas” é mostrada no gráfico 2.7, junto com os gráficos obtidos depois de aplicar os testes.

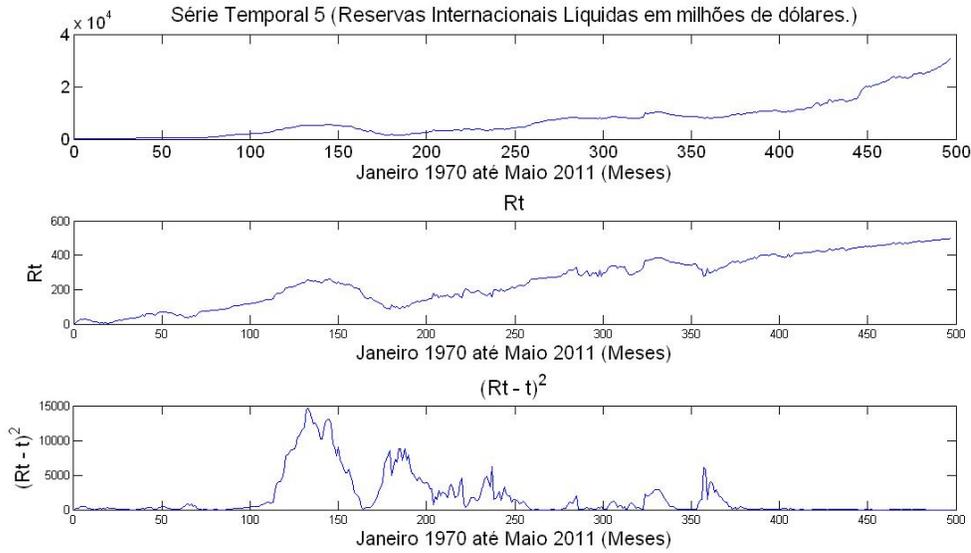


Figura 2.7: Série 5 (Reservas Internacionais Líquidas em milhões de dólares. ) e teste de Tendência.

Resultado do teste de seqüências (Wald-Wolfowitz):

$$m = 5381$$

$$n_1 = 249$$

$$m_1 = 248$$

$$N = 497$$

$$T_1 = 0$$

Rejeitamos  $H_0$  dado que existem poucas seqüências, ou seja se  $T_1$  é pequeno.

Dado que  $n_1$  é maior que 20 podemos usar a aproximação normal  $T_1 \sim N(\mu, \sigma^2)$  com:

$$\mu = \frac{2n_1n_2}{N} + 1 = 249,4990$$

$$\sigma^2 = \frac{2n_1n_2(2n_1n_2 - N)}{N^2(N - 1)} = 123,9985$$

Resultados do teste baseado no coeficiente de correlação de Spearman:

$$T_3 = \sum_{t=1}^N (R_t - t)^2 = 844760$$

Como definimos anteriormente na equação 2.17:  $R_t$  como as posições de cada  $Z_t$ , dentre as  $N$  observações e  $t = 1, 2, \dots, N$  as posições “naturais” dos instantes de tempo. Sendo então  $T_3$  a somatória da diferença da posição da série e a posição natural ao quadrado, para a série trabalhada. No caso de uma tendência pura positiva temos  $T_3 = 0$  e para uma tendência pura negativa temos  $T_3 = 40920992$ . O resultado para a série trabalhada neste caso representa um  $\%T_3 = 2,06\%$  do máximo.

Rejeitamos  $H_0$  dado que  $T_3 \leq c_1$  ou  $T_3 \geq c_2$ ; os valores de  $c_1$  e  $c_2$  são definidos em [11].

Calculamos o coeficiente de correlação de Spearman, obtendo:

$$\rho = 1 - \frac{6T_3}{N(N^2 - 1)} = 0,9587$$

- A sexta série temporal “Volume de exportações de café verde Colombiano” é mostrada no gráfico 2.8, junto com os gráficos obtidos depois de aplicar os testes.

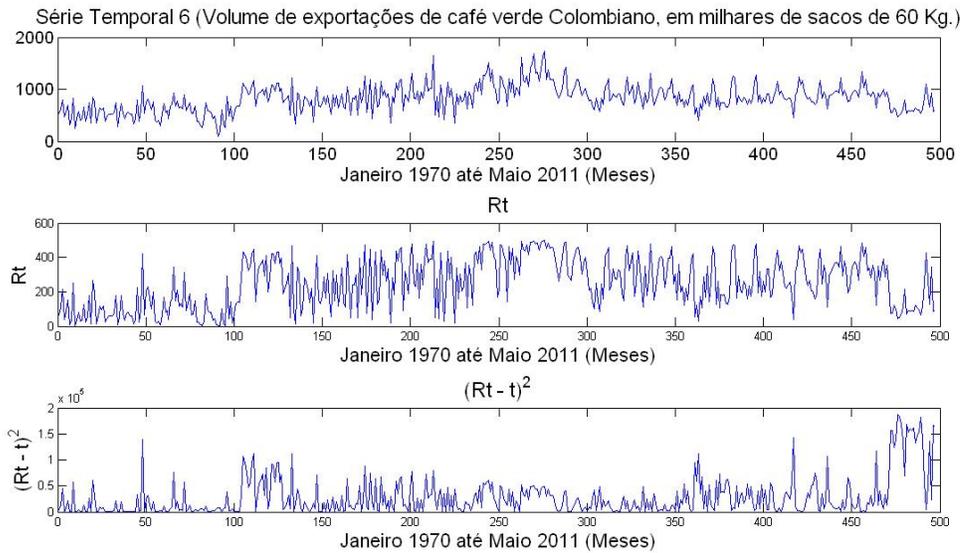


Figura 2.8: Série 6 (Volume de exportações de café verde Colombiano, em milhares de sacos de 60 Kg.) e teste de Tendência.

Resultado do teste de sequências (Wald-Wolfowitz):

$$m = 827$$

$$n_1 = 250$$

$$m_1 = 247$$

$$N = 497$$

$$T_1 = 1$$

Rejeitamos  $H_0$  dado que existem poucas sequências, ou seja se  $T_1$  é pequeno.

Dado que  $n_1$  é maior que 20 podemos usar a aproximação normal  $T_1 \sim N(\mu, \sigma^2)$  com:

$$\mu = \frac{2n_1n_2}{N} + 1 = 249,4990$$

$$\sigma^2 = \frac{2n_1n_2(2n_1n_2 - N)}{N^2(N - 1)} = 123,9985$$

Resultados do teste baseado no coeficiente de correlação de Spearman:

$$T_3 = \sum_{t=1}^N (R_t - t)^2 = 13942312$$

Como definimos anteriormente na equação 2.17:  $R_t$  como as posições de cada  $Z_t$ , dentre as  $N$  observações e  $t = 1, 2, \dots, N$  as posições “naturais” dos instantes de tempo. Sendo então  $T_3$  a somatória da diferença da posição da série e a posição natural ao quadrado, para a série trabalhada. No caso de uma tendência pura positiva temos  $T_3 = 0$  e para uma tendência pura negativa temos  $T_3 = 40920992$ . O resultado para a série trabalhada neste caso representa um  $\%T_3 = 34,07\%$  do máximo.

Rejeitamos  $H_0$  dado que  $T_3 \leq c_1$  ou  $T_3 \geq c_2$ ; os valores de  $c_1$  e  $c_2$  são definidos em [11].

Calculamos o coeficiente de correlação de Spearman, obtendo:

$$\rho = 1 - \frac{6T_3}{N(N^2 - 1)} = 0,3186$$

## 2.6 Modelos de séries temporais.

É fundamental entender o conceito de modelo, para isto define-se modelo como “uma representação compacta dos aspectos essenciais de um sistema existente que descreverá como ele em atividade irá se comportar”. E no âmbito matemático o modelo é “uma representação matemática com aspectos essenciais de um sistema existente, apresentando conhecimento de forma compacta”. As duas definições anteriores estão vinculadas com sistema, então se define sistema como: “um conjunto integrado de uma ou diversas estruturas que interagem entre si gerando um conjunto de dados possibilitando uma relação com outro conjunto de dados” [13].

Introduza-se brevemente a preocupação de muitas áreas de aplicação, as quais querem obter uma representação de processos físicos com modelos matemáticos. Querendo finalmente prever os resultados dos processos. Se o modelo matemático de um processo físico pode prever o resultado desse processo exatamente, ele é chamado um modelo determinista. [9]. A discussão de quanto um modelo determinístico é realista, não faz parte de nosso trabalho. O que se podemos falar é que, nos casos que não seja possível obter um modelo determinístico, a melhor representação do processo pode ser feita com um modelo estocástico ou probabilístico. Para o qual neste caso “devemos tratar os dados observados como uma realização de uma população infinita de séries temporais, de tal modo que poderia ter sido gerado pelo processo estocástico. Por isso, não é possível prever exatamente a saída do processo futuro. Em vez disso, pode-se proporcionar um intervalo de predição e a probabilidade com que a observação futura recairá naquele intervalo. No entanto, a fim de fazer isso é necessário assumir que determinadas propriedades do nosso processo permanecem constantes no tempo.” Daqui temos a noção de estacionariedade: se quisermos fazer previsões, então claramente temos que considerar que algo não varie com o tempo.” [14].

Podem ocorrer vários problemas com as observações das séries temporais, e medidas devem ser tomadas para evitá-los ou amenizá-los, as quais vamos a enunciar algumas neste trabalho:

- Estacionariedade: É frequente supor respeito de uma série temporal que ela seja estacionária, ou seja, ela se desenvolve no tempo aleatoriamente ao redor de uma média constante, refletindo alguma forma de equilíbrio estável. A maior parte das séries que encontramos na prática apresentam alguma forma de não-estacionariedade.

- Transformações: As transformações nos dados originais são necessárias quando temos presença de não estacionariedade. A transformação mais comum consiste em tomar diferenças sucessivas da série original, até obter-se uma estacionariedade. A primeira diferença da série  $Z(t)$  é definida assim:

$$\Delta Z(t) = Z(t) - Z(t - 1) \quad (2.19)$$

a segunda diferença é:

$$\Delta^2 Z(t) = \Delta [\Delta Z(t)] = \Delta [Z(t) - Z(t - 1)] \quad (2.20)$$

ou seja

$$\Delta^2 Z(t) = Z(t) - 2Z(t - 1) + Z(t - 2) \quad (2.21)$$

De modo geral,  $n$ -ésima diferença de  $Z(t)$  é:

$$\Delta^n Z(t) = \Delta [\Delta^{n-1} Z(t)] \quad (2.22)$$

- Observações perdidas e irregulares: Refere-se quando na série temporal temos observações perdidas, isto pode acontecer porque o instrumento de medição quebra em algum momento. Quando o número de observações perdidas é pequeno e não concentrado em um intervalo, podemos usar interpolação. Para outros casos temos que usar técnicas de análise específicas.
- Atípicos: Refere-se aos valores de uma série temporal que são aberrantes, valores com muita diferença das outras observações. Os quais podem ser originados por erros. No caso, que tenham algum significado, temos que identificar as observações aberrantes e incorporá-las na análise.
- Registros curtos: Nos casos que temos processos que tem poucas observações para análise. Quando acontece, temos que usar técnicas apropriadas na análise.

### 2.6.1 Séries temporais estacionárias e não-estacionárias.

O conceito de estacionariedade é uma base fundamental na análise de séries temporais, e trabalhando neste conceito podemos achar previsões. Agora, voltando com o gráfico 2.1 de reservas internacionais de Colômbia em milhões de dólares, exaltamos o seguinte “No entanto, é a série temporal não-estacionária, que é o mais interessante, em muitas aplicações, especialmente em administração e economia. Da mesma forma, em aplicações industriais, quando os processos são deixados sozinhos, eles são esperados para mostrar o comportamento não estacionário simplesmente seguindo a segunda lei da termodinâmica. De fato, na maioria dos processos de estacionariedade só é assegurada pela ação de controle efetuada a intervalos regulares e/ou de manutenção continuada dos componentes do sistema. Sem essas ações deliberadas para tornar os processos estacionários, uma solução não se deve concentrar nos dados originais que exibe estacionariedade, mas em vez disso, por exemplo, sobre as mudanças nas observações sucessivas”

[9]. Utilizando os exemplos gráficos descritos, apresento nas figuras 2.9 e 2.10, o comportamento estacionário e não estacionário respectivamente.

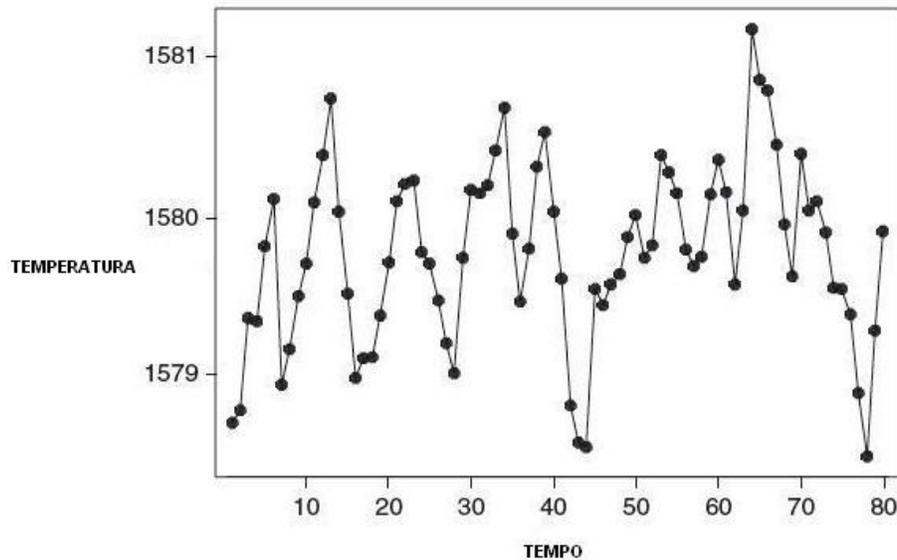


Figura 2.9: Medição de temperatura para um processo químico (comportamento estacionário).

O processo da figura 2.9 é rigorosamente controlado para ser em torno de um valor de referência determinado.

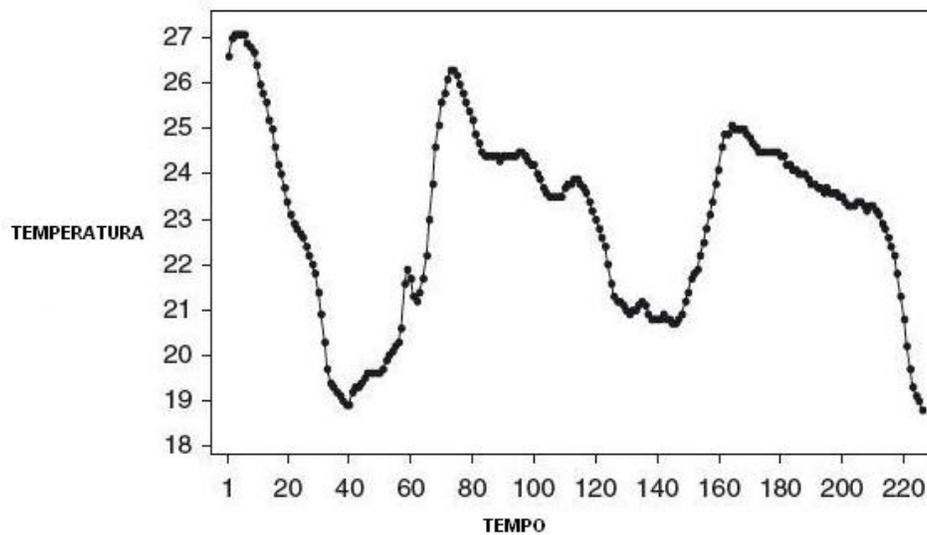


Figura 2.10: Medição de temperatura para um processo químico (comportamento não-estacionário).

O processo da figura 2.10 não é controlado de modo a que as medições de temperatura não variam em torno de um valor alvo.

No gráfico 2.9 apresentado podemos observar, a função específica que exerce o controlador físico para conseguir o comportamento estacionário. O qual tenta manter ao redor de uma média constante a temperatura desejada.

Na prática grande parte das séries temporais têm alguma não estacionariedade. “Assim, as séries econômicas apresentam em geral tendências, sendo o caso mais simples aquele em que a série flutua ao redor de uma reta, com inclinação positiva a negativa (tendência linear)” [6].

Para deixar o conceito de estacionariedade completo vamos dar sua definição estrita, descrita assim: “uma série temporal é chamada de estacionária se a distribuição de probabilidade conjunta de todas as  $n$  observações,  $\{z_{t+1}, z_{t+2}, \dots, z_{t+n}\}$  desta série permanecem as mesmas que outro conjunto de  $n$  observações deslocadas por  $k$  unidades de tempo, isto é,  $\{z_{t+1+k}, z_{t+2+k}, \dots, z_{t+n+k}\}$ . Para fins práticos, no entanto, nos definimos uma série temporal estacionária como um processo de variância finita onde a média e a variância é constante no tempo, e a correlação entre as observações de diferentes pontos no tempo fica dependente.” [9].

### 2.6.2 Função de Autocorrelação

Usando a definição estrita de estacionariedade, mencionada anteriormente, uma série temporal monovariável estacionária  $z_t$  que tem variância finita com média constante e variação ao longo do tempo  $t$  pode-se escrever como:

$$E_{(z_t)} = \mu_t = \mu \quad (2.23)$$

Com a estimativa para a média dada por:

$$\bar{z} = \frac{1}{T} \sum_{t=1}^T z_t \quad (2.24)$$

Da mesma forma a variância pode ser calculada usando:

$$\hat{\sigma}_z^2 = \frac{1}{T-1} \sum_{t=1}^T (z_t - \bar{z})^2 \quad (2.25)$$

E para um estudo melhor de séries temporais estacionarias, introduzimos a autocovariância e autocorrelação. Em geral, define-se a covariância entre duas variáveis aleatórias,  $X$  e  $Y$ , conforme:

$$Cov(X, Y) = E[(X - E(X))(Y - E(Y))] \quad (2.26)$$

e sua correlação é simplesmente o escalar das covariâncias por seu desvio padrão

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sqrt{\sigma_X^2} \sqrt{\sigma_Y^2}} \quad (2.27)$$

Definimos a função de autocovariância como:

$$\nu(k) = E[(z_{t+k} - \mu)(z_t - \mu)] \quad (2.28)$$

Note-se que a variância da série temporal é  $\nu(0)$ . Seguindo a equação 2.27, definimos a função de autocorrelação (ACF) para uma série temporal estacionária como

$$\rho(k) = \frac{\nu(k)}{\sqrt{\nu(0)}\sqrt{\nu(0)}} = \frac{\nu(k)}{\nu(0)} \quad (2.29)$$

A ACF desempenha um papel extremamente importante na identificação de modelos de séries de temporais. É claro que na vida real não podemos saber o verdadeiro valor da ACF, mas nós estimamos a partir dos dados à mão, usando.

$$\hat{\nu}(k) = \frac{1}{T} \sum_{t=1}^{T-k} (z_{t+k} - \bar{z})(z_t - \bar{z}) \quad (2.30)$$

e

$$\hat{\rho}(k) = \frac{\hat{\nu}(k)}{\hat{\nu}(0)} \quad (2.31)$$

Esta será uma ferramenta principal que iremos usar para verificar a autocorrelação dos dados, bem como os resíduos obtidos na nossa modelagem.

Outras considerações e outros modelos podem ser tidos em conta, para procurar ter uma melhor análise. Os estudos mais detalhados dos modelos não estacionários podem ser observados nas referências [9, 6, 15].

### 2.6.3 Seleção do Modelo para múltiplas séries temporais de entrada.

“Uma das tarefas mais difíceis na análise de séries temporais é a seleção do modelo e a escolha de um modelo adequado para um determinado conjunto de dados. Mesmo analistas treinados em série temporais, muitas vezes, é difícil escolher um dos modelos” [9]. A anterior idéia deixa evidente que ainda tendo um modelo de dados bem ajustado, temos que ter critérios para saber se ele é bom. Na atualidade temos muitos critérios para saber se um modelo de dados é adequado, o qual ficaria complicado querer explicar todos eles neste trabalho. Nos próximos capítulos vamos a construir nosso modelo para múltiplas séries temporais de entrada y assim poder aplicar o método de MQR; agora só queremos exemplificar como a escolha errada de um modelo pode ter previsões pouco acertadas da realidade. Vamos apresentar um exemplo do que acontece quando o critério de validação aplicado numa série temporal não é adequado: “Uma vez que o modelo é uma aproximação, é possível que vários modelos possam ser mais ou menos equivalentes. Isso levanta a questão de qual deles prefere. Aqui, pode-se aplicar uma série de critérios de seleção do modelo. Uma seria de tentar encontrar o modelo que minimiza a soma dos quadrados dos resíduos ou maximiza o  $R^2$ . No entanto, tal como na análise de regressão, podemos fazer a soma dos quadrados dos resíduos pequenos ou o  $R^2$  tão grande quanto quisermos, adicionando mais termos no modelo. Se um modelo estatístico envolve muitos termos, os parâmetros serão geralmente mal estimados, uma vez que a informação contida nos dados será dissipada ao longo de muitos deles. Modelos que também envolvem parâmetros supérfluos, como regra geral, produzem previsões pobres. Isso é ilustrado na figura 2.11, onde montado um polinômio de décimo ordem para a população dos EUA, pelo censo oficial a cada

10 anos, a partir de 1900 ao 2000 e extrapolada para 2010. Obtivemos um ajuste perfeito para as 10 observações, mas uma previsão pobre e pouco realista para 2010. Super ajuste, como este é chamado, muitas vezes leva a modelos que "localmente" encaixam quase perfeitamente, "globalmente" executar-se mal" [9].

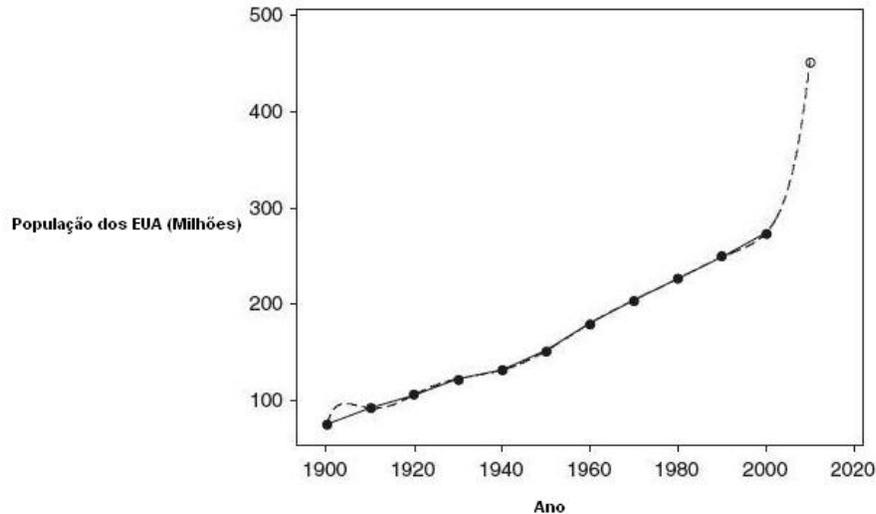


Figura 2.11: A população dos EUA, Censo população x Tempo de 1900 a 2000.

Com o exemplo anterior, temos um caso bem comum no análise de séries temporais. Onde conseguimos boas aproximações de nosso modelo mas na hora de validar ele, os critérios de validação usados apresentam grandes erros nas saídas. Tendo finalmente previsões erradas que invalidam o uso do modelo. Neste caso podemos fazer ajustes no modelo ou trabalhar com outros critérios para a seleção do modelo que alguns casos não é dependente dos dados, o número de observações, ou o número de parâmetros.

Implementar um modelo não garante, necessariamente, um processo de previsão. “Será necessário especificar, além do modelo, uma função perda, para se chegar ao procedimento. Uma função perda que é utilizada frequentemente é o erro quadrático médio (EQM), embora em algumas ocasiões outros critérios ou funções-perdas sejam mais apropriados” [9].

# Mínimos quadrados

## 3.1 Introdução.

O método de mínimos quadrados (MQ) foi utilizado para obter estimativas com o mínimo erro médio quadrático. Este método tem duas características fundamentais para escolha de seu uso neste trabalho: tem um baixo consumo computacional e permite processar múltiplas séries temporais. Para obter o menor erro de previsão de saída para um processo, faz-se a diferença entre o valor de saída real e o valor estimado. O critério do erro de previsão de saída é “uma alternativa quando o erro não pode ser controlado, e eventualmente tem-se um sinal de ruído” [13].

## 3.2 Método de mínimos quadrados - MQ

Com as definições antes apresentadas, vamos a construir um sistema discreto onde temos que o conjunto de entradas  $u_k$  é  $u_0, u_1, \dots, u_N$ , e o conjunto de saídas  $y_k$  é  $y_0, y_1, \dots, y_N$ .

Então, seja um sistema discreto cuja relação entrada-saída, supondo um erro  $e_k$  no  $k^{\text{ésimo}}$  instante, é:

$$y_k = b_0 u_k + b_1 u_{k-1} + b_2 u_{k-2} + \dots + b_k u_{k-n} + \\ - a_1 y_{k-1} - a_2 y_{k-2} - \dots - a_n y_{k-n} + e_k$$

Para  $N$  medidas de saída temos

$$y_1 = b_0 u_{1-0} + b_1 u_{1-1} + b_2 u_{1-2} + \dots + b_k u_{1-n} + \tag{3.1}$$

$$- a_1 y_{1-1} - a_2 y_{1-2} - \dots - a_n y_{1-n} + e_1 \tag{3.2}$$

$$y_2 = b_0 u_{2-0} + b_1 u_{2-1} + b_2 u_{2-2} + \dots + b_k u_{2-n} + \tag{3.3}$$

$$- a_1 y_{2-1} - a_2 y_{2-2} - \dots - a_n y_{2-n} + e_2 \tag{3.4}$$

$\vdots \quad \vdots \quad \vdots$

$$y_N = b_0 u_{N-0} + b_1 u_{N-1} + b_2 u_{N-2} + \cdots + b_n u_{N-n} + \quad (3.5)$$

$$-a_1 y_{N-1} - a_2 y_{N-2} - \cdots - a_n y_{N-n} + e_N \quad (3.6)$$

onde  $N$  refere-se aos dados atuais.

Fazendo  $Y$  o vetor de  $N$  medidas de saída:

$$Y = [y_1, y_2, \dots, y_N]' \quad (3.7)$$

$\Psi$  a matriz de dados :

$$\Psi = \begin{bmatrix} u_{1-0} & u_{1-1} & \cdots & u_{1-n} & -y_{1-1} & -y_{1-2} & \cdots & y_{1-n} \\ u_{2-0} & u_{2-1} & \cdots & u_{2-n} & -y_{2-1} & -y_{2-2} & \cdots & y_{2-n} \\ \vdots & \vdots & \cdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ u_{N-0} & u_{N-1} & \cdots & u_{N-n} & -y_{N-1} & -y_{N-2} & \cdots & y_{N-n} \end{bmatrix} \quad (3.8)$$

$\Theta$  o vetor de parâmetros a ser estimado:

$$\Theta = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \quad (3.9)$$

$\varepsilon$  o vetor de erros, definido como:

$$\varepsilon = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} \quad (3.10)$$

pode-se reescrever as equações 3.1, 3.3 e 3.5 como:

$$Y = \Psi\Theta + \varepsilon \quad (3.11)$$

Se há erros,  $\varepsilon = Y - \Psi\Theta$  podemos enunciar o problema de mínimos quadrados como:

Dados  $Y$  e  $\Psi$ , determine o estimador  $\hat{\Theta}$  de  $\Theta$  que minimize o critério  $J$ :

$$J = \frac{1}{2} \|\varepsilon\|^2 = \frac{1}{2} (Y - \Psi\Theta)' (Y - \Psi\Theta)$$

Da condição necessária de mínimo,  $\frac{\partial J}{\partial \Theta} = 0$  tem-se para a minimização da forma quadrática:

$$J = (Y - \Psi\Theta)'(Y - \Psi\Theta)$$

$$J = (Y' - (\Psi\Theta)')(Y - \Psi\Theta)$$

$$J = Y'Y - \Theta'\Psi'Y - Y'\Psi\Theta + \Theta'\Psi'\Psi\Theta$$

que a derivada de  $J$  em relação a  $\Theta$  é:

$$\frac{\partial J}{\partial \Theta} = -Y'\Psi - Y'\Psi + 2\Theta'\Psi'\Psi = 0$$

Deste modo podemos escrever a equação normal:

$$\Psi'\Psi\hat{\Theta} = \Psi'Y \quad (3.12)$$

donde

$$\hat{\Theta} = (\Psi'\Psi)^{-1}\Psi'Y \hat{\Theta} \equiv \Psi^\dagger Y \quad (3.13)$$

é o estimador ótimo de mínimos quadrados dos parâmetros do sistema e  $\Psi^\dagger = (\Psi'\Psi)^{-1}\Psi'$  é a matriz pseudo-inversa de  $\Psi$ .

$$\hat{\Theta} \equiv \Psi^\dagger Y \quad (3.14)$$

O desempenho deste método é baseado nas medidas efetuadas, deste modo podemos ter diferentes erros para cada medida efetuada. Para melhorar o desempenho do método usasse um tratamento conhecido como mínimos quadrados ponderado, onde dados melhores ou mais recentes terão uma influência maior do cálculo dos parâmetros; para tanto se define o critério:

$$J(\Theta) = \sum_{k=1}^N w(k)\varepsilon^2(k; \Theta) \quad (3.15)$$

sendo  $w$  o valor da influência de cada medida para o cálculo dos parâmetros.

Neste caso o estimador  $\hat{\Theta}$  de mínimos quadrados é: [13]

$$\hat{\Theta} = (\Psi'w\Psi)^{-1}w\Psi'Y \quad (3.16)$$

### 3.3 Exemplos numéricos com MQ.

#### 3.3.1 Exemplo com modelo simples

Considere um sistema estático em que se deseja obter o modelo mais adequado com às seguintes medidas:

$$t = 1 : \quad \psi(1) = 0; \quad y(1) = 0$$

$$t = 2 : \quad \psi(2) = 1; \quad y(2) = 0,9$$

$$t = 3 : \quad \psi(3) = 2; \quad y(3) = 2,1$$

Considerando um modelo simples:

$$y(t) = \theta_1$$

então temos que:

$$\Psi = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$Y = \begin{bmatrix} 0 \\ 0,9 \\ 2,1 \end{bmatrix}$$

Usando o estimador de mínimos quadrados  $\hat{\Theta} = (\Psi' \Psi)^{-1} \Psi' Y$  e  $w(k) = 1$ , obtemos:

$$\hat{\Theta}_1 = 3^{-1}(0 + 0,9 + 2,1) = 1$$

com custo associado à obtenção desse modelo:

$$J(\hat{\Theta}) = \sum_{k=1}^1 w(k) \varepsilon^2(k; \Theta)$$

$$\varepsilon = Y - \Psi \hat{\Theta}$$

$$\varepsilon = \begin{bmatrix} 0 \\ 0,9 \\ 2,1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot 1$$

$$J(\hat{\Theta}) = \sum_{k=1}^1 w(k) \left( \begin{bmatrix} 0 \\ 0,9 \\ 2,1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot 1 \right)^2$$

$$J(\hat{\Theta}) = 1 + 0,9^2 + 1,1^2 = 2,22$$

Mostrando a tendência polinomial, com a equação 2.10, temos que o estimador de tendência  $T_t$  é:

$$\hat{T}_t = \hat{\theta}_1 = 1 \quad (3.17)$$

Utilizando o modelo  $y(t) = \theta_1$  e o modelo estimado para  $T_t$ , podemos fazer previsões de valores futuros da série com:

$$\hat{y}_1(h) = \hat{T}_{1+h}, \quad h = 1, 2, \dots$$

$$\hat{y}_1(h) = 1$$

Neste caso com o modelo simples, temos o valor previsto  $h$ , passos à frente.

Tabela 3.1: Valores reais e previstos para exemplo modelo simples.

h	y(1)	$\hat{y}(1+h)$	Erro de previsão
1	0,9	1	0,1
2	2,1	1	1,1

### 3.3.2 Exemplo de modelo linear.

Considere um sistema estático em que se deseja obter o modelo mais adequado com às seguintes medidas:

$$t = 1 : \quad \psi(1) = 0; \quad y(1) = 0$$

$$t = 2 : \quad \psi(2) = 1; \quad y(2) = 0,9$$

$$t = 3 : \quad \psi(3) = 2; \quad y(3) = 2,1$$

Considerando um modelo linear:

$$y(t) = \theta_0 + \theta_1 \psi(t)$$

na forma matricial

$$y(t) = [ 1 \quad \psi(t) ] \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix}$$

então temos que:

$$\Psi = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$$

$$Y = \begin{bmatrix} 0 \\ 0,9 \\ 2,1 \end{bmatrix}$$

Usando o estimador de mínimos quadrados  $\hat{\Theta} = (\Psi' \Psi)^{-1} \Psi' Y$  e  $w(k) = 1$ , obtemos:

$$\hat{\Theta} = \left( \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}' \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}' \begin{bmatrix} 0 \\ 0,9 \\ 2,1 \end{bmatrix}$$

$$\hat{\Theta} = \begin{bmatrix} -0,05 \\ 1,05 \end{bmatrix}$$

com custo associado à obtenção desse modelo:

$$J(\hat{\Theta}) = \sum_{k=1}^2 w(k) \varepsilon^2(k; \Theta)$$

$$\varepsilon = Y - \Psi \hat{\Theta}$$

$$\varepsilon = \begin{bmatrix} 0 \\ 0,9 \\ 2,1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} -0,05 \\ 1,05 \end{bmatrix}$$

$$J(\hat{\Theta}) = \sum_{k=1}^2 w(k) \left( \begin{bmatrix} 0 \\ 0,9 \\ 2,1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} -0,05 \\ 1,05 \end{bmatrix} \right)^2$$

$$J(\hat{\Theta}) = 0,05^2 + (-0,1)^2 + 0,05 = 0,0025 + 0,01 + 0,0025 = 0,015$$

Mostrando a tendência polinomial, com a equação 2.10, temos que o estimador de tendência  $T_t$  é:

$$\hat{T}_t = \theta_0 + \theta_1 \psi(t) = -0,05 + 1,05 \psi(t) \tag{3.18}$$

Utilizando o modelo  $y(t) = \theta_0 + \theta_1 \psi(t)$  e o modelo estimado para  $T_t$ , podemos fazer previsões de valores futuros da série com:

$$\hat{y}_2(h) = \hat{T}_{2+h}, \quad h = 1, 2, \dots$$

$$\hat{y}_2(h) = -0,05 + 1,05 \psi(2+h)$$

Neste caso com o modelo lineal, temos o valor previsto  $h$ , passos à frente.

Tabela 3.2: Valores reais e previstos para exemplo modelo linear.

h	y(2+h)	$\hat{y}(2+h)$	Erro de previsão
1	2,1	2,05	0,05

### 3.3.3 Exemplo de modelo não linear.

Considere um sistema estático em que se deseja obter o modelo mais adequado com às seguintes medidas:

$$t = 1 : \quad \psi(1) = 0; \quad y(1) = 0$$

$$t = 2 : \quad \psi(2) = 1; \quad y(2) = 0,9$$

$$t = 3 : \quad \psi(3) = 2; \quad y(3) = 2,1$$

Considerando um modelo não linear:

$$y(t) = \theta_0 + \theta_1\psi(t) + \theta_2\psi^2(t)$$

na forma matricial

$$y(t) = [ 1 \quad \psi(t) \quad \psi^2(t) ] \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{bmatrix}$$

então temos que:

$$\Psi = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix}$$

$$Y = \begin{bmatrix} 0 \\ 0,9 \\ 2,1 \end{bmatrix}$$

Usando o estimador de mínimos quadrados  $\hat{\Theta} = (\Psi' \Psi)^{-1} \Psi' Y$  e  $w(k) = 1$ , obtemos:

$$\hat{\Theta} = \left( \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix}' \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix}' \begin{bmatrix} 0 \\ 0,9 \\ 2,1 \end{bmatrix}$$

$$\hat{\Theta}_0 = \begin{bmatrix} 0 \\ 0,75 \\ 0,15 \end{bmatrix}$$

com custo associado à obtenção desse modelo:

$$J(\hat{\Theta}) = \sum_{k=1}^2 w(k) \varepsilon^2(k; \Theta)$$

$$\varepsilon = Y - \Psi \hat{\Theta}$$

$$\varepsilon = \begin{bmatrix} 0 \\ 0,9 \\ 2,1 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 0 \\ 0,75 \\ 0,15 \end{bmatrix}$$

$$J(\hat{\Theta}) = \sum_{k=1}^3 w(k) \left( \begin{bmatrix} 0 \\ 0,9 \\ 2,1 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 0 \\ 0,75 \\ 0,15 \end{bmatrix} \right)^2$$

$$J(\hat{\Theta}) = 0 + 0 + 0 = 0$$

A condição de custo zero ocorre quando o número de medidas é igual ao número de parâmetros a serem estimados e a matriz  $\Psi$  é não singular.

Mostrando a tendência polinomial, com a equação 2.10, temos que o estimador de tendência  $T_t$  é:

$$\hat{T}_t = \theta_0 + \theta_1\psi(t) + \theta_2\psi^2(t) = 0,75\psi(t) + 0,15\psi^2(t) \quad (3.19)$$

Utilizando o modelo  $y(t) = \theta_0 + \theta_1\psi(t) + \theta_2\psi^2(t)$  e o modelo estimado para  $T_t$ , podemos fazer previsões de valores futuros da série com:

$$\hat{y}_3(h) = \hat{T}_{3+h}, \quad h = 1, 2, \dots$$

$$\hat{y}_3(h) = 0 + 0,75\psi(3+h) + 0,15\psi^2(3+h)$$

Neste caso com o modelo não linear, temos o valor previsto  $h$ , passos à frente.

# Método de mínimos quadrados recursivo e múltiplas séries temporais.

## 4.1 Introdução.

Neste capítulo serão desenvolvidos os conceitos de recursividade e análise multivariada de dados para o método de mínimos quadrados apresentado no capítulo anterior. Lembrando que o método de mínimos quadrados visto, faz um tratamento em bloco das observações efetuadas (equação 3.13) gerando um custo computacional alto na hora de efetuar a inversão da matriz.

## 4.2 Análise multivariada de dados

No primeiro lugar vamos a enunciar uma definição básica de análise multivariada, donde refere-se: todos os métodos estatísticos que simultaneamente analisam múltiplas medidas sobre cada individuo ou objeto sob investigação. Qualquer análise simultânea de mais de duas variáveis de certo modo pode ser considerada análise multivariada [5].

Os seguintes conceitos são básicos para a análise multivariada:

- Variável estatística.
- Escalas de medida não métrica
- Escalas de medida métrica
- Erro de medida e medida multivariada

Na análise multivariada de dados temos diferentes tipos de técnicas, que dependendo de nosso objetivo e do tipo de dados podemos classificá-las, exaltamos as seguintes técnicas:

- Análise de componentes principais e análise dos fatores comuns.
- Regressão múltipla.
- Análise discriminante múltipla.

- Análise multivariada de variância e covariância.
- Análise conjunta.
- Correlação canônica.
- Análise de agrupamentos.
- Escalonamento multidimensional.
- Análise de correspondência.
- Modelos lineares de probabilidade.
- Modelagem de equações estruturais.

Com a análise multivariada temos que conseguir uma significância prática, conhecer os dados, examinar os erros e validar os dados[5].

Finalmente, enunciamos um processo de seis passos para construir modelos multivariados. Onde o resultado final do modelo desenvolvido depende de um conjunto de critérios para ter em conta. Os passos são os seguintes:

1. Definição do problema, dos objetivos e da técnica multivariada.
2. Desenvolvimento do plano de análise.
3. Avaliação das suposições inerentes à técnica multivariada.
4. Estimção do modelo multivariado e avaliação do ajuste geral do modelo.
5. Interpretação das variáveis estatísticas.
6. Validação do modelo multivariado.

### 4.3 Método de mínimos quadrados recursivo - MQR.

Nesta seção vamos trabalhar o método de mínimos quadrados levando só em consideração a chegada da nova observação, para assim poder atualizar os parâmetros estimados. Chamado de método de mínimos quadrados recursivo e apresentado com mais detalhe em [16] e [17].

As equações que apresentamos a seguir mostram como tornar recursiva a equação 3.13. Por exemplo, se desejamos estimar o valor de uma constante:

$$y_k = b \quad (4.1)$$

de forma recursiva, podemos efetuar uma série de medidas. Neste caso tem-se:

$$\psi_k = 1 \quad e \quad \Theta = b \quad (4.2)$$

donde

$$\Psi_k = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad (4.3)$$

Fazendo as substituições necessárias na equação 3.13 temos:

$$\hat{\Theta} = \frac{1}{N} [y_1 + y_2 + \dots + y_N] \quad (4.4)$$

Então podemos generalizar o exemplo acima para:

$$\hat{\Theta} = \frac{1}{k} \sum_{s=1}^k y_s \quad (4.5)$$

Construimos agora, um algoritmo recursivo a partir da equação anterior 4.5 do seguinte modo:

$$\hat{\Theta}_k = \frac{1}{k} \left[ \sum_{s=1}^{k-1} y_s + y_k \right] \quad (4.6)$$

$$\hat{\Theta}_k = \frac{1}{k} \left[ (k-1)\hat{\Theta}_{k-1} + y_k \right] \quad (4.7)$$

$$\hat{\Theta}_k = \hat{\Theta}_{k-1} + \frac{1}{k} \left[ y_k - \hat{\Theta}_{k-1} \right] \quad (4.8)$$

Por tanto a estimativa de  $\hat{\Theta}_k$  no  $k$ -ésimo instante é dependente de  $\hat{\Theta}_{k-1}$  e de um termo de correção. Note que neste caso o fator  $\frac{1}{k}$  pode ser visto como um fator de esquecimento pois quanto maior for  $k$  menor será o efeito do termo de correção. Para o exemplo fazemos :

$$P_k = \Psi' \Psi = \frac{1}{k} \quad (4.9)$$

onde,

$$\Psi' = (1, 1, 1, \dots, 1) \quad e \quad k = 1, 2, \dots, N \quad (4.10)$$

ou seja

$$P_k^{-1} = k = P_{k-1}^{-1} + 1 \quad (4.11)$$

logo

$$P_k = \frac{1}{P_{k-1}^{-1} + 1} = \frac{P_{k-1}}{1 + P_{k-1}} \quad (4.12)$$

Para uma discussão formal do procedimento de mínimos quadrados recursivos reescrevamos a equação 3.13, como:

$$\hat{\Theta}_k = (\Psi'_k \Psi_k)^{-1} \Psi'_k Y_k \quad (4.13)$$

Quando da adição de novas medidas,  $\Psi_{k+1}$  pode ser reescrita como:

$$\Psi_{k+1} = \begin{bmatrix} \Psi_k \\ \psi_{k+1} \end{bmatrix}, \quad Y_{k+1} = \begin{bmatrix} Y_k \\ y_{k+1} \end{bmatrix} \quad (4.14)$$

onde  $\psi_{k+1}$  é o vetor com as novas medidas de entradas e saídas no sistema.

A estimativa de  $\hat{\Theta}_{k+1}$  pode ser reescrita como:

$$\hat{\Theta}_{k+1} = (\Psi'_{k+1} \Psi_{k+1})^{-1} \Psi'_{k+1} Y_{k+1} \quad (4.15)$$

$$\hat{\Theta}_{k+1} = \left( \begin{bmatrix} \Psi'_k \\ \psi'_{k+1} \end{bmatrix}' \begin{bmatrix} \Psi_{k+1} \\ \psi_{k+1} \end{bmatrix} \right)^{-1} \begin{bmatrix} \Psi'_k \\ \psi'_{k+1} \end{bmatrix}' \begin{bmatrix} Y_k \\ y_{k+1} \end{bmatrix} \quad (4.16)$$

que pode ser reescrita assim:

$$\hat{\Theta}_{k+1} = \left[ \sum_{s=1}^{k+1} \psi_s \psi'_s \right]^{-1} \left[ \sum_{s=1}^{k+1} \psi_s y_s \right] \quad (4.17)$$

Introduzindo a notação:

$$P_{k+1} = \left[ \sum_{s=1}^{k+1} \psi_s \psi'_s \right]^{-1} \quad (4.18)$$

podemos reescrever a equação 4.18, como:

$$P_{k+1}^{-1} = P_k^{-1} + \psi_{k+1} \psi'_{k+1} \quad (4.19)$$

Substituindo em 4.17, 4.18 e 4.19 temos:

$$\hat{\Theta}_{k+1} = \hat{\Theta}_k + P_{k+1} \psi_{k+1} \left[ y_{k+1} - \psi'_{k+1} \hat{\Theta}_k \right] \quad (4.20)$$

Assim podemos escrever:

$$K_{k+1} = P_{k+1}\psi_{k+1} \quad (4.21)$$

$$\varepsilon_{k+1} = \left[ y_{k+1} - \psi'_{k+1}\hat{\Theta}_k \right] \quad (4.22)$$

$$\hat{\Theta}_{k+1} = \hat{\Theta}_k + K_{k+1}\varepsilon_{k+1} \quad (4.23)$$

O termo  $\varepsilon_{k+1}$  pode ser interpretado como um preditor de erro de saída, isto é a diferença entre a saída  $y_{k+1}$  e o valor estimado desta saída dado por  $\psi'_{k+1}\hat{\Theta}_k$ . O vetor  $K_{k+1}$  pode ser interpretado como um fator de compensação ou ganho das medidas [18]. Para o cálculo da equação 4.21 precisamos na equação 4.19 invertir uma matriz a cada nova medida; este procedimento poderá consumir um tempo de processamento desnecessário; para melhorar isto vamos a tomar o resultado desenvolvido com o lema de inversão matricial dado por [19], então temos:

$$P_{k+1} = P_k + \frac{P_k\psi_{k+1}\psi'_{k+1}P_k}{1 + \psi'_{k+1}P_k\psi_{k+1}} \quad (4.24)$$

Então podemos reescrever  $K_{k+1}$  dado na equação 4.21, como:

$$K_{k+1} = \left[ P_k + \frac{P_k\psi_{k+1}\psi'_{k+1}P_k}{1 + \psi'_{k+1}P_k\psi_{k+1}} \right] \psi_{k+1} \quad (4.25)$$

$$K_{k+1} = \frac{P_k\psi_{k+1}}{1 + \psi'_{k+1}P_k\psi_{k+1}} \quad (4.26)$$

O método de mínimos quadrados recursivo esta completo. [13]

Nos próximos capítulo iremos apresentar a aplicação deste algoritmo para múltiplas séries temporais.



# Aplicação do método MQR a múltiplas séries temporais.

## 5.1 Introdução

Tendo estudado nos capítulos anteriores o método de mínimos quadrados recursivo e os conceitos de modelagem de dados e séries temporais, damos início ao trabalho específico de construir nosso modelo de dados usando como exemplo séries temporais relacionadas com as exportações de café na Colômbia [20]. Neste capítulo iremos observar a aplicação do método MQR em diferentes modelos de dados, procurando achar o melhor estimador. Os resultados gerados no final deste capítulo pelo método MQR, servirão para realizar comparações com os resultados obtidos pelo método MQR3 do próximo capítulo.

Para iniciar vamos apresentar a totalidade das múltiplas séries temporais que usaremos na aplicação. As séries de entrada são as seguintes:

- Produção de café verde Colombiano, medida em milhares de sacos de café.
- Câmbio de Dólar a peso Colombiano.
- Exportações de petróleo e derivados em milhões de dólares (FOB).
- Exportações de café em milhões de dólares (FOB).
- Reservas Internacionais Líquidas em milhões de dólares.

A série de temporal de saída é:

- Volume de exportações de café verde Colombiano, em milhares de sacos de 60 Kg.

As séries temporais apresentadas têm um total de 497 dados cada uma, onde cada valor representa o resultado de um mês [21]. A quantidade de dados mencionados inclui os valores para validação do método (para os dados de validação são usados os últimos 15 meses de cada série temporal). Assim nossa matriz de dados tem 482 observações por cada série temporal [22].

### 5.1.1 MQR aplicado nas exportações de café da Colômbia.

O comportamento de cada uma das séries temporais é fundamental em nossa análise, então mostramos graficamente cada uma das sequências de dados ao longo do tempo. Uma análise das informações contidas nas séries temporais foi feita no capítulo dois. Apresentamos as figuras 5.1, 5.2 e 5.3 que incluem a totalidade dos dados das séries temporais tratadas neste capítulo.

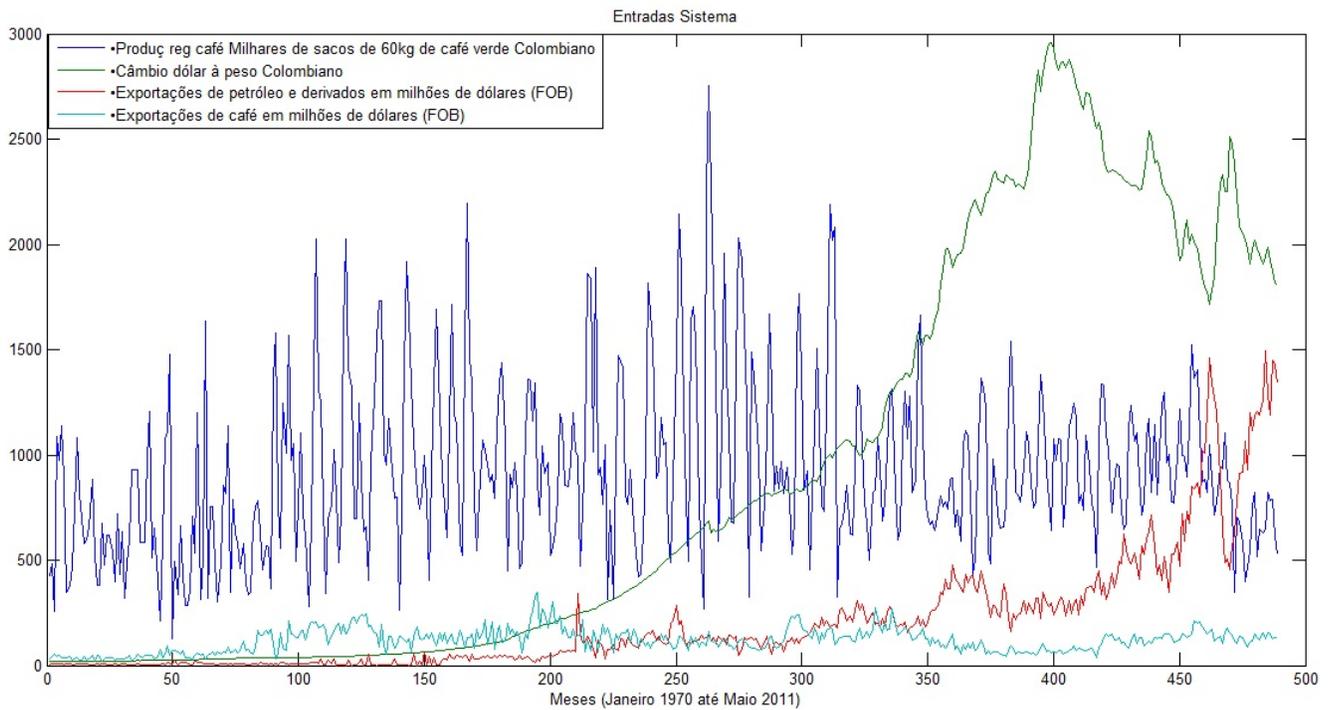


Figura 5.1: Séries temporais de entrada (janeiro do ano 1970 até maio do ano 2011)

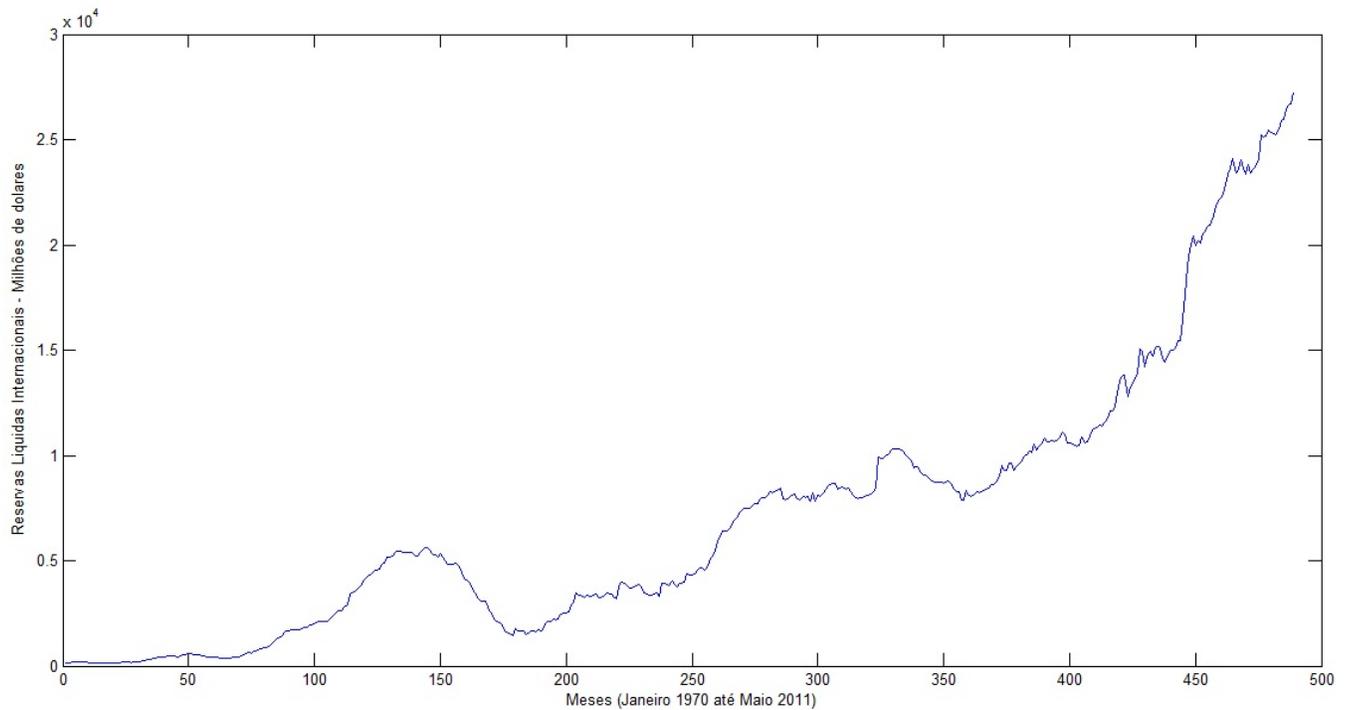


Figura 5.2: Série temporal de entrada (janeiro do ano 1970 até maio do ano 2011)

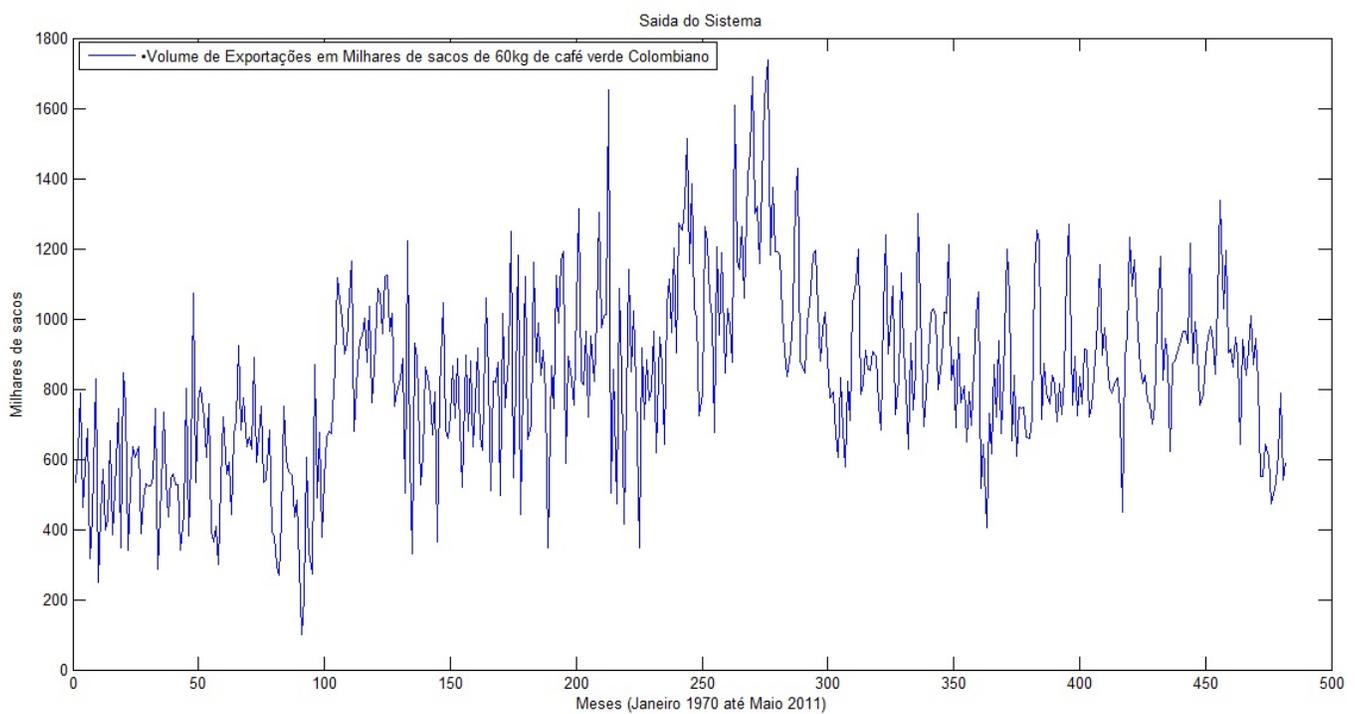


Figura 5.3: Série temporal de saída (janeiro do ano 1970 até maio do ano 2011).

Para aplicar o método MQR nestas séries temporais precisamos ter uma estrutura do modelo de dados, a qual se pode obter relacionando os dados de entrada com os dados de saída:  $n$  passos à frente (figura 5.5), ou com  $n$  passos para atrás ou relacionando a entrada de um mês com a saída do mesmo mês  $n = 0$  (figura 5.4). A seleção do modelo de dados depende do objetivo; vamos apresentar duas possíveis relações entre os dados de entrada e saída nas figuras 5.4 e 5.5. Nesta tese usaremos a estrutura de modelo de dados,  $n = 0$  passos para frente, figura 5.4.

As figuras 5.4 e 5.5 mostram: ao lado esquerdo o modelo de dados com as séries temporais de entrada e saída, na metade representamos a aplicação do método MQR no modelo de dados para estimar a série temporal de saída e do lado esquerdo da figura mostramos a relação dos dados de entrada com as saídas estimadas.

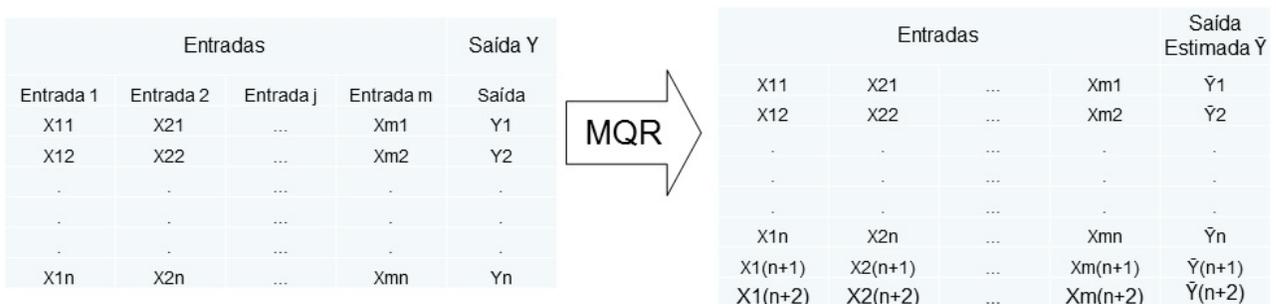


Figura 5.4: Estrutura do Modelo de dados  $n=0$  passos para frente.

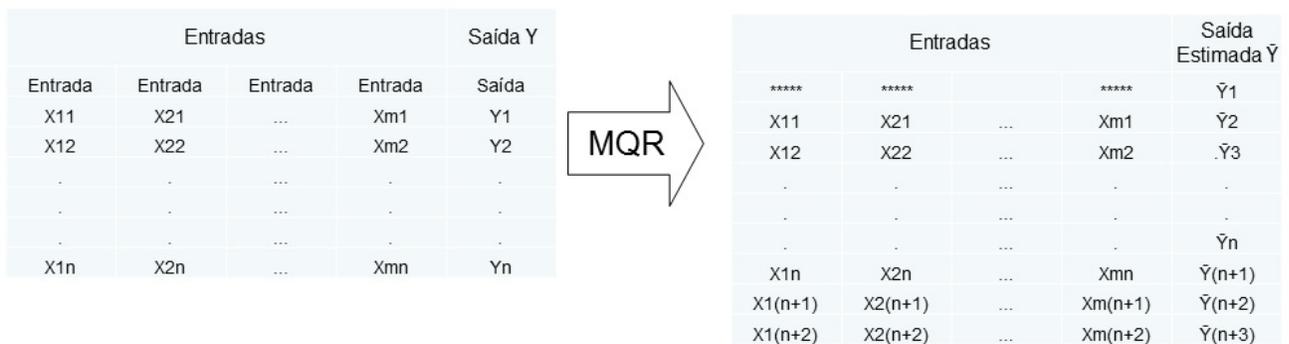


Figura 5.5: Estrutura do Modelo de dados  $n=1$  passos para frente.

A escolha do modelo de dados para  $n = 0$  apresentado na figura 5.4 para o exemplo do café, é feita porque os resultados obtidos, com o método MQR nestas estruturas de modelos de dados, foram suficientes em comparação com as obtidas para outros valores de  $n$ . A estrutura do modelo de dados da figura 5.4 com  $n = 0$  prevê os valores da saída de um mês para os dados de entrada nesse mesmo mês, caso diferente no modelo de dados da figura 5.5, que prevê os dados da saída de um mês a partir dos dados de entrada do mês anterior. Os dois modelos de dados são válidos, só que no primeiro temos melhores previsões. Nas figuras 5.6 e 5.7 pode-se observar como o erro de saída, depois de aplicar o método MQR nas duas estruturas de modelos de dados apresentadas, é menor para o modelo de dados com  $n = 0$  passos para frente.

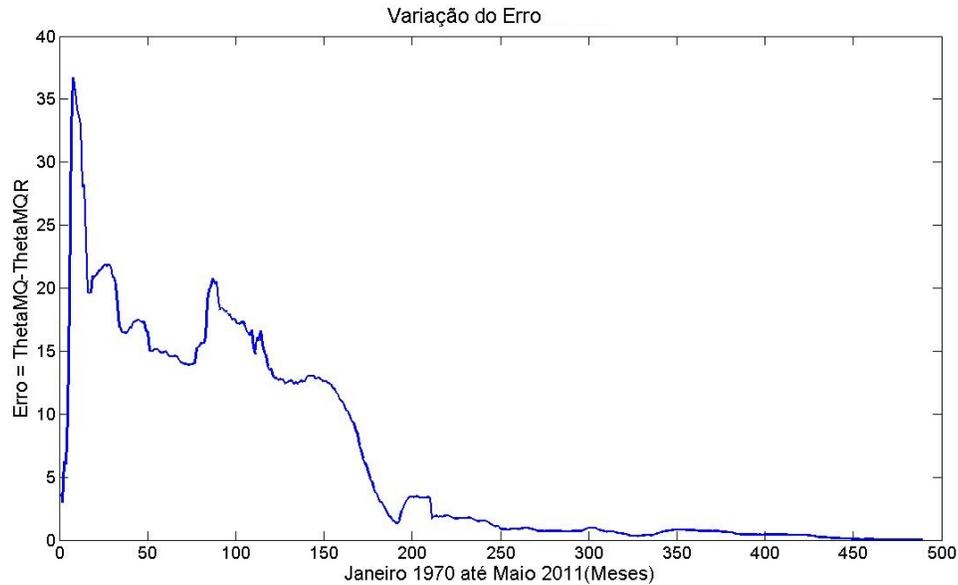


Figura 5.6: Erro de saída com modelo de dados  $n=0$  passos para frente.

O resultado para o erro médio de estimação da saída aplicando o método MQR (equação 4.22) com o modelo de dados  $n = 0$  passos para frente da figura 5.6, as cinco séries temporais de entrada e à série de saída foi de 11,32% para os últimos 100 dados.

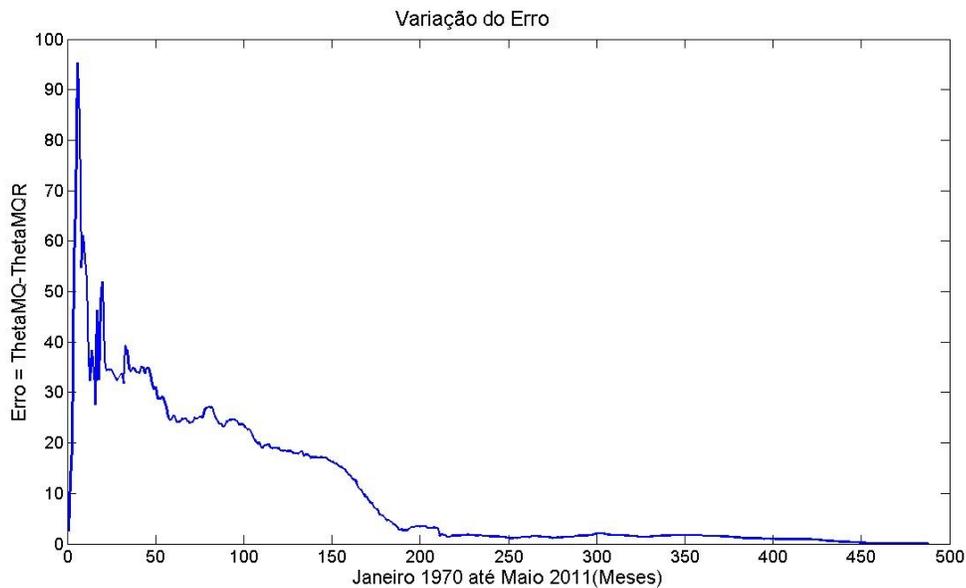


Figura 5.7: Erro de saída com modelo de dados  $n=1$  passos para frente.

O resultado para o erro médio de estimação da saída aplicando o método MQR (equação 4.22) com o modelo de dados  $n = 1$  passos para frente da figura 5.7, as cinco séries temporais de entrada e à série de saída foi de 20,61% para os últimos 100 dados.

Para representar a série temporal de saída fazemos seis diferentes combinações das series

temporais de entrada e analisamos os resultados da aplicação do MQR a cada uma destas combinações. Obtendo o modelo de dados que melhor representa nossa saída. Na tabela 5.1.1 e na tabela 5.2 podemos encontrar os resultados das seis combinações feitas com as séries temporais de entrada apresentadas anteriormente, as quais também mostram a percentagem do erro médio na saída dos últimos 100 dados e a quantidade de dados usados.

Tabela 5.1: Modelos de dados com diferentes séries temporais de entrada

Modelo de Dados N°	1	2	3
Número de entradas	5	4	3
Entradas	<ul style="list-style-type: none"> <li>• Produção de café verde Colombiano, medida em milhares de sacos de café.</li> <li>• Câmbio de Dólar a peso Colombiano.</li> <li>• Exportações de petróleo e derivados em milhões de dólares (FOB).</li> <li>• Exportações de café em milhões de dólares (FOB).</li> <li>• Reservas Internacionais Líquidas em milhões de dólares.</li> </ul>	<ul style="list-style-type: none"> <li>• Produção de café verde Colombiano, medida em milhares de sacos de café.</li> <li>• Câmbio de Dólar a peso Colombiano.</li> <li>• Exportações de café em milhões de dólares (FOB).</li> <li>• Reservas Internacionais Líquidas em milhões de dólares.</li> </ul>	<ul style="list-style-type: none"> <li>• Produção de café verde Colombiano, medida em milhares de sacos de café.</li> <li>• Câmbio de Dólar a peso Colombiano.</li> <li>• Exportações de café em milhões de dólares (FOB).</li> </ul>
% Erro Médio de saída para os últimos 100 dados	11,32%	13,18%	32,15%
N° de Dados de entrada	2445	1956	1467

Tabela 5.2: Modelos de dados com diferentes séries temporais de entrada (2).

Modelo de Dados N°	4	5	6
Número de entradas	2	2	1
Entradas	<ul style="list-style-type: none"> <li>• Produção de café verde Colombiano, medida em milhares de sacos de café.</li> <li>• Câmbio de Dólar a peso Colombiano.</li> </ul>	<ul style="list-style-type: none"> <li>• Produção de café verde Colombiano, medida em milhares de sacos de café.</li> <li>• Exportações de café em milhões de dólares (FOB)</li> </ul>	<ul style="list-style-type: none"> <li>• Produção de café verde Colombiano, medida em milhares de sacos de café.</li> </ul>
% Erro Médio de saída para os últimos 100 dados	15,74%	16,66%	17,44%
N° de Dados de entrada	978	978	489

Das tabelas 5.1.1 e 5.2, vemos que os três modelos de dados que tiveram menor percentagem de erro médio de saída para os últimos 100 dados foram os de números 1, 2 e 4 respectivamente. Agora para saber qual dos três modelos nos vamos escolher, aplicamos o critério da melhor relação entre o “% Erro Médio de saída para os últimos 100 dados” e a quantidade de “Dados de entrada”. Apresentamos os resultados na tabela 5.3.

Tabela 5.3: Critério: % Erro Médio multiplicado pelos Dados de entrada.

Modelo de Dados N°	1	2	4
% Erro Médio de saída para os últimos 100 dados	11,32%	13,18%	15,74%
N° de Dados de entrada	2445	1956	978
% Erro Médio x Dados de entrada	276,77	257,80	153,93

O melhor modelo de dados para nosso caso, depois de aplicar o método MQR é o de número 4. Percebemos nos resultados da Tabela 5.3 uma diferença bem significativa para o melhor modelo de dados.

Finalmente, o modelo escolhido nesta aplicação para a previsão do erro médio de saída final com o método MQR tem duas séries temporais de entrada e uma de saída, com 497 dados em total por cada série temporal.

Portanto as séries temporais de entrada e saída que utilizaremos neste trabalho são:

Entrada:

- Produção de café verde Colombiano, medida em milhares de sacos de café.
- Câmbio de dólar a peso colombiano.

Saída:

- Volume de exportações de café verde Colombiano, em milhares de sacos de 60 kg.

com a estrutura de modelo de dados para as séries temporais apresentada na tabela 5.4.

Nosso modelo de dados, que é fundamental para este trabalho, é apresentado na tabela 5.4:

Tabela 5.4: Modelo de dados.

Descrição	Entrada ( $M1_{482 \times 1}$ )	Entrada ( $M1_{482 \times 2}$ )	Saída ( $M1_{482 \times 3}$ )
	Produção de café verde Colombiano, medida em milhares de sacos de café.	Câmbio de Dolar a peso Colombiano	Volume de exportações de café verde Colombiano, em milhares de sacos de 60 Kg
N° Dados para o modelo.	482	482	482
N° Dados de validação.	15	15	15
N° Total de Dados.	497	497	497

## 5.2 Validação do método MQR.

Nesta seção vamos tomar os resultados obtidos com nosso procedimento: a estrutura do modelo de dados da figura 5.4 e a matriz final de dados da tabela 5.4, para validar o método MQR. Os 45 dados de validação do modelo de dados da figura 5.4, terão a finalidade de mostrar estatísticas nos erros de previsão. Para o exemplo das exportações de café da Colômbia os dados de validação representam um ano e três meses de exportações futuras. Na tabela 5.5 apresentamos os valores das séries temporais para validação da série temporal de saída.

Tabela 5.5: Dados de validação.

Mês	Entrada	Entrada	Saída
	Produção de café verde Colombiano, medida em milhares de sacos de café.	Câmbio de Dolar a peso Colombiano	Volume de exportações de café verde Colombiano, em milhares de sacos de 60 Kg
Mar-10	629	1909,1	576
Abr-10	647	1940,36	588
Mai-10	822	1984,36	576
Jun-10	780	1925,9	648
Jul-10	787	1874,52	612
Ago-10	615	1819,06	599
Set-10	530	1805,6	530
Out-10	807	1808,46	632
Nov-10	979	1863,67	827
Dic-10	1164	1925,86	1105
Jan-11	908	1866,64	848
Fev-11	764	1882,61	651
Mar-11	779	1884,38	924
Abr-11	523	1812,77	580
Mai-11	673	1801,65	593

O passo seguinte é aplicar o método MQR no modelo de dados da tabela 5.4. Depois de ter armazenado e processado todos os dados, encontramos  $\theta_1 MQR$  final que é o melhor estimador da saída:

$$\theta_1 MQR = \begin{bmatrix} 0,6965 \\ 0,1138 \end{bmatrix} \quad (5.1)$$

Com o método MQR aplicado ao modelo e conseguindo o melhor estimador em 5.1, podemos observar o comportamento da previsão de saída durante o processamento da matriz de dados (Tabela 5.4). A figura 5.8 mostra como o comportamento da previsão de saída melhora quando o número de dados é aumentado.

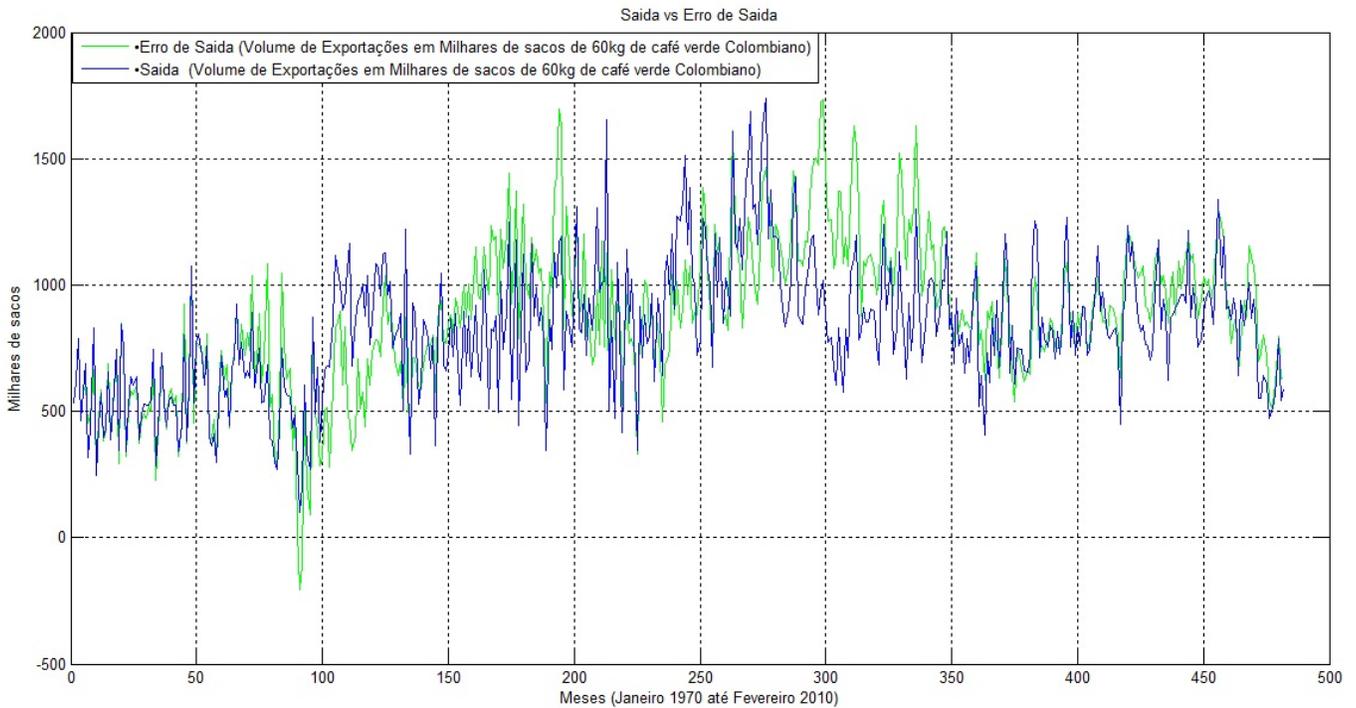


Figura 5.8: Série temporal de Saída x Série temporal Estimada. (janeiro do ano 1970 até fevereiro do ano 2010).

Uma observação importante para ter em conta na análise dos resultados de nossa aplicação do método MQR esta implícita na figura 5.8: o estimador  $\theta_1 MQR$  com cada iteração apresenta uma boa convergência para o estimador obtido com método de mínimos quadrados não recursivo ( $\theta MQ$  equação 3.13). Na figura 5.9 ilustra-se a convergencia do estimador  $\theta_1 MQR$ .

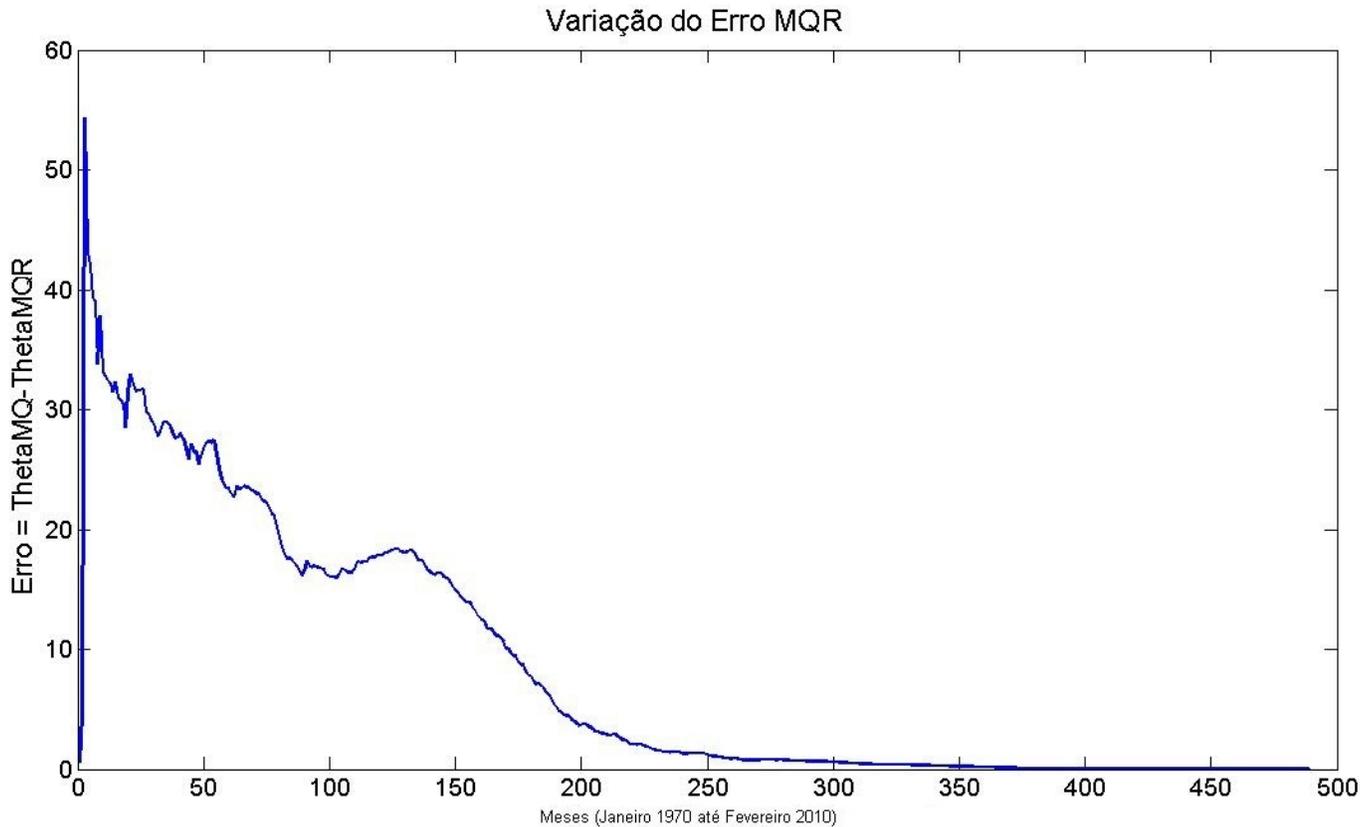


Figura 5.9: Convergência de  $\theta_1 MQR$  a  $\theta MQ$

Para iniciar nosso processo de validação com o estimador da equação 5.1, usamos os primeiros dados de validação indicados na tabela 5.6. Os dados de validação de entrada multiplicados com o estimador  $\theta_1 MQR$  da equação 5.1 geram a primeira previsão, equação 5.2.

Tabela 5.6: Primeiros dados de validação.

Mês	Entrada	Entrada	Saída
	Produção de café verde Colombiano, medida em milhares de sacos de cafe.	Câmbio de Dolar a peso Colombiano.	Volume de exportações de cafe verde Colombiano, em milhares de sacos de 60 Kg.
Mar-10	629	1909,1	576

$$PrevisãoMQR = [ 629 \quad 1909.1 ] * \theta_1 MQR$$

$$PrevisãoMQR = [ 629 \quad 1909.1 ] * \begin{bmatrix} 0.6965 \\ 0.1138 \end{bmatrix}$$

$$PrevisãoMQR = 655,38 \quad (5.2)$$

O resultado obtido na equação 5.2 é fundamental para comparações futuras. Vamos apresentar as 15 previsões feitas com todos os dados de validação da tabela 5.5 e o  $\theta_1MQR$  da equação 5.1, resultados obtidos do mesmo modo que a primeira previsão na equação 5.2. Observa-se nas tabelas 5.7 e 5.8 os valores de previsão obtidos.

Tabela 5.7: Resultados de previsão com MQR (Tabela 1).

Mês	Previsão MQR
Mar-10	655,4
Abr-10	671,5
Mai-10	798,4
Jun-10	762,5
Jul-10	761,5
Ago-10	635,4
Set-10	574,7

Tabela 5.8: Resultados de previsão com MQR (Tabela 2).

Mês	Previsão MQR
Out-10	767,9
Nov-10	894
Dic-10	1029,9
Jan-11	844,9
Fev-11	746,4
Mar-11	757
Abr-11	570,6
Mai-11	673,8

Com os resultados de validação das tabelas 5.7 e 5.8 podemos obter os erros de saída nas previsões para cada mês, o qual é feito calculando a percentagem de erro entre as previsões e as saídas reais da tabela 5.5 para cada mês. Mostra-se nas tabelas 5.9 e 5.10 as percentagens de erro de saída nas previsões para cada mês.

Tabela 5.9: Erro de previsão com MQR (Tabela 1).

Mês	% Erro Previsão MQR
Mar-10	13,78%
Abr-10	14,20%
Mai-10	38,61%
Jun-10	17,67%
Jul-10	24,43%
Ago-10	6,07%
Set-10	8,43%

Tabela 5.10: Erro de previsão com MQR (Tabela 2).

Mês	% Erro Previsão MQR
Out-10	21,50%
Nov-10	8,10%
Dic-10	6,79%
Jan-11	0,37%
Fev-11	14,65%
Mar-11	18,07%
Abr-11	1,62%
Mai-11	13,63%

Para concluir o capítulo apresentamos a percentagem do erro médio de previsão para todos os dados de validação na tabela 5.11.

Tabela 5.11: Percentagem do erro médio de previsão com MQR.

% Erro Médio de Previsão MQR	13,86%
------------------------------	--------

O resultado final da percentagem do erro médio de previsão depois de validar o método MQR com as estimativas das exportações de café na Colômbia completa o objetivo deste capítulo. Realizamos comparações nos próximos capítulos com os resultados de outros métodos.



## Contribuição computacional para o método de mínimos quadrados recursivo - MQR3

### 6.1 Introdução

Procurando obter uma melhor estimativa com o método MQR e executando o mesmo modelo de dados de séries temporais do capítulo anterior, apresentamos um procedimento que pode melhorar as estimativas. O procedimento MQR3 nasceu depois de aplicar diversas metodologias nos dados e de realizar uma análise rigorosa nas séries temporais, para assim melhorar o erro médio de saída nas estimativas.

### 6.2 Diagrama implementado para o método de mínimos quadrados recursivo - MQR

As equações do método MQR usual já foram apresentadas e exemplificadas nos capítulos anteriores, agora para um melhor entendimento do mesmo mostramos na figura 6.1 um possível diagrama do algoritmo com as equações 4.26, 4.22 e 4.23. Definimos os vetores de sinais de entrada  $u_{k+1}$ , saída  $y_{k+1}$ , estimador de MQR  $\hat{\theta}_{1k}$ , erro de previsão  $e_{k+1}$  e a matriz de dados  $M1$  de dimensão  $s \times d$ .

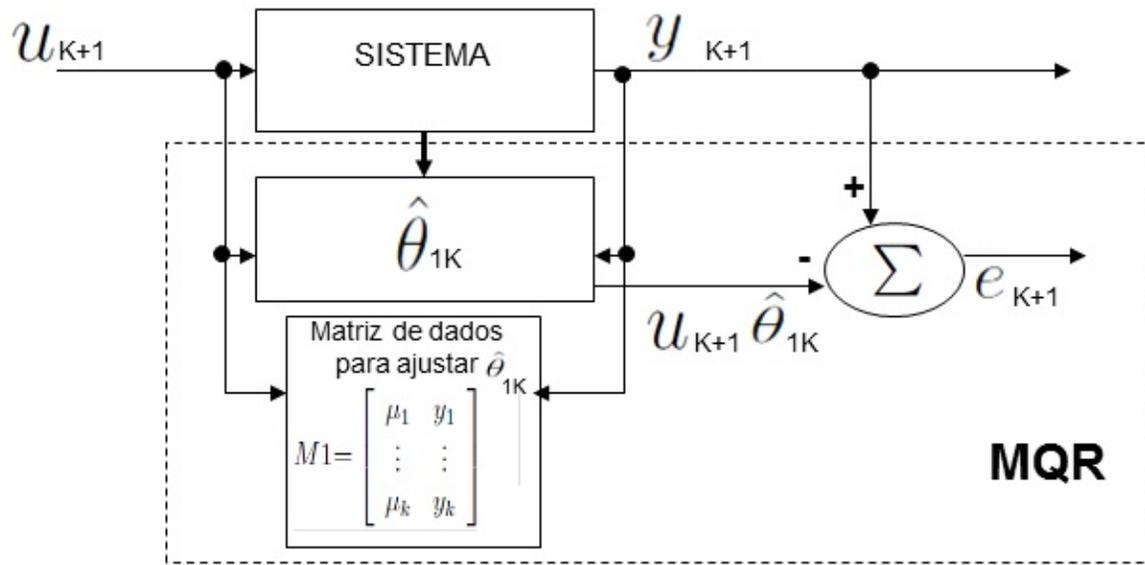


Figura 6.1: Diagrama para o método MQR.

$M1_{s \times d}$ : é a matriz de dados de entrada e saída com um número de observações  $s$  e séries temporais  $d$ , que garante a convergência do preditor de erro de saída apresentado na equação 4.22. Para não ter problemas de armazenamento, a matriz  $M1$  guarda cada nova observação e apaga o registro mais velho. Um possível exemplo de aplicação da matriz  $M1$ , é quando desejamos obter previsões de um sistema em tempo real, e não podemos aguardar para que o método ajuste o estimador  $\hat{\theta}_{1k}$  com cada nova observação; então usamos as observações passadas já armazenadas para ajustar rapidamente  $\hat{\theta}_{1k}$ .

$\hat{\theta}_{1k}$ : é o estimador do método MQR apresentado nas equações 4.26, 4.22 e 4.23, gerado a partir da matriz de dados  $M1$  e da nova observação.

$e_{k+1}$ : é o preditor de erro de saída apresentado na equação 4.22 do método de mínimos quadrados recursivo.

### 6.3 Diagrama implementado para o método de mínimos quadrados recursivo 3 - MQR3

Mostramos através de um diagrama, qual é o procedimento para aplicar o método MQR, juntamente com a contribuição computacional desenvolvida para melhorar o erro médio de saída das estimativas para um sistema com múltiplas séries temporais. Ilustramos e descrevemos a continuação o procedimento do método MQR3.

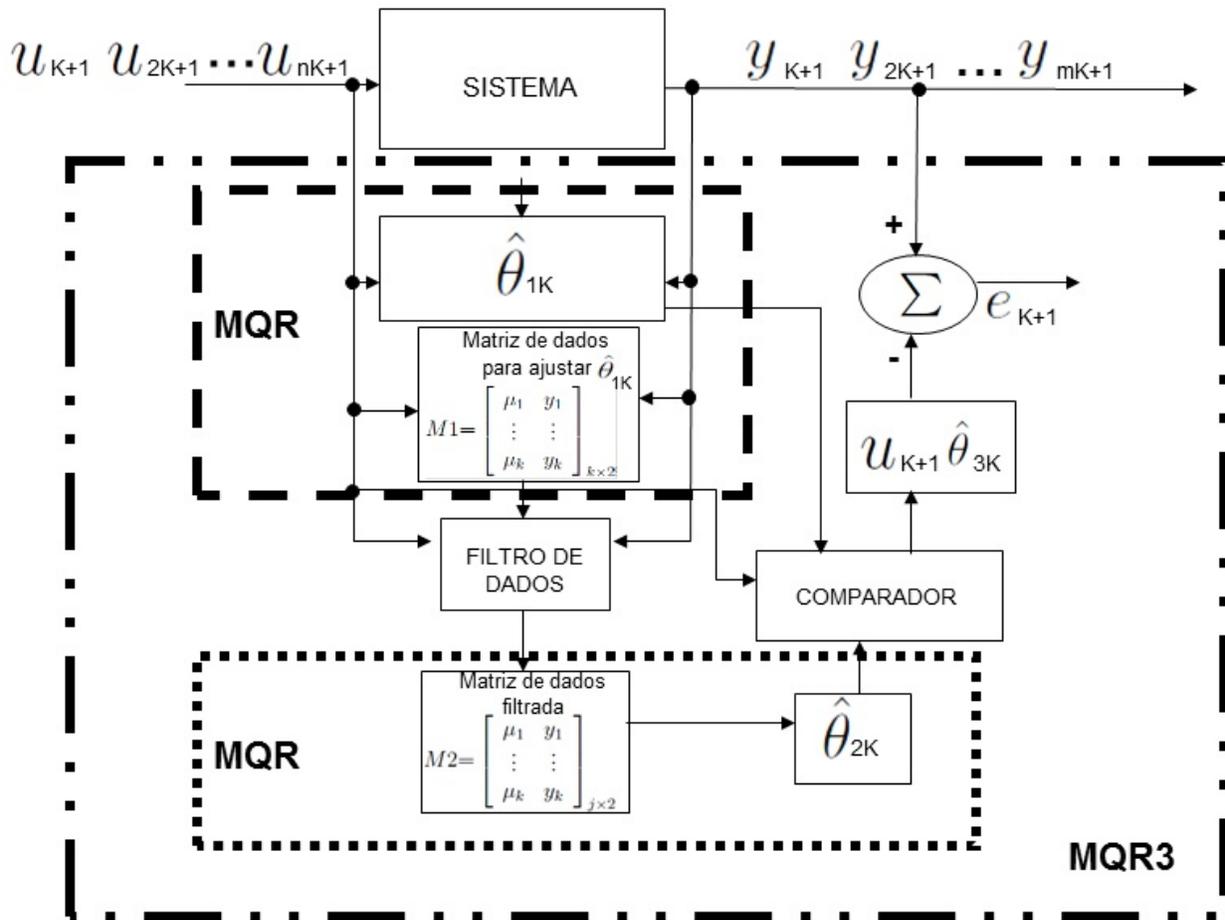


Figura 6.2: Diagrama geral do método MQR3.

Na figura 6.2 mostra-se de forma geral e marcado com linhas descontínuas, como o método MQR3 é composto por dois procedimentos principais baseados na execução do método MQR em duas matrizes de dados diferentes do mesmo modelo, que geram assim, dois estimadores associados. Observamos outros elementos no diagrama que também fazem parte fundamental do método MQR3 e que serão descritos a continuação.

A figura 6.3 mostra o diagrama completo para executar o método MQR3 proposto. Adicionamos os seguintes blocos em comparação com o diagrama anterior da figura 6.1: o segundo estimador por MQR  $\hat{\theta}_{2k}$ , Filtro de Dados, a Matriz de dados  $M2_{g \times d}$ , o Comparador de estimadores e o terceiro estimador por comparações  $\hat{\theta}_{3k}$ .

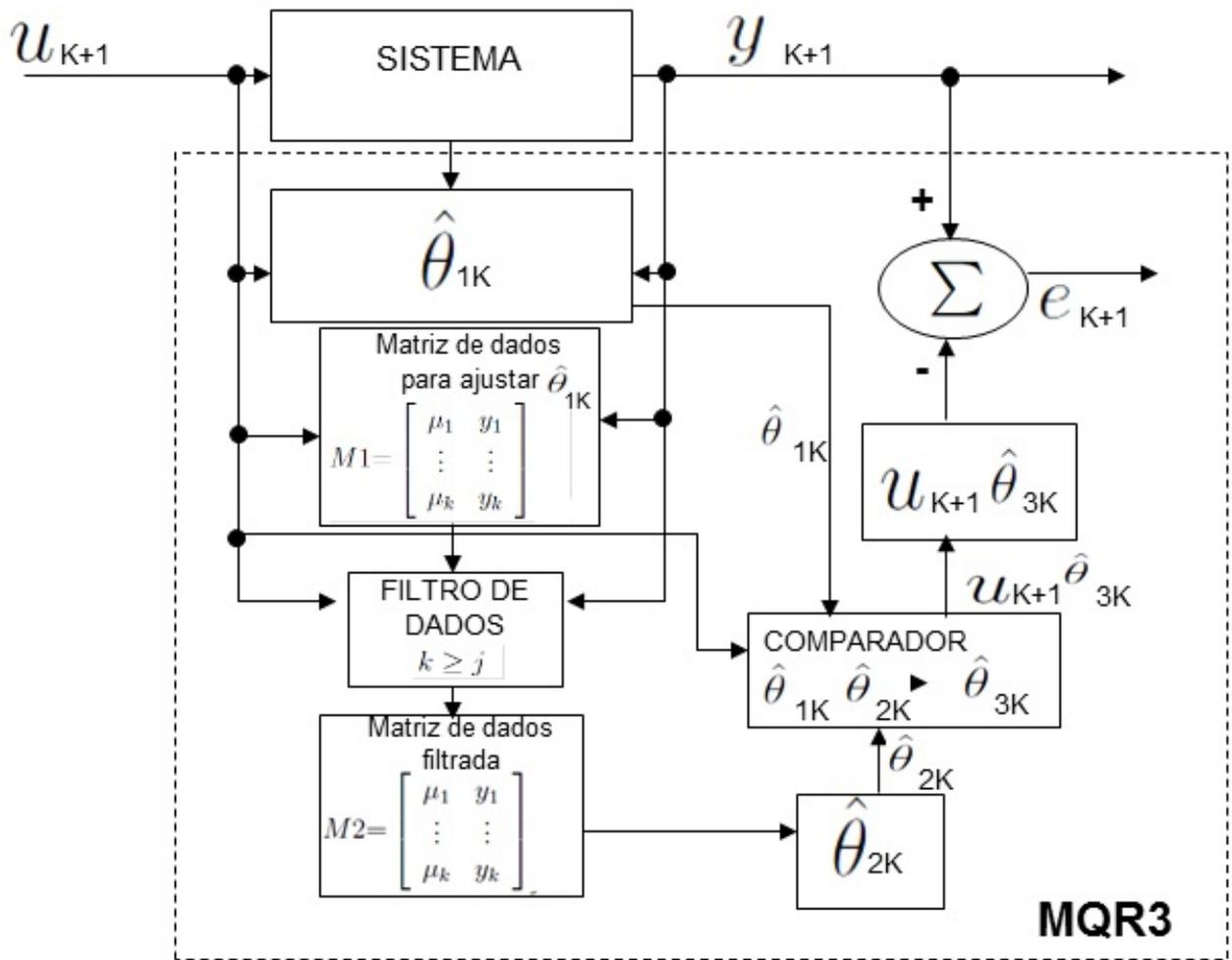


Figura 6.3: Diagrama para o método MQR3.

$\hat{\theta}_{2k}$ : é o estimador do método MQR apresentado nas equações 4.26, 4.22 e 4.23, gerado a partir da matriz de dados filtrada  $M2_{g \times d}$  e da nova observação.

Algoritmo para o Filtro de dados:

---

**Algorithm 6.1** Matriz Filtrada para o método MQR3

---

1. Definimos; Nova entrada  $:u_{k+1}$ ;  $M1$  de dimensão  $s \times d$  ;  $s = 1$ ;  $M2$  de dimensão  $g \times d$  ;  $g = 1$ .
  2. Tomamos o vetor de dados da nova entrada  $u_{k+1}$ .
  3. Calculamos a diferença entre o vetor de dados tomado em “2” e o vetor de dados de entrada da matriz  $M1_{s \times d}$ .
  4. Calculamos o valor absoluto do resultado em “3”.
  5. Multiplicamos os dados do vetor de entrada de “2” por 0,8 (é a percentagem máxima de diferença que procuramos entre os dados para nosso caso).
  6. Comparamos se o resultado em “4” é menor ou igual que em “5”. Se sim, então  $M1_{s \times d} = M2_{g \times d}$ ;  $s = s + 1$ ;  $g = g + 1$ . Se não, então  $M2_{g \times d} = M2_{g \times d}$ ;  $s = s + 1$ .
  7. Repetimos “3” - “4” - “5” - “6” até  $s = \#máximo\ de\ filas$ .
- 

$M2_{g \times d}$ : é a matriz de dados de entrada e saída, resultante do algoritmo para o Filtro de dados, com um número de observações  $g$  e séries temporais  $d$ .

*COMPARADOR*: Com  $\hat{\theta}_{1k}$  e  $\hat{\theta}_{2k}$  fazemos  $PrevisãoMQR = u_{k+1} \times \hat{\theta}_{1k}$  e  $PrevisãoparcialMQR3 = u_{k+1} \times \hat{\theta}_{2k}$ , então:

---

**Algorithm 6.2** Critério de comparação para obter a previsão MQR3

---

i) Calculamos

$$\frac{PrevisãoMQR - PrevisãoparcialMQR3}{PrevisãoMQR} \quad (6.1)$$

ii) Se 6.1 maior que 30% então a melhor previsão é:

$$PrevisãoFinalMQR3 = PrevisãoparcialMQR3$$

iii) Se 6.1 estiver entre 3% e 10%, então a melhor previsão é:

$$PrevisãoFinalMQR3 = PrevisãoMQR - \|PrevisãoMQR - PrevisãoparcialMQR3\|$$

iv) Se não ii) e iii), então a melhor previsão é:

$$PrevisãoFinalMQR3 = PrevisãoMQR$$


---

Vemos como o “comparador” trabalha com  $\hat{\theta}_{1k}$  e  $\hat{\theta}_{2k}$  aplicando o algoritmo enunciado em 6.2, para obter o terceiro estimador  $\hat{\theta}_{3k}$  que é a previsão final dada por  $PrevisãoFinalMQR3 = u_{k+1} \times \hat{\theta}_{3k}$ .

Finalmente, podemos calcular o erro de saída  $e_{k+1}$ .

$e_{k+1}$ : é o preditor de erro de saída apresentado na equação 4.22 do método de mínimos quadrados recursivo, gerado a partir da  $PrevisãoFinalMQR3$  e da saída  $y_{k+1}$ .

O método MQR3 apresentado anteriormente tem dois procedimentos principais que procuram melhorar as estimativas. O filtro de dados é o primeiro, e tem o objetivo de achar as observações do modelo de dados que mais se aproximem com a nova entrada, gerando assim uma nova matriz. O valor de percentagem do filtro de dados pode ser modificada dependendo das necessidades e características das séries temporais.

O segundo procedimento a destacar deste método é o Comparador, este funciona estabelecendo critérios entre os dois estimadores obtidos  $\hat{\theta}_{1k}$  e  $\hat{\theta}_{2k}$ , e assim calcular ou selecionar o terceiro estimador  $\hat{\theta}_{3k}$ . Os critérios do comparador podem variar dependendo do conhecimento e da natureza das séries temporais. No seguinte capítulo vamos aplicar o método MQR3 com as especificações apresentadas neste capítulo.

# Aplicação do método MQR3 a múltiplas séries temporais.

## 7.1 Introdução

Em capítulos anteriores aplicamos o método MQR mostrado na figura 6.1 em nosso modelo de dados da tabela 5.4, obtendo os correspondentes resultados de validação. Neste capítulo aplicaremos o método MQR3 da figura 6.3, ao mesmo modelo de dados da tabela 5.4 para assim proceder a validar o método.

Continuando com o diagrama implementado para o método MQR3 da figura 6.3, marcamos os blocos já calculados no capítulo 5 na figura 7.1.

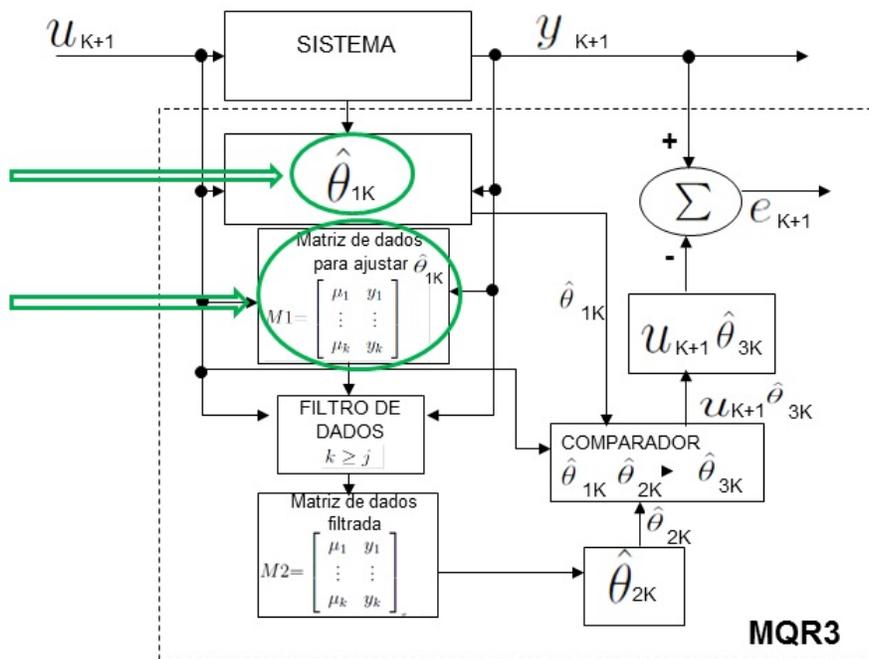


Figura 7.1: Blocos calculados do diagrama para o método MQR3.

Os blocos calculados são: a matriz de dados de entrada e saída  $M_{1482 \times 3}$  da tabela 5.4 e o estimador do método MQR  $\hat{\theta}_{1k}$  da equação 5.1.

## 7.2 MQR3 aplicado nas exportações de café na Colômbia.

- Filtro de dados.

Prosseguindo com o método MQR3, realizamos o filtro de dados com o algoritmo 6.1 usando:  $M_{1482 \times 3}$  da tabela 5.4 e as entradas  $u_{k+1}$  que para o exemplo é a matriz de dados de validação da tabela 5.5. Com  $u_{k+1}$  sendo uma matriz de 15 filas ou observações, geramos 15 matrizes  $M_{2g \times d}$  de dados de entrada e saída, resultantes do Filtro de dados. O bloco calculado é exaltado na figura 7.2.

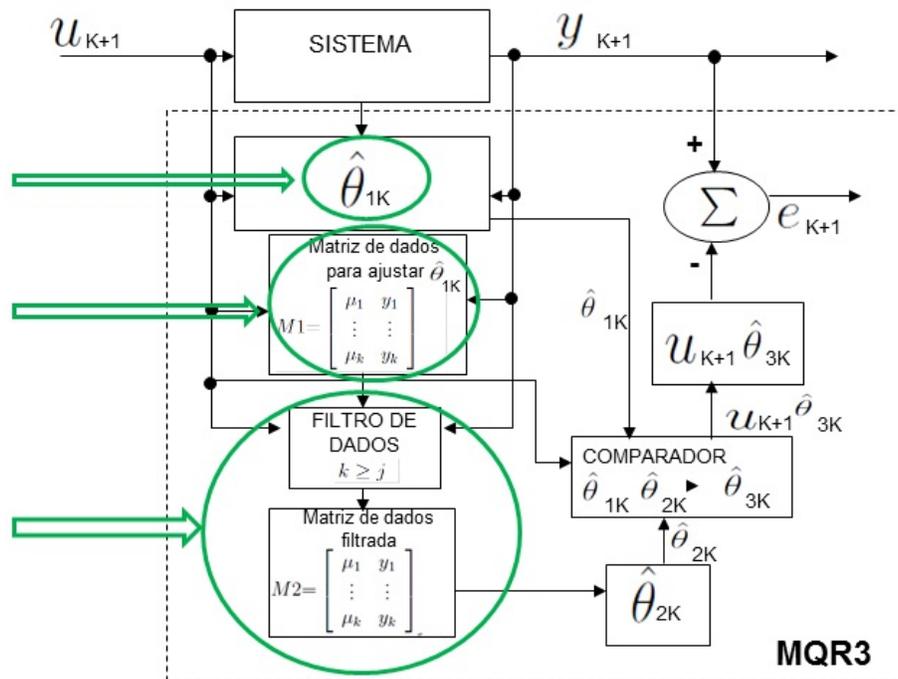


Figura 7.2: Blocos calculados do diagrama para o método MQR3.

- MQR aplicado em  $M_{2g \times d}$

Aplicamos o método MQR da figura 6.1, na matriz  $M_{2g \times d}$  resultante do Filtro de dados. Obtendo o estimador  $\hat{\theta}_{2k}$  correspondente, que chamaremos estimador  $\hat{\theta}_{2k}^{MQR3}$ , apresentado na equação 7.1. O bloco calculado é exaltado na figura 7.3.

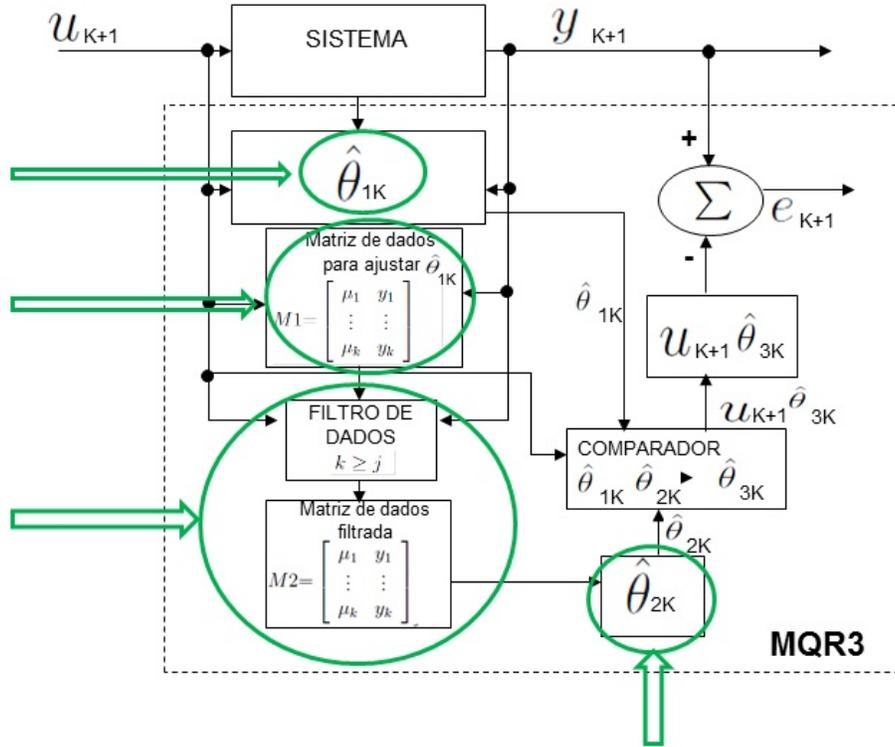


Figura 7.3: Blocos calculados do diagrama para o método MQR3.

Exemplo: usamos a primeira matriz  $M2_{g \times d}$  que gera o primeiro  $\hat{\theta}_{2k}$ , então:

$$\hat{\theta}_{21}MQR3 = \begin{bmatrix} 1.0627 \\ -0.0252 \end{bmatrix} \quad (7.1)$$

(Nota-se a diferença em relação ao  $\theta_1MQR$  da equação 5.1)

Para nosso exemplo  $u_{k+1}$  é uma matriz de 15 linhas ou observações, gerando 15 estimadores  $\hat{\theta}_{2k}MQR3$ .

- Previsão e Comparador.

Calculamos a previsão inicial de saída que chamaremos *PrevisãoparcialMQR3* com: o estimador  $\hat{\theta}_{2k}MQR3$  e os dados de entrada de validação da Matriz  $u_{k+1}$  da tabela 5.5, para assim prever a saída do mês.

Exemplo: usamos a primeira fila da Matriz  $u_{k+1}$  e o primeiro  $\hat{\theta}_{21}MQR3$  da equação 7.1.

A equação 7.2 mostra o resultado da primeira previsão de saída.

$$PrevisãoparcialMQR3 = [ 629 \quad 1909.1 ] * \hat{\theta}_{21}MQR3$$

$$PrevisãoparcialMQR3 = [ 629 \quad 1909.1 ] * \begin{bmatrix} 1.0627 \\ -0.0252 \end{bmatrix}$$

$$PrevisãoparcialMQR3 = 620,43 \quad (7.2)$$

Aplicamos o algoritmo de comparação da tabela 6.2 com as duas previsões calculadas anteriormente: *PrevisãoMQR* da equação 5.2 e *PrevisãoparcialMQR3* da equação 7.2, para assim obter a previsão final *PrevisãoFinalMQR3* e o estimador  $\hat{\theta}_{3k}$  correspondente. O bloco calculado é exaltado na figura 7.4.

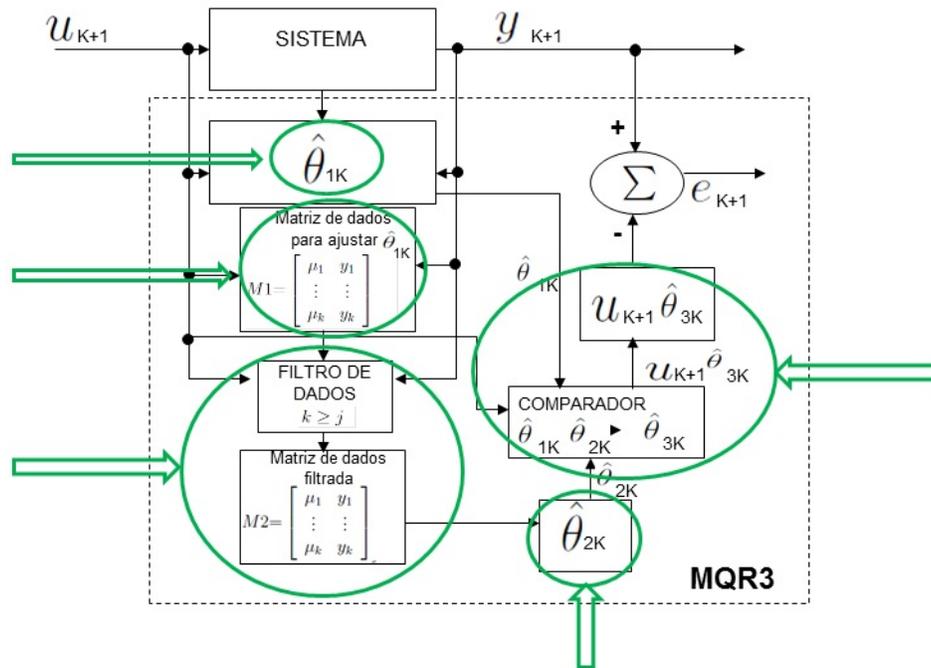


Figura 7.4: Blocos calculados do diagrama para o método MQR3.

Para nosso exemplo  $u_{k+1}$  é uma matriz de 15 linhas ou observações, gerando 15 previsões *PrevisãoFinalMQR3*.

O procedimento e aplicação do método MQR3 a múltiplas séries temporais está concluído. No próximo capítulo iremos fazer uma análise dos resultados obtidos com o método.

### 7.3 Validação do método MQR3

Apresentamos os resultados da validação do método MQR3 da mesma forma que foi feito anteriormente para o método MQR. Mostramos então as 15 previsões obtidas (*PrevisãoFinalMQR3* da tabela 6.2) com os dados de validação da tabela 5.5, resultados gerados com o procedimento completo MQR3 anterior. Observa-se nas tabelas 7.1 e 7.2 os valores de previsão obtidos.

Tabela 7.1: Resultados de previsão final com MQR3 (1).

Mês	Previsão Final MQR3
Mar-10	620,4
Abr-10	635,4
Mai-10	798,4
Jun-10	762,5
Jul-10	761,5
Ago-10	613,3
Set-10	552,6

Tabela 7.2: Resultados de previsão final com MQR3 (2).

Mês	Previsão Final MQR3
Out-10	767,9
Nov-10	894
Dic-10	1029,9
Jan-11	844,9
Fev-11	746,4
Mar-11	757
Abr-11	544,6
Mai-11	673,8

Também apresentamos com os resultados de validação das tabelas de previsão final 7.1 e 7.2, os erros de saída nas previsões para cada mês. Isto é feito calculando a porcentagem de erro entre as duas tabelas anteriores e as saídas reais da tabela de validação 5.5, para cada mês. Mostra-se nas tabelas 7.3 e 7.4 as porcentagens de erro de saída nas previsões para cada mês.

Tabela 7.3: % Erro de previsão final com MQR3 (1).

Mês	% Erro Previsão Final MQR3
Mar-10	7,72%
Abr-10	8,06%
Mai-10	38,61%
Jun-10	17,67%
Jul-10	24,43%
Ago-10	2,39%
Set-10	4,27%

Tabela 7.4: % Erro de previsão final com MQR3 (2).

Mês	% Erro Previsão Final MQR3
Out-10	21,50%
Nov-10	8,10%
Dic-10	6,79%
Jan-11	0,37%
Fev-11	14,65%
Mar-11	18,07%
Abr-11	6,10%
Mai-11	13,63%

Para concluir o capítulo e obter finalmente um resultado geral, da mesma forma mostrada na tabela 5.11 para o método MQR, apresentamos a percentagem do erro médio de previsão para os 15 dados de validação na tabela 7.5 com o método MQR3.

Tabela 7.5: Percentagem do erro médio de previsão com MQR3.

% Erro Médio Previsão Final MQR3	12,82%
----------------------------------	--------

O resultado do erro médio de previsão final obtido, depois de validar o método MQR3 com o modelo de dados, indica um erro médio na previsão de saída melhor que o método MQR, tabela 5.11. Comparações e análises são feitas no próximo capítulo.

## Análise e resultados das aplicações dos métodos MQR e MQR3 à múltiplas séries temporais.

### 8.1 Introdução

Em capítulos anteriores estudamos os métodos MQR e MQR3, sendo estes aplicados para prever os dados das exportações de café verde na Colômbia. Os resultados serão analisados neste capítulo. Esclarecemos que os métodos mostrados também podem ser aplicados em diversas áreas e para diferentes séries temporais.

### 8.2 Análise e resultados dos métodos MQR e MQR3.

Queremos apresentar as comparações entre os dois métodos de uma forma amigável para o leitor, usando gráficos que mostrem as diferenças dos resultados. Apresentamos no gráfico da figura 8.1 o comportamento dos dados de validação da tabela 5.5 da série temporal de saída.

O comportamento dos quinze dados é uma pequena amostra da série temporal de saída “volume de exportações de café verde Colombiano”, que vamos a comparar com os resultados dos dois métodos validados. Na figura 8.2 apresentamos os resultados obtidos nas tabelas 5.7 e 5.8 com o método MQR e nas tabelas 7.1 e 7.2 com o método MQR3.

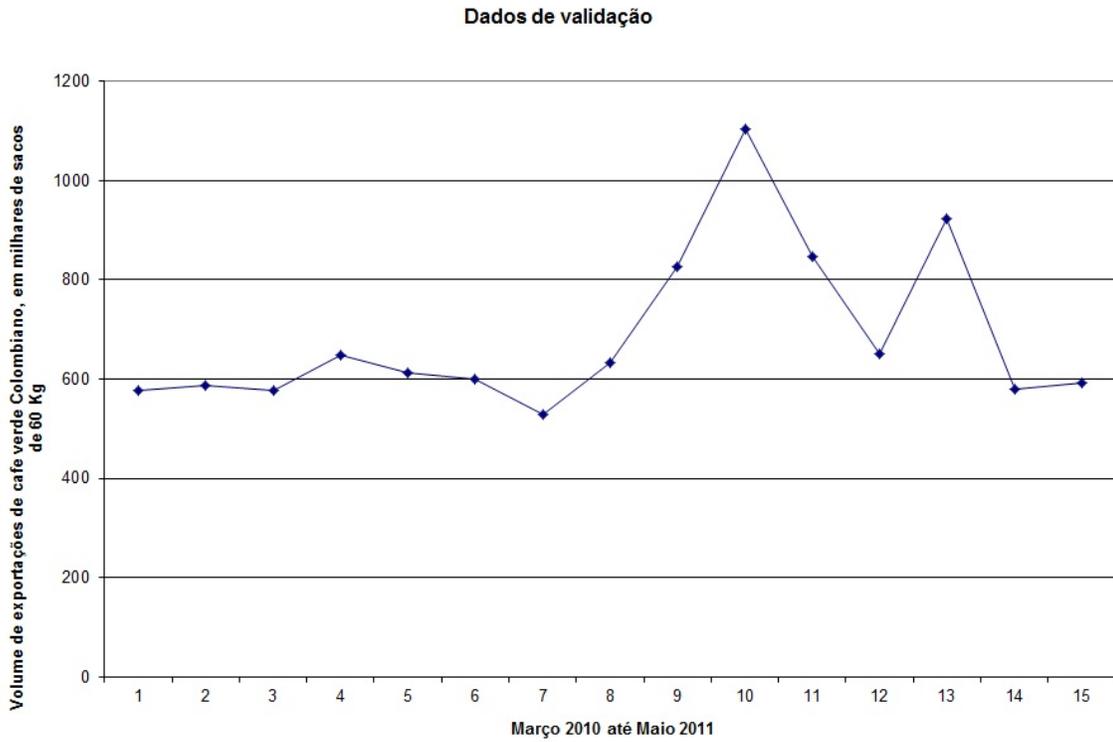


Figura 8.1: Dados de Validação.

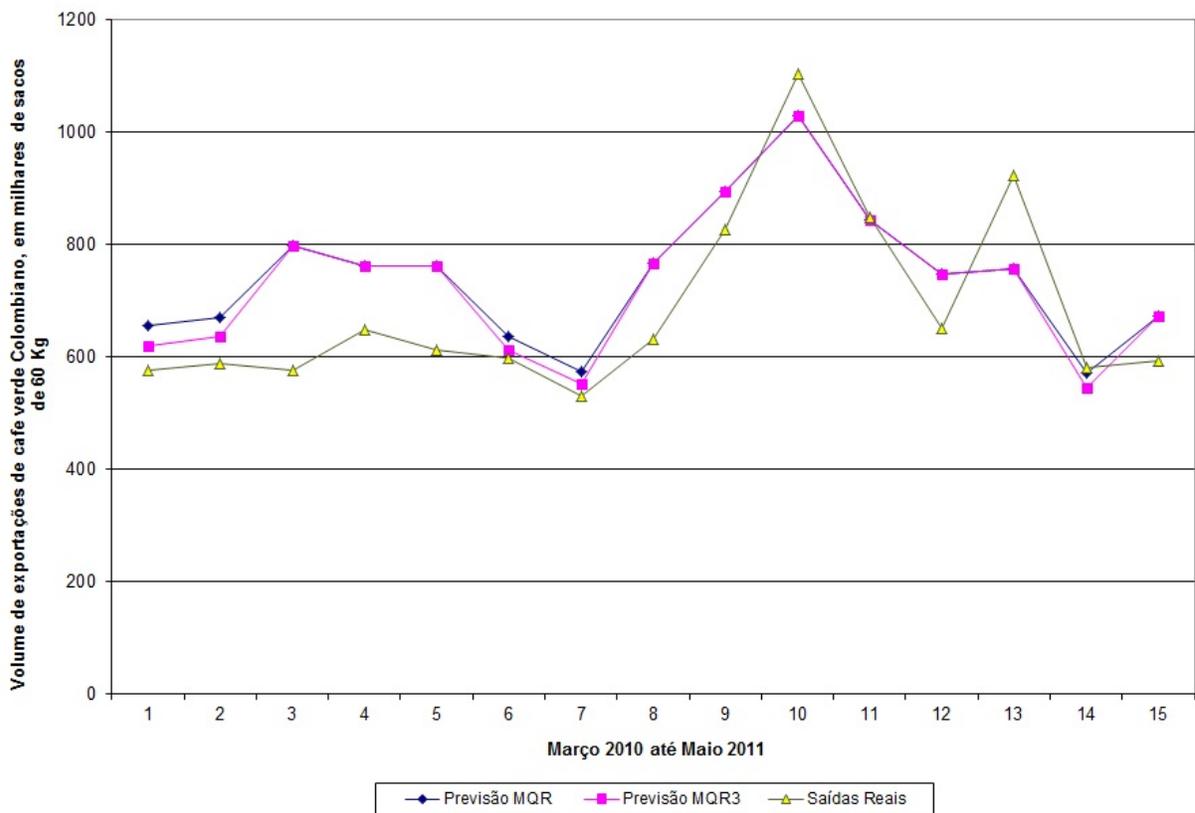


Figura 8.2: Saída real (dados de validação), Método MQR3 e Método MQR.

Na figura 8.2, pode-se ver como as duas previsões dos métodos MQR e MQR3 acompanham a saída real. Nos casos em que a saída real tem declividade considerável, o método MQR3 é igual ou bem mais próximo ao método MQR. Para os casos em que as variações de declividade da saída real são suaves, o método MQR3 tende a ficar mais próximo da saída real ou seja dar uma melhor previsão. A diferença em percentagem entre as duas previsões para os 15 dados de validação é calculada e apresenta da tabela 8.1, deixando uma ideia mais analítica do que acontece com os métodos e seus resultados.

Tabela 8.1: %Diferença entre as duas previsões

Mês	% Erro Previsão entre MQR e MQR3	Mês	% Erro Previsão entre MQR e MQR3
Mar-10	6,08%	Out-10	0,00%
Abr-10	6,14%	Nov-10	0,00%
Mai-10	0,00%	Dic-10	0,00%
Jun-10	0,00%	Jan-11	0,00%
Jul-10	0,00%	Fev-11	0,00%
Ago-10	3,69%	Mar-11	0,00%
Set-10	4,17%	Abr-11	-4,48%
		Mai-11	0,00%

O significado das porcentagens positivas e negativas da tabela 8.1, são: No caso positivo a previsão do método MQR3 é melhor e mais próxima da saída real. No caso negativo a previsão com o método MQR3 foi mais distante da saída real. Valores de zero mostram que as duas previsões são iguais nos dois métodos.

Os valores da diferença em percentagem para cada previsão respeito a saída real, são mostrados na figura 8.3.

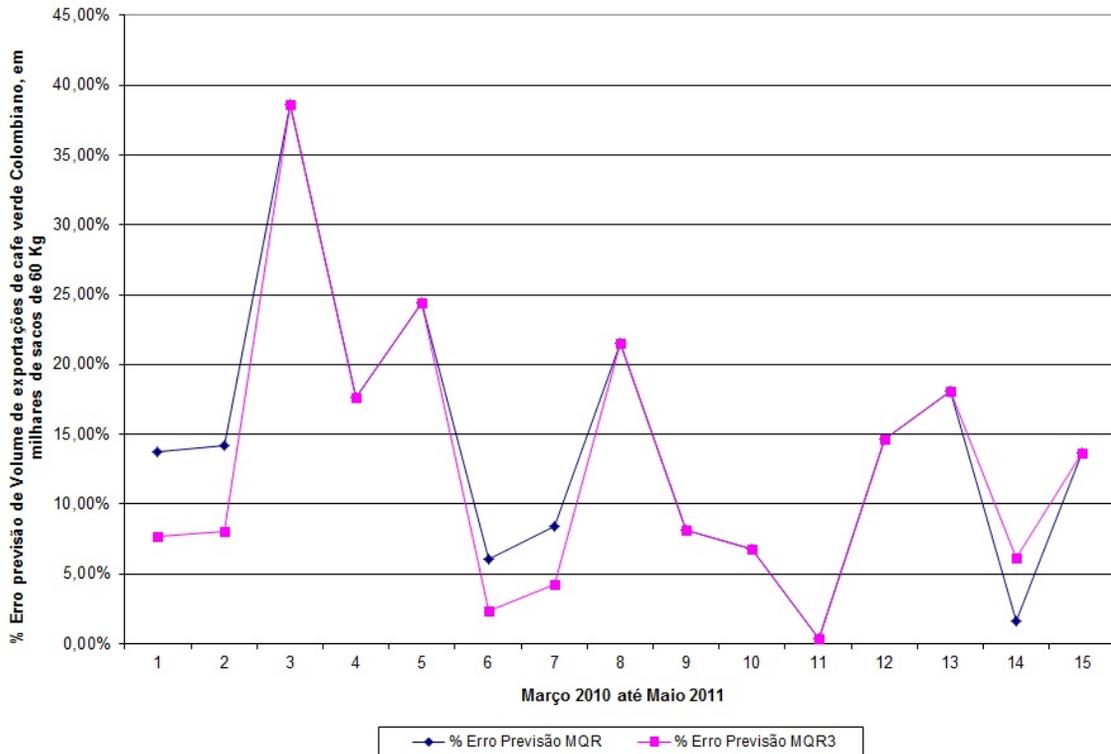


Figura 8.3: Percentagem de Erro de saída entre as duas previsões.

Para os 15 dados de validação trabalhados, pode-se observar uma tendência do método MQR3 à ter valores menores e mais aproximados da saída real (ou seja, uma percentagem de erro menor). É importante notar que o mesmo trabalho de análise feita anteriormente para 15 dados de validação, reproduziria um comportamento geral bem similar para uma maior quantidade de dados.

Lembremos que nosso trabalho consiste em mostrar que o método MQR3, aplicado nas exportações de café na Colômbia calcula para um conjunto de dados de validação, um melhor erro médio das estimativas de saída que a do método MQR. Isto pode-se ver com os resultados obtidos nas tabelas 5.11 e 7.5. Apresentamos a tabela 8.2 com os valores obtidos com os dois métodos trabalhados, e adicionalmente mostramos para sua análise, o resultado das previsões feitas com o método MQR aplicado na matriz de dados filtrada 6.1 e o estimador 7.1, chamado de Previsão parcial MQR3.

Tabela 8.2: Percentagem de erro médio de previsão com os métodos: MQR3, MQR e MQR3 parcial.

% Erro Médio de Previsão Final MQR3	12,82%
% Erro Médio de Previsão MQR	13,86%
% Erro Médio de Previsão parcial MQR3	12,57%

Da tabela anterior é evidente que a percentagem do erro médio de previsão final do método MQR3 é melhor que a do método MQR, e o resultado do erro médio de previsão parcial do

método MQR3 é melhor que os dois anteriores. Então temos que analisar qual resultado entre as previsões parcial e final do método MQR3 é mais confiável, a análise vai depender do modelo de dados escolhido. No grafico 8.4 mostra-se como a previsão parcial acompanha os dados de saída; neste caso específico ela tem uma melhor previsão.

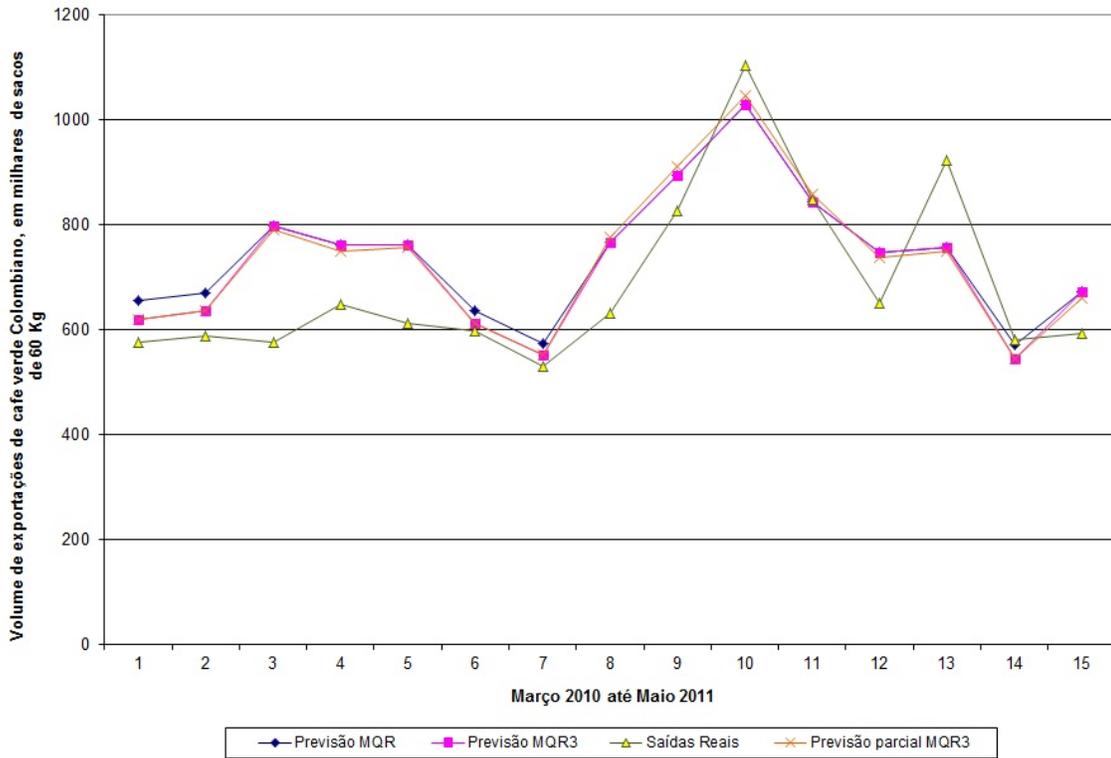


Figura 8.4: Saída real (dados de validação), Método MQR3, MQR e MQR3 parcial .

Os resultados da tabela 8.2 mostram para nosso exemplo de exportações de café na Colômbia uma diferença real entre o método MQR e MQR3 de 1.04%, sendo as previsões do método MQR3 mais próximas das saídas reais. O que representa uma melhora na previsão de 7.130 sacos de café verde Colombiano de 60 Kg exportados. Outro dado para ter em conta do resultado final com o método MQR3 na tabela 8.2, é quantidade de café que implicam os 12,82% do erro médio de previsão, representando este valor a quantidade de 87.961 sacos de café verde Colombiano de 60 Kg exportados. Lembrando que nossos dados de validação de saída, têm um mínimo de 530.000 e um máximo 1.105.000 de sacos de café verde Colombiano de 60 Kg exportados.

### 8.3 Análise do método MQR3.

O objetivo é apresentar uma comparação do erro médio de saída das previsões obtidas com o método MQR3 aplicado nas exportações de café, e o método usado para a previsão mensal do ICMS do estado de Maranhão para o ano de 2008. “ICMS é a sigla que identifica o Imposto sobre Operações relativas à Circulação de Mercadorias e sobre Prestações de Serviços de Transporte Interestadual e Intermunicipal e de Comunicação. É um imposto que cada um dos Estados e o Distrito Federal podem instituir como determina a Constituição Federal Brasileira de 1988”.

Das comparações feitas com os dois métodos a seguir, queremos mostrar que apesar de usar diferentes métodos de previsão de séries temporais, este trabalho esta dentro dos limites usados na atualidade para estimar os erros médios de previsão com nossa contribuição computacional, método MQR3.

### 8.3.1 Previsão mensal do ICMS do estado de Maranhão para o ano de 2008.

De acordo com os cadernos IMESC o trabalho feito por Alan Vasconcelos Santos [23], apresenta o seguinte: , “O presente trabalho tem por finalidade desenvolver um modelo de previsão mensal, para o ano de 2008, da principal receita tributária do Estado do Maranhão, o ICMS. A metodologia empregada consistirá na aplicação de um modelo de série temporal determinístico, denominado alisamento exponencial. De forma mais precisa, como o ICMS apresenta, ao longo do tempo, tendência e sazonalidade, serão utilizados os modelos de alisamento exponencial sazonal de Holt- Winters aditivo e multiplicativo”.

A preocupação que levou em fazer as previsões do ICMS é explicada no mesmo trabalho “A escolha de se fazer previsões da variável ICMS decorre do fato de que este imposto representa a maior fonte de arrecadação do Governo Estadual (91,66% do total das receitas tributárias). Desta forma, dado que a previsão constitui-se em um meio de fornecer informações para uma conseqüente tomada de decisões, as previsões desse tributo podem auxiliar o Governo em suas decisões que visem o aumento do bem-estar econômico.”

O trabalho completo mostra as questões teóricas dos modelos econométricos utilizados e especificamente os métodos de alisamento exponencial (Simples, Duplo e Sazonal de Holt-Winters). Em uma última parte apresenta os resultados empiricos de previsão obtidos, “Para cada técnica mostram-se os critérios utilizados para a seleção dos modelos capazes de gerar previsões e os seus principais resultados. Assim, após a escolha do modelo de previsão, serão ilustrados os resultados da previsão mensal oriundos dos métodos determinísticos do ICMS para o ano de 2008”. Os resultados de previsão mensal são nosso principal enfoque do trabalho feito por Alan Vasconcelos Santos, os quais vão ser uma referência para a porcentagem de erro nas previsões obtidas como o método MQR3.

O autor do trabalho deixa claro que: “A escolha do modelo mais adequado para gerar as previsões mensais desejadas do ICMS, embora existam outros critérios disponíveis, foi feita com base na comparação do erro percentual absoluto médio de cada modelo (EPAM) dos dois métodos”. Os resultados obtidos são apresentados na tabela 8.3.

O cálculo do EPAM é elaborado de acordo da seguinte fórmula:

$$EPAM = \frac{\sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{y_t}}{n} \quad (8.1)$$

com  $y_t$  saída real,  $\hat{y}_t$  previsão de saída e observações feitas em  $t = 1, 2, 3, \dots, n$ .

Tabela 8.3: Comparação dos EPAM para a escolha do modelo de alisamento exponencial.

<i>PERÍODO</i>	<i>ICMS</i>	<i>HW - AD</i>	<i>HW - MU</i>	<i>EPAM AD</i>	<i>EPAM MU</i>
jan/07	191.597.951,28	185.828.670,85	189.631.369,79	3,01	1,03
fev/07	161.288.444,04	195.211.486,22	204.168.793,01	21,03	26,59
mar/07	170.366.853,51	180.123.059,41	179.561.151,52	5,73	5,40
abr/07	166.089.270,97	177.939.765,34	175.064.281,29	7,14	5,40
mai/07	166.164.489,45	175.570.253,32	173.763.637,05	5,66	4,57
jun/07	173.289.288,52	176.525.668,51	172.985.359,93	1,87	0,18
jul/07	171.118.409,47	195.299.865,40	200.002.563,40	14,13	16,88
ago/07	173.975.902,69	188.392.817,91	189.010.090,48	8,29	8,64
set/07	178.737.603,10	193.784.373,76	192.524.371,49	8,42	7,71
out/07	179.177.244,20	201.578.747,84	206.373.987,81	12,50	15,18
nov/07	186.818.917,23	190.006.652,58	187.002.629,82	1,71	0,10
dez/07	169.735.896,00	192.119.814,69	189.483.144,38	13,19	11,63
<b>MÉDIA</b>				<b>8,56</b>	<b>8,61</b>

O autor da análise das previsões do ICMS do Maranhão, deixa a seguinte consideração final: “O objetivo deste trabalho foi realizar previsões mensais da série do imposto sobre operações relativas à circulação de mercadorias e sobre prestações de serviços de transporte interestadual, intermunicipal e de comunicação (ICMS) para o ano de 2008. Na busca desse objetivo, utilizou-se a técnica do alisamento exponencial sazonal de Holt-Winters aditivo, pois, como visto no decorrer do estudo, os resultados da previsão ex-post, através da verificação do erro percentual absoluto médio, mostraram-se superiores ao modelo de alisamento exponencial de Holt- Winters multiplicativo.”

Apresentamos a tabela 8.4, que tem como objetivo principal mostrar e comparar que apesar de usar diferentes métodos de previsão de séries temporais, estamos dentro dos limites usados na atualidade para calcular os erros de previsão. Deixando um caminho livre para aplicar o método proposto de MQR3 em diversas áreas.

Tabela 8.4: Quadro comparativo entre os exemplos e resultados da previsão para o ICMS e Exportações de café.

	Quantidade de Dados de validação.	Método usado para previsão.	% Erro Médio de previsão obtido.	Valor Máximo de validação (VM).	Valor mínimo de validação (Vm).	% de variação máxima no exemplo. (VM-Vm)/(VM)
Exportações de café.	15	(MQR).3p	12,82%	1105	530	52,04%
ICMS	12	EPAM AD	8,56%	191597951,28	161288444,04	15,82%
Diferenças entre os exemplos	3		4,27%			36,22%

Um resultado importante da tabela 8.4, é a proximidade dos resultados do “erro médio de previsão” nos dois métodos, comparado com a ampla diferença do “% de variação máxima” nos dados, sendo isto uma razão do poder de calculo que temos na aplicação do método MQR3.

A seguir mostramos nas figuras 8.5 e 8.6, o comportamento no tempo junto dos parâmetros de tendência da equação 2.17, para cada um dos dados de validação das séries temporais do “ICMS do Maranhão” e o “Volume de exportações de café verde Colombiano” respectivamente.

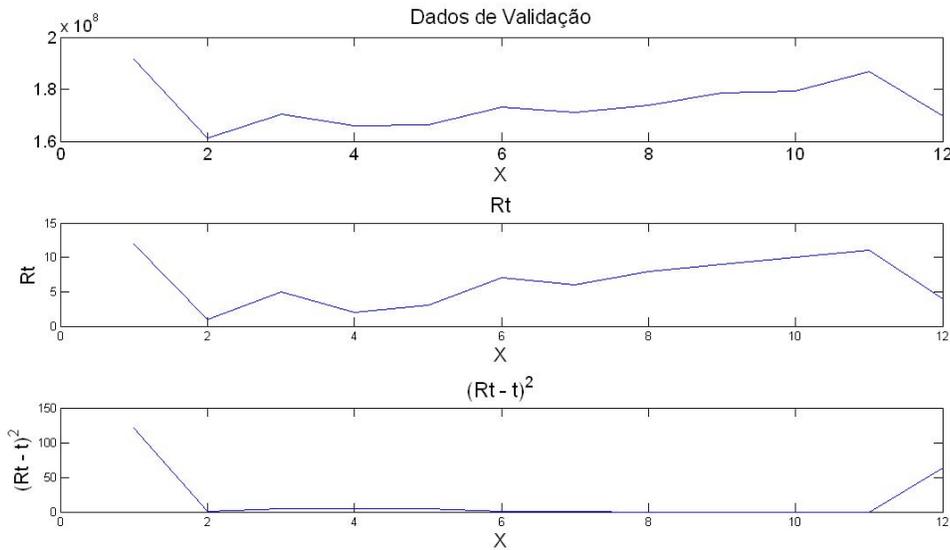


Figura 8.5: Teste para tendência em dados validação série ICMS -Maranhão (X=observações)

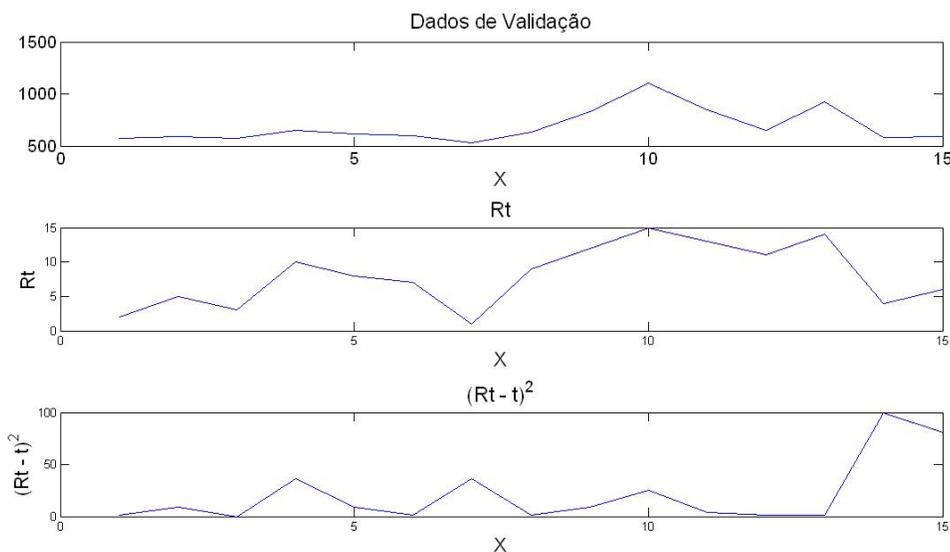


Figura 8.6: Teste para tendência em dados validação série “Volume de exportações de café verde Colombiano” (X=observações)

Finalmente, dos gráficos apresentados observamos como os dados do ICMS têm um comportamento mais uniforme no tempo em comparação com os dados das exportações de café.



## Exemplos: aplicações dos métodos MQR e MQR3 a séries temporais com previsão de um passo.

### 9.1 Introdução

Para finalizar e deixar claro que o método MQR3, é uma contribuição confiável para melhorar as estimativas do método MQR. Apresentamos os resultados da utilização dos dois métodos em seis séries temporais, comparando as estimativas e apresentando os erros médios de previsão.

### 9.2 Resultados com os métodos MQR e MQR3.

Usando cada uma das séries temporais já apresentadas nesta tese, vamos a executar os dois métodos em cada uma de elas para obter previsões para frente. Apresentando os gráficos e as tabelas de resultados.

Para aplicar os métodos MQR e MQR3 nas seguintes séries temporais vamos usar a estrutura do modelo de dados para frente apresentada na figura 5.5, esta relação de dados prevê os dados da saída de um mês com os dados de entrada do mês anterior.

#### 9.2.1 Série 1: Produção de café verde Colombiano, medida em milhares de sacos de café.

Com o método MQR aplicado ao modelo de dados, podemos observar o comportamento da saída estimada a medida que são processados os dados da série. A figura 9.1 mostra como o comportamento da saída estimada melhora enquanto os dados de entrada são aumentados, acompanhando os dados de saída reais para cada iteração.

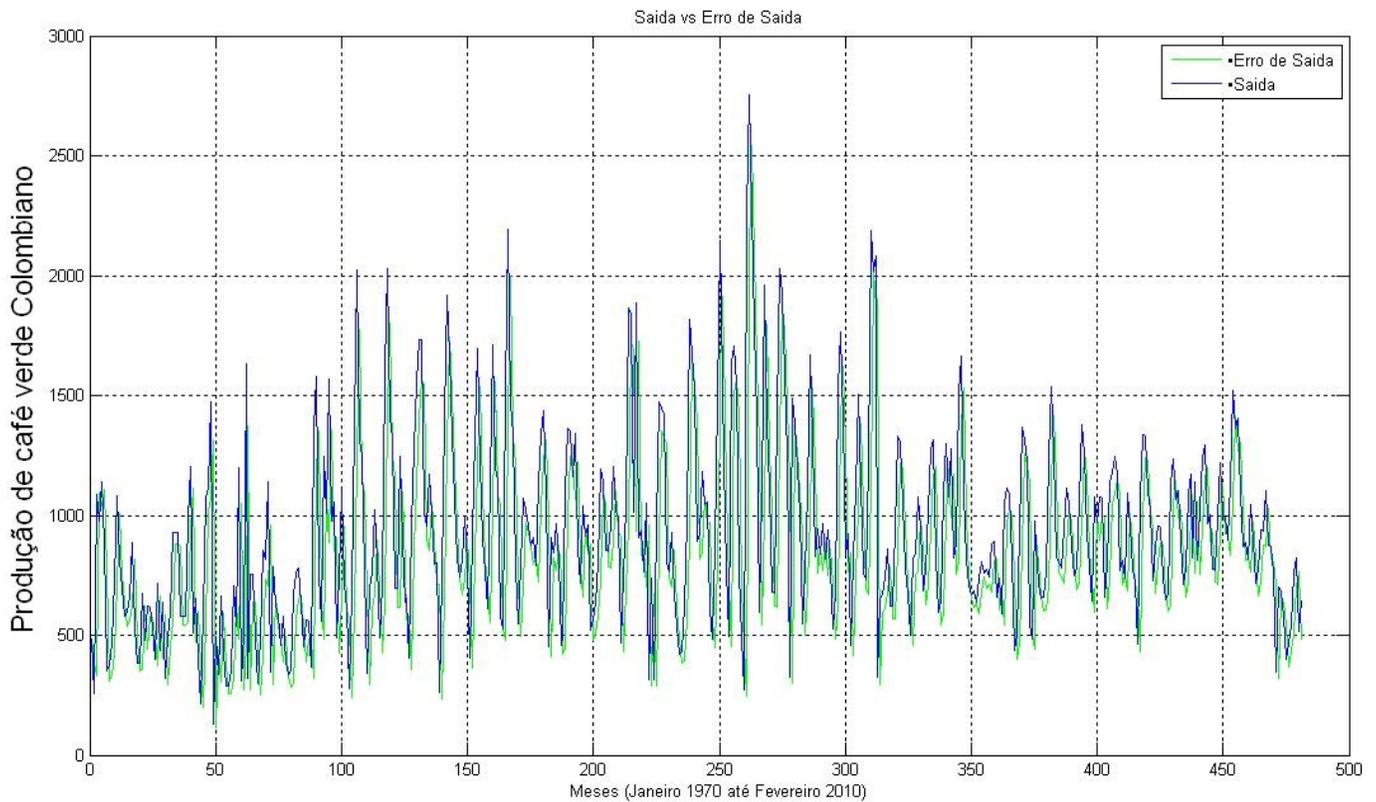


Figura 9.1: Série 1 Saída do Sistema x Saída estimada.

Outra observação importante para ter em conta na análise dos resultados do método MQR relacionado com a figura 9.1, é o comportamento do estimador  $\theta_{MQR}$  para cada iteração e apresentado na equação 4.23, o qual precisa ter uma boa convergência ao estimador obtido do método de mínimos quadrados ( $\theta_{MQ}$ ) da equação 3.13. Na figura 9.2 ilustra-se a variação de  $\theta_{MQR}$ .

Na figura 9.2 pode-se observar como a curva apresenta um erro menor a 0,05 nos últimos 100 dados depois de aplicar o método MQR .

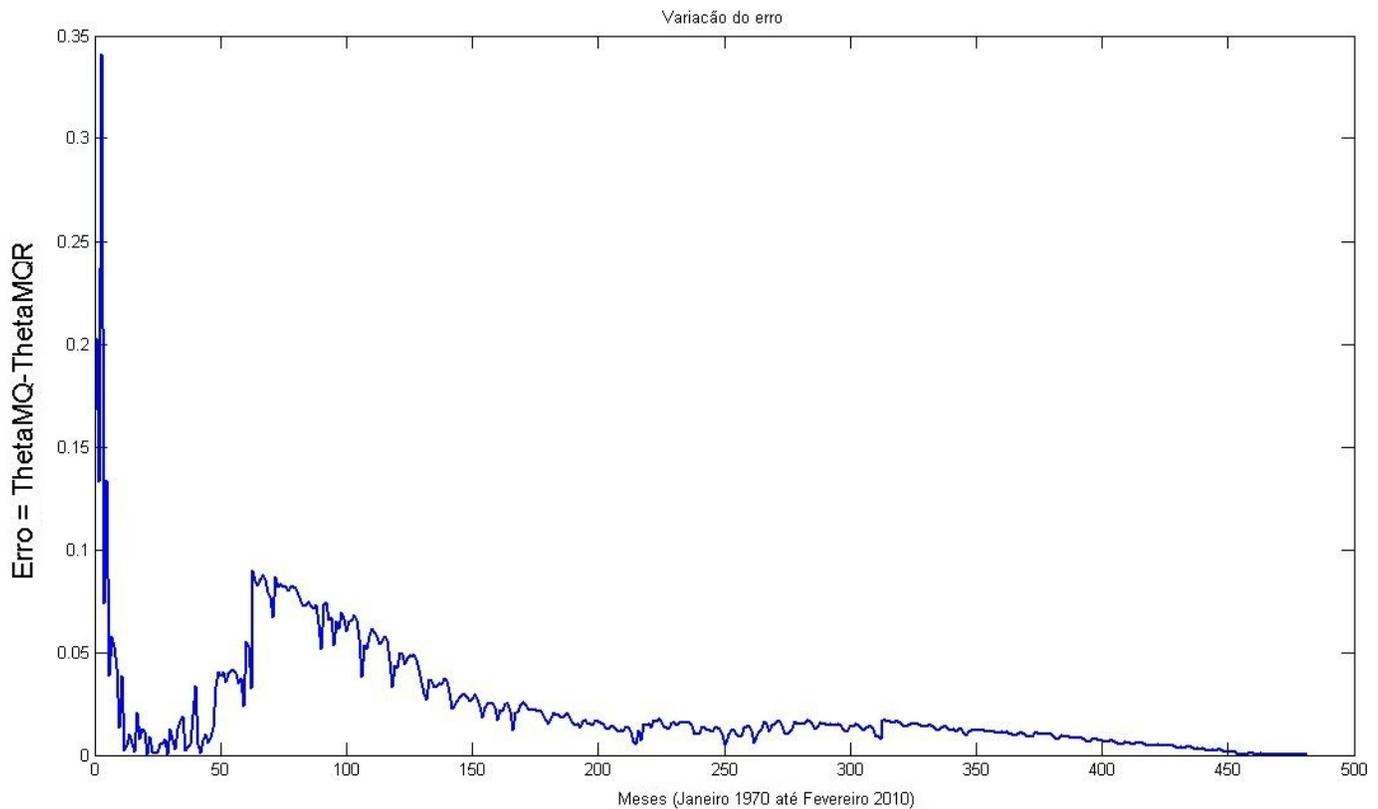


Figura 9.2: Série 1. Convergência de  $\theta MQR$  a  $\theta MQ$

Apresentamos a tabela 9.1 com os resultados obtidos depois de aplicar os métodos MQR e MQR3, de forma análoga com a tabela 8.2. Também mostramos o resultado das previsões feitas com o método MQR aplicado na matriz de dados filtrada 6.1, chamado de Previsão parcial MQR3.

Tabela 9.1: Série 1. Percentagem do erro médio de previsão com os métodos de : MQR3, MQR e MQR parcial.

% Erro Médio Previsão Final MQR3	9,88%
% Erro Médio Previsão MQR	6,93%
% Erro Médio Previsão parcial MQR3	4,03%

Da tabela anterior é evidente que a percentagem do erro médio de previsão parcial do método MQR3 é a melhor.

### 9.2.2 Série 2: Câmbio de Dólar a peso Colombiano.

Com o método MQR aplicado ao modelo de dados, podemos observar o comportamento da saída estimada a medida que são processados os dados da série. A figura 9.3 mostra como

o comportamento da saída estimada melhora enquanto os dados de entrada são aumentados, acompanhando os dados de saída reais para cada iteração.

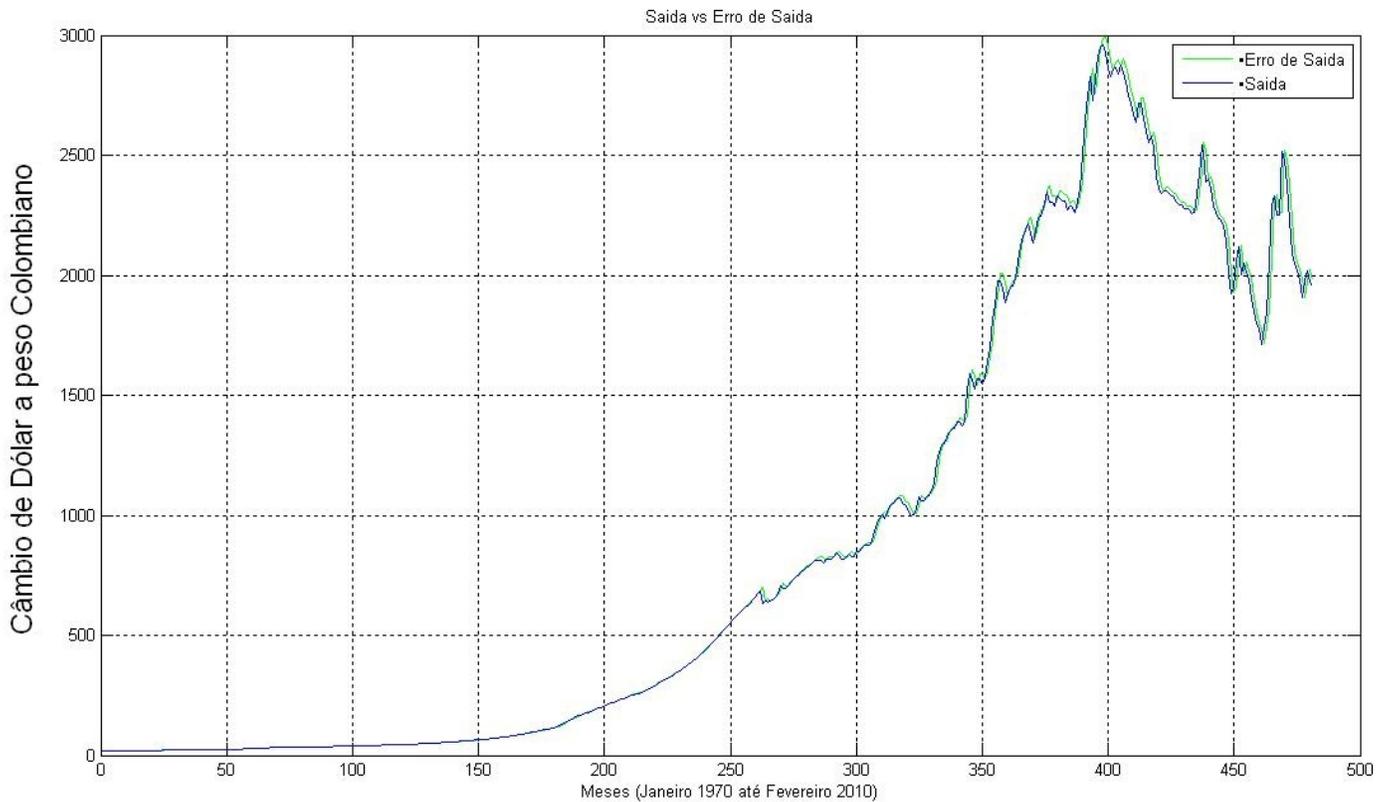


Figura 9.3: Série 2 Saída do Sistema x Saída estimada.

Outra observação importante para ter em conta na análise dos resultados do método MQR relacionado com a figura 9.3, é o comportamento do estimador  $\theta_{MQR}$  para cada iteração e apresentado na equação 4.23, o qual precisa ter uma boa convergência ao estimador obtido do método de mínimos quadrados ( $\theta_{MQ}$ ) da equação 3.13. Na figura 9.4 ilustra-se a variação de  $\theta_{MQR}$ .

Na figura 9.4 pode-se observar como a curva apresenta um erro menor a 0,005 nos últimos 100 dados depois de aplicar o método MQR .

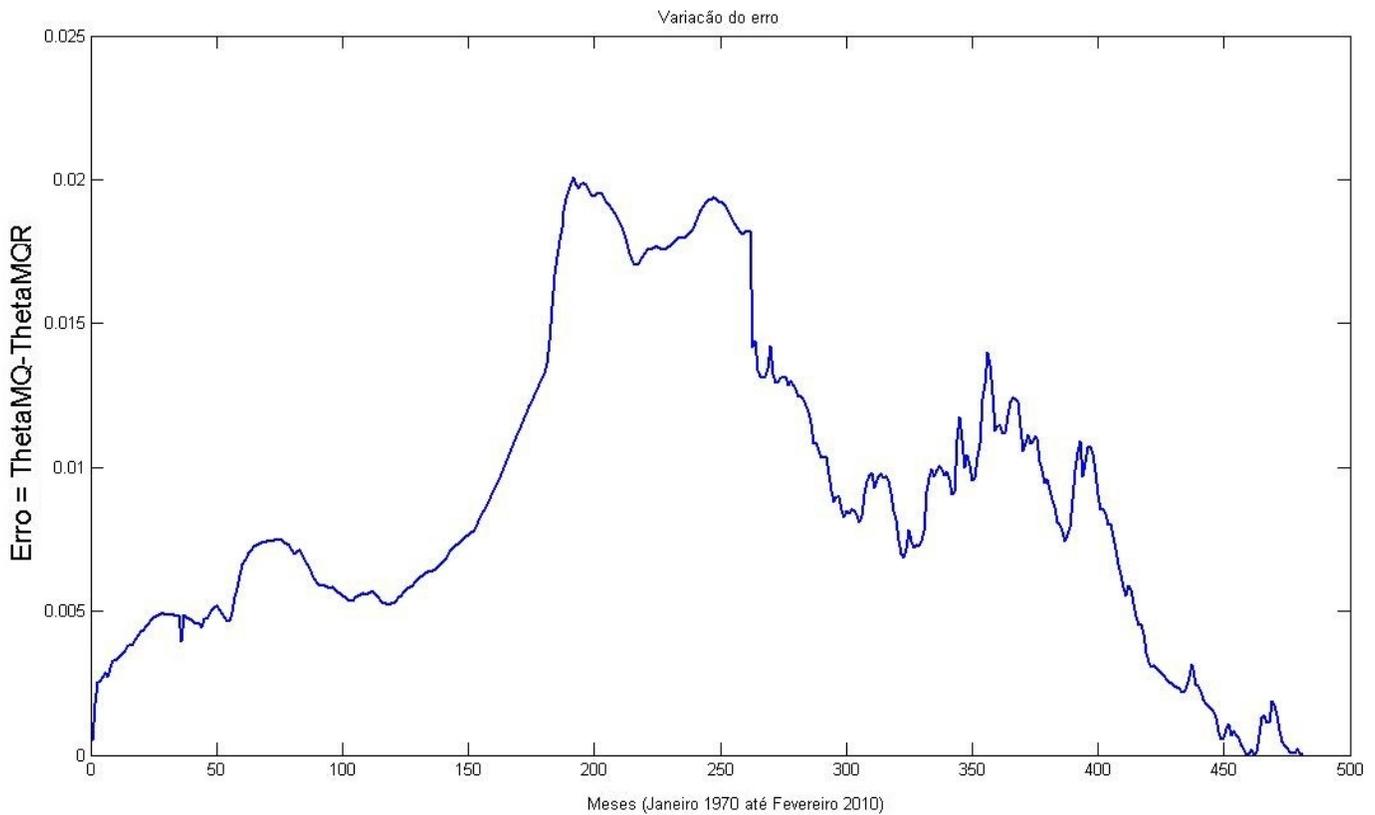


Figura 9.4: Série 2. Convergência de  $\theta MQR$  a  $\theta MQ$

Apresentamos a tabela 9.2 com os resultados obtidos depois de aplicar os métodos MQR e MQR3, de forma análoga com a tabela 8.2. Também mostramos o resultado das previsões feitas com o método MQR aplicado na matriz de dados filtrada 6.1, chamado de Previsão parcial MQR3.

Tabela 9.2: Série 2. Percentagem do erro médio de previsão com os métodos de : MQR3, MQR e MQR parcial.

% Erro Médio Previsão Final MQR3	0,2%
% Erro Médio Previsão MQR	0,2%
% Erro Médio Previsão parcial MQR3	0,19%

Da tabela anterior a percentagem do erro médio de previsão parcial do método MQR3 é a melhor.

### 9.2.3 Série 3: Exportações de petróleo e derivados em milhões de dólares (FOB).

Com o método MQR aplicado ao modelo de dados, podemos observar o comportamento da saída estimada a medida que são processados os dados da série. A figura 9.5 mostra como

o comportamento da saída estimada melhora enquanto os dados de entrada são aumentados, acompanhando os dados de saída reais para cada iteração.

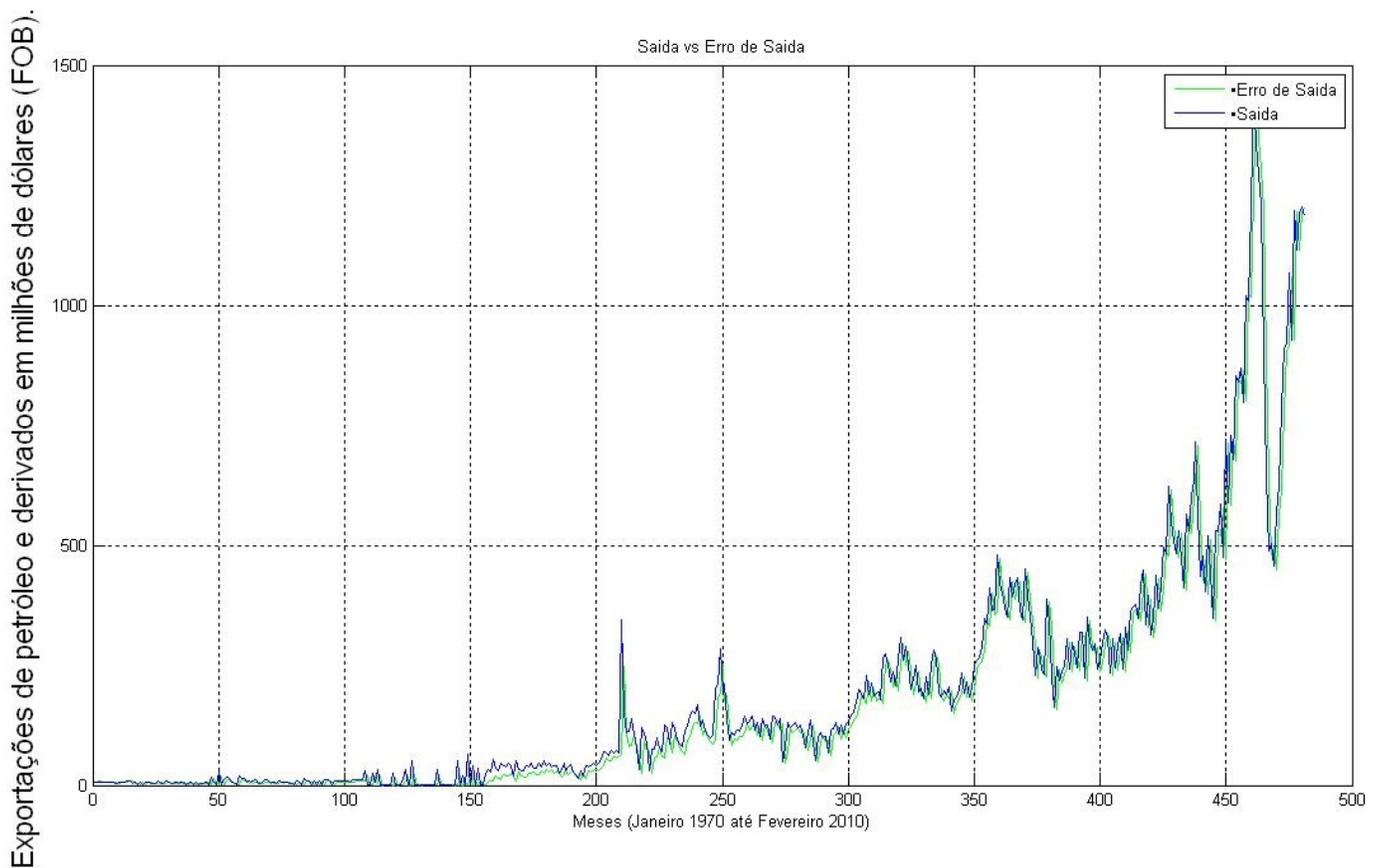


Figura 9.5: Série 3 Saída do Sistema x Saída estimada.

Outra observação importante para ter em conta na análise dos resultados do método MQR relacionado com a figura 9.5, é o comportamento do estimador  $\theta_{MQR}$  para cada iteração e apresentado na equação 4.23, o qual precisa ter uma boa convergência ao estimador obtido do método de mínimos quadrados ( $\theta_{MQ}$ ) da equação 3.13. Na figura 9.6 ilustra-se a variação de  $\theta_{MQR}$ .

Na figura 9.6 pode-se observar como a curva apresenta um erro menor a 0,05 nos últimos 100 dados depois de aplicar o método MQR .

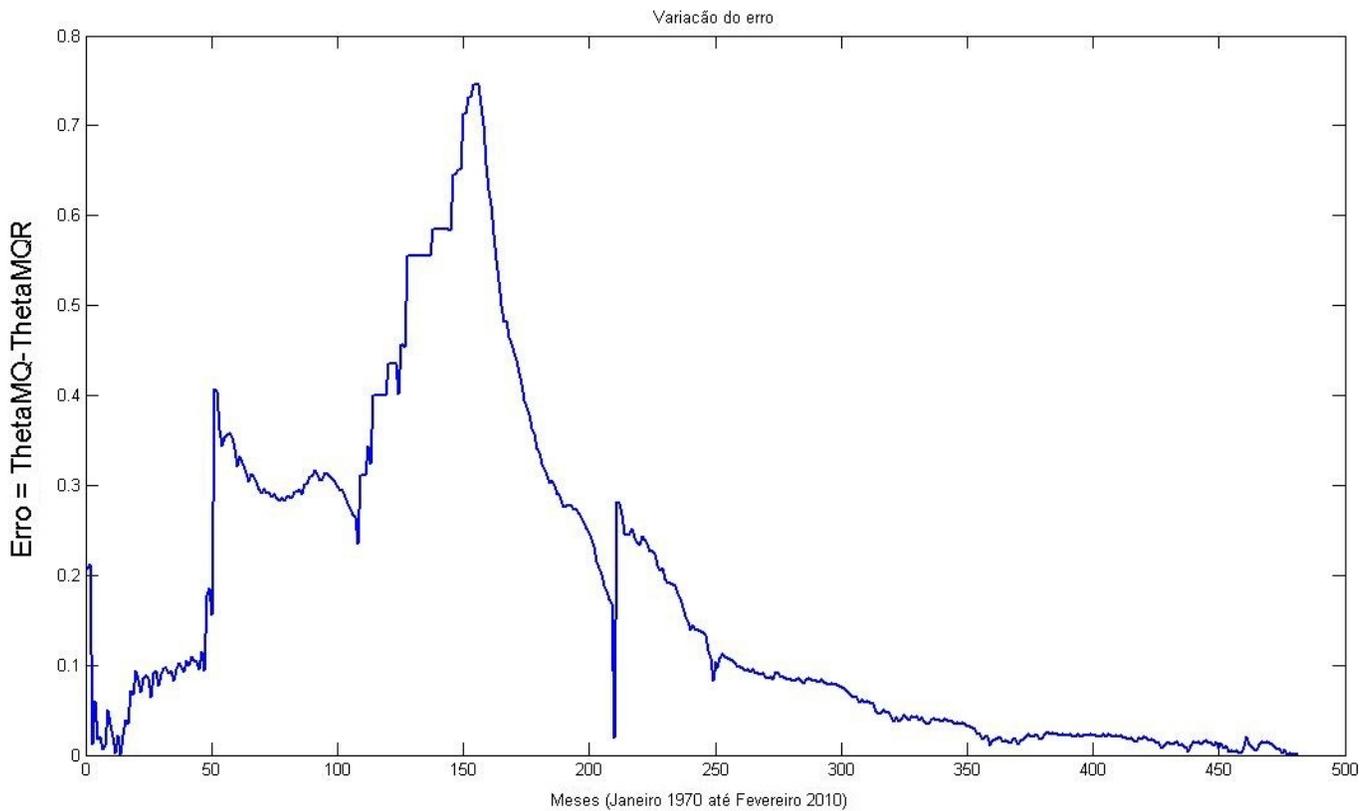


Figura 9.6: Série 3. Convergência de  $\theta MQR$  a  $\theta MQ$

Apresentamos a tabela 9.3 com os resultados obtidos depois de aplicar os métodos MQR e MQR3, de forma análoga com a tabela 8.2. Também mostramos o resultado das previsões feitas com o método MQR aplicado na matriz de dados filtrada 6.1, chamado de Previsão parcial MQR3.

Tabela 9.3: Série 3. Percentagem do erro médio de previsão com os métodos de : MQR3, MQR e MQR parcial.

% Erro Médio Previsão Final MQR3	0,06%
% Erro Médio Previsão MQR	0,06%
% Erro Médio Previsão parcial MQR3	0,4%

Da tabela anterior a percentagem do erro médio de previsão final do método MQR3 é a melhor.

#### 9.2.4 Série 4: Exportações de café em milhões de dólares (FOB).

Com o método MQR aplicado ao modelo de dados, podemos observar o comportamento da saída estimada a medida que são processados os dados da série. A figura 9.7 mostra como

o comportamento da saída estimada melhora enquanto os dados de entrada são aumentados, acompanhando os dados de saída reais para cada iteração.

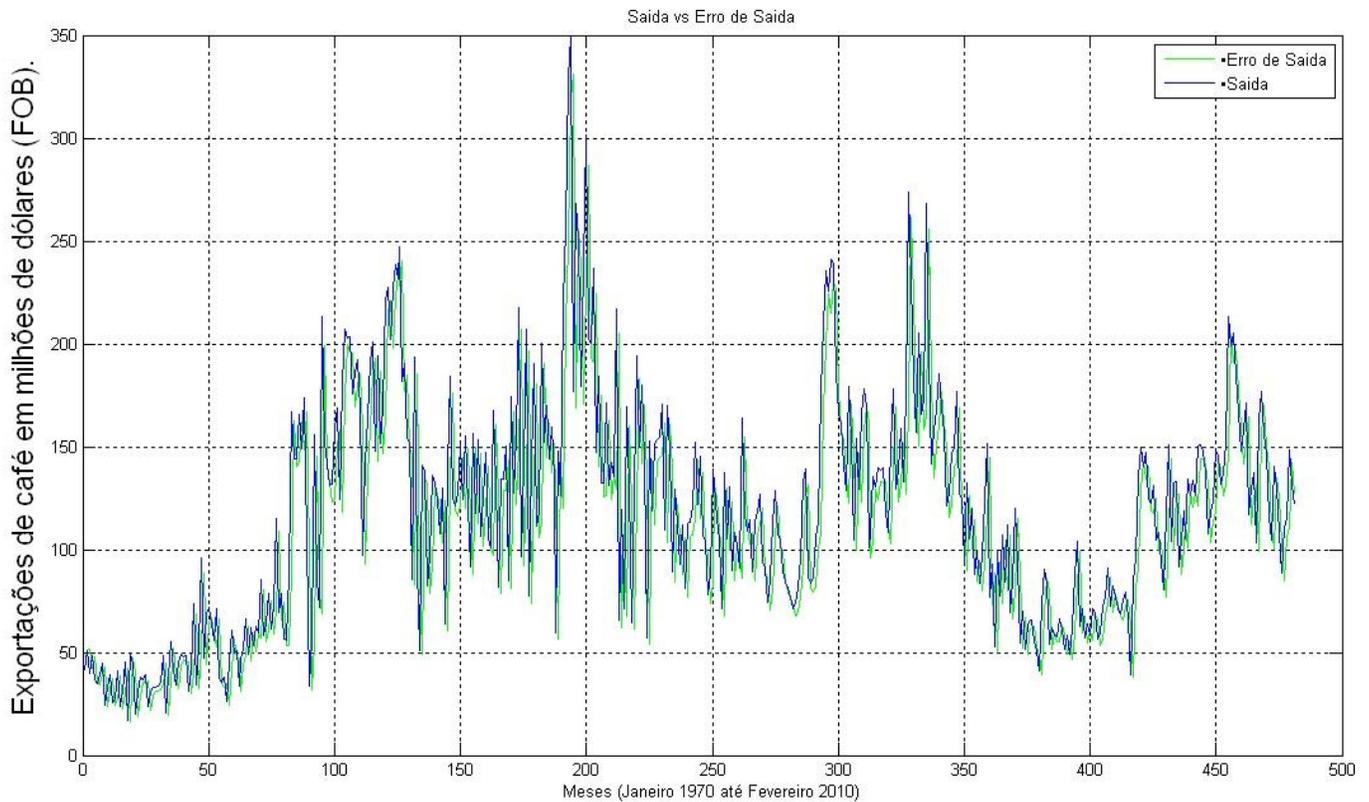


Figura 9.7: Série 4 Saída do Sistema x Saída estimada.

Outra observação importante para ter em conta na análise dos resultados do método MQR relacionado com a figura 9.7, é o comportamento do estimador  $\theta_{MQR}$  para cada iteração e apresentado na equação 4.23, o qual precisa ter uma boa convergência ao estimador obtido do método de mínimos quadrados ( $\theta_{MQ}$ ) da equação 3.13. Na figura 9.8 ilustra-se a variação de  $\theta_{MQR}$ .

Na figura 9.8 pode-se observar como a curva apresenta um erro menor a 0,025 nos últimos 100 dados depois de aplicar o método MQR .

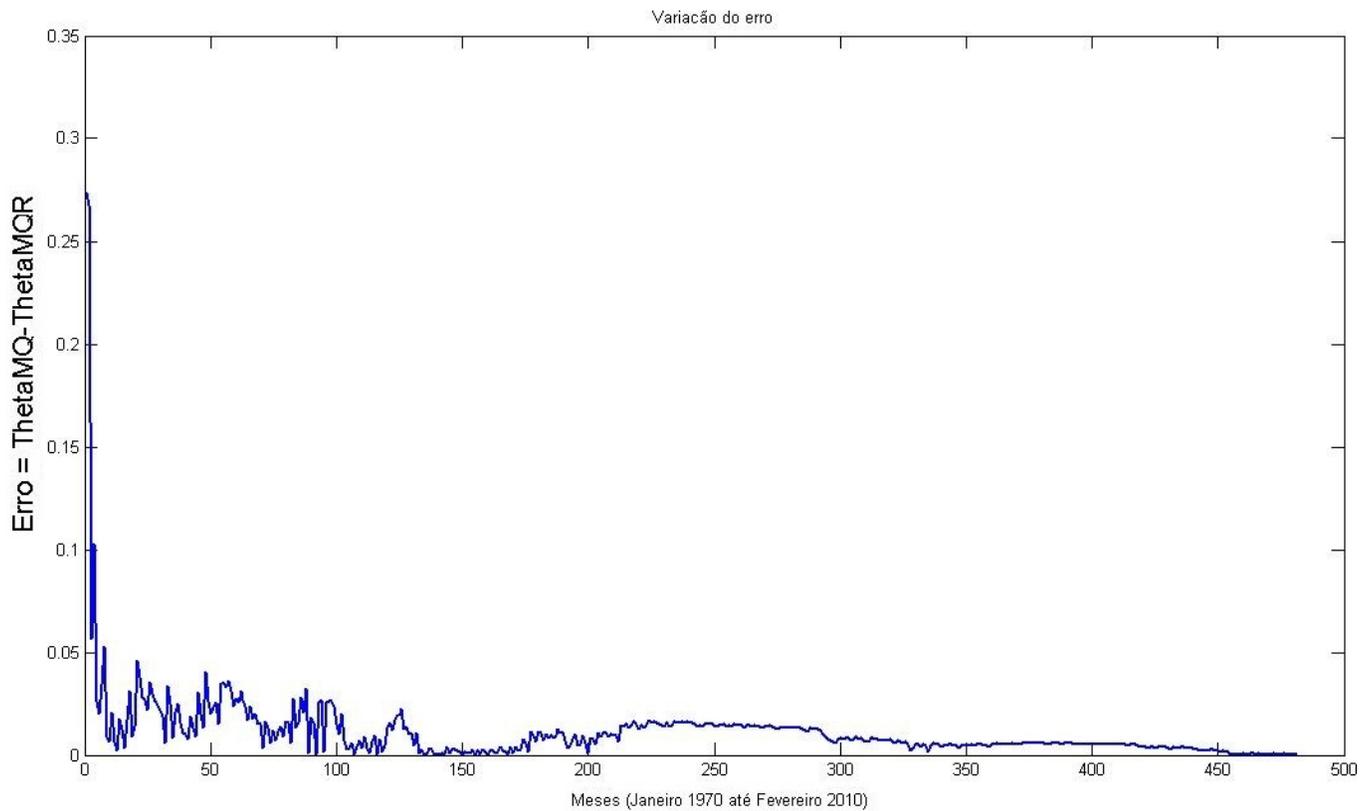


Figura 9.8: Série 4. Convergência de  $\theta MQR$  a  $\theta MQ$

Apresentamos a tabela 9.4 com os resultados obtidos depois de aplicar os métodos MQR e MQR3, de forma análoga com a tabela 8.2. Também mostramos o resultado das previsões feitas com o método MQR aplicado na matriz de dados filtrada 6.1, chamado de Previsão parcial MQR3.

Tabela 9.4: Série 4. Percentagem do erro médio de previsão com os métodos de : MQR3, MQR e MQR da matriz filtrada.

% Erro Médio Previsão Final MQR3	4,04%
% Erro Médio Previsão MQR	4,04%
% Erro Médio Previsão parcial MQR3	3,45%

Da tabela anterior a percentagem do erro médio de previsão parcial do método MQR3 é a melhor.

### 9.2.5 Série 5: Reservas Internacionais Líquidas em milhões de dólares.

Com o método MQR aplicado ao modelo de dados, podemos observar o comportamento da saída estimada a medida que são processados os dados da série. A figura 9.9 mostra como

o comportamento da saída estimada melhora enquanto os dados de entrada são aumentados, acompanhando os dados de saída reais para cada iteração.

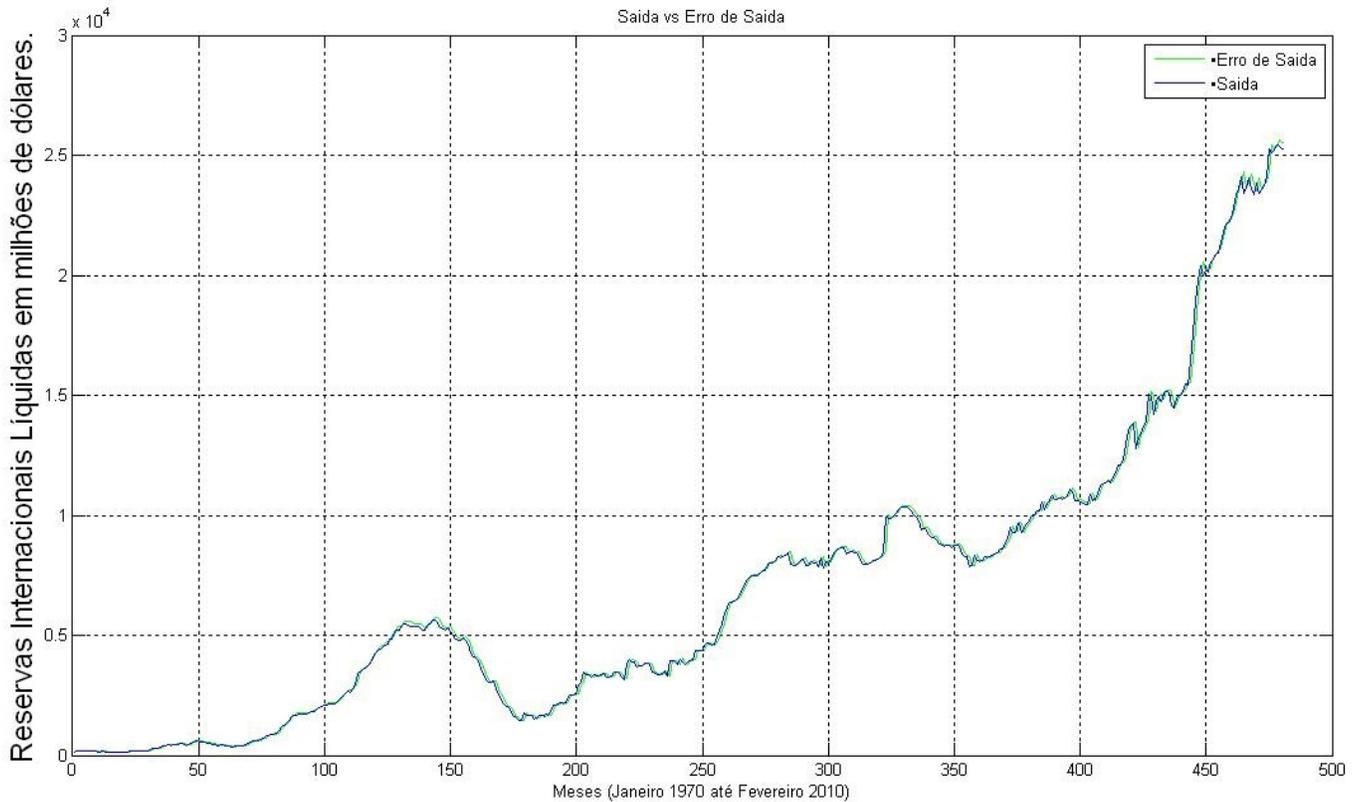


Figura 9.9: Série 5 Saída do Sistema x Saída estimada.

Outra observação importante para ter em conta na análise dos resultados do método MQR relacionado com a figura 9.9, é o comportamento do estimador  $\theta_{MQR}$  para cada iteração e apresentado na equação 4.23, o qual precisa ter uma boa convergência ao estimador obtido do método de mínimos quadrados ( $\theta_{MQ}$ ) da equação 3.13. Na figura 9.10 ilustra-se a variação de  $\theta_{MQR}$ .

Na figura 9.10 pode-se observar como a curva apresenta um erro menor a 0,01 nos últimos 100 dados depois de aplicar o método MQR .

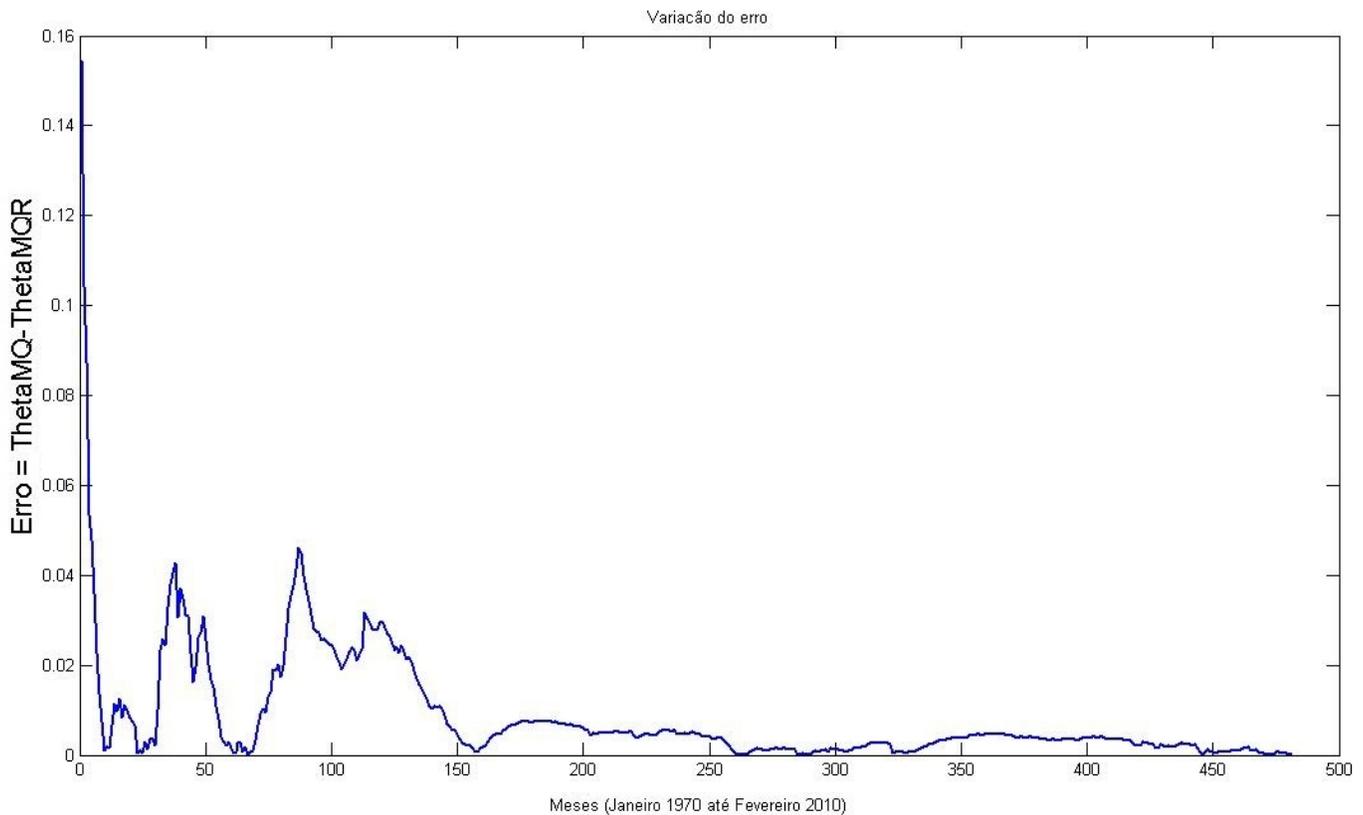


Figura 9.10: Série 5. Convergência de  $\theta MQR$  a  $\theta MQ$

Apresentamos a tabela 9.5 com os resultados obtidos depois de aplicar os métodos MQR e MQR3, de forma análoga com a tabela 8.2. Também mostramos o resultado das previsões feitas com o método MQR aplicado na matriz de dados filtrada 6.1, chamado de Previsão parcial MQR3.

Tabela 9.5: Série 5. Percentagem do erro médio de previsão com os métodos de : MQR3, MQR e MQR da matriz filtrada.

% Erro Médio Previsão Final MQR3	0,77%
% Erro Médio Previsão MQR	0,77%
% Erro Médio Previsão parcial MQR3	0,77%

Da tabela anterior as percentagens do erro médio de previsão são iguais.

### 9.2.6 Série 6: Volume de exportações de café verde Colombiano, em milhares de sacos de 60 Kg.

Com o método MQR aplicado ao modelo de dados, podemos observar o comportamento da saída estimada a medida que são processados os dados da série. A figura 9.11 mostra como

o comportamento da saída estimada melhora enquanto os dados de entrada são aumentados, acompanhando os dados de saída reais para cada iteração.

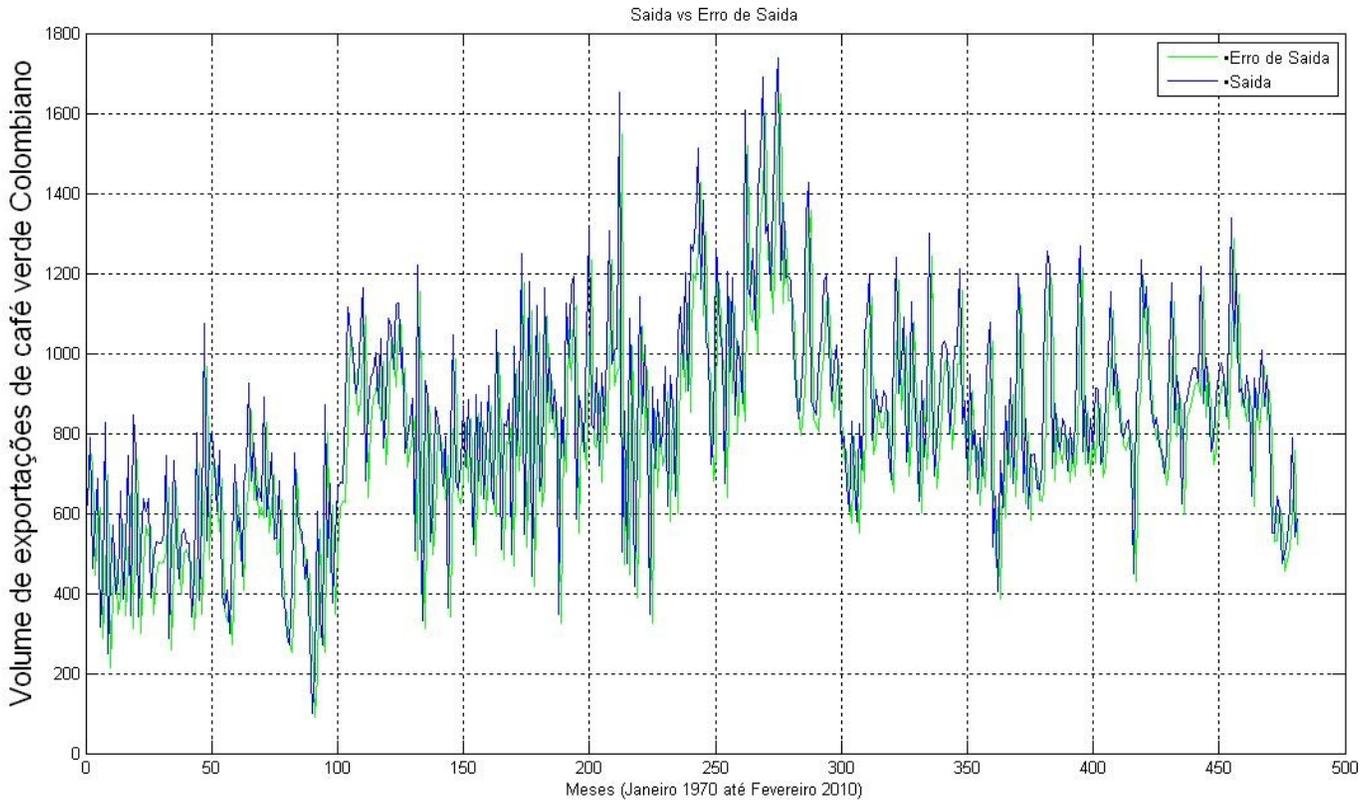


Figura 9.11: Série 6 Saída do Sistema x Saída estimada.

Outra observação importante para ter em conta na análise dos resultados do método MQR relacionado com a figura 9.11, é o comportamento do estimador  $\theta_{MQR}$  para cada iteração e apresentado na equação 4.23, o qual precisa ter uma boa convergência ao estimador obtido do método de mínimos quadrados ( $\theta_{MQ}$ ) da equação 3.13. Na figura 9.12 ilustra-se a variação de  $\theta_{MQR}$ .

Na figura 9.12 pode-se observar como a curva apresenta um erro menor a 0,025 nos últimos 100 dados depois de aplicar o método MQR .

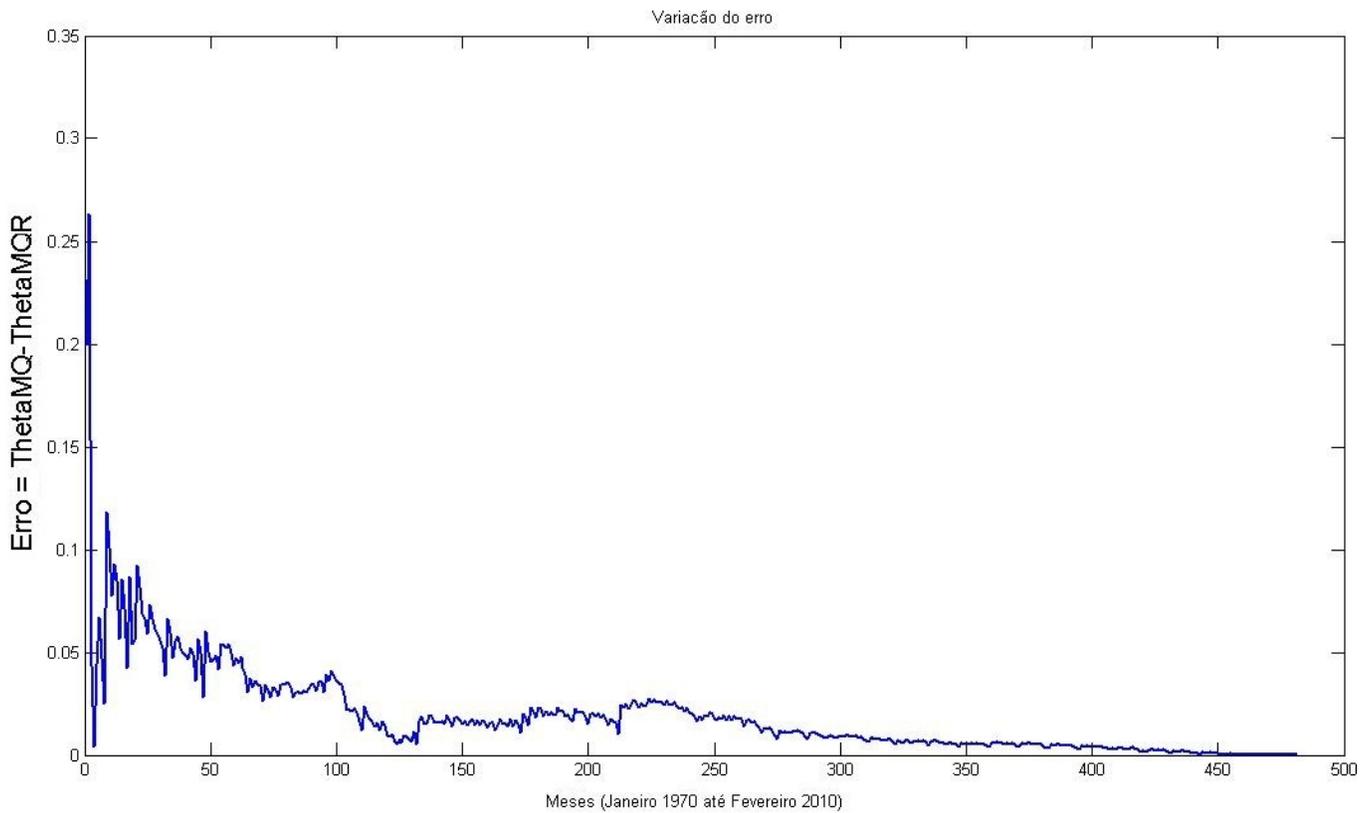


Figura 9.12: Série 6. Convergência de  $\theta_{MQR}$  a  $\theta_{MQ}$

Apresentamos a tabela 9.6 com os resultados obtidos depois de aplicar os métodos MQR e MQR3, de forma análoga com a tabela 8.2. Também mostramos o resultado das previsões feitas com o método MQR aplicado na matriz de dados filtrada 6.1, chamado de Previsão parcial MQR3.

Tabela 9.6: Série 6. Percentagem do erro médio de previsão com os métodos de : MQR3, MQR e MQR da matriz filtrada.

% Erro Médio Previsão Final MQR3	8,32%
% Erro Médio Previsão MQR	3,88%
% Erro Médio Previsão parcial MQR3	2,46%

Da tabela anterior a percentagem do erro médio de previsão parcial do método MQR3 é a melhor.



## Conclusões.

Aplicar os métodos MQR e MQR3 à múltiplas séries temporais reais, mostra a capacidade computacional dos mesmos para obter estimadores e previsões com menor erro. Acreditamos que a capacidade computacional do método MQR3 é uma das principais características deste trabalho. Como importantes conclusões podemos citar:

- O erro médio de previsão de saída calculado com o método MQR3 é menor que o erro médio obtido com o método MQR, tornando os resultados finais da tabela 8.2 com o método MQR3 mais confiáveis para fazer previsões. No caso onde a serie temporal de saída oscila muito é importante conhecer e analisar a diferença entre os três erros da tabela 8.2.
- A matriz de dados  $M1$  do método MQR3, apresentada nas figuras 6.1 e 6.3, não precisa incluir todas as observações existentes. A matriz só necessita armazenar o número suficiente de dados (atualizados a cada nova observação) para ajustar o estimador  $\hat{\theta}_{1k}$ , como é mostrado na figura 5.9. Desenvolvemos assim, um método que continua trabalhando com baixa memória de armazenamento.
- Apresentar o método desenvolvido MQR3 em forma esquemática facilita o entendimento do leitor e mostra que ele trabalha usando basicamente, e duas vezes, o método MQR. Assim o método MQR3 é considerado uma contribuição para o MQR.
- Com os exemplos apresentados no capítulo nove, concluímos que aplicando o método MQR3 em modelos de dados com só uma série temporal de entrada e uma de saída, obtemos um melhor erro médio de saída com a previsão parcial de MQR3, gerada a partir da matriz de dados filtrada  $M2_{g \times d}$  e a nova observação, tabela 9.1. Nos casos de múltiplas séries temporais os resultados da previsão final de MQR3 são mais confiáveis de usar, isto pela redundância de dados, tabela 8.2.
- Com os resultados de validação e a análise feita ao método MQR3, pode-se concluir que o erro médio de previsão para a saída de volume nas exportações de café verde colombiano por mês, está dentro do intervalo de percentagem de erro de saída estimada utilizado na atualidade para fazer estimativas econômicas mensais, tabela 8.4.

- A contribuição computacional apresentada para o método MQR e validada com a série temporal de saída de exportações de café na Colômbia, deixa um caminho aberto para explorar as vantagens do método. Ela Oferece uma possibilidade para melhorar os erros médios de previsão em procedimentos que atualmente usam o método MQR.
- Os tipos de séries temporais para os quais o método de MQR3 apresenta melhor desempenho, não são motivo de estudo ou comparações neste trabalho. Isto porque estamos focados em principio, na redução do erro médio de previsão de saída e validação do método para qualquer série temporal.
- Os resultados obtidos utilizando o método MQR3, ainda estando muito próximos aos resultados do método MQR tabela 8.1, são melhores. A diferença de 1.04% nos resultados pode melhorar as estatísticas de erro médio de previsão em sistemas de tempo real que realizam muitas estimativas por segundo.
- A aplicação do método MQR3 ao modelo de dados apresentado na tabela 5.4, estimou adequadamente o comportamento da série temporal de saída, resultando uma previsão de saída com erro médio final de 12,82%, para os 15 meses de validação, tabela 5.5. Conseguimos desta forma uma previsão de saída com erro médio menor em 1,04% que o obtido com o método MQR. Neste caso isto, corresponde a 7,13 milhares de sacos de café verde Colombiano exportados em um mês.
- O método MQR3 melhora significativamente os resultados do erro médio de previsão, quando os dados da série temporal de saída apresentam muita oscilação no tempo. Isto, porque a matriz filtrada M2 pode discriminar com maior precisão os dados atípicos e obter um melhor erro médio de previsão de saída, conforme observado nos resultados dos exemplos do capítulo nove.
- É possível obter resultados diferentes para o exemplo de séries temporais de saída apresentado e com a teoria deste trabalho, apenas mudando algumas das séries temporais de entrada de nosso modelo de dados. Os resultados podem ser melhores se as séries temporais de entrada conseguem melhor explicar os dados da série temporal de saída.

## Proposta de Trabalho Futuro.

Um dos problemas a ser estudado será desenvolver um algoritmo de realização estocástica, que permita calcular estimativas precisas das covariâncias com base em medições de dados finitos.

Com ele obter-se-á um modelo inovativo

$$x_{(t+1)} = Ax_{(t)} + Ke_{(t)} \quad (10.1)$$

$$y_{(t)} = Cx_{(t)} + e_{(t)} \quad (10.2)$$

onde  $y$  é o vetor de saída,  $x$  o vetor de estado, e  $e$  é o processo de inovação.

Com algoritmo de realização estocástica baseado em dados de covariância finitos, usando a teoria da realização estocástica no espaço de estado, a qual fornece um método de construção de modelos de Markov que simula um processo estocástico estacionário com a matriz de covariância prescrita, e serve como uma base para os métodos de identificação em subespaços, e poder-se-á aplica-lo a múltiplas séries temporais tanto as aqui estudadas, como no caso de múltiplas séries temporais de saída [24].



# Bibliografía

- [1] . Seal, L., *Studies in History of Probability and Statistics*, Bio., Ed. XV -The historical development of the Gauss linear model., 1967.
- [2] . Plackett, R. L., *Studies in History of Probability and Statistics*, Biometrika, Ed. XXIX - The discovery of the method of least squares., 1972.
- [3] . Stigler, Stephen., “Gauss and the invention of least squares.” *The Annals of Statistics.*, vol. 9, pp. 465–474, 1981.
- [4] . Young, Peter C., *Recursive Estimation and Time-Series Analysis.*, . Green Meadows, Stanmore Drive., Ed. Springer., 2011.
- [5] . Hair, Jr J.F., *Análise Multivariada de Dados*, . Adonai Schlup Sant’ Anna e Anselmo Chaves Neto., Ed. Bookman, 2005.
- [6] . Morettin, P. A., *Séries Temporais*, S. Hazzan, Ed. Atual editora, 1986.
- [7] . Berenson, M., *Análisis de Series de Tiempo en Estadística para la Administración.* Editorial Prentice Hall, 2001.
- [8] . Wold, H.O.A., *A Study in the Analysis of Stationary Time Series*, Almqvist and Wiksell, Eds. Uppsala, 1938.
- [9] . Bisgaard, S., *Time Series Analysis and Forecasting by Example*, D. Balding, Ed. Wiley series in probability and statistics, 2011.
- [10] . Box G.E.P., Jenkins G.M., *Time Series Analysis: Forecasting and Control*, E. revista, Ed. Holden-Day, 1976.
- [11] . Conover, W. J., *Practical Nonparametric Statistics*, J. Wiley and Sons, Eds. John Wiley and Sons, 1971.
- [12] . Banco de la republica de Colombia. (1993) Principales indicadores económicos, 1923-1992. [Online]. Available: [http://www.banrep.gov.co/seriesestadisticas/see\\_s\\_externo.htm#cambiaría](http://www.banrep.gov.co/seriesestadisticas/see_s_externo.htm#cambiaría)

- [13] . Barreto, G., “Modelagem computacional distribuída e paralela de sistemas e de séries temporais multivariáveis no espaço e estado.” Ph.D. dissertation, Universidad Estadual de Campinas, 2002.
- [14] . Brockwell, P. J., *Introduction to Time Series and Forecasting*, 2nd Edition, Ed. New York: Springer, 2002.
- [15] . Kalman, R. E., *Mathematical description of linear dynamical systems*. SIAM Journal on Control, 1963.
- [16] . Goodwin, G.C., *Dinamic System Identification*, A. press, Ed. Academic press, 1977.
- [17] . Ljung, L., *Theory and Practice of Recursive Identification*, . The MIT press., Ed., 1983.
- [18] . Caradori do Amaral, W., “Identificação recursiva de sistemas multivariáveis,” Ph.D. dissertation, Unicamp, 1980.
- [19] . Henderson, H. V., *On Deriving The Inverse Of a Sum of Matrices*, S. Review, Ed. Siam Review, 1981.
- [20] . Baker, P., “La broca del cafe en colombia (chinchiná).” *Cenicafé*, vol. 1, p. 154, 1999.
- [21] . Junguito, Roberto y Pizano, *Producción de café en Colombia*. Fondo Cultural Cafetero y Fedesarrollo, 1991.
- [22] CEPAL, “Análisis y proyecciones del desarrollo estadístico en colombia (el desarrollo económico de colombia),” *DANE*, vol. 1, p. 30, 1957.
- [23] . Vasconcelos, Santos A., *Análise de modelos de séries temporárias para a previsão mensal do ICMS do Estado do Maranhão.*, . Raimundo Nonato Palhano Silva, Jhonatan U. P. Sousa., Ed. São Luís: IMESC., 2008.
- [24] . Katayama, Tohru, *Subspace Methods for System Identification*, . E.D. Sontag., M. Thoma., Ed. Springer-Verlag, 2005.