

**LUZIA VIDAL DE SOUZA**

**PROGRAMAÇÃO GENÉTICA E COMBINAÇÃO DE  
PREDITORES PARA PREVISÃO  
DE SÉRIES TEMPORAIS**

Tese de doutorado apresentada como requisito parcial à obtenção do grau de Doutora em Ciências. Programa de Pós Graduação em Métodos Numéricos em Engenharia – Programação Matemática, Setores de Tecnologia e de Ciências Exatas, Universidade Federal do Paraná.

Orientadora: Prof<sup>a</sup>. Dra. Aurora T. R. Pozo  
Co-orientador: Prof. Dr. Anselmo C. Neto

Curitiba  
Outubro – 2006

# **Livros Grátis**

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

# TERMO DE APROVAÇÃO

LUZIA VIDAL DE SOUZA

## PROGRAMAÇÃO GENÉTICA E COMBINAÇÃO DE PREDITORES PARA PREVISÃO DE SÉRIES TEMPORAIS

Tese aprovada como requisito parcial para obtenção do grau de Doutora em Ciências, no Programa de Pós-Graduação em Métodos Numéricos em Engenharia – Programação Matemática da Universidade Federal do Paraná, pela seguinte banca examinadora:

Orientador:

---

Prof.<sup>a</sup>. Dra. Aurora Trinidad Ramirez Pozo  
Departamento de Construção Informática, UFPR

---

Prof. Dr. Anselmo Chaves Neto  
Departamento de Estatística, UFPR

---

Prof. Dr. Joel M. Corrêa da Rosa  
Departamento de Estatística, UFPR

---

Prof. Dr. Leandro dos Santos Coelho  
Engenharia de Produção e Sistemas, PUC-PR

---

Prof.<sup>a</sup>. Dra. Silvia Modesto Nassar  
Departamento de Informática e Estatística, UFSC-SC

Curitiba, 30 de outubro de 2006.

Dedico este trabalho aos meus queridos filhos  
Caroline e Gustavo e à minha mãe Maria Vidal.

## AGRADECIMENTOS

Meu agradecimento especial à professora Aurora Trinidad Ramirez Pozo, pela orientação durante a realização deste trabalho, pela confiança depositada em mim na escolha do tema, e principalmente pelo companheirismo, encorajamento e dedicação ao meu trabalho durante todas as fases do curso. Sua orientação foi preciosa, sem a qual a realização deste trabalho não teria sido possível.

Ao professor Anselmo Chaves Neto, pela co-orientação deste trabalho, pelas valiosas sugestões, e pelo acompanhamento constante na obtenção dos resultados deste trabalho.

Ao professor Joel Corrêa da Rosa, por suas valiosas contribuições e atendimento às dúvidas que surgiram durante desenvolvimento desta tese.

Aos meus filhos Caroline e Gustavo, por compreenderem tão bem minha dedicação ao trabalho e todos os momentos de ausência durante a realização deste curso.

À minha mãe, Maria Vidal de Souza, por todo apoio, compreensão e auxílio sempre presentes em minha vida.

Ao meu amigo e companheiro Paulo Henrique Siqueira, pelo apoio, companheirismo e por sua prontidão em me ajudar em tudo quanto foi possível para a realização deste trabalho.

Aos meus amigos e companheiros de curso Sonia Isoldi Gama Muller, Ademir Alves Ribeiro e demais colegas pelo apoio e companheirismo em grupos de estudo e seminários.

Ao meu colega de trabalho e companheiro do curso de Mestrado Arinei Lindbeck por suas preciosas contribuições em relação às dificuldades de programação encontradas durante a fase de implementação computacional.

Aos Alunos do Curso de Mestrado em Ciência da Computação: Eduardo Costa e Homero de Cuffa, e especialmente ao aluno bolsista do PET/Informática Fabiano Engler os quais foram importantes colaboradores na fase de implementação dos algoritmos estudados.

Aos professores do curso: Maria Teresinha Arns Steiner, Celso Carnieri, Liliana M. Gramani Cumin, Jair Mendes Marques, Mildred Ballin Hecke e Sérgio Scheer, pelos ensinamentos transmitidos durante o curso.

À secretária do CESEC, Maristela Bandil, pela motivação e grande eficiência nos serviços prestados durante o curso.

Aos meus colegas do Departamento de Desenho, da UFPR, pela licença concedida e pelo incentivo durante a realização deste curso.

À professora Simone da Silva Soria Medina, chefe do departamento de Desenho, pelo apoio e pela concessão de espaço e condições para a realização deste trabalho.

À todos os meus amigos que sempre estiveram presentes, me apoiando, dando forças e compreendendo minha ausência durante este período.

A todos, que de alguma forma contribuíram para a realização deste trabalho.

## SUMÁRIO

LISTA DE TABELAS.....	ix
LISTA DE QUADROS.....	xi
LISTA DE FIGURAS.....	xii
LISTA DE SIGLAS.....	xiii
RESUMO.....	xiv
ABSTRACT.....	xv
 1. INTRODUÇÃO.....	 1
1.2 OBJETIVO.....	2
1.3 IMPORTÂNCIA DO TRABALHO.....	3
1.4 DIFICULDADES ENCONTRADAS NO TRABALHO.....	4
1.5 INOVAÇÕES PROPOSTAS NESTE TRABALHO.....	5
1.6 ESTRUTURA DO TRABALHO.....	6
 2. REVISÃO DE LITERATURA.....	 8
2.1 INTRODUÇÃO.....	8
2.2 BREVE HISTÓRICO DA ANÁLISE DE SÉRIES TEMPORAIS.....	9
2.3 TRABALHOS RELACIONADOS.....	10
2.4 DEFINIÇÕES E OBJETIVOS DA ANÁLISE DE SÉRIES TEMPORAIS.....	13
2.4.1 Processos Estocásticos.....	14
2.4.2 Processo Estocástico Ergódico.....	14
2.4.3 Processo Estocástico Estacionário.....	15
2.4.4 Série Temporal Estacionária.....	15
2.4.5 Média de um Processo Estocástico.....	15
2.4.6 Autocovariância do Processo Estocástico.....	16
2.4.7 Variância do Processo Estocástico.....	
2.4.8 Função de Autocorrelação do Processo Estocástico.....	
2.5 MODELOS E MÉTODOS DE PREVISÃO.....	17
2.5.1 Modelos Auto-Regressivos.....	18
2.5.2 Modelos de Médias Móveis.....	19
2.5.3 Modelos Auto-Regressivos e de Médias Móveis (ARMA).....	20
2.6 METODOLOGIA BOX & JENKINS .....	21
2.7 MÉTODOS DE PREVISÃO.....	23
 3. MÉTODOS UTILIZADOS NA PREVISÃO DE SÉRIES TEMPORAIS.....	 26
3.1 PROGRAMAÇÃO GENÉTICA.....	27
3.1.1 Descrição Geral do Algoritmo de Programação Genética.....	27
3.1.2 Representação dos Programas.....	29
3.1.3 Definição dos conjuntos de Funções e Terminais.....	29
3.1.4 Propriedades: Fechamento e Suficiência.....	30
3.1.5 População Inicial.....	31
3.1.6 Função de Aptidão.....	34
3.1.7 Métodos de Seleção.....	35
3.1.8 Operadores Genéticos.....	38
3.1.9 Parâmetros utilizados na Programação Genética.....	40

3.1.10 Critério de Parada.....	41
3.2 EXEMPLO DO ALGORITMO DE PROGRAMAÇÃO GENÉTICA.....	41
4. CONJUNTO DE PREDITORES.....	46
4.1 BREVE HISTÓRICO DO ALGORITMO BOOSTING.....	46
4.2 ADABOOST PARA PROBLEMAS DE CLASSIFICAÇÃO BINÁRIO.....	47
4.3 BOOSTING PARA PROBLEMAS DE REGRESSÃO.....	49
4.3.1 O Algoritmo <i>Adaboost.R</i> .....	50
4.3.2 O Algoritmo <i>Adaboost.RT</i> .....	52
4.3.3 O Algoritmo <i>GPBoost</i> .....	54
4.3.4 Exemplo do Algoritmo <i>GPBoost</i> .....	55
4.4 ADAPTAÇÃO DO ALGORITMO <i>BOOSTING</i> UTILIZANDO COEFICIENTE DE CORRELAÇÃO.....	58
4.4.1 Algoritmo <i>Boosting</i> utilizando o Coeficiente de Correlação ( <i>BCI – Boosting Correlation Improvement</i> ).....	59
5. TESTES DE HIPÓTESES.....	63
5.1 TESTE “ <i>t</i> ” - DESEMPENHO DE ALGORITMOS.....	63
5.2 COMPARAÇÃO DOS ERROS NO CONJUNTO DE TESTE.....	65
5.3 DESEMPENHO DOS ALGORITMOS - TESTE ANOVA.....	66
5.3.1 Teste F para comparar duas ou mais Médias Populacionais.....	66
5.3.2 Teste Tukey-Kramer.....	68
5.3.3 Pressupostos para se utilizar a ANOVA.....	69
6. EXPERIMENTOS E SIMULAÇÕES.....	70
6.1 EXPERIMENTO UTILIZANDO SÉRIES TEMPORAIS REAIS.....	70
6.1.1 Configuração da Programação Genética.....	71
6.1.2 Configuração do algoritmo <i>GPBoost</i> .....	72
6.1.3 Configuração da metodologia Box & Jenkins.....	73
6.2 DESCRIÇÃO DOS EXPERIMENTOS.....	73
6.2.1 Dados de Entrada.....	73
6.2.2 Experimento I.....	74
6.2.3 Experimento II.....	76
6.2.4 Análise do desempenho do algoritmo <i>BCIGP</i> utilizando o teste “ <i>t</i> ”.....	78
6.3 INTERPRETAÇÃO E ANÁLISE DAS SÉRIES FINANCEIRAS – UMA APLICAÇÃO REAL.....	78
6.3.1 Descrição das séries financeiras utilizadas.....	79
6.3.2 Retornos.....	80
6.3.3 Avaliação da estratégia de ação baseada nas previsões.....	81
7. SIMULAÇÃO MONTE CARLO.....	84
7.1 DESCRIÇÃO DA SIMULAÇÃO.....	84
7.1.1 Definição dos Parâmetros.....	86
7.1.2 Espaço Paramétrico das Estruturas <i>AR(1)</i> e <i>MA(1)</i> .....	86
7.1.3 Espaço Paramétrico das Estruturas <i>AR(2)</i> e <i>MA(2)</i> .....	87
7.1.4 Espaço Paramétrico da Estrutura <i>ARMA(1, 1)</i> .....	88
7.1.5 Simulação das Séries.....	89
7.1.6 Banco de Dados.....	89



7.2 EXECUÇÃO DO EXPERIMENTO.....	90
7.2.1 Configuração da Programação Genética e do <i>GPBoost</i> .....	90
7.2.2 Ambiente Computacional.....	90
7.2.3 Análise dos Resultados.....	91
7.3 EXPERIMENTO UTILIZANDO PESOS 2.....	99
7.4 CONCLUSÕES.....	104
8. EXTENSÃO DO MÉTODO BCI À PROBLEMAS DE REGRESSÃO MÚLTIPLA	106
8.1 BREVE HISTÓRICO SOBRE PROBLEMAS DE REGRESSÃO.....	106
8.2 APLICAÇÃO DO MÉTODO BCI A PROBLEMAS DE REGRESSÃO.....	107
8.2.1 Descrição das Bases de Dados.....	108
8.2.2 Preparação dos Dados.....	109
8.2.3 Atualização dos pesos do algoritmo BCI para problemas de regressão.....	109
9. CONCLUSÃO E TRABALHOS FUTUROS	112
9.1 CONCLUSÃO.....	112
9.2 FUTUROS TRABALHOS.....	113
REFERÊNCIAS.....	115
APÊNDICE A.....	121
APÊNDICE B.....	123
APÊNDICE C.....	124

## LISTA DE TABELAS

TABELA 3.1 – VALORES DE ENTRADA E SAÍDA NO CONJUNTO DE TREINAMENTO	42
TABELA 3.2 – PARÂMETROS DA PG.....	43
TABELA 3.3 – VALORES DE SAÍDA E MELHORES INDIVÍDUOS ENCONTRADOS NAS GERAÇÕES DE 0 A 3.....	44
TABELA 4.1 – PARÂMETROS PARA O GPBOOST	56
TABELA 4.2 - EVOLUÇÃO DOS PESOS NAS TRÊS PRIMEIRAS EXECUÇÕES DO BOOSTING.....	56
TABELA 4.3 – RESULTADOS OBTIDOS DAS EXECUÇÕES BOOSTING.....	57
TABELA 6.1 – BASES DE DADOS.....	74
TABELA 6.2 – BASES DE DADOS SÉRIES FINANCEIRAS.....	74
TABELA 6.3 - COMPARAÇÃO DAS PREVISÕES ARMA, PG E GPBOOST.....	75
TABELA 6.4 - COMPARAÇÃO DO RMSE NO CONJUNTO DE TESTE PARA CADA MÉTODO.....	76
TABELA 6.5 - COMPARAÇÃO DAS DIFERENÇAS ENTRE OS PARES DE ALGORITMOS.....	78
TABELA 6.6 – RETORNO FINANCEIRO NO PERÍODO 110 DIAS.....	82
TABELA 6.7 – RETORNO FINANCEIRO ANUALIZADO.....	82
TABELA 6.8 – NÚMERO DE TRANSAÇÕES NO PERÍODO.....	83
TABELA 7.1 – NÚMERO DE PARÂMETROS E PARÂMETROS E SÉRIES.....	89
TABELA 7.2 – NÚMERO DE PARÂMETROS E PARÂMETROS E SÉRIES.....	89
TABELA 7.3 – MSE MÉDIO DAS 500 SÉRIES PARA A ESTRUTURA AR(1).....	92
TABELA 7.4 – MÉDIA DOS MSE DOS PARÂMETROS PARA CADA ESTRUTURA.....	95
TABELA 7.5 – RESULTADOS DA ANOVA E TESTE TUKEY-KRAMER PARA A ESTRUTURA MA(1).....	96
TABELA 7.6 – VALOR-p AR(1) E AR(2) .....	98
TABELA 7.7 – VALOR - p – MA (1) E VALOR - P – MA (2).....	98
TABELA 7.8 – VALOR - P – ARMA (1,1) .....	99
TABELA 7.10 – MSE MÉDIO PARA A ESTRUTURA MA(1) – ATUALIZAÇÃO DE PESOS 2.....	100
TABELA 7.11 - RESULTADOS DA ANOVA E TESTE TUKEY-KRAMER PARA A ESTRUTURA MA(1) – PESOS 2.....	101
TABELA 7.12 – VALOR - p MA(1) PESOS 2 .....	104
TABELA 8.1 – CONJUNTOS DE DADOS.....	108
TABELA 8.2 – NÚMERO E TIPO DE VARIÁVEIS.....	108
TABELA 8.3 – CONJUNTOS DE DADOS.....	110
TABELA 8.4 – COMPARAÇÃO DO RMSE NOS 10 CONJUNTOS DE TESTE.....	111
TABELA C1 – RESULTADOS DA DO MSE MÉDIO PARA AS 500 SÉRIES DA ESTRUTURA MA(1) .....	124
TABELA C2 – RESULTADOS DA DO MSE MÉDIO PARA AS 500 SÉRIES DA ESTRUTURA AR(2) .....	125
TABELA C3 – RESULTADOS DA DO MSE MÉDIO PARA AS 500 SÉRIES DA ESTRUTURA MA(2) .....	126
TABELA C4 – RESULTADOS DA ANOVA E TESTE TUKEY-KRAMER PARA A ESTRUTURA AR(1) .....	131
TABELA C5 – RESULTADOS DA ANOVA E TESTE TUKEY-KRAMER PARA A ESTRUTURA AR(2) .....	133

TABELA C6 – RESULTADOS DA ANOVA E TESTE TUKEY-KRAMER PARA A ESTRUTURA MA(2) .....	135
TABELA C7 – RESULTADOS DA ANOVA E TESTE TUKEY-KRAMER PARA A ESTRUTURA ARMA(1,1).....	137

## LISTA DE QUADROS

QUADRO 4.1 – ALGORTIMO ADABOOST.....	49
QUADRO 4.2 – ALGORITMO ADABOOST.R.....	52
QUADRO 4.3 – ALGORITMO ADABOOST.RT.....	55
QUADRO 4.4 – ALGORITMO BCI.....	61
QUADRO 6.1 – PARÂMETROS DE CONFIGURAÇÃO DA PG.....	71
QUADRO 6.2 – MELHOR INDIVÍDUO GERADO PELA PG.....	72

## LISTA DE FIGURAS

FIGURA 2.1 – EXEMPLO DE UMA SÉRIE TEMPORAL.....	13
FIGURA 2.2 – ESTRUTURA DE FUNCIONAMENTO DOS MODELOS ARMA.....	23
FIGURA 3.1 – ALGORITMO BÁSICO DE PROGRAMAÇÃO GENÉTICA.....	28
FIGURA 3.2 – ÁRVORE DE SINTAXE DA PROGRAMAÇÃO GENÉTICA.....	29
FIGURA 3.3 – ÁRVORE COM PROFUNDIDADE MÁXIMA QUATRO (4) INICIALIZADA PELO MÉTODO GROW.....	32
FIGURA 3.4 – ÁRVORE DE PROFUNDIDADE MÁXIMA TRÊS INICIALIZADA PELO MÉTODO FULL.....	32
FIGURA 3.5 – REPRESENTAÇÃO DO OPERADOR DE CRUZAMENTO DE UM PONTO.....	39
FIGURA 3.6 – REPRESENTAÇÃO DO OPERADOR DE MUTAÇÃO.....	40
FIGURA 3.7 - MELHOR INDIVÍDUO NA GERAÇÃO 0.....	43
FIGURA 3.8 - REPRESENTAÇÃO GRÁFICA DOS MELHORES INDIVÍDUOS DAS GERAÇÕES 0-3.....	45
FIGURA 5.1 – NÍVEL DE CONFIANÇA P E NÍVEL DE SIGNIFICÂNCIA .....	64
FIGURA 6.1 – COMPARAÇÃO DO RMSE DA PG E ARMA.....	75
FIGURA 6.2 – COMPARAÇÃO DO RMSE DA PG, GPBOOST, BCIGP E ARMA....	77
FIGURA 6.3 – VALORES DE PREVISÃO OBTIDOS PELA PG E BCIGP PARA A SÉRIE IPI.....	77
FIGURA 6.4 – RETORNOS DIÁRIOS DA IBOVESPA (03/01/1995 – 27/12/2000)....	81
FIGURA 7.1 – FLUXOGRAMA DA SIMULAÇÃO MONTE CARLO.....	85
FIGURA 7.2 – REGIÃO DE ESTACIONARIDADE PARA OS MODELOS AR(1) E MA(1).....	86
FIGURA 7.3 – REGIÃO DE ESTACIONARIDADE PARA UM MODELO AR(2).....	87
FIGURA 7.4 – REGIÃO DE ESTACIONARIDADE PARA UM MODELO ARMA(1,1) .....	88
FIGURA 7.5 – COMPORTAMENTO DO MSE PARA OS MÉTODOS – <i>e136</i> .....	93
FIGURA 7.6 – COMPORTAMENTO DO MSE PARA OS MÉTODOS - <i>e150</i> .....	93
FIGURA 8.1 – PARTIÇÃO DO VETOR DE PESOS.....	110
FIGURA C1 – COMPORTAMENTO DOS MÉTODOS PARA O 1º VALOR PREVISTO <i>e136</i> – AR(2).....	127
FIGURA C2 – COMPORTAMENTO DOS MÉTODOS PARA O ÚLTIMO VALOR PREVISTO <i>e150</i> – AR(2) .....	127
FIGURA C3 – COMPORTAMENTO DOS MÉTODOS PARA O 1º VALOR PREVISTO <i>e136</i> – MA(1) .....	128
FIGURA C4 – COMPORTAMENTO DOS MÉTODOS PARA O ÚLTIMO VALOR PREVISTO <i>e150</i> – MA(1) .....	128
FIGURA C5 – COMPORTAMENTO DOS MÉTODOS PARA O 1º VALOR PREVISTO <i>e136</i> – MA(2) .....	129
FIGURA C6 – COMPORTAMENTO DOS MÉTODOS PARA O ÚLTIMO VALOR PREVISTO <i>e150</i> – MA(2).....	129
FIGURA C7 – COMPORTAMENTO DOS MÉTODOS PARA O 1º VALOR PREVISTO <i>e136</i> – ARMA(1, 1) .....	130
FIGURA C8 – COMPORTAMENTO DOS MÉTODOS PARA O ÚLTIMO VALOR PREVISTO <i>e150</i> – ARMA(1, 1).....	130

## LISTA DE SIGLAS

PG	- Programação Genética
RMSE	- Median Square Error
MSE	- Median Square Error
CE	- Computação Evolucionária
BCI	- Boosting Correlation Improvement
BCIGP	- Boosting Correlation Improvement Genetic Programming
ARMA	- Autoregressive and Moving Average
AR	- Autoregressive
MA	- Moving Average
UFPR	- Universidade Federal do Paraná
IA	- Inteligência Artificial
RNA	- Redes Neurais Artificiais
AG	- Algoritmos Genéticos
AE	- Algoritmo Evolucionário
ANOVA	- Análise de Variância
CART	- Classification and Regression Tree

## RESUMO

Nas Ciências Econômicas, assim como na Engenharia e nas Ciências Naturais, ocorrem fenômenos que dependem da observação de dados em intervalos de tempo, durante um período específico. Estas observações de valores são denominadas de Séries Temporais. As técnicas disponíveis para analisar estas observações são denominadas de Análise de Séries Temporais. Este conjunto de técnicas tem por objetivo construir um modelo com número adequado de parâmetros estimados de forma a ajustar o modelo à série temporal. A obtenção de um modelo adequado é de extrema importância, pois pode revelar algumas características da série temporal que ajudam na previsão de seus valores futuro, ou simplesmente descrevem seu comportamento. Esta habilidade de prever valores futuros para uma série temporal tem grande relevância prática. A construção de um modelo estatístico adequado para ajustar os dados apresenta um grau de dificuldade razoável. Se o modelo escolhido não fornecer um bom ajuste para a série, a uma previsão enganosa e ineficiente. Para a construção destes modelos, novas técnicas da área de Inteligência Artificial vêm sendo aplicadas e têm apresentado um bom desempenho, dentre estas técnicas destacam-se: Redes Neurais Artificiais, Algoritmos Evolucionários Sistemas Nebulosos. A contribuição original deste trabalho é a proposta de uma metodologia de combinação de preditores utilizando os coeficientes de correlação (*BCI – Boosting Correlation Improvement*) entre os valores observados e os valores previstos por um algoritmo básico de previsão. Após vários preditores terem sido gerados, os mesmos são combinados entre si utilizando o coeficiente de correlação, para gerar um único preditor. A previsão é efetuada  $n$  passos à frente. O método é testado em vários experimentos utilizando séries reais e séries reais financeiras. Com o objetivo de validar a metodologia na Previsão de Séries Temporais é também realizada uma simulação Monte Carlo, onde são geradas séries artificiais, distribuídas em todo o espaço paramétrico das principais estruturas dos modelos ARMA, que são AR(1), AR(2), MA(1), MA(2) e ARMA(1, 1). O banco de dados que foi gerado contém 214.000 séries temporais geradas artificialmente, onde cada uma das séries possui 150 valores, dos quais 90% formam o conjunto de treinamento ficando os outros 10% para o conjunto de teste. Os resultados obtidos através do algoritmo BCI foram comparados aos resultados obtidos através dos modelos ARMA, Programação Genética e *GPBoost*. Todas as comparações foram feitas no conjunto de teste e mostraram que o algoritmo BCI fornece melhores resultados que os demais métodos analisados. O algoritmo proposto é também testado nos problemas de regressão múltipla. Os resultados obtidos são comparados à outros métodos de previsão, tais como a metodologia Box & Jenkins, a PG tradicional, e PG com *Boosting*. Na tarefa de regressão os resultados são comparados aos resultados obtidos através de Redes Neurais Artificiais, Modelos de regressão utilizando árvores de indução e *Boosting*. Para comparar o desempenho obtido pelo algoritmo proposto e os demais métodos, foram utilizadas diversas medidas estatísticas, tais como o Erro Médio Quadrático, a Raiz Quadrada do Erro Médio Quadrático e os testes de hipótese (Teste “ $t$ ”; ANOVA e Teste Tukey). Nos testes realizados, a metodologia proposta fornece previsões com erros de previsão menores do que os obtidos com a utilização das outras técnicas.

## ABSTRACT

In Economic Sciences, Engineering and Natural Sciences, there is a phenomena that depends of the data observation in time intervals during a specific period. This collection of observed data is called Time Series and the methodology to analyze this data is known as Time Series Analysis. To analyze time series there is a techniques set used to construct models with reduced number of estimated parameters to fit the observed data. This model must be able to fit the data set in the best way. The model can show some characteristics of the series that can help in the forecast of the future values of the time series or simply describing its behavior. This ability to forecast the future values of the time series has a great practical relevance. The task to obtain a good statistical model to fit the time series data is not too easy. A bad choice of a model can lead a bad and inefficient prediction. In way to construct these models, new techniques have been developed in the Artificial Intelligence and the results have been good, between these techniques the most important are: Artificial Neural Network, Evolutionary Algorithms and Hybrid Algorithms. The original contribution of this work is the proposal of a new methodology (BCI – Boosting Correlation Improvement) to combine predictors, this is based in the correlation coefficients between the observed and predicted values obtained from a basic prediction algorithm. After generate some predictors, they are combined between itself using the correlation coefficients to get an only predictor, that supplies better results than if is used only one predictor. The method is tested in some experiments using real and financial time series. To validate this methodology in time series forecasting, a Monte Carlo simulation was made. In this simulation, artificial series were generated from the parametric space of the principal ARMA models, they are AR(1), AR(2), MA(1), MA(2) e ARMA(1,1). The database generated was composed by 214.000 time series with 150 observations each one. The training set is composed by 90% of date and the others 10% composes the test set. The results were compared out of sample and the BCI showed better performance than ARMA methodology, Genetic Programming and *GPBoost*. The BCI algorithm was also applied to multiple regressions problem and the results obtained from this method were compared with the results from Artificial Neural Network, Model Tree and Boosting. This comparison showed that the BCI supplied better results than other ones. In way compare the performance of the BCI methodology with other methods, many statistical tests were performed such as Median Square Error (MSE), Root Median Square Error (RMSE) and hypothesis tests (*t*-Test, ANOVA and Tukey-Test). In all the tests the results were compared out of sample and the BCI methodology showed to be able to improve the forecasts.



# Capítulo 1

## Introdução

Com o crescimento industrial, econômico e tecnológico, a necessidade de se efetuar previsões que auxiliem no planejamento empresarial torna-se cada vez mais importante. A previsão de determinados fatos auxilia a tomada de decisões que poderão melhorar o desempenho das empresas ou até mesmo minimizar prejuízos. Alguns exemplos da utilidade das previsões são:

- A minimização dos prejuízos causados por catástrofes, como enchentes ou terremotos, caso estas possam ser previstas com um intervalo de tempo razoável;
- Uma melhor taxa de retorno de investimento de aplicações, caso os investidores possuam uma previsão do preço das ações;
- Adequação da produção à demanda do mercado através de um planejamento da produção, tendo em mãos a previsão de utilização, venda e consumo dos produtos.

Enfim, muitas são as formas de se utilizar a previsão de modo a obter melhores resultados em problemas reais.

Diversas técnicas para realizar a previsão de séries temporais são utilizadas, porém todas possuem um erro de previsão. O desafio é, então minimizar o erro de previsão, construir um modelo que melhor se ajuste aos dados e utilizar um número parcimonioso de parâmetros. A construção deste modelo não é uma tarefa simples e muitas vezes conduz a uma previsão pouco eficiente. Os métodos convencionais de previsão fornecem resultados precisos quando

os dados estudados apresentam um comportamento linear, porém quando há um grau elevado de não linearidade, estes métodos passam a ser pouco eficientes (CORTEZ, 2002).

Os métodos mais difundidos são os modelos Auto-Regressivos (AR), modelos de Médias Móveis (MA - *Moving Average*) e os modelos Auto-Regressivos e de Médias Móveis (ARMA – *Auto Regressive Moving Average*). A metodologia Box & Jenkins (BOX & JENKINS, 1970) é a mais eficiente e a mais utilizada para a previsão de séries temporais, porém sua aplicação envolve uma teoria de alta complexidade e a tarefa de identificação do melhor modelo a ser utilizado não é simples. Além disso, a estimação dos parâmetros envolve métodos de programação não-linear. Outra dificuldade que se pode citar na aplicação desta técnica é o tamanho da série, que deve conter pelo menos 50 observações [CHAVES, 1991].

Neste sentido, pesquisas vêm sendo desenvolvidas, nas mais diversas áreas, para se obter modelos de previsão com o menor erro possível sem exigir interferências de um analista.

## 1.2 OBJETIVO

O objetivo geral deste trabalho é apresentar uma metodologia alternativa para a previsão de séries temporais, utilizando um conjunto de técnicas da Programação Genética (PG) e em seguida uma combinação de preditores (*Boosting*) com o intuito de minimizar os erros de previsão. Para atingir esta meta, os objetivos específicos são:

- Generalizar a metodologia de forma a tratar séries temporais que apresentam características diversas tais como: volatilidade<sup>1</sup>, tendência ou sazonalidade;
- Apresentar uma nova metodologia de combinação final de preditores, fornecendo assim previsões mais eficientes, especialmente quando a análise dos resultados é feita fora da amostra, ou seja, no conjunto de teste;
- Apresentar uma nova metodologia na atualização dos pesos dos algoritmos de *boosting*;
- Aplicar o algoritmo proposto a séries financeiras, analisando também a estratégia de ação dos investidores indicando qual é o momento de movimentar as ações;

---

<sup>1</sup> Volatilidade é a variância condicional de uma série temporal.

- Avaliar a metodologia proposta através de:
  - testes com séries reais financeiras e algumas séries reais encontradas da literatura (MORETIN & TOLOI, 2004);
  - Simulação Monte Carlo analisando todas as séries geradas dentro do espaço paramétrico da região de estacionariedade da metodologia Box & Jenkins.
- Verificar o desempenho do algoritmo proposto (BCI) quando aplicado aos problemas de regressão múltipla;

### 1.3 IMPORTÂNCIA DO TRABALHO

Diversas organizações estão interessadas em planejar sua estratégia de ação, baseadas em previsões efetivas que sejam capazes de fornecer dados suficientes para guiar o processo de tomada de decisão. No planejamento de produção, pode-se prever o impacto dos recursos financeiros, de *marketing* e de recursos humanos; nas aplicações financeiras, pode-se prever taxa de inflação, taxas de juros e valores de indicadores; previsões sobre o número de manchas solares na tomada de decisões sobre missões espaciais e sobre a colocação de satélites em órbita. Se as previsões obtidas possuem um alto grau de confiabilidade, as organizações poderão planejar seus empreendimentos de uma maneira mais segura e efetiva. Embora os erros de previsão sejam inevitáveis, os métodos de previsão buscam minimizar estes erros e fornecer dados que possam servir de suporte ao planejamento e estratégias de ação.

Os métodos convencionais de previsão de séries temporais, na sua maioria, exigem a estacionariedade da série, o que não ocorre com as séries financeiras e econômicas, tornando-se assim, difícil efetuar uma previsão com valores mais próximos dos reais desejados.

Neste trabalho apresenta-se uma metodologia alternativa para a previsão de séries temporais, baseada na combinação de preditores, obtidos através de um algoritmo básico de previsão. Para fazer esta combinação utiliza-se o coeficiente de correlação entre os valores observados no conjunto de treinamento e os valores previstos. O objetivo é de minimizar ou até mesmo superar as dificuldades encontradas na aplicação de outras técnicas e métodos de

previsão, especialmente quanto à dificuldade da obtenção de modelos adequados às séries não estacionárias e que apresentem um alto grau de volatilidade.

#### 1.4 DIFICULDADES ENCONTRADAS NO TRABALHO

Após o algoritmo proposto ter sido aplicado com sucesso em experimentos utilizando algumas séries reais, verificou-se a necessidade de estudar o seu comportamento na previsão de outros tipos de séries. Para isto foi escolhido trabalhar com séries artificiais. Foi então realizada uma simulação Monte Carlo, na qual foram geradas séries sintéticas pertencentes à estrutura ARMA.

Na geração destas séries foi considerado todo o espaço paramétrico das estruturas mais utilizadas no estudo de séries temporais, que são as estruturas AR(1), AR(2), MA(1), MA(2) e ARMA(1,1). O espaço paramétrico de estacionariedade de cada uma destas estruturas foi dividido com incremento de 0,1 unidades nos eixos  $x$  e  $y$ . Seguindo esta metodologia, foram geradas 214.000 séries com 150 observações cada uma, tendo sido utilizada uma semente aleatória para a PG e dez execuções do algoritmo de *Boosting*. O detalhamento desta simulação é apresentado no capítulo V desta tese. Para a execução de todo este banco de dados, foi necessária a otimização de utilização dos programas desenvolvidos de forma a obter os resultados num período de tempo viável. Foi utilizado um grupo de computadores e todo o processo foi gerenciado de forma que nenhum dado fosse perdido. Os dados foram processados utilizando um grupo de 64 computadores pertencentes ao *cluster* da Universidade Federal do Paraná (UFPR).

Os estudos realizados na busca de uma nova forma de atualização dos pesos do algoritmo de *Boosting* e da combinação final dos preditores, de modo a obter previsões mais precisas, não foi uma tarefa simples, já que a PG não é considerada um algoritmo base<sup>2</sup>, sendo assim maior a dificuldade de melhoria nos resultados obtidos pela aplicação da PG tradicional.

Para se chegar a uma boa definição dos parâmetros a serem utilizados, tais como número de sementes iniciais da PG, número de variáveis a serem utilizadas, taxas de

---

<sup>2</sup> Os algoritmos de *Boosting* são em geral aplicados juntamente com algoritmos “fracos” ou “base” da área de Aprendizado de Máquina.

aplicação dos operadores genéticos e número de execuções do algoritmo *boosting* a serem utilizados, foi baseada na realização de muitos experimentos. Na fase de análise dos resultados e verificação da relevância do algoritmo proposto, foram realizados diferentes testes estatísticos para a validação da metodologia proposta. O detalhamento destes testes encontra-se no capítulo V desta tese.

A implementação do método para problemas de regressão multivariada exigiu várias modificações na metodologia proposta, de forma a adaptá-la a problemas que possuem mais de uma variável de entrada.

### 1.5 INOVAÇÕES PROPOSTAS NESTE TRABALHO

A previsão de todas as séries artificiais obtidas através da simulação Monte Carlo, utilizada neste trabalho, variando os parâmetros das estruturas AR(1), AR(2), MA(1), MA(2) e ARMA(1,1) nos seus respectivos espaços paramétricos efetuando a comparação através de 4 diferentes métodos.

Diversas adaptações foram realizadas para a utilização das técnicas de *Boosting* e da PG associadas na Previsão de Séries Temporais. A técnica de *Boosting* foi até então, utilizada somente para problemas de regressão e classificação.

A maior contribuição deste trabalho se dá na apresentação de uma nova metodologia para o algoritmo de *Boosting*, a qual é baseada no coeficiente de correlação entre os valores observados no conjunto de treinamento e os valores previstos pela PG.

Em sua maioria, os algoritmos de *Boosting* existentes, são aplicados à problemas de classificação. Poucos trabalhos na área de regressão têm sido realizados utilizando estes algoritmos, dentre os quais pode-se citar Drucker (1997). Assim, foi realizada uma extensão da metodologia à problemas de regressão múltipla, para analisar o desempenho do algoritmo proposto.

A acurácia do algoritmo proposto é considerada satisfatória, pois apresentou bons resultados quando comparado aos resultados obtidos através de outras metodologias encontradas na literatura. A maioria dos métodos existentes apresenta bom desempenho

somente no conjunto de treinamento e quando aplicados ao conjunto de teste, as taxas de erro aumentam substancialmente.

## 1.6 METODOLOGIA UTILIZADA NA PESQUISA E ESTRUTURA DO TRABALHO

Nesta seção são apresentadas a metodologia utilizada nesta pesquisa, bem como a estrutura do trabalho. A primeira etapa da realização desta pesquisa, apresentada no capítulo II, é o estudo teórico sobre a Análise de Séries Temporais e os métodos de previsão clássicos utilizados. Neste capítulo é também apresentado um breve histórico sobre Previsão de Séries Temporais e os trabalhos mais recentemente publicados nesta área. A metodologia de previsão clássica escolhida para ser comparada aos novos métodos de previsão é a metodologia Box & Jenkins. Os modelos utilizados são os modelos Auto-Regressivos (AR), de Médias Móveis (MA) e os modelos mistos Auto-Regressivos e de Médias Móveis (ARMA).

No capítulo III é apresentada a técnica de Programação Genética e sua adaptação para a utilização em problemas de Previsão de Séries Temporais. No capítulo IV é apresentado um estudo teórico sobre combinação de preditores utilizando o algoritmo *Boosting* e são apresentadas as várias versões existentes deste algoritmo. Na conclusão do capítulo e dos estudos teóricos é apresentada a nova metodologia do algoritmo de *Boosting*, contribuição principal desta tese. Este novo algoritmo, para a combinação de preditores, utiliza os coeficientes de correlação entre os valores observados e os valores previstos por algum preditor básico, esta nova técnica e os algoritmos são detalhados.

Para a verificação do desempenho do algoritmo proposto, são apresentados no capítulo V os testes de hipótese, o teste “*t*” e de Análise da Variância (ANOVA). Para a validação da metodologia proposta são apresentados no capítulo VI os experimentos realizados. A partir do primeiro experimento, realizado com previsão de séries reais, foram definidos os parâmetros para a Programação Genética e o número de algoritmos *Boosting* que são necessários para a obtenção de melhorias nos resultados obtidos, ao se utilizar a combinação de preditores. No segundo experimento os métodos clássicos de previsão são comparados aos obtidos pelo algoritmo proposto nesta tese (BCI – *Boosting Correlation Improvement*), utilizando a raiz quadrada do erro quadrático médio (RMSE) e também

fazendo a análise do valor-p obtido através do teste “ $t$ ”. Com base na análise dos resultados obtidos no capítulo VI, é realizada uma Simulação Monte Carlo para a verificação do desempenho do algoritmo proposto numa base de dados maior. Para realizar esta simulação são geradas séries artificiais analisando o espaço paramétrico da região de estacionariedade das estruturas AR(1), AR(2), MA(1), MA(2) e ARMA(1, 1). Para cada série gerada, é feita a previsão  $n$  passos à frente, utilizando o modelo ARMA( $p, q$ ) através do qual a série é gerada; o algoritmo da PG tradicional; o algoritmo de *Boosting* que utiliza a PG como algoritmo base (*GPBoost*) e finalmente o algoritmo proposto nesta tese (BCIGP) que também utiliza como algoritmo básico a PG. Os resultados são comparados utilizando o erro médio quadrático (MSE), o teste “ $t$ ” e ANOVA.

Tendo em vista que o algoritmo de *Boosting* é um algoritmo que pode ser utilizado não somente para previsão de séries temporais, como também para problemas de regressão multivariada, é realizado um experimento utilizando o algoritmo BCI para este tipo de problema. O detalhamento deste experimento encontra-se no capítulo VIII, desta tese. Finalmente, no capítulo IX são apresentadas as conclusões e sugestões para futuros trabalhos.

## Capítulo 2

### Revisão de Literatura

Neste capítulo é apresentada uma breve introdução sobre a Análise de Séries Temporais. Além disso, um histórico no qual são citados alguns dos principais trabalhos encontrados na literatura sobre Análise de Séries Temporais e trabalhos relacionados ao tema desta tese.

#### 2.1 INTRODUÇÃO

O principal objetivo da Análise de Séries Temporais é prever valores futuros da série, que podem ser: a curto prazo, como no caso de vendas e produção ou estoque; ou a longo prazo, como para séries populacionais de produtividade e outros (MORETTIN; TOLOI, 1985). A tomada de decisões baseada em previsões abrange várias áreas de Gestão, tais como: *Marketing*, Produção, Área Financeira, Gestão Empresarial e Recursos Humanos (MAKRIDAKIS; WHEELWRIGHT, 1989). As empresas devem tomar o cuidado de utilizar as previsões como base de suas decisões, considerando que as mesmas estão sujeitas à erros de previsão inerentes à qualquer experiência aleatória. Qualquer que seja a metodologia utilizada o objetivo é minimizar estes erros de previsão (CORTEZ, 2002).



Cortez cita que: “O desejo de compreender o passado e prever o futuro impulsiona a procura por leis que expliquem o comportamento de certos fenômenos ou acontecimentos. Se as equações que os determinam são conhecidas, é possível utilizar as mesmas para prever o resultado de um determinado experimento, desde que sejam conhecidas as condições iniciais. Na ausência de regras que definam o comportamento de um sistema, procura-se determinar o seu comportamento futuro a partir de observações concretizadas no passado”.

Uma das técnicas mais comuns para efetuar previsões é aquela baseada em observações cronologicamente ordenadas da variável em questão (CHATFIELD, 2000).

Uma previsão é uma manifestação relativa a sucessos ou insucessos desconhecidos em um futuro determinado (BARBANCHO, 1970). A previsão não constitui um fim em si, mas um meio de fornecer informações e subsídios para uma conseqüente tomada de decisões, visando atingir determinados objetivos.

## 2.2 BREVE HISTÓRICO DA ANÁLISE DE SÉRIES TEMPORAIS

Em geral, os métodos de previsão partem de valores passados para prever valores futuros. Até 1920, a previsão era efetuada através de extrapolação simples de um valor global, ajustado em função do tempo. Só em 1927, Yule criou o modelo Auto-Regressivo (AR), onde o valor previsto dependia de valores passados. Por duas décadas seguintes, os pesquisadores da área de Estatística, acreditaram que o comportamento de uma série dependia dos modelos lineares e do ruído. Com o passar do tempo, porém, percebeu-se que modelos lineares não eram suficientes para a análise de Séries Temporais, já que a maioria das séries reais apresentam fortes tendências de não linearidade (CHAVES, 1991).

A partir de 1950, novos estudos foram realizados na área da Previsão de Séries Temporais. Nesta época foi criada a técnica do Alisamento Exponencial, que exigia poucos cálculos e era de fácil utilização. Com o advento do computador nos anos 50, houve uma revolução nos métodos de previsão.

Em 1976, surgiu a metodologia Box & Jenkins, trata-se de uma técnica completa de identificação do modelo que melhor se ajusta aos dados, de estimação dos parâmetros e da verificação da validade do modelo empregado. Esta técnica fornece previsões através de um modelo linear probabilístico, embora apresente algumas restrições de utilização, tais como o

tamanho da amostra (que deve ser de no mínimo 50 observações) e a imprescindível interferência de um analista, na fase de identificação do modelo (CHAVES, 1991).

Desde então, muitas técnicas têm sido testadas com o intuito de se efetuar previsões mais precisas e preferencialmente automatizadas. Muitos estudos na área da Inteligência Artificial (IA) foram dirigidos a este fim. Modelos inspirados na natureza, como Redes Neurais Artificiais (RNA), Algoritmos Genéticos (AG) e Algoritmos Evolutivos (AE) e, dentre estes, a Programação Genética (KABOUDAN, 2002), apresentaram resultados promissores nesta área.

## 2.3 TRABALHOS RELACIONADOS

Alguns dos trabalhos encontrados na literatura que estão relacionados com o presente trabalho são o de Povinelli (1999), que utiliza conceitos de Mineração de Dados para analisar os dados de uma Série Temporal e identificar eventos<sup>3</sup> que possam ser significativos nos dados observados. O autor utilizou conceitos básicos de Algoritmos Genéticos, compostos com uma busca Monte Carlo para gerar a população inicial para o AG. Esta técnica é denominada *TSDM (Time Series Data Mining)*, cria um conjunto de métodos que revelam padrões temporais escondidos que são característicos e preditivos de eventos em uma Série Temporal. O método foi aplicado com sucesso em Séries Temporais complexas e não estacionárias e teve sua aplicação principal no domínio das séries financeiras.

Kaboudan (1999) apresentou um trabalho no qual utilizou a Programação Genética para estimar a previsibilidade de séries temporais de índices financeiros de bolsa de valores. Procurou encontrar o melhor modelo de ajuste para as mesmas usando PG e minimizando a soma dos erros quadráticos. Sua métrica de previsibilidade foi estabelecida com base na comparação da soma dos erros quadráticos entre a série original e o resultado obtido pela PG.

Kaboudan (2000) utilizou a PG na previsão de Séries Temporais Reais, como a quantidade de manchas solares. O autor criou um *software* (TSGP)<sup>4</sup>, que realiza previsões de séries temporais utilizando a PG e seus resultados são comparados aos resultados obtidos

---

<sup>3</sup> Evento – uma ocorrência importante; por exemplo numa série temporal sísmica, um terremoto é um evento.

<sup>4</sup> TSGP – fonte: [http://bulldog2.redlands.edu/fac/mak\\_kaboudan/](http://bulldog2.redlands.edu/fac/mak_kaboudan/)

utilizando os modelos ARMA e os erros de previsão obtidos através da PG mostram erros de previsão menores. Vale salientar que em seu trabalho, Kaboudan utilizou a Programação Genética pura; já a intenção neste trabalho é utilizar, além do algoritmo da Programação Genética, algumas técnicas de combinação de preditores na tentativa de melhorar os resultados ainda mais, ou seja, reduzir os erros de previsão.

Duan (2001) propôs uma nova métrica para medir a previsibilidade de uma Série Temporal, baseado no trabalho de Kaboudan (2000) e o aplicou ao investimento em ações. Seu trabalho é uma extensão do trabalho de Kaboudan e supera algumas limitações daquele trabalho como a de que o valor da métrica depende do tamanho da série.

Cortez (2002) propôs uma abordagem híbrida entre Programação Genética e Redes Neurais na previsão de séries temporais, em sua tese de doutorado. Os resultados foram comparados com métodos de previsão tradicionais, como alisamento exponencial e a metodologia Box & Jenkins, tendo concluído que os modelos inspirados na natureza produzem melhores resultados.

Recentes publicações na área de Aprendizado de Máquina mostram que efetuando uma combinação de preditores, pode-se gerar melhores resultados do que a utilização de um único preditor. Uma das mais conhecidas técnicas é o *Boosting*. Em seu trabalho, Paris; Robilliard; Fonlupt (2001) propõem a utilização do algoritmo de *Boosting* juntamente com a Programação Genética. Esta técnica é denominada de *GPBoost*, que utiliza a idéia de combinar classificadores através de pesos na busca de um classificador melhor, fornecendo resultados melhores do que os obtidos através da PG tradicional. O algoritmo foi testado em problemas binários e de regressão. Ainda não há trabalhos referentes à utilização do *Boosting* na combinação de modelos de Previsão de Séries Temporais. Outros trabalhos relacionados à técnica de *Boosting* são apresentados no capítulo IV.

Um trabalho sobre a Seleção de Modelos, publicado por Santos; Ludemir e Prudêncio (2004) propõe formar um *ranking* de algoritmos candidatos a solucionar um problema com base nas suas informações de desempenho em problemas anteriores. O algoritmo *Zoomed Ranking* é composto de duas fases; na primeira fase é selecionado um sub-conjunto de problemas similares ao problema em questão e na segunda fase é gerado o *ranking* dos modelos candidatos, baseados em seu desempenho. O algoritmo foi aplicado à seleção de modelos para a Previsão de Séries Temporais onde o principal fator era a precisão da previsão. O método *Zoomed Ranking*, originalmente proposto para seleção de algoritmos de

classificação, utiliza as informações de desempenho dos modelos candidatos em séries anteriormente processadas na geração de *ranking* desses modelos. Sua contribuição se dá na área de Meta-Aprendizado e na Previsão de Séries Temporais.

Uma outra abordagem é feita na combinação de modelos para Previsão de Séries Temporais por Zou e Yang (2004), que propõe uma combinação convexa de modelos candidatos à solução do problema ao invés de utilizar um único modelo. A idéia é que, quando existe muita incerteza na escolha do melhor modelo, a combinação de vários deles poderá reduzir a instabilidade da previsão e, portanto, melhorar a precisão da predição. Ele propõe um algoritmo, AFTER, no qual os pesos são sequencialmente atualizados após cada observação adicional. O método foi estudado teoricamente por Zou e Yang (2004) e esse estudo mostrou que modelos combinados produzem melhores taxas de convergência do que as encontradas por modelos individuais. Através de seus experimentos, afirma que não é uma boa idéia combinar cegamente, todos os possíveis modelos disponíveis, mas analisá-los preliminarmente e criar uma lista com os melhores modelos disponíveis.

Em 2003, Hui publicou um trabalho no qual analisava a performance da Programação Genética na análise de séries temporais, construindo um programa capaz de produzir indivíduos que fornecessem uma previsão para aplicações financeiras futuras baseadas em  $N$  períodos passados, onde  $N$  seria uma variável a ser definida pelo usuário. Tendo realizado várias experiências variando  $N$  e o tamanho da população, chegou a uma configuração capaz de produzir indivíduos que forneciam previsões razoáveis, porém considerando a previsão um passo à frente.

O primeiro experimento realizado nesta tese, utilizou a PG para analisar o seu desempenho na Previsão de Séries Temporais, utilizando algumas séries reais encontradas na literatura. Os resultados obtidos mostraram a boa performance da PG na previsão de valores futuros das séries, fornecendo erros de predição menores do que os obtidos por modelos tradicionais (SOUZA; COSTA; POZO, 2005a).

Souza, Costa e Pozo (2005b) realizaram um trabalho de comparação entre o desempenho da PG quando comparada aos métodos de Alisamento Exponencial, Tendência Polinomial e com os modelos ARMA, tendo obtido resultados com erros de previsão menores, utilizando a metodologia proposta.

## 2.4 DEFINIÇÕES E OBJETIVOS DA ANÁLISE DE SÉRIES TEMPORAIS

Uma Série Temporal é qualquer conjunto  $Z$  de observações ordenadas no tempo,

$$Z = \{Z_t, t = 1, 2, \dots, N\}$$

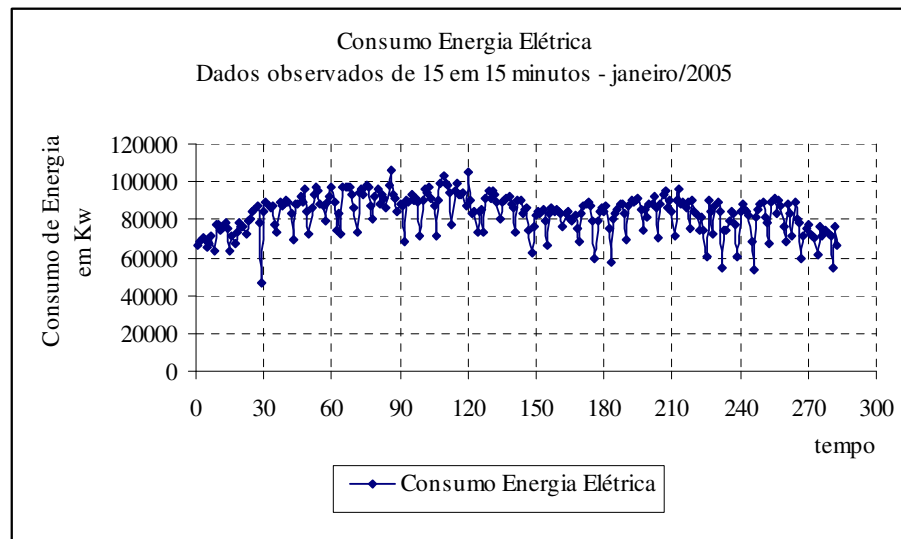
onde  $t$  é o índice de tempo e  $N$  é o número de observações.

Pode-se citar como exemplos de Séries Temporais:

- consumo de Energia Elétrica de uma residência;
- valores mensais de vendas de veículos no Brasil;
- índices diários de bolsa de valores, registro de marés, e outros.

As Séries Temporais podem ser contínuas ou discretas. No caso de um conjunto de dados com valores contínuos, este deve ser discretizado no tempo e com intervalos iguais (MORETTIN; TOLOI, 1985).

Na figura (2.1) é apresentado um exemplo de uma Série Temporal que representa o consumo diário de energia elétrica na cidade de Curitiba, as observações são diárias e tomadas de 15 em 15 minutos, no mês de janeiro de 2005.



**FIGURA 2.1 – EXEMPLO DE UMA SÉRIE TEMPORAL**

Os principais objetivos da Análise de Séries Temporais são (MORETTIN; TOLOI; 1985):

- identificar o mecanismo gerador da série;
- descrever o comportamento da série, verificando a existência de tendência e ciclos de variações sazonais;
- procurar periodicidades relevantes nos dados;
- encontrar um modelo matemático que seja capaz de efetuar previsão de valores futuros das séries.

As definições dadas a seguir são necessárias ao estudo e Análise de Séries Temporais.

#### 2.4.1 Processos Estocásticos

Uma série temporal é uma observação de um processo estocástico. Suponha então, uma amostra de valores observados de tamanho  $T$  da variável aleatória  $Z_t$ , tal que:  $\{Z_1, Z_2, \dots, Z_T\}$  e considere uma coleção de variáveis independentes e identicamente distribuídas  $\varepsilon_t$  tal que:  $\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_T\}$   $\varepsilon_t$  tem distribuição normal com média 0 (zero) e variância  $\sigma^2$  ( $\varepsilon_t \sim N(0, \sigma^2)$ ), então  $\varepsilon_t$  é considerado um processo *Gaussiano* com ruído branco. A sequência de valores observados,  $Z_t$ , representa uma série temporal. Desta forma, uma série temporal é uma realização de um processo estocástico (MORETTIN & TOLOI, 2004).

#### 2.4.2 Processo Estocástico Ergódico

A caracterização de um Processo Estocástico (PE) exige o conhecimento de todas as suas funções amostras (realizações, trajetórias). Isto permite determinar a função da média,  $\mu(t)$  e a função de autocorrelação,  $\rho(t)$  do processo. Quando estes parâmetros são determinados a partir de apenas uma realização (função amostra), este processo é um processo Ergódico.

### 2.4.3 Processo Estocástico Estacionário

Considere um PE,  $Z(w, t)$  e considere os instantes  $t_1, t_2, \dots, t_{n1}$  pertencentes ao conjunto de instantes  $T$ , se para qualquer número  $n$  de variáveis aleatórias  $Z_1, Z_2, \dots, Z_n$  a função da média,  $\mu(t) = E[Z_t] = \mu$  e  $E[Z_t^2] < \infty$ , ou seja, a esperança é finita e a função de autocovariância,  $\gamma(t_1, t_k) = E[(Z_{t+k} - \mu)(Z_t - \mu)] = \gamma_k$  é uma função da defasagem  $k$ , então o PE é denominado estacionário no sentido amplo.

### 2.4.4 Série Temporal Estacionária

Uma série temporal é dita estacionária quando possui média e variância constantes e a função de autocovariância entre dois períodos distintos depende apenas da defasagem de tempo (*lag*) entre os períodos, como pode ser visto na equação (2.1):

$$E(Z_t) = \mu; E(Z_{t-\mu})^2 = \sigma^2 < \infty \text{ e } E(Z_t - \mu)(Z_{t+k} - \mu) = \gamma_k, \forall t \in T \quad (2.1)$$

### 2.4.5 Média de um Processo Estocástico

A média de um PE é representada por  $\mu(t) = E[Z(t)], \forall t \in T$  e se o processo é estacionário, tem-se que  $\mu(t) = E[Z(t)] = \mu$ . A média pode ser estimada conforme equação (2.2).

$$\bar{Z}_t = \frac{\sum_{t=1}^n Z_t}{n} \quad (2.2)$$

#### 2.4.6 Autocovariância do Processo Estocástico

A função de autocovariância do processo do PE  $Z(w, t)$ , representada por  $\gamma_k$  (ordem  $k$ ) é o momento central conjunto das variáveis aleatórias  $Z_t$  e  $Z_{t+k}$ , e é calculada conforme equação (2.3), tal que:

$$\hat{\gamma}_k = E[(Z_t - \mu(t))(Z_{t+k} - \mu(t+k))] \quad (2.3)$$

onde  $z_t$  é a observação no instante  $t$  e  $\bar{z}$  é a média da amostra. O estimador para a autocovariância é dado pela equação (2.4).

$$\hat{\gamma}_k = \frac{1}{n} \sum_{t=1}^n (z_t - \bar{z})(z_{t+k} - \bar{z}) \quad (2.4)$$

#### 2.4.7 Variância do Processo Estocástico

A variância do PE,  $Z(w, t)$ , representada por  $E(Z_t - \mu)^2 = \sigma^2 = \gamma_0$  é obtida da função de autocovariância quando  $k = 0$  e pode ser calculada através do seu estimador, conforme equação (2.5).

$$\hat{\gamma}_0 = \frac{1}{n} \sum_{t=1}^n (z_t - \bar{z})^2 = \hat{\sigma}_z^2 \quad (2.5)$$

#### 2.4.8 Função de Autocorrelação do Processo Estocástico

A função de autocorrelação do PE,  $Z(w, t)$ , representada por  $\rho_k$ , para a ordem  $k$ , é a razão,  $\rho_k = \frac{\gamma_k}{\gamma_0}$  entre a autocovariância de ordem  $k$  do PE,  $\gamma_k$ , e a variância do processo,  $\gamma_0$ .

O estimador para esta função é dado pela equação (2.6).



$$\hat{\rho}_k = \frac{\sum_{t=1}^n (z_t - \bar{z})(z_{t+k} - \bar{z})}{\sum_{t=1}^n (z_t - \bar{z})^2} \quad (2.6)$$

## 2.5 MODELOS E MÉTODOS DE PREVISÃO

Um modelo é uma descrição de uma série temporal e a forma de utilização deste modelo depende do objetivo da análise. Nem sempre um modelo conduz a uma boa fórmula de previsão. Muitas vezes os valores estimados apresentam valores muito distantes da realidade, o erro de previsão é a quantificação desta distância.

Para ter um controle destes erros é importante que se defina uma função de perda, sendo que a mais utilizada é a do erro quadrático médio (*MSE - Mean Square Error*) (MORETTIN; TOLOI, 1985). Os erros de previsões são causados, essencialmente, pela falta de confiabilidade dos dados históricos e pela falta de um modelo que possua um melhor ajuste, considerando a instabilidade do mercado (CORRÊA; GIANESI; CAON, 2000).

A precisão da previsão é fortemente influenciada pelas características das observações realizadas e também pelo horizonte de previsão. A combinação de previsões de mais de um modelo aumenta a confiabilidade da predição (MAKRIDAKIS; WHEELWRIGHT, 1989).

Uma das considerações mais frequentes a respeito de uma série temporal é a de que ela seja estacionária, ou seja, ela se desenvolve no tempo aleatoriamente ao redor de uma média e variância constantes, refletindo alguma forma de equilíbrio estável (MORETTIN; TOLOI, 1985).

Na prática, no entanto, as séries temporais apresentam alguma forma de não estacionariedade. 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 ou negativa (*tendência linear*). Pode-se ter, também, uma forma de não estacionariedade explosiva, como o crescimento de uma colônia de bactérias. Alguns dos métodos estatísticos para Análise ou Previsão de Séries Temporais exige que a série seja estacionária, caso contrário há

necessidade de se transformar os dados originais através de: diferenças sucessivas<sup>5</sup> ou transformações de variáveis, até que se obtenha uma série estacionária. Em geral é suficiente se tomar uma ou duas diferenças para que a série se torne estacionária (MORETTIN; TOLOI, 2004).

A seguir são brevemente descritos os modelos Auto-Regressivos (AR), os modelos de Médias Móveis (MA), os modelos mistos Auto-Regressivos e de Médias Móveis (ARMA).

### 2.5.1 Modelos Auto-Regressivos

Um modelo auto-regressivo de ordem  $p$ ,  $AR(p)$ , pode ser representado matematicamente pela expressão (2.7):

$$Z_t = \delta + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t \quad (2.7)$$

$Z_t$  - valor observado da série temporal no instante  $t$ ;

$\phi_i$  - representa o  $i$ -ésimo parâmetro auto-regressivo a ser estimado,  $i = 1, 2, \dots, p$ ;

$a_t$  - ruído branco, ou seja,  $a_t \sim N(0, \sigma_a^2)$  no instante  $t$ ;

$a_{t-1}$  - ruído branco, ou seja,  $a_{t-1} \sim N(0, \sigma_a^2)$  no instante  $t-1$ ;

$\delta$  - valor constante

A expressão (2.7) é uma soma ponderada de valores passados, adicionados a um ruído aleatório e a um valor constante  $\delta$ .

A função de autocovariância do processo  $AR(p)$ , considera o modelo em função dos desvios e da média e é dada pela expressão (2.8):

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p} \quad (2.8)$$

A média para a estrutura  $AR(p)$  é obtida, tomando-se a esperança condicional do processo  $Z_t$ , e está representada pela expressão (2.9).

$$\mu = \frac{\delta}{1 - \phi_1 - \phi_2 - \dots - \phi_p} \quad (2.9)$$

A expressão da variância para a estrutura  $AR(p)$  é dada pela equação (2.10):

<sup>5</sup> Diferença sucessiva - 1ª diferença:  $\Delta Z(t) = Z(t) - Z(t-1)$ ; 2ª diferença:  $\Delta^2 Z(t) = \Delta[\Delta Z(t)]$

$$\gamma_0 = \sigma_w^2 = \frac{\sigma_a^2}{1 - \phi_1 \rho_1 - \phi_2 \rho_2 - \dots - \phi_p \rho_p} \quad (2.10)$$

onde  $\rho_i$  é a função da autocovariância do processo AR( $p$ ).

O caso mais simples de um modelo auto-regressivo de ordem  $p = 1$  é o AR(1), representado pela equação (2.11):

$$Z_t = \delta + \phi_1 Z_{t-1} + a_t \quad (2.11)$$

onde o valor de  $Z_t$ , depende apenas de  $Z_{t-1}$  e do ruído no instante  $t$  e considerando o termo constante  $\delta$ .

Os modelos AR mais utilizados são os de ordem 1 e 2 (modelos parcimoniosos) (MAKRIDAKIS; WHEELWRIGHT; MCGEE, 1983). Se os valores da série puderem ser representados na forma dos modelos AR( $p$ ), após ser identificada a estrutura do processo gerador da série e estimados os parâmetros, o modelo será utilizado para prever valores futuros da série analisada.

### 2.5.2 Modelos de Médias Móveis

Um modelo de médias móveis é um modelo que efetua uma média ponderada de valores dos ruídos, observados nos períodos precedentes. O modelo de ordem  $q$ , MA( $q$ ), é representado pela expressão (2.12):

$$Z_t = \delta + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (2.12)$$

onde:  $Z_t$  – valor da série temporal no instante  $t$ ;

$\delta$  - termo constante da série (nível);

$\theta_i$  –  $i$ -ésimo parâmetro de média móvel;

$a_t$  – ruído branco no instante  $t$ .

A média do processo MA( $q$ ) é dada pela equação (2.13):

$$\mu = E(Z_t) = \delta \quad (2.13)$$

A função de autocovariância do processo MA( $q$ ) é dada pela expressão (2.14):

$$\gamma_k = \left( -\theta_k + \sum_{l=1}^{q-k} \theta_l \theta_{l+k} \right) \sigma_a^2, \quad k = 0, 1, 2, \dots, q \quad (2.14)$$

a variância do processo é dada pela equação (2.15), onde  $\sigma_a^2$  é a variância de  $a_t$ :

$$\gamma_0 = \left( -\theta_0 + \sum_{l=1}^q \theta_l \theta_l \right) \sigma_a^2 = \left( 1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2 \right) \sigma_a^2 \quad (2.15)$$

e a função de autocorrelação dos modelos MA( $q$ ) é dada pela equação (2.16):

$$\rho_k = \begin{cases} \frac{\gamma_k}{\gamma_0} = \frac{-\theta_k + \sum_{l=1}^{q-k} \theta_l \theta_{l+k}}{1 + \sum_{l=1}^q \theta_l^2}, & k = 1, 2, \dots, q \\ 0, & k > q \end{cases} \quad (2.16)$$

As previsões podem então, ser geradas por este modelo, desde que os valores da série estejam bem ajustados por ele.

### 2.5.3 Modelos Auto-Regressivos e de Médias Móveis (ARMA)

A combinação de modelos auto-regressivos e de médias móveis pode conduzir a um ajuste mais parcimonioso, ou seja, que contenha um número inferior de parâmetros a serem estimados, do que se fosse utilizado um modelo AR ou MA puros. Um modelo da estrutura ARMA( $p, q$ ), é representado matematicamente pela equação (2.17).

$$Z_t = \delta + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} + a_t \quad (2.17)$$

onde  $p$  e  $q$  representam, respectivamente, a ordem dos modelos AR e MA. O modelo relaciona os valores passados observados e os erros obtidos em cada período (Trentin, 2002).

A expressão da média do processo ARMA( $p, q$ ) é dada pela equação (2.18) (HAMILTON, 1994):

$$E(Z_t) = \mu = \frac{\delta}{1 - \phi_1 - \phi_2 - \dots - \phi_p} \quad (2.18)$$

Assim, a estacionariedade do processo ARMA depende inteiramente dos parâmetros auto-regressivos ( $\phi_1, \phi_2, \dots, \phi_p$ ) portanto é conveniente escrevê-lo em termos do desvio da média (Eq. 2.19):

$$Z_t - \mu = \phi_1 (Z_{t-1} - \mu) + \phi_2 (Z_{t-2} - \mu) + \dots + \phi_p (Z_{t-p} - \mu) + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q} \quad (2.19)$$

$$k = q+1, q+2, \dots$$

a função de autocovariância é obtida multiplicando-se ambos os lados da equação (2.19) por  $(Z_{t-k} - \mu)$  e tomando as esperanças. Para  $k > q$ , a equação resultante tem a forma representada na equação (2.20).

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p} - \theta_1 \gamma_{Z_a \gamma_{k-1}} - \dots - \theta_q \gamma_{Z_a \gamma_{k-q}} \quad (2.20)$$

a expressão da variância do processo é dada pela equação (2.21) quando  $k = 0$ :

$$\gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \dots + \phi_p \gamma_p - \theta_1 \gamma_{Z_a} - \dots - \theta_q \gamma_{Z_a} + \sigma_a^2 \quad (2.21)$$

a expressão da função de autocorrelação,  $\rho_k$  do processo é dada pela equação (2.22):

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad (2.22)$$

## 2.6 METODOLOGIA BOX & JENKINS

A metodologia Box & Jenkins é um dos trabalho de maior importância e reconhecimento na área de Previsão de Séries Temporais. O estudo feito pelos pesquisadores George Box e Gwilyn Jenkins (1970) é baseado no importante resultado de Wold (1954) que provou que qualquer série temporal pode ser representada por uma estrutura de médias móveis infinita (CHAVES, 1991). A metodologia proposta consiste em ajustar modelos Auto-Regressivos e de Médias Móveis, ARMA( $p, q$ ) à série temporal.

Quando comparada aos demais métodos de previsão, a metodologia Box & Jenkins apresenta resultados mais precisos e os modelos contém um número pequeno de parâmetros. Uma desvantagem da utilização deste método é que sua aplicação requer um conhecimento que vai além do uso de um pacote computacional (MORETTIN; TOLOI, 2004), ou seja, faz-se necessário a intervenção de um analista especializado.

Na construção de um modelo ARMA, deve-se seguir os passos da metodologia Box & Jenkins, que são (MORETTIN; TOLOI, 1985):

- identificação do modelo ARMA a ser ajustado aos dados - esta escolha é feita com base nas autocorrelações e autocorrelações parciais estimadas. É considerada a fase mais crítica do método. A partir de 1970 foram propostos procedimentos para identificação do modelo dentre os quais se pode citar: o Critério de Informação de Akaike (AIC) (AKAIKE, 1973, 1974) e o Critério Bayesiano (BIC) (AKAIKE, 1977), (RISSANEN & SCHWARZ, 1978). Estes critérios visam identificar a ordem do modelo ARMA a ser utilizado, minimizando uma função que penaliza o ajuste de modelos não parcimoniosos, ou seja, que possuem uma grande quantidade de parâmetros. Neste trabalho é utilizado o critério AIC, cujo detalhamento encontra-se no Apêndice A. A escolha de utilização deste critério foi feita pelo fato de que o mesmo identifica modelos com o menor número de parâmetros possível;
- estimação dos parâmetros – tendo sido identificado o modelo provisório para a série temporal, o passo seguinte é estimação de seus parâmetros. Para fazer a estimativa por Mínimos Quadrados Não Lineares ou Máxima Verossimilhança, utiliza-se o algoritmo de Marquardt (CHAVES, 1991);
- verificação da adequabilidade do modelo – após a estimação dos parâmetros, deve-se verificar se o modelo se ajusta bem aos dados, para isto é feita uma análise nos resíduos. Uma técnica que pode ser utilizada, caso o modelo se mostre inadequado é o super-ajustamento, no qual são adicionados parâmetros e o modelo é estimado. Examina-se então se estes parâmetros são significativos no ajuste (teste “ $t$ ”);
- previsão – Na fase final, após se ter selecionado o melhor modelo para a série temporal e se o modelo escolhido for considerado adequado, ele será usado para fazer previsões de valores futuros da série.

Na figura (2.2) é apresentada a estrutura de aplicação da metodologia Box & Jenkins (BOX & JENKINS, 1970).

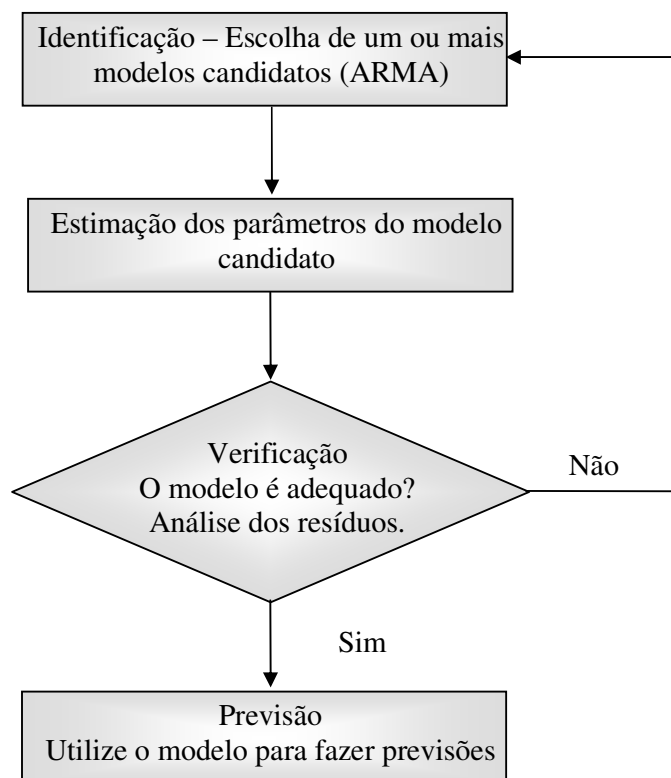


FIGURA 2.2 – FLUXOGRAMA DA METODOLOGIA BOX &amp; JENKINS

## 2.7 MÉTODOS DE PREVISÃO

Os pesquisadores Box & Jenkins (1970) propuseram a classe dos modelos ARMA (*Autoregressive Moving Average*) para uma série temporal  $\{Z_t, t = 1, 2, \dots, n\}$ . A utilização da metodologia ARMA é aplicada para a previsão de valores futuros de uma série temporal. As formas de previsão utilizadas são: forma de equações de diferenças, formas de choques aleatórios e forma invertida (CHAVES, 1991).

Ao se realizar a previsão, o interesse se concentra nos valores futuros da variável observada. Sendo  $t$  o período de tempo atual, deseja-se obter o valor futuro da série observada,  $Z_{t+h}$ , onde  $h \geq 1$ . O período  $t$  é dito origem da previsão e  $h$  é o horizonte de previsão. A previsão de  $Z_{t+h}$  é denotada por  $\hat{Z}_{t+h}$ , e a esperança condicional de  $Z_{t+h}$  é representada pela equação (2.23)

$$\hat{Z}_t(h) = E(Z_{t+h} | I_t) \quad (2.23)$$

onde  $I_t$  é o conjunto dos valores passados da série observada  $\{..., Z_{t-3}, Z_{t-2}, Z_{t-1}, Z_t\}$ . Tomando-se a esperança condicional de  $Z_{t+h}$ , na equação (2.24):

$$Z_{t+h}(l) = \xi_1 Z_{t+h-1} + \xi_2 Z_{t+h-2} + \dots + \xi_{p+d} Z_{t+h-p-d} - \theta_1 a_{t+h-1} - \theta_2 a_{t+h-2} - \dots - \theta_q a_{t+h-q} + a_{t+h} \quad (2.24)$$

onde:  $\xi(B) = \phi(B)\nabla^d = \phi(B)(1-B)^d$ , obtém-se a equação de previsão (2.25), onde  $B$  é o operador de translação para o passado<sup>6</sup>.

$$\hat{Z}_t(l) = E(Z_{t+h} | I_t) = \xi_1 [Z_{t+h-1}] + \dots + \xi_{p+d} [Z_{t+h-p-d}] - \theta_1 [a_{t+h-1}] - \dots - \theta_q [a_{t+h-q}] + [a_{t+h}], \quad h \geq 1 \quad (2.25)$$

e considerando-se as equações (2.26):

$$\begin{aligned} [Z_{t+h}] &= \hat{Z}_t(h), \quad h > 0 \\ [Z_{t+h}] &= \hat{Z}_{t+h}, \quad h \leq 0 \\ [a_{t+h}] &= 0, \quad h > 0 \\ [a_{t+h}] &= a_{t+h}, \quad h \leq 0 \end{aligned} \quad (2.26)$$

A variância do erro de previsão é estabelecida partindo-se da forma de choques aleatórios, isto é, substituindo-se os termos AR (auto-regressivos) por uma série infinita de termos MA (médias móveis), representado nas equações (2.27).

$$Z_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots = \mu + \sum_{i=0}^{\infty} \psi_i a_{t-i}, \quad \psi_0 = 1 \quad (2.27)$$

$$Z_t = \mu + (1 + \psi_1 B + \psi_2 B^2 + \dots) a_t$$

$$Z_t = \mu + \psi(B) a_t$$

Supondo-se que a origem de previsão seja  $t$  e que pretende-se fazer a previsão  $h$  períodos à frente, ou seja  $\hat{Z}_t(h)$  irá prever o valor de  $Z_{t+h}$ , onde  $Z_t$  é uma função linear de origem  $Z_t$  e das observações anteriores  $Z_{t-1}, Z_{t-2}, Z_{t-3}, \dots$ , sendo assim, uma função linear dos choques aleatórios  $a_t, a_{t-1}, a_{t-2}, \dots$  e a melhor previsão será então dada pela equação (2.28).

$$\hat{Z}_t(h) = \psi_l^* a_t + \psi_{l+1}^* a_{t-1} + \psi_{l+2}^* a_{t-2} + \dots \quad (2.28)$$

---

<sup>6</sup>  $BZ_t = Z_{t-1}$  e  $B^m Z_t = Z_{t-m}$



onde  $\psi_h^*, \psi_{h+1}^*, \psi_{h+2}^*$  são os pesos para os choques aleatórios, com a soma infinita dos valores atuais e dos choques aleatórios, representados na equação (2.29).

$$\hat{Z}_{t+l} = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t+h-i} \quad (2.29)$$

Tem-se que o erro quadrático será dado pela equação (2.30)

$$E[Z_{t+h} - \hat{Z}_t(h)]^2 = (1 + \psi_1^2 + \psi_2^2 + \dots + \psi_{h-1}^2) \sigma_\varepsilon^2 + \sum_{i=0}^{\infty} \{\psi_{h+j} - \psi_{h+j}^*\}^2 \sigma_\varepsilon^2 \quad (2.30)$$

e quando  $\psi_{h+j} = \psi_{h+j}^*$ , obtém-se o valor mínimo de  $E[Z_{t+h} - \hat{Z}_t(h)]^2$ . Conseqüentemente,  $Z_{t+h} = \hat{Z}_t(h) + e_t(h)$ , onde  $e_t(h)$  é o erro de previsão, a esperança e a variância deste erro estão representados pelas equações (2.31).

$$E[e_t(h)] = 0 \text{ e } V[e_t(h)] = (1 + \psi_1^2 + \psi_2^2 + \dots + \psi_{h-1}^2) \sigma_\varepsilon^2 \quad (2.31)$$

Finalmente, os resíduos um passo a frente serão dados como na equação (2.32).

$$e_t(1) = Z_{t+1} - \hat{Z}_t(1) = \varepsilon_{t+1} \quad (2.32)$$

Sendo o intervalo de confiança de  $(1 - \alpha)\%$  para  $Z_{t+l}$  dado pela equação (2.33)

$$P\left(\hat{Z}_t(h) - Z_{1-\alpha/2} \sqrt{V[e_t(h)] \hat{\sigma}_\varepsilon^2} \leq Z_{t+h} \leq \hat{Z}_t(h) + Z_{1-\alpha/2} \sqrt{V[e_t(h)] \hat{\sigma}_\varepsilon^2}\right) = 1 - \alpha \quad (2.33)$$

Como se pode observar, a variância aumenta, conforme aumenta o horizonte de previsão  $h$ , quando nos afastamos da origem  $t$ , a amplitude do intervalo aumenta, caracterizando o aumento na incerteza das previsões para  $h$  passos à frente.

## Capítulo 3

### MÉTODOS UTILIZADOS NA PREVISÃO DE SÉRIES TEMPORAIS

A Computação Evolucionária (CE) é uma das áreas da Inteligência Artificial, que engloba um conjunto de métodos computacionais, inspirados na Teoria da Evolução das Espécies de Charles Darwin (DARWIN, 2000) para a solução de problemas. Segundo sua teoria, na natureza sobrevivem os indivíduos que possuem maior capacidade de se adaptarem ao meio ambiente, suas características genéticas são repassadas para as gerações seguintes e melhoradas. Assim a nova geração será composta por indivíduos com material genético melhor do que os da população anterior.

John Holland (1975) propôs um algoritmo baseado neste princípio com o objetivo de estudar os fenômenos relacionados à evolução das espécies e da seleção natural que ocorre na natureza. Holland foi gradualmente refinando suas idéias e em 1975 publicou seu livro (HOLLAND, 1975) no qual apresentou uma maneira de implementar computacionalmente este algoritmo. Sua técnica ficou conhecida como Algoritmos Genéticos (AG) e desde então vem sendo objeto de estudo e aprimoramento de muitos pesquisadores

As principais áreas dentro da CE são: Programação Evolutiva, Estratégias Evolutivas, Algoritmos Genéticos e Programação Genética. Na seção (3.1) é apresentada uma descrição detalhada do algoritmo da PG que é utilizada neste trabalho.

### 3.1 PROGRAMAÇÃO GENÉTICA

A Programação Genética é uma das técnicas da Computação Evolucionária na qual os indivíduos são programas computacionais. Sua teoria foi desenvolvida por John Koza (KOZA, 1989) e assim como os AG's, está baseada no princípio da Evolução de Darwin, na qual sobrevivem os indivíduos que possuem maior capacidade de se adaptarem ao meio ambiente. A carga genética destes indivíduos será passada aos seus descendentes e sofrerá modificações com o objetivo de melhor se ajustar ao meio, o resultado é que ao final de várias gerações, obter-se-á uma população de indivíduos com características naturalmente selecionadas, ocorrendo desta forma uma evolução natural da população.

Na Programação Genética, o Algoritmo Evolutivo opera numa população de programas computacionais que variam de forma e tamanho (KOZA, 1992). Esta população de indivíduos será evoluída de modo a gerar uma nova população constituída por indivíduos melhores, utilizando operadores de reprodução, cruzamento e mutação. O processo é guiado por uma função de aptidão (*fitness*) que mede o quanto o indivíduo está próximo da solução do problema. Indivíduos que possuem maior capacidade de adaptação têm melhores chances de sobreviver. A Programação Genética vem sendo aplicada em diversas áreas do conhecimento, tais como Engenharia de Software, Circuitos Digitais, Mineração de Dados, Previsão de Séries Temporais e outras (GECCO, 2002-2006).

#### 3.1.1 Descrição Geral do Algoritmo de Programação Genética

O algoritmo básico da Programação Genética está representado na figura (3.1). Inicialmente cria-se uma população inicial aleatória, com base nesta população o algoritmo irá criar novas populações até encontrar a solução desejada. Esta população deverá ter diversidade suficientemente grande para que se encontre nos indivíduos todas as características que sejam necessárias para solucionar o problema, pois as características não existentes na população inicial dificilmente aparecem durante o processo evolutivo. Numa segunda etapa a população é avaliada, para isso é associado a cada indivíduo um valor de

aptidão que indicará o quão próximo da solução o indivíduo se encontra. A definição desta função de aptidão varia de acordo com o problema em questão. Após a população ter sido avaliada, os indivíduos com melhores valores de aptidão são selecionados para que a eles sejam aplicados os operadores genéticos (mutação, reprodução e cruzamento). Estes indivíduos irão compor a nova população, o algoritmo continua até que um critério de parada seja atingido (RODRIGUES, 2002).

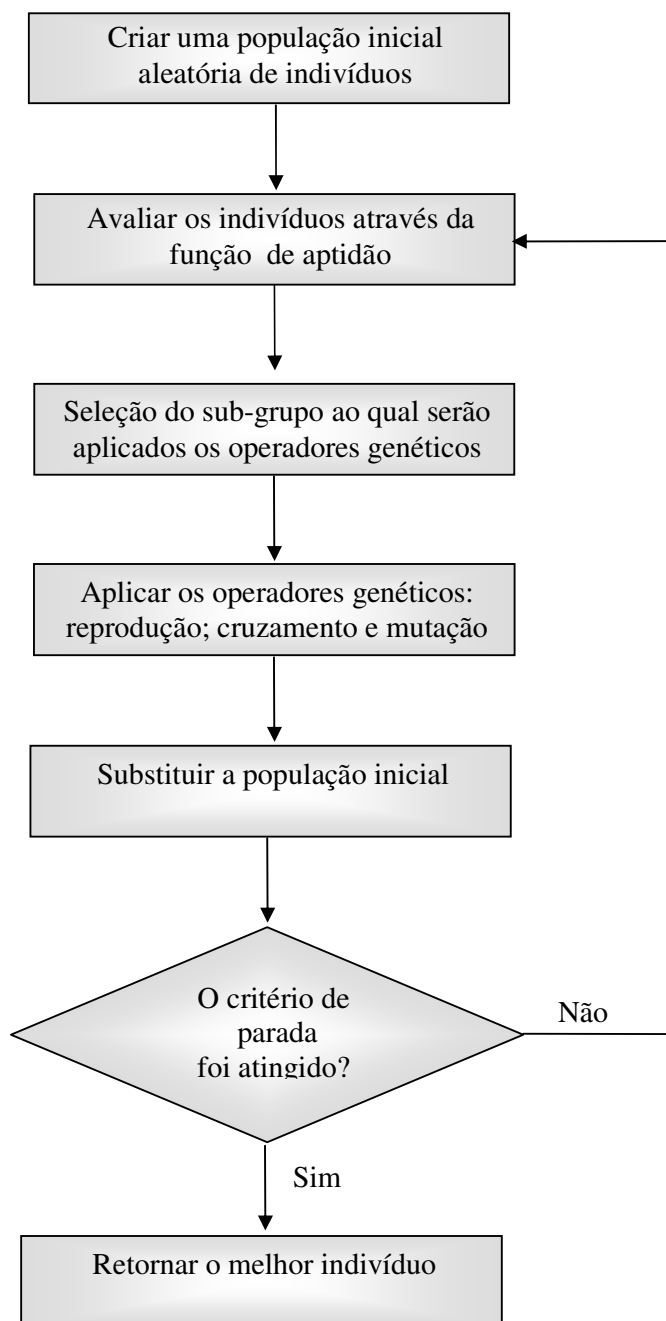


FIGURA 3.1 – ALGORITMO BÁSICO DE PROGRAMAÇÃO GENÉTICA

### 3.1.2 Representação dos Programas

Na Programação Genética, os indivíduos são representados por árvores de sintaxe, ou seja, são formados por uma combinação dos conjuntos de Funções (F) e Terminais (T), de acordo com o domínio do problema. Um indivíduo da população que tem a forma:  $x^2 + y$ , é representado na notação pré-fixa, utilizada pela PG de acordo com a equação (3.1) e sua representação em forma de árvore de sintaxe está mostrada na figura (3.2).

$$(+ (* x x) y) \quad (3.1)$$

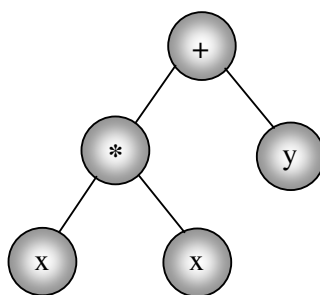


FIGURA 3.2 – ÁRVORE DE SINTAXE DA PROGRAMAÇÃO GENÉTICA

### 3.1.3 Definição dos Conjuntos de Funções e Terminais

Em todo algoritmo de Programação Genética deve-se definir inicialmente os conjuntos F, de funções e T, de terminais. No conjunto F, define-se os operadores aritméticos, funções matemáticas, operadores lógicos, entre outros. O conjunto T é composto pelas variáveis e constantes e fornece um valor para o sistema, enquanto que o conjunto de funções processa os valores no sistema. Juntos, os conjuntos de funções e terminais representam os nós.

Pode-se citar como exemplo, o conjunto F, dos operadores aritméticos, de aridade<sup>7</sup> dois (2), e o conjunto T, de terminais, da seguinte forma:

$$F = \{+, -, *, /\} \text{ e } T = \{x, y\}$$

<sup>7</sup> Número de argumentos de uma função

Um indivíduo resultante da combinação destes dois conjuntos pode ser o indivíduo apresentado na equação (3.1). A escolha dos conjuntos F e T influencia, consideravelmente, na solução apresentada pela Programação Genética. Se no conjunto F houver poucos operadores disponíveis, a Programação Genética provavelmente não será capaz de apresentar uma boa solução para o problema, por outro lado, ao disponibilizar muitas operações, o programa poderá ficar extenso, provocando esforço computacional desnecessário. O mais aconselhável é iniciar com os operadores básicos, tais como: adição, subtração, multiplicação, divisão, conjunção, disjunção e negação e ir adicionando outros operadores caso a solução apresentada não seja suficientemente boa. Da mesma forma deve-se ter cuidado ao formar o conjunto das variáveis e constantes, pois o algoritmo de Programação Genética tem habilidade de combinar as variáveis, transformando-as em novas variáveis (BANZHAF,1998).

O espaço de busca da PG é constituído por todas as árvores que possam ser construídas através da combinação dos conjuntos F e T.

#### 3.1.4 Propriedades: Fechamento e Suficiência

As propriedades do Fechamento (*closure*) e da suficiência (*sufficiency*) (KOZA, 1992), foram definidas por John Koza, com o objetivo de garantir soluções viáveis para o problema.

A propriedade do Fechamento garante que qualquer função do conjunto F deve ser capaz de operar com todos os valores recebidos como entrada. Isso garante que sejam geradas árvores sintaticamente viáveis.

Um exemplo comum de uma função que não atende à propriedade do fechamento é a divisão por zero. O operador divisão, não pode aceitar zero como entrada. A divisão por zero, faz com que o programa pare, apresentando um erro. Assim, ao invés de um operador comum de divisão, pode-se definir um novo operador denominado de divisão protegida, que é a divisão normal, porém com a diferença de que quando surge uma divisão por zero, o operador retorna o valor um. Todas as outras funções (como por exemplo: raiz quadrada, logaritmo), devem aceitar todas as possíveis entradas para que o sistema não acuse mensagem de erro e possa continuar sua busca pela melhor solução (BANZHAF,1998).

A propriedade da suficiência garante a convergência do sistema, fazendo com que os conjuntos F e T sejam capazes de representar uma solução viável para o problema em questão.

### 3.1.5 População Inicial

O primeiro passo na inicialização de uma PG é definir sua população inicial, ou seja, deve-se criar uma população de estruturas, ou estruturas de programas para posterior evolução.

Um dos principais parâmetros da Programação Genética é o tamanho máximo permitido para um programa, na PG este parâmetro é definido como sendo a profundidade máxima da árvore, ou seja, o número máximo de nós da árvore. Esta profundidade é a maior profundidade que será permitida entre a raiz e todos os nós terminais de um mesmo indivíduo (TERADA, 1991). A profundidade de um nó em uma árvore é a distância do nó raiz até o nó  $n$ . A estrutura de árvore é construída a partir dos conjuntos F e T, para maior clareza, considere o exemplo (BANZHAF, 1998) onde os conjuntos F e T são definidos como sendo:  $F = \{+, -, *, \%\}$  e  $T = \{a, b, c, d, e\}$ .

Existem vários métodos para inicializar uma população em estrutura de árvores, os mais comuns são (LUKE; PAINAT, 2001): *Full*, *Grow*, *ramped-half-and-half* (KOZA, 1992), que é uma combinação dos métodos *Full* e *Grow*, *random-branch* (CHELLAPILLA, 1997), *uniform* (BOHM; GEYER-SCHULZ, 1996), que são brevemente descritos a seguir.

- **Método Grow**: os nós são selecionados aleatoriamente dos conjuntos F e T (exceto para o nó raiz que é retirado do conjunto F), por este motivo o método produz árvores de formatos irregulares. Se uma ramificação contém um nó terminal, esta ramificação pára, mesmo que a profundidade máxima não tenha sido atingida. Na figura (3.3), o nó d tem profundidade três (3).

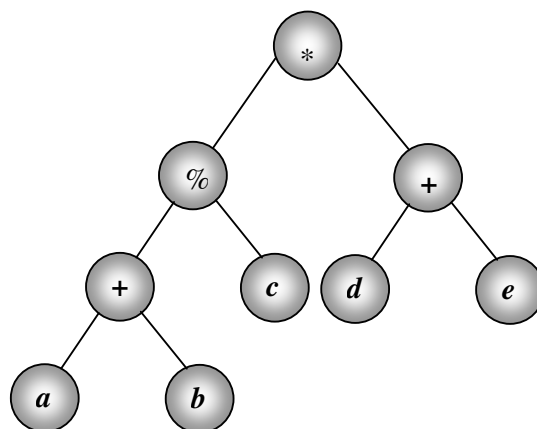


FIGURA 3.3 – ÁRVORE COM PROFUNDIDADE MÁXIMA QUATRO (4) –  
INICIALIZADA PELO MÉTODO GROW

- **Método Full:** Ao invés de escolher aleatoriamente os nós do conjunto de funções e de terminais, o método *Full*, escolhe somente funções até que um nó de profundidade máxima seja selecionado, então ele passa a escolher somente terminais (BANZHAF, 1998). O resultado disso é que cada árvore atinge a profundidade máxima. A figura (3.4) mostra uma árvore inicializada pelo método *Full*.

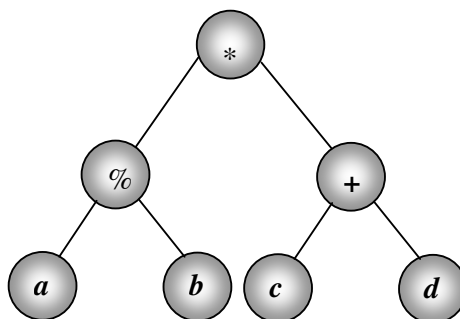


FIGURA 3.4 – ÁRVORE DE PROFUNDIDADE MÁXIMA TRÊS INICIALIZADA PELO MÉTODO FULL

- **Método Half-and-half:** o método *Half-and-half* é uma combinação dos métodos *Grow* e *Full*, ou seja, utiliza o método *Full* em 50% das vezes e o método *Grow* nas outras 50%, tem por objetivo gerar um número igual de árvores para cada profundidade (KOZA, 1992). Supondo, por exemplo, uma árvore de profundidade máxima seis (6), a população é igualmente dividida em árvores com profundidade dois, três, quatro, cinco e seis, ou seja, 20% terão profundidade dois, 20% terão profundidade três e assim sucessivamente. Em cada grupo, metade das árvores são geradas pelo método *Full* e a



outra metade pelo método *Grow*. As desvantagens deste método, segundo Luke, (LUKE; PAINAT, 2000) são citadas a seguir:

- Se o conjunto de funções é maior que o conjunto de terminais, a tendência será de gerar a maior árvore possível;
  - A escolha do parâmetro de profundidade máxima da árvore é realizada de forma proporcional e não aleatória;
  - A faixa de profundidade é fixa (usualmente entre 2 e 6), independente do tamanho da árvore e dependendo do número de argumentos (aridade) de cada função, mesmo tendo a mesma profundidade, as árvores geradas pode ter tamanhos<sup>8</sup> muito diferentes.
- **Método Random-Branch**: neste método, ao invés de se informar a profundidade máxima da árvore, é informado seu tamanho máximo, S, este valor é igualmente dividido dentre as árvores de um nó-pai não terminal, o que faz com que muitas árvores não viáveis sejam geradas (CHELLAPILLA,1997), este método é menos restritivo, apesar de ter complexidade linear (LUKE; PAINAT, 2001);
  - **Método Uniform**: Criado por Bohm, o método *uniform* foi desenvolvido com o objetivo de criar árvores uniformes, geradas a partir do conjunto de todas as árvores possíveis (BOHM. 1996). O algoritmo calcula várias vezes quantas árvores poderão ser geradas para cada tamanho desejado, por este motivo o método possui um alto custo computacional. Uma aplicação do método foi feita por Helmut Horner (1996) numa ferramenta denominada GPK.

---

<sup>8</sup> O tamanho de uma árvore é definido pelo número de nós que a compõe (KOZA, 1992).

### 3.1.6 Função de Aptidão

A função de aptidão é a medida utilizada pela PG durante o processo de evolução, que irá dizer quão bem o programa aprendeu a prever as saídas dentro de um processo de aprendizagem (BANZHAF, 1998).

A definição de uma função de aptidão é feita de acordo com o domínio do problema. Em geral, nos problemas de otimização esta função é definida como sendo a função objetivo, porém nada impede que se defina uma outra função. Uma boa escolha da função de aptidão pode ser responsável pelo bom funcionamento do algoritmo da PG. Especificamente, no caso de Séries Temporais, pode-se utilizar como função de aptidão, a função que mede o erro calculado entre o valor previsto e o valor real, como por exemplo, o erro médio quadrático. Quanto menor for o erro obtido, melhor será o ajuste do modelo de previsão. O que se deseja, portanto, é minimizar a função de aptidão ou função objetivo.

A função de aptidão é uma forma de se diferenciar os melhores dos piores indivíduos. Se esta função for bem definida há uma grande probabilidade de que o algoritmo gere uma solução muito próxima da solução ótima. Algumas das funções de aptidão mais utilizadas são (KOZA, 1992):

1. **Aptidão Nata (*raw fitness*):** é uma maneira de calcular qualquer melhoria que o programa tenha obtido no conjunto de treinamento. A continuidade é uma propriedade importante da função de aptidão, pois isto permite que a PG melhore os programas iterativamente. O método mais comum de aptidão nata é através da avaliação do erro cometido.
2. **Aptidão Padronizada (*standardized fitness*):** é uma função transformada da função de aptidão nata, na qual o valor zero é o valor designado ao melhor indivíduo.
3. **Aptidão Ajustada (*adjusted fitness*):** é obtida a partir da aptidão padronizada, seu valor varia entre zero e um, onde os maiores valores são associados aos melhores indivíduos. Se  $f(i, t)$  é a aptidão padronizada do indivíduo  $i$  na geração  $t$ , a aptidão ajustada,  $a(i, t)$ , é calculada pela equação (3.2):

$$a(i, t) = \frac{1}{1 + f(i, t)} \quad (3.2)$$

4. **Aptidão Normalizada (*normalized fitness*):** é uma função de *aptidão* na qual o seu valor está entre zero e um. A soma de todas as funções normalizadas dentro de uma população deve ser igual a um. Se  $a(i,t)$  é a aptidão ajustada do indivíduo  $i$  na geração  $t$ , sua aptidão normalizada,  $n(i, t)$ , será dada de acordo com a equação (3.3).

$$n(i,t) = \frac{a(i,t)}{\sum_{k=1}^m a(k,t)} \quad (3.3)$$

### 3.1.7 MÉTODOS DE SELEÇÃO

Para selecionar quais indivíduos da população farão parte de uma nova geração e quais deles sofrerão alterações, através dos operadores genéticos (reprodução, cruzamento e mutação), é necessário que se tenha um critério de seleção que garanta que uma boa escolha seja realizada. Um dos métodos mais utilizados para se efetuar esta seleção, baseia-se no valor de aptidão de cada indivíduo, os indivíduos selecionados deverão ser aqueles que apresentam melhores valores de aptidão.

Depois de se ter definido a qualidade de cada indivíduo, através do seu valor de aptidão, são escolhidos os indivíduos que sofrerão intervenção dos operadores genéticos e quais deles deverão permanecer na população utilizando para isto o operador de seleção.

Existem diferentes operadores de seleção e a decisão de qual destes operadores será utilizado pela PG é uma tarefa importante durante a utilização do algoritmo. O método de seleção é responsável pela velocidade da evolução e geralmente citado como responsável pelos casos de convergência prematura que poderão determinar o sucesso do algoritmo evolucionário (BANZHAF, 1998). Alguns destes métodos são descritos a seguir.

- **Seleção Proporcional (*Proportional Selection*):** este método de seleção é aplicado aos Algoritmos Evolutivos e especifica a probabilidade de que cada indivíduo seja selecionado para a próxima geração. Para o indivíduo  $i$ , a probabilidade de ser selecionado para a próxima geração é dada pela equação (3.4).

$$p_i = \frac{f_i}{\sum_j f_j} \quad (3.4)$$

onde  $f_j$  representa o valor de aptidão do indivíduo e  $\sum_j f_j$  representa o valor acumulado de aptidão. Os indivíduos que possuem maior aptidão possuem uma probabilidade maior de serem selecionados para a próxima geração, este pode ser considerado, então, como um problema de maximização. Em geral o melhor indivíduo da população é copiado para a população seguinte, a esta escolha dá-se o nome de elitismo, que tem por objetivo privilegiar a melhor solução, de forma que este indivíduo propague suas características para a população seguinte. No entanto, se um indivíduo possui uma alta aptidão em relação aos demais, a probabilidade de que ele seja selecionado tende a ser alta, e como os demais indivíduos da população possuem uma aptidão bem menor, a tendência é que o indivíduo que possui maior valor de probabilidade, seja selecionado muitas vezes, fazendo com que haja uma convergência prematura da solução que poderá não ser a solução ótima, por outro lado, se os indivíduos apresentarem aptidões muito próximas, sua probabilidade de serem selecionados é a mesma, assim a população seguinte será basicamente a mesma, não havendo evolução (BANZHAF, 1998);

- **Truncamento** (*truncation selection*): este é o segundo método mais popular utilizado para seleção e provém dos algoritmos de Estratégias Evolucionárias (SCHWEFEL, 1995), com base em um valor de limiar (*threshold*)  $T$  que está no intervalo entre 0 e 1, a seleção é feita aleatoriamente entre os  $T$  melhores indivíduos (MUHLENBEIN; SCHIERKAMP-VOSSSEN, 1993). Se, por exemplo,  $T = 0,6$ , isto significa que a seleção é feita entre os 60% melhores indivíduos e os demais são descartados;
- **Ranqueamento** (*Ranking Selection*): no método de seleção por ranking (GREFENSTETTE; BAKER, 1989) (WHITLEY, 1989) os indivíduos são ordenados de forma crescente de acordo com seu valor de aptidão. Assim, a cada indivíduo, é atribuído um número inteiro de acordo com sua posição no ranking, quanto melhor o ranking do indivíduo, melhor sua aptidão em relação aos demais indivíduos da população e, portanto, melhores são suas chances de ser sorteado. Pode ser utilizado

o ranking linear ou exponencial. Para o ranking linear, a probabilidade é uma função linear dada pela equação (3.6):

$$p_i = \frac{1}{N} \left[ p^- + (p^+ - p^-) \frac{i-1}{N-1} \right] \quad (3.6)$$

onde  $p^-/N$  é a probabilidade do pior indivíduo ser selecionado,  $p^+/N$  é a probabilidade do melhor indivíduo ser selecionado e a equação (3.7) assegura que a ordem da população se mantenha constante

$$p^- + p^+ = 2 \quad (3.7)$$

Para o *ranking* exponencial, a probabilidade é exponencialmente ponderada é dada pela equação (3.8):

$$p_i = \frac{c-1}{c^{N-1}} c^N - i, \text{ com } 0 < c < 1 \quad (3.8)$$

- **Torneio:** a seleção por torneio não é baseada na competição dentro da geração completa, mas apenas num sub-conjunto da população. Um certo número de indivíduos, que é o tamanho do torneio, é selecionado aleatoriamente, e é realizada uma competição seletiva. As características dos melhores indivíduos no torneio são substituídas pelas características dos piores indivíduos. No menor torneio possível é permitido que dois indivíduos participem da reprodução. O resultado da reprodução retorna à população substituindo o perdedor do torneio (BANZHAF, 1998). Sua ordem de complexidade é linearmente proporcional ao tamanho da população, pois independe de uma ordenação prévia dos elementos e do cálculo das probabilidades de seleção (BICKLE, 1995).

### 3.1.8 Operadores Genéticos

Após os indivíduos terem sido selecionados por um dos métodos de seleção, os operadores genéticos são aplicados a estes indivíduos para então gerar a nova população. Diversos operadores genéticos foram criados, porém os mais importantes e mais utilizados são (KOZA, 1992): cruzamento, mutação e reprodução, que serão descritos a seguir:

- **Reprodução:** seleciona-se um indivíduo da população atual e o mesmo é copiado para a próxima geração sem nenhuma alteração em sua estrutura, em seguida este indivíduo é re-inserido na população, ou seja, haverá duas versões do mesmo indivíduo na população;
- **Cruzamento:** dois indivíduos pais<sup>9</sup> são selecionados e seu material genético é combinado, permutando uma parte de um dos pais por uma parte do outro, gerando assim um novo indivíduo, que espera-se que seja melhor do que os anteriores, pois foram criados a partir da combinação das melhores partes de cada indivíduo. O operador de cruzamento visa guiar a solução de maneira a combinar as melhores soluções na busca da solução ótima. Basicamente, o operador funciona da seguinte maneira:
  - Escolhe dois indivíduos através do valor de sua função de aptidão;
  - Seleciona aleatoriamente, em cada indivíduo, um ponto de cruzamento;
  - Permuta as sub-árvores dos dois indivíduos gerando os filhos, que farão parte da nova população. Um exemplo do operador de cruzamento pode ser visto na figura (3.5.).
- **Mutação:** o operador efetua alterações em um indivíduo somente, ou seja, nenhum novo indivíduo é gerado. Normalmente, após ter sido efetuado um cruzamento, a probabilidade de que o indivíduo gerado seja submetido a uma operação de mutação é baixa, em geral, este índice varia na faixa de 0,1% a 0,5%. A probabilidade de mutação é um parâmetro definido em cada execução. Nesta operação, seleciona-se aleatoriamente, um ponto do indivíduo e substitui-se a sub-árvore, cujo nó foi selecionado por uma nova sub-árvore gerada aleatoriamente. Esta sub-árvore está sujeita às mesmas limitações quanto à profundidade e tamanho, do que as árvores

<sup>9</sup> Pai – os indivíduos de uma população, pai e filhos são programas computacionais gerados pela PG.

geradas na população inicial. Este novo indivíduo é inserido novamente na população. A função deste operador é inserir diversidade na população, fazendo com que os novos indivíduos explorem novas áreas do espaço de busca (MICHALEWICZ, 1997), evitando máximos e mínimos locais. Porém ao se inserir muita diversidade numa população, a mesma poderá não convergir para um ótimo global, ou mesmo poderá não convergir, oscilando indefinidamente, por este motivo a taxa de mutação deve ser baixa. Um exemplo do funcionamento do operador de mutação é apresentado na figura (3.6).

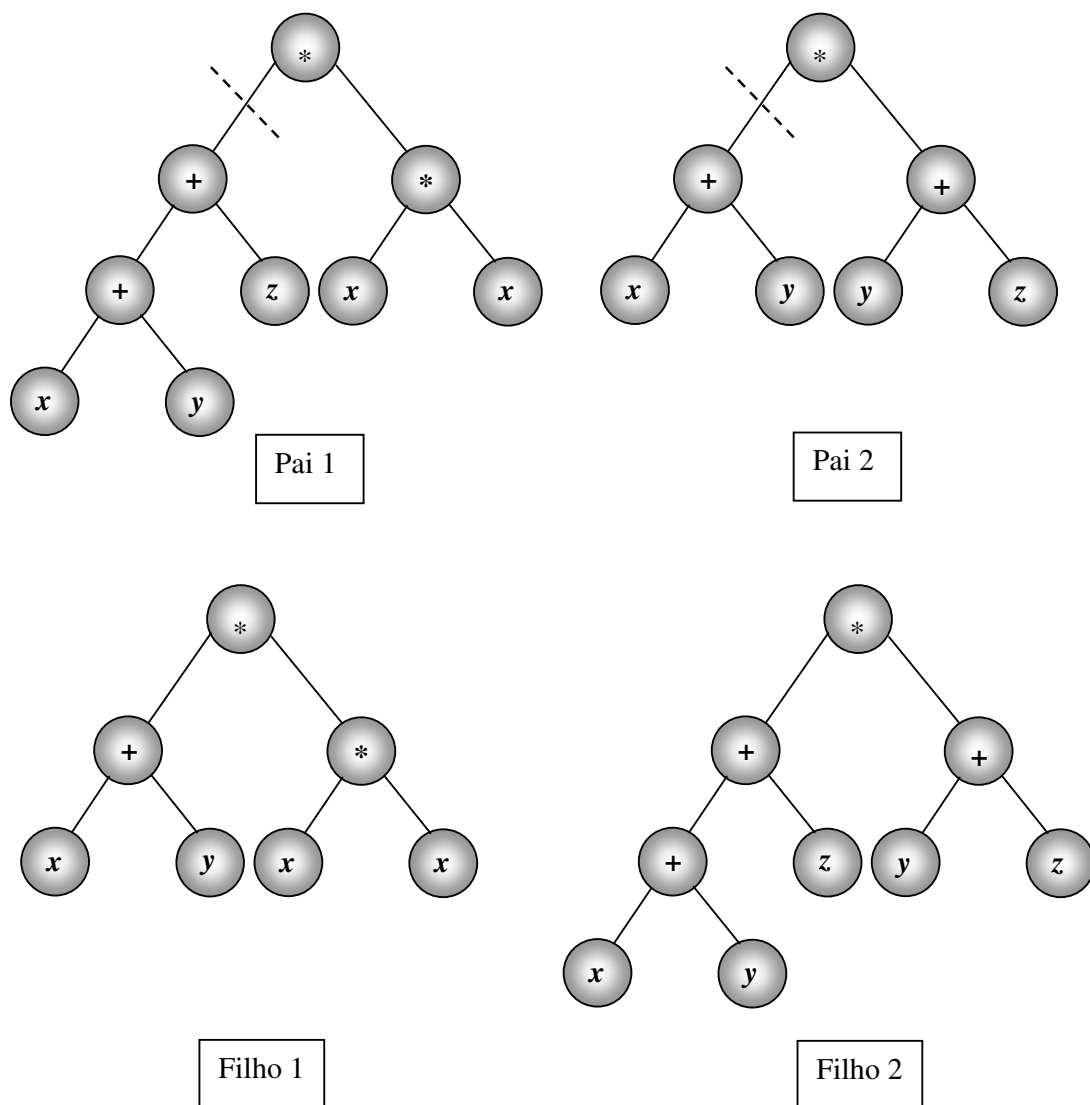


FIGURA 3.5 – REPRESENTAÇÃO DO OPERADOR DE CRUZAMENTO DE UM PONTO

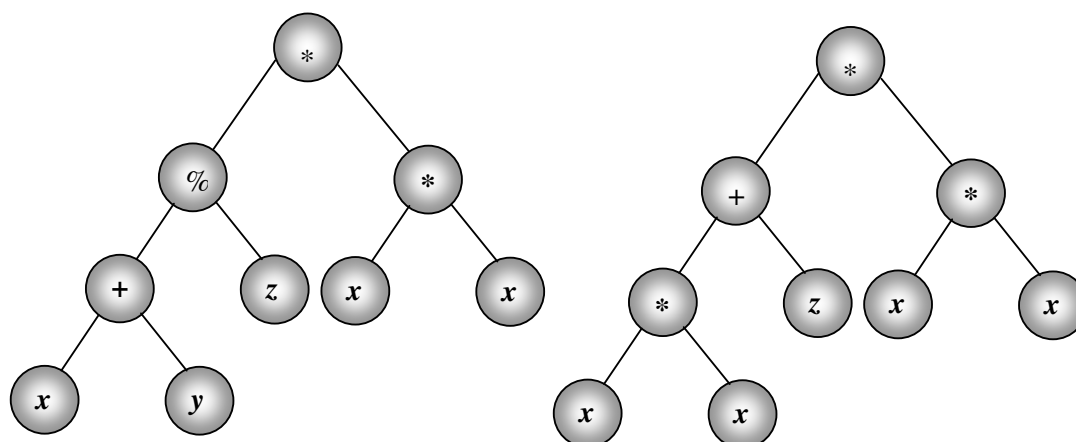


FIGURA 3.6 – REPRESENTAÇÃO DO OPERADOR DE MUTAÇÃO

### 3.1.9 Parâmetros utilizados na Programação Genética

A definição dos parâmetros a serem utilizados podem ter grande influência nos resultados obtidos através da PG, estes parâmetros também influenciam no tempo computacional. Os parâmetros a serem definidos são:

- **Tamanho da População:** neste parâmetro deve ser informado o número de indivíduos que a população deverá conter. A escolha deste parâmetro deve ser criteriosa, pois de sua escolha depende a qualidade dos resultados. Uma população pequena restringe o espaço de busca, enquanto que se a população for grande, poderá estar provocando um esforço computacional excessivo, sem grandes alterações nos resultados (BANZHAF, 1998);
- **Taxa de cruzamento:** este parâmetro define a taxa de utilização do operador de cruzamento que deve ocorrer em cada geração. Se esta taxa for alta, pode haver uma convergência pré-matura do algoritmo, por outro lado se for muito pequena, o algoritmo poderá levar muito tempo para obter a convergência necessária para que uma boa solução para problema;



- Taxa de mutação: este parâmetro define o percentual de mutações que deverão ocorrer para cada geração, se esta taxa for alta poderá tornar a busca completamente aleatória;
- Número de gerações: o número de gerações define quantas vezes o processo evolutivo será executado, pode funcionar como critério de parada.

### 3.1.10 Critério de Parada

O critério mais utilizado é limitar o número máximo de gerações ou até que uma boa solução seja encontrada (KOZA,1992), porém existem outros critérios baseados no acompanhamento do processo evolutivo, ou seja, enquanto houver melhoria na população, o processo evolutivo prossegue (KRAMER; ZHANG; GAPS, 2000).

## 3.2 EXEMPLO DO ALGORITMO DE PROGRAMAÇÃO GENÉTICA

Nesta seção é apresentado um exemplo de aplicação do algoritmo de Programação Genética (BANZHAF, 1998). Considere a função (3.9), formada por operadores básicos, a tarefa é encontrar um modelo de regressão para a função dada. Dez casos de treinamento foram utilizados para a tarefa de regressão, tomando  $x$  no intervalo (0 1) como mostrado na tabela (3.1).

$$y = f(x) = \frac{x^2}{2} \quad (3.9)$$

TABELA 3.1 – VALORES DE ENTRADA E SAÍDA NO CONJUNTO DE TREINAMENTO

Exemplos de treinamento	Entrada	Saída
1	0,000	0,000
2	0,100	0,005
3	0,200	0,020
4	0,300	0,045
5	0,400	0,080
6	0,500	0,125
7	0,600	0,180
8	0,700	0,245
9	0,800	0,320
10	0,900	0,405

A definição dos conjuntos e parâmetros necessários é definida da seguinte forma:

1. Conjunto de terminais:  $\{x, (-5, 5)\}$  – onde  $x$  é a variável utilizada e  $(-5, 5)$ , números inteiros entre -5 e 5 são as constantes utilizadas que uma vez definidas, não são mais modificadas durante a execução da PG;
2. Conjunto de funções:  $\{+, -, *, \%^{10}\}$  – são utilizadas operações básicas, porém se estas não forem suficientes para obter uma boa aproximação, pode-se inserir novas funções no conjunto;
3. Função de aptidão – neste caso a função de aptidão utilizada é a raiz quadrada do erro médio quadrático (Eq. )(RMSE – *Root Mean Square Error*);

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (x_i - \hat{x}_i)^2}{n}} \quad (3.10)$$

onde  $x_i$  são os valores observados e são  $\hat{x}_i$  os valores previstos.

4. Os parâmetros utilizados são: tamanho da população, profundidade máxima da árvore, probabilidade de cruzamento, método de seleção e critério de parada (por ex: número máximo de gerações). Os parâmetros estão definidos na tabela (3.2).

---

<sup>10</sup> % é a divisão protegida

TABELA 3.2 – PARÂMETROS DA PG

Parâmetros	Valores
Objetivo:	Evolui a função ajustando valores da tabela dos exemplos de treinamento
Conjunto de Terminais:	x, inteiros entre -5 e 5
Conjunto de Funções:	+, -, *, %
Tamanho da População:	600
Taxa de Cruzamento:	90%
Taxa de Mutação:	5%
Seleção:	Torneio, tamanho 4
Critério de Parada:	Nenhum
Número máx de gerações:	100
Profundidade máxima da árvore após o cruzamento:	200
Profundidade máxima após mutação:	4
Método de Inicialização:	Grow

Na geração 0, o melhor indivíduo encontrado foi a função  $f_0$ . A disposição em árvore está apresentada na figura (3.7).

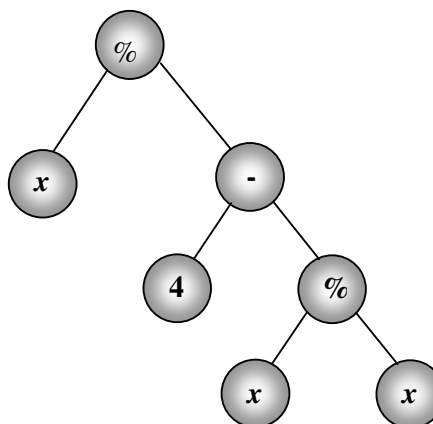


FIGURA 3.7 - MELHOR INDIVÍDUO NA GERAÇÃO 0

$$f_0(x) = \frac{x}{3} \quad (3.11)$$

O melhor indivíduo das gerações 1 e 2 são denominados de  $f_1(x)$  e  $f_2(x)$ . Assim, obtém-se as funções representadas nas equações (3.12) e (3.13):

$$f_1(x) = \frac{x}{6-3x} \quad (3.12)$$

$$f_2(x) = \frac{x}{x(x-4)-1+\frac{4}{x}-\frac{9(x-1)+x}{5x(6-3x)}} \quad (3.13)$$

Na geração 3, o melhor indivíduo encontrado apresenta uma forma simples, dada pela função (3.14):

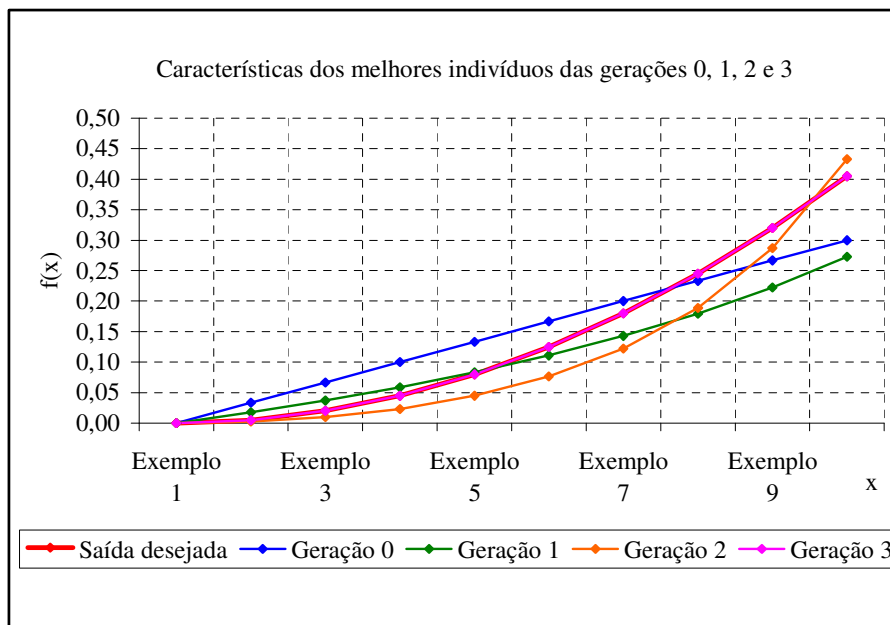
$$f_3(x) = \frac{x^2}{2} \quad (3.14)$$

As gerações seguintes combinam esta solução com outras, como consequência o tamanho do melhor indivíduo aumenta novamente, pelo fato de que não se estar armazenando o melhor indivíduo encontrado, (o que pode ser feito através de uma estratégia denominada de elitismo), porém a qualidade poderá piorar. A tabela (3.3) mostra os valores obtidos nas gerações de 0 a 3:

**TABELA 3.3 – VALORES DE SAÍDA E MELHORES INDIVÍDUOS ENCONTRADOS NAS GERAÇÕES DE 0 A 3.**

<b>Exemplos de Treinamento</b>	<b>Saída desejada</b>	<b>Geração 0</b>	<b>Geração 1</b>	<b>Geração 2</b>	<b>Geração 3</b>
Exemplo 1	0,000000	0,000000	0,000000	0,000000	0,000000
Exemplo 2	0,005000	0,033333	0,017544	0,002375	0,005000
Exemplo 3	0,020000	0,066667	0,037037	0,009863	0,020000
Exemplo 4	0,045000	0,100000	0,058824	0,023416	0,045000
Exemplo 5	0,080000	0,133333	0,083333	0,044664	0,080000
Exemplo 6	0,125000	0,166667	0,111111	0,076207	0,125000
Exemplo 7	0,180000	0,200000	0,142857	0,122140	0,180000
Exemplo 8	0,245000	0,233333	0,179487	0,188952	0,245000
Exemplo 9	0,320000	0,266667	0,222222	0,287024	0,320000
Exemplo 10	0,405000	0,300000	0,272727	0,432966	0,405000

A figura (3.8) mostra os valores de aptidão média da população inteira e do melhor indivíduo, escolhido para a execução em cada geração.



**FIGURA 3.8 - REPRESENTAÇÃO GRÁFICA DOS MELHORES INDIVÍDUOS DAS GERAÇÕES 0-3**

Verifica-se que um simples processo de PG, com poucas funções, terminais e com alguns operadores como mutação e cruzamento são capazes de produzir boas soluções. Esta natureza dinâmica do processo de PG mostra a eficiência do método.

Neste capítulo foi apresentada a fundamentação teórica da Programação Genética os aspectos necessários à aplicação desta técnica. No capítulo IV será apresentado o algoritmo *Boosting*, um algoritmo que é capaz de efetuar melhorias aos algoritmos de aprendizado de máquina. São apresentadas algumas das principais versões do algoritmo e também a nova metodologia proposta nesta tese.

## Capítulo 4

### CONJUNTO DE PREDITORES

Neste capítulo é abordado o algoritmo *Boosting* e são apresentadas algumas variações do algoritmo. Esta técnica vem sendo utilizada, com frequência, por pesquisadores na área de Aprendizagem de Máquina com o objetivo de obter melhorias nos resultados obtidos, através da utilização de algoritmos tradicionais de classificação e, mais recentemente, a técnica vem sendo aplicada com sucesso a problemas de regressão. É também apresentada a nova metodologia de *Boosting* baseada nos coeficientes de correlação entre os valores do treinamento e os valores obtidos pelo algoritmo base utilizado.

#### 4.1 BREVE HISTÓRICO DO ALGORITMO *BOOSTING*

A principal idéia da técnica *Boosting* é a de combinar a saída obtida por vários algoritmos “base” para produzir um comitê de preditores. A técnica consiste em efetuar repetidas execuções de um algoritmo de aprendizagem básico, modificando a distribuição de pesos no conjunto de treinamento e combinando os preditores obtidos, num único preditor eficiente.

A primeira proposta do algoritmo foi apresentada por Schapire (1990) para problemas de classificação binários. Em 1996, Freund e Schapire (1996) propuseram uma nova versão

do algoritmo denominada de *Adaboost*, na qual o algoritmo ajusta um modelo logístico aditivo, o número de iterações é o número de funções usadas na representação aditiva, ou seja, o número de funções que serão utilizadas para aproximar a função estimada. O algoritmo gera em cada passo uma distribuição sobre as observações da amostra, dando um peso maior às observações classificadas incorretamente no passo anterior. Outras variações deste método (FREUND & SCHAPIRE, 1999), denominadas *AdaBoost.M1* e *AdaBoost.M2*, para problemas com múltiplas classes. Neste mesmo trabalho é proposta a metodologia de *Boosting* para problemas de regressão, surge então, a versão *AdaBoost.R*. Uma implementação *ad hoc* do algoritmo denominada *AdaBoost.R* é apresentada por Drucker (1997), no qual realizou alguns experimentos para problemas de regressão tendo obtido resultados promissores. Recentemente, vários pesquisadores como Friedman, Hastie e Tibshirani (1998; 2001), Duffy (2000), Zemel e Elmasri, (2001) e Ridgway (1999) comparam o algoritmo *Boosting* com um algoritmo de gradiente descendente, que otimiza a função de perda (resíduos). Eles mostraram que quando a função de perda utilizada é a exponencial, o algoritmo *Adaboost* se assemelha ao método de Newton. A proposta de Solomatine e Shrestha (2004) propõe uma forma de atualização dos pesos utilizando o erro relativo ao invés do erro absoluto.

Recentemente as estratégias *Boosting* foram usadas em vários problemas reais, muitas destas aplicações e referências, podem ser encontradas em *Boosting* (2006). O detalhamento do algoritmo *AdaBoost* e algumas versões são apresentadas nas seções seguintes.

## 4.2 ADABOOST PARA PROBLEMAS DE CLASSIFICAÇÃO BINÁRIO

Em um problema de classificação binário, tem-se um conjunto de objetos que devem ser separados em duas classes, de forma que esta classificação seja consistente. A classificação é obtida com base em alguns dados que se tem sobre o problema em questão.

O algoritmo *Adaboost* utiliza como entrada um conjunto de treinamento  $(x_1, y_1), \dots, (x_m, y_m)$  onde cada  $x_i$  é um vetor de características, pertencente a algum espaço  $X$  e  $y_i$  pertence ao conjunto  $Y = \{-1, +1\}$ , ou seja, o algoritmo é utilizado para problemas de classificação com duas classes. O *Adaboost* chama um classificador base, repetidamente, num conjunto de  $t$

execuções,  $t = 1, \dots, T$ . Uma das principais idéias do algoritmo é modificar a distribuição, ou conjunto de pesos, sobre o conjunto de treinamento. O peso desta distribuição no exemplo de treinamento  $i$  na execução  $t$  é denotado por  $D_t(i)$ . Inicialmente os pesos são todos iguais, mas a cada execução, os pesos dos exemplos classificados incorretamente são incrementados de forma que o classificador “base” seja forçado a atuar com maior intensidade sobre estes exemplos no conjunto de treinamento, a taxa de erro é calculada contando o número de classificações incorretas (Eq. 4.1). Observa-se que o erro é calculado de acordo com a distribuição  $D_t$  sobre a qual o classificador fraco foi treinado. O pseudo-código do algoritmo está apresentado no quadro (4.1) (FREUND & SCHAPIRE, 1996).

$$e_t = \Pr_{i \sim D_t} [h_t(x_i) \neq y_i] = \sum_{i: h_t(x_i) \neq y_i} D_t(i) \quad (4.1)$$

Uma vez que a hipótese  $h_t$  tenha sido recebida, o *Adaboost* escolhe um parâmetro  $\alpha_t$  pertencente ao conjunto dos números reais, o qual mede a importância que é dada à hipótese  $h_t$ . O valor de  $\alpha_t$  para problemas binários, bem como a atualização do vetor de pesos  $D_t$ , estão detalhados na descrição do algoritmo apresentada no quadro (4.1).

Para formular a hipótese final  $H$ , cada hipótese  $h_t$  contribui com uma certa confiança dada por  $\alpha_t$ .



**QUADRO 4.1 – ALGORITMO ADABOOST**

**Dado:**  $S = \{(x_1, y_1), \dots, (x_m, y_m); x_i \in X, y_i \in \{-1, +1\}\}$

**Algoritmo Adaboost:**

Inicialize  $D_1(i) := \frac{1}{m} \quad \forall (x_i, y_i) \in S$

**For**  $t = 1, \dots, T$ , faça

Treine o classificador base usando a distribuição  $D_t$

Obtenha a hipótese fraca  $h_t: X \rightarrow \{-1, +1\}$

Calcule  $\alpha_t$ :  $\alpha_t = \frac{1}{2} \ln \frac{1 - e_t}{e_t}$ , onde  $e_t$  é a taxa de erro do classificador  $h_t$ .

Atualize a distribuição:

$$D_{t+1}(i) = \frac{D_t(i)}{Z_t} = \begin{cases} e^{-\alpha_t}, & \text{se } h_t(x_i) = y_i \\ e^{\alpha_t}, & \text{se } h_t(x_i) \neq y_i \end{cases}$$

$$= \frac{D_t(i) \exp(-\alpha_t y_i h_t(x_i))}{Z_t}$$

Onde  $Z_t$  é o fator de normalização

**End For**

**Saída:** Hipótese Final:  $H(x) = \text{sign}\left(\sum_{t=1}^T \alpha_t * h_t(x)\right)$

### 4.3 BOOSTING PARA PROBLEMAS DE REGRESSÃO

Nos problemas de regressão o que se deseja é encontrar uma aproximação para determinados valores de uma variável contínua. Uma função  $f(x)$ , num certo intervalo, ou seja, dada uma função  $f(x)$ , deve-se encontrar uma função  $g(x)$ , tal que  $f(x_i) = g(x_i) \quad \forall x_i \in X$ , onde  $X$  é um conjunto de valores do intervalo considerado. A função  $f(x)$  não é necessariamente conhecida, o que se conhece é um conjunto de pontos  $\{x, f(x) | x \in X\}$ . Sobre um conjunto de treinamento  $S = \{(x_1, g(x_1)), \dots, (x_m, g(x_m))\}$ , aplica-se um algoritmo básico de aprendizagem, na tentativa de encontrar uma função  $g$  próxima da função  $f$  no domínio desejado. Há muitas maneiras de se medir o quão próxima a função  $g$  está da função  $f$ . Uma maneira de verificar esta proximidade é verificar quando o erro médio quadrático se torna pequeno, ou garantir que a função  $g$  esteja próxima da função  $f$  em todo o domínio. Ao se utilizar o algoritmo *Boosting*, pode-se obter uma aproximação ainda maior. Após se utilizar um algoritmo de aprendizado

básico repetidas vezes, em diferentes amostras, obtém-se várias hipóteses que serão então combinadas, de forma a obter uma melhor aproximação para a função  $f$ .

#### 4.3.1 O Algoritmo *Adaboost.R*

O algoritmo *AdaBoost.R* (FREUND & SCHAPIRE, 1998), funciona de maneira semelhante ao *AdaBoost*. A idéia é usar alguma ordenação dos valores para diferenciar as predições corretas das incorretas, como se fosse um problema de classificação. Nos problemas de regressão não é possível obter valores de predição exatos como nos problemas de classificação. As discrepâncias entre os valores preditos e observados são inevitáveis. A medida desta discrepância é que vai permitir dizer se o valor previsto é ou não aceitável.

Da mesma forma que no algoritmo *Adaboost*, o algoritmo base recebe os exemplos  $(x_i, y_i)$  do conjunto de treinamento para gerar a hipótese  $h: X \rightarrow Y$ , onde  $Y = [0, 1]$ , tal que para cada valor  $x$  o preditor forneça um valor aproximado para  $y$ , ou seja, o algoritmo procura gerar uma hipótese  $h$ , com o menor erro de previsão possível. Assume-se que são conhecidos os exemplos do conjunto de treinamento:  $S = \{(x_1, y_1), \dots, (x_m, y_m)\}$ . O objetivo é então minimizar a soma dos erros médios quadráticos obtidos (*MSE*), conforme a equação (4.2).

$$MSE = \frac{1}{N} \sum_{i=1}^N (h(x_i) - y_i)^2 \quad (4.2)$$

Os problemas de regressão são então reduzidos à problemas de classificação binários (ALLWEIN; SCHAPIRE; SINGER, 2000) para então se aplicar o *Adaboost*. Para cada  $(x_i, y_i)$  do conjunto de treinamento define-se um conjunto contínuo de exemplos indexado pelos pares  $(i, y)$  para todo  $y \in [0, 1]$ , tal que a instância é  $\tilde{x}_{i,y} = (x_i, y)$  e a classe é  $\tilde{y}_{i,y} = [y \geq y_i]$ . Assim, informalmente tem-se que, cada instância  $(x_i, y_i)$  é designada a um conjunto de questões binárias, uma para cada  $y \in Y$ , e cada uma da forma: “ $y_i$  é maior ou menor que  $y$ ?”. Ou de maneira similar, pode-se dizer que cada hipótese  $h: X \rightarrow Y$  é reduzida à uma hipótese de valor binário  $\tilde{h}: X \times Y \rightarrow \{0, 1\}$  definida pela regra (4.3):

$$\tilde{h}(x, y) = [y \geq h(x)] \quad (4.3)$$

Neste caso,  $\tilde{h}$  tenta responder a questão binária usando o valor estimado de  $h(x)$ . Como nos problemas de classificação, o vetor de pesos  $D$  é inicializado tal que  $D(i) = 1/m$

sobre o conjunto de treinamento, porém para o problema de regressão, esta distribuição é designada à distribuição de densidade  $\tilde{D}$  sobre os pares  $(i, y)$  de maneira que ao minimizar os erros de classificação, minimiza-se também o *MSE* do problema original. Assim, define-se a distribuição  $\tilde{D}$  de acordo com a equação (4.4),

$$\tilde{D}(i, y) = \frac{D(i)|y - y_i|}{Z} \quad (4.4)$$

onde  $Z$  é uma constante de normalização calculada conforme a equação (4.5)

$$Z = \sum_{i=1}^N D(i) \int_0^1 |y - y_i| dy \quad (4.5)$$

O algoritmo mantém um peso  $w_{i,y}^t$  para cada instância  $i$  e classe  $y \in Y$ . O vetor de pesos inicial é a densidade  $\tilde{D}$  definida acima, normalizando os pesos  $w^t$ , a densidade  $p^t$  é definida de acordo com a equação (4.6) e é utilizada para calcular a perda média de  $h_t$ , definida pela equação (4.7).

$$p^t = \frac{w^t}{\sum_{i=1}^m \int w_{i,y}^t dy} \quad (4.6)$$

$$\mathcal{E}_t = \sum_{i=1}^m \left| \int_{y_i}^{h_t(x_i)} p_{i,y}^t dy \right| \quad (4.7)$$

O objetivo do algoritmo base é encontrar a hipótese  $h: X \rightarrow Y$  que minimiza a perda  $\mathcal{E}_t$ . Finalmente a hipótese final  $h_f$  é determinada, calculando a mediana ponderada das hipóteses fracas, conforme equação (4.8).

$$h_f(x) = \inf \left\{ y \in Y : \sum_{t: h_t(x) \leq y} \log(1/\beta_t) \geq \frac{1}{2} \sum_i \log(1/\beta_t) \right\} \quad (4.8)$$

Como a hipótese final é binária, para cada  $x$  há somente um valor de  $y$  para o qual  $\tilde{h}_f(x, y') = 0$  para todo  $y' < y$  e  $\tilde{h}_f(x, y') = 1$  para todo  $y' > y$ , que será exatamente o valor de  $y$  calculado pela hipótese  $h_f(x)$  (Eq. 4.8).

**QUADRO 4.2 – ALGORITMO ADABOOST.R**

**Dado:** um conjunto  $S = \{(x_1, y_1), \dots, (x_m, y_m)\}; x_i \in X, y_i \in [0, 1]$

Considere um algoritmo de regressão básico e a distribuição  $D$  sobre o conjunto  $S$ .

**Algoritmo:**

$D_1(i)$  é o peso da amostra  $(x_i, y_i)$

Inicialize  $D_1(i) := \frac{1}{m} \quad \forall (i, y)$  associado a cada exemplo  $(x_i, y_i)$  e  $\forall y \in [0, 1]$ .

Inicialize o vetor de pesos:  $w_{i,y}^1 = \frac{D(i)|y - y_i|}{Z}$ , para  $i = 1, \dots, m$  e  $y \in Y$  e  $Z$  é um fator de normalização.

**For**  $t = 1$  até  $T$ :

$$1. \text{ Faa: } p^t = \frac{w^t}{\sum_{i=1}^m \int w_{i,y}^t dy}$$

2. Chame o algoritmo de regressão base, com a função de densidade  $p^t$ ; e retorne a hipótese  $h_t: X \times Y$

3. Calcule a perda de  $h_t$ :

$$\epsilon_t = \sum_{i=1}^m \left| \int_{y_i}^{h_t(x_i)} p_{i,y}^t dy \right|$$

Se  $\epsilon_t > 1/2$  então faça  $T = t-1$  e saia do *loop*.

$$4. \text{ Faa } \beta_t = \epsilon_t / (1 - \epsilon_t)$$

5. Faça o novo vetor de pesos ser:

$$w_{i,y}^{t+1} = \begin{cases} w_{i,y}^t, & \text{se } y_i \leq y \leq h_t(x_i) \text{ ou } h_t(x_i) \leq y \leq y_i \\ w_{i,y}^t, & \text{caso contrário} \end{cases}$$

**End for**

Hipótese Final:  $h_t(x) = \inf\{y \in Y : \sum_{t: h_t(x) \leq y} \log(1/\beta_t) \geq 1/2 \sum_t 1/\beta_t\}$

#### 4.3.2 O Algoritmo *Adaboost.RT*

Uma proposta de modificação no algoritmo de *Adaboost.R* para problemas de regressão é apresentada por Solomatine e Shrestha (2004). Em seu algoritmo, denominado

*AdaBoost.RT*, os autores propõem uma metodologia, cuja idéia principal, é a introdução de uma constante  $\phi$ , como um valor limiar de erro relativo que deve classificar a predição como correta ou incorreta, esta taxa de erro é calculada contando o número de predições corretas e incorretas, conforme equação (4.9).

$$i : \left| \frac{f_t(x_i) - y_i}{y_i} \right| > \phi \quad (4.9)$$

Este valor de  $\phi$  é então utilizado para a atualização do vetor de pesos, equação (4.10).

$$D_{t+1}(i) = \frac{D_t(i)}{Z_t} x \begin{cases} \alpha_t, se \left| \frac{f_t(x_i) - y_i}{y_i} \right| \leq \phi \\ 1, caso contrário \end{cases} \quad (4.10)$$

A principal diferença entre o algoritmo *Adaboost.RT* e os demais algoritmos de *Boosting* é no cálculo da função de perda, que usa o erro relativo ao invés do erro absoluto, o que faz com seja dada maior ênfase aos exemplos que possuem maiores erros de previsão, de forma que na próxima iteração estes erros sejam reduzidos. A outra diferença deste algoritmo em relação aos demais algoritmos de *Boosting* é quanto ao critério de parada, neste caso pode-se definir o número de iterações necessárias, enquanto que na maioria dos demais algoritmos, o critério de parada utilizado, é o de verificar quando a taxa de erro torna-se maior ou igual a 0.5, por fim a combinação final dos preditores obtidos é feita pela média ponderada dos preditores obtidos e do valor de confiança  $\alpha$  dado à cada preditor, enquanto que nos outros algoritmos a combinação final é feita utilizando a mediana geométrica. A fórmula de combinação final dos preditores está representada na equação (4.11).

$$f_{fin}(x) = \sum_t \log\left(\frac{1}{\alpha_t}\right) \frac{f_t(x)}{\sum_t \log\left(\frac{1}{\alpha_t}\right)} \quad (4.11)$$

Uma desvantagem deste algoritmo é a necessidade de se encontrar um valor ótimo para  $\phi$ , se os valores de  $\phi$  forem muito baixos, poucos exemplos serão corretamente classificados, por outro lado, se os valores de  $\phi$  forem muito altos, poucos exemplos obterão uma boa previsão. Assim, para encontrar valores ótimos para  $\phi$ , é necessário utilizar algum procedimento de minimização do erro relativo absoluto, antes de se aplicar o algoritmo *AdaBoost.RT*. Os experimentos realizados pelos autores desta técnica mostraram que o algoritmo *Adaboost.RT* efetua predições melhores que as realizadas por um único preditor

com uma confiança de 99%, porém para que seja comparado aos demais algoritmos de *Boosting* propostos, há necessidade de se realizar mais experimentos.

#### 4.3.3 O Algoritmo *GPBoost*

Uma extensão da Programação Genética foi proposta por Iba (1999), através de técnicas de reamostragem, na qual a população inicial é dividida em sub-populações e cada uma destas é evoluída utilizando o algoritmo de *Boosting*. Este algoritmo foi chamado por Iba de *GPBoost*. Baseado neste trabalho, Paris, Robilliard e Folunpt (2001) propuseram a utilização do algoritmo *GPBoost* com algumas modificações, especialmente na interpretação da distribuição de pesos, para problemas de classificação e de regressão. Existem algumas dificuldades ao se aplicar a técnica de *Boosting* à Programação Genética, e a mais importante delas é que a PG não é considerada um algoritmo básico de aprendizagem, desta forma não é esperada uma grande redução de erros, embora vários testes tenham mostrado resultados promissores.

O algoritmo funciona da seguinte maneira: para obter a imagem de  $x$ , cada função  $f_i$  fornece um valor  $f_i(x)$ . Estes valores são ordenados e então é calculada a mediana geométrica que será a função de aproximação final,  $F(x)$ . A função de aptidão é definida como sendo a soma das diferenças absolutas, multiplicadas pelos valores dos pesos da distribuição. O valor desta função é multiplicado por  $m$  (Eq. 4.12), assim a primeira execução do algoritmo *Boosting* é a idêntica à primeira execução da PG. A função de aptidão pode ser definida de acordo com o problema em questão e uma fórmula bastante utilizada para esta função é a Raiz quadrada do Erro Médio Quadrático (*RMSE*). As demais etapas do algoritmo são semelhantes ao *Adaboost.R*. O algoritmo completo está apresentado no quadro (4.2).

$$fit = \sum_{i=1}^m \left( \left| f(x_i) - y_i \right| * D_i(i) \right) * m \quad (4.12)$$

**QUADRO 4.3 – ALGORITMO ADABOOST.RT**

**Dado:** um conjunto  $S = \{(x_1, y_1), \dots, (x_m, y_m)\}$ ;  $x_i \in X, y_i \in R$ , PG é o algoritmo base.

**Algoritmo *GPboost*:**

$D_1(i)$  é o peso da amostra  $(x_i, y_i)$ ,  $D_1(i) := \frac{1}{m} \quad \forall (x_i, y_i) \in S$

**For**  $t = 1$  até  $T$  faça

Execute a PG sobre  $D_t$  com a função de aptidão:

$fit = \sum_{i=1}^m (|f(x_i) - y_i| * D_t(i)) * m$ , onde  $f$  é uma função na população PG.

A melhor função encontrada é  $f_t$ .

Calcule a perda para cada amostra:  $L_i = \frac{|f_t(x_i) - y_i|}{\max_{i=1 \dots m} |f_t(x_i) - y_i|}$

Calcule a perda média:  $\bar{L} = \sum_{i=1}^m L_i D_i$  e

Faça  $\beta_t = \frac{\bar{L}}{1 - \bar{L}}$  ser a confiança dada à função  $f_t$  e atualize a distribuição:

$D_{t+1}(i) := \frac{D_t(i)^{1-L_i}}{Z_t}$ , com  $Z_t$  um fator de normalização.

**End For**

**Saída:** Hipótese Final:

$$F(x) = \min\{y \in R : \sum_{t: f_t(x) \leq y} \log\left(\frac{1}{\beta_t}\right) \geq \frac{1}{2} \sum_{t=1}^m \log\left(\frac{1}{\beta_t}\right)\}$$

#### 4.3.4 Exemplo do algoritmo *GPBoost*

Nesta seção é apresentado um exemplo de um problema de regressão para mostrar o funcionamento do algoritmo *GPBoost* (PARIS; ROBILLIARD; FONLUPT, 2001). Os parâmetros estão definidos na tabela (4.1).

TABELA 4.1 – PARÂMETROS PARA O *GPBOOST*

Terminais	X
Funções	+, -, *, /
Conjunto de treinamento	5 exemplos da função $y = x^2$ $\{(-0,1;0,005), (-0,5; 0,125), (1; 0,5), (0,2; 0,02), (0,7;0,245)\}$
Aptidão	$\sum_{i=1}^5 ( f(x_i) - y_i  * D_i) * 5$
População	20
Gerações	10

FONTE: RETIRADO DO ARTIGO (PARIS; ROBILLIARD; FONLUPT, 2001)

A evolução dos pesos durante as três primeiras execuções do algoritmo *Boosting* é apresentada na tabela (4.2).

TABELA 4.2 - EVOLUÇÃO DOS PESOS NAS TRÊS PRIMEIRAS EXECUÇÕES DO *BOOSTING*

Amostra		Distribuição		
X	Y	$D_1$	$D_2$	$D_3$
-0,1	0,01	0,2	0,135645	0,091587
-0,5	0,25	0,2	0,155274	0,097203
1	1	0,2	<b>0,433427</b>	0,271331
0,2	0,04	0,2	0,140010	0,116996
0,7	0,49	0,2	0,135645	<b>0,422883</b>

Na primeira execução, os pesos são iguais para todos os exemplos da amostra, ou seja, tem-se uma execução da PG tradicional. A melhor função obtida na primeira execução do *GPBoost* é  $f_1(x) = x^4$ , a qual não é uma boa aproximação para a função desejada. Na próxima etapa, calcula-se a atualização no vetor de pesos, obtendo  $D_2$ . Uma nova execução de PG é realizada, e como resultado tem-se a função  $f_2(x) = \frac{x^4}{1+x}$ . Observa-se, que esta função fornece a melhor aproximação para o exemplo que possui o maior peso, no caso (1,1), cujo peso é 0,433427. Na tabela (4.3), encontram-se as melhores funções obtidas e a confiança das hipóteses para todas as execuções.



**TABELA 4.3 – RESULTADOS OBTIDOS DAS EXECUÇÕES *BOOSTING***

Hipóteses	Confiança
$f_1(x) = x^4$	0,309380
$f_2(x) = \frac{x^4}{1+x}$	0,200801
$f_3(x) = x^4$	0,414084
$f_4(x) = \frac{x^4(x+2)}{(3x^3-x+1)(x+1)}$	0,720217
$f_5(x) = \frac{x^2}{1+x}$	0,158855

Hipótese Final: Para cada valor dado  $x$ , cada função  $f_i$  fornece um valor de  $f_i(x)$ . Obtém-se  $F(x)$ , calculando a mediana geométrica de  $f_i(x)$  com os pesos sendo seus respectivos coeficientes de confiança.

$$\begin{aligned}
 \text{Se } x = 0,5 \quad \rightarrow \quad & f_1(x) = 0,0625 \\
 & f_2(x) = 0,041667 \\
 & f_3(x) = 0,0625 \\
 & f_4(x) = 0,116305 \\
 & f_5(x) = 0,166667
 \end{aligned}$$

A mediana geométrica é dada por  $\frac{1}{2} \sum_{t=1}^m \log\left(\frac{1}{\beta_t}\right) = 2,907245$  e os valores são ordenados em ordem crescente. Para  $x = 0,5$ , tem-se:  $f_2(x) \leq f_1(x) \leq f_3(x) \leq f_4(x) \leq f_5(x)$ . Então ordena-se os termos  $\log\left(\frac{1}{\beta_t}\right)$  na mesma ordem que as  $f_t$ :

$$\log\left(\frac{1}{\beta_2}\right) = 1,60544$$

$$1,60544 + \log\left(\frac{1}{\beta_1}\right) = 2,77862$$

$$2,77862 + \log\left(\frac{1}{\beta_3}\right) = 3,660312 \geq 2,907245$$

Quando o valor obtido é maior ou igual à mediana geométrica, tem-se a hipótese final. Neste caso é dada por  $f_3(x)$ . Assim, tem-se que:  $F(0,5) = f_3(0,5) = 0,0625$ .

Os resultados deste experimento mostraram uma melhoria de 20% nos resultados em comparação aos resultados obtidos utilizando somente a PG tradicional.

Na seção (4.4) é apresentada uma adaptação do algoritmo *Boosting* para a sua utilização na previsão de séries temporais, bem como é proposta uma modificação do algoritmo, na atualização dos pesos e na forma de combinação final das funções, o que caracteriza uma das contribuições originais desta tese.

#### 4.4 ADAPTAÇÃO DO ALGORITMO BOOSTING UTILIZANDO COEFICIENTE DE CORRELAÇÃO

Após ter sido realizada a revisão de literatura a respeito dos algoritmos de *Boosting* e as várias modificações propostas por diversos autores, concluiu-se que estes algoritmos têm sido bastante utilizados para problemas de classificação, porém menor ênfase tem sido dada quanto à sua aplicação aos problemas de regressão. Verificou-se também, que o algoritmo não tem sido utilizado juntamente com métodos de previsão de séries temporais. Assim, surgiu a idéia de verificar o desempenho dos algoritmos de *Boosting* já existentes, para melhorar os resultados de previsão obtidos através da aplicação de alguns métodos de previsão. O método foi proposto para melhorar o desempenho dos preditores obtidos, utilizando como algoritmo base a PG.

A metodologia proposta é uma nova abordagem do algoritmo *Boosting*, e tem uma fundamentação empírica ao utilizar o coeficiente de correlação para a atualização do vetor de pesos, o que influencia diretamente na minimização da função de perda, pois este coeficiente fornece uma relação entre as variáveis em questão, que neste caso são os valores previstos e valores observados no conjunto de treinamento. O mesmo coeficiente foi também utilizado na combinação final dos preditores obtidos por um algoritmo base. A definição do coeficiente de correlação é dada a seguir.

**Definição:** diz-se que existe correlação entre duas variáveis quando as alterações sofridas por uma delas são acompanhadas por modificações nas outras. Ou seja, no caso de

duas variáveis  $X$  e  $Y$ , verifica-se se as variações que ocorrem em  $X$ , correspondem às mesmas variações que ocorrem em  $Y$ . Assim, a correlação revela se existe uma relação funcional entre uma variável e as demais. A fórmula da correlação entre as variáveis  $x$  e  $y$  é dada pela equação (4.13)

$$\rho(x, y) = \frac{\sum_{i=1}^m (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^m (x_i - \bar{x})^2 \sum_{i=1}^m (y_i - \bar{y})^2}} \quad (4.13)$$

onde  $m$  é o número de observações da amostra,  $\bar{x}$  é a média da amostra  $X$  e  $\bar{y}$  é a média da amostra  $Y$  e  $\rho$  é o coeficiente de correlação linear de Pearson.

#### 4.4.1 Algoritmo *Boosting* utilizando o Coeficiente de Correlação (*BCI – Boosting Correlation Improvement*)

Considerando que o coeficiente de correlação relaciona duas variáveis, ou seja, de acordo com seu valor, é possível verificar quão próxima uma variável está da outra, o estudo foi empiricamente direcionado de forma a obter uma combinação final de preditores, utilizando estes coeficientes de correlação entre os valores observados e os valores previstos, pois quanto mais próximos estes valores estiverem uns dos outros, melhores serão os valores de previsão obtidos.

O método proposto utiliza estes coeficientes de correlação na atualização da distribuição de pesos do algoritmo de *Boosting* e na combinação final dos preditores obtidos e segue a metodologia do algoritmo *Adaboost.R*. Inicializa a distribuição de pesos com valores iguais para todos os exemplos da amostra, ou seja, dado o conjunto de observações no tempo  $T = \{x_1, x_2, \dots, x_m\}$ , o vetor de pesos para a primeira iteração é:  $P_1(i) = 1/m$  para todo  $x_i \in T$ . Utiliza-se então um algoritmo de aprendizagem básico para obter os valores previstos. O erro é então calculado, utilizando a função de perda exponencial apresentada na equação (4.14).

$$L_t(x_i) = 1 - \exp\left(-\frac{|f_t(x_i) - y(x_i)|}{\max_{i=1 \dots m} |f_t(x_i) - y(x_i)|}\right) \quad (4.14)$$

Onde  $f_t(x_i)$  representa os valores previstos pelo algoritmo base na iteração  $t$  e  $y(x_i)$  são os valores observados. A escolha da utilização da função exponencial para calcular a perda obtida pelo algoritmo se deu em função dos resultados obtidos em vários experimentos que comparavam o desempenho do algoritmo utilizando as funções de perda linear (Eq. 4.15), quadrática (Eq. 4.16) e exponencial. Além disso, pesquisas recentes (HASTIE; TIBSHIRANI; FRIEDMAN, 2001) mostraram o melhor desempenho do algoritmo *Adaboost* quando a função de perda utilizada é a exponencial.

$$L_i = \frac{|f_t(x_i) - y(x_i)|}{\max_{i=1, \dots, m} |f_t(x_i) - y(x_i)|} \quad (4.15)$$

$$L_i = \frac{|f_t(x_i) - y(x_i)|^2}{\max_{i=1, \dots, m} |f_t(x_i) - y(x_i)|^2} \quad (4.16)$$

O cálculo dos coeficientes de correlação entre os valores observados e os valores previstos são feitos de acordo com a equação (4.17).

$$\rho_t(f_t(x_i), y(x_i)) = \frac{\sum_{i=1}^m (f_t(x_i) - f_t(\bar{x}))(y(x_i) - y(\bar{x}))}{\sqrt{\sum_{i=1}^m (f_t(x_i) - f_t(\bar{x}))^2 \sum_{i=1}^m (y(x_i) - y(\bar{x}))^2}} \quad (4.17)$$

onde  $m$  é o número de observações do conjunto de treinamento,  $f_t(\bar{x})$  é a média da amostra obtida por  $f_t(x)$  e  $y(\bar{x})$  é a média dos valores observados no conjunto de treinamento. A atualização dos pesos é feita multiplicando-se o coeficiente de correlação obtido na iteração  $t$  do algoritmo *BCI* pelos vetores de peso  $P_t$  e de perda  $L_t$ , a atualização é feita de acordo com a equação (4.18).

$$P_{t+1}(x_i) = \rho_t(f_t(x_i), y(x_i)) * P_t * L_t(x_i) \quad (4.18)$$

Finalmente, após terem sido completadas as  $T$  iterações do algoritmo *BCI*, a combinação final dos preditores é feita conforme a equação (4.19).

$$F(x_{i-1}) = \frac{\sum_{t=1}^T \rho_t(f_t(x_i), y(x_i)) * f_t(x_i)}{\sum_{t=1}^T \rho_t(f_t(x_i), y(x_i))}, \quad i = 1, \dots, m-1$$

$$F(x_m) = \frac{\sum_{t=1}^T f_t(x_i)}{T}, \quad i = m \quad (4.19)$$

O pseudo-código do algoritmo BCI está apresentado no quadro (4.4).

**QUADRO 4.4 – ALGORITMO BCI**

Dado  $T = \{x_1, x_2, \dots, x_m\}$

Inicializar  $P_1(i) = 1/m$ , para todo  $x_i \in T$

**For**  $t = 1, \dots, T$

Execute um algoritmo básico de aprendizado, e determine  $f_t$ , modelo de previsão para a iteração  $t$ .

Calcule a perda para cada exemplo:

$$L_t = 1 - \exp\left(-\frac{|f_t(x_i) - y(x_i)|}{\max_{i=1..m} |f_t(x_i) - y(x_i)|}\right)$$

onde  $y(x_i)$  são os valores observados.

Calcule o coeficiente de correlação entre  $f_t(x_i)$  e  $y(x_i)$ .

$$\rho_t(f_t(x_i), y(x_i)) = \frac{\sum_{i=1}^m (f_t(x_i) - f_t(\bar{x}))(y(x_i) - y(\bar{x}))}{\sqrt{\sum_{i=1}^m (f_t(x_i) - f_t(\bar{x}))^2 \sum_{i=1}^m (y(x_i) - y(\bar{x}))^2}}$$

Faça a atualização dos pesos:  $P_{t+1}(x_i) = \rho_t(f_t(x_i), y(x_i)) * P_t * L_t(x_i)$

**End for**

**Saída Final:** obter a combinação final dos preditores,  $F(x)$ :

$$F(x_{i-1}) = \frac{\sum_{t=1}^T \rho_t(f_t(x_i), y(x_i)) * f_t(x_i)}{\sum_{t=1}^T \rho_t(f_t(x_i), y(x_i))}, i = 1, \dots, m-1$$

$$F(x_m) = \frac{\sum_{t=1}^T f_t(x_i)}{T}, i = m$$

Com o objetivo de verificar o desempenho do algoritmo BCI, vários experimentos foram realizados. Estes experimentos envolvem previsão de séries temporais financeiras, algumas séries temporais reais e séries artificialmente simuladas. O algoritmo também foi testado para problemas de regressão múltipla. O detalhamento destes experimentos é

apresentado no capítulo VI, em sua grande maioria, nestes experimentos, o algoritmo base utilizado foi a PG, assim o método proposto aqui neste capítulo é denominado de *BCIGP* (*Boosting Correlation Improvement using Genetic Programming*).

# Capítulo 5

## TESTES DE HIPÓTESES

**A**o se comparar dois algoritmos, nem sempre é uma tarefa simples saber qual deles apresenta o melhor desempenho utilizando apenas as taxas de erros. O desvio padrão pode ser considerado como a imagem da robustez do algoritmo, ou seja, se os erros calculados sobre diferentes conjuntos de teste, provenientes de diferentes conjuntos de treinamento da mesma amostra, forem muito diferentes de um experimento para o outro, diz-se que o algoritmo não é robusto a mudanças no conjunto de treinamento proveniente de uma mesma distribuição. Assim, para verificar dentre dois algoritmos, qual deles possui o melhor desempenho, basta verificar se a diferença entre os algoritmos é ou não significativa (WEISS; INDURKHYA, 1999). Neste capítulo são apresentados os testes estatísticos realizados para a verificação da validade e significância dos modelos obtidos através da metodologia proposta nesta tese.

### 5.1 TESTE “*t*” - DESEMPENHO DE ALGORITMOS

Os testes estatísticos de hipóteses julgam uma afirmação quanto ao valor numérico de um parâmetro de uma população com base em uma amostra. Os testes “*t*” fazem parte da

família destes testes e têm como objetivo comparar duas médias, ou uma média com um valor padrão. Há várias abordagens na utilização do teste t, neste trabalho, utiliza-se o teste para dados emparelhados, ou seja, aqueles que são obtidos aos pares, em indivíduos ou ensaios relacionados, decorrentes do fato de que as observações são obtidas de medições repetidas a partir de um mesmo conjunto de indivíduos. Na execução do teste, avalia-se o valor-p que é o valor de  $\alpha$  quando  $t = t_{crítico}$ . É o valor de  $\alpha$  em que ocorre entre a transição entre a aceitação e a rejeição da hipótese nula. O cálculo do valor-p permite uma conclusão direta sobre a hipótese nula, caso o valor-p seja menor do que o nível de significância escolhido rejeita-se a hipótese nula.

As hipóteses consideradas são as seguintes:

- $H_0$ : Hipótese nula é a hipótese da igualdade ou hipótese da diferença nula. É sempre a hipótese testada. Esta hipótese poderá ser aceita ou rejeitada pelo teste.
- $H_1$ : Hipótese alternativa é a hipótese que será aceita se o teste rejeitar a  $H_0$ . Esta hipótese poderá ter configurações diferentes, conforme o teste seja unilateral ou bilateral.
  - $H_1: \mu_A \neq \mu_B$  representa o teste bilateral, onde se verifica se a média da amostra A difere da média da amostra B;
  - $H_1: \mu_A > \mu_B$  é o teste unilateral ou unicaudal.

Outro fator importante a ser considerado, é o nível de confiança  $P = 1 - \alpha$ , este valor é definido pelo usuário do teste e representa a probabilidade de aceitar a  $H_0$ , quando esta é verdadeira. Delimita a região de aceitação da hipótese  $H_0$ , como mostrado na figura (5.1).

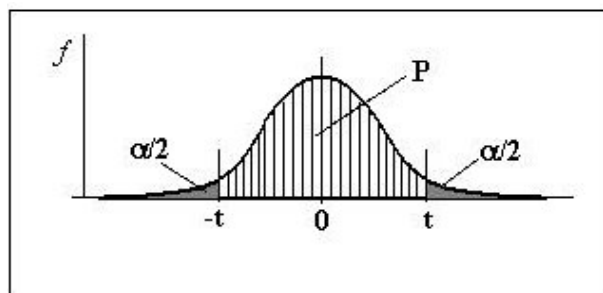


FIGURA 5.1 – NÍVEL DE CONFIANÇA  $P$  E NÍVEL DE SIGNIFICÂNCIA  $\alpha$

O nível de significância  $\alpha$  representa o risco de rejeitar a hipótese  $H_0$ , sendo esta verdadeira.

A estatística “ $t$ ” para dados emparelhados, é calculada através da equação (5.1)



$$t = \frac{\bar{d}\sqrt{n}}{s_d} \sim t_{n-1} \quad (5.1)$$

onde  $\bar{d}$  é a média das diferenças entre os pares,  $n$  é o número de pares e  $s_d$  é o desvio padrão das diferenças entre os pares.

## 5.2 COMPARAÇÃO DOS ERROS NO CONJUNTO DE TESTE

Outra maneira de verificar as discrepâncias existentes entre os valores reais observados  $x_i$  e os valores previstos  $\hat{x}_i$  é utilizar uma das medidas de erro apresentadas nas equações de (5.2) a (5.5).

$$\text{Erro Médio:} \quad ME = \frac{\sum_{i=1}^n (x_i - \hat{x}_i)}{n} \quad (5.2)$$

$$\text{Erro Absoluto Médio:} \quad MAE = \frac{\sum_{i=1}^n |x_i - \hat{x}_i|}{n} \quad (5.3)$$

$$\text{Erro Quadrático Médio:} \quad MSE = \frac{\sum_{i=1}^n (x_i - \hat{x}_i)^2}{n} \quad (5.4)$$

$$\text{Raiz Quadrada do EQM:} \quad RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i - \hat{x}_i)^2}{n}} \quad (5.5)$$

onde:

$x_i$  é o valor da observação no instante  $i$ ;

$\hat{x}_i$  é o valor previsto para o instante  $i$ ;

$n$  é o número de observações.

Quanto menor for o erro obtido, melhor será o ajuste do modelo de previsão. O método mais utilizado é do Erro Quadrático Médio ( $MSE$ ).

Para comparar o desempenho de todos os algoritmos utilizados nos experimentos foi realizado o teste ANOVA, descrito a seguir.

### 5.3 DESEMPENHO DOS ALGORITMOS – TESTE ANOVA

Uma técnica estatística bastante utilizada para testar a hipótese de igualdade na média entre vários algoritmos é o teste de Análise de Variância (ANOVA), que verifica se existe ou não uma diferença significativa entre os métodos propostos utilizados. A ANOVA utiliza o teste F descrito na seção seguinte.

#### 5.3.1 Teste F para comparar duas ou mais Médias Populacionais

Inicialmente é aplicada a Análise de variância (ANOVA) para investigar se existe diferença significativa entre as médias dos algoritmos utilizados. Na técnica da ANOVA, a variação total na mensuração dos resultados é subdividida em variações atribuídas a diferenças entre os grupos e variações dentro dos grupos. A variação dentro dos grupos é considerada erro experimental enquanto que a variação entre os grupos é atribuída a efeitos de tratamento.

Admitindo  $c$  populações normalmente distribuídas com variâncias iguais, a hipótese nula, mostrada em (5.12), de igualdade das médias das populações é testada contra a hipótese

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_c \quad (5.12)$$

alternativa de que existe pelo menos uma das médias  $\mu_j$  que possui valor diferente das demais, a hipótese alternativa  $H_1$ , está representada na equação (5.13)

$$H_1: \text{nem todas as } \mu_j \text{ são iguais } (j = 1, 2, \dots, c) \quad (5.13)$$

onde  $c$  representa o número de grupos que estão sendo comparados.

Uma medida de observação geral entre todas as observações é obtida através da soma dos quadrados das diferenças entre cada observação individual e a média geral. A variação total (STQ), que é a soma total dos quadrados, é calculada de acordo com a equação (5.14).

$$STQ = \sum_{j=1}^c \sum_{i=1}^{n_j} (X_{ij} - \bar{X})^2, \text{ onde } \bar{X} = \frac{\sum_{j=1}^c \sum_{i=1}^{n_j} X_{ij}}{n} \quad (5.14)$$

De maneira que:

$\bar{\bar{X}}$  é o estimador da média geral

$X_{ij}$  é a  $i$ -ésima observação no grupo  $j$ ;

$n_j$  é o número de observações no grupo  $j$ ;

$n$  é o número total de observações em todos os grupos;

$c$  é o número de grupos.

A variação entre os grupos, denominada de soma dos quadrados entre os grupos ( $SQE$ ) é calculada de acordo com a equação (5.15).

$$SQE = \sum_{j=1}^c n_j (\bar{X}_j - \bar{\bar{X}})^2 \quad (5.15)$$

onde  $\bar{X}_j$  é a média aritmética da amostra do grupo  $j$  e as demais variáveis são as mesmas da equação (5.14).

A variação dentro do grupo, denominada de soma dos quadrados dentro dos grupos ( $SQD$ ) é calculada de acordo com a equação (5.16).

$$SQD = \sum_{j=1}^c \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_j)^2 \quad (5.16)$$

Desde que  $c$  grupos estejam sendo comparados, existem  $(c-1)$  graus de liberdade associados à soma dos quadrados entre os grupos e como cada um dos grupos contribui com  $(n_j - 1)$  graus de liberdade, existem  $(n - c)$  graus de liberdade associados à soma dos quadrados dentro dos grupos, além disso, existem  $(n - 1)$  graus de liberdade associados à soma total dos quadrados, pois cada observação está sendo comparada com a média geral, baseada nas  $n$  observações.

Assim, ao dividir cada uma das somas dos quadrados pelos seus respectivos graus de liberdade, são obtidas três variâncias, que estão representadas na equação (5.17).

$$\begin{aligned} MQE &= \frac{SQE}{c-1} \\ MQD &= \frac{SQD}{n-c} \\ MTQ &= \frac{STQ}{n-1} \end{aligned} \quad (5.17)$$

Para testar a hipótese nula,  $H_0$ , em relação à hipótese alternativa  $H_1$ , a estatística do teste  $F$  é calculada como sendo a razão entre duas das variâncias,  $MQE$  e  $MQD$ , de acordo com a equação (5.18).

$$F = \frac{MQE}{MQD} \quad (5.18)$$

A estatística do teste  $F$ , segue uma distribuição  $F$ , com  $(c - 1)$  graus de liberdade que correspondem a  $MQE$  no numerador e  $(n - c)$  graus de liberdade correspondentes ao  $MQD$  no denominador.

Assim, se o nível de significância é  $\alpha$ , a hipótese nula será rejeitada se a estatística do teste  $F$  for maior do que o valor crítico  $F_s$ , tomado a partir da distribuição  $F$ , tendo  $(c - 1)$  graus de liberdade no numerador e  $(n - c)$  graus de liberdade no denominador. Desta forma, a regra de decisão é a seguinte:

$$\text{Se } F > F_s \Rightarrow \text{Rejeitar } H_0$$

Caso a hipótese nula do teste  $F$ , seja verdadeira, espera-se que a estatística do teste  $F$  calculada possua um valor próximo de 1, se a hipótese nula é rejeitada, isto é, se existirem diferenças significativas entre as médias, espera-se que a estatística calculada  $F$ , tenha um valor substancialmente maior do que 1. Assim, a metodologia ANOVA, fornece um teste  $F$ , no qual a hipótese nula é rejeitada para um dado nível de significância,  $\alpha$ , somente se a estatística calculada  $F$  for maior do que o valor crítico  $F_s$ , que possui  $(c - 1)$  e  $(n - c)$  graus de liberdade.

### 5.3.2 Teste Tukey-Kramer

Se existir diferença significativa entre os métodos analisados, deve-se verificar qual dos grupos apresenta o melhor desempenho. Um dos procedimentos que se pode utilizar é o teste Tukey-Kramer (LEVINE, et al., 2002), que possibilita examinar, simultaneamente, comparações entre todos os pares de grupos.

O primeiro passo do procedimento é calcular as diferenças das médias,  $|\bar{X}_j - \bar{X}_{j'}|$ , entre os  $(c(c-1)/2)$  pares de grupos. Em seguida calcula-se o intervalo crítico para o teste Tukey-Kramer de acordo com a equação (5.19)

$$IC = Q_s \sqrt{\frac{MQD}{2} \left( \frac{1}{n_j} + \frac{1}{n_{j'}} \right)} \quad (5.19)$$

onde  $Q_s$  é o valor crítico obtido da distribuição de intervalos de *Student*, com  $c$  graus de liberdade no numerador e  $(n - c)$  graus de liberdade no denominador. A análise final é feita comparando cada diferença entre as médias aritméticas dos algoritmos com o valor do intervalo crítico, caso alguma das diferenças entre as médias,  $|\bar{X}_j - \bar{X}_{j'}|$ , supere o valor do intervalo crítico, conclui-se que existe diferença significativa entre os grupos.

### 5.3.3 Pressupostos para se utilizar a ANOVA

Ao se aplicar a ANOVA as premissas em relação aos dados que estão sendo investigados devem ser assumidas. Essas premissas são: independência, normalidade e homogeneidade das variâncias (homoscedasticidade). No caso da normalidade, o teste F ANOVA é relativamente robusto em relação a distanciamientos da distribuição normal, ou seja, o nível de significância do teste ANOVA em geral não é afetado pela ausência da normalidade, especialmente quando a amostra é grande. A premissa de homoscedasticidade afirma que as variâncias são iguais nas populações. Se as amostras possuem tamanhos iguais em cada um dos grupos, as inferências baseadas na distribuição F não são seriamente afetadas em função de variâncias desiguais (LEVINE, et al., 2002).

Neste capítulo foram apresentados os testes estatísticos utilizados para a comparação do desempenho dos métodos apresentados nos experimentos dos capítulos VI e VII desta tese.

## Capítulo 6

### EXPERIMENTOS E SIMULAÇÕES

Neste capítulo são relatados os experimentos realizados com a utilização da metodologia de *Boosting* proposta nesta tese. Num primeiro experimento é feita a comparação dos resultados obtidos através da técnica da PG tradicional e dos algoritmos de *Boosting*: *GPBoost*, *GPBoost* e *AdaboostRT*. Em seguida são apresentados os experimentos utilizando o algoritmo BCI e os resultados são comparados aos resultados obtidos através da metodologia ARMA e da PG tradicional e do *GPBoost*. O método *BCI* utiliza como algoritmo base, em todos os experimentos relatados neste capítulo, o algoritmo da PG, desta forma, o método é aqui denominado de *BCIGP*.

#### 6.1 EXPERIMENTO UTILIZANDO SÉRIES TEMPORAIS REAIS

No primeiro experimento, são realizados testes com algumas séries temporais reais, as quais foram retiradas do livro: *Análise de Séries Temporais* (MORETTIN; TOLOI, 2004), e algumas séries financeiras obtidas em (ECONOMATICA, 2006). Estas séries foram base dos estudos realizados com diversas medidas estatísticas de forma a se chegar ao algoritmo proposto no capítulo IV. Com base nas análises dos resultados obtidos com este experimento, foram estabelecidos os parâmetros da PG utilizados no banco de dados que contém as séries

simuladas. As descrições destas séries estão apresentadas no Apêndice B desta tese. Os resultados destes experimentos foram publicados em (SOUZA; COSTA; POZO, 2005).

### 6.1.1 Configuração da Programação Genética

A Programação Genética para Previsão de Séries Temporais foi implementada utilizando a ferramenta Lil-gp1.1 (ZONGKER; PUNCH, 1995). Inicialmente foram realizados experimentos com número variável de execuções da PG, nas quais era alterada somente a semente inicial<sup>11</sup>. O algoritmo utiliza até os últimos quatro valores da série temporal para gerar os próximos valores ( $Z_{t-1}$ ,  $Z_{t-2}$ ,  $Z_{t-3}$ ,  $Z_{t-4}$ ), as funções utilizadas no conjunto  $F$  (sen, cos, exp, raiz) possibilitam que modelos não lineares sejam ajustados aos dados. Estes modelos, em geral, se ajustam melhor aos dados, pois normalmente as séries analisadas não são oriundas de modelos lineares. O tamanho da população, o número de gerações e as taxas de reprodução, mutação e cruzamento foram obtidos experimentalmente, assim como o método de inicialização da população inicial, o método de seleção e a função de aptidão. A configuração que foi utilizada para a PG está no quadro (6.1).

**QUADRO 6.1 – PARÂMETROS DE CONFIGURAÇÃO DA PG**

$T = \{Z_{t-1}, Z_{t-2}, Z_{t-3}, Z_{t-4}\}$ $F = \{+, -, *, /, \log, \text{sen}, \text{cos}, \text{exp}, \text{raiz}\}$ População: 1000 Gerações: 200 Taxa de Cruzamento: 80% Taxa de Mutação: 10% Taxa de Reprodução: 10% Método de Inicialização da População Inicial: <i>Grow</i> Método de Seleção: <i>Best</i> Função de aptidão: <i>RMSE</i>
--

É importante ressaltar que o modelo é gerado utilizando as observações do conjunto de treinamento, este modelo é então aplicado para efetuar a previsão no conjunto de testes e as medidas de erro apresentadas são sempre calculadas no conjunto de teste.

<sup>11</sup> Semente Inicial – valor utilizado pelo algoritmo para replicar os experimentos.

No quadro (6.2) é apresentado um exemplo de um indivíduo gerado pela PG.

**QUADRO 6.2 – MELHOR INDIVÍDUO GERADO PELA PG**

```

=== BEST-OF-RUN ===
generation: 41
nodes: 40
depth: 7
TOP INDIVIDUAL:
raw aptidão: 17050
standardized aptidão: 17050
adjusted aptidão: 0.3697
TREE:
(* -0.09724(+ (- (* (rlog (+ Zt-3 Zt-1))(+ Zt-4 Zt-4))(+ Zt-1 Zt-1))(+ (* (exp (- 0.09934 Zt-
1))(* -0.09724(rlog Zt-1)))(+ (* (exp (- 0.88495 Zt-1))(* -0.09724(rlog Zt-1)))(- 0.17202(+
Zt-1 Zt-1))))))

```

### 6.1.2 Configuração do algoritmo *GPBoost*

Após estudo e implementação do *GPBoost*, foram realizados experimentos no sentido de determinar o número ótimo de algoritmos de *Boosting* a ser utilizado, de forma a obter resultados satisfatórios sem provocar aumento excessivo no custo computacional. Foram realizados testes com números de 5, 8, 10, 12, 15, 18 e 20 algoritmos de *Boosting*. Os resultados obtidos com estes experimentos mostraram que utilizando 10 execuções do algoritmo de *Boosting*, obtém-se melhorias nas previsões com redução do RMSE que chegam a 20% quando comparado aos demais números de *Boosting*. Foram testadas também as três formas de se calcular a perda obtida pelo preditor: exponencial, quadrática e linear, os melhores resultados foram obtidos com a utilização da fórmula exponencial conforme algoritmo apresentado no quadro (4.3).



### 6.1.3 Configuração da metodologia Box & Jenkins

A implementação da metodologia Box & Jenkins foi feita utilizando o *software* estatístico livre<sup>12</sup> R, que seleciona o melhor modelo ARMA(p, q) cuja ordem dos modelos varia de 0 a 4. O critério para seleção e ajuste do modelo utilizado, foi o critério *Akaike (AIC)*, o detalhamento deste critério encontra-se no Apêndice A desta tese. Para a verificação da adequabilidade do modelo, neste primeiro experimento, foi utilizado o *RMSE* calculado sobre o conjunto de treinamento. O modelo selecionado foi aquele que apresentou o menor *RMSE*, este modelo foi então utilizado para efetuar a previsão dos valores futuros das séries no conjunto de teste.

## 6.2 DESCRIÇÃO DOS EXPERIMENTOS

Nesta seção são apresentados os resultados obtidos utilizando a metodologia proposta na seção (6.1) para algumas séries reais e financeiras. Os valores utilizados são os valores reais das séries, nenhum procedimento de tratamento inicial dos dados foi aplicado.

### 6.2.1 Dados de Entrada

As séries utilizadas neste experimento estão apresentadas nas tabelas (6.1) e (6.2). De cada uma das séries foi escolhido um conjunto de treinamento contendo 90% dos dados observados e um conjunto de teste contendo 10% dos dados. Outros testes foram realizados, utilizando 70% dos dados para o conjunto de treinamento e 30% para o teste, porém os resultados obtidos utilizando 90% para treinamento apresentaram menor *RMSE*. A previsão

---

<sup>12</sup> <http://www.r-project.org/>

foi realizada  $n$  passos à frente, ou seja, o modelo foi gerado no conjunto de treinamento e então utilizado para a previsão dos valores futuros da série.

**TABELA 6.1 – BASES DE DADOS**

Série	Conjunto Total (100%)	Treinamento (90%)	Teste (10%)
Atmosfera	365	329	36
Bebida	187	169	18
Consumo	154	139	15
Fortaleza	149	135	14
ICV	126	114	12
IPI	187	169	18
Lavras	384	346	38
Manchas	176	159	17

**FONTE: SÉRIES TEMPORAIS – (MORETTIN & TOLOI, 2004)**

**TABELA 6.2 – BASES DE DADOS SÉRIES FINANCEIRAS**

Série	Conjunto Total (100%)	Treinamento (90%)	Teste (10%)
Djiad	1100	990	110
Ibovespa	1100	990	110
Nasdaq	1100	990	110

**FONTE: [WWW.ECONOMATICA.COM](http://WWW.ECONOMATICA.COM)**

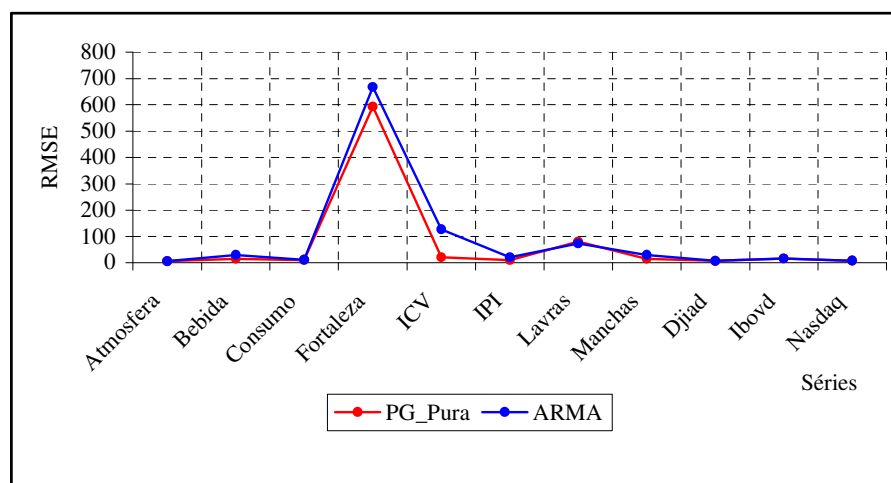
### 6.2.2 Experimento

Este experimento teve como objetivo comparar o desempenho dos algoritmos de *PG*, *GPBoost*, *Adaboost.RT* na previsão de Séries Temporais com a metodologia Box & Jenkins. Os testes foram realizados utilizando 10 sementes aleatórias iniciais para o algoritmo de PG. O critério de comparação utilizado foi o *RMSE* calculado no conjunto de testes, foi considerado o melhor método, aquele que apresentou o menor *RMSE* no conjunto de testes de cada uma das séries analisadas. O *RMSE* apresentado na tabela (6.3), foi calculado como sendo a média dos *RMSEs* de dez algoritmos de PG utilizados.

TABELA 6.3 - COMPARAÇÃO DAS PREVISÕES ARMA, PG E *GPBoost*

Série	PG Tradicional	<i>GPBoost</i>	<i>Adaboost.RT</i>	ARMA		PG x ARMA
Atmosfera	5,938722	5,844508	<b>5,741550</b>	6,244017	ARMA(2,1)	4,89%
Bebida	14,794792	<b>14,716823</b>	14,861233	29,747780	ARMA(2,2)	50,27%
Consumo	<b>10,323963</b>	11,250817	10,856045	11,758310	ARMA(4,3)	12,20%
Fortaleza	<b>592,983924</b>	656,548041	656,819081	667,600500	ARMA(1,0)	11,18%
ICV	20,247427	19,796984	<b>19,779801</b>	126,431500	ARMA(1,0)	83,99%
IPI	<b>9,875241</b>	10,578101	10,518967	20,449340	ARMA(3,2)	51,71%
Lavras	<b>80,101648</b>	82,982135	83,921154	<b>73,047860</b>	ARMA(3,2)	-9,66%
Manchas	<b>14,697371</b>	17,806627	18,081780	29,776310	ARMA(2,1)	50,64%
Djiad	0,007267	<b>0,007235</b>	0,007722	0,007292	ARMA(2,2)	0,35%
Ibovd	0,016115	0,016097	<b>0,016097</b>	0,016375	ARMA(2,3)	1,59%
Nasdaq	<b>0,007895</b>	0,007875	0,008490	0,007964	ARMA(2,2)	0,86%

Na última coluna da tabela são apresentados os resultados comparativos entre a previsão obtida pela PG e pelos modelos ARMA. Observa-se, que na maioria dos casos, a PG obtém o menor *RMSE*, com melhorias que chegam a 84% no caso da série ICV. Para as demais séries, a PG também apresentou bons resultados com melhoria de 51,7% para a série IPI e 50% para a série Manchas. Apenas no caso da série Lavras, o modelo ARMA apresentou resultado melhor em relação à PG. Para as séries Djiad e Ibovd, o *Boosting* apresentou uma pequena melhoria no *RMSE* quando comparado à PG.

FIGURA 6.1 COMPARAÇÃO DO *RMSE* DA PG E ARMA

O fato do algoritmo *Boosting* não ter apresentado melhorias significativas, apenas em torno de 1%, em relação aos resultados obtidos através da PG, foi motivação para buscar uma forma de utilização do algoritmo de *Boosting*, que pudesse fornecer melhores resultados.

### 6.2.3 Experimento II

As configurações utilizadas para este experimento são as mesmas do experimento anterior. Os valores apresentados na tabela (6.4) representam a média dos *RMSE* de 10 PGs obtidas com sementes iniciais diferentes.

Neste experimento o algoritmo base utilizado foi a PG, assim o método que utiliza os coeficientes de correlação para a combinação final dos preditores foi denominada de *BCIGP* (*Boosting Correlation Improvement using Genetic Programming*). Nas colunas de 2 a 6 da tabela (6.4) são apresentados os valores médios de *RMSE*, analisando estes valores obtidos com as séries reais, somente em dois casos, o algoritmo *BCIGP* obteve um resultado pior do que os outros dois algoritmos de *Boosting* analisados e embora não tenha se comportado bem no caso da série ICV, na qual obteve um *RMSE* bem superior ao encontrado pelos demais métodos, as melhorias obtidas para as outras séries mostraram a eficiência do método. Os valores de *RMSE* apresentados são calculados fora da amostra em todas as séries, ou seja, no conjunto de teste. Nas colunas 7 e 8 da tabela é feita uma comparação do desempenho do algoritmo *BCIGP* com a PG e com os modelos ARMA, os valores representam em percentual, a melhoria obtida pelo método *BCIGP*, somente para a série ICV, a PG obteve uma melhoria de 63% em relação ao *BCIGP* e 0,12% na série *Nasdaq*. Comparando com os modelos ARMA, o método *BCIGP* mostrou-se mais eficiente.

**TABELA 6.4 - COMPARAÇÃO DO *RMSE* NO CONJUNTO DE TESTE PARA CADA MÉTODO**

Série	PG Tradicional	GPBoost	Adaboost_RT	BCIGP	ARMA	PG x BCIGP	ARMA x BCIGP
Atmosfera	5,938722	5,844508	5,741550	<b>2,467783</b>	6,244017	58,45%	153,02%
Bebida	14,794792	14,716823	14,861233	<b>6,613493</b>	29,747780	55,30%	349,80%
Consumo	10,323963	11,250817	10,856045	<b>7,896470</b>	11,758310	23,51%	48,91%
Fortaleza	592,983924	656,548041	656,819081	<b>427,114760</b>	667,600500	27,97%	56,30%
ICV	20,247427	19,796984	<b>19,779801</b>	33,058154	126,431500	-63,27%	282,45%
IPI	9,875241	10,578101	10,518967	<b>3,554103</b>	20,449340	64,01%	475,37%
Lavras	80,101648	82,982135	83,921154	<b>53,661973</b>	73,047860	33,01%	36,13%
Manchas	14,697371	17,806627	18,081780	<b>14,571732</b>	29,776310	0,85%	104,34%
Djiad	7,266819	7,235000	7,722000	<b>7,206467</b>	7,292027	0,83%	1,19%
Ibovd	16,115177	16,097000	16,097000	<b>15,200588</b>	16,375060	5,68%	7,73%
Nasdaq	7,895456	<b>7,875000</b>	8,490000	7,904953	7,964157	-0,12%	0,75%

Na figura (6.2) são apresentados, graficamente, o *RMSE* obtido por todos métodos estudados para as séries analisadas. Observa-se que o *RMSE* obtido pelo método *BCIGP* está

quase sempre abaixo das linhas do  $RMSE$  obtido na previsão feita através da PG tradicional e dos modelos ARMA.

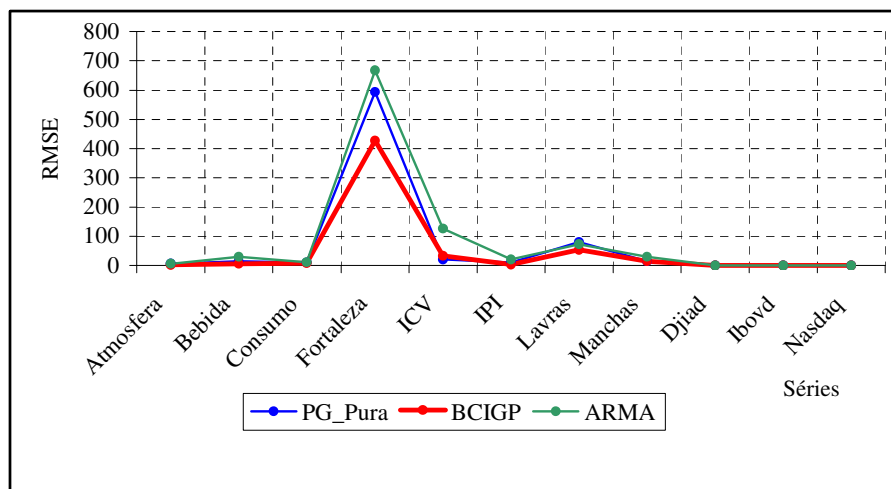


FIGURA 6.2– COMPARAÇÃO DO  $RMSE$  DA PG,  $BCIGP$  E ARMA

Na figura (6.3) são apresentados os valores de previsão obtidos pela PG e pelo  $BCIGP$  para a série IPI. No gráfico pode-se perceber que os valores obtidos pelo método proposto estão mais próximos dos valores reais observados do que os valores obtidos pela PG.

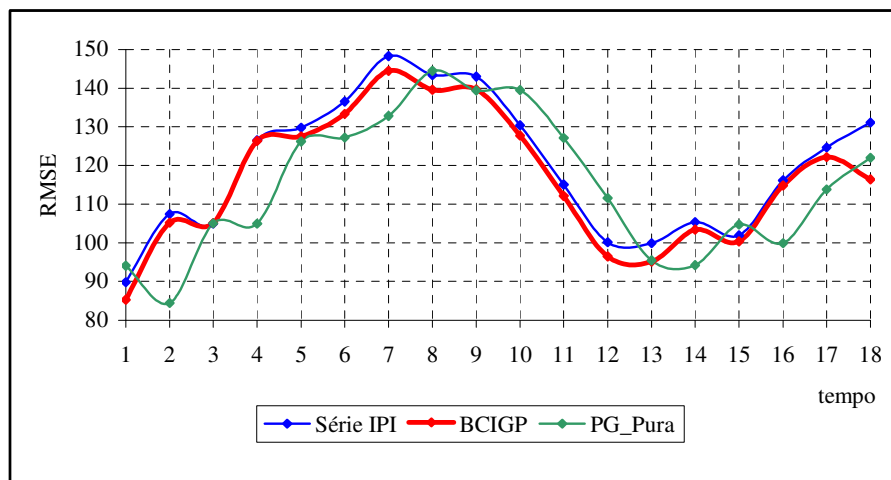


FIGURA 6.3– VALORES DE PREVISÃO OBTIDOS PELA PG E  $BCIGP$  PARA A SÉRIE IPI

### 6.2.4 Análise do desempenho do algoritmo BCIGP utilizando o teste “t”

A comparação é feita utilizando a mesma base de dados. Os resultados são calculados entre os resultados obtidos das 10 diferentes sementes iniciais para cada execução do algoritmo da PG. A hipótese nula é a de que a diferença média entre os MSE dos algoritmos é zero, contra a alternativa de que este valor é diferente de zero. Foi estabelecido o nível de significância de 1% para a rejeição da hipótese nula. Nas colunas 2, 3 e 4 da tabela (6.5) estão os valores-p das comparações entre os algoritmos. Observa-se que o valor-p é na maioria dos casos inferior ao nível de significância fixado. Na comparação com a PG (coluna 2), somente em dois casos, para as séries Manchas e Djiad, o valor-p foi superior a 0,01; na comparação com o *GPBoost*, o valor-p foi superior a 0,01 para as séries ICV e Manchas e na comparação com os modelos ARMA o valor-p foi superior a 0,01 para as séries Lavras e Djiad. Em todas as análises, o método BCIGP superou a PG e os modelos ARMA na maioria dos casos.

**TABELA 6.5 – VALOR – p - TESTE “t”**

Séries	PG x BCIGP	<i>GPBoost</i> x BCIGP	ARMA x BCIGP
Atmosfera	3,7905E-13	1,00883E-16	4,15965E-18
Bebida	4,7524E-10	3,04027E-11	4,88797E-16
Consumo	0,00914777	1,64447E-12	2,59222E-15
Fortaleza	9,9558E-06	7,38805E-10	1,11589E-12
ICV	0,01226633	0,012144333	8,21525E-13
IPI	4,3907E-13	3,35132E-16	2,60942E-26
Lavras	2,9221E-07	2,14086E-07	0,011376221
Manchas	0,61925777	0,358819112	2,55251E-10
Djiad	0,06501221	0,005344011	0,085420922
Ibovespa	1,8239E-09	1,77275E-09	3,35255E-11
Nasdaq	0,00164798	2,15808E-05	0,001647981

### 6.3 INTERPRETAÇÃO E ANÁLISE DAS SÉRIES FINANCEIRAS – UMA APLICAÇÃO REAL

A análise das séries financeiras é considerada um desafio para os pesquisadores, no que se refere à construção de um modelo que se ajuste bem aos dados, gerando previsões que

produzam ganhos financeiros. Em geral as séries financeiras, principalmente as obtidas pelos retornos de ativos financeiros, são processos estocásticos cuja estrutura de dependência é complexa, com alto grau de não linearidade e os métodos existentes, geralmente lineares, acabam por não obter boas previsões, especialmente quando o horizonte de previsão é grande.

### 6.3.1 Descrição das séries financeiras utilizadas

A série Ibovd representa os índices da BOVESPA – Bolsa de Valores de São Paulo. Foi fundada em 1890. Até meados de 1960, a BOVESPA e as demais bolsas brasileiras eram entidades oficiais corporativas, vinculadas às secretarias de finanças dos governos estaduais e compostas por corretores nomeados pelo poder público. A partir de 1965, as bolsas transformaram-se em associações civis sem fins lucrativos, com autonomia administrativa, financeira e patrimonial. Foi a primeira bolsa brasileira a implantar o pregão automatizado com informações disponíveis em tempo real. Em 1990, foram iniciadas as negociações através do sistema de negociação eletrônica - CATS (*Computer Assisted Trading System*) que operava simultaneamente com o sistema tradicional de Pregão Viva Voz. O sistema eletrônico foi implantado em 1997 e em 1999 foram criados o *Home Broker* e o *After Market*, ambos criados para tornar possível a participação do pequeno e médio investidor, através da *internet*. Atualmente, a BOVESPA é o maior centro de negociação com ações da América Latina.

A *National Association of Security Dealers Automated Quotation System (NASDAQ)* foi criada nos Estados Unidos, em 1971 e foi a primeira bolsa do mundo a negociar exclusivamente ações de empresas de Internet, informática e tecnologia. Como as operações de compra e venda são feitas por meio de computadores no mercado de balcão, a *Nasdaq* ficou conhecida como a "bolsa eletrônica". A *Nasdaq* é conhecida por negociar ações das maiores empresas de tecnologia, como por exemplo a Microsoft, Intel, *Dell Computer*, *Yahoo*, *Amazon.com*, etc. Recentemente, a *Nasdaq* uniu-se à *American Stock Exchange (AMEX)*, formando o *Nasdaq-Amex Market Group*.

A série *Djiad* indica os índices da *Dow Jones*, fundada em 1882 com a finalidade de divulgar cotações das ações e notícias econômicas do mercado de *New York*. Em 1896 foi criada uma média específica para ações emitidas por empresas industriais, *Dow-Jones*

*Industrial Average (DJIA)*, porém as divisões que foram ocorrendo fizeram com que fosse criado um multiplicador para anular o efeito das variações bruscas de preços.

### 6.3.2 Retornos

Em geral, na previsão de séries financeiras, é preferível trabalhar com retornos do que com preços, pois estes são livres de escalas e são estacionários. Denotando por  $P_t$ , o preço de um ativo<sup>13</sup> no instante  $t$  (normalmente um instante é considerado um dia de negociação), a variação de preços entre os instantes  $(t-1)$  e  $t$  é dada pela equação (6.8):

$$\Delta P_t = P_t - P_{t-1} \quad (6.8)$$

e a variação relativa de preços, ou retorno líquido simples deste ativo, entre os mesmos instantes é definido como na equação (6.9):

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{\Delta P_t}{P_{t-1}} \quad (6.9)$$

Assim, na equação (6.10), tem-se que:

$$R_t + 1 = \frac{P_t}{P_{t-1}} \quad (6.9)$$

$R_t + 1$  é denominado de retorno bruto simples. Usualmente,  $R_t$  é expresso em percentual, relativo ao período (dia, mês ou ano), e é também chamado de taxa de retorno.

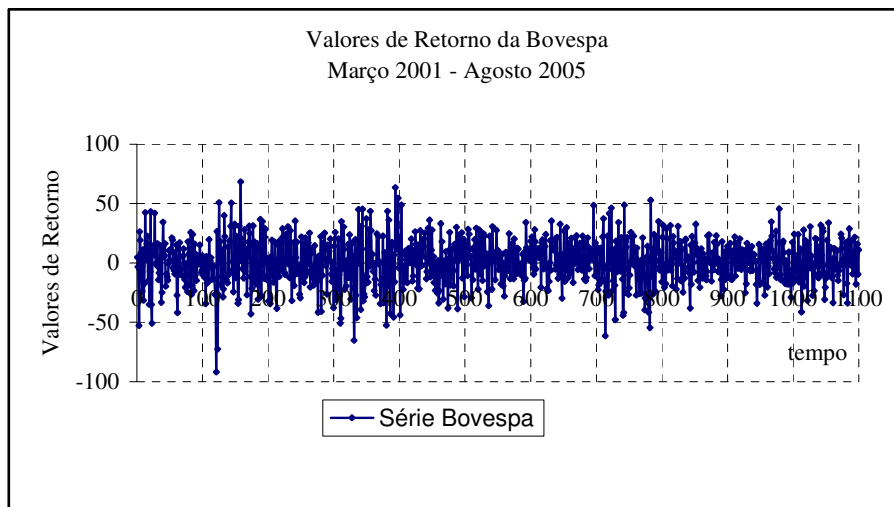
Os retornos financeiros apresentam algumas peculiaridades, que muitas séries não apresentam. Retornos raramente apresentam tendências ou sazonalidades (MORETTIN; TOLOI, 2004).

Na figura (6.4) está representada a série de retornos diários do IBOVESPA (Índice da Bolsa de Valores de São Paulo), no período de 3 de janeiro de 1995 a 27 de dezembro de 2000. Observa-se que a série é estacionária com média ao redor de zero e agrupamentos de volatilidade. Os períodos que apresentam alta volatilidade foram períodos em que ocorreram crises em diversos países, inclusive no Brasil e que tiveram influência sobre o mercado financeiro brasileiro, entre as quais pode-se citar a crise do México nos meses de fevereiro e

<sup>13</sup> Todo e qualquer título representativo de parte patrimonial ou dívida.



março de 1995, a crise na Ásia, em outubro de 1997, moratória na Rússia em agosto de 1998, a desvalorização do Real em janeiro de 1999 e a queda da bolsa Nasdaq, em abril do ano 2000.



**FIGURA 6.4– RETORNOS DIÁRIOS DA IBOVESPA (01/03/1995 – 27/08/2000)**

O principal objetivo da previsão das séries financeiras é obter uma estratégia de ação baseada na qual o investidor possa tomar uma posição no mercado financeiro, se irá comprar, vender ou manter a posse de suas ações. Em geral, essas previsões são diárias, ou seja, o horizonte de previsão considerado é um.

### 6.3.3 Avaliação da estratégia de ação baseada nas previsões

Após ter sido feita a previsão de retorno das séries financeiras Ibovd, Nasdaq e Djiad, foi aplicado também ao problema, uma estratégia de *trading* (negociação das ações) baseada somente no sinal da previsão e não levando em conta o custo da operação, adaptando a metodologia proposta em Dunis e Jalilov (2002). Se a previsão é positiva, uma estratégia de compra é realizada, caso negativa, o investidor vende as suas ações. Se o sinal confirmar a estratégia adotada no dia anterior, o investidor mantém a sua posição, isto implica em permanecer “comprado”, com as ações na mão, ou “vendido”, sem a posse das ações. Em ambos os casos, é importante ressaltar que ao manter a sua posição, o lucro naquele instante é

computado como zero. Os valores de retorno são obtidos, considerando o preço médio das ações em dois dias consecutivos, conforme a equação (6.10)

$$r_t = \frac{p_t}{p_{t-1}} \quad (6.10)$$

onde  $p_t$  é o valor da ação no instante  $t$  e  $p_{t-1}$  é o valor da ação no instante  $t-1$ .

Na tabela (6.6) são apresentados valores comparativos dos lucros obtidos utilizando as estratégias de ação fornecidas por previsões através da utilização dos modelos ARMA e *BCIGP*. O horizonte de previsão utilizado para ambos os métodos foi de 110 dias, o que representa 10% do conjunto dos valores observados.

Na tabela (6.6) são apresentados valores comparativos dos lucros obtidos utilizando as estratégias de ação fornecidas por previsões através da utilização dos modelos ARMA e *BCIGP*. O horizonte de previsão utilizado para ambos os métodos foi de 110 dias, o que representa 10% do conjunto dos valores observados. Os dados observados são relativos ao período de 13/08/2001 a 17/08/2005. Na tabela (6.7) estão apresentados os lucros anualizados<sup>14</sup>, ou seja, projeta-se o lucro com a utilização da estratégia para o período de um ano comercial.

**TABELA 6.6 – RETORNO FINANCEIRO NO PERÍODO 110 DIAS**

Método	<i>BCIGP</i>	ARMA	PG	<i>GPBoost</i>
Djiad	<b>2,02%</b>	0,45%	-1,02%	-3,48%
Ibovd	<b>16,88%</b>	-1,69%	-8,67%	-6,85%
Nasdaq	<b>7,29%</b>	-5,87%	-8,82%	-3,69%

**TABELA 6.7 – RETORNO FINANCEIRO ANUALIZADO**

Método	<i>BCIGP</i>	ARMA	PG	<i>GPBoost</i>
Djiad	<b>4,62%</b>	1,03%	-2,34%	-7,98%
Ibovd	<b>76,64%</b>	-3,88%	-19,85%	-15,69%
Nasdaq	<b>16,69%</b>	-13,46%	-20,21%	-8,45%

Observa-se que em todas as situações, os resultados obtidos com a utilização do método *BCIGP* foram superiores aos obtidos pela aplicação das demais metodologias avaliadas. Em todos os períodos os resultados foram favoráveis ao investidor, ou seja, a estratégia de ação gerou um retorno positivo em todos os casos. O fato do horizonte de

<sup>14</sup>Retorno anualizado é obtido multiplicando-se a média do período por 252 – número de dias de aplicação no período de um ano.

previsão ser grande, deve ter contribuído para os resultados superiores da utilização da metodologia *BCIGP*, já que para horizontes de previsão grandes.

Neste caso, não foram consideradas as despesas com as operações financeiras, porém há um custo operacional para cada transação realizada. Assim, foram contadas também, o número de transações realizadas para cada método no período de 110 dias, ou seja, quantas vezes houve compra ou venda de ações. Os valores estão mostrados na tabela (6.8).

**TABELA 6.8 – NÚMERO DE TRANSAÇÕES NO PERÍODO**

Método	<i>BCIGP</i>	ARMA	PG	B1
Djiad	43	84	54	43
Ibovd	47	53	53	49
Nasdaq	60	60	54	50

Observa-se que, em quase todos os casos o método *BCIGP* realizou um número menor de operações, se for levado em consideração o custo das operações financeiras, o lucro pode ser considerado ainda maior quando se adota as estratégias obtidas através desta metodologia.

# Capítulo 7

## SIMULAÇÃO MONTE CARLO

Neste capítulo são apresentados os resultados obtidos com a aplicação do método *BCIGP* para séries geradas através de uma simulação Monte Carlo. O objetivo desta simulação é mostrar que o método proposto apresenta um bom desempenho para a previsão de séries temporais, independente do domínio ao qual elas pertençam. Os resultados obtidos são comparados aos da PG tradicional, do *GPBoost* e dos modelos ARMA para Previsão de Séries Temporais. O detalhamento desta simulação é apresentado, bem como os resultados obtidos, as comparações com os métodos estatísticos e as conclusões.

### 7.1 DESCRIÇÃO DA SIMULAÇÃO

As etapas utilizadas para a realização da simulação estão descritas a seguir:

- Definição dos parâmetros a serem utilizados em cada uma das cinco estruturas;
- Geração das séries artificiais para cada parâmetro definido em cada uma das estruturas;
- Ajuste do modelo ARMA;

- Previsão utilizando a metodologia ARMA, PG, *GPBoost*, *AdaBoost.RT* e *BCIGP*.

O fluxograma contendo as etapas de realização da simulação é apresentado na figura (7.1).

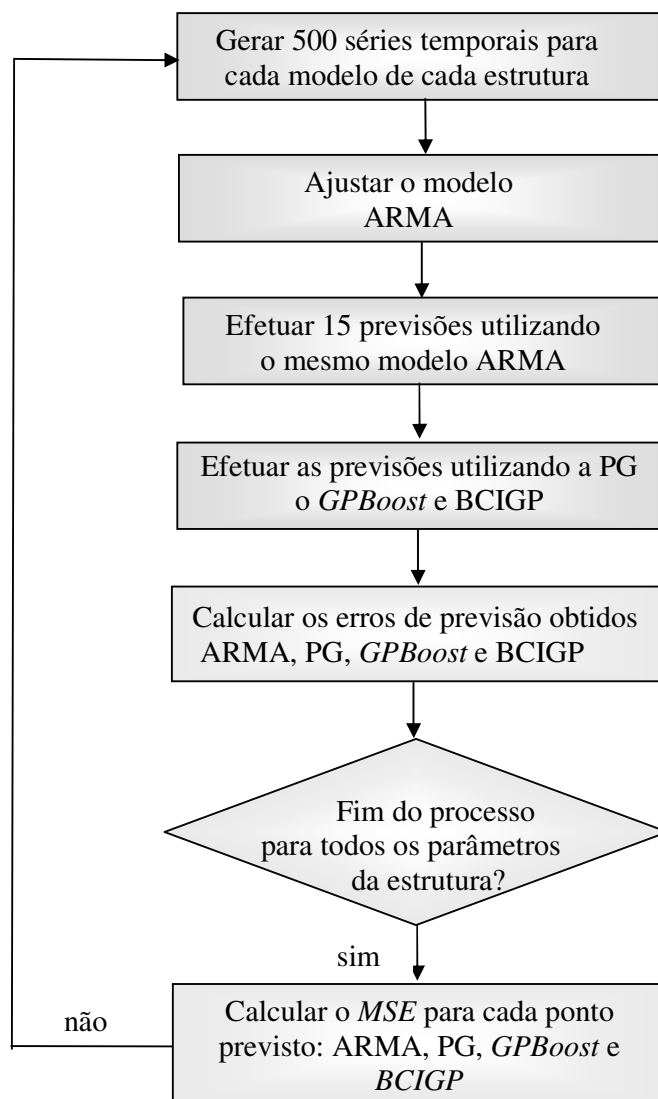


FIGURA 7.1 – FLUXOGRAMA DA SIMULAÇÃO MONTE CARLO

### 7.1.1 Definição dos Parâmetros

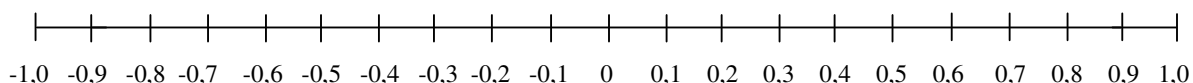
Para definir os parâmetros utilizados nos modelos da simulação das séries, tomou-se no espaço paramétrico a região de estacionariedade das principais estruturas dos modelos ARMA: AR(1), AR(2), MA(1), MA(2) e ARMA(1, 1).

### 7.1.2 Espaço Paramétrico das Estruturas AR(1) e MA(1)

Para que um modelo da estrutura AR(1) seja estacionário, é necessário que o parâmetro auto-regressivo,  $\phi_I$ , esteja compreendido na região do conjunto dos números reais  $\mathbb{R}$ , entre -1 e 1, ou seja, o parâmetro  $\phi_I$  deve satisfazer a inequação (7.1). De forma análoga, para que um modelo da estrutura MA(1) seja invertível é necessário que o parâmetro de médias móveis  $\theta_I$  satisfaça à equação (7.2) (BOX; JENKINS, 1970). O espaço paramétrico da região de estacionariedade da estrutura AR(1) e de invertibilidade da estrutura MA(1) estão representados na figura (7.2). Este espaço foi dividido utilizando passo de 0,1.

$$-1 < \phi_I < 1 \quad (7.1)$$

$$-1 < \theta_I < 1 \quad (7.2)$$



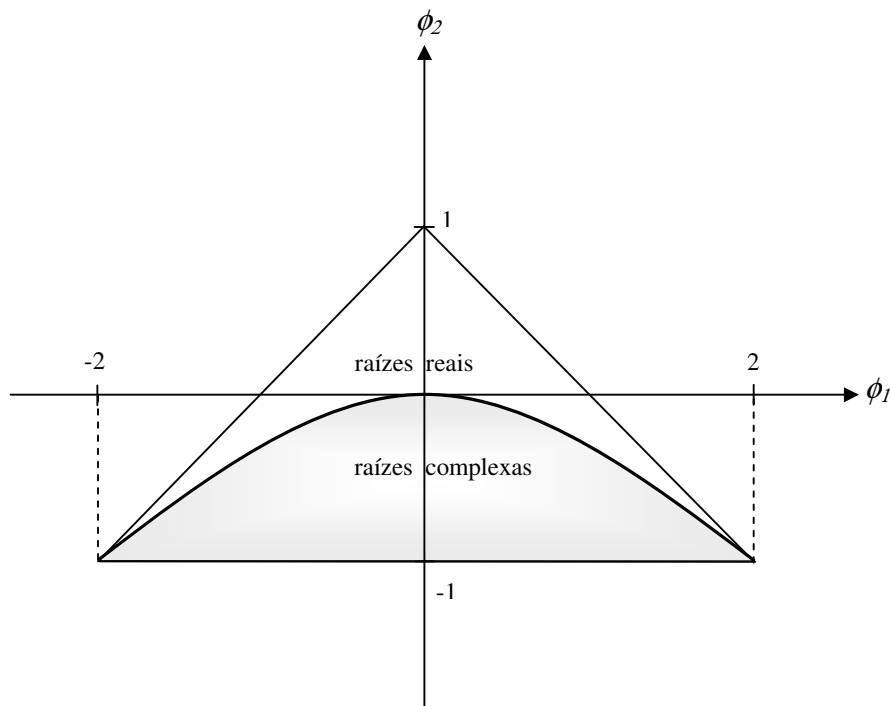
**FIGURA 7.2 – REGIÃO DE ESTACIONARIEDADE PARA OS MODELOS AR(1) E INVERTIBILIDADE PARA MA(1)**  
EIXO DOS PARÂMETROS  $\phi_I$  E  $\theta_I$

### 7.1.3 Espaço Paramétrico das Estruturas AR(2) e MA(2)

O espaço paramétrico da região de estacionariedade da estrutura AR(2) está representado na figura (7.3). Nestas estruturas, os parâmetros  $\phi_1$  e  $\phi_2$  devem satisfazer às inequações (7.3).

$$\begin{aligned}\phi_1 + \phi_2 &< 1 \\ \phi_2 - \phi_1 &< 1 \\ -1 &< \phi_2 < 1\end{aligned}\tag{7.3}$$

Para a estrutura MA(2) não há restrição sobre os parâmetros para que o processo seja estacionário (MORETTIN; TOLOI, 2004), porém sua região de invertibilidade é a mesma região de estacionariedade do processo AR(2) (Fig. 7.3).

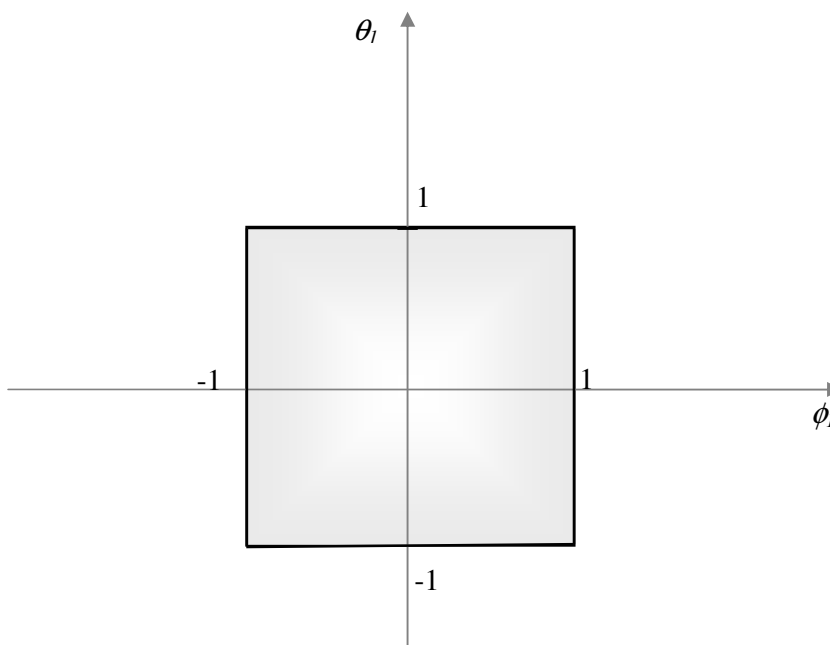


**FIGURA 7.3 – REGIÃO DE ESTACIONARIDADE PARA UM MODELO AR(2) E INVERTIBILIDADE PARA MA(2)**

#### 7.1.4 Espaço Paramétrico da Estrutura ARMA(1, 1)

O espaço paramétrico da região de estacionariedade e invertibilidade da estrutura ARMA(1,1) está representado na figura (7.4), os valores dos parâmetros devem satisfazer às equações (7.4). A região foi dividida utilizando passo de 0,2 nos eixos x e y (BOX; JENKINS, 1976).

$$\begin{aligned} \theta_1 + \phi_1 &< 1 \\ -1 &< \theta_1 < 1 \\ -1 &< \phi_1 < 1 \end{aligned} \quad (7.4)$$



**FIGURA 7.4 – REGIÃO DE ESTACIONARIDADE PARA UM MODELO ARMA(1,1)**

Utilizando este critério, o número de parâmetros definidos para cada estrutura está apresentado na tabela (7.1).



**TABELA 7.1 – NÚMERO DE PARÂMETROS E SÉRIES**

<b>Estrutura</b>	<b>Número de Parâmetros</b>
AR(1)	19
AR(2)	90
MA(1)	19
MA(2)	200
ARMA(1,1)	100

### 7.1.5 Simulação das Séries

Para gerar as séries sintéticas a serem utilizadas no experimento, após terem sido definidos os parâmetros de cada estrutura, foi feito um programa computacional utilizando o software R. Para cada parâmetro foram geradas 500 séries com 150 observações cada uma. Na simulação das séries, os 150 primeiros valores gerados são descartados para evitar vício na geração dos dados. As séries são geradas com ruído que possui distribuição normal com média zero e desvio padrão um.

### 7.1.6 Banco de Dados

O banco de dados gerado com os parâmetros descritos na seção anterior ficou constituído de 214.000 séries, distribuídas para cada estrutura, da forma mostrada na tabela (7.2).

**TABELA 7.2 – NÚMERO DE PARÂMETROS E SÉRIES**

<b>Estrutura</b>	<b>Número de parâmetros</b>	<b>Número de séries</b>
AR(1)	19	9.500
AR(2)	90	45.000
MA(1)	19	9.500
MA(2)	200	100.000
ARMA(1,1)	100	50.000

## 7.2 EXECUÇÃO DO EXPERIMENTO

Na etapa de execução do banco de dados, foram obtidos os resultados de Previsão das Séries Temporais, artificialmente geradas, utilizando os métodos da Programação Genética tradicional (PG-Pura); Programação Genética usando *GPBoost* (B1), modelos ARMA e *BCIGP*. Neste experimento, os coeficientes de correlação foram utilizados somente na combinação final dos preditores obtidos. Na seção (7.3) são relatados os resultados dos experimentos utilizando o algoritmo *BCIGP* completo, tanto na combinação final dos preditores quanto na atualização de pesos, o que foi denominado aqui de “Pesos 2”. Os resultados obtidos são comparados aos demais métodos e também ao método *BCIGP* utilizado somente na combinação final dos preditores.

### 7.2.1 Configuração da Programação Genética e do *GPBoost*

Neste experimento foi escolhido utilizar somente uma semente inicial para a execução da PG, pois considerando que o número de séries é grande, o tempo de execução se tornaria impraticável. Nos experimentos relatados no capítulo VI, verificou-se que a utilização de 10 execuções do algoritmo de *Boosting*, fornecem resultados satisfatórios, com menores erros de previsão no conjunto de teste. Portanto, esta mesma configuração foi utilizada na simulação. A função de perda utilizada é a função de perda exponencial apresentada na equação (4.11).

### 7.2.2 Ambiente Computacional

Para a execução do Banco de Dados, foi criado um ambiente computacional para o processamento do experimento. Foi utilizado um grupo contendo 42 computadores, dos quais 16 possuem processador duplo com 2 GHz e 2 GB de memória RAM e os outros 26 possuem processador simples com 1.7 GHz e 2 GB de memória RAM, pertencentes ao Cluster da Universidade Federal do Paraná (UFPR). O ambiente de execução utilizado foi o LINUX, e

os programas foram implementados utilizando a linguagem de programação C++. O banco de dados foi dividido em grupos de 500 séries, que eram as séries geradas para cada parâmetro utilizado, formando um total de 428 grupos. Para cada um destes grupos são executados os 4 algoritmos que são: PG, *GPBoost*, *Adaboost.RT* e *BCIGP*, para cada um destes algoritmos é gerado um arquivo que contém os valores previstos para cada série, para cada um dos métodos utilizados. O tempo de processamento para cada um destes 428 grupos ficou em torno de 32 horas.

### 7.2.3 Análise dos Resultados

Após a execução do procedimento para todas séries geradas, os resultados foram analisados calculando o erro médio quadrático, fora da amostra, ou seja, o modelo foi gerado utilizando as observações do conjunto de treinamento e em seguida foi feita a previsão para os valores do conjunto de teste. O cálculo do erro foi feito de acordo com as equações (7.5),

$$erro(h_i) = \begin{cases} \frac{1}{500} \sum_{i=1}^{500} (série(x_i) - ARMA(x_i))^2 \\ \frac{1}{500} \sum_{i=1}^{500} (série(x_i) - PG(x_i))^2 \\ \frac{1}{500} \sum_{i=1}^{500} (série(x_i) - GPBoost(x_i))^2 \\ \frac{1}{500} \sum_{i=1}^{500} (série(x_i) - BCIGP(x_i))^2 \end{cases} \quad (7.5)$$

onde:  $série(x_i)$  são os valores reais observados da série;  $ARMA(x_i)$  são os valores previstos pelos modelos ARMA;  $GPBoost(x_i)$  são os valores previstos utilizando o algoritmo GPBoost e  $BCIGP(x_i)$  são os valores obtidos pelo algoritmo BCIGP. Os valores são calculados no conjunto de teste.

Na tabela (7.3) estão os valores do *MSE* das 500 séries geradas para cada parâmetro dentro da estrutura AR(1). Cada uma das colunas  $P_1, P_2, \dots, P_j$ , representam os parâmetros considerados para esta estrutura, partindo da esquerda para a direita, com passo de 0,1 e os valores *e135, e136, ..., e150* representam os horizontes de previsão analisados.

TABELA 7.3 – MSE MÉDIO DAS 500 SÉRIES PARA A ESTRUTURA AR(1)

MSE	AR(1)	P1	P2	P3	P4	P5	P8	P7	P8	P9	P15	P18	P19	Média
e136	ARMA	<b>0,9522</b>	0,4589	<b>0,1025</b>	<b>0,0660</b>	0,6113	<b>0,6602</b>	1,0153	<b>0,9761</b>	0,5312	5,8741	4,1962	3,8506	1,6079
	PG-Pura	1,0233	0,4567	0,1154	0,0791	0,6459	0,7537	1,1055	1,1393	0,5829	2,2558	1,7052	3,2271	1,0908
	B1	0,9663	0,4656	0,1130	0,0857	0,6881	0,7713	1,0665	1,1393	0,5833	2,2965	1,7177	3,2726	1,0972
	BCIGP	1,1425	<b>0,4441</b>	0,1046	0,0802	<b>0,5840</b>	0,7048	<b>0,9590</b>	1,0254	<b>0,5280</b>	<b>2,0544</b>	<b>1,6020</b>	<b>2,9453</b>	<b>1,0145</b>
e137	ARMA	1,8065	0,8463	0,1738	0,0676	0,7828	0,9061	1,0265	<b>0,8788</b>	0,5062	4,1043	3,4076	3,7389	1,5204
	PG-Pura	1,2669	0,5503	0,1480	0,0474	0,6524	12,5038	1,0222	0,9915	0,6065	0,8255	1,1275	3,1441	1,9072
	B1	<b>1,0413</b>	0,5367	0,1512	0,0525	0,6380	<b>0,8541</b>	0,9726	0,9915	0,5634	0,7436	1,1990	3,2176	0,9135
	BCIGP	1,0684	<b>0,5113</b>	<b>0,1314</b>	<b>0,0459</b>	<b>0,5576</b>	0,9410	<b>0,8828</b>	0,8923	<b>0,4994</b>	<b>0,7323</b>	<b>1,1199</b>	<b>2,8960</b>	<b>0,8565</b>
e138	ARMA	2,4164	1,0428	0,2113	0,1128	0,7123	1,0325	1,1560	<b>0,9652</b>	<b>0,5498</b>	5,2608	2,7322	2,5220	1,5595
	PG-Pura	1,1425	0,5486	0,1404	0,0968	0,6123	0,9478	1,1935	1,0948	0,6255	2,9702	1,5241	2,5000	1,1164
	B1	1,0667	0,5618	0,1541	0,0905	0,6403	<b>0,9027</b>	1,0981	1,0948	0,6332	2,9462	1,4847	2,5194	1,0994
	BCIGP	<b>0,9680</b>	<b>0,4971</b>	<b>0,1324</b>	<b>0,0880</b>	<b>0,5763</b>	1,2107	<b>1,0068</b>	0,9854	0,5683	<b>2,7342</b>	<b>1,3485</b>	<b>2,2923</b>	<b>1,0340</b>
e139	ARMA	3,1069	1,1962	0,2061	0,0445	0,6827	1,0322	1,1911	<b>0,9672</b>	<b>0,4976</b>	3,6281	3,3153	2,3939	1,5218
	PG-Pura	1,1387	0,5274	0,1476	<b>0,0411</b>	0,5341	0,9792	1,4551	1,0956	0,5029	1,3335	<b>1,8864</b>	2,5226	1,0137
	B1	1,0469	0,5175	0,1541	0,0604	0,5567	0,9050	<b>1,1239</b>	1,0956	0,5552	1,3064	1,9197	2,5443	0,9821
	BCIGP	<b>1,0013</b>	<b>0,4794</b>	<b>0,1402</b>	0,0544	<b>0,4884</b>	<b>0,8025</b>	1,4582	0,9860	0,4990	<b>1,2007</b>	2,1576	<b>2,3260</b>	<b>0,9661</b>
e140	ARMA	3,2906	1,2308	0,2416	0,1056	0,7351	0,9074	1,1902	<b>0,8383</b>	0,4626	2,4915	2,1326	<b>2,3229</b>	1,3291
	PG-Pura	1,0665	0,5146	0,1320	<b>0,0818</b>	0,6668	0,8651	1,1983	0,9995	0,4954	0,9462	1,2941	2,5147	0,8979
	B1	1,0021	0,5000	0,1325	0,1005	0,6477	0,8374	1,1504	0,9995	0,5165	0,8839	1,3378	2,6584	0,8972
	BCIGP	<b>0,9717</b>	<b>0,4621</b>	<b>0,1252</b>	0,0821	<b>0,5842</b>	<b>0,7341</b>	<b>1,0452</b>	0,8995	<b>0,4616</b>	<b>0,7969</b>	<b>1,2192</b>	<b>2,4455</b>	<b>0,8189</b>
e141	ARMA	3,2970	1,2763	0,2233	0,1528	0,6986	0,9472	1,0985	<b>0,8969</b>	0,4724	2,0265	1,9687	3,2547	1,3594
	PG-Pura	1,1616	0,5086	0,1523	0,0879	0,6027	0,9458	1,3945	1,3158	0,5743	0,7421	1,2459	3,2535	0,9987
	B1	1,0821	0,4511	0,1429	0,1038	0,5843	0,8977	1,0768	1,3158	0,5705	0,7226	<b>1,2412</b>	3,3381	<b>0,9606</b>
	BCIGP	<b>1,0313</b>	<b>0,4051</b>	<b>0,1357</b>	<b>0,0819</b>	<b>0,5414</b>	<b>0,8280</b>	<b>0,9625</b>	1,1842	<b>0,4614</b>	<b>0,6542</b>	6,6515	<b>3,0304</b>	1,3306
e142	ARMA	3,7910	1,2337	0,2574	0,1554	0,7694	1,0540	1,0239	<b>0,9728</b>	<b>0,5179</b>	2,8715	1,5340	<b>2,2673</b>	1,3707
	PG-Pura	0,9814	0,4768	0,1367	<b>0,1089</b>	0,6220	0,9713	1,2814	1,3179	22,1876	1,7936	1,3587	2,9584	2,8496
	B1	0,9244	<b>0,4607</b>	0,1295	0,1166	0,6441	0,9261	0,9908	1,3179	0,5933	1,8613	1,3120	3,0688	1,0288
	BCIGP	<b>0,8910</b>	1,5279	<b>0,1135</b>	0,1256	<b>0,5163</b>	<b>0,8577</b>	<b>0,9484</b>	1,2861	0,7022	<b>1,6828</b>	<b>1,2257</b>	2,8278	<b>1,0588</b>
e143	ARMA	4,1995	1,2061	0,2606	0,0969	0,7172	1,0102	1,1003	0,8576	0,5755	1,9112	2,9293	<b>1,6482</b>	1,3761
	PG-Pura	1,0518	0,4932	0,1474	0,0729	0,6618	0,9882	1,1356	0,9096	1,8902	18,6550	2,8987	2,1351	2,5866
	B1	1,0084	0,4654	0,1462	0,0676	0,6166	0,9008	1,0177	0,9096	0,6442	<b>1,0903</b>	2,8846	2,2364	0,9990
	BCIGP	<b>0,9187</b>	<b>0,4244</b>	<b>0,1309</b>	<b>0,0548</b>	<b>0,5698</b>	<b>0,7980</b>	<b>0,9164</b>	<b>0,8186</b>	<b>0,5461</b>	1,3028	<b>2,5884</b>	2,0104	<b>0,9233</b>
e144	ARMA	4,4809	1,2337	0,3183	0,1204	0,6258	0,9303	1,1644	1,0240	0,5115	1,0850	2,4556	<b>1,6141</b>	1,2970
	PG-Pura	1,0845	0,4874	<b>0,1178</b>	0,0722	0,4940	0,8264	1,2448	1,1045	0,5893	0,4895	2,5229	2,5122	0,9621
	B1	1,0689	0,4515	0,1302	0,0714	0,5090	0,8369	1,1072	1,1045	0,5428	0,4789	2,4714	2,5395	0,9427
	BCIGP	<b>0,9639</b>	<b>0,4087</b>	0,1179	<b>0,0701</b>	<b>0,4533</b>	<b>0,7552</b>	<b>1,0497</b>	<b>0,9940</b>	<b>0,5012</b>	<b>0,4725</b>	<b>2,2935</b>	2,3189	<b>0,8666</b>
e145	ARMA	4,5527	1,0910	0,3332	0,0692	0,7542	0,8363	1,0868	<b>0,9477</b>	0,5860	1,2869	1,8975	<b>2,8628</b>	1,3587
	PG-Pura	1,0719	0,4458	0,1235	<b>0,0605</b>	0,6531	0,7935	1,1768	1,0552	0,6464	0,9279	1,8491	3,8813	1,0571
	B1	1,0203	0,4705	0,1561	0,0712	0,6249	0,7937	<b>1,0474</b>	1,0552	0,6353	0,8658	1,8685	3,9282	<b>1,0448</b>
	BCIGP	<b>0,9749</b>	<b>0,4223</b>	<b>0,1133</b>	0,0686	<b>0,5724</b>	<b>0,7124</b>	8,3915	0,9496	<b>0,5600</b>	<b>0,8223</b>	<b>1,7258</b>	3,5887	1,5751
e146	ARMA	4,3709	1,2117	0,2882	0,0857	0,7443	1,0218	1,0927	<b>0,8802</b>	<b>0,5099</b>	1,5789	4,2805	<b>4,9902</b>	1,7546
	PG-Pura	1,1581	0,4996	0,2163	0,0635	0,5866	0,9807	1,2069	1,0162	0,6733	1,2449	4,2139	6,2016	1,5051
	B1	0,9448	0,5286	<b>0,1724</b>	<b>0,0625</b>	0,6055	0,9464	1,1240	1,0162	0,5727	1,1843	4,1517	6,2879	1,4664
	BCIGP	<b>0,8912</b>	<b>0,4909</b>	0,3130	0,0690	<b>0,5267</b>	<b>0,8816</b>	<b>1,0079</b>	0,9146	0,5323	<b>1,1067</b>	<b>3,8479</b>	5,7261	<b>1,3590</b>
e147	ARMA	4,5453	1,1528	0,2557	<b>0,0774</b>	0,7593	0,9743	0,9481	<b>1,0134</b>	<b>0,5533</b>	1,4147	<b>2,0436</b>	<b>4,6041</b>	1,5285
	PG-Pura	1,0135	0,4781	0,1556	0,0882	0,6602	0,9001	0,9746	1,3674	0,6477	1,1749	2,2620	5,7313	1,2878
	B1	0,9539	0,4659	0,1440	0,0870	0,5914	0,8559	0,9153	1,3674	0,6569	1,1453	2,2565	5,8288	1,2724
	BCIGP	<b>0,8781</b>	<b>0,4353</b>	<b>0,1219</b>	0,0828	<b>0,5216</b>	<b>0,7738</b>	<b>0,8146</b>	1,2306	0,5780	<b>1,0595</b>	2,0660	5,3036	<b>1,1555</b>
e148	ARMA	4,5619	1,1798	0,2315	0,0787	0,6408	0,8525	1,3024	<b>0,9444</b>	0,4995	2,7966	3,8039	<b>3,0263</b>	1,6599
	PG-Pura	1,2477	0,4932	0,2131	0,0559	0,5591	0,8033	1,2652	1,1180	0,6749	2,5767	3,9575	4,0344	1,4166
	B1	<b>0,9491</b>	0,5095	0,1775	0,0528	0,5536	0,8322	1,2297	1,1180	0,5384	2,5772	3,9907	4,1444	1,3894
	BCIGP	0,9533	<b>0,4497</b>	<b>0,1619</b>	<b>0,0449</b>	<b>0,5135</b>	<b>0,7246</b>	<b>1,1702</b>	1,0062	<b>0,4640</b>	<b>2,3187</b>	<b>3,7305</b>	3,7620	<b>1,2750</b>
e149	ARMA	4,5275	1,0555	0,2801	0,0683	0,7142	0,8994	1,2101	<b>0,8105</b>	<b>0,5696</b>	2,2555	3,4711	<b>1,1786</b>	1,4200
	PG-Pura	0,9923	<b>0,4615</b>	<b>0,1440</b>	0,0473	0,6460	0,9172	1,2330	0,9552	0,7169	2,2123	3,7843	2,1049	1,1846
	B1	0,9487	0,4827	0,1541	0,0474	0,6490	0,9046	1,1403	0,9552	0,6993	2,2239	3,8933	2,1286	1,1856
	BCIGP	<b>0,9224</b>	0,4906	0,1683	<b>0,0432</b>	<b>0,5568</b>	<b>0,7868</b>	<b>1,0319</b>	0,8597	0,5843	<b>2,1602</b>	<b>3,4347</b>	1,9228	<b>1,0801</b>
e150	ARMA	4,7239	1,0751	0,2600	0,0564	0,7383	0,9307	1,1548	1,0807	<b>0,4749</b>	6,3139	<b>5,8322</b>	<b>3,2017</b>	2,1536
	PG-Pura	1,1353	0,4104	0,1188	0,0400	0,6444	1,1732	1,1503	1,1604	0,8354	5,9629	6,4714	7,7428	2,2371
	B1	<b>1,0586</b>	0,3809	0,1180	0,0437	0,6316	0,9200	1,1113	1,1604	0,5599	5,9857	6,4942	4,1128	<b>1,8814</b>
	BCIGP	5,2786	<b>0,3527</b>	<b>0,0975</b>	<b>0,0299</b>	<b>0,5693</b>	<b>0,8189</b>	<b>1,0452</b>	<b>1,0444</b>	0,4831	<b>5,4569</b>	5,8879	3,6751	2,0616

Algumas colunas foram ocultadas, simplesmente para efeito de enquadramento na página. Na última coluna os valores correspondem à média dos 19 parâmetros dos *MSE* para cada horizonte de previsão. Os resultados obtidos para esta estrutura, mostram que o método *BCIGP* pode ser considerado como um método alternativo de previsões, já que os resultados obtidos apresentam erros de previsão menores do que os obtidos através dos outros métodos analisados. Para as demais estruturas: AR(2), MA(1), MA(2) e ARMA(1, 1), os resultados também foram satisfatórios, pois os erros de previsão obtidos (fora da amostra) são em geral, menores do que os obtidos pelos demais métodos. As outras tabelas contendo os resultados obtidos para as demais estruturas estão apresentadas no Apêndice C desta tese.

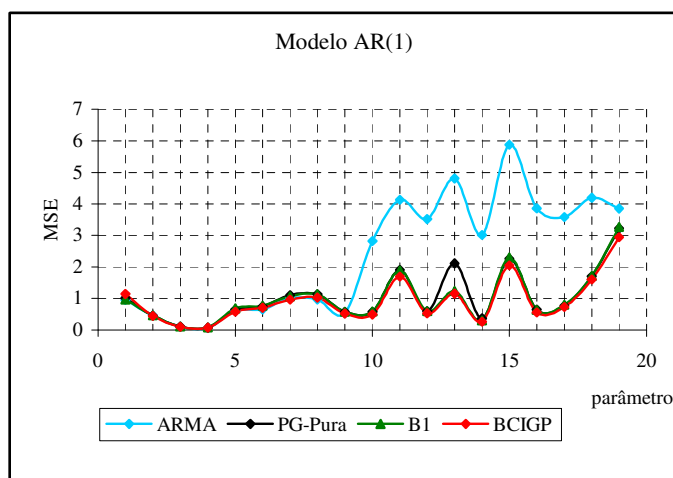


FIGURA 7.5 – COMPORTAMENTO DOS MÉTODOS PARA O 1º VALOR PREVISTO *e136* – MODELO AR(1)

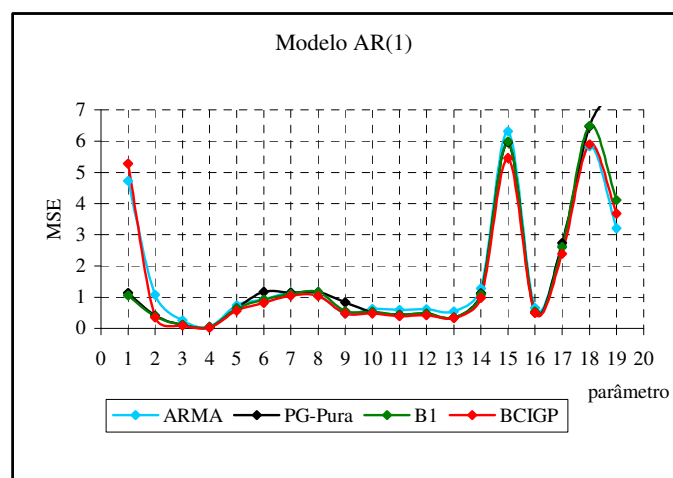


FIGURA 7.6 – COMPORTAMENTO DOS MÉTODOS PARA O O ÚLTIMO VALOR PREVISTO *e150* - MODELO AR(1)

A figura (7.5) mostra o comportamento dos algoritmos implementados, de acordo com os valores do *MSE* obtidos no conjunto de teste para o horizonte de previsão um (*e136*), das séries geradas para a estrutura AR(1) e a figura (7.6) mostra o comportamento dos mesmos algoritmos, porém considerando o horizonte de previsão 15 (*e150*) para o mesmo teste. Os gráficos representativos das demais estruturas, para os horizontes de previsão um e quinze estão apresentados no apêndice C desta tese.

Na tabela (7.4) são apresentados as médias dos *MSE*'s dos parâmetros de cada estrutura. Observa-se que os valores médios dos *MSE* são na maioria das vezes menor para o algoritmo BCIGP, sendo em alguns casos maior do que os obtidos pelo *GPBoost*, que é o terceiro algoritmo analisado.

Nas tabela (7.5) são apresentados os resultados comparativos do desempenho dos algoritmos utilizados através da Análise de Variância (ANOVA) e do teste Tukey-Kramer, descritos no capítulo V, para as séries geradas. Na coluna 2, são apresentados os grupos (ou métodos) analisados, numerados de 1 a 4, onde o número 1 corresponde à metodologia ARMA, o número 2, corresponde ao algoritmo de PG tradicional, o número 3 corresponde ao *GPBoost* e o número 4 é o algoritmo BCIGP; na coluna 3, são apresentadas as médias dos *MSE* de cada grupo para a quantidade de parâmetros de cada estrutura; na coluna 4, são apresentadas as médias gerais entre os 4 grupos. A estatística F é apresentada na coluna 5. Para o nível de significância  $\alpha = 0.05$ , a hipótese nula será rejeitada, se a estatística do teste F, calculada, for maior do que o valor crítico  $F_s$ . Na coluna 6 é apresentado o valor-p, que representa a probabilidade de ser obtida uma estatística F tão grande ou maior do que aquela obtida, se a hipótese nula for verdadeira. Analisando o valor-p, pode-se concluir se a hipótese nula deve ou não ser rejeitada, sem a necessidade de recorrer à tabela dos valores críticos de F. Se o valor-p for menor do que o nível de significância  $\alpha$ , a hipótese nula é rejeitada. A partir da coluna 7, estão os resultados do teste Tukey-Kramer, nesta coluna são mostrados os grupos que estão sendo comparados. Na coluna 9 estão os valores do intervalo crítico Q, se os valores mostrados na coluna 8 forem superiores ao valor do intervalo crítico, existe diferença entre os métodos. O melhor dos métodos, quando a diferença existe, está representado na última coluna. A tabela apresentada mostra os resultados obtidos para a estrutura MA(1). As tabelas que mostram os resultados da análise de variância das demais estruturas estão apresentadas no Apêndice C desta tese.

TABELA 7.4 – MÉDIA DOS *MSE* DOS PARÂMETROS PARA CADA ESTRUTURA

Valor Previsto	<i>MSE</i> Médio	AR(1)	AR(2)	MA(1)	MA(2)	ARMA(1, 1)
e136	ARMA	2,3702	4,3479	2,3150	3,0335	1,7917
	PG-Pura	1,0567	1,3176	5,6063	2,2456	1,8133
	B1	1,0118	1,1906	1,1318	2,0985	1,1327
	BCIGP	<b>0,9282</b>	<b>1,0997</b>	<b>1,0781</b>	<b>1,9234</b>	<b>1,1087</b>
e137	ARMA	2,3527	4,1828	1,8223	2,7355	2,0809
	PG-Pura	1,9907	1,5090	1,6281	2,1690	1,2657
	B1	1,1842	1,3537	1,1394	2,0773	1,1427
	BCIGP	<b>1,0924</b>	<b>1,2563</b>	<b>1,0260</b>	<b>1,8843</b>	<b>1,0553</b>
e138	ARMA	2,0450	4,0820	1,8098	2,6617	2,2494
	PG-Pura	1,0470	1,3674	1,1456	3,2616	1,3300
	B1	1,0311	<b>1,1733</b>	1,0935	2,0698	1,1277
	BCIGP	<b>0,9583</b>	1,4734	<b>0,9838</b>	<b>1,9117</b>	<b>1,0897</b>
e139	ARMA	2,1817	4,1294	1,9132	2,6473	2,3892
	PG-Pura	1,3314	1,4605	1,2253	3,0242	1,6402
	B1	1,2955	1,2963	1,1898	2,0711	1,1417
	BCIGP	<b>1,2194</b>	<b>1,5489</b>	<b>1,0972</b>	<b>1,9318</b>	<b>1,0871</b>
e140	ARMA	2,0486	4,2354	1,7691	2,6666	2,4768
	PG-Pura	1,5099	1,6394	1,2392	2,1816	1,3522
	B1	1,3779	<b>1,4485</b>	1,1130	2,0828	1,1416
	BCIGP	<b>1,2547</b>	1,4844	<b>1,0077</b>	<b>1,8970</b>	<b>1,0806</b>
e141	ARMA	1,9550	4,5942	1,8326	2,6444	2,5734
	PG-Pura	1,4487	1,5345	1,1601	2,3654	1,2922
	B1	<b>1,2907</b>	1,3558	1,1285	2,0613	1,1399
	BCIGP	1,4669	<b>1,3466</b>	<b>1,0123</b>	<b>1,9397</b>	<b>1,1006</b>
e142	ARMA	1,6495	4,5034	1,8526	2,6580	2,6567
	PG-Pura	2,2842	1,6545	1,2233	2,3011	1,6019
	B1	1,1341	<b>1,1711</b>	1,1996	2,0800	<b>1,1467</b>
	BCIGP	<b>1,0966</b>	1,2458	<b>1,0695</b>	<b>1,9264</b>	1,2413
e143	ARMA	1,4880	4,7462	1,8323	2,6514	2,6738
	PG-Pura	2,0030	1,7681	1,2425	3,1361	1,2913
	B1	0,9834	<b>1,3549</b>	1,1424	<b>2,0581</b>	1,1375
	BCIGP	<b>0,9053</b>	1,4738	<b>1,0188</b>	4,4017	<b>1,0651</b>
e144	ARMA	1,2274	3,6964	1,7596	2,6497	2,7013
	PG-Pura	0,8461	5,0540	1,3345	2,7877	1,3619
	B1	0,8106	<b>1,2776</b>	1,1668	<b>2,2478</b>	1,1386
	BCIGP	<b>0,7475</b>	1,5608	<b>1,0914</b>	2,6743	<b>1,0664</b>
e145	ARMA	1,4540	3,6346	1,8127	2,6596	2,7278
	PG-Pura	1,1212	1,4859	1,4670	3,0190	1,3692
	B1	<b>1,1029</b>	1,2636	1,1852	<b>2,0717</b>	<b>1,1365</b>
	BCIGP	1,3978	<b>1,2616</b>	<b>1,0703</b>	2,4415	1,4385
e146	ARMA	1,6769	3,3792	1,8306	2,6673	2,7513
	PG-Pura	1,4171	2,4862	1,1538	2,5006	10,2940
	B1	1,3755	1,2689	1,1349	<b>2,0858</b>	1,1417
	BCIGP	<b>1,2677</b>	<b>1,2078</b>	<b>1,0377</b>	2,4169	<b>1,1677</b>
e147	ARMA	1,2946	3,4343	1,8601	2,6690	2,7759
	PG-Pura	1,0576	1,5164	1,2334	2,7557	1,8779
	B1	1,0276	1,3071	1,1368	<b>2,0776</b>	<b>1,1450</b>
	BCIGP	<b>0,9326</b>	<b>1,2561</b>	<b>1,0224</b>	2,5989	1,1471
e148	ARMA	1,3387	3,3868	1,7664	2,6551	2,7658
	PG-Pura	1,1235	2,2805	1,2059	2,7182	6,5438
	B1	1,0967	1,2308	1,0886	<b>2,0734</b>	1,2860
	BCIGP	<b>1,0049</b>	<b>1,1477</b>	<b>1,0428</b>	2,1184	<b>1,2057</b>
e149	ARMA	1,2239	3,5028	1,7232	2,6740	2,8048
	PG-Pura	1,0304	1,5118	1,1609	2,7705	1,3648
	B1	1,0237	<b>1,2823</b>	1,1196	2,1262	1,1386
	BCIGP	<b>0,9286</b>	1,8203	<b>1,0069</b>	<b>1,9977</b>	<b>1,0578</b>
e150	ARMA	1,7268	3,4916	1,7436	2,6571	2,8316
	PG-Pura	1,7379	1,6753	1,1366	2,5503	3,4325
	B1	<b>1,5064</b>	<b>1,4174</b>	1,0936	<b>2,0637</b>	1,1398
	BCIGP	1,5922	1,4630	<b>0,9879</b>	2,0916	<b>1,0977</b>

TABELA 7.5 – RESULTADOS DA ANOVA E TESTE TUKEY-KRAMER PARA A ESTRUTURA MA(1)

Valor Previsto	Grupos	Média Grupo	Média Geral	Estatís. F	p-valor	Grupos	Dif. Abs.	Interv. Crítico	Comp. médias	Melhor Grupo
e136	1	2,3150	2,53284	0,9553	0,4139	Grupo1 e 2	3,2913	8,0918	Iguais	X
	2	5,6063				Grupo1 e 3	1,1832		Iguais	X
	3	1,1318				Grupo1 e 4	1,2369		Iguais	X
	4	1,0781				Grupo2 e 3	4,4745		Iguais	X
						Grupo2 e 4	4,5282		Iguais	X
						Grupo3 e 4	0,0537		Iguais	X
e137	1	1,8223	1,4040	2,1062	0,0989	Grupo1 e 2	0,1943	0,9791	Iguais	X
	2	1,6281				Grupo1 e 3	0,6829		Iguais	X
	3	1,1394				Grupo1 e 4	0,7964		Iguais	X
	4	1,0260				Grupo2 e 3	0,4887		Iguais	X
						Grupo2 e 4	0,6021		Iguais	X
						Grupo3 e 4	0,1134		Iguais	X
e138	1	1,8098	1,2582	6,0609	0,0005	Grupo1 e 2	0,6642	0,5647	Diferentes	2
	2	1,1456				Grupo1 e 3	0,7163		Diferentes	3
	3	1,0935				Grupo1 e 4	0,8260		Diferentes	4
	4	0,9838				Grupo2 e 3	0,0521		Iguais	X
						Grupo2 e 4	0,1618		Iguais	X
						Grupo3 e 4	0,1097		Iguais	X
e139	1	1,9132	1,3564	6,5183	0,0003	Grupo1 e 2	0,6879	0,5463	Diferentes	2
	2	1,2253				Grupo1 e 3	0,7234		Diferentes	3
	3	1,1898				Grupo1 e 4	0,8160		Diferentes	4
	4	1,0972				Grupo2 e 3	0,0355		Iguais	X
						Grupo2 e 4	0,1281		Iguais	X
						Grupo3 e 4	0,0926		Iguais	X
e140	1	1,7691	1,2822	5,6142	0,0009	Grupo1 e 2	0,5299	0,5306	Iguais	X
	2	1,2392				Grupo1 e 3	0,6561		Diferentes	3
	3	1,1130				Grupo1 e 4	0,7614		Diferentes	4
	4	1,0077				Grupo2 e 3	0,1261		Iguais	X
						Grupo2 e 4	0,2315		Iguais	X
						Grupo3 e 4	0,1054		Iguais	X
e141	1	1,8326	1,2834	5,6546	0,0008	Grupo1 e 2	0,6725	0,5811	Diferentes	2
	2	1,1601				Grupo1 e 3	0,7042		Diferentes	3
	3	1,1285				Grupo1 e 4	0,8204		Diferentes	4
	4	1,0123				Grupo2 e 3	0,0317		Iguais	X
						Grupo2 e 4	0,1478		Iguais	X
						Grupo3 e 4	0,1162		Iguais	X
e142	1	1,8526	1,3363	5,2527	0,0015	Grupo1 e 2	0,6294	0,5692	Diferentes	2
	2	1,2233				Grupo1 e 3	0,6530		Diferentes	3
	3	1,1996				Grupo1 e 4	0,7831		Diferentes	4
	4	1,0695				Grupo2 e 3	0,0236		Iguais	X
						Grupo2 e 4	0,1537		Iguais	X
						Grupo3 e 4	0,1301		Iguais	X

**Nota:** A legenda apresentada a seguir tem o mesmo significado para todas as tabelas do teste ANOVA, inclusive para as tabelas que se encontram nos apêndices.

1 – Modelos ARMA; 2 – Programação Genética; 3 – Algoritmo *GPBoost*; 4 – Algoritmo BCIGP



**TABELA 7.5 – RESULTADOS DA ANOVA E TESTE TUKEY-KRAMER PARA A ESTRUTURA MA(1) (CONTINUAÇÃO)**

Valor Previsto	Grupos	Média Grupo	Média Geral	Estatís. F	p-valor	Grupos	Dif. Abs.	Interv. Crítico	Comp. médias	Melhor Grupo
e143	1	1,8323	1,3090	7,2735	0,0001	Grupo1 e 2	0,5898	0,4972	Diferentes	2
	2	1,2425				Grupo1 e 3	0,6899		Diferentes	3
	3	1,1424				Grupo1 e 4	0,8135		Diferentes	4
	4	1,0188				Grupo2 e 3	0,1000		Iguais	X
						Grupo2 e 4	0,2237		Iguais	X
						Grupo3 e 4	0,1236		Iguais	X
e144	1	1,7596	1,3381	3,9063	0,0090	Grupo1 e 2	0,4251	0,5622	Iguais	X
	2	1,3345				Grupo1 e 3	0,5928		Diferentes	3
	3	1,1668				Grupo1 e 4	0,6682		Diferentes	4
	4	1,0914				Grupo2 e 3	0,1677		Iguais	X
						Grupo2 e 4	0,2430		Iguais	X
						Grupo3 e 4	0,0753		Iguais	X
e145	1	1,8127	1,3838	4,0350	0,0076	Grupo1 e 2	0,3457	0,6126	Iguais	X
	2	1,4670				Grupo1 e 3	0,6275		Diferentes	3
	3	1,1852				Grupo1 e 4	0,7424		Diferentes	4
	4	1,0703				Grupo2 e 3	0,2818		Iguais	X
						Grupo2 e 4	0,3967		Iguais	X
						Grupo3 e 4	0,1149		Iguais	X
e146	1	1,8306	1,2893	6,0779	0,0005	Grupo1 e 2	0,6768	0,5497	Diferentes	2
	2	1,1538				Grupo1 e 3	0,6957		Diferentes	3
	3	1,1349				Grupo1 e 4	0,7929		Diferentes	4
	4	1,0377				Grupo2 e 3	0,0189		Iguais	X
						Grupo2 e 4	0,1161		Iguais	X
						Grupo3 e 4	0,0972		Iguais	X
e147	1	1,8601	1,3132	5,4731	0,0011	Grupo1 e 2	0,6267	0,5955	Diferentes	2
	2	1,2334				Grupo1 e 3	0,7233		Diferentes	3
	3	1,1368				Grupo1 e 4	0,8377		Diferentes	4
	4	1,0224				Grupo2 e 3	0,0966		Iguais	X
						Grupo2 e 4	0,2110		Iguais	X
						Grupo3 e 4	0,1144		Iguais	X
e148	1	1,7664	1,27594	5,3376	0,0013	Grupo1 e 2	0,5604	0,5377	Diferentes	2
	2	1,2059				Grupo1 e 3	0,6777		Diferentes	3
	3	1,0886				Grupo1 e 4	0,7235		Diferentes	4
	4	1,0428				Grupo2 e 3	0,1173		Iguais	X
						Grupo2 e 4	0,1631		Iguais	X
						Grupo3 e 4	0,0458		Iguais	X
e149	1	1,7232	1,2527	5,5767	0,0009	Grupo1 e 2	0,5623	0,5045	Diferentes	2
	2	1,1609				Grupo1 e 3	0,6037		Diferentes	3
	3	1,1196				Grupo1 e 4	0,7163		Diferentes	4
	4	1,0069				Grupo2 e 3	0,0414		Iguais	X
						Grupo2 e 4	0,1541		Iguais	X
						Grupo3 e 4	0,1127		Iguais	X
e150	1	1,7436	1,2404	5,1803	0,0016	Grupo1 e 2	0,6070	0,5575	Diferentes	2
	2	1,1366				Grupo1 e 3	0,6500		Diferentes	3
	3	1,0936				Grupo1 e 4	0,7557		Diferentes	4
	4	0,9879				Grupo2 e 3	0,0430		Iguais	X
						Grupo2 e 4	0,1487		Iguais	X
						Grupo3 e 4	0,1057		Iguais	X

Um resumo dos resultados é mostrado nas tabelas (7.6) a (7.9), são apresentados os valores-p de cada um dos horizontes de previsão, para as cinco estruturas e o melhor algoritmo está representado em negrito. Nas linhas onde aparece a letra “x” não existe diferença significativa entre os algoritmos e que qualquer um deles pode ser selecionado ao acaso para efetuar as previsões.

TABELA 7.6 – VALOR - p – AR (1)

AR(1)		
Valor Previsto	p-valor ANOVA	Melhor Grupo
<b>e136</b>	0,0758955	4
<b>e137</b>	1,46E-07	X
<b>e138</b>	6,71E-10	4
<b>e139</b>	1,57E-08	X
<b>e140</b>	3,63E-13	X
<b>e141</b>	1,14E-14	X
<b>e142</b>	2,60E-06	X
<b>e143</b>	6,88E-15	X
<b>e144</b>	1,38E-14	X
<b>e145</b>	2,81E-08	X
<b>e146</b>	0,1778037	X
<b>e147</b>	1,25E-06	X
<b>e148</b>	0,3044825	X
<b>e149</b>	4,36E-14	X
<b>e150</b>	0,0209229	X

E VALOR - p – AR (2)

AR(2)		
Valor Previsto	p-valor ANOVA	Melhor Grupo
<b>e136</b>	2,52E-12	4
<b>e137</b>	2,67E-11	4
<b>e138</b>	4,89E-11	4
<b>e139</b>	2,90E-09	4
<b>e140</b>	1,89E-08	4
<b>e141</b>	1,34E-09	4
<b>e142</b>	1,08E-06	4
<b>e143</b>	3,50E-05	4
<b>e144</b>	0,064436	x
<b>e145</b>	9,88E-15	4
<b>e146</b>	1,62E-06	4
<b>e147</b>	1,04E-13	4
<b>e148</b>	2,93E-05	4
<b>e149</b>	7,70E-06	4
<b>e150</b>	7,08E-12	4

TABELA 7.7 – VALOR - p – MA (1)

MA(1)		
Valor Previsto	p-valor ANOVA	Melhor Grupo
<b>e136</b>	0,4138558	X
<b>e137</b>	0,098902	X
<b>e138</b>	0,0004836	X
<b>e139</b>	0,0002593	4
<b>e140</b>	0,0008891	4
<b>e141</b>	0,0008415	4
<b>e142</b>	0,0014546	4
<b>e143</b>	9,27E-05	4
<b>e144</b>	0,0090384	4
<b>e145</b>	0,0075967	4
<b>e146</b>	0,0004726	4
<b>e147</b>	0,0010775	4
<b>e148</b>	0,0012959	4
<b>e149</b>	0,0009356	4
<b>e150</b>	0,0016053	4

E VALOR - p – MA (2)

MA(2)		
Valor Previsto	p-valor ANOVA	Melhor Grupo
<b>e136</b>	9,35E-13	4
<b>e137</b>	6,04E-11	4
<b>e138</b>	0,004464	4
<b>e139</b>	0,070179	x
<b>e140</b>	9,51E-09	4
<b>e141</b>	2,55E-06	4
<b>e142</b>	1,05E-06	4
<b>e143</b>	0,533148	x
<b>e144</b>	0,669801	x
<b>e145</b>	0,341698	x
<b>e146</b>	0,287049	x
<b>e147</b>	0,465983	x
<b>e148</b>	0,022435	x
<b>e149</b>	0,000156	4
<b>e150</b>	0,037747	x

TABELA 7.8 – VALOR - p – ARMA (1,1)

ARMA(1, 1)		
Valor Previsto	p-valor ANOVA	Melhor Grupo
e136	0,0758955	x
e137	1,46E-07	4
e138	6,71E-10	4
e139	1,57E-08	4
e140	3,63E-13	4
e141	1,14E-14	4
e142	2,60E-06	4
e143	6,88E-15	4
e144	1,38E-14	4
e145	2,81E-08	4
e146	0,1778037	x
e147	1,25E-06	4
e148	0,3044825	x
e149	4,36E-14	4
e150	0,0209229	x

Para as estruturas AR(1) e ARMA(1,1) verifica-se que o valor-p é menor que 0,01 em 74% dos casos (horizontes de previsão); para a estrutura AR(2) o valor-p é menor que 0,01 em 94% dos casos; para MA(1) o valor-p é menor que 0,01 para 87% dos casos e para a estrutura MA(2) o valor-p é menor que 0,01 para 54% dos casos, ou seja, para o nível de significância de 1%, a hipótese nula de que a diferença média entre os MSE dos algoritmos é zero, é rejeitada.

### 7.3 EXPERIMENTO UTILIZANDO PESOS 2

Nesta seção serão apresentados os resultados dos experimentos realizados utilizando a atualização dos pesos proposta no algoritmo apresentado no capítulo V.

O experimento utilizando a atualização de pesos foi aplicado a somente uma das estruturas, pois o tempo de processamento para todos os dados propostos na seção anterior foi consideravelmente alto e a utilização do *cluster* de computadores da Universidade Federal do Paraná já estava em seu limite máximo. Assim, optou-se por verificar o desempenho do algoritmo completo em uma das bases de dados. A estrutura escolhida ao acaso foi a MA(1). As mesmas configurações do teste anterior foram utilizadas, ou seja, para cada um dos parâmetros da estrutura MA(1), foram geradas 500 séries e para cada série, foram feitas as

previsões para os 4 algoritmos analisados. Os resultados da média dos *MSE* das 500 séries encontram-se dispostos na tabela (7.10). Verifica-se que o método teve um bom comportamento, tendo apresentado erros de previsão menores do que os outros métodos analisados.

**TABELA 7.10 – MSE MÉDIO PARA A ESTRUTURA MA(1) – ATUALIZAÇÃO DE PESOS 2**

<i>MSE</i>	MA(1)P2	P1	P2	P3	P4	P5	P8	P7	P11	P16	P17	P18	P19	MÉDIA
e136	ARMA	1,9426	1,4588	1,4788	1,0890	<b>1,0997</b>	<b>1,0491</b>	1,0542	2,6897	3,3434	3,0012	4,1974	6,8716	2,4396
	PG-Pura	0,9832	0,9889	1,1166	1,0069	1,2894	2,2564	<b>0,9621</b>	4,6824	1,3333	1,2750	1,0728	1,5770	1,5453
	B1	0,9408	<b>0,9767</b>	1,2052	1,0109	1,3169	1,1673	1,2763	1,1558	1,2170	1,3750	1,0358	1,5757	<b>1,1878</b>
	BCIGP	<b>0,8358</b>	1,9808	<b>1,0820</b>	<b>0,8352</b>	1,1020	1,0972	2,5313	<b>0,9598</b>	<b>1,0214</b>	<b>1,1308</b>	<b>0,9621</b>	<b>1,4128</b>	1,2459
e137	ARMA	5,3663	2,2476	2,4543	1,8508	1,5834	1,0368	1,2010	<b>1,0186</b>	1,7992	1,8284	2,3638	4,7280	2,2899
	PG-Pura	0,9459	0,9886	1,4495	1,3925	1,4447	1,0178	1,1218	1,3464	0,9417	1,2518	1,0040	1,5451	1,2042
	B1	<b>0,9381</b>	0,9407	1,2631	0,9875	<b>1,3266</b>	0,9132	1,1490	1,1829	0,9125	1,1860	0,9059	1,7657	1,1226
	BCIGP	1,1990	<b>0,8691</b>	<b>1,2108</b>	<b>0,8936</b>	1,5668	<b>0,7990</b>	<b>1,0068</b>	1,0532	<b>0,7933</b>	<b>1,1677</b>	<b>0,8365</b>	<b>1,4851</b>	<b>1,0734</b>
e138	ARMA	5,9171	2,4518	1,9754	1,6532	1,3401	1,0976	1,2134	0,9323	1,2916	2,5770	2,3512	4,8031	2,3003
	PG-Pura	1,1319	1,3280	1,1158	0,9596	1,1995	1,0738	1,1238	1,0301	1,1054	1,2718	0,8756	<b>1,4896</b>	1,1421
	B1	1,0968	1,0996	1,0750	<b>0,8529</b>	1,2039	1,0432	1,3326	1,0415	0,9240	1,1784	<b>0,7717</b>	1,6003	1,1017
	BCIGP	<b>0,9635</b>	<b>1,0363</b>	<b>0,9698</b>	1,8557	<b>1,0140</b>	<b>0,9148</b>	<b>1,0701</b>	<b>0,8861</b>	<b>0,8461</b>	<b>1,0414</b>	0,9998	1,5264	<b>1,0937</b>
e139	ARMA	5,7537	2,4986	2,1311	1,9588	1,3385	1,1592	1,4828	1,1612	1,4842	2,6675	2,3068	4,6620	2,3837
	PG-Pura	1,4775	0,9748	1,5084	5,6884	1,0746	1,1341	1,3467	1,2129	1,2230	1,2790	1,0218	<b>1,0722</b>	1,5845
	B1	1,3733	0,9815	1,1759	1,2935	1,1326	1,1302	1,3819	1,1151	1,1019	1,1238	0,9212	1,2139	1,1621
	BCIGP	<b>1,2381</b>	<b>0,9651</b>	<b>1,1281</b>	<b>1,1888</b>	<b>0,9598</b>	<b>0,9695</b>	<b>1,2619</b>	<b>1,0088</b>	<b>0,9593</b>	<b>0,9703</b>	<b>0,8713</b>	1,1293	<b>1,0542</b>
e140	ARMA	5,2151	2,4750	1,8091	1,5190	1,3689	1,1005	1,1835	<b>1,0461</b>	1,3098	2,2871	2,6425	4,4511	2,2007
	PG-Pura	1,0420	1,2125	1,4967	1,5000	1,3249	1,1262	1,2167	1,1308	1,1398	<b>1,0038</b>	1,0574	1,1438	1,1995
	B1	1,0710	1,1537	1,2139	1,1613	1,1855	1,0571	1,2517	1,2196	1,0886	1,1524	0,9314	1,2452	1,1443
	BCIGP	<b>0,9410</b>	<b>1,0937</b>	<b>1,0555</b>	<b>0,9756</b>	<b>1,0824</b>	<b>1,0312</b>	<b>1,0398</b>	1,0899	<b>0,9250</b>	1,0274	<b>0,8469</b>	<b>1,0072</b>	<b>1,0096</b>
e141	ARMA	5,1577	2,4440	1,7982	1,5003	0,8914	1,2026	1,1462	1,1999	1,3167	2,3317	2,7971	5,5038	2,2741
	PG-Pura	0,7914	1,2679	1,0698	1,7027	0,9725	1,0997	1,3131	1,3062	1,0568	1,0480	1,0906	1,4507	1,1808
	B1	0,9467	1,2235	1,0140	<b>0,9845</b>	0,9043	1,1975	1,2685	1,2786	1,0230	1,0392	1,0896	1,4824	1,1210
	BCIGP	<b>0,7814</b>	<b>1,1059</b>	<b>0,9114</b>	1,6178	<b>0,8308</b>	<b>1,0776</b>	<b>1,0565</b>	<b>1,1727</b>	<b>0,9885</b>	<b>0,9875</b>	<b>0,9790</b>	<b>1,3159</b>	<b>1,0688</b>
e142	ARMA	5,2085	2,3991	1,7273	1,5967	1,1387	1,3733	1,1085	<b>1,0241</b>	1,5628	2,4785	2,2297	5,4516	2,2749
	PG-Pura	0,9065	0,9225	0,9721	1,1787	1,0028	1,1551	1,2959	1,2473	1,4681	1,3905	0,9575	1,4764	1,1644
	B1	0,9055	0,8562	0,9182	1,0990	1,0444	1,0902	1,1544	1,1364	1,3639	1,3911	0,8823	1,5373	1,1149
	BCIGP	<b>0,7205</b>	<b>0,7597</b>	<b>0,8756</b>	<b>0,9245</b>	<b>0,8449</b>	<b>0,8970</b>	<b>1,0070</b>	1,1013	<b>1,1762</b>	<b>1,3026</b>	<b>0,8007</b>	<b>1,3541</b>	<b>0,9803</b>
e143	ARMA	4,9180	2,8184	1,8137	1,1805	1,7742	1,4047	1,3348	1,1548	1,3964	2,1014	2,3998	4,7049	2,2501
	PG-Pura	0,9310	1,1980	1,0519	7,2529	1,4178	1,3388	1,2811	1,3678	1,0434	<b>0,8348</b>	1,1756	1,5025	1,6996
	B1	0,9856	1,1200	1,0063	0,8925	1,3454	1,3533	1,2802	1,3102	1,1062	1,0128	<b>1,1344</b>	1,3920	1,1616
	BCIGP	<b>0,9138</b>	<b>1,0134</b>	<b>0,9371</b>	<b>0,8080</b>	<b>1,2060</b>	<b>1,2066</b>	<b>1,1559</b>	<b>1,0821</b>	<b>0,9901</b>	0,8407	1,1579	<b>1,1483</b>	<b>1,0383</b>
e144	ARMA	4,9709	2,5869	1,4667	1,4940	1,1077	1,1533	<b>0,9163</b>	1,1368	1,5427	2,3169	2,2021	5,0428	2,1614
	PG-Pura	1,4054	1,4991	0,9266	1,0879	0,9877	0,8664	15,0044	1,2391	1,1738	1,2968	1,0991	1,6376	2,3520
	B1	1,3013	1,2571	0,9560	1,1975	0,7641	0,9843	0,9810	1,2373	1,2260	1,3410	1,0312	1,6032	1,1567
	BCIGP	<b>1,2072</b>	<b>1,1598</b>	<b>0,8796</b>	<b>1,0026</b>	<b>0,7038</b>	<b>0,8657</b>	0,9402	<b>1,0916</b>	<b>1,1064</b>	<b>1,2240</b>	<b>0,9294</b>	<b>1,4638</b>	<b>1,0478</b>
e145	ARMA	5,2889	2,5349	1,6630	1,6558	1,1189	1,2713	<b>0,8812</b>	1,2233	1,6205	2,2071	2,3572	5,0183	2,2367
	PG-Pura	1,3618	1,2180	0,9709	1,2091	1,2028	1,1001	1,0406	11,5206	1,3925	1,2868	1,0409	1,4845	2,0691
	B1	1,3473	1,1649	0,8936	1,2899	1,2179	1,1475	1,0229	1,3251	1,2481	1,3529	0,9801	1,5133	1,2086
	BCIGP	<b>1,2944</b>	<b>1,0879</b>	<b>0,8349</b>	<b>1,1046</b>	<b>1,0991</b>	<b>0,9859</b>	0,9501	<b>1,1974</b>	<b>1,1702</b>	<b>1,1845</b>	<b>0,9031</b>	<b>1,3815</b>	<b>1,0994</b>
e146	ARMA	5,9532	2,5756	1,7463	1,9438	0,9938	1,3436	1,2046	<b>1,0071</b>	1,4098	2,1513	2,7226	4,8211	2,3227
	PG-Pura	<b>1,2714</b>	1,2225	1,0391	1,2254	1,1028	1,3127	1,2376	1,1730	1,0789	<b>1,1178</b>	1,1513	1,1718	1,1754
	B1	1,4319	1,0085	1,0353	1,3022	0,9662	1,2680	1,2023	1,1727	0,8963	1,2324	1,1230	1,1211	1,1467
	BCIGP	1,3508	<b>0,9204</b>	<b>0,9397</b>	<b>1,1277</b>	<b>0,9609</b>	<b>1,1354</b>	<b>1,0883</b>	1,0137	<b>0,8594</b>	1,1604	<b>1,0208</b>	<b>0,9798</b>	<b>1,0464</b>
e147	ARMA	6,1890	2,9263	1,9945	1,4392	1,2397	1,2708	<b>1,2562</b>	1,1705	1,6236	1,9823	2,0924	5,0530	2,3531
	PG-Pura	1,1846	1,2213	0,9281	1,1812	1,2746	1,3516	1,6572	1,4767	1,1871	1,0466	0,8259	1,1807	1,2096
	B1	1,1689	1,1496	0,9537	1,0010	1,0867	1,1695	1,3481	1,1635	1,1973	1,2148	0,7509	1,2375	1,1201
	BCIGP	<b>0,9558</b>	<b>1,0548</b>	<b>0,8576</b>	<b>0,9122</b>	<b>0,9458</b>	<b>1,0779</b>	1,8812	<b>1,0746</b>	<b>1,0759</b>	<b>1,0067</b>	<b>0,6660</b>	<b>1,1496</b>	<b>1,0548</b>
e150	ARMA	5,2021	2,4992	1,5836	1,3056	1,2423	1,1928	0,9967	<b>0,9838</b>	1,3518	1,7588	3,2617	4,7288	2,1756
	PG-Pura	1,2339	1,4391	0,6600	1,0475	0,9909	1,2216	1,0032	1,2475	1,2511	1,0563	1,4852	<b>1,1743</b>	1,1509
	B1	1,2771	1,2410	0,6633	1,0068	1,0978	1,1579	1,0191	1,0922	1,2929	1,0258	1,3901	1,2095	1,1228
	BCIGP	<b>1,1904</b>	<b>1,1062</b>	<b>0,5971</b>	<b>0,8918</b>	<b>0,9238</b>	<b>1,0437</b>	<b>0,9265</b>	0,9893	<b>1,1653</b>	<b>0,8910</b>	<b>1,2553</b>	1,4870	<b>1,0389</b>

Na tabela (7.11) estão os resultados da comparação dos cinco algoritmos, utilizando o teste F ANOVA, seguido do teste Tukey-Kramer.

**TABELA 7.11 - RESULTADOS DA ANOVA E TESTE TUKEY-KRAMER PARA A ESTRUTURA MA(1) – PESOS 2**

Valor Previsto	Grupos	Média Grupo	Média Geral	Estatística F	p-valor	Grupos	Dif. Absoluta	Intervalo Crítico	Comp. médias	Melhor Grupo
e136	1	2,3150	2,2592	0,9939	0,4151	Grupo1 e 2	3,2913	7,2062	Iguais	X
	2	5,6063				Grupo1 e 3	1,1832		Iguais	X
	3	1,1318				Grupo1 e 4	1,1504		Iguais	X
	4	1,1646				Grupo1 e 5	1,2369		Iguais	X
	5	1,0781				Grupo2 e 3	4,4745		Iguais	X
						Grupo2 e 4	4,4417		Iguais	X
						Grupo2 e 5	4,5282		Iguais	X
						Grupo3 e 4	0,0328		Iguais	X
						Grupo3 e 5	0,0537		Iguais	X
						Grupo4 e 5	0,0865		Iguais	X
e137	1	1,8223	1,3341	2,3911	0,0565	Grupo1 e 2	0,1943	0,8760	Iguais	X
	2	1,6281				Grupo1 e 3	0,6829		Iguais	X
	3	1,1394				Grupo1 e 4	0,7677		Iguais	X
	4	1,0547				Grupo1 e 5	0,7964		Iguais	X
	5	1,0260				Grupo2 e 3	0,4887		Iguais	X
						Grupo2 e 4	0,5734		Iguais	X
						Grupo2 e 5	0,6021		Iguais	X
						Grupo3 e 4	0,0847		Iguais	X
						Grupo3 e 5	0,1134		Iguais	X
						Grupo4 e 5	0,0287		Iguais	X
e138	1	1,8098	1,2167	5,9412	0,0003	Grupo1 e 2	0,6642	0,5115	Diferentes	2
	2	1,1456				Grupo1 e 3	0,7163		Diferentes	3
	3	1,0935				Grupo1 e 4	0,7590		Diferentes	4
	4	1,0508				Grupo1 e 5	0,8260		Diferentes	5
	5	0,9838				Grupo2 e 3	0,0521		Iguais	X
						Grupo2 e 4	0,0948		Iguais	X
						Grupo2 e 5	0,1618		Iguais	X
						Grupo3 e 4	0,0427		Iguais	X
						Grupo3 e 5	0,1097		Iguais	X
						Grupo4 e 5	0,0670		Iguais	X
e139	1	1,9132	1,2913	7,2474	0,0000	Grupo1 e 2	0,6879	0,4894	Diferentes	2
	2	1,2253				Grupo1 e 3	0,7234		Diferentes	3
	3	1,1898				Grupo1 e 4	0,8820		Diferentes	4
	4	1,0312				Grupo1 e 5	0,8160		Diferentes	5
	5	1,0972				Grupo2 e 3	0,0355		Iguais	X
						Grupo2 e 4	0,1941		Iguais	X
						Grupo2 e 5	0,1281		Iguais	X
						Grupo3 e 4	0,1586		Iguais	X
						Grupo3 e 5	0,0926		Iguais	X
						Grupo4 e 5	0,0660		Iguais	X
e140	1	1,7691	1,2253	6,1981	0,0002	Grupo1 e 2	0,5299	0,4746	Diferentes	2
	2	1,2392				Grupo1 e 3	0,6561		Diferentes	3
	3	1,1130				Grupo1 e 4	0,7715		Diferentes	4
	4	0,9976				Grupo1 e 5	0,7614		Diferentes	5
	5	1,0077				Grupo2 e 3	0,1261		Iguais	X
						Grupo2 e 4	0,2416		Iguais	X
						Grupo2 e 5	0,2315		Iguais	X
						Grupo3 e 4	0,1155		Iguais	X
						Grupo3 e 5	0,1054		Iguais	X
						Grupo4 e 5	0,0101		Iguais	X

**TABELA 7.11 - RESULTADOS DA ANOVA E TESTE TUKEY-KRAMER PARA A ESTRUTURA MA(1) – PESOS 2 (CONTINUAÇÃO)**

Valor Previsto	Grupos	Média Grupo	Média Geral	Estatística F	p-valor	Grupos	Dif. Absoluta	Intervalo Crítico	Comp. médias	Melhor Grupo
e141	1	1,8326	1,2353	5,7674	0,0004	Grupo1 e 2	0,6725	0,5230	Diferentes	2
	2	1,1601				Grupo1 e 3	0,7042		Diferentes	3
	3	1,1285				Grupo1 e 4	0,7895		Diferentes	4
	4	1,0431				Grupo1 e 5	0,8204		Diferentes	5
	5	1,0123				Grupo2 e 3	0,0317		Iguais	X
						Grupo2 e 4	0,1170		Iguais	X
						Grupo2 e 5	0,1478		Iguais	X
						Grupo3 e 4	0,0853		Iguais	X
						Grupo3 e 5	0,1162		Iguais	X
						Grupo4 e 5	0,0309		Iguais	X
e142	1	1,8526	1,2709	5,9546	0,0003	Grupo1 e 2	0,6294	0,5114	Diferentes	2
	2	1,2233				Grupo1 e 3	0,6530		Diferentes	3
	3	1,1996				Grupo1 e 4	0,8434		Diferentes	4
	4	1,0093				Grupo1 e 5	0,7831		Diferentes	5
	5	1,0695				Grupo2 e 3	0,0236		Iguais	X
						Grupo2 e 4	0,2140		Iguais	X
						Grupo2 e 5	0,1537		Iguais	X
						Grupo3 e 4	0,1903		Iguais	X
						Grupo3 e 5	0,1301		Iguais	X
						Grupo4 e 5	0,0603		Iguais	X
e143	1	1,8323	1,2513	7,9020	0,0000	Grupo1 e 2	0,5898	0,4449	Diferentes	2
	2	1,2425				Grupo1 e 3	0,6899		Diferentes	3
	3	1,1424				Grupo1 e 4	0,8116		Diferentes	4
	4	1,0207				Grupo1 e 5	0,8135		Diferentes	5
	5	1,0188				Grupo2 e 3	0,1000		Iguais	X
						Grupo2 e 4	0,2218		Iguais	X
						Grupo2 e 5	0,2237		Iguais	X
						Grupo3 e 4	0,1217		Iguais	X
						Grupo3 e 5	0,1236		Iguais	X
						Grupo4 e 5	0,0019		Iguais	X
e144	1	1,7596	1,2829	4,4160	0,0026	Grupo1 e 2	0,4251	0,5050	Iguais	X
	2	1,3345				Grupo1 e 3	0,5928		Diferentes	3
	3	1,1668				Grupo1 e 4	0,6975		Diferentes	4
	4	1,0621				Grupo1 e 5	0,6682		Diferentes	5
	5	1,0914				Grupo2 e 3	0,1677		Iguais	X
						Grupo2 e 4	0,2724		Iguais	X
						Grupo2 e 5	0,2430		Iguais	X
						Grupo3 e 4	0,1047		Iguais	X
						Grupo3 e 5	0,0753		Iguais	X
						Grupo4 e 5	0,0294		Iguais	X
e145	1	1,8127	1,3226	4,5910	0,0020	Grupo1 e 2	0,3457	0,5488	Iguais	X
	2	1,4670				Grupo1 e 3	0,6275		Diferentes	3
	3	1,1852				Grupo1 e 4	0,7352		Diferentes	4
	4	1,0775				Grupo1 e 5	0,7424		Diferentes	5
	5	1,0703				Grupo2 e 3	0,2818		Iguais	X
						Grupo2 e 4	0,3896		Iguais	X
						Grupo2 e 5	0,3967		Iguais	X
						Grupo3 e 4	0,1078		Iguais	X
						Grupo3 e 5	0,1149		Iguais	X
						Grupo4 e 5	0,0072		Iguais	X

**TABELA 7.11 - RESULTADOS DA ANOVA E TESTE TUKEY-KRAMER PARA A ESTRUTURA MA(1) – PESOS 2 (CONTINUAÇÃO)**

Valor Previsto	Grupos	Média Grupo	Média Geral	Estatística F	p-valor	Grupos	Dif. Absoluta	Intervalo Crítico	Comp. médias	Melhor Grupo
e146	1	1,8306	1,2372	6,3888	0,0001	Grupo1 e 2	0,6768	0,4926	Diferentes	2
	2	1,1538				Grupo1 e 3	0,6957		Diferentes	3
	3	1,1349				Grupo1 e 4	0,8015		Diferentes	4
	4	1,0291				Grupo1 e 5	0,7929		Diferentes	5
	5	1,0377				Grupo2 e 3	0,0189		Iguais	X
						Grupo2 e 4	0,1247		Iguais	X
						Grupo2 e 5	0,1161		Iguais	X
						Grupo3 e 4	0,1058		Iguais	X
						Grupo3 e 5	0,0972		Iguais	X
						Grupo4 e 5	0,0086		Iguais	X
e147	1	1,8601	1,2628	5,5715	0,0005	Grupo1 e 2	0,6267	0,5386	Diferentes	2
	2	1,2334				Grupo1 e 3	0,7233		Diferentes	3
	3	1,1368				Grupo1 e 4	0,7991		Diferentes	4
	4	1,0610				Grupo1 e 5	0,8377		Diferentes	5
	5	1,0224				Grupo2 e 3	0,0966		Iguais	X
						Grupo2 e 4	0,1724		Iguais	X
						Grupo2 e 5	0,2110		Iguais	X
						Grupo3 e 4	0,0758		Iguais	X
						Grupo3 e 5	0,1144		Iguais	X
						Grupo4 e 5	0,0386		Iguais	X
e148	1	1,7664	1,2246	5,6645	0,0004	Grupo1 e 2	0,5604	0,4840	Diferentes	2
	2	1,2059				Grupo1 e 3	0,6777		Diferentes	3
	3	1,0886				Grupo1 e 4	0,7469		Diferentes	4
	4	1,0195				Grupo1 e 5	0,7235		Diferentes	5
	5	1,0428				Grupo2 e 3	0,1173		Iguais	X
						Grupo2 e 4	0,1865		Iguais	X
						Grupo2 e 5	0,1631		Iguais	X
						Grupo3 e 4	0,0692		Iguais	X
						Grupo3 e 5	0,0458		Iguais	X
						Grupo4 e 5	0,0234		Iguais	X
e149	1	1,7232	1,2049	5,8585	0,0003	Grupo1 e 2	0,5623	0,4546	Diferentes	2
	2	1,1609				Grupo1 e 3	0,6037		Diferentes	3
	3	1,1196				Grupo1 e 4	0,7091		Diferentes	4
	4	1,0141				Grupo1 e 5	0,7163		Diferentes	5
	5	1,0069				Grupo2 e 3	0,0414		Iguais	X
						Grupo2 e 4	0,1468		Iguais	X
						Grupo2 e 5	0,1541		Iguais	X
						Grupo3 e 4	0,1055		Iguais	X
						Grupo3 e 5	0,1127		Iguais	X
						Grupo4 e 5	0,0072		Iguais	X
e150	1	1,7436	1,2087	4,8767	0,0013	Grupo1 e 2	0,6070	0,5093	Diferentes	2
	2	1,1366				Grupo1 e 3	0,6500		Diferentes	3
	3	1,0936				Grupo1 e 4	0,6616		Diferentes	4
	4	1,0820				Grupo1 e 5	0,7557		Diferentes	5
	5	0,9879				Grupo2 e 3	0,0430		Iguais	X
						Grupo2 e 4	0,0546		Iguais	X
						Grupo2 e 5	0,1487		Iguais	X
						Grupo3 e 4	0,0116		Iguais	X
						Grupo3 e 5	0,1057		Iguais	X
						Grupo4 e 5	0,0941		Iguais	X

Na última coluna da tabela (7.11) estão representados os grupos que possuem diferença significativa dos demais e o melhor dos grupos, ou seja, aquele que apresentou o melhor desempenho está destacado em negrito. Pode-se verificar que em todos os casos, quando a diferença foi detectada, o algoritmo que obteve o melhor desempenho foi o algoritmo 5, que é o algoritmo *BCIGP* utilizando a atualização de pesos proposta no capítulo IV. Na tabela (7.12) são apresentados os resultados obtidos com o teste F ANOVA seguido do teste Tukey-Kramer, para a estrutura MA(1), utilizando o algoritmo completo proposto no capítulo IV, *BCIGP*. Os resultados obtidos mostram que o método *BCIGP* obteve o melhor desempenho dentre os algoritmos que foram utilizados para fazer esta comparação. Para apenas dois horizontes de previsão, a metodologia não foi considerada significativamente diferente das demais. Em todos os outros casos, o método se mostrou mais eficiente.

**TABELA 7.12 – VALOR - p MA(1) PESOS 2**

MA(1) - Pesos 2		
Valor Previsto	p-valor ANOVA	Melhor Grupo
<b>e136</b>	0,4151	x
<b>e137</b>	0,0565	x
<b>e138</b>	0,0003	5
<b>e139</b>	0,0000	5
<b>e140</b>	0,0002	5
<b>e141</b>	0,0004	5
<b>e142</b>	0,0003	5
<b>e143</b>	0,0000	5
<b>e144</b>	0,0026	5
<b>e145</b>	0,0020	5
<b>e146</b>	0,0001	5
<b>e147</b>	0,0005	5
<b>e148</b>	0,0004	5
<b>e149</b>	0,0003	5
<b>e150</b>	0,0013	5

## 7.4 CONCLUSÕES

Após a realização de todos os testes com séries reais, financeiras e simuladas e analisar os resultados, observou-se que o algoritmo *BCIGP* proposto nesta tese obteve um bom desempenho na tarefa de previsão de séries temporais. Em todos os testes realizados, seja



considerando o *RMSE*, o *MSE* médio na simulação Monte Carlo, o teste “t”, o teste F ANOVA seguido de Tukey-Kramer, o algoritmo mostrou-se bastante eficiente, tendo em todos os casos, obtido erros de previsão menores do que os obtidos pelos demais métodos analisados.

## Capítulo 8

### EXTENSÃO DO MÉTODO BCI À PROBLEMAS DE REGRESSÃO MÚLTIPLA

Neste capítulo é implementada a versão do algoritmo BCI para os problemas de regressão multivariada. Os algoritmos de *Boosting* existentes são propostos para problemas de classificação e de regressão, desta forma, uma comparação do desempenho do algoritmo com os algoritmos de *Boosting* existentes mostra que o método pode ser utilizado, não apenas na previsão de séries temporais, mas em outros tipos de problemas que possam ser refinados através desta técnica. É apresentado um breve histórico relativo dos problemas de regressão e os resultados obtidos através do algoritmo BCI são comparados aos obtidos por outros métodos, tais como Redes Neurais Artificiais, M5 *Model Tree* e os algoritmos de *Boosting* tradicionais.

#### 8.1 BREVE HISTÓRICO SOBRE PROBLEMAS DE REGRESSÃO

Num problema de regressão, o que se deseja é obter valores aproximados de uma variável contínua. Dada uma amostra com variáveis de entrada, ou variáveis preditoras,  $x = \{x_1, x_2, \dots, x_n\}$  e a variável de saída  $y$ , a tarefa de um regressor é encontrar uma função  $y = f(x)$ , dentro do espaço de possibilidades. O problema de regressão difere dos problemas de

classificação pelo fato de que a variável de saída  $y$  nos problemas de regressão é contínua enquanto que nos problemas de classificação é discreta (WEISS; INDURKHYA, 1999). O procedimento tradicional, geralmente utilizado para problemas de regressão, é o problema clássico de mínimos quadrados para regressão (SCHEFFE, 1959). Os métodos de regressão linear simples possuem suas limitações, porém os métodos vêm sendo refinados ao longo dos anos, tornando-se cada vez mais capazes de se ajustar aos dados. Naturalmente os modelos mais complexos possuem um ajuste melhor. Com o avanço da tecnologia dos computadores, e com a possibilidade de se analisar maior volume de dados, o interesse dos pesquisadores tem se voltado para modelos complexos não-lineares, destas pesquisas têm surgido vários métodos novos (EFRON, 1988), tais como *Projection Pursuit* (FRIEDMAN; STUETZLE, 1981) e *Multivariate Adaptive Regression Splines* (MARS) (FRIEDMAN, 1991). Novos métodos têm também sido estudados e desenvolvidos em outras áreas, como Redes Neurais utilizando *Back-Propagation* (McCLELLAND; RUMELHART, 1988).

Muitos problemas reais que podem ser resolvidos através de métodos de regressão, em especial, problemas que envolvem séries temporais multivariadas também podem ser resolvidos através de métodos de regressão múltipla.

Os métodos de regressão através de árvores apresentam um bom desempenho quando comparados a outros métodos de regressão (BREIMAN et al., 1997). As árvores de regressão mostram-se eficientes quando existe um alto grau de dependência entre as variáveis (FRIEDMAN, 1991).

## 8.2 APLICAÇÃO DO MÉTODO BCI À PROBLEMAS DE REGRESSÃO

Após a verificação do bom desempenho do algoritmo BCI na Previsão de Séries Temporais (capítulos VI e VII) o mesmo foi implementado para os problemas de regressão multivariada.

Para efetuar uma avaliação do método, foi utilizada a mesma base de dados do trabalho de Solomatine (2003), retiradas do *UCI repository* (BLAKE, 1998) e do trabalho de Drucker (1997). Desta maneira os resultados puderam ser comparados aos resultados obtidos

através das Redes Neurais Artificiais (multilayer perceptron), *M5 Model Tree* e *Adaboost.RT*. As bases de dados utilizadas foram: *Housing*, *Auto-Mpg* e CPU e Friedman #1.

### 8.2.1 Descrição das Bases de Dados

Os conjuntos de dados foram divididos em treinamento e teste, conforme é apresentado na tabela (8.1).

**TABELA 8.1 – CONJUNTOS DE DADOS**

Conjuntos de Dados	Instâncias	Treinamento (70%)	Teste (30%)
CPU	209	137	72
Housing	506	337	169
Auto-Mpg	398	262	136
Friedman #1	1500	990	510

Na tabela (8.2) está representado o número de variáveis preditoras, número de variáveis não preditoras (classes) e o tipo da variável de saída (Prevista).

**TABELA 8.2 – NÚMERO E TIPO DE VARIÁVEIS**

Base de Dados	No de Atributos (Preditores)	No de Atributos (não Preditores)	Tipo da Variável Prevista
CPU	6	2	Contínua
Housing	13	0	Contínua
Auto_Mpg	7	1	Contínua
Friedman	5	1	Contínua

A base de dados denominada CPU, refere-se à performance relativa de CPUs de 29 fabricantes diferentes de computadores. A base de dados *Housing* possui dados referentes aos valores medianos de moradias no subúrbio da cidade de Boston. A base *Auto-Mpg* contém dados para a previsão do consumo de combustível por milha, medida em galões, de acordo com as características de cada automóvel, tais como número de cilindros, potência do motor, aceleração e outros itens. A base de dados *Friedman* foi gerada de acordo com os trabalhos de Friedman (1991) e Drucker (1997), é um problema de predição não linear, com 10 variáveis

independentes geradas a partir da distribuição uniforme  $[0,1]$  e gerada de acordo com a equação (8.2).

$$y = 10\sin(\pi x_1 x_2) + 20(x_3 - 0.5)^2 + 10x_4 + 5x_5 + n \quad (8.2)$$

Onde  $n$  possui uma distribuição Normal com média zero e desvio padrão 1. Somente 5 variáveis preditoras são necessárias, mas o preditor tem a função de distinguir as variáveis que não possuem habilidade de previsão ( $x_6$  a  $x_{10}$ ) de acordo com as variáveis preditoras ( $x_1$  a  $x_5$ ).

### 8.2.2 Preparação dos Dados

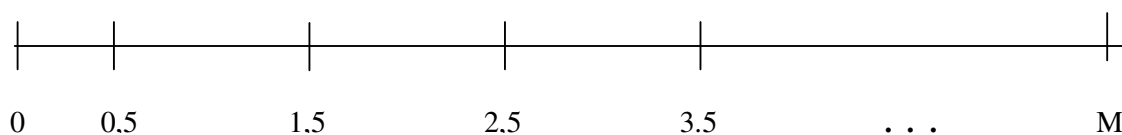
Os conjuntos de treinamento e teste foram selecionados aleatoriamente e sem reposição do conjunto de dados original e de acordo com os dados da tabela (8.1). Foram gerados 10 conjuntos de treinamento e teste. O algoritmo BCI foi implementado utilizando o *software R* e o algoritmo utilizado para gerar as árvores de indução é o *CART (Classification and Regression Tree)* descrito em Breiman et al. (1984).

### 8.2.3 Atualização dos pesos do algoritmo BCI para problemas de regressão

No capítulo V foi apresentado o algoritmo *BCI*, a atualização dos pesos é feita utilizando o coeficiente de correlação, conforme algoritmo apresentado no capítulo IV (eq. 4.11), porém a proposta deste algoritmo foi feita para ser aplicada inicialmente a problemas de previsão com uma única variável. Desta forma foi necessário fazer uma adaptação do algoritmo para que o mesmo pudesse ser aplicado a problemas de previsão com mais de uma variável e para que pudesse ser utilizado em árvores de indução.

A atualização dos pesos segue a mesma metodologia de cálculos exposta no capítulo IV, porém após a obtenção do primeiro vetor de pesos o conjunto de treinamento é modificado de forma que todos os atributos do conjunto de treinamento possam ser atualizados.

Cada exemplo aparecerá no conjunto de treinamento proporcionalmente ao seu peso. Seja  $P_t = \{p_1, p_2, \dots, p_m\}$ , o vetor de pesos para a iteração *Boosting*  $t$ ,  $m$  é o número de amostras do conjunto de treinamento e seja  $r_i$ , uma partição do conjunto  $P$ , com  $i = \{1, \dots, M\}$  onde  $M$  é o máximo do conjunto. Os valores de  $P$  menores que 0.5 são colocados na partição  $r_0$  e o vetor de pesos  $P$ , começa a ser dividido, com intervalos de tamanho um, a partir do ponto 0.5, conforme mostrado na figura (8.1).



**FIGURA 8.1 – PARTIÇÃO DO VETOR DE PESOS**

Cada linha do conjunto de treinamento, foi re-inserida no conjunto de acordo com seu valor de inserção, proporcionalmente a  $\alpha$ . Os valores de  $\alpha$ , são estabelecidos de acordo com o intervalo a que  $p_i$  pertence. Os valores estão apresentados na tabela (8.3).

**TABELA 8.3 – CONJUNTOS DE DADOS**

Intervalo	$\alpha$
0,0 - 0,5	--
0,5 - 1,5	1
1,5 - 2,5	2
2,5 - 3,5	3
3,5 - 4,5	4
.	.
.	.

O novo conjunto de treinamento gerado é então utilizado para gerar o novo modelo de regressão. Com o modelo obtido, são gerados os valores de previsão para o conjunto de teste. Este procedimento foi repetido 10 vezes para cada conjunto de treinamento e teste, isto é, foi utilizada validação cruzada com 10 subconjuntos (*folds*). O número de iterações de *Boosting* também é 10. Assim, foram executados 10 algoritmos de *Boosting* para cada um dos 10 conjuntos de treinamento. Após terem sido realizadas as 10 execuções de *Boosting*, os valores previstos são combinados de acordo com a equação proposta no algoritmo *BCI* (Quadro 4.3).

Os resultados foram comparados aos obtidos por *MT* (*Model Tree*), *ANN* (*Artificial Neural Network*), *Bagging*, *ANN*, *Adaboost.R* e *Adaboost.RT*. Na tabela (8.4) são

apresentados os resultados obtidos da média dos RMSE de 10 diferentes conjuntos, independentes de dados.

**TABELA 8.4 – COMPARAÇÃO DO RMSE NOS 10 CONJUNTOS DE TESTE**

<b>Bases</b>	<b><i>BCI</i></b>	<b><i>MT</i></b>	<b><i>Bagging</i></b>	<b><i>ANN</i></b>	<b><i>AdaBoost.R</i></b>	<b><i>AdaBoost.RT</i></b>
CPU	<b>22,48</b>	34,65	32,64	<b>13,91</b>	24,45	26,52
Housing	<b>1,12</b>	3,62	3,24	3,54	3,23	3,23
Auto-Mpg	<b>0,85</b>	3,01	2,86	3,79	2,84	2,96
Friedman #1	<b>0,70</b>	2,19	2,06	1,51	1,82	1,72

Como se pode observar na tabela (8.4), o algoritmo *BCIGP* apresenta o menor *RMSE* dos conjuntos de teste, exceto para a base de dados CPU, na qual o algoritmo *ANN* apresenta um *RMSE* menor, porém apresenta resultado melhor quando, comparado aos demais algoritmos de *boosting*.

## Capítulo 9

### CONCLUSÃO E TRABALHOS FUTUROS

Neste trabalho foi proposta uma nova metodologia de *Boosting* utilizando a Programação Genética como algoritmo base e os coeficientes de correlação entre os valores observados e os valores previstos através da PG. O algoritmo *BCI* proposto nesta tese foi inicialmente desenvolvido, para a previsão de séries temporais. Os primeiros experimentos realizados utilizando os coeficientes de correlação mostraram a acurácia da técnica na tarefa de previsão. Assim a metodologia foi sendo aprimorada e foram realizados testes no sentido de validar a técnica também em séries geradas artificialmente, nas estratégias de *trading* para séries financeiras e para os problemas de regressão multivariada.

#### 9.1 CONCLUSÃO

A relevância da metodologia proposta deve-se, especialmente ao fato, de que os resultados obtidos são testados fora da amostra em todos os casos, e são melhores do que os obtidos pelos demais métodos que foram testados para realizar a mesma tarefa. Em geral, os modelos obtidos, analisando os conjuntos de treinamento, fornecem modelos muito bons que se ajustam bem aos dados, porém ao serem aplicados a novos exemplos não produzem os mesmos resultados.

O algoritmo foi testado para realizar previsões nos seguintes experimentos:



- séries reais encontradas na literatura da previsão de séries temporais;
- séries financeiras, tendo sido feita a análise de *trading* baseada nos resultados de previsão obtidas através do procedimento proposto;
- simulação de séries artificiais utilizando uma Simulação Monte Carlo, na qual foram testadas 214.000 séries temporais, contendo 150 valores observados, tendo sido utilizado para o conjunto de treinamento 90% dos dados observados e 10% para teste. O modelo gerado no conjunto de treinamento foi utilizado para a gerar os valores de previsão do conjunto de teste. As séries foram geradas com ruído que possui distribuição normal com média zero e desvio padrão um;
- uma extensão da metodologia *BCI* foi proposta para problemas de regressão multivariada e os resultados foram comparados aos obtidos por outras técnicas, tendo apresentado resultados melhores.

A partir da realização destes experimentos, conclui-se que a metodologia *BCI* proposta mostrou-se bastante eficiente na tarefa de previsão. Os resultados obtidos foram comparados a outros métodos de previsão utilizado a Programação Genética tradicional, o algoritmo *GPBoost* e a metodologia Box & Jenkins. Além de terem sido comparados os erros de previsão obtidos (*RMSE* e *MSE*), foram também aplicados testes de hipóteses para a avaliação do algoritmo proposto. Em todas as análises, o algoritmo mostrou ser uma técnica alternativa para problemas de previsão por apresentar resultados que apresentam erros de previsão menores do que os erros obtidos através de outras técnicas, mesmo quando o horizonte de previsão considerado é grande.

## 9.2 FUTUROS TRABALHOS

Como sugestão para futuros trabalhos sugere-se:

- A implementação de um critério de seleção dos preditores antes de ser aplicada a metodologia proposta, pois observou-se, nos experimentos que existem preditores ruins, que acabam por influenciar nos resultados finais. A geração destes preditores ruins deve-se às sementes aleatórias que são utilizadas na PG. A utilização de mais de uma semente inicial para a PG, poderia ser uma forma para resolver este problema;

- O estudo preliminar do tipo de série a ser analisada, como verificação de tendência e sazonalidade, poderia ser uma boa estratégia para a escolha do número de variáveis a serem utilizadas na geração do modelo de PG, porém isto poderia acarretar um custo computacional elevado;
- A implementação da previsão um passo à frente, para aplicação ao mercado financeiro, já que os métodos ingênuos de previsão utilizam este critério de previsão e os resultados obtidos com a utilização destes modelos são muito bons, com erros de previsão pequenos;

## REFERÊNCIAS

- AKAIKE, H., **Maximum likelihood identification of Gaussian autoregressive moving average models**. Biometrika, v. 60, p. 255-265, 1973.
- AKAIKE, H., **A new look at the statistical model identification**. IEEE Transactions on Automatic Control, AC-19, p.716-723, 1974.
- AKAIKE, H., **On entropy maximization principle**. In: Applications of Statistics, Amsterdam. Ed. P.R. Krishnaiah. p. 27-41, 1977.
- ALLWEIN, E. L, SCHAPIRE, R. E. & SINGER, Y., **Reducing multiclass to binary: A unifying approach for margin classifiers**. Journal of Machine Learning Research, v.1, p. 113-141, 2000.
- BANZHAF; W. NORDIN, P.; KELLER, R. E. & FRANCONI, F. D., **Genetic Programming an introduction**. Morgan Kaufmann, 1998.
- BARBANCHO, A. G., **Fundamentos e possibilidades da econometria**. Rio de Janeiro. Fórum Editora, p. 18 – 32, 1970.
- BICKLE, Y. & THIELE, L., **A Mathematical Analysis of Tournament Selection**. In: ICGA95. L. J. Eshelman. Ed. San Francisco: Morgan Kaufmann Publishers, p. 9-16, 1995.
- BLAKE, C. L. & MERZ, C. J., **UCI Repository of machine learning databases**, Irvine, CA: University of California, Dep. of Information and Computer Science, 1998. Available: <http://www.ics.uci.edu/~mllearn/MLRepository.html>. Consulta em agosto de 2006.
- BOHM, W. & GEYER-SCHULZ, **A Exact uniform initialization for genetic program**. Foundations of Genetic Algorithms IV. p 379-407. Morgan Kaufmann, 1996.
- BOX, G. E. P & JENKINS G. M., **Time series analysis: forecasting and control**, Ed. Rev.(1976). San Francisco: Holden-Day, 1970.
- BOOSTING. Disponível em: <http://www.boosting.org/applications>, 2006. Acesso em agosto de 2006.
- BREIMAN, L., FRIEDMAN, J., OLSHEN, R. & STONE, C., **Classification and Regression Trees**, Wadsworth, Monterrey, CA, 1984.
- BREIMAN L. **Prediction Games and Arcing Algorithms** . Neural Computation, v. 11, n. 7, p. 1493-1518, 1997.
- CHATFIELD, C., **The analysis of time series**. Chapman & Hall, New York, 6th Ed. 2000.
- CHAVES, A. N., **Bootstrap em Séries Temporais**. Rio de Janeiro. 208f. Tese (Doutorado em Engenharia Elétrica), Universidade Federal do Rio de Janeiro, 1991.

CHELLAPILLA, K., **Evolutionary programming with tree mutations: Evolving computer programs without sub-tree crossover**. Genetic Programming. In: Proc. Second Annual Conference of San Francisco, CA. Morgan Kauffmann, p. 432-438, 1997.

CORRÊA, H. L.; GIANESI, I. G. N. & CAON M., **Planejamento, programação e controle da produção: MRP II/EPR: conceitos, uso e implantação**. 3ª ed. São Paulo: Giancesi Corrêa & associados: Atlas, 2000.

CORTEZ, P. A. R., **Modelos Inspirados na Natureza para a Previsão de Séries Temporais**. Braga, 2002. 188 f. Tese (Doutorado em Informática) – Departamento de Informática, Universidade do Minho.

DARWIN, C., **A origem das Espécies e a Seleção Natural**. Ed. Hemus, 5ª Ed., 2000.

DRUCKER H., **Improving regression using boosting techniques**. In: International Conference on Machine Learning, 1997, Orlando: Proceeding of International Conference on Machine Learning, ICML97.

DRUCKER H., **Boosting using Neural Networks**. In A. J. C, Sharkey (ed.), In Combining Artificial Neural Nets. p. 51-77, London, Springer-Verlag, 1999.

DUAN, M. & POVINELLI, R. **Estimating Time Series Predictability Using Genetic Programming**. In: Cihan H. Dagli *editor*, Intelligent Engineering Systems Through Artificial Neural Networks (ANNIE 2001), pages 215-220, St. Louis, Missouri, USA, 2001

DUFFY, N. & HELMBOLD, D. P., **Leveraging for regression**. Procedures of the 13th Annual Conference on Computer Learning Theory, p. 208-219, San Francisco: Morgan Kaufmann, 2000.

DUNIS, C & JALILOV, J., **Neural Network regression and alternative forecasting techniques for predicting financial variables**. Neural Network World, v. 2, p. 113-139, 2002.

ECONOMATICA, Disponível em: <http://www.economatica>, 2006. Acesso em agosto de 2006. [www.economatica.com](http://www.economatica.com)

EFRON, B., **Computer-intensive Methods in Statistical Regression**. SIAM Review, v. 30 n. 3, 421-449, 1988.

FREUND Y. & SCHAPIRE, R. E., **Experiments with a new boosting algorithm**. Machine Learning. Proceedings of the Thirteenth Conference, ed: L. Saitta, Morgan Kaufmann, p. 148-156, 1996.

FREUND Y. & SCHAPIRE, R.E., **A decision-theoretic generalization of on-line learning and an application to boosting**. Journal of Computer and System Sciences, v. 55 n. p. 119-139, 1997.

FRIEDMAN, J., **Multivariate Adaptive Regression Splines**. Annals of Statistics, v.19 n.1, p.1-141, 1991.

FRIEDMAN, J.; HASTIE, T.; TIBSHIRANI, R. **Additive Logistic Regression: a Statistical view of boosting**. Thecnical Report . Stanford University Statistics Department, 1998.

FRIEDMAN, J.; HASTIE, T. & TIBSHIRANI. **Greedy Function Approximation: A Gradient Boosting Machine**. In: Annals of Statistics, v. 29, n. 5, p. 1189-1232, 2001.

GECCO, Genetic and Evolutionary Computation Conference. 2002, 2006.

GREFENSTETTE, J. J. & BAKER, J. E., **How genetic algorithms work: A critical look at implicit parallelism**. In Proc. 3<sup>rd</sup> International Conference on Genetic Algorithms, p. 20-27. San Mateo. CA. Morgan Kaufmann. San Francisco, CA, 1989.

HAMILTON, J. D., **Time Series Analysis**. Princeton University Press. New Jersey, 2004.

HASTIE, T.; TIBSHIRANI, R. & FRIEDMAN, J., **The elements of Statistical Learning**. Springer Science+Business Media. New York, 10013, USA, 2001.

HOLLAND J. H., **Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence**. Ann Arbor, MI: University of Michigan Press. 1975.

HORNER, H. **A C++ class library for genetic programming**. Technical Report. The Vienna University of Economics, Vienna, Austria, 1996.

HUI, A., **Using Programming to perform Time-Series Forecasting of Stock Prices**. Book: Genetic Algorithms and Genetic Programming. Editor: John Koza, p. 83-90. Stanford, California. 2003.

IBA H., **Bagging, boosting, and bloting in genetic programming**. In: BDE, p. 1053-1060, 1999.

KABOUDAN, M., **A Measure of Time Series Predictability Using Genetic Programming Applied to Stock Returns**. Journal of Forecasting, v. 18, p. 345 - 357, 1999.

KABOUDAN, M., **Genetic Programming Prediction of Stock Prices**. Computational Economics. v. 16, n.3, p. 207 – 236, 2000.

KABOUDAN, M., **TSGP – Time Series Genetic Programming**. School of Business, university of Redlands, 2002. [www.compumetrica.com](http://www.compumetrica.com)

KOZA, J. R. **Hierarquical genetic algorithms operating on populations of computer programs**. Proceedings of the 11<sup>th</sup> International Joint Conference on Artificial Intelligent (IJCAI-89). Detroit, MI. Pp 768-774. Morgan Kaufmann, 1989.

KOZA, J. R., **Genetic Programming: On the Programming of Computers by Means of Natural Selection**. MIT Press, 1992.

KRAMER, M. D. & ZHANG, D. **A Genetic Programming System**. In: The 24<sup>th</sup> Annual International Computer Software and Applications Conference, p. 614-619, IEEE Press, 2000.

LEVINE, D. M.; STEPHAN, D.; KREHBIEL, T. C. & BERENSON, M. L.; **Statistics for Managers using MS EXCEL**, 3<sup>rd</sup> Edition, 2002.

LUKE, S., **Two fast tree-creation algorithms for genetic programming**. IEEE Transactions in Evolutionary Computation, v. 4, n.3, p. 274-283. IEEE Press. September, 2000.

LUKE, S. & PAINAT, L., **A survey and comparison of tree generation algorithms**. Proceedings of the 6th Annual Conference in Genetic Programming (GECCO 2001). Springer-Verlag, 2001.

MAKRIDAKIS, S. & WHEELWRIGHT, S., **Forecasting Methods for Management**. John Wiley & Sons, New York, Fifth edition, 1989.

MAKRIDAKIS, S. WHEELWRIGHT, S. C. & McGEE, V. E. **Forecasting: Methods and Applications**. 2nd ed. New York: John Wille & Sons, 1983.

McCLELLAND, J. & RUMELHART, D., **Explorations in Parallel Distributed Processing**, MIT Press, Cambridge, Ma, 1988.

MICHALEWICZ, Z. **Genetic Algorithms + Data Structures = Evolution Programs**. IE-Springer-Verlag, 3<sup>rd</sup> Ed. New York, 1997.

MORETTIN; TOLOI, P. A. & TOLOI, C. M. C. **Modelos para previsão de séries temporais**. In : 13<sup>o</sup> Colóquio Brasileiro de Matemática. Rio de Janeiro, 1981.

MORETTIN; TOLOI, P. A. & TOLOI, C. M. C. **Análise de séries temporais**. Ed. Edgard Blucher LTDA. São Paulo, 2004.

MUHLENBEIN, H. & SCHIERKAMP-VOOSEN, D. **Predictive models for the breeder genetic algorithms**. Evolutionary Computation, v.1, n.1, p. 25-49. MIT Press, 1993.

OLIVEIRA, S. R., **Sistemas Inteligentes: fundamentos e aplicações**. p. 54-56. Barueri: Monole, 2003.

PARIS; ROBILLIARD; FONLUPT, G.; ROBILLIARD D. & FONLUPT C. **Applying Boosting Techniques to Genetic Programming**. Lecture Notes In Computer Science, London, v. 2310, p. 267-280, 2001.

POVINELLI, R. J., **Identifying Temporal Patterns for Characterization and Prediction of Financial Time Series Events**. Ph.D. Dissertation, Marquette University, 1999.

QUINLAN, J. R. **Bagging, boosting e C4.5.**, Procedures of the 13th National Conference on Artificial Intelligence, Portland, OR, p.725-730, 1996.

RISSANEN, J., **Modelling by shortest data description.** Automatica, v.14, p.465-471, 1978.

RODRIGUES, E. L. M. **Evolução de funções em programação genética orientada a gramáticas.** Dissertação (Mestrado em Informática). Universidade Federal do Paraná, 2002.

SANTOS P. M.; LUDEMIR, T. B. & PRUDÊNCIO, R.B.C. **Seleção de Modelos de Previsão de Séries Temporais baseada em Informações de Performance.** In: VIII Simpósio Brasileiro de Redes Neurais, São Luis, MA, 2004.

SCHAPIRE R. E., **The strenght of weak learnability.** In: Machine Learning, p.197-227, 1990.

SCHAPIRE R. E. & FREUND Y. **Experiments with a new boosting algorithm.** In: MACHINE LEARNING: Proceedings of the Thirteenth International Conference, p.148-156, 1996.

SCHAPIRE R. E. & SINGER., **Improved boosting algorithms using confidence – rated predictions.** In: Proceedings of the Eleventh Annual Conference on Computational Learning Theory, p.80 – 91, 1998.

SCHAPIRE R. E., **Theoretical views of boosting.** In: Computational Learning Theory: Fourth European Conference, EuroCOLT99, p.1-10, 1999.

SCHAPIRE R. E. & FREUND, Y., **A short introduction to boosting.** Journal of Japanese Society for Artificial Intelligence, v.14, n.5, p.771-780, 1999b.

SCHEFFE, H., **The Analysis of Variance.** John Wiley, New York, 1999.

SCHWEFEL, H., **Evolution and optimum seeking.** Sixth-Generation Computer Technology Series. John Wiley & Sons. New York, 1995.

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

SOLOMATINE, D. P. & DULAL, K. N., **Model trees an alternative to neural networks in rainfall-runoff modelling.** Hydrological Sciences. Journal-des Sciences Hydrologiques, v. 48 n.3, p. 339-441, 2003

SOLOMATINE, D. P. & SHRESTHA, D. L., **Adaboost.RT: a Boosting Algorithm for Regression Problems.** IEEE, p. 1163-1168, 2004.

SOUZA, L. V.; COSTA, E. O. & POZO, A. T. R., **Análise da Capacidade da Programação Genética na Previsão de Séries Temporais.** In: Congresso de Métodos Numéricos en Ingeniería, Anais do Congresso de Métodos Numéricos en Ingeniería. Granada, Espanha, p. 18, 2005a.

SOUZA, L. V.; COSTA, E. O. & POZO, A. T. R., **Previsão de Séries Temporais utilizando Programação Genética**. In: XXXVII Simpósio Brasileiro de Pesquisa Operacional. Gramado, RS, Brasil, setembro 2005b.

TERADA, P. D., **Compilers and Compiler Generators, an introduction with C++**. McGraw-Hill, Makron, 1991

TRENTIN, M. G., **Planejamento Estratégico de Materiais: Uma aplicação dos métodos automáticos de previsão de séries temporais**. Curitiba, 2002. Dissertação (Mestrado em Métodos Numéricos em Engenharia). UFPR.

WEISS, S. M. & INDURKHYA, N. **Estimating performance for voted decision trees**. IBM Research Division Technical Report. In: Intelligent data Analysis (IDA), 1999.

WHITLEY, D., **The genitor algorithm and selection pressure: Why rank-based allocation of reproductive trial is best**. In: Schaffer, J. D., editor, Proc. 3<sup>rd</sup> Int. Conference on Genetic Algorithm, pp 116-121, San Mateo, CA. Morgan Kaufmann, San Francisco, CA, 1989.

WOLD, H., **A Study in the ysis of Stationary Time Series**. Almgvist & Wiksell. 1<sup>st</sup>. ed., Stocolm, 1938.

ZEMEL, R. & PITASSI, T., **A gradient-based boosting algorithm for regression problems**. Leen, T. K. Dietterich, T. G., & Tresp, V. (Eds.), Advances in Neural Information Processing Systems 13. MIT press, 2001.

ZOU H. & YANG, Y., **Combining time series models for forecasting**. International Journal of forecasting, v.20, p. 69-84, 2004.

ZONGKER, D. & PUNCH, B., **Lil-gp 1.0 User's manual**. Michigan State University, USA, 1995.



## APÊNDICE A

### CRITÉRIO DE INFORMAÇÕES DE AKAIKE (AIC)

O critério de informações de Akaike é um procedimento de identificação do modelo a ser utilizado para uma determinada série. Em seu trabalho, Akaike (AKAIKE, 1973/1974) sugere que o modelo a ser escolhido seja aquele cujas ordens  $k$  e  $l$  minimizem o critério.

$$AIC(k, d, l) = N \ln \sigma_\varepsilon^2 + \frac{N}{N-d} 2(k+l+1+\sigma_{d0}) + N \ln 2\pi + N \quad (\text{A.1})$$

Onde:

$$\sigma_{d0} = \begin{cases} 1, d = 0 \\ 0, d \neq 0 \end{cases}$$

e  $\hat{\sigma}_\varepsilon^2$  é o estimador de máxima verossimilhança de  $\sigma_\varepsilon^2$ . Quando a comparação é feita entre muitos modelos, os dois últimos termos da equação (1) podem ser eliminados. Assim, considerando-se a série apropriadamente diferenciada, o critério de Akaike passa a ser representado pela equação (A.2)

$$AIC(k, l) = N \ln \sigma_a^2 + 2(k+l+2) \quad (\text{A.2})$$

como critério para determinação das ordens  $p$  e  $q$  do modelo ARMA a ser utilizado. Em seguida, designa-se valores para  $K$  e  $L$  que serão seus limites superiores e são realizadas todas as combinações possíveis de  $k$  e  $l$ , com  $0 \leq k \leq K$  e  $0 \leq l \leq L$ , geralmente  $K$  e  $L$  são funções de  $N$ , por exemplo  $K = L = \ln(N)$ .

Reescrevendo a equação (A.1), tem-se:

$$AIC(k, l) = \ln \sigma_a^2 + 2(k+l) \quad (\text{A.3})$$

Os valores que minimizam (A.1) são os mesmos que minimizam (A.3). Dependendo dos valores estabelecidos para  $K$  e  $L$ , muitos ajustes deverão ser feitos a fim de se obter o

mínimo de AIC. Alguns exemplos de aplicação do critério AIC podem ser encontrados em (MORETTIN; TOLOI, 1979).

No caso dos modelos auto-regressivos de ordem  $p$ ,  $AR(p)$ , o critério AIC reduz-se a:

$$AIC(k) = \ln \hat{\sigma}_k^2 + 2K, k \leq K \quad (\text{A.4})$$

Em 1976, Shibata demonstra que o critério MAICE fornece estimativas inconsistentes do processo AR. Em 1980, Hannan generalizou o resultado de Shibata para o processo ARMA  $(p, q)$  (MORETTIN; TOLOI, 2004). Várias correções foram propostas para melhorar o processo AIC, de forma que o modelo selecionado não possua uma ordem superior à necessária para a série que está sendo analisada. Uma destas correções foi proposta por Hurvich e Tsai em 1989, sua proposta apresenta o critério da forma apresentada na equação (A.5).

$$AIC_c(k) = AIC(k) + \frac{2(k+1)(k+2)}{N-k+2}, k \leq K \quad (\text{A.5})$$

Eles mostraram que quando  $N$  é pequeno, ou quando  $K$  é uma fração “moderadamente grande” de  $N$ , esta correção é útil.

Em 1979, Akaike propõe uma extensão desta correção que está representada na equação (A.6)

$$AIC_\alpha(k) = N \ln \hat{\sigma}_k^2 + \alpha K, k \leq K \quad (\text{A.6})$$

onde  $\alpha$  é uma constante.

Em 1977, Tong apresenta uma discussão na qual mostra que, assintoticamente, a probabilidade de selecionar um modelo de ordem correta, quando se minimiza  $AIC_\alpha(k)$ , aumenta quando  $\alpha$  cresce. Além disso, Hannan mostra em 1980 que o critério  $AIC_\alpha(k)$  com  $\alpha = \alpha(N)$  é fortemente consistente para qualquer  $\alpha(N) > 2 \ln \ln(N)$  (MORETTIN; TOLOI, 2004).

## APÊNDICE B

### DESCRIÇÃO DAS SÉRIES REAIS

Neste apêndice será feita uma descrição das séries reais utilizadas no primeiro experimento do capítulo VI. Os dados destas séries temporais podem ser obtidas em <http://www.ime.usp.br/~pam/ST.html>

**Atmosfera** – Registram as temperaturas em graus centígrados ( $^{\circ}\text{C}$ ). Observações diárias tomadas às 12:00 na cidade de São Paulo, durante o período de 1 $^{\circ}$  de janeiro a 31 de dezembro de 1997;

**Bebida** – Produção física industrial (PIB). Observações mensais, no período de janeiro de 1985 a julho de 2000.

**Consumo** – Vendas físicas na região metropolitana de São Paulo. Observações mensais no período de janeiro de 1984 a outubro de 1996.

**ICV** – Índice do custo de vida no município de São Paulo. Observações mensais de janeiro de 1970 a junho de 1980.

**Lavras** – Precipitação atmosférica no município de Lavras – MG. Observações mensais de janeiro de 1966 a dezembro de 1997.

**Manchas** – Número de Manchas solares de Wolfer. Observações anuais de 1724 a 1924.

## APÊNDICE C

Neste apêndice são apresentadas as tabelas com *MSE* médio das 500 séries para cada um dos parâmetros e para os 15 horizontes de previsão. Algumas linhas e colunas destas tabelas foram suprimidas, meramente para fins de apresentação.

**TABELA C1 – RESULTADOS DA DO *MSE* MÉDIO PARA AS 500 SÉRIES DA ESTRUTURA MA(1)**

<i>MSE</i>	MA(1)	P1	P2	P3	P4	P5	P8	P14	P15	P16	P17	P18	P19	Média
e136	ARMA	1,8073	1,4279	1,4798	1,1015	<b>0,9595</b>	1,0654	2,5044	2,1605	3,3434	3,0012	4,1974	6,8773	2,4938
	PG-Pura	83,7022	0,9901	1,0939	1,9166	1,3962	1,1162	0,9151	1,1270	1,2577	1,2491	1,1470	1,5686	8,1233
	B1	<b>0,9301</b>	0,9874	1,1435	0,9163	1,3913	<b>1,0444</b>	0,9698	0,9312	1,1945	1,2338	1,0709	1,5208	1,1112
	BCIGP	1,2387	<b>0,8402</b>	<b>1,0595</b>	<b>0,8421</b>	1,4341	1,5482	<b>0,8914</b>	<b>0,8268</b>	<b>1,0711</b>	<b>1,1036</b>	<b>0,9531</b>	<b>1,3918</b>	<b>1,1000</b>
e137	ARMA	5,1475	2,1638	2,4561	1,8508	<b>1,0263</b>	1,0737	1,1365	1,0127	1,7992	1,8284	2,3638	4,7337	2,2161
	PG-Pura	<b>0,9462</b>	0,9954	1,3734	1,0785	1,5078	0,9730	1,0055	1,0341	1,0529	1,4271	1,0728	1,4601	1,1606
	B1	1,0820	0,9628	1,2975	1,0032	1,5524	0,9350	0,9109	0,9340	0,9361	1,3293	0,8779	1,4497	1,1059
	BCIGP	0,9472	<b>0,8170</b>	<b>1,2490</b>	<b>0,8915</b>	1,3497	<b>0,8249</b>	<b>0,8280</b>	<b>0,8111</b>	<b>0,8158</b>	<b>1,2105</b>	<b>0,8778</b>	<b>1,3310</b>	<b>0,9961</b>
e138	ARMA	5,4861	2,3233	1,9689	1,6546	<b>1,0542</b>	1,1271	1,1794	1,3157	1,2916	2,5770	2,3512	4,8036	2,2611
	PG-Pura	1,0459	1,1426	1,0960	0,9468	1,3490	1,0036	1,0626	1,0981	0,9296	1,1880	0,7776	1,5345	1,0979
	B1	0,9858	1,2715	1,0814	1,0632	1,2488	1,0030	0,9955	1,1078	0,8305	1,1830	0,7558	1,4739	1,0834
	BCIGP	<b>0,8867</b>	<b>1,0620</b>	<b>0,9899</b>	<b>0,9182</b>	1,2210	<b>0,8959</b>	<b>0,8891</b>	<b>0,9555</b>	<b>0,8041</b>	<b>1,0970</b>	<b>0,6849</b>	<b>1,3110</b>	<b>0,9763</b>
e139	ARMA	5,4291	2,4519	2,1267	1,9668	<b>1,3012</b>	1,1654	1,0526	1,4367	1,4842	2,6675	2,3068	4,6677	2,3381
	PG-Pura	1,5529	1,0655	1,4009	1,2177	1,7326	1,0677	0,9415	0,9938	1,1197	1,2947	0,9937	1,2537	1,2195
	B1	1,3865	1,0489	1,1526	1,4378	1,7777	1,0334	0,9004	0,9908	1,1281	1,0822	1,0119	<b>1,1511</b>	1,1751
	BCIGP	<b>1,2215</b>	<b>0,9205</b>	<b>1,0820</b>	<b>1,1256</b>	1,6619	<b>0,9416</b>	<b>0,7871</b>	<b>0,8832</b>	<b>1,0108</b>	<b>0,9882</b>	<b>0,9176</b>	1,7575	<b>1,1081</b>
e140	ARMA	5,0006	2,4324	1,8302	1,5192	<b>1,5237</b>	1,1011	1,1258	1,1765	1,3098	2,2871	2,6425	4,4513	2,2000
	PG-Pura	1,0142	1,1759	1,3514	1,0117	1,9032	1,3695	1,1182	1,2931	1,2297	1,0487	0,9388	1,2301	1,2237
	B1	1,0414	1,1470	1,2000	1,0589	1,8265	1,1717	1,0515	1,0470	1,1097	0,9744	0,9017	1,0432	1,1311
	BCIGP	<b>1,0091</b>	<b>1,0223</b>	<b>1,0955</b>	<b>0,9011</b>	1,6721	<b>1,0545</b>	<b>0,9209</b>	<b>0,9330</b>	<b>0,9964</b>	<b>0,8255</b>	<b>0,8469</b>	<b>0,9990</b>	<b>1,0230</b>
e141	ARMA	4,9626	2,4068	1,7932	1,5148	<b>1,0847</b>	1,2202	0,9705	1,0564	1,3167	2,3317	2,7971	5,5159	2,2475
	PG-Pura	0,8320	1,2880	1,0470	0,9410	1,6660	1,1933	0,9140	0,9568	1,1688	0,9909	1,1870	1,5593	1,1453
	B1	0,8408	1,3203	1,0223	1,0827	1,5798	1,1558	0,8759	0,9242	1,0887	0,9998	1,0766	1,3470	1,1095
	BCIGP	<b>0,7579</b>	<b>1,1399</b>	<b>0,9320</b>	<b>0,8564</b>	1,5489	<b>1,0625</b>	<b>0,8450</b>	<b>0,8023</b>	<b>0,9717</b>	<b>0,9486</b>	<b>0,9903</b>	<b>1,1948</b>	<b>1,0042</b>
e142	ARMA	4,8753	2,3547	1,7383	1,6458	<b>1,2775</b>	1,3871	1,4036	1,1277	1,5628	2,4785	2,2297	5,5081	2,2991
	PG-Pura	0,8437	<b>0,7989</b>	0,9757	1,1336	1,7653	1,1572	1,3311	1,2063	1,4157	1,3279	0,8899	1,4852	1,1942
	B1	0,9010	0,9403	0,9446	1,3329	1,9855	1,0519	1,2729	1,0693	1,3635	1,3389	0,8536	1,5304	1,2154
	BCIGP	<b>0,7259</b>	0,8443	<b>0,8607</b>	<b>1,0923</b>	1,7769	<b>0,9667</b>	<b>1,1398</b>	<b>0,9461</b>	<b>1,1583</b>	<b>1,2022</b>	<b>0,7737</b>	<b>1,3899</b>	<b>1,0731</b>
e143	ARMA	4,4121	2,7973	1,8116	1,2313	1,6922	1,4455	1,3716	1,4336	1,3964	2,1014	2,3998	4,8008	2,2411
	PG-Pura	0,9978	1,1547	1,0482	0,8313	1,8248	1,2684	1,2622	1,0756	1,1279	0,8636	1,1588	1,3249	1,1615
	B1	1,0699	1,1402	1,0001	0,8445	1,9419	1,2216	1,0630	1,0834	1,0575	0,8534	1,0786	1,1860	1,1283
	BCIGP	<b>0,9404</b>	<b>0,9738</b>	<b>0,8780</b>	<b>0,7905</b>	<b>1,6834</b>	<b>1,1271</b>	<b>0,9476</b>	<b>0,9798</b>	<b>0,9658</b>	<b>0,7751</b>	<b>1,0155</b>	<b>1,0889</b>	<b>1,0138</b>
e148	ARMA	4,7903	2,2956	1,9993	1,6461	<b>1,2517</b>	1,0316	1,1270	1,2831	1,2952	1,8431	2,7011	4,7902	2,1712
	PG-Pura	<b>1,2806</b>	0,7871	1,0173	1,1691	1,4820	0,9537	1,3551	1,0228	0,9725	1,0939	1,4778	1,2213	1,1528
	B1	1,3768	0,8934	1,0797	1,2342	1,3999	0,8757	<b>0,9988</b>	0,9730	0,8833	1,0119	1,2769	1,1385	1,0952
	BCIGP	1,3425	<b>0,7820</b>	<b>0,9593</b>	<b>1,0232</b>	1,3084	<b>0,7744</b>	1,9210	<b>0,8919</b>	<b>0,8584</b>	<b>0,9288</b>	<b>1,1487</b>	<b>1,0354</b>	<b>1,0812</b>
e149	ARMA	4,1228	2,0101	1,8843	1,5562	<b>1,1588</b>	0,9320	1,1669	1,1122	1,4404	1,8083	2,8186	5,0569	2,0890
	PG-Pura	0,9435	0,7197	1,0674	1,0787	1,6308	0,9913	1,2977	0,8525	1,0803	1,1768	1,2285	1,5380	1,1338
	B1	1,0325	0,7453	1,0224	0,9994	1,6732	0,8815	1,2091	0,7962	1,0880	1,1447	1,1746	1,4711	1,1032
	BCIGP	<b>0,8882</b>	<b>0,6694</b>	<b>0,9476</b>	<b>0,9509</b>	1,4910	<b>0,8279</b>	<b>1,0595</b>	<b>0,7270</b>	<b>0,9969</b>	<b>1,0401</b>	<b>1,0600</b>	<b>1,4409</b>	<b>1,0083</b>
e150	ARMA	4,9377	2,4864	1,5950	1,3445	<b>1,1586</b>	1,1936	1,1953	1,6063	1,3518	1,7588	3,2617	4,7427	2,2194
	PG-Pura	1,2370	1,3181	0,7286	0,9478	1,3618	1,1444	1,2802	1,4672	1,3320	0,9693	1,5271	1,1969	1,2092
	B1	1,3226	1,2112	<b>0,6644</b>	0,9697	1,4661	1,0859	<b>1,0982</b>	1,3951	1,2271	0,9343	1,4507	1,0122	1,1531
	BCIGP	<b>1,1355</b>	<b>1,1446</b>	0,7202	<b>0,8728</b>	1,2805	<b>0,9932</b>	1,1041	<b>1,2280</b>	<b>1,1336</b>	<b>0,8347</b>	<b>1,2847</b>	<b>0,9354</b>	<b>1,0556</b>

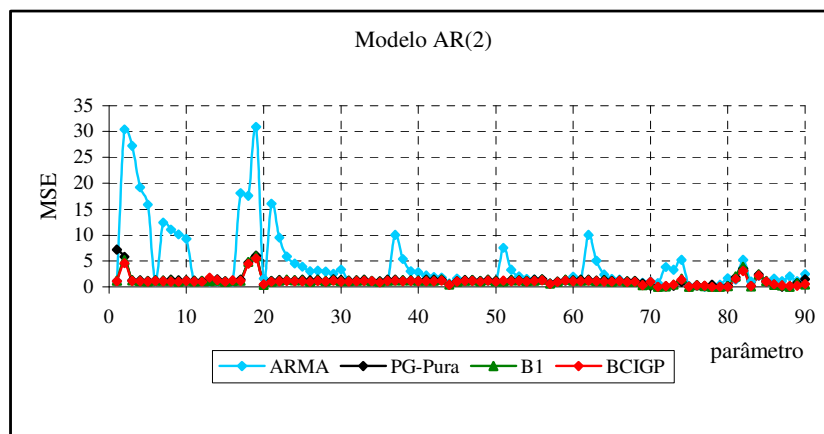
TABELA C2 – RESULTADOS DA DO MSE MÉDIO PARA AS 500 SÉRIES DA ESTRUTURA AR(2)

MSE	AR(2)	P1	P2	P3	P4	P5	P6	P7	P8	P9	P11	P12	Média
e136	ARMA	<b>1,0802</b>	30,4184	27,2204	19,1999	15,9212	<b>1,1224</b>	12,4174	11,1114	10,1633	<b>1,0091</b>	<b>0,9146</b>	4,3479
	PG-Pura	7,2094	5,7822	1,3014	1,2973	1,1840	1,2542	1,2402	1,3701	1,2667	1,1965	1,1477	1,3176
	B1	1,2341	5,1413	1,2588	1,1543	1,1235	1,2357	1,2418	1,2001	1,0551	1,1270	1,1374	1,1906
	BCIGP	1,0855	<b>4,5957</b>	<b>1,1383</b>	<b>1,0782</b>	<b>1,0653</b>	1,1588	<b>1,1102</b>	<b>1,0571</b>	<b>1,0057</b>	1,0576	0,9879	<b>1,0997</b>
e137	ARMA	4,8487	26,2853	26,6214	17,5784	14,7372	<b>1,1453</b>	11,2089	11,3507	9,7828	1,0563	<b>1,1107</b>	4,1828
	PG-Pura	2,3544	6,1254	1,3099	1,4221	1,2871	1,6537	1,2946	1,2462	1,2073	1,1823	1,2025	1,5090
	B1	1,2032	<b>5,4885</b>	1,1850	1,2552	1,2177	1,2575	1,3240	1,2796	1,1745	1,1461	1,1891	1,3537
	BCIGP	<b>1,0394</b>	6,5075	<b>1,0681</b>	<b>1,1602</b>	<b>1,1144</b>	1,1519	<b>1,1782</b>	<b>1,1662</b>	<b>1,0862</b>	<b>1,0517</b>	1,2760	<b>1,2563</b>
e138	ARMA	12,8973	24,7732	25,6531	18,7672	14,0630	1,8867	11,7266	10,6651	9,3283	1,8587	1,9538	4,0820
	PG-Pura	1,5221	9,3548	1,2829	1,4114	1,4412	1,2350	1,1564	1,3404	1,2227	1,2654	1,3917	1,3674
	B1	<b>1,2717</b>	5,0143	1,1403	1,2026	1,3595	1,1937	1,1689	1,3574	<b>1,1527</b>	1,2351	1,4682	<b>1,1733</b>
	BCIGP	1,2972	<b>4,5473</b>	<b>1,0242</b>	<b>1,1098</b>	<b>1,3494</b>	<b>1,0824</b>	<b>1,0941</b>	<b>1,2456</b>	1,7626	<b>1,1619</b>	<b>1,2349</b>	1,4734
e139	ARMA	24,8881	22,6164	23,8962	18,6023	13,7678	2,1022	10,4568	10,0440	8,8330	2,0606	2,1276	4,1294
	PG-Pura	2,7826	5,6362	1,2856	1,3203	1,2176	1,1218	1,3319	1,2932	1,2969	1,2193	1,5794	1,4605
	B1	1,3638	5,2595	1,1810	<b>1,1476</b>	1,2154	1,0808	1,3017	<b>1,2693</b>	1,1985	1,1577	1,3708	<b>1,2963</b>
	BCIGP	<b>1,2213</b>	<b>4,8055</b>	<b>1,0972</b>	1,1498	<b>1,0721</b>	<b>0,9959</b>	<b>1,1539</b>	2,4237	<b>1,0658</b>	<b>1,0637</b>	<b>1,1815</b>	1,5489
e140	ARMA	40,1308	19,9142	21,4903	16,5470	12,9552	2,6010	10,5260	9,4413	9,1637	2,5578	2,4234	4,2354
	PG-Pura	1,4557	5,1691	1,2463	1,1257	1,1421	1,1823	<b>1,1643</b>	1,2257	1,1716	2,4261	1,3668	1,6394
	B1	1,3233	4,7760	1,0880	0,9736	1,1588	1,1376	1,1696	1,2580	1,1229	<b>1,1083</b>	1,2723	<b>1,4485</b>
	BCIGP	<b>1,1855</b>	<b>4,5534</b>	<b>0,9989</b>	<b>0,8949</b>	<b>1,0658</b>	<b>1,0453</b>	1,1869	<b>1,1128</b>	<b>1,0089</b>	12,2212	<b>1,0942</b>	1,4844
e141	ARMA	57,9008	21,1050	21,1283	15,4964	12,4722	2,9042	9,8654	9,5064	7,8496	2,5637	3,0887	4,5942
	PG-Pura	2,8662	5,6992	1,1974	1,2540	1,1746	1,2178	1,2292	1,2779	1,3622	1,2084	1,3480	1,5345
	B1	1,2027	<b>5,1438</b>	1,0731	1,1025	1,2491	1,1260	1,2377	1,2950	1,2422	1,1708	1,3394	1,3558
	BCIGP	<b>1,1989</b>	6,4460	<b>0,9907</b>	<b>1,0162</b>	<b>1,1651</b>	<b>1,0519</b>	<b>1,1200</b>	<b>1,1885</b>	<b>1,1237</b>	<b>1,1469</b>	<b>1,2254</b>	<b>1,3466</b>
e142	ARMA	76,3452	18,5512	20,9183	16,2920	11,6596	2,8273	8,9483	9,0586	8,6535	3,0144	2,7725	4,5034
	PG-Pura	1,7251	5,6469	1,2412	1,2672	1,1906	1,3700	1,3071	1,1374	1,4149	1,3995	1,5712	1,6545
	B1	1,1303	5,1020	1,1079	1,1208	1,2436	1,2844	1,3179	1,1794	1,2238	1,2454	1,1701	<b>1,1711</b>
	BCIGP	<b>1,0961</b>	<b>4,6036</b>	<b>1,0283</b>	<b>1,0285</b>	<b>1,1308</b>	<b>1,1669</b>	<b>1,1647</b>	<b>1,0543</b>	<b>1,1131</b>	<b>1,1514</b>	<b>1,0561</b>	1,2458
e143	ARMA	92,8756	15,9869	19,9769	16,2265	11,7152	3,5491	9,0154	8,8684	7,1558	3,0709	3,5252	4,7462
	PG-Pura	22,7279	5,1451	1,2343	3,8372	1,3756	<b>1,3261</b>	<b>1,3286</b>	1,1978	1,4064	1,2838	1,1424	1,7681
	B1	<b>1,2480</b>	4,9167	1,0414	<b>1,1356</b>	1,3301	8,4110	1,3364	1,2025	1,2151	<b>1,2021</b>	1,1835	<b>1,3549</b>
	BCIGP	1,3090	<b>4,4090</b>	<b>0,9611</b>	1,4604	<b>1,1569</b>	5,8323	1,5620	<b>1,0774</b>	<b>1,1200</b>	1,9981	<b>1,0517</b>	1,4738
e144	ARMA	1,5260	16,7639	19,8103	14,3815	11,7778	3,2493	8,2570	8,6419	8,1574	3,5778	3,4480	3,6964
	PG-Pura	1,2953	5,9317	1,2876	1,2673	1,2569	1,1417	1,2092	<b>1,1925</b>	1,3707	1,3370	182,9982	5,0540
	B1	<b>1,2926</b>	5,5162	1,1843	1,1866	<b>1,2087</b>	<b>1,0898</b>	1,2357	1,2565	1,2709	1,2722	<b>1,3003</b>	<b>1,2776</b>
	BCIGP	1,3838	<b>5,1028</b>	<b>1,0920</b>	<b>1,1042</b>	1,6563	1,1655	<b>1,1278</b>	1,3326	<b>1,1564</b>	<b>1,1959</b>	2,7954	1,5608
e145	ARMA	1,7414	14,3084	19,8260	13,2495	10,7788	3,6646	8,8848	8,2555	7,3526	3,7720	4,2307	3,6346
	PG-Pura	1,3318	5,3188	1,5140	1,2831	1,2400	1,1490	1,2575	1,2755	1,3172	2,5584	1,3855	1,4859
	B1	1,2868	<b>4,8494</b>	1,2742	1,1664	1,1989	1,1127	1,1883	1,1776	1,1575	1,1092	1,4729	1,2636
	BCIGP	<b>1,2622</b>	10,9423	<b>1,1365</b>	<b>1,0342</b>	<b>1,0491</b>	<b>1,0291</b>	<b>1,0955</b>	<b>1,0693</b>	<b>1,0620</b>	<b>1,0458</b>	<b>1,3482</b>	<b>1,2616</b>
e146	ARMA	1,4519	14,3757	18,0474	13,9645	10,3335	3,9041	8,3059	7,8002	7,4655	3,8769	4,0398	3,3792
	PG-Pura	1,2932	5,7363	1,2243	1,2801	1,2085	1,3783	<b>1,2495</b>	1,2384	1,2155	2,3435	1,2296	2,4862
	B1	1,1749	5,4490	1,0262	1,1103	1,1718	1,1495	1,3103	1,1904	1,1465	1,1934	1,2053	1,2689
	BCIGP	<b>1,1431</b>	<b>4,8539</b>	<b>0,9346</b>	<b>1,0337</b>	<b>1,0861</b>	<b>1,0756</b>	1,7796	<b>1,1099</b>	<b>1,0426</b>	<b>1,1175</b>	<b>1,1124</b>	<b>1,2078</b>
e147	ARMA	1,7248	14,3862	16,8247	13,7947	10,1737	3,9345	7,7295	7,5343	7,2597	3,8329	4,3897	3,4343
	PG-Pura	1,4267	6,2440	1,3643	1,2421	1,1664	1,2452	1,2831	1,3762	1,1993	1,1735	1,2556	1,5164
	B1	<b>1,3185</b>	5,8859	<b>1,2136</b>	1,1768	1,1574	1,2003	1,2975	1,3397	1,0814	1,0950	1,3075	1,3071
	BCIGP	3,0574	<b>5,5592</b>	1,2431	<b>1,0731</b>	<b>1,0589</b>	<b>1,0837</b>	<b>1,1843</b>	<b>1,2099</b>	<b>0,9655</b>	<b>1,0011</b>	<b>1,1877</b>	<b>1,2561</b>
e148	ARMA	1,6248	11,9728	16,8477	12,7637	9,8661	4,2312	8,3458	7,7868	6,7437	4,3591	4,4856	3,3868
	PG-Pura	1,3822	5,9932	1,2929	1,5128	1,2434	1,2944	1,6200	1,2691	1,4438	1,4024	1,1987	2,2805
	B1	1,3798	5,7450	<b>1,1975</b>	<b>1,3116</b>	1,2999	1,2297	1,1762	1,2400	<b>1,2643</b>	1,2336	1,1809	1,2308
	BCIGP	<b>1,2171</b>	<b>5,3115</b>	1,8183	1,3924	<b>1,1412</b>	<b>1,1435</b>	<b>1,0690</b>	<b>1,1407</b>	1,6593	<b>1,1274</b>	<b>1,0637</b>	<b>1,1477</b>
e149	ARMA	1,5818	12,3031	17,6594	12,5267	9,9771	4,5445	7,7927	7,3195	7,1576	4,3245	4,5140	3,5028
	PG-Pura	1,2915	5,7968	1,4642	1,2896	1,7125	1,2725	1,1876	1,1802	1,2745	1,2525	1,9175	1,5118
	B1	1,2895	5,4314	<b>1,1464</b>	1,1090	1,2718	1,1745	1,2176	1,2497	1,1587	1,1476	1,3103	<b>1,2823</b>
	BCIGP	<b>1,1744</b>	<b>4,9147</b>	1,1558	<b>1,0719</b>	<b>1,2074</b>	<b>1,0586</b>	<b>1,0957</b>	<b>1,1311</b>	<b>1,0620</b>	<b>1,0709</b>	<b>1,2071</b>	1,8203
e150	ARMA	1,9268	10,9129	17,0416	13,0269	9,4549	4,3041	7,9229	7,1098	6,0829	4,6121	5,0135	3,4916
	PG-Pura	1,4565	5,3119	1,3237	1,2440	1,1526	1,2865	1,2510	1,2406	1,2411	1,1382	1,3005	1,6753
	B1	<b>1,2757</b>	4,9397	1,2433	1,0928	1,1409	<b>1,0403</b>	1,1761	1,3001	<b>1,1600</b>	<b>1,0783</b>	1,3088	<b>1,4174</b>
	BCIGP	3,3563	<b>4,4785</b>	<b>1,1414</b>	<b>0,9972</b>	<b>1,0389</b>	1,5726	<b>1,1614</b>	<b>1,1430</b>	4,1752	1,3417	<b>1,1988</b>	1,4630

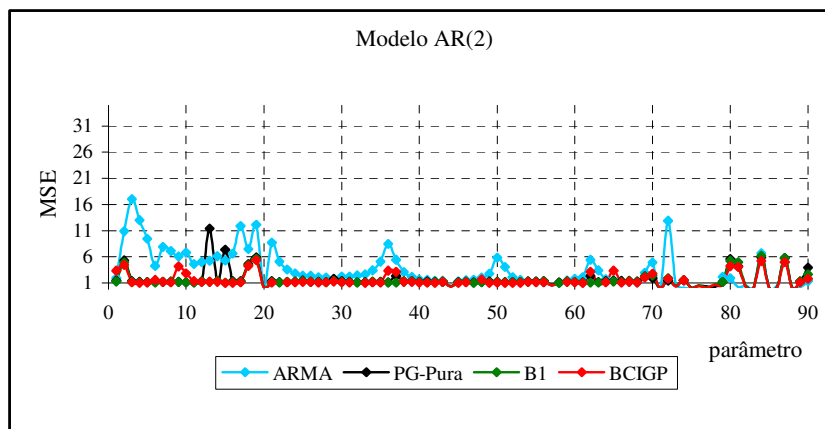
**TABELA C3 – RESULTADOS DA DO MSE MÉDIO PARA AS 500 SÉRIES DA ESTRUTURA MA(2)**

MSE	MA(2)	P1	P2	P3	P4	P5	P8	P7	P8	P9	P10	P11	Média
e136	ARMA	5,4594	<b>4,8009</b>	<b>1,0284</b>	3,6715	3,6004	5,5681	<b>1,8078</b>	1,6048	<b>1,2183</b>	2,3741	2,5054	3,0581
	PG-Pura	6,0223	5,0640	1,4835	3,7383	3,3566	4,9471	2,1112	1,8420	1,9154	1,4451	1,2795	3,0186
	B1	6,0982	5,1096	1,4604	3,7673	3,1004	4,9008	2,1099	1,7300	1,4438	1,2078	1,1607	2,9172
	BCIGP	<b>5,4091</b>	5,1738	1,5305	<b>3,4506</b>	<b>2,8332</b>	<b>4,4084</b>	1,8607	<b>1,5394</b>	1,2937	<b>1,0649</b>	<b>1,0506</b>	<b>2,6923</b>
e137	ARMA	5,0318	<b>4,5225</b>	2,0600	3,8187	4,0626	4,7411	<b>1,9265</b>	<b>1,6948</b>	1,3091	2,4049	2,5760	3,1044
	PG-Pura	5,5272	5,0498	1,4823	3,6222	3,7215	4,9376	2,0866	1,9647	1,5636	1,5222	1,4392	2,9924
	B1	5,6263	4,9852	1,5018	3,5851	3,3060	4,6315	2,2429	1,8701	1,4683	1,4170	1,3931	2,9116
	BCIGP	<b>5,0096</b>	4,7712	<b>1,2940</b>	<b>3,3396</b>	<b>3,0279</b>	<b>4,1936</b>	2,0167	1,7049	<b>1,2926</b>	<b>1,2321</b>	<b>1,2334</b>	<b>2,6469</b>
e138	ARMA	<b>5,0623</b>	<b>4,9960</b>	2,6580	<b>3,6677</b>	2,8745	4,5620	2,3240	1,9058	2,0380	1,7015	1,7188	3,0462
	PG-Pura	5,9637	5,3801	1,5202	3,9252	3,4105	4,4963	2,3435	2,3411	1,7167	1,4709	1,3436	3,0829
	B1	5,6849	5,4076	1,5264	4,5971	2,8718	4,5226	2,2018	1,5396	1,5698	1,4006	1,3430	2,9696
	BCIGP	5,2235	5,1070	<b>1,3070</b>	3,9293	<b>2,5572</b>	<b>4,3429</b>	<b>1,9833</b>	<b>1,4771</b>	<b>1,4009</b>	<b>1,2630</b>	<b>1,1997</b>	<b>2,7083</b>
e139	ARMA	<b>5,8466</b>	<b>4,4018</b>	3,3906	4,1714	3,3106	4,5128	2,0900	2,1286	1,8751	1,9356	1,9467	3,2373
	PG-Pura	7,5813	5,1918	1,6535	4,3501	3,7962	4,7896	1,8854	1,7065	1,5989	1,4180	1,4779	3,2227
	B1	6,3823	5,0360	1,5829	4,3711	<b>3,1298</b>	4,4398	1,8869	1,6502	1,5027	1,3776	1,3508	2,9737
	BCIGP	5,9161	4,5898	<b>1,3924</b>	<b>3,9812</b>	4,6292	<b>3,9586</b>	<b>1,6989</b>	<b>1,5234</b>	<b>1,3577</b>	<b>1,2842</b>	<b>1,1962</b>	<b>2,8661</b>
e140	ARMA	<b>5,4165</b>	<b>4,6826</b>	3,4089	<b>3,5268</b>	3,2395	<b>4,7385</b>	2,3319	1,8340	1,9197	1,8497	1,8351	3,1621
	PG-Pura	5,7407	5,4421	1,4618	3,5871	3,2852	5,0845	2,3673	1,6110	1,5554	<b>1,3729</b>	1,3432	2,9865
	B1	6,1127	5,1042	1,4815	3,7694	3,0418	5,1178	2,0987	1,5576	1,5030	1,4107	1,2576	2,9505
	BCIGP	5,4851	4,7612	<b>1,2526</b>	3,6740	<b>2,7297</b>	5,9195	<b>1,8800</b>	<b>1,3809</b>	<b>1,3527</b>	1,9813	<b>1,1259</b>	<b>2,8675</b>
e141	ARMA	<b>5,1189</b>	4,8354	2,9969	3,7578	3,2047	4,3695	2,2785	2,0438	1,8353	1,7411	1,6844	3,0788
	PG-Pura	6,0379	5,5150	1,4428	4,1115	3,0724	4,6498	2,1108	1,7029	1,5242	1,5190	1,3608	3,0043
	B1	5,9841	5,3549	1,4888	3,7972	2,9931	4,8267	2,1601	1,6195	1,3948	1,4590	1,2952	2,9430
	BCIGP	5,4048	<b>4,8129</b>	<b>1,2978</b>	<b>3,4125</b>	<b>2,7098</b>	<b>4,2923</b>	<b>1,9085</b>	<b>1,4584</b>	<b>1,2487</b>	<b>1,3165</b>	<b>1,1932</b>	<b>2,6414</b>
e142	ARMA	<b>5,4936</b>	<b>4,7302</b>	2,6478	3,4588	2,7453	<b>4,1495</b>	2,4092	2,1193	1,7650	1,6833	1,6623	2,9877
	PG-Pura	6,0012	5,4193	1,5690	3,5727	2,8506	4,5518	2,2705	1,9819	1,4653	1,4311	1,5479	2,9692
	B1	7,1691	5,2741	1,5728	3,6298	2,6952	4,3382	2,1987	1,8044	1,3941	1,4283	<b>1,3663</b>	2,9883
	BCIGP	6,0887	4,8483	<b>1,3695</b>	<b>3,2812</b>	<b>2,4062</b>	4,7140	<b>1,9547</b>	<b>1,6183</b>	<b>1,2491</b>	<b>1,2784</b>	2,0811	<b>2,8081</b>
e143	ARMA	<b>4,9845</b>	4,4928	2,9623	3,7131	3,0555	4,6118	2,1175	1,8681	1,8620	1,8165	1,8349	3,0290
	PG-Pura	5,4914	5,0442	1,5200	5,2046	3,2519	4,7496	5,1162	1,6014	1,5787	1,3483	1,3477	3,2958
	B1	5,3728	4,9064	1,5143	3,9475	3,0152	4,5497	2,0068	1,5699	<b>1,4171</b>	1,3983	1,2913	<b>2,8172</b>
	BCIGP	5,1058	<b>4,2915</b>	<b>1,3157</b>	<b>3,5404</b>	<b>2,6498</b>	<b>4,2377</b>	<b>1,8541</b>	<b>1,4111</b>	5,8488	<b>1,1769</b>	<b>1,1424</b>	2,9613
e144	ARMA	<b>5,5831</b>	<b>4,4468</b>	3,1804	<b>3,4794</b>	3,0329	4,3849	2,3393	2,0929	1,7802	1,9079	1,8573	3,0986
	PG-Pura	6,3746	4,9296	1,9550	3,8388	3,1769	4,6553	2,2137	1,8276	1,4604	1,7118	1,4071	3,0501
	B1	5,8153	4,8898	1,4983	3,8967	2,8843	4,4882	2,1255	<b>1,6910</b>	1,3999	1,5679	1,3881	<b>2,8768</b>
	BCIGP	5,6573	4,4641	<b>1,3059</b>	3,5474	<b>2,5627</b>	<b>4,3798</b>	<b>1,9156</b>	2,3412	<b>1,2547</b>	<b>1,3369</b>	<b>1,3655</b>	2,7392
e145	ARMA	5,2977	<b>4,1004</b>	2,9206	<b>3,2318</b>	2,9088	4,6347	2,2128	2,1117	1,6556	1,9360	1,9213	2,9938
	PG-Pura	5,6472	5,0257	1,4885	3,8395	2,9926	4,4469	2,1080	1,8471	1,3786	1,8088	1,4417	2,9113
	B1	5,6054	4,7905	1,4805	3,7222	2,8156	4,6280	2,1783	<b>1,7725</b>	1,3151	1,5229	1,3628	<b>2,8358</b>
	BCIGP	<b>5,1691</b>	4,5341	<b>1,3220</b>	3,3961	<b>2,5151</b>	<b>4,1133</b>	<b>2,0038</b>	2,6830	<b>1,1590</b>	<b>1,3472</b>	<b>1,2202</b>	2,6784
e146	ARMA	<b>5,3492</b>	4,7882	3,0257	3,3060	3,0926	4,8387	2,3415	1,9418	1,8679	1,7823	1,7221	3,0960
	PG-Pura	6,4009	5,0611	1,5516	3,6148	3,3126	5,2284	2,3050	1,7578	1,5788	1,3757	1,3622	3,0499
	B1	5,9971	5,0849	1,5015	7,8041	<b>2,9086</b>	<b>4,7045</b>	2,2912	1,6363	1,4251	1,3630	<b>1,2469</b>	<b>3,2694</b>
	BCIGP	5,3648	<b>4,4839</b>	<b>1,3011</b>	<b>3,1797</b>	3,7847	8,6834	<b>2,0513</b>	<b>1,4747</b>	<b>1,2946</b>	<b>1,2053</b>	1,6203	3,1313
e147	ARMA	<b>5,3690</b>	<b>4,3878</b>	2,9568	<b>3,3791</b>	3,1202	4,7786	2,1658	2,1155	1,8170	1,7592	1,6932	3,0493
	PG-Pura	6,2145	4,8319	1,4266	4,3725	3,4735	5,1689	1,9567	1,9153	1,5800	1,4425	1,3927	3,0705
	B1	5,9349	4,9676	1,4014	3,6349	<b>3,0941</b>	5,1408	1,9469	1,7532	1,4274	1,4167	<b>1,3513</b>	<b>2,9154</b>
	BCIGP	5,3765	4,4514	<b>1,2028</b>	3,3853	3,1981	<b>4,6958</b>	<b>1,7327</b>	<b>1,5232</b>	<b>1,2943</b>	<b>1,2322</b>	1,4445	2,6852
e148	ARMA	<b>5,2617</b>	<b>4,2105</b>	2,8659	3,6397	3,1329	4,5550	2,3194	2,0182	1,7621	1,9242	1,8515	3,0492
	PG-Pura	5,5399	4,6902	2,3636	4,0420	3,3190	<b>4,2684</b>	2,1353	1,6930	1,5223	1,4909	1,5830	2,9680
	B1	5,6786	4,5872	<b>1,3885</b>	3,7301	3,0029	4,4418	2,1118	1,5521	1,4051	1,5754	<b>1,4434</b>	<b>2,8106</b>
	BCIGP	5,3160	4,4014	2,4220	<b>3,4403</b>	<b>2,6906</b>	4,2868	<b>1,8499</b>	<b>1,4081</b>	<b>1,2525</b>	<b>1,4293</b>	2,2771	2,7976
e149	ARMA	<b>5,7131</b>	4,7585	2,9308	3,4957	2,8962	5,1384	2,1831	2,1260	1,9216	1,6443	1,7532	3,1419
	PG-Pura	6,2896	5,2288	1,5066	3,8465	3,1592	5,3245	1,9895	1,8283	1,5794	1,4002	1,3442	3,0452
	B1	6,7234	5,0059	1,5675	3,7197	2,8493	5,3490	2,0385	1,7329	1,4515	1,3762	1,2694	3,0076
	BCIGP	6,4839	<b>4,6709</b>	<b>1,3427</b>	<b>3,3764</b>	<b>2,5716</b>	<b>5,0149</b>	<b>1,8705</b>	<b>1,5703</b>	<b>1,2897</b>	<b>1,1971</b>	<b>1,1152</b>	<b>2,7730</b>
e150	ARMA	<b>5,7051</b>	4,8405	3,1043	3,9386	2,8196	4,7408	2,2447	2,1117	1,8661	1,8467	1,8183	3,1851
	PG-Pura	5,8408	5,3527	2,5827	4,0424	2,8517	4,5980	2,2869	1,7357	1,6710	1,3964	1,3389	3,0634
	B1	5,9287	5,1876	1,5002	3,9027	2,7305	4,9105	2,1987	1,6635	1,4677	1,4041	1,2712	<b>2,9241</b>
	BCIGP	5,9638	<b>4,7971</b>	<b>1,3103</b>	<b>3,5387</b>	<b>2,4695</b>	<b>4,4304</b>	<b>2,0360</b>	<b>1,5037</b>	<b>1,3142</b>	<b>1,1535</b>	<b>1,1342</b>	2,6956

Nas figuras C1 e C2 encontram-se os gráficos representativos dos erros de previsão para os horizontes um e quinze, das séries geradas pela estrutura AR(2).

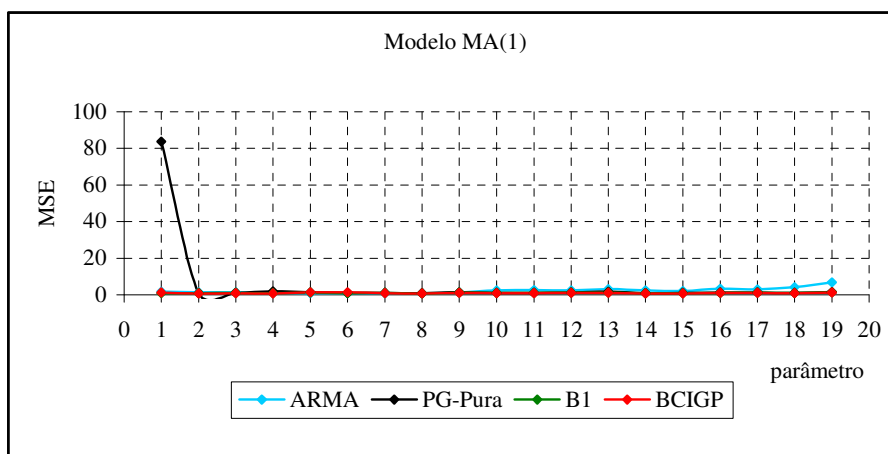


**FIGURA C1 COMPORTAMENTO DOS MÉTODOS PARA O 1º VALOR PREVISTO  $e_{136}$  – MODELO AR(2)**

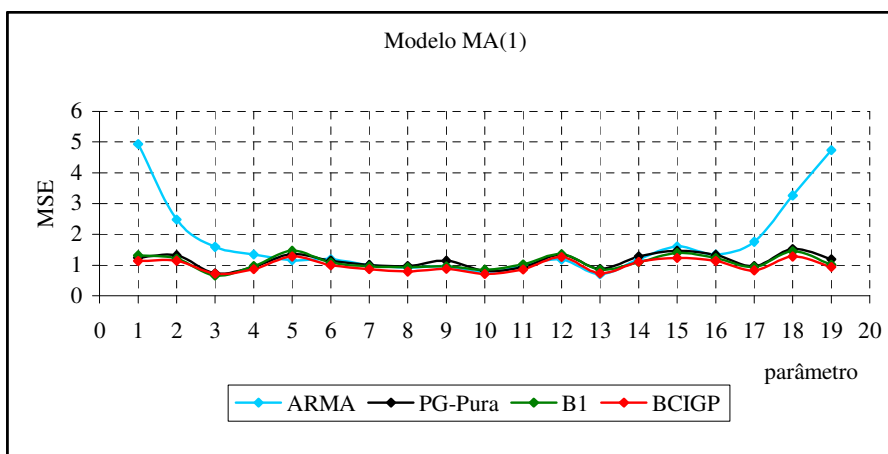


**FIGURA C2 – COMPORTAMENTO DOS MÉTODOS PARA O ÚLTIMO VALOR PREVISTO  $e_{150}$  – MODELO AR(2)**

Nas figuras C3 e C4 encontram-se os gráficos representativos dos erros de previsão para os horizontes um e quinze, das séries geradas pela estrutura MA(1).



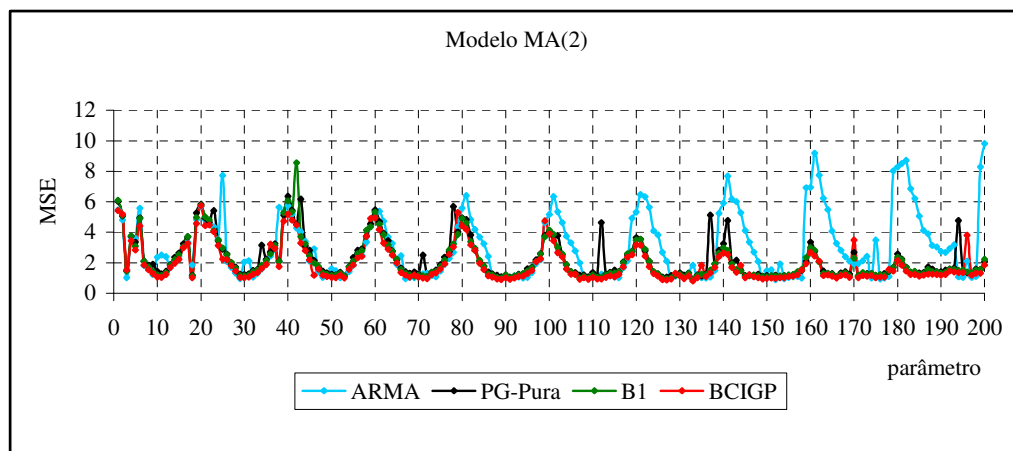
**FIGURA C3 – COMPORTAMENTO DOS MÉTODOS PARA O 1º VALOR PREVISTO  $e_{136}$  – MODELO MA(1)**



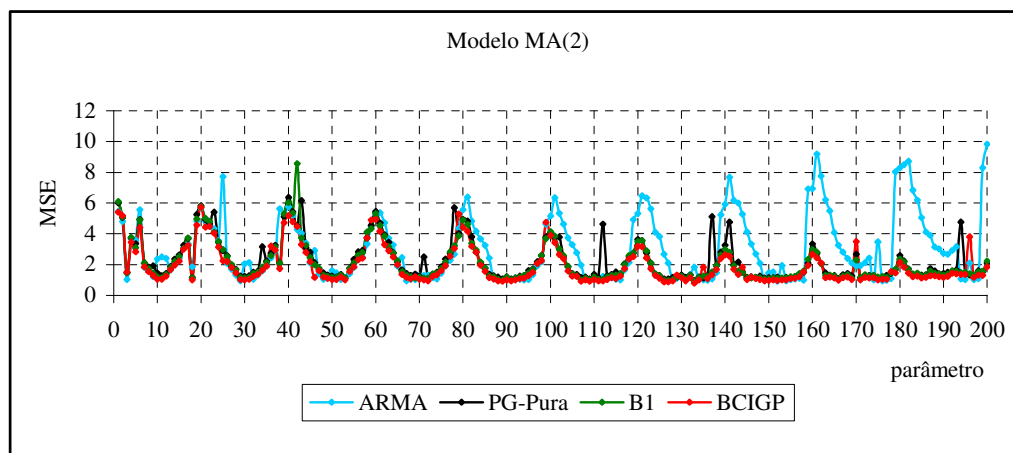
**FIGURA C4 – COMPORTAMENTO DOS MÉTODOS PARA O ÚLTIMO VALOR PREVISTO  $e_{150}$  – MODELO MA(1)**



Nas figuras C5 e C6 encontram-se os gráficos representativos dos erros de previsão para os horizontes um e quinze, das séries geradas pela estrutura MA(2).

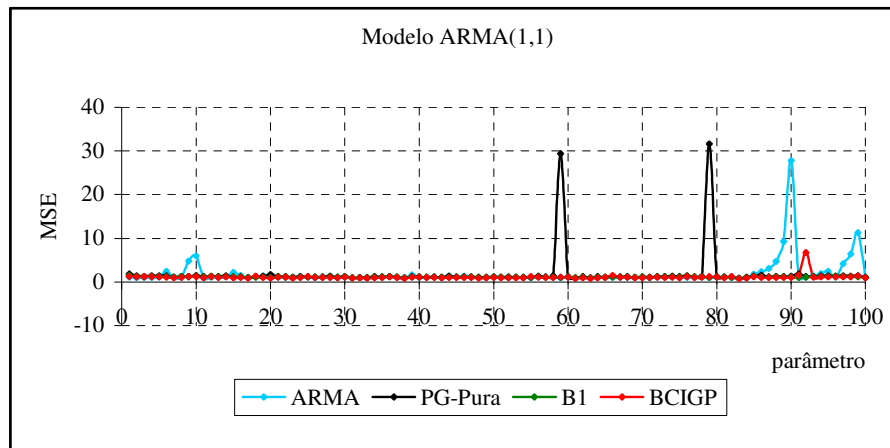


**FIGURA C5 – COMPORTAMENTO DOS MÉTODOS PARA O 1º VALOR PREVISTO  $e_{136}$  – MODELO AR(1)**

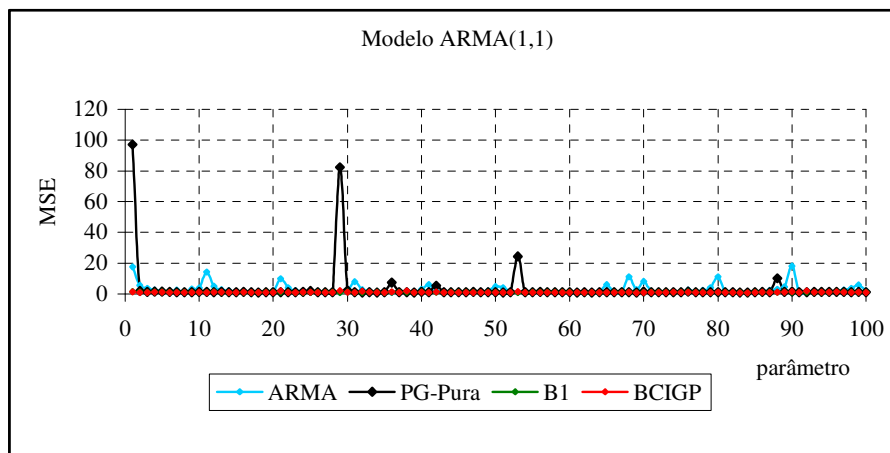


**FIGURA C6 - COMPORTAMENTO DOS MÉTODOS PARA O ÚLTIMO VALOR PREVISTO  $e_{150}$  – MODELO MA(2)**

Nas figuras C7 e C8 encontram-se os gráficos representativos dos erros de previsão para os horizontes um e quinze, das séries geradas pela estrutura ARMA(1, 1).



**FIGURA C7 – COMPORTAMENTO DOS MÉTODOS PARA O 1º VALOR PREVISTO  $e_{136}$  – MODELO ARMA(1, 1)**



**FIGURA C8 – COMPORTAMENTO DOS MÉTODOS PARA O ÚLTMO VALOR PREVISTO  $e_{150}$  – MODELO ARMA(1, 1)**

As tabelas de C4 a C7 contém os resultados do teste F ANOVA e do teste Tukey-Kramer.

**TABELA C4 – RESULTADOS DA ANOVA E TESTE TUKEY-KRAMER PARA A ESTRUTURA AR(1)**

Valor Previsto	Grupos	Média Grupo	Média Geral	Estatís. F	p-valor	Grupos	Dif. Abs.	Interv. Crítico	Comp. médias	Melhor Grupo
e136	1	2,3702	1,3417	6,8196	0,0759	Grupo1 e 2	1,3135	0,9792	Diferentes	2
	2	1,0567				Grupo1 e 3	1,3584		Diferentes	3
	3	1,0118				Grupo1 e 4	1,4420		Diferentes	4
	4	0,9282				Grupo2 e 3	0,0450		Iguais	X
						Grupo2 e 4	0,1285		Iguais	X
						Grupo3 e 4	0,0836		Iguais	X
e137	1	2,3527	1,6550	2,2025	0,0000	Grupo1 e 2	0,3620	1,5429	Iguais	X
	2	1,9907				Grupo1 e 3	1,1685		Iguais	X
	3	1,1842				Grupo1 e 4	1,2603		Iguais	X
	4	1,0924				Grupo2 e 3	0,8065		Iguais	X
						Grupo2 e 4	0,8983		Iguais	X
						Grupo3 e 4	0,0918		Iguais	X
e138	1	2,0450	1,2704	5,8125	0,0000	Grupo1 e 2	0,9980	0,7987	Diferentes	2
	2	1,0470				Grupo1 e 3	1,0139		Diferentes	3
	3	1,0311				Grupo1 e 4	1,0867		Diferentes	4
	4	0,9583				Grupo2 e 3	0,0159		Iguais	X
						Grupo2 e 4	0,0887		Iguais	X
						Grupo3 e 4	0,0727		Iguais	X
e139	1	2,1817	1,5070	1,8080	0,0000	Grupo1 e 2	0,8503	1,2505	Iguais	X
	2	1,3314				Grupo1 e 3	0,8862		Iguais	X
	3	1,2955				Grupo1 e 4	0,9623		Iguais	X
	4	1,2194				Grupo2 e 3	0,0359		Iguais	X
						Grupo2 e 4	0,1120		Iguais	X
						Grupo3 e 4	0,0761		Iguais	X
e140	1	2,0486	1,5478	0,9972	0,0000	Grupo1 e 2	0,5387	1,3023	Iguais	X
	2	1,5099				Grupo1 e 3	0,6707		Iguais	X
	3	1,3779				Grupo1 e 4	0,7939		Iguais	X
	4	1,2547				Grupo2 e 3	0,1320		Iguais	X
						Grupo2 e 4	0,2552		Iguais	X
						Grupo3 e 4	0,1232		Iguais	X
e141	1	1,9550	1,5403	0,3820	0,0000	Grupo1 e 2	0,5063	1,7299	Iguais	X
	2	1,4487				Grupo1 e 3	0,6643		Iguais	X
	3	1,2907				Grupo1 e 4	0,4881		Iguais	X
	4	1,4669				Grupo2 e 3	0,1580		Iguais	X
						Grupo2 e 4	0,0182		Iguais	X
						Grupo3 e 4	0,1762		Iguais	X
e142	1	1,6495	1,5411	0,8351	0,0000	Grupo1 e 2	0,6347	2,2621	Iguais	X
	2	2,2842				Grupo1 e 3	0,5154		Iguais	X
	3	1,1341				Grupo1 e 4	0,5529		Iguais	X
	4	1,0966				Grupo2 e 3	1,1501		Iguais	X
						Grupo2 e 4	1,1876		Iguais	X
						Grupo3 e 4	0,0375		Iguais	X
e143	1	1,4880	1,3449	1,0502	0,0000	Grupo1 e 2	0,5150	1,8471	Iguais	X
	2	2,0030				Grupo1 e 3	0,5047		Iguais	X
	3	0,9834				Grupo1 e 4	0,5827		Iguais	X
	4	0,9053				Grupo2 e 3	1,0196		Iguais	X
						Grupo2 e 4	1,0977		Iguais	X
						Grupo3 e 4	0,0781		Iguais	X

**TABELA C4 – RESULTADOS DA ANOVA E TESTE TUKEY-KRAMER PARA A ESTRUTURA AR(1) (CONTINUAÇÃO)**

Valor Previsto	Grupos	Média Grupo	Média Geral	Estatís. F	p-valor	Grupos	Dif. Abs.	Interv. Crítico	Comp. médias	Melhor Grupo
e144	1	1,2274	0,9079	1,5880	0,0000	Grupo1 e 2	0,3814	0,6399	Iguais	X
	2	0,8461				Grupo1 e 3	0,4168		Iguais	X
	3	0,8106				Grupo1 e 4	0,4799		Iguais	X
	4	0,7475				Grupo2 e 3	0,0355		Iguais	X
						Grupo2 e 4	0,0985		Iguais	X
						Grupo3 e 4	0,0631		Iguais	X
e145	1	1,4540	1,2690	0,3824	0,0000	Grupo1 e 2	0,3328	1,0992	Iguais	X
	2	1,1212				Grupo1 e 3	0,3511		Iguais	X
	3	1,1029				Grupo1 e 4	0,0562		Iguais	X
	4	1,3978				Grupo2 e 3	0,0183		Iguais	X
						Grupo2 e 4	0,2766		Iguais	X
						Grupo3 e 4	0,2949		Iguais	X
e146	1	1,6769	1,4343	0,2636	0,1778	Grupo1 e 2	0,2598	1,2568	Iguais	X
	2	1,4171				Grupo1 e 3	0,3014		Iguais	X
	3	1,3755				Grupo1 e 4	0,4091		Iguais	X
	4	1,2677				Grupo2 e 3	0,0416		Iguais	X
						Grupo2 e 4	0,1493		Iguais	X
						Grupo3 e 4	0,1078		Iguais	X
e147	1	1,2946	1,0781	0,3021	0,0000	Grupo1 e 2	0,2370	1,0409	Iguais	X
	2	1,0576				Grupo1 e 3	0,2670		Iguais	X
	3	1,0276				Grupo1 e 4	0,3620		Iguais	X
	4	0,9326				Grupo2 e 3	0,0300		Iguais	X
						Grupo2 e 4	0,1250		Iguais	X
						Grupo3 e 4	0,0950		Iguais	X
e148	1	1,3387	1,1409	0,2759	0,3045	Grupo1 e 2	0,2152	1,0000	Iguais	X
	2	1,1235				Grupo1 e 3	0,2420		Iguais	X
	3	1,0967				Grupo1 e 4	0,3338		Iguais	X
	4	1,0049				Grupo2 e 3	0,0268		Iguais	X
						Grupo2 e 4	0,1186		Iguais	X
						Grupo3 e 4	0,0918		Iguais	X
e149	1	1,2239	1,0517	0,3284	0,0000	Grupo1 e 2	0,1934	0,8036	Iguais	X
	2	1,0304				Grupo1 e 3	0,2001		Iguais	X
	3	1,0237				Grupo1 e 4	0,2952		Iguais	X
	4	0,9286				Grupo2 e 3	0,0067		Iguais	X
						Grupo2 e 4	0,1018		Iguais	X
						Grupo3 e 4	0,0951		Iguais	X
e150	1	1,7268	1,6408	0,0572	0,0209	Grupo1 e 2	0,0111	1,7333	Iguais	X
	2	1,7379				Grupo1 e 3	0,2204		Iguais	X
	3	1,5064				Grupo1 e 4	0,1347		Iguais	X
	4	1,5922				Grupo2 e 3	0,2315		Iguais	X
						Grupo2 e 4	0,1458		Iguais	X
						Grupo3 e 4	0,0857		Iguais	X

**TABELA C5 – RESULTADOS DA ANOVA E TESTE TUKEY-KRAMER PARA A ESTRUTURA AR(2)**

Valor Previsto	Grupos	Média Grupo	Média Geral	Estatís. F	p-valor	Grupos	Dif. Abs.	Interv. Crítico	Comp. médias	Melhor Grupo
e136	1	4,3479	1,9889	20,4244	2,52E-12	Grupo1 e 2	3,0304	1,2723	Diferentes	2
	2	1,3176				Grupo1 e 3	3,1574		Diferentes	3
	3	1,1906				Grupo1 e 4	3,2482		Diferentes	4
	4	1,0997				Grupo2 e 3	0,1270		Iguais	X
						Grupo2 e 4	0,2178		Iguais	X
						Grupo3 e 4	0,0908		Iguais	X
e137	1	4,1828	2,0754	18,5868	2,67E-11	Grupo1 e 2	2,6737	1,1928	Diferentes	2
	2	1,5090				Grupo1 e 3	2,8290		Diferentes	3
	3	1,3537				Grupo1 e 4	2,9265		Diferentes	4
	4	1,2563				Grupo2 e 3	0,1553		Iguais	X
						Grupo2 e 4	0,2528		Iguais	X
						Grupo3 e 4	0,0975		Iguais	X
e138	1	4,0820	2,0240	18,1191	4,89E-11	Grupo1 e 2	2,7146	1,1814	Diferentes	2
	2	1,3674				Grupo1 e 3	2,9087		Diferentes	3
	3	1,1733				Grupo1 e 4	2,6086		Diferentes	4
	4	1,4734				Grupo2 e 3	0,1941		Iguais	X
						Grupo2 e 4	0,1060		Iguais	X
						Grupo3 e 4	0,3001		Iguais	X
e139	1	4,1294	2,1087	14,9967	2,90E-09	Grupo1 e 2	2,6688	1,2736	Diferentes	2
	2	1,4605				Grupo1 e 3	2,8330		Diferentes	3
	3	1,2963				Grupo1 e 4	2,5805		Diferentes	4
	4	1,5489				Grupo2 e 3	0,1642		Iguais	X
						Grupo2 e 4	0,0883		Iguais	X
						Grupo3 e 4	0,2525		Iguais	X
e140	1	4,2354	2,2019	13,5796	1,89E-08	Grupo1 e 2	2,5960	1,3454	Diferentes	2
	2	1,6394				Grupo1 e 3	2,7869		Diferentes	3
	3	1,4485				Grupo1 e 4	2,7510		Diferentes	4
	4	1,4844				Grupo2 e 3	0,1909		Iguais	X
						Grupo2 e 4	0,1550		Iguais	X
						Grupo3 e 4	0,0359		Iguais	X
e141	1	4,5942	2,2077	15,5852	1,34E-09	Grupo1 e 2	3,0597	1,4733	Diferentes	2
	2	1,5345				Grupo1 e 3	3,2385		Diferentes	3
	3	1,3558				Grupo1 e 4	3,2476		Diferentes	4
	4	1,3466				Grupo2 e 3	0,1788		Iguais	X
						Grupo2 e 4	0,1879		Iguais	X
						Grupo3 e 4	0,0092		Iguais	X
e142	1	4,5034	2,1436	10,5537	1,08E-06	Grupo1 e 2	2,8489	1,7837	Diferentes	2
	2	1,6545				Grupo1 e 3	3,3324		Diferentes	3
	3	1,1711				Grupo1 e 4	3,2576		Diferentes	4
	4	1,2458				Grupo2 e 3	0,4834		Iguais	X
						Grupo2 e 4	0,4087		Iguais	X
						Grupo3 e 4	0,0748		Iguais	X
e143	1	4,7462	2,3357	7,9880	3,50E-05	Grupo1 e 2	2,9781	2,0876	Diferentes	2
	2	1,7681				Grupo1 e 3	3,3914		Diferentes	3
	3	1,3549				Grupo1 e 4	3,2724		Diferentes	4
	4	1,4738				Grupo2 e 3	0,4132		Iguais	X
						Grupo2 e 4	0,2943		Iguais	X
						Grupo3 e 4	0,1189		Iguais	X

**TABELA C5 – RESULTADOS DA ANOVA E TESTE TUKEY-KRAMER PARA A ESTRUTURA AR(2) (CONTINUAÇÃO)**

Valor Previsto	Grupos	Média Grupo	Média Geral	Estatís. F	p-valor	Grupos	Dif. Abs.	Interv. Crítico	Comp. médias	Melhor Grupo
e144	1	3,6964	2,8971	2,4347	0,0644	Grupo1 e 2	1,3576	4,2065	Iguais	X
	2	5,0540				Grupo1 e 3	2,4188		Iguais	X
	3	1,2776				Grupo1 e 4	2,1356		Iguais	X
	4	1,5608				Grupo2 e 3	3,7764		Iguais	X
						Grupo2 e 4	3,4932		Iguais	X
						Grupo3 e 4	0,2832		Iguais	X
e145	1	3,6346	1,9114	24,8156	9,88E-15	Grupo1 e 2	2,1488	0,8453	Diferentes	2
	2	1,4859				Grupo1 e 3	2,3711		Diferentes	3
	3	1,2636				Grupo1 e 4	2,3730		Diferentes	4
	4	1,2616				Grupo2 e 3	0,2223		Iguais	NA
						Grupo2 e 4	0,2243		Iguais	NA
						Grupo3 e 4	0,0020		Iguais	NA
e146	1	3,3792	2,0855	10,2561	1,62E-06	Grupo1 e 2	0,8929	1,1903	Iguais	2
	2	2,4862				Grupo1 e 3	2,1103		Diferentes	3
	3	1,2689				Grupo1 e 4	2,1714		Diferentes	4
	4	1,2078				Grupo2 e 3	1,2173		Diferentes	3
						Grupo2 e 4	1,2785		Diferentes	4
						Grupo3 e 4	0,0611		Iguais	NA
e147	1	3,4343	1,8784	22,9336	1,04E-13	Grupo1 e 2	1,9179	0,7952	Diferentes	2
	2	1,5164				Grupo1 e 3	2,1272		Diferentes	3
	3	1,3071				Grupo1 e 4	2,1781		Diferentes	4
	4	1,2561				Grupo2 e 3	0,2093		Iguais	NA
						Grupo2 e 4	0,2603		Iguais	NA
						Grupo3 e 4	0,0509		Iguais	NA
e148	1	3,3868	2,0114	8,1202	2,93E-05	Grupo1 e 2	1,9910	1,3475	Iguais	2
	2	2,2805				Grupo1 e 3	2,2205		Diferentes	3
	3	1,2308				Grupo1 e 4	1,6825		Diferentes	4
	4	1,1477				Grupo2 e 3	0,2295		Iguais	NA
						Grupo2 e 4	0,3085		Iguais	NA
						Grupo3 e 4	0,5380		Iguais	NA
e149	1	3,5028	2,0293	9,1038	7,70E-06	Grupo1 e 2	1,8163	1,2180	Diferentes	2
	2	1,5118				Grupo1 e 3	2,0742		Diferentes	3
	3	1,2823				Grupo1 e 4	2,0286		Diferentes	4
	4	1,8203				Grupo2 e 3	0,2579		Iguais	NA
						Grupo2 e 4	0,2123		Iguais	NA
						Grupo3 e 4	0,0456		Iguais	NA
e150	1	3,4916	2,0117	19,6169	7,08E-12	Grupo1 e 2	1,8163	0,8183	Diferentes	2
	2	1,6753				Grupo1 e 3	2,0742		Diferentes	3
	3	1,4174				Grupo1 e 4	2,0286		Diferentes	4
	4	1,4630				Grupo2 e 3	0,2579		Iguais	NA
						Grupo2 e 4	0,2123		Iguais	NA
						Grupo3 e 4	0,0456		Iguais	NA

**TABELA C6 – RESULTADOS DA ANOVA E TESTE TUKEY-KRAMER PARA A ESTRUTURA MA(2)**

Valor Previsto	Grupos	Média Grupo	Média Geral	Estatís. F	p-valor	Grupos	Dif. Abs.	Interv. Crítico	Comp. médias	Melhor Grupo
e136	1	3,0335	2,3252	21,2007	9,35E-13	Grupo1 e 2	0,7879	0,3876	Diferentes	2
	2	2,2456				Grupo1 e 3	0,9350		Diferentes	3
	3	2,0985				Grupo1 e 4	1,1101		Diferentes	4
	4	1,9234				Grupo2 e 3	0,1470		Iguais	X
						Grupo2 e 4	0,3222		Iguais	X
						Grupo3 e 4	0,1751		Iguais	X
e137	1	2,7355	2,2165	17,9565	6,04E-11	Grupo1 e 2	0,5665	0,3143	Diferentes	2
	2	2,1690				Grupo1 e 3	0,6582		Diferentes	3
	3	2,0773				Grupo1 e 4	0,8512		Diferentes	4
	4	1,8843				Grupo2 e 3	0,0917		Iguais	X
						Grupo2 e 4	0,2847		Iguais	X
						Grupo3 e 4	0,1930		Iguais	X
e138	1	2,6617	2,4762	4,4276	4,46E-03	Grupo1 e 2	0,5999	1,0643	Iguais	X
	2	3,2616				Grupo1 e 3	0,5919		Iguais	X
	3	2,0698				Grupo1 e 4	0,7499		Iguais	X
	4	1,9117				Grupo2 e 3	1,1918		Diferentes	3
						Grupo2 e 4	1,3499		Diferentes	4
						Grupo3 e 4	0,1580		Iguais	X
e139	1	2,6473	2,4186	2,3696	7,02E-02	Grupo1 e 2	0,3769	1,2036	Iguais	X
	2	3,0242				Grupo1 e 3	0,5762		Iguais	X
	3	2,0711				Grupo1 e 4	0,7155		Iguais	X
	4	1,9318				Grupo2 e 3	0,9531		Iguais	X
						Grupo2 e 4	1,0924		Iguais	X
						Grupo3 e 4	0,1393		Iguais	X
e140	1	2,6666	2,2070	14,0977	9,51E-09	Grupo1 e 2	0,4851	0,3184	Diferentes	2
	2	2,1816				Grupo1 e 3	0,5838		Diferentes	3
	3	2,0828				Grupo1 e 4	0,7696		Diferentes	4
	4	1,8970				Grupo2 e 3	0,0988		Iguais	X
						Grupo2 e 4	0,2846		Iguais	X
						Grupo3 e 4	0,1858		Iguais	X
e141	1	2,6444	2,2527	9,9198	2,55E-06	Grupo1 e 2	0,2790	0,3660	Iguais	X
	2	2,3654				Grupo1 e 3	0,5831		Diferentes	3
	3	2,0613				Grupo1 e 4	0,7046		Diferentes	4
	4	1,9397				Grupo2 e 3	0,3041		Iguais	X
						Grupo2 e 4	0,4256		Diferentes	4
						Grupo3 e 4	0,1215		Iguais	X
e142	1	2,6580	2,2414	10,5802	1,05E-06	Grupo1 e 2	0,3569	0,3553	Diferentes	2
	2	2,3011				Grupo1 e 3	0,5780		Diferentes	3
	3	2,0800				Grupo1 e 4	0,7315		Diferentes	4
	4	1,9264				Grupo2 e 3	0,2211		Iguais	X
						Grupo2 e 4	0,3747		Diferentes	4
						Grupo3 e 4	0,1536		Iguais	X
e143	1	2,6514	3,0618	0,7324	5,33E-01	Grupo1 e 2	0,4846	4,2378	Iguais	X
	2	3,1361				Grupo1 e 3	0,5934		Iguais	X
	3	2,0581				Grupo1 e 4	1,7503		Iguais	X
	4	4,4017				Grupo2 e 3	1,0780		Iguais	X
						Grupo2 e 4	1,2657		Iguais	X
						Grupo3 e 4	2,3437		Iguais	X

**TABELA C6 – RESULTADOS DA ANOVA E TESTE TUKEY-KRAMER PARA A ESTRUTURA MA(2) (CONTINUAÇÃO)**

Valor Previsto	Grupos	Média Grupo	Média Geral	Estatís. F	p-valor	Grupos	Dif. Abs.	Interv. Crítico	Comp. médias	Melhor Grupo
e144	1	2,6497	2,5899	0,5185	6,70E-01	Grupo1 e 2	0,1380	1,1924	Iguais	X
	2	2,7877				Grupo1 e 3	0,4019		Iguais	X
	3	2,2478				Grupo1 e 4	0,0246		Iguais	X
	4	2,6743				Grupo2 e 3	0,5398		Iguais	X
						Grupo2 e 4	0,1133		Iguais	X
						Grupo3 e 4	0,4265		Iguais	X
e145	1	2,6596	2,5480	1,1176	3,42E-01	Grupo1 e 2	0,3594	1,3667	Iguais	X
	2	3,0190				Grupo1 e 3	0,5879		Iguais	X
	3	2,0717				Grupo1 e 4	0,2181		Iguais	X
	4	2,4415				Grupo2 e 3	0,9473		Iguais	X
						Grupo2 e 4	0,5775		Iguais	X
						Grupo3 e 4	0,3698		Iguais	X
e146	1	2,6673	2,4177	1,2621	2,87E-01	Grupo1 e 2	0,1667	0,7924	Iguais	X
	2	2,5006				Grupo1 e 3	0,5815		Iguais	X
	3	2,0858				Grupo1 e 4	0,2504		Iguais	X
	4	2,4169				Grupo2 e 3	0,4148		Iguais	X
						Grupo2 e 4	0,0837		Iguais	X
						Grupo3 e 4	0,3311		Iguais	X
e147	1	2,6690	2,5253	0,8523	4,66E-01	Grupo1 e 2	0,0867	1,2041	Iguais	X
	2	2,7557				Grupo1 e 3	0,5915		Iguais	X
	3	2,0776				Grupo1 e 4	0,0701		Iguais	X
	4	2,5989				Grupo2 e 3	0,6782		Iguais	X
						Grupo2 e 4	0,1569		Iguais	X
						Grupo3 e 4	0,5213		Iguais	X
e148	1	2,6551	2,3913	3,2302	2,24E-02	Grupo1 e 2	0,0631	0,6939	Iguais	X
	2	2,7182				Grupo1 e 3	0,5817		Iguais	X
	3	2,0734				Grupo1 e 4	0,5367		Iguais	X
	4	2,1184				Grupo2 e 3	0,6448		Iguais	X
						Grupo2 e 4	0,5997		Iguais	X
						Grupo3 e 4	0,0451		Iguais	X
e149	1	2,6740	2,3921	5,2012	1,56E-03	Grupo1 e 2	0,0965	0,6175	Iguais	X
	2	2,7705				Grupo1 e 3	0,5478		Iguais	X
	3	2,1262				Grupo1 e 4	0,6763		Diferentes	4
	4	1,9977				Grupo2 e 3	0,6443		Diferentes	3
						Grupo2 e 4	0,7728		Diferentes	4
						Grupo3 e 4	0,1285		Iguais	X
e150	1	2,6571	2,8398	3,77E-02		Grupo1 e 2	0,1069	0,6634	Iguais	X
	2	2,5503				Grupo1 e 3	0,5934		Iguais	X
	3	2,0637				Grupo1 e 4	0,5656		Iguais	X
	4	2,0916				Grupo2 e 3	0,4865		Iguais	X
						Grupo2 e 4	0,4587		Iguais	X
						Grupo3 e 4	0,0278		Iguais	X



**TABELA C7 – RESULTADOS DA ANOVA E TESTE TUKEY-KRAMER PARA A ESTRUTURA ARMA(1,1)**

Valor Previsto	Grupos	Média Grupo	Média Geral	Estatís. F	p-valor	Grupos	Dif. Abs.	Interv. Crítico	Comp. médias	Melhor Grupo
e136	1	1,7917	1,4616	2,3097	7,59E-02	Grupo1 e 2	0,0216	0,9456	Iguais	X
	2	1,8133				Grupo1 e 3	0,6590		Iguais	X
	3	1,1327				Grupo1 e 4	0,6830		Iguais	X
	4	1,1087				Grupo2 e 3	0,6807		Iguais	X
						Grupo2 e 4	0,7047		Iguais	X
						Grupo3 e 4	0,0240		Iguais	X
e137	1	2,0809	1,3861	12,0455	1,46E-07	Grupo1 e 2	0,8152	0,4953	Diferentes	2
	2	1,2657				Grupo1 e 3	0,9382		Diferentes	3
	3	1,1427				Grupo1 e 4	1,0256		Diferentes	4
	4	1,0553				Grupo2 e 3	0,1230		Iguais	X
						Grupo2 e 4	0,2104		Iguais	X
						Grupo3 e 4	0,0874		Iguais	X
e138	1	2,2494	1,4492	16,1096	6,71E-10	Grupo1 e 2	0,9194	0,4943	Diferentes	2
	2	1,3300				Grupo1 e 3	1,1217		Diferentes	3
	3	1,1277				Grupo1 e 4	1,1597		Diferentes	4
	4	1,0897				Grupo2 e 3	0,2024		Iguais	X
						Grupo2 e 4	0,2404		Iguais	X
						Grupo3 e 4	0,0380		Iguais	X
e139	1	2,3892	1,5646	13,7188	1,57E-08	Grupo1 e 2	0,7490	0,5945	Diferentes	2
	2	1,6402				Grupo1 e 3	1,2475		Diferentes	3
	3	1,1417				Grupo1 e 4	1,3021		Diferentes	4
	4	1,0871				Grupo2 e 3	0,4985		Iguais	X
						Grupo2 e 4	0,5531		Iguais	X
						Grupo3 e 4	0,0545		Iguais	X
e140	1	2,4768	1,5128	21,9461	3,63E-13	Grupo1 e 2	1,1245	0,5087	Diferentes	2
	2	1,3522				Grupo1 e 3	1,3352		Diferentes	3
	3	1,1416				Grupo1 e 4	1,3961		Diferentes	4
	4	1,0806				Grupo2 e 3	0,2107		Iguais	X
						Grupo2 e 4	0,2716		Iguais	X
						Grupo3 e 4	0,0609		Iguais	X
e141	1	2,5734	1,5265	24,6964	1,14E-14	Grupo1 e 2	1,2812	0,5160	Diferentes	2
	2	1,2922				Grupo1 e 3	1,4335		Diferentes	3
	3	1,1399				Grupo1 e 4	1,4728		Diferentes	4
	4	1,1006				Grupo2 e 3	0,1523		Iguais	X
						Grupo2 e 4	0,1916		Iguais	X
						Grupo3 e 4	0,0393		Iguais	NA
e142	1	2,6567	1,6616	9,9050	2,60E-06	Grupo1 e 2	1,0548	0,8019	Diferentes	2
	2	1,6019				Grupo1 e 3	1,5100		Diferentes	3
	3	1,1467				Grupo1 e 4	1,4153		Diferentes	4
	4	1,2413				Grupo2 e 3	0,4552		Iguais	X
						Grupo2 e 4	0,3606		Iguais	X
						Grupo3 e 4	0,0947		Iguais	X
e143	1	2,6738	1,5419	25,0980	6,88E-15	Grupo1 e 2	1,3824	0,5538	Diferentes	2
	2	1,2913				Grupo1 e 3	1,5362		Diferentes	3
	3	1,1375				Grupo1 e 4	1,6086		Diferentes	4
	4	1,0651				Grupo2 e 3	0,1538		Iguais	X
						Grupo2 e 4	0,2262		Iguais	X
						Grupo3 e 4	0,0724		Iguais	X

**TABELA C7 – RESULTADOS DA ANOVA E TESTE TUKEY-KRAMER PARA A ESTRUTURA ARMA(1,1) (CONTINUAÇÃO)**

Valor Previsto	Grupos	Média Grupo	Média Geral	Estatística F	p-valor	Grupos	Dif. Abs.	Intervalo Crítico	Comp. médias	Melhor Grupo
e144	1	2,7013	1,5670	24,5447	1,38E-14	Grupo1 e 2	1,3394	0,5645	Diferentes	2
	2	1,3619				Grupo1 e 3	1,5627		Diferentes	3
	3	1,1386				Grupo1 e 4	1,6349		Diferentes	4
	4	1,0664				Grupo2 e 3	0,2232		Iguais	X
						Grupo2 e 4	0,2954		Iguais	X
						Grupo3 e 4	0,0722		Iguais	X
e145	1	2,7278	1,6680	13,2822	2,81E-08	Grupo1 e 2	1,3587	0,7191	Diferentes	2
	2	1,3692				Grupo1 e 3	1,5914		Diferentes	3
	3	1,1365				Grupo1 e 4	1,2893		Diferentes	4
	4	1,4385				Grupo2 e 3	0,2327		Iguais	X
						Grupo2 e 4	0,0693		Iguais	X
						Grupo3 e 4	0,3021		Iguais	X
e146	1	2,7513	3,8386	1,6479	1,78E-01	Grupo1 e 2	7,5427	12,4174	Iguais	X
	2	10,2940				Grupo1 e 3	1,6096		Iguais	X
	3	1,1417				Grupo1 e 4	1,5836		Iguais	X
	4	1,1677				Grupo2 e 3	9,1523		Iguais	X
						Grupo2 e 4	9,1263		Iguais	X
						Grupo3 e 4	0,0260		Iguais	X
e147	1	2,7759	1,7365	10,4476	1,25E-06	Grupo1 e 2	0,8981	0,8738	Diferentes	2
	2	1,8779				Grupo1 e 3	1,6309		Diferentes	3
	3	1,1450				Grupo1 e 4	1,6289		Diferentes	4
	4	1,1471				Grupo2 e 3	0,7328		Iguais	X
						Grupo2 e 4	0,7308		Iguais	X
						Grupo3 e 4	0,0020		Iguais	X
e148	1	2,7658	2,9503	1,2135	3,04E-01	Grupo1 e 2	3,7781	8,2828	Iguais	X
	2	6,5438				Grupo1 e 3	1,4798		Iguais	X
	3	1,2860				Grupo1 e 4	1,5601		Iguais	X
	4	1,2057				Grupo2 e 3	5,2578		Iguais	X
						Grupo2 e 4	5,3382		Iguais	X
						Grupo3 e 4	0,0803		Iguais	X
e149	1	2,8048	1,5915	23,6259	4,36E-14	Grupo1 e 2	1,4401	0,6150	Diferentes	2
	2	1,3648				Grupo1 e 3	1,6662		Diferentes	3
	3	1,1386				Grupo1 e 4	1,7471		Diferentes	4
	4	1,0578				Grupo2 e 3	0,2262		Iguais	X
						Grupo2 e 4	0,3070		Iguais	X
						Grupo3 e 4	0,0809		Iguais	X
e150	1	2,8316	2,1254	3,2823	2,09E-02	Grupo1 e 2	0,6009	2,3928	Iguais	X
	2	3,4325				Grupo1 e 3	1,6918		Iguais	X
	3	1,1398				Grupo1 e 4	1,7339		Iguais	X
	4	1,0977				Grupo2 e 3	2,2927		Iguais	X
						Grupo2 e 4	2,3349		Iguais	X
						Grupo3 e 4	0,0421		Iguais	X



# Livros Grátis

( <http://www.livrosgratis.com.br> )

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)  
[Baixar livros de Literatura de Cordel](#)  
[Baixar livros de Literatura Infantil](#)  
[Baixar livros de Matemática](#)  
[Baixar livros de Medicina](#)  
[Baixar livros de Medicina Veterinária](#)  
[Baixar livros de Meio Ambiente](#)  
[Baixar livros de Meteorologia](#)  
[Baixar Monografias e TCC](#)  
[Baixar livros Multidisciplinar](#)  
[Baixar livros de Música](#)  
[Baixar livros de Psicologia](#)  
[Baixar livros de Química](#)  
[Baixar livros de Saúde Coletiva](#)  
[Baixar livros de Serviço Social](#)  
[Baixar livros de Sociologia](#)  
[Baixar livros de Teologia](#)  
[Baixar livros de Trabalho](#)  
[Baixar livros de Turismo](#)