



UNIVERSIDADE FEDERAL DO PARÁ
CENTRO DE CIÊNCIAS EXATAS E NATURAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA E ESTATÍSTICA

Félix Lélis da Silva

*Monitoramento da Fase de Filtração e Secagem da Polpa de
Cristais de Hidrato no Processo de Extração da Alumina
Através de Abordagens Dinâmicas Multivariadas*

Belém/PA
2007

Livros Grátis

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

Milhares de livros grátis para download.

UNIVERSIDADE FEDERAL DO PARÁ
CENTRO DE CIÊNCIAS EXATAS E NATURAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA E ESTATÍSTICA

Félix Lélis da Silva

*Monitoramento da Fase de Filtração e Secagem da Polpa de
Cristais de Hidrato no Processo de Extração da Alumina
Através de Abordagens Dinâmicas Multivariadas*

Dissertação de Conclusão de Curso apresentada
ao Programa de Pós-Graduação em Matemática e
Estatística da Universidade Federal do Pará para
obtenção do Título de Mestre em Matemática e
Estatística.

Orientador: Prof. Dr. Joaquim Carlos Barbosa Queiroz

Belém/PA
2007

UNIVERSIDADE FEDERAL DO PARÁ
CENTRO DE CIÊNCIAS EXATAS E NATURAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA E ESTATÍSTICA

FÉLIX LÉLIS DA SILVA

*Monitoramento da Fase de Filtração e Secagem da Polpa de
Cristais de Hidrato no Processo de Extração da Alumina
Através de Abordagens Dinâmicas Multivariadas*

Prof. Dr. Marcus Pinto da Costa da Rocha
Coord. do Programa de Pós-Graduação em Matemática e Estatística - UFPA

Banca Examinadora

Prof. Dr. Joaquim Carlos Barbosa Queiroz
Faculdade de Estatística - UFPA
Orientador

Prof. Dr. Antônio Moraes da Silveira
Faculdade de Ciência da Computação - UFPA
Co-Orientador

Prof. Dr. Heliton Ribeiro Tavares
Faculdade de Estatística - UFPA
Examinador

Prof. Dr. Carlos Tavares Júnior
Faculdade de Engenharia Elétrica - UFPA
Examinador

DEDICATÓRIA

*Este trabalho é fruto de um sonho que até alguns anos parecia estar tão longe, porém no alto de minha persistência jamais deixei de sonhá-lo, do contrário, estaria fadado as angustias e a insignificância de um perdedor que tem como adversário a si próprio. Portanto, o mesmo, em toda a sua extensão, é dedicado a mim e a minha família em especial a minha esposa Andréa Pereira Lélis (**AMOR**) e meus filhos Gabriel Lélis (**BIEL**), Maria Eduarda Lélis (**DUDA**), Carolina M. Braga (**CAROL**) e aos amigos que de alguma forma contribuíram para a construção deste apanhado.*

AGRADECIMENTOS

À Universidade Federal do Pará;

À minha família Lélis, em especial a minha queridíssima mãe Maria Esmeni dos Santos Lélis ao meu pai Manoel Farias pelos ensinamentos de valores e aos irmãos de grandes personalidades: Flávio Lélis, Francisco Lélis e Franklin Lélis;

Ao amigo, professor e orientador Joaquim Carlos Barbosa Queiroz, pela busca incessante do conhecimento e pelo incrível dinamismo educacional do ensino superior contribuindo de forma coerente, eficaz e eficiente no aprendizado de seus alunos;

Ao professor e co-orientador Antônio Morais pelos ensinamentos;

A todos os professores do Programa de Pós Graduação em Matemática e Estatística;

Ao amigo Msc. e Doutorando em Engenharia Elétrica Vanilson Gomes Pereira, pelas idéias na estruturação das Redes Neurais, as quais foram de grandiosa e excelente contribuição para este trabalho;

Ao Departamento de Estatística;

Ao Departamento de Matemática;

Ao Programa de Pós Graduação em Matemática e Estatística - PPGME;

À Américo José Preto Borges da Albras pela cessão dos dados;

E a todos, os seres iluminados que de alguma forma, contribuíram para a realização deste trabalho.

Algo que jamais um ser de fazer é desistir dos seus sonhos, pois estará assim desistindo de si mesmo!

Félix Lélis

Resumo

SILVA F. L. Monitoramento da Fase de Filtração e Secagem da Polpa de Cristais de Hidrato no Processo de Extração da Alumina Através de Abordagens Dinâmicas Multivariadas. Dissertação - 2007 - PPGME/UFPA, Belém - PA, Brasil.

As conseqüências do desenvolvimento industrial levam à aplicação de novas ou melhoradas tecnologias, as quais exigem avaliações criteriosas para que sejam implementadas de forma adequada às necessidades da natureza do processo produtivo. Assim, através de métodos de análise e diagnóstico, pretende-se estruturar processos cada vez mais eficientes e até mesmo inteligentes, seja no controle ou monitoramento da produção.

É fato que, a evolução do processo produtivo está relacionada diretamente à evolução do processo tecnológico, pois nas últimas décadas com o emergente crescimento da produção de sistemas computacionais, houve a retomada de muitas pesquisas que haviam sido abandonadas por falta de máquinas capazes de processar algoritmos mais complexos. Sendo assim, o crescimento do processo de industrialização trouxe consigo a necessidade da implementação de sistemas inteligentes, caracterizando um processo de formulação da inteligência artificial (**I.A**). Com isso, criou-se máquinas e equipamentos capazes de processar, avaliar e acompanhar tarefas antes somente gerenciadas por seres humanos.

Devido necessidades, estas tarefas de processar, avaliar e acompanhar processos, estendeu-se do processo tecnológico as mais diversas áreas do conhecimento, como por exemplo: ***Estudos de Fatores Climáticos, Geociências no auxílio do processo de perfuração de poços e Edificações, Saúde, Modelagens Estatísticas, Oceanografia, Biologia, etc.*** Todo este processo teve como base a evolução dos sistemas tecnológicos computacionais fruto do processo dinâmico da revolução industrial, pois a difusão das tecnologias digitais de informação contribuíram com aparelhos eletrônicos cada vez mais eficazes e com grande capacidade de processamento.

Portanto, na busca da otimização da ***Fase de Filtração e Secagem da Polpa de Cristais de Hidrato no Processo de Extração da Alumina***, neste trabalho se ajustou e avaliou modelos inteligentes de Redes Neurais e de Modelos de Séries Temporais Multivariadas com Função de Transferência, cujos apresentaram excelente capacidade de monitorar e avaliar as respostas obtidas do processo. No entanto, com os resultados validados se concluiu que as ***RNA Multicamadas Diretas com algoritmo de treinamento Levenberg Marquardt Backpropagation*** apresentam melhores resultados na identificação do padrão gerador das séries representativas das variáveis estudadas, em comparação aos resultados obtidos através dos Modelos Multivariados com ***FT***, caracterizando-as com maior capacidade de monitoramento e avaliação do processo.

Palavra-Chave: Redes Neurais, Funções de Transferência Multivariáveis, Alumina.

Abstract

SILVA F. L. Monitoring the filtration stage of the pulp of crystals hydrate in the process of extraction of alumina through dynamic multivariate approaches. Dissertation - 2007 - PPGME/UFPA, Belém - PA, Brazil.

The consequences of industrial development leading to implementation of new technologies, which require wise tests to be implemented adequately the needs of the nature of the production process. Thus, through methods of review and diagnosis, it is intended structuring processes increasingly efficient and even intelligent, either in the control or monitoring of production.

It is a fact that, the evolution of the production process is linked directly to technological developments of the process, because in the last decades with the emerging growth of the production of computer systems, there was a resumption of many searches that were forgotten for lack of machines capable of processing algorithms more complex. Thus, the growth of the process of industrialization brought with the need for the implementation of intelligent systems, featuring a process of formulation of artificial intelligence (A.I.).

In this way, It has been able machines and equipment to process, evaluate and monitor tasks before only managed by humans. Due to needs, these tasks to process, evaluate and monitor processes, extended the technological process the most diverse areas of knowledge, such as: *studying of climatic factors, geoscience helping in the process of drilling of wells and buildings, health, statistical modeling, oceanography, biology, etc.* This whole process was based on the evolution of technological computer systems fruit of the process of industrial revolution, as the digital dissemination of information contributed with electronic devices increasingly effective and with great capacity for processing.

Therefore, in search of the optimization of *phase filtration and drying of the flesh of crystals hydrate in the process of extraction of alumina*, this work fit and evaluated intelligent models of neural networks models and series of storms to monitor and evaluate the responses obtained in the process. However, with the results validated, It's concluded that the *RNA Multilayers with the training algorithm Levenberg Marquadt Backpropagation* show better results in the identification of the pattern generator of the series representing variables studied, compared to the results obtained through the models multivariate with *FT*, characterizing them with greater capacity for monitoring and evaluation of the process

Key-words: Neural Networks, Functions of Multivariate Transfer, Alumina.

Índice

Resumo	vii
Abstract	viii
Lista de Tabelas	xiii
Lista de Figuras	xv
1 Introdução	1
1.1 Principais Considerações	2
1.2 Problema do Trabalho	5
1.3 Definição do Objetivo	5
1.4 Objetivos Especificos	5
1.5 Justificativa e Importância do Trabalho	6
1.6 A Hipótese Básica da Tese	6
1.7 Estrutura do Trabalho	7
2 Modelos Probabilísticos de Séries Temporais em Processos Dinâmicos Univariados	8
2.1 Introdução	8
2.2 Processos Estocásticos	9
2.2.1 Características de um Processos Estocásticos	9
2.3 A estacionariedade	12
2.4 Processos Estacionários	12
2.5 Processos Não-Estacionários	13
2.5.1 Modelos <i>ARIMA</i> (p, d, q) de Box & Jenkins	13
2.5.2 Análises das Funções de Autocorrelações	17
2.5.3 Análise das Funções de Autocorrelações Parciais	18
2.5.4 Critérios de Seleção - AIC e SBC	18
3 Modelos Probabilísticos de Séries Temporais em Processos Dinâmicos Multivariados	20
3.1 Introdução	20
3.2 Propriedades de Segunda Ordem de Séries Temporais Multivariada	20
3.3 Modelamento com Processo <i>AR</i> Multivariado	22
3.4 Modelamento com Processo <i>MA</i> Multivariado	23

3.5	Processos <i>ARMA</i> Multivariados	23
3.6	Processos <i>ARIMA</i> Multivariados	25
3.7	Modelos de Função de Transferência (MFT) ou Modelos Causais	25
3.8	Funções de Transferências de Única Entrada	26
3.9	Algumas Funções Típicas da Resposta do Impulso	28
3.10	Funções de Correlação Cruzada e Modelos de Função de Transferência . . .	29
3.10.1	Funções de Correlação Cruzada - FCC	29
3.10.2	Relação entre Função de Correlação Cruzada e Função de Transferência	30
3.11	Construção do Modelo de Função de Transferência	32
3.11.1	Função de Correlação Cruzada Amostral	32
3.11.2	Identificação do Modelo de Função de Transferência	33
3.11.3	Identificação do Modelo de Ruído	34
3.11.4	Estimativas dos Parâmetros do Modelo de Função de Transferência .	35
3.11.5	Diagnóstico dos Modelos de Função de Transferência	36
3.11.6	Estimação do Modelo de Função de Transferência	38
3.12	Validação de Modelos de Função de Transferência	38
3.12.1	Erro Quadrático Médio de Séries de Entrada e Saída Estacionárias .	38
3.12.2	Resumo	40
4	Sistemas de Implementações Inteligentes - Redes Neurais	41
4.1	Introdução	41
4.2	Redes Neurais Artificiais - RNA's	42
4.2.1	Funções de ativação	44
4.3	Algumas Teorias Referentes a Algoritmos de Redes Neurais	48
4.3.1	Modelos de Redes Neurais de Warren McCulloch e Pitts	48
4.3.2	Neurônios de Remmelhart, Hilton e Williams - RHW	49
4.3.3	Modelos de Saída Binária - Binário	50
4.3.4	RNA's com Multicamadas de Neurônios	51
4.3.5	Propagação do erro	52
4.3.6	Retropropagação do erro	53
4.3.7	Processo de Aprendizado por Correção de Erros	53
4.3.8	Redes Neurais: Pontos Fortes e Limitações	55
4.4	A Estatística do Processo de Aprendizagem	56
4.4.1	O Erro de Expectativa (ϵ)	58
4.4.2	Verificação dos Termos "Bias" e Variância	60
4.5	Algoritmo de Levenberg Marquardt - <i>LMA</i>	62
4.6	Normalização dos Dados de Entrada	64
4.6.1	Normalização por Mínimo e Máximo	64
4.6.2	Normalização por Desvio Padrão	64
4.7	Desenvolvimento e Aplicações de Redes Neurais	65
4.8	Redes Neurais em Sistemas de Controle	68
4.8.1	Sistemas de Controle Dinâmicos	69
4.8.2	Identificação de Sistemas Dinâmicos	70
4.8.3	Retropropagação Dinâmica (Dynamic Back Propagation)	71

4.8.4	Sistemas de Controle de Processos	71
4.8.5	Análise de Dados	72
4.8.6	Previsões com Redes Neurais	73
5	Análise de Dados	74
5.1	Processo de filtração dos cristais de hidrato	74
5.1.1	Filtros de discos rotativos	76
5.1.2	Cr�terios de Eventuais Paradas no Processo	77
5.1.3	Efici�ncia do Processo	77
5.2	An�lise dos Dados Atrav�s de Modelo de Rede Neural	78
5.3	Estrutura��o e Treinamento da Rede - 1� Ensaio	79
5.3.1	Simula��o de Treinamento da Vari�vel N�vel de Umidade- 1� Ensaio .	81
5.3.2	Simula��o de Treinamento da Vari�vel N�vel de Soda- 1� Ensaio . . .	82
5.3.3	Simula��o de Valida��o da Vari�vel N�vel de Umidade- 1� Ensaio . . .	83
5.3.4	Simula��o de Valida��o da Vari�vel N�vel de Soda- 1� Ensaio	84
5.3.5	Simula��o das Valida��es Desnormalizadas- 1� Ensaio	85
5.4	Estrutura��o e Treinamento da Rede - 2� Ensaio	86
5.4.1	Simula��o de Treinamento da Rede - 2� Ensaio	87
5.4.2	Simula��o de Treinamento da Vari�vel N�vel de Umidade - 2� Ensaio .	88
5.4.3	Simula��o de Treinamento da Vari�vel N�vel de Soda - 2� Ensaio . . .	89
5.4.4	Simula��o de Valida��o da Vari�vel N�vel de Umidade - 2� Ensaio . .	90
5.4.5	Simula��o de Valida��o da Vari�vel N�vel de Soda - 2� Ensaio	91
5.4.6	Simula��o das Valida��es Desnormalizadas - 2� Ensaio	92
5.5	Estrutura��o e Treinamento da Rede - 3� Ensaio	93
5.5.1	Treinamento da Rede - 3� Ensaio	94
5.5.2	Simula��o de Treinamento da Vari�vel Umidade - 3� Ensaio	95
5.5.3	Simula��o de Treinamento da Vari�vel Soda - 3� Ensaio	96
5.5.4	Simula��o de Valida��o da Vari�vel Umidade - 3� Ensaio	97
5.5.5	Simula��o de Valida��o da Vari�vel Soda - 3� Ensaio	98
5.5.6	Simula��o das Valida��es Desnormalizadas - 3� Ensaio	99
5.5.7	S�nteses dos Modelos de RNA Ajustados	100
5.6	Modelagem Multivariada com Fun��o de Transfer�ncia	101
5.6.1	Modelagem para a Vari�vel N�vel de Umidade	103
5.6.2	An�lise das <i>FAC</i> e <i>FACP</i>	103
5.7	Ajuste de Modelos ARIMA	105
5.7.1	Ajuste de Modelos ARIMA para a Vari�vel X_1	105
5.7.2	Ajuste de Modelos ARIMA para a Vari�vel X_2	106
5.7.3	Ajuste de Modelos ARIMA para a Vari�vel X_3	107
5.8	Correla��es Cruzadas	108
5.8.1	An�lise da Correla��o Cruzada entre Y_1 vs X_1	108
5.8.2	An�lise da Correla��o Cruzada entre Y_1 vs X_2	109
5.8.3	An�lise da Correla��o Cruzada entre Y_1 vs X_3	109
5.8.4	Ajuste dos Modelos de Fun��o de Transfer�ncia	109
5.8.5	Estimativas para a Estrutura��o do Modelo 1 em 5.7	110
5.8.6	An�lise Res�dual para o Modelo de F.T em 5.7	111

5.8.7	Estimativas para a Estruturação do Modelo 2	112
5.8.8	Análise Residual para o Modelo de F.T em 5.8	112
5.8.9	Estimativas para a Estruturação do Modelo 3 em 5.9	113
5.8.10	Análise Residual para o Modelo 3 de F.T em 5.9	114
5.8.11	Seleção do Melhor Modelo para o Nível de Umidade	115
5.8.12	Teste de Validação para os Modelos do Nível de Umidade	116
5.9	Modelagem para a Variável Nível de Soda	118
5.9.1	Análise das <i>FAC</i> e <i>FACP</i>	118
5.9.2	Ajuste do Modelo de Função de Transferência - Soda	119
5.9.3	Estimativas para a Estruturação do Modelo 1 em 5.13	119
5.9.4	Análise Residual para o Modelo de F.T em 5.15	120
5.9.5	Estimativas para a Estruturação do Modelo 2 em 5.14	121
5.9.6	Análise Residual para o Modelo de F.T em 5.16	122
5.9.7	Seleção do Melhor Modelo para o Nível de Soda	123
5.9.8	Teste de Validação para os Modelos do Nível de Soda	124
5.10	Conclusões	125
6	Considerações Finais	126
A	ANEXO	128
A.1	Introdução	128
A.1.1	Algoritmo para Obtenção de Rede Neural - MATLAB 7.0	128
A.1.2	Algoritmo para Obtenção do Modelo de Séries Temporais - SAS-9	137
A.1.3	Tabela do Banco de Dados	148
	Bibliografia	155

Lista de Tabelas

5.1	Comparativo para Decisão entre as Três Redes Neurais Propostas	100
5.2	Cálculo e Avaliação dos Parâmetros da Variável X1	105
5.3	Estudo das Auto correlações dos Resíduos	105
5.4	Cálculo e Avaliação dos Parâmetros da Variável X2	106
5.5	Estudo das Auto correlações dos Resíduos	106
5.6	Cálculo e Avaliação dos Parâmetros da Variável X3	107
5.7	Estudo das Auto correlações dos Resíduos	107
5.8	Cálculo dos Parâmetros Estimados para o Modelo de F.T.	110
5.9	Estudo das Auto correlações dos Resíduos	111
5.10	Estudo das Correlação Cruzada dos Resíduos vs X1	111
5.11	Estudo das Correlação Cruzada dos Resíduos vs X2	111
5.12	Estudo das Correlação Cruzada dos Resíduos vs X3	111
5.13	Cálculo dos Parâmetros Estimados para o Modelo de F.T.	112
5.14	Estudo das Auto correlações dos Resíduos	112
5.15	Estudo das Correlação Cruzada dos Resíduos vs X1	113
5.16	Estudo das Correlação Cruzada dos Resíduos vs X3	113
5.17	Cálculo dos Parâmetros Estimados para o Modelo de F.T.	113
5.18	Estudo das Auto correlações dos Resíduos	114
5.19	Estudo das Correlação Cruzada dos Resíduos vs X1	114
5.20	Estudo das Correlação Cruzada dos Resíduos vs X3	114
5.21	Critérios para Seleção do Melhor Modelo em Ajuste Ajustado	116
5.22	Critérios para Seleção do Melhor Modelo em Validação	117
5.23	Cálculo dos Parâmetros Estimados para o Modelo de F.T.	119
5.24	Estudo das Autocorrelação dos Resíduos do Modelo	120
5.25	Estudo das Correlação Cruzada dos Resíduos vs X2	120
5.26	Estudo das Correlação Cruzada dos Resíduos vs X3	120
5.27	Cálculo dos Parâmetros Estimados para o Modelo de F.T.	121
5.28	Estudo das Autocorrelação dos Resíduos do Modelo	122
5.29	Estudo das Correlação Cruzada dos Resíduos vs X2	122
5.30	Estudo das Correlação Cruzada dos Resíduos vs X3	122
5.31	Critérios para Seleção do Melhor Modelo Ajustado	123
5.32	Critérios para Seleção do Melhor Modelo Validado	124
A.1	Valores Fornecidos pelo Processo	148
A.2	Continuação da Tabela A.1	149
A.3	Continuação da Tabela A.2	150

A.4	Continuação da Tabela A.3	151
A.5	Continuação da Tabela A.4	152
A.6	Continuação da Tabela A.5	153
A.7	Continuação da Tabela A.6	154

Lista de Figuras

2.1	Estrutura de Modelos Univariados	8
3.1	Estrutura de Modelos de Função de Transferência	26
3.2	Sistema de Função de Transferência Dinâmico	27
4.1	Estrutura de Neurônio Biológico	43
4.2	Neurônios Artificial apresentado por McCulloch	44
4.3	Função de ativação linear	44
4.4	Função de ativação degrau	45
4.5	Função de ativação rampa	45
4.6	Função de ativação sigmóide -(logsig)	47
4.7	Função de ativação tangente hiperbólica -(tansig)	47
4.8	Rede com neurônio (<i>RHW</i>)	49
4.9	Erro de aproximação do (<i>RHW</i>)	49
4.10	Neurônio artificial do tipo (<i>Adaline</i>)	50
4.11	Estrutura de uma Rede com Multicamadas ou Redes Neurais Complexas .	51
4.12	Propagação do erro na rede	52
4.13	Retropropagação do erro na rede	53
4.14	Superfície de erro com mínimos locais e o mínimo global	54
4.15	Modelo Matemático Regressivo	57
4.16	Modelo Físico de Rede Neural	58
4.17	Efeito da normalização dos dados de entrada	65
4.18	Esquema de um sistema de controle dinâmico	69
4.19	Identificação de um sistema dinâmico	70
4.20	Treinamento de uma rede neural para diagnóstico de falhas	72
4.21	Valores Previstos	73
5.1	Processo de lavagem e filtração da polpa de hidrato	74
5.2	Emissões de Gases por Indústrias	76
5.3	Emissões de Gases por Queimadas	76
5.4	Filtro de disco rotativo	77
5.5	Substituição do Tecido Filtrante	77
5.6	Comportamentos das Variáveis do processo	78
5.7	Evolução da Função Erro $E(W)$	80
5.8	Treinamento da Variável Umidade	81
5.9	Erro de Treinamento da Variável Umidade	81

5.10	Treinamento da Variável Soda	82
5.11	Erro de Treinamento da Variável Soda	82
5.12	Validação da Variável Nível de Umidade	83
5.13	Erro de Validação da Variável Nível de Umidade	83
5.14	Validação da Variável Nível de Soda	84
5.15	Erro de Validação da Variável Nível de Soda	84
5.16	Validação Desnormalizada da Variável Umidade	85
5.17	Validação Desnormalizada Variável Soda	85
5.18	Evolução da Função Erro $E(W)$	87
5.19	Treinamento da Variável Umidade	88
5.20	Erro de Treinamento da Variável Umidade	88
5.21	Treinamento da Variável Soda	89
5.22	Erro de Treinamento da Variável Soda	89
5.23	Validação da Variável Nível de Umidade	90
5.24	Erro de Validação da Variável Nível de Umidade	90
5.25	Validação da Variável Nível de Soda	91
5.26	Erro de Validação da Variável Nível de Soda	91
5.27	Validação Desnormalizada Variável Umidade	92
5.28	Validação Desnormalizada Variável Soda	92
5.29	Evolução da Função Erro $E(W)$	94
5.30	Treinamento da Variável Umidade	95
5.31	Erro de Treinamento da Variável Umidade	95
5.32	Treinamento da Variável Nível de Soda	96
5.33	Erro de Treinamento da Variável Soda	96
5.34	Validação da Variável Nível de Umidade	97
5.35	Validação da Variável Nível de Soda	97
5.36	Erro de Validação da Variável Nível de Umidade	98
5.37	Erro de Validação da Variável Nível de Soda	98
5.38	Validação Desnormalizada Variável Umidade	99
5.39	Validação Desnormalizada Variável Soda	99
5.40	Níveis de Umidade na Torta Y_1	102
5.41	Níveis de Soda na Torta Y_2	102
5.42	Vazão de Condensado X_1	102
5.43	Vazão de Vapor X_2	102
5.44	Rotação do Tambor X_3	102
5.45	FAC para a Variável X_1	103
5.46	FACP para a Variável X_1	103
5.47	FAC para a Variável X_2	103
5.48	FACP para a Variável X_2	103
5.49	FAC para a Variável X_3	104
5.50	FACP para a Variável X_3	104
5.51	Função de Correlação Cruzada Y_1 vs X_1	108
5.52	Função de Correlação Cruzada Y_1 vs X_2	108
5.53	Função de Correlação Cruzada Y_1 vs X_3	108

5.54	Ajuste do Modelo 5.10 de Função de Transferência	115
5.55	Ajuste do Modelo 5.11 de Função de Transferência	115
5.56	Ajuste do Modelo 5.12 de Função de Transferência	115
5.57	Teste de Validação do 1º Modelo	116
5.58	Teste de Validação do 2º Modelo	116
5.59	Teste de Validação do 3º Modelo	117
5.60	Função de Correlação Cruzada Y_2 vs X_1	118
5.61	Função de Correlação Cruzada Y_2 vs X_2	118
5.62	Função de Correlação Cruzada Y_2 vs X_3	118
5.63	Ajuste do Modelo 5.15 de Função de Transferência	123
5.64	Ajuste do Modelo 5.16 de Função de Transferência	123
5.65	Teste de Validação do 1º Modelo	124
5.66	Teste de Validação do 2º Modelo	124
A.1	Gera a Normalização por Desvio Padrão	128
A.2	Padrões apresentados a rede	129
A.3	Função e parâmetros de treinamento apresentados a rede	130
A.4	Gera a Normalização no Espaço $(-1, 1)$	131
A.5	Gera a Análise Gráfica de Treinamento	132
A.6	Gera a Análise Gráfica de Validação	133
A.7	Gera a Análise Gráfica da Validação Desnormalizada	134
A.8	Algoritmo para iniciar o treinamento e validação	135
A.9	Projeto para seleção dos dados de treinamento e validação	136
A.10	Gera as Séries das Variáveis	137
A.11	Gera a FAC e FACP das Séries	138
A.12	Gera os Modelos ARIMA	139
A.13	Gera os Testes dos Modelos ARIMA	140
A.14	Gera as Funções de Correlação Cruzada	141
A.15	Teste do Modelo 1 de F.T - Umidade	142
A.16	Teste do Modelo 2 de F.T - Umidade	143
A.17	Teste do Modelo 3 de F.T - Umidade	144
A.18	Gera Modelo de F.T - Soda	145
A.19	Teste do Modelo 1 de F.T - Soda	146
A.20	Teste do Modelo 2 de F.T - Soda	147

Capítulo 1

Introdução

O grande interesse neste trabalho está em realizar uma descrição de algumas técnicas mais empregada na atualidade para análise de séries temporais, identificação, monitoramento e controle industrial. Serão mostradas diferentes abordagens e suas particularidades no tratamento das séries temporais multivariadas. Entre essas técnicas se destaca a aplicação de Redes Neurais Artificiais do Tipo Bacpropagation e a modelagem por Função de Transferência baseada na metodologia de Box e Jenkins. Especificamente, essas abordagens serão aplicadas no ajuste de séries temporais multivariadas fornecidas pelo processo industrial responsável pela filtração e secagem da polpa de cristais de hidrato na fase de extração da alumina. Onde os resultados serão utilizados para verificar a melhor adequação em relação ao problema e objetivos vinculado ao estudo, ou seja, identificação e monitoramento do processo.

As técnicas de Redes Neurais, estão sendo cada vez mais empregadas, principalmente quando se pretende modelar memórias associativas, reconhecimento de padrões, representação de funções Booleanas e representação de funções contínuas. Esta técnica é constantemente aplicada com o objetivo de Construção de Modelos, Projeções e Controle nas mais variadas áreas das ciências médicas, biomédicas, da robótica, das telecomunicações, da indústria de transporte, da indústria de prospecção petrolífera, dentre outras relacionada ao conhecimento Humano.

Considerada uma das ferramentas mais eficientes para o estudo referente à modelagem de séries de tempo e com maior utilização pela comunidade científica, a metodologia de Box e Jenkins será aplicada neste trabalho com a finalidade de se construir um modelo que auxilie no processo de comparação entre as técnicas apresentadas. A idéia não é a de dispor as técnicas como antagônicas, mas apresentar como elas podem ser usadas

alternativamente ou conjuntamente para as análises de séries temporais e o controle de processos industriais.

1.1 Principais Considerações

Pode-se considerar que os modelos de Box e Jenkins foram uma grande descoberta em relação ao estudo de séries de tempo. Segundo Morettin e Toloi (2004), determinada série descrita nos instantes t_1, \dots, t_n pode fornecer diversas características do processo que o gerou, tais como:

- a - Investigação do processo gerador da série;
- b - Realização de previsões de valores futuros da série;
- c - Descrição do comportamento da série;
- d - Procura periódica de relevância nos dados.

Conforme Ajoy K. Palit e Dobrivoje Popovic, em seu livro, “*Computational Intelligence in Time Series Forecasting Theory and Engineering Applications*”), uma série temporal é uma seqüência ordenada no tempo de observação (valores) de natureza física ou financeira variável feitas em intervalos igualmente espaçados tempo delta t , representado como um conjunto de valores discretos x_1, x_2, x_3, \dots , etc. Na engenharia p.e: a seqüência de valores é obtidos a partir de sensores por amostragem relacionada aos sinais contínuos, sendo baseados em valores medidos são normalmente comprometidos ruído, em séries temporais sempre é incluído componentes determinísticas, ou seja, um sinal e uma componente estocástica representando o ruído (resíduos) que provocam interferências estatísticas (flutuações) em torno do valores previamente determinados. (*tradução do autor*)

Aduz ainda que, a análise da série em um determinado momento é essencialmente necessário verificar a ocorrência na estrutura da série das componentes (autocorrelação, tendência, sazonalidade, etc), para um possível entendimento do processo dinâmico o qual gerou a mesma no tempo t . No processo Controle de processos ou simplesmente monitoramentos ou avaliações, o tempo previsto a análise de séries temporais é essencial

para ajuste e tomada de decisão e conseqüentemente direcionam as ações a serem tomadas. A análise de séries temporais abrange atividades como:

- definição, classificação, e descrição de séries temporais;
- modelagem com base na utilização de valores selecionados da própria série temporal em estudo;
- previsão de valores futuros.

onde, basicamente, existem duas abordagens da metodologia da análise de séries temporais, são elas:

- abordagem do domínio do tempo, principalmente baseado no uso da função covariância da série cronológica;
- abordagem no domínio da frequência, baseada principalmente na análise espectral de potência ou transformada de Fourier da função autocorrelação.

Para os modelos de função de transferência ($\mathbf{F.T}$), segundo os comentários de José Renes Pinheiro, quando se trata de um sistema é uma propriedade que independe da natureza e da magnitude da entrada; além de possibilitar um sistema dinâmico ser representado por expressões algébricas da variável complexa " \mathbf{S} ". Um detalhe importante é que a FT não fornece informações a respeito da estrutura física do sistema. As mesmas podem ser idênticas quando forem de sistemas fisicamente diferentes, se a FT de um sistema é conhecida, a resposta do mesmo pode ser analisada para diferentes formas de excitação (entrada), com a finalidade de compreender a natureza e o comportamento do sistema, caso a FT possa ser obtida experimentalmente pela introdução de sinais de entrada conhecidos e estudando-se as respostas obtidas, portanto a função de transferência fornece ou pode fornecer uma descrição completa das características dinâmicas do sistema.

Já as redes neurais (RNA), segundo Dahmer (1998), apresentam inúmeras características que estimulam a sua utilização nas mais diversas pesquisas. Entre elas pode-se citar:

- Processamento altamente paralelo e distribuído;

- Capacidade de aprendizagem através da observação de um conjunto de exemplos, sendo esses, com ou sem informação de respostas desejadas a cada estímulo;
- Robustez, isto é, a queda do desempenho gradual no eventual funcionamento irregular ou, até mesmo, inutilização de parte da rede;
- Manutenção do desempenho na presença de ruído e capacidade de lidar com dados incompletos.

Segundo Hecht-Nielsen(1990), uma RNA é uma estrutura que processa informações de forma paralela e distribuída e que consiste de unidades computacionais, as quais podem possuir uma memória local e podem executar operações locais, interconectadas por canais unidirecionais chamadas de *sinapses*, em que cada unidade computacional possui uma única conexão de saída que pode ser dividida em quantas conexões laterais se fizer necessário, sendo que cada uma delas transporta o mesmo sinal (HECHT-NIELSEN, 1990).

O sinal de saída pode ser contínuo ou discreto e o processo executado por cada unidade pode ser definido arbitrariamente com a restrição de que ele deve ser completamente local, isto é, deve depender somente dos valores atuais dos sinais de entrada que chegam até a unidade através das conexões e dos valores armazenados na memória local.

Para De Azevedo (1993), uma rede neural artificial pode ser especificada como um sistema complexo, representado por um grafo arco-rotulado, na qual, cada vértice é um neurônio artificial com estrutura dinâmica e, o grafo apresenta um conjunto finito, não vazio, de vértices juntos com um conjunto, não ordenado, de arcos conectando pares de vértices.

Conforme Campelo, (2005), os sistemas inteligentes vêm sendo aplicados nas mais diversas áreas do conhecimento, tais como, reconhecimento de padrões, mineração de dados, tomada de decisão, identificação, mapeamento e controle de processos dinâmicos, etc. Entre os sistemas inteligentes pode-se destacar dois modelos não lineares: os Modelos Fuzzy e Modelos de Redes Neurais. Ambos apresentam uma propriedade fundamental que impulsionou seu desenvolvimento: são aproximadores universais. Em se tratando de mapeamento Souza et al (1999), descreve que, as redes neurais artificiais (RNA) **Multi**

Layer Perceptron - MLP podem ser utilizadas para mapear os relacionamentos entre as diversas variáveis associadas com o processo de especificação (identificação).

Fernandes et al, (1995), cita que a principal dificuldade na utilização de redes neurais artificiais na previsão de séries temporais, ainda é a determinação da arquitetura ótima da rede, pois ainda não se possui uma metodologia consistente que apresente a melhor configuração da rede para cada série proposta. Baseando-se nessa perspectiva, este trabalho propõe um modelo que otimiza o processo de filtração e secagem da polpa de cristais de hidrato, sendo este, passível de controle através de sistemas inteligentes de Redes Neurais Artificiais.

1.2 Problema do Trabalho

O alto nível de **umidade** e o alto percentual de **soda** na polpa de cristais de hidrato, no processo de fabricação da alumina, é um fator culminante no elevado consumo de combustível, energia elétrica e na redução da qualidade final do produto.

1.3 Definição do Objetivo

Neste trabalho, pretende-se apresentar modelos dinâmicos de Redes Neurais e de Séries Temporais para mapear e monitorar as multivariáveis que influenciam diretamente no processo industrial de Filtração e Secagem da Polpa de Cristais de Hidrato no Processo de Extração da Alumina. E, dentre os modelos apresentados, realizar uma avaliação da robustez, considerando suas características relacionadas à descrição, controle e otimização de processos.

1.4 Objetivos Especificos

1. Construir um modelo dinâmico não linear de Rede Neural Artificial Multicamada Direta com treinamento supervisionado utilizando o algoritmo **Backpropagation** para monitorar o processo de filtração e secagem da polpa de cristais de hidrato, que apresente a melhor aproximação (reconhecimento do padrão do processo) possível.
2. Construir um modelo dinâmico linear com base na metodologia de Box e Jenkins

para *Séries Temporais Multivariadas*, utilizando *Função de Transferência*, para ajustar da forma mais aproximada possível o processo de filtração e secagem da polpa de cristais de hidrato. Apresentando assim, capacidade de minitoramento do processo.

3. Avaliar, através do teste de validação, o modelo mais representativo para o monitoramento do processo de extração da alumina.

1.5 Justificativa e Importância do Trabalho

No decorrer dos anos foram implementados e apresentados no meio científico, vários estudos sobre novas técnicas de modelagem de séries de temporais, cujas finalidades são estudar periodicidades e descrever comportamentos, objetivando melhorar as investigações e suposições feitas sobre as mesmas. Isto é, através da formulação de modelos que apresentem projeções com menor variabilidade possível, para poder explicar a real evolução da série no decorrer do tempo t .

Este trabalho possui sua relevância por mostrar que apesar da grande distinção entre os métodos de RNA's e a metodologia de Box e Jenkins, os mesmos apresentam excelentes resultados, quando estão direcionados à um objetivo comum, como por exemplo, avaliar e controlar processos através de análises prévias de ajuste do modelo. Destaque-se que, dependendo das situações, é mais favorável a aplicação de determinada metodologia de forma conjugada do que individualizada para se obter melhores resultados.

1.6 A Hipótese Básica da Tese

Será que as respostas obtidas através da modelagem de *Redes Neurais Artificiais* - RNA's, para o controle de umidade e do nível de soda cáustica presente na polpa de cristais de hidrato no processo de fabricação de alumina são mais significativas, que as obtidas por Modelos de *Séries Temporais Multivariados*, devido a sua alta capacidade de reconhecimento de padrões.

1.7 Estrutura do Trabalho

Este trabalho encontra-se dividido em seis capítulos e um anexo, com o seguinte conteúdo:

- **Capítulo 1:** Neste capítulo, estão as idéias iniciais referentes ao projeto, onde são englobadas as principais considerações, as justificativas e a relevância do trabalho, a hipótese básica da dissertação e os objetivos;
- **Capítulo 2:** Neste capítulo, inicia-se uma introdução aos conceitos básicos dos modelos probabilísticos de séries temporais univariados, dando mais ênfase aos modelos referentes à metodologia de Box e Jenkins;
- **Capítulo 3:** Neste capítulo, faz-se referência aos modelos probabilísticos de séries temporais multivariados ressaltando as funções de transferência, modelos multivariados e função de correlação cruzada;
- **Capítulo 4 :** Neste capítulo, são apresentados os Sistemas Inteligentes, mais especificamente os sistemas de Redes Neurais Artificiais. São demonstradas algumas contribuições referentes a algoritmos de redes neurais, pontos fortes e limitações, além da utilização de funções de transferência (funções de ativação).
- **Capítulo 5:** Neste capítulo, são apresentados os resultados obtidos através dos modelos de Redes Neurais e de Função de Transferência, ambos ajustados para as multivariáveis representativas do processo;
- **Capítulo 6:** Neste capítulo, é desenvolvido e apresentado o estudo referente a aplicação das teorias citadas nos capítulos anteriores, na avaliação do controle de filtração da polpa de cristais de hidrato no processo de fabricação de alumina;
- **Capítulo 7:** Neste capítulo, são relatadas as considerações finais e recomendações para futuros trabalhos.
- **Anexos:** Neste, são apresentados as tabelas que contém todos os algoritmos utilizados para a construção das Redes Neurais Artificiais e para a construção do Modelo Multivariado com auxílio de Função de Transferência, e o banco de dados utilizado.

Capítulo 2

Modelos Probabilísticos de Séries Temporais em Processos Dinâmicos Univariados

2.1 Introdução

A classe dos Modelos Univariados limita-se a estudar uma única série histórica cujas características de interesse são explicadas exclusivamente a partir do seu próprio comportamento de evolução, ou seja, os valores futuros de uma série são explicados a partir de seus valores passados. Estes modelos supõem que a série é gerada através de um filtro linear, cuja entrada é ruído branco (Moretin e Toloi, 2004), apresentado por Figura 2.1,

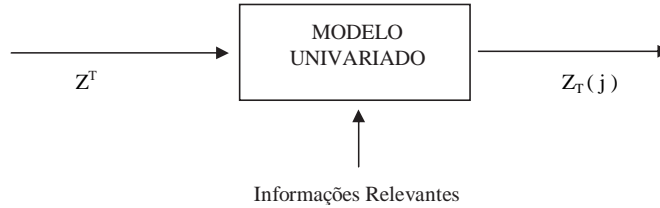


Figura 2.1 *Estrutura de Modelos Univariados*

Onde a composição da estrutura do modelo, é definida da seguinte forma:

- $Z^t \rightarrow$ entrada com característica de ruído branco. Portanto, $Z^t = a_t$
- Modelo Univariado \rightarrow Processo de filtro linear univariado, também considerado como uma função de transferência $\psi(B)$,
- $Z_t(j) \rightarrow$ Série de saída que descreve um processo linear (discreto), onde pode-se escrever:

$$Z_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots = \mu + \psi(B) a_t, \quad (2.1)$$

em que

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots, \quad (2.2)$$

- B é o operador deslocamento unitário discreto;
- e μ será um parâmetro de determinação do nível da série, caso a sequência de pesos $\{\psi_j, j \geq 1\}$ for infinita e convergente, onde o filtro será estável (somável) e Z_t é estacionária. Sendo assim,

$$E(a_t) = 0 \quad \text{para todo } t;$$

$$Var(a_t) = \sigma_a^2, \quad \text{para todo } t;$$

$$E(a_t a_s) = 0 \quad s \neq t.$$

Considerando, $\tilde{Z} = Z_t - \mu$, temos que:

$$\tilde{Z} = \psi(B)a_t \quad (2.3)$$

Dentro desta classe de modelos pode-se citar:

- Os métodos de decomposição e os métodos de amortecimento exponencial;
- Os modelos automáticos que incorporam modelos de regressão e o modelos de ajustamento sazonal;
- E os modelos univariados de Box & Jenkins. Esta classe de modelos foi implementada por Box & Jenkins (1970), e consiste de uma classe geral de modelos dinâmicos lineares conhecidos como Autoregressivo Integrado Médias Móveis ou modelos **ARIMA**.

2.2 Processos Estocásticos

Definição 2.1: Seja T um conjunto arbitrário. Um processo estocástico é uma família $Z = \{Z(t), t \in T\}$, tal que, para cada $t \in T$, Z_t é uma variável aleatória (v.a).

2.2.1 Características de um Processos Estocásticos

Sejam t_1, t_2, \dots, t_n elementos quaisquer de T e consideremos,

$$F(z_1, \dots, z_n; t_1, \dots, t_n) = P\{Z(t_1) \leq x_1, \dots, Z(t_n) \leq z_n\} \quad (2.4)$$

O processo estocástico $X = \{Z(t), t \in T\}$ estará especificado se conhecermos as distribuições finito-dimensionais em 2.4, para todo $n \geq 1$ onde:

- Para $n=1$, se conhece as distribuições unidimensionais da v.a $Z(t_1), t \in T$;
- Para $n=2$, se conhece as distribuições unidimensionais da v.a $(Z(t_1), Z(t_2)), t_1, t_2 \in T$, e assim por diante.

Conforme Moretin e Toloi, (2004) algumas condições devem serem satisfeitas pela função de distribuição (*fd*) descrita em 2.4, são elas :

1. **Condição de Simetria:** para qualquer permutação j_i, \dots, j_n , dos índices $1, 2, \dots, n$, temos:

$$F(x_{j_1}, \dots, x_{j_n}; t_{j_1}, \dots, t_{j_n}) = F(z_1, \dots, z_n; t_1, \dots, t_n) \quad (2.5)$$

2. **Condição de Compatibilidade:** para $m < n$,

$$F(x_1, \dots, x_m, \dots, +\infty; t_1, \dots, t_m, t_{m+1}, \dots, t_n) = F(z_1, \dots, z_m; t_1, \dots, t_m) \quad (2.6)$$

observe que o membro esquerdo da Equação (2.6), deve ser entendido como:

$$\lim F(x_1, \dots, x_m; x_{m+1}, \dots, x_n; t_1, \dots, t_n), \quad (2.7)$$

para $x_{m+1} \rightarrow +\infty, \dots, x_n \rightarrow +\infty$.

Um processo estocástico $X_t; t \in T$ é dito ser estacionário de segunda ordem ou fracamente estacionário se a sua função média é constante e sua função de autocovariância depende apenas da defasagem. Por definição temos:

Definição 2.2: Seja T um conjunto arbitrário. Um processo estocástico é uma família $X = \{X(t), t \in T\}$, tal que, para cada $t \in T$, $X(t)$ é uma variável aleatória. Deste modo, um processo estocástico é uma família de variáveis aleatórias (v.a). Por suposição, estão definidas em um espaço de probabilidade, do tipo (Ω, A, P) , e onde o conjunto T é normalmente tomado como o conjunto dos inteiros $T = 0, \pm 1, \pm 2, \dots$, no caso discreto, ou o conjunto dos reais R , no caso contínuo. Observe que, para todo $t \in T$, $Z(t)$ é uma variável aleatória definida sobre o espaço Ω , e na realidade $Z(t)$ é uma função de dois argumentos, $Z(t, \omega)$, $t \in T$ e $\omega \in \Omega$.

Alguns estudos podem ser restritos a momentos de baixa ordem e, em particular, para uma classe de processos os quais interessarão a este trabalho, os chamados processos estacionários de segunda ordem, onde será considerado os momentos de primeira e segunda ordem, onde: A função média, ou simplesmente média, de \mathbf{X} é:

$$\mu(1, t) = \mu(t) = E[X(t)] = \int_{-\infty}^{\infty} zf(z, t)dz \quad (2.8)$$

E a função de autocovariância (facv) de \mathbf{X} é:

$$\mu(1, 1; t_1 t_2) - \mu(1; t_1)(1; t_2) = \gamma(t_1, t_2) = \gamma(\tau) \quad (2.9)$$

$$\gamma(t_1, t_2) = E\{X(t_1)X(t_2)\} - E\{X(t_1)\}E\{X(t_2)\}, \quad t_1, t_2, \in T \quad (2.10)$$

Observe que $\mu(t)$ é uma função de $t \in T$ e que $\gamma(t_1, t_2)$ depende de dois argumentos, t_1 e t_2 e em particular $t_1 = t_2 = t$, a função (2.9) nos dará:

$$\gamma(t, t) = Var\{X(t)\} = E\{X^2(t)\} - E^2\{X(t)\} \quad (2.11)$$

especificada como a função variância do processo \mathbf{X} e será indicada por $V(t)$. Então:

$$E[X(t)] = \mu(t) \quad (2.12)$$

$$Cov[X(t_1); X(t_2)] = \gamma(t_1, t_2) \quad (2.13)$$

Nenhuma outra suposição é feita a respeito dos momentos de ordem mais alta. Além disso, fazendo $\tau = 0$ segue que $Var[X_t] = 0$, ou seja, a variância do processo assim como a média também é constante. Note também que tanto a média quanto a variância precisam ser finitos. Esta definição mais fraca de estacionariedade será utilizada daqui em diante já que muitas propriedades dos processos estacionários dependem apenas da estrutura especificada pelo primeiro e segundo momentos. Uma classe importante de processos em que isto se verifica é a classe dos processos normais ou Gaussianos cuja a distribuição conjunta de $X(t_1), \dots, (t_k)$ é normal multivariada para todo conjunto t_1, \dots, t_k . A distribuição normal multivariada fica completamente caracterizada pelo primeiro e segundo momentos, i.e. por $\mu(t)$ e $\gamma(t_1; t_2)$. Assim, estacionariedade fraca implica em estacionariedade estrita para processos normais. Por outro lado, μ e $\gamma(\tau)$ podem não descrever adequadamente processos que sejam muito "não-normais" (Wei, 1990).

2.3 A estacionariedade

A estacionariedade numa série temporal significa que os dados oscilam sobre uma média constante, independente do tempo, com a variância das flutuações permanecendo essencialmente a mesma. Diniz (1998), afirma que uma série temporal é estacionária se o processo aleatório oscilar em torno de um nível médio constante. Séries temporais com tendência linear ou exponencial são exemplos de séries temporais com comportamento não estacionário. A condição de estacionariedade de 2ª ordem implica que tanto a média do processo, como a variância são constantes.

2.4 Processos Estacionários

Definição 2.3: Um processo estocástico $X = \{Z(t), t \in T\}$ é dito estritamente estacionário se todas as distribuições finito-dimensionais:

$$F(z_1, \dots, z_n; t_1, \dots, t_n) = P[Z(t_1) \leq z_1, \dots, Z(t_n) \leq z_n]. \quad (2.14)$$

permanecem as mesmas sob translações no tempo, (Moretin e Toloi, 2004), ou seja:

$$F(x_1, \dots, x_n; t_1 + \tau, \dots, t_n + \tau) = F(x_1, \dots, x_n; t_1, \dots, t_n), \quad (2.15)$$

para quaisquer t_1, \dots, t_n, τ de T .

Uma série temporal é dita estritamente estacionária se a distribuição de probabilidade conjunta de $X(t_1), \dots, X(t_k)$ é a mesma de $X(t_1 + \tau), \dots, X(t_k + \tau)$. Ou seja, o deslocamento da origem dos tempos por uma quantidade τ não tem efeito na distribuição conjunta que portanto depende apenas dos intervalos entre t_1, \dots, t_k . Em particular, para $k = 1$ a estacionariedade estrita implica que a distribuição de X_t é a mesma para todo t de modo que, se os dois primeiros momentos forem finitos, temos que:

$$\mu(t) = \mu \quad (2.16)$$

$$v(t) = \sigma^2 \quad (2.17)$$

são constantes e não dependem de t . Para $k = 2$ a distribuição conjunta de $X(t_1)$ e $X(t_2)$ depende apenas da distância $t_2 - t_1$, chamada defasagem. A função de autocovariância $\gamma(t_1; t_2)$ também depende apenas de $t_2 - t_1$ e pode ser escrita como $\gamma(\tau)$ onde, o coeficiente de autocovariância na defasagem τ é dada por:

$$\gamma(\tau) = E[X(t) - \mu][X(t + \tau) - \mu] \quad (2.18)$$

$$\gamma(\tau) = Cov[X(t), X(t + \tau)] \quad (2.19)$$

Observe, que para melhor interpretação, pode-se denominar a Equação (2.19) de autocovariância, e assim, dando origem a função de autocorrelação $\rho(\tau)$:

$$\rho(\tau) = \frac{\gamma(\tau)}{\gamma(0)} \quad (2.20)$$

Essa função explica o nível de correlação existente entre $X(t)$ e $X(t + \tau)$.

2.5 Processos Não-Estacionários

2.5.1 Modelos *ARIMA* (p, d, q) de Box & Jenkins

Estes modelos foram propostos no decorrer da década de 1960, pelos professores George E. P. Box e Gwilym M. Jenkins e a partir de 1970 foi publicado na literatura ***Time Series Analysis, forecasting and control*** (MORETTIN e TOLOI, 2004). Esta é a metodologia mais utilizada pela comunidade científica para a análise de séries temporais e formulação de modelos com maior capacidade de gerar previsões significativas.

Segundo *Box e Jenkins (1976)*; *Souza e Camargo (1996)* uma série temporal é um conjunto de observações de uma dada variável ordenada segundo o parâmetro tempo em intervalos iguais. Se o processo estocástico que gerou a série de observações é invariante com respeito ao tempo, diz-se que o processo é estacionário. Se as características do processo se alteram no decorrer do tempo, o mesmo é caracterizado como não estacionário.

A importância do conhecimento da série, ser ou não estacionária, reside no fato de que quando se trabalha com uma série estacionária, se está em presença de uma função amostral do processo que tem a mesma forma em todos os instantes do tempo $t \in N$, onde N é o número de observações existentes, acarretando na facilidade de obtenção de estimativas do processo.

Segundo Morettin e Tolo (2004), uma série temporal $W_t = \Delta^d Z_t$ apresenta as características de estacionariedade, este fato, se deve ao operador diferença Δ^d . Então podemos escrever W_t por um modelo **ARMA**(p, q), da forma:

$$\phi(B)W_t = \theta(B)a_t. \quad (2.21)$$

onde, $\phi(B) = 1 - \phi_1(B) - \dots$ e $\theta(B) = 1 - \theta_1(B) - \dots$

sendo:

B : Operador Atraso, tal que, $B^K X_t = X_{t-k}$

$\theta(B)$: Polinômio médias móveis (MA)

$\phi(B)$: Polinômio auto regressivo (AR)

Δ : Operador diferença simples, tal que, $\Delta Z_t = Z_t - Z_{t-1}$

No entanto, se W_t for uma diferença de Z_t , conseqüentemente Z_t é uma integral de W_t , daí diz-se que Z_t apresenta característica de um modelo auto-regressivo, integrado, médias móveis ou simplesmente modelo **ARIMA**. Os termos auto-regressivos correspondem a defasagens da série transformada, ou seja, a série estacionária obtida por intermédio de diferenciações e as médias móveis a defasagens dos erros aleatórios e o termo integrado refere-se a ordem de diferenciações sofridas pela série para tornar-se estacionária.

Esta metodologia tem como premissa fundamental que as séries temporais são geradas por processos estocásticos cuja natureza pode ser representada por um modelo matemático, com a seguinte estrutura.

$$\phi(B)\Delta^d Z_t = \theta(B)a_t. \quad (2.22)$$

de ordem (p,d,q) e escrevemos **ARIMA**(p, d, q), onde p e q são as ordens do parâmetro autoregressivo $\phi(B)$ e do parâmetro média móvel $\theta(B)$, respectivamente.

Portanto, Segundo *Morettin & Tolo* (2004) o modelo especificado em (2.22), supõe que a **d-ésima** diferença da série Z_t pode ser representado por um modelo **ARMA**, estacionário e inversível. Na maioria dos casos usuais, $d = 1$ ou $d = 2$, que correspondem a dois casos interessantes e comuns de não estacionariedade homogênea, são eles:

- a) Séries não-estacionárias quanto ao nível: oscilam ao redor de um nível médio durante algum tempo e depois saltam para outro nível temporário. Para torná-las estacionárias é suficiente tomar uma única diferença; este é o caso típico de séries econômicas.

$$\phi(B)\Delta^1 Z_t = \theta(B)a_t.$$

- b) Séries não-estacionárias quanto a inclinação: oscilam numa direção por algum tempo e depois mudam para outra direção temporária. Para torná-las estacionárias é suficiente tomar a segunda diferença.

$$\phi(B)\Delta^2 Z_t = \theta(B)a_t.$$

O modelo **ARIMA**(p, d, q) é um caso especial de um processo integrado. Em geral, diz-se que X_t é integrável de ordem d se $\Delta^d X_t$ for estacionário, e escrevemos $X_t \sim I(d)$. Se $d=0$, conseqüentemente X_t será considerado estacionário.

A análise no domínio do tempo é baseada em um modelo paramétrico, utilizando-se as funções de autocovariância e autocorrelação. A autocorrelação é a autocovariância padronizada, que serve para medir a extensão de um processo para o qual o valor tomado no tempo t , depende daquele tomado no tempo $t-k$. Define-se a autocorrelação de ordem k como:

$$\rho_{kk} = \frac{\gamma_k}{\gamma_0} = \frac{Cov[Z_t, Z_{t+k}]}{\sqrt{Var(Z_t)Var(Z_{t+k})}} \quad (2.23)$$

Onde: $Var(Z_t) = Var(Z_{t+k}) = 0$ o que é igual a variância do processo $\rho_0 = 1$ e $\rho_k = \rho_{-k}$

A autocorrelação pode ser estendida. Se for medida a correlação entre duas observações seriais Z_t e Z_{t+k} , eliminando-se a dependência dos termos intermediários, Z_{t+1}, Z_{t+k-1} , tem-se a autocorrelação parcial, representada por:

$$Cor \frac{Z_t, Z_{t+k}}{Z_{t+1}, \dots, Z_{t+k-1}} \quad (2.24)$$

A função de autocorrelação pode ser obtida considerando-se um modelo de regressão para um processo estacionário com média zero. A variável dependente Z_{t+k} dependente das variáveis; $Z_{t+k-1}, Z_{t+k-2}, \dots, Z_{t+k-n}$. Deste modo; temos:

$$Z_{t-k} = \phi_{k1}Z_{t+k-1} + \phi_{k2}Z_{t+k-2} + \dots + \phi_{KK}Z_t + a_{t-k} \quad (2.25)$$

Onde: ϕ_{ki} - i - ésimo parâmetro da regressão

a_{t+k} - é o termo de erro descorrelatado com Z_{t+k-j} para $j \geq 1$

De acordo com Moretin e Toloi (2004), os *Modelos Autoregressivos Puros* $AR(p, \theta)$ são aqueles cujo polinômio $\phi(B) = 1$, e onde as estimativas de Z para um tempo t dependem de uma combinação linear de p termos da série observada, incluindo o termo aleatório $a(t)$ de ruído branco (erros de estimação com distribuição normal de média zero e variância constante e não correlacionadas). Deste modo tem-se:

$$Z_t = \sum_{i=1}^p \phi_i Z(t-i) + a_t \quad (2.26)$$

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t \quad (2.27)$$

onde ϕ_i são parâmetros que ponderam os valores de Z_t , do instante imediatamente anterior $t-1$ até o mais distante $t-p$, sendo determinados através de técnicas de minimização de erros.

Os *Modelos Médias Móveis puros* $MA(\theta, q)$ são aqueles cujo polinômio $\theta(B) = 1$, este modelo assume que a série modelada é gerada através de uma combinação linear de q sinais de ruído $a(t-i)$, aleatórios e independentes entre si. Assim temos:

$$Z_t = - \sum_{i=1}^q \theta_i a(t-i) + a_t \quad (2.28)$$

$$Z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_p a_{t-p} \quad (2.29)$$

Os *Modelos Autoregressivos-Médias-Móveis Puros* $ARMA$, com $AR(p, \theta)$ e $MA(\theta, q)$ são aqueles cujo polinômio $\phi(B) = 1$ e $\theta(B) = 1$. Deste modo temos:

$$Z_t = - \sum_{i=1}^p \phi_i Z(t-i) - \sum_{i=1}^q \theta_i a(t-i) + a_t \quad (2.30)$$

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_p a_{t-p} \quad (2.31)$$

Para que o polinômio $\phi(B)$ seja estacionário, suas raízes têm de estar fora do círculo unitário, e para que $\theta(B)$ seja inversível, suas raízes devem se encontrar fora do círculo unitário.

Para se fazer uma previsão de uma série temporal através dos modelos *ARIMA*, torna-se necessário identificar a ordem dos parâmetros p, d, q . Sendo assim, o primeiro parâmetro a ser identificado é o grau de diferenciação d necessário à estabilização dos dados. Isto é feito através de um exame de correlograma, ou seja, do diagrama da função de autocorrelação (*FAC*), no qual são apresentados os valores das autocorrelações em relação aos *lags* k .

Se as autocorrelações decrescem de forma linear, realizam-se diferenciação na série, até que o diagrama apresente um corte abrupto para um valor qualquer de autocorrelação, a partir daí a série será considerada estacionária. A ordem autoregressiva p é determinada pela verificação da função de autocorrelação parcial (*FACP*) ϕ_{kk} da série que está sendo analisada. Se a série for unicamente autoregressiva *ARIMA* $(p, d, 0)$, sua função de autocorrelação parcial apresentará uma queda rápida após o *lag* k .

Segundo (*Bartlett, 1946; Box & Jenkins, 1976*) caso não utilize a *FACP*, deve-se efetuar uma análise dos estimadores ϕ_{kk} para verificar até que ordem de defasagem do correlograma desta função ele é estatisticamente significativo. Após indentificá-la ela será considerada autoregressiva.

2.5.2 Análises das Funções de Autocorrelações

Os processos $AR(p)$, $MA(q)$ e $ARMA(p, q)$ apresentam as funções de autocorrelações com as seguintes características:

- Um processo $AR(p)$ apresenta *fac* com decaimento de acordo com exponenciais e/ou senóides amortecidas, infinitas em extensão;
- Um processo $MA(q)$ apresenta *fac* finita, no sentido que ela apresenta um corte após o *lag* q ;
- Um processo $ARMA(p, q)$ apresenta *fac* infinita em extensão, a qual decai de acordo com exponenciais e/ou senóides amortecidas após o *lag* $q - p$.

2.5.3 Análise das Funções de Autocorrelações Parciais

Os processos AR(p), MA(q) e ARMA(p,q) apresentam as funções de autocorrelações parciais com as seguintes características:

- Um processo AR(p) apresenta facp $\phi_{kk} \neq 0$, para $k \leq p$ e $\phi_{kk} = 0$ para $k > p$;
- Um processo MA(q) apresenta facp de maneira similar à fac de um processo AR(p): com exponenciais e/ou senóides amortecidas;
- Um processo ARMA(p,q) apresenta facp com as características da fac de um processo MA puro e/ou senóides amortecidas após o lag $q - p$.

2.5.4 Critérios de Seleção - AIC e SBC

Existem vários critérios que permitem identificar modelos ARMA, sugeridos a partir de 1970, cuja idéia era escolher as ordens p e q que fossem capazes de minimizar a expressão:

$$P(p, q) = \ln \hat{\sigma}_{p,q}^2 + (p, q) \frac{C(N)}{N} \quad (2.32)$$

em que $\hat{\sigma}_{p,q}^2$ é uma estimativa da variância residual obtida ajustando um modelo ARMA(p,q) às N observações da série e C(N) é uma função do tamanho da série.

- **Critério de Informação de Akaike - AIC.**

Akaike (1973 e 1974) afirma que a escolha de modelos mais significativos, implica na escolha de modelos, cujas as estimativas para os parâmetros (p,q), venham minimizar a expressão.

$$AIC(p, d, q) = \ln \hat{\sigma}_a^2 + \frac{N}{N-d} \cdot 2(p + q + 1 + \delta_{d0}) + N \cdot \ln 2\pi + N \quad (2.33)$$

em que:

$$\delta_{d0} = \begin{cases} 1, & d = 0 \\ 0, & d \neq 0 \end{cases} \quad (2.34)$$

e $\hat{\sigma}_a^2$ é o estimador de máxima verossimilhança de σ_a^2 .

Uma forma mais simples de se calcular este critério é:

$$AIC = N \cdot \ln\left(\frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2\right) + 2 \cdot \omega.$$

- **Critério de Informação Bayesiano - BIC ou SBC**

Sugerido por Akaike(1977), Rissanen (1978) e Schwartz (1978), este critério, para o caso de um modelo ARMA(p,q), implica em minimizar a expressão:

$$SBC(p, q) = \ln \hat{\sigma}_{p,q}^2 + (k + l) \frac{\log N}{N} \quad (2.35)$$

onde, $\hat{\sigma}_{p,q}^2$ é o estimativa de máxima verossimilhança da variância residual do modelo ARMA(p,q). Uma forma mais simples de se calcular este critério é:

$$SBC = N \cdot \ln \frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2 + \omega \cdot \ln(N)$$

Onde, ω = número de parâmetros estimados; N = número de observações utilizadas.

Capítulo 3

Modelos Probabilísticos de Séries Temporais em Processos Dinâmicos Multivariados

3.1 Introdução

Para Souza(1989), os Modelos Multivariados são modelos capazes de realizar várias previsões ao mesmo tempo, onde as séries de interesse são explicadas em função não apenas do comportamento da própria série mas também de outras séries, permitindo, desta forma, relações de interdependência e causalidade. Esses modelos apresentam uma estrutura capaz de estimar a curto, médio e longo prazo diversas séries simultaneamente.

Na prática, muitas das séries temporais são avaliadas melhor como uma componente de algum vetor multivariado de uma série de tempo X_t com dependência existente dentro de cada componennte X_{tj} , mas também deve-se avaliar a interdependência entre as diferentes componentes X_{tj} e X_{ti} . Deve-se observar que muitas das teorias e suposições levantadas na análise univariada de série de tempo estende-se de forma similar para a análise de séries de tempo multivariada (Wylomanska, 2005).

3.2 Propriedades de Segunda Ordem de Séries Temporais Multivariada

Considere m séries de tempo $X_{ti}, t = 0, \pm 1, \pm 2, \dots, i = 1, 2, \dots, m$ com $E(X_{ti}^2) < \infty$ para todo i . Se todas as distribuições finitas dimensionais das variáveis aleatórias X_{ti} forem do tipo normal multivariada, as distribuições de X_{ti} serão completamente determinada pelas médias.

$$\mu_{ti} = E(X_{ti}) \quad (3.1)$$

e as covariâncias por:

$$\gamma_{ij}(t+h, t) = E[(X_{t+h, i} - \mu_{ti})(X_{tj} - \mu_{tj})] \quad (3.2)$$

Caso as observações X_t não apresentem uma distribuição normal, as quantidades μ_{ti} e $\mu_{ij}(t+h, t)$ especificam a propriedade de segunda ordem, e a covariância é uma medida de dependência, não somente entre as observações da mesma série, mas também em observações de séries diferentes.

Quando se trata de m séries, costuma-se utilizar a notação de vetor, definido da seguinte forma:

$$X_t = [X_{t1}, \dots, X_{tm}]', \quad t = 0, \pm 1, \pm 2, \dots \quad (3.3)$$

As propriedades de segunda ordem das séries de tempo multivariada X_t são especificadas pelos vetores de médias, na forma:

$$\mu_t = [\mu_{t1}, \dots, \mu_{tm}]' \quad (3.4)$$

Definição 1 : Uma série m -variada X_t é dita fracamente estacionária se μ_t for independente de t e $\Gamma(t+h, t)$ for independente de t , para todo h . Para séries estacionárias temos:

$$\mu = E(X_t \Gamma(h)) = E[(X_{t+h} - \mu)(X_t - \mu)]. \quad (3.5)$$

Definição 2 : A série temporal m -variada Z_t é definida como (**White Noise**) ou ruído branco com média 0 e matriz de covariância $\sum(Z_t \sim WN(0, \Sigma))$ se Z_t é estacionária com média d vetor igual a zero e a função de matriz de covariância dada por:

$$\Gamma(h) = \begin{cases} \Sigma, & h = 0 \\ 0, & \text{caso contrário} \end{cases} \quad (3.6)$$

Definição 3 : Uma série m -variada Z_t é dita série de ruído branco Independente e Identicamente Distribuído - IID com média 0 e matriz de covariância $\sum(Z_t \sim IID(0, \Sigma))$ se o vetor aleatório Z_t for independente e identicamente distribuído com média 0 e matriz de covariância σ^2 .

Definição 4 : Uma série m -variada X_t é um processo linear, se apresentar a seguinte representação.

$$X_t = \sum_{j=-\infty}^{\infty} C_j Z_{t-j}, \quad Z_t \sim WN(0, \Sigma) \quad (3.7)$$

Onde C_j é uma sequência de matrizes $m \times m$ com componentes absolutamente somável.

Observação 1: Um processo da forma descrita em (3.7) é estacionário com média 0 e função de covariância

$$\Gamma(h) = \sum_{j=-\infty}^{\infty} C_{j+h} \Sigma C_j', \quad h = 0, \pm 1, \pm 2, \dots \quad (3.8)$$

3.3 Modelamento com Processo *AR* Multivariado

Segundo Wilomanska(2005), se X_t é um processo $AR(p)$ multivariado causal definido pela equação diferença:

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + Z_t, \quad WN(0, \Sigma), \quad (3.9)$$

ou ainda:

$$X_t = \mu + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \epsilon_t \quad (3.10)$$

$$\phi(L)X_t = \mu + \epsilon_t \quad (3.11)$$

onde, $\epsilon_t \sim WN(0, \sigma_\epsilon^2)$ (Ruído Branco), isto é:

$$E(\epsilon_t) = 0,$$

$$E(\epsilon_t \epsilon_s) = \begin{cases} \sigma_\epsilon^2, & t = s \\ 0, & c.c \end{cases} \quad (3.12)$$

com, $1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$ Lag polinomial de ordem p , então:

$$L^k X_t = y_{t-k} \quad (3.13)$$

com operador Lag ($L^0 y_t = y_t$).

Com a condição de estacionariedade satisfeita, tem-se:

$$\phi(L) = 0, \quad (3.14)$$

Com as raízes fora do círculo unitário, desta forma;

$$z^p - \phi_1 z^{p-1} - \dots - \phi_{p-1} z - \phi_p = 0 \quad (3.15)$$

3.4 Modelamento com Processo *MA* Multivariado

Segundo Wilomanska(2005), se X_t é um processo *MA*(q) multivariado é definido:

$$X_t = \mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q} = \mu + \theta(L) \cdot \epsilon_t, \quad (3.16)$$

com, $\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q$ Lag polinomial de ordem q, com:

$$\epsilon_t \sim WN(0, \sigma_\epsilon^2), \quad (3.17)$$

Um processo *MA* estacionário e com as condições de invertibilidade satisfeita, apresenta a característica polinomial $\theta(L) = 0$, com raízes fora do círculo unitário.

3.5 Processos *ARMA* Multivariados

Como definido no processo univariado, pode-se definir uma classe útil de processos estacionários multivariados X_t . Desta forma, X_t deve satisfazer uma classe de equações de diferenças lineares com coeficientes constantes. Z_t é uma série multivariada de ruído branco, fundamental para a construção de processos *ARMA*.

Definição 5 : As séries são caracterizadas como um processo *ARMA*(p, q) se X_t é estacionária e para cada t , ocorrer.

$$\phi(B)X_t = \theta(B)Z_t, \quad Z_t \sim WM(0, \Sigma), \quad (3.18)$$

onde $\phi(z) = I - \phi_1 z - \dots - \phi_p z^p$, $\theta_z = I + \theta_1 z + \dots + \theta_q z^q$ são matrizes com valores polinomiais, I é uma matriz identidade $\mathbf{m} \times \mathbf{m}$.

Definição 6 : Um processo *ARMA*(p, q) para uma série X_t é causal ou uma função causal de Z_t , se existirem matrizes ψ_j com componentes absolutamente somáveis, de tal modo que:

$$\phi(L)Y_t = \mu + \theta(L)\epsilon_t, \quad (3.19)$$

ou

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j} \quad (3.20)$$

que é equivalente à condição $\det\phi(z) \neq 0$, para todo $z \in C$, de tal modo que, $|z| \leq 1$. A matriz ψ_j é encontrada através da equação.

$$\psi_j = \theta_j + \sum_{k=1}^{\infty} \psi_{j-k} \phi_k \quad (3.21)$$

$$\phi_j = 0 \text{ para } j > p, \quad \psi_j = 0 \text{ para } j < 0. \quad (3.22)$$

onde $\theta_0 = I$, $\theta_j = 0$ para $j > q$, $\phi_j = 0$ para $j > p$ e $\psi_j = 0$ para $j < 0$.

Definição 7 : Um processo *ARMA*(p, q) para uma série X_t é invertível se existirem matrizes Π_j com componentes absolutamente somáveis, de tal modo que:

$$\Pi_j = -\phi_j - \sum_{k=1}^{\infty} \theta_k \Pi_{j-k}, \quad j = 0, 1, \dots, \sum_{j=0}^{\infty} \psi_j Z_{t-j} \quad (3.23)$$

onde $\phi_0 = -I$, $\phi_j = 0$ para $j > p$, $\theta_j = 0$ e $\Pi_j = 0$ para $j > 0$.

Para um processo AR(1) multivariado com o auxílio da equação (3.23) dada na definição de causalidade, tem-se:

$$\begin{aligned}\psi_0 &= I \\ \psi_1 &= \phi\psi_0 = \phi \\ \psi_2 &= \phi\psi_1 = \phi^2 \\ &\vdots \\ \psi_j &= \phi\psi_{j-1} = \phi^j, \quad j \geq 3\end{aligned}$$

3.6 Processos *ARIMA* Multivariados

Definição 8 : Uma série é definida integrada de ordem (d), ou seja, $X_t \sim I(d)$, se tornando estacionário após d diferenciação. Portanto o processo *ARIMA*(p,d,q) para uma série X_t é definido, da seguinte forma:

$$(1 - L)^d X_t \sim ARMA(p, q), \quad (3.24)$$

Diz-se que $X_t \sim ARIMA(p, d, q)$, onde \mathbf{p} , denota a ordem dos *AR-lags* e \mathbf{q} , denota a ordem dos *MA-lags* e \mathbf{d} as diferenciações.

3.7 Modelos de Função de Transferência (MFT) ou Modelos Causais

Nos Modelos de Função de Transferência uma série de saída é relacionada a uma ou mais séries da entrada, ou seja, o interesse agora é estudar uma única saída em um sistema linear de entrada múltipla. Nestes tipos de modelos os valores futuros de uma série são explicados não somente pelos valores passados da mesma, mas também por séries que de alguma forma possuam relação entre si, (Fig. 3.1).

onde:

$Z_{1,T}, \dots, Z_{k-1,T}$ são séries de entrada;

$Z_{K,T}$ séries de saída, pós modelo de função de transferência; Por exemplo, vendas podem estar relacionadas a gastos com propaganda; consumo diário de eletricidade podem estar relacionados a certas variáveis ambientais tais como máxima temperatura externa

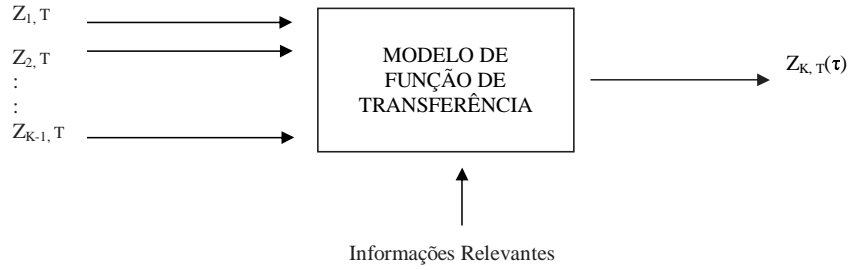


Figura 3.1 Estrutura de Modelos de Função de Transferência

e unidade relativa. Serão vistos modelos lineares com saída simples com entrada simples e múltipla. Depois do estudo das características básicas dos modelos de função de transferência serão discutidas a identificação, estimação e checagem do diagnóstico desses modelos e ainda previsões para modelos com função de transferência.

3.8 Funções de Transferências de Única Entrada

Assumindo que X_t e Y_t são séries corretamente transformadas, de modo que, ambas são estacionárias. Em um sistema linear com entrada e saída única, a série de entrada X_t e de saída Y_t são relacionadas através do seguinte filtro linear:

$$Y_t = v(B)X_t + N_t \quad (3.25)$$

Onde:

$v(B) = \sum_{j=-\infty}^{\infty} v_j B^j$ se refere a um filtro de função de transferência descrito por Box e Jenkins (1976), e N_t é uma série de ruído do sistema que é independente da série de entrada X_t .

Box e Jenkins descreveu a função (3.25), como um modelo de função de transferência que apresenta a capacidade de descrever funções resposta de frequência. Quando X_t e Y_t seguem um modelo **ARMA**, esta Função (3.25) também é conhecida como um modelo, **AutoRegressive Moving Average with exogenous signal ARMAX**, ou seja, **Autoregressivo com Média Móvel e Entradas Exogenas**.

Os coeficientes do modelos de função de transferência em (3.25) são geralmente conhecidos como pesos de resposta do impulso. Em função de j , v_j também é conhecida como função de resposta de impulsos. O modelo de função de transferência é dito ser estável se a sequência desses pesos puder assumir a seguinte função:

$$\sum |v_j| < \infty \quad (3.26)$$

Para Wei(1990), um sistema estável com uma entrada limitada sempre produz uma saída limitada. O modelo de função de transferência é dito ser causal se $v_j = 0$ para $j < 0$. Assim, em um modelo causal, o sistema não funciona para a série de entrada até que as mesmas tenham sido realmente aplicadas no sistema.

Um modelo casual e estável é geralmente descrito como sendo da seguinte forma racional:

$$Y_t = v_0 X_t + v_1 X_{t-1} + v_2 X_{t-2} + \dots + N_t \quad (3.27)$$

$$Y_t = v(B)X_t + N_t \quad (3.28)$$

onde; $v(B) = \sum_{j=0}^{\infty} v_j B^j$, $\sum_{j=0}^{\infty} |v_j| < \infty$, onde X_t e N_t são independentes.

Segundo Wei(1990), um sistema de função de transferência dinâmico pode ser escrito esquematicamente da seguinte forma:

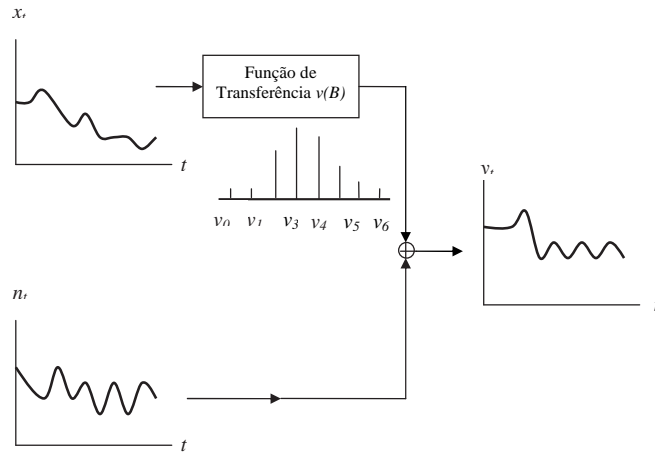


Figura 3.2 Sistema de Função de Transferência Dinâmico

O grande objetivo da modelagem com funções de transferência é o de identificar e estimar a função de transferência $v(B)$ e modelar o ruído N_t , baseado na informação disponível das séries de entrada X_t e da série de saída Y_t . A maior dificuldade encontrada ao se trabalhar os MFT é que a informação sobre X_t e de Y_t é finita e como a função de transferência $v(B)$ em (3.27) pode conter um número infinito de coeficientes, isto muitas vezes torna-se não trivial.

Para reduzir estas dificuldades pode-se representar a função de transferência $v(B)$, descrita na Figura 3.2, através da seguinte equação:

$$v(B) = \frac{w_s(B)}{\delta_r(B)} \cdot B^b \quad (3.29)$$

onde; $w_s(B) = w_0 - w_1(B) - \dots - w_s B^s$, $\delta_r(B) = 1 - \delta_1(B) - \dots - \delta_r B^r$ e b é um parâmetro de defasagem que representa o **lag** do tempo presente que decorre antes que o impulso da variável de entrada produza um efeito sobre a variável de saída. Para um sistema estável, é assumido que as raízes $\delta_r(B)$ encontram-se dentro do círculo unitário.

Uma vez que $W_s(B)$, $\Delta_r(B)$ e b são definidos, o peso v_j da resposta do impulso pode ser obtido igualando-se os coeficientes de B^j em ambos os lados da seguinte equação:

$$\delta_r(B)v(B) = w_s(B) \cdot B^b \quad (3.30)$$

então tem-se:

$$[1 - \delta_1(B) - \dots - \delta_r(B)^r][v_0 + v_1(B) + v_2(B)^2 + \dots] = [w_0 - w_1(B) - \dots - w_s B^s]B^b \quad (3.31)$$

3.9 Algumas Funções Típicas da Resposta do Impulso

Na prática, os valores de r e s na Equação (3.31), raramente excedem 2. Por este fato, será descrito o comportamento de três tipo de funções típicas de transferência:

tipo 1 : $r = 0$, neste caso, a função de transferência contém somente um número finito dos pesos de resposta do impulso que começam em $v_b = w_0$ e que terminam em $v_{b+s} = -w_s$

tipo 2 : $r = 1$, neste caso, os pesos da resposta de impulso exibem um decaimento exponencial que inicia em, v_b se $s = 0$, para um v_{b+1} se $s = 1$, e de v_{b+2} se $s = 2$. Portanto, só irá acontecer se as raízes estiverem dentro do círculo unitário.

tipo 3 : $r = 2$, neste caso, os pesos da resposta de impulso exibem uma onda exponencial ou um senóide amortecido (seno), dependendo da natureza das raízes do polinômio $\delta_2(B) = (1 - \delta_1(B) - \delta_2(B^2)) = 0$, este segue um decaimento exponencial, se as raízes forem do tipo reais, ou seja, se $\delta_1^2 - 4\delta_2 \geq 0$. Caso as raízes sejam complexas, ou seja, se apresentarem $\delta_1^2 - 4\delta_2 < 0$, então o valor de s é descrito similar ao descrito anteriormente. Se as raízes de $\delta_r(B)$ estiverem dentro do círculo unitário.

3.10 Funções de Correlação Cruzada e Modelos de Função de Transferência

3.10.1 Funções de Correlação Cruzada - FCC

Wei 1990, descreve que a função de correlação cruzada é utilizada para medir a relação existente entre duas variáveis aleatórias. Dado dois processos estocásticos X_t e Y_t para $t = 0, \pm 1, \pm 2, \dots$, diz-se que x_t e y_t são conjuntamente estacionárias se x_t e y_t são processos estacionários univariados com covariância cruzada entre X_t e Y_t , ou seja, $\text{Cov}(X_t, Y_t)$, é uma função somente da diferença de tempo ($x-t$). Neste caso, tem-se a seguinte covariância cruzada entre X_t e Y_t :

$$\gamma_{XY}(k) = E[(X_t - \mu_x)][(Y_{t+k} - \mu_y)] \quad (3.32)$$

para $k = 0, \pm 1, \pm 2, \dots$, onde padronizando a equação (3.32), tem-se a seguinte função de correlação cruzada:

$$\rho_{xy}(k) = \frac{\gamma_{xy}(k)}{\sigma_x \cdot \sigma_y} \quad (3.33)$$

para $k = 0, \pm 1, \pm 2, \dots$, onde σ_x e σ_y são os desvios padrões de X_t , Y_t respectivamente. É importante observar que a função de covariância cruzada $\gamma_{xy}(k)$ e a função de correlação cruzada $\rho_{xy}(k)$ são generalização das funções de auto-covariâncias e auto-correlações devido $\gamma_{xx}(k) = \gamma_x(k)$ e $\rho_{xx}(k) = \rho_x(k)$.

Observa-se que, ao contrário da função de autocorrelação $\rho_x(k)$, que é simétrica em torno da origem, isto é, $\rho_x(k) = \rho_x(-k)$ a função de correlação cruzada não é simétrica, ou seja, $\rho_{xy}(k) \neq \rho_{xy}(-k)$. Por outro lado, devido $E(x_t - \mu_x)(y_{t+k} - \mu_y) = E(y_{t+k} - \mu_y)(x_t - \mu_x) = \gamma_{xy}(-k)$, tem-se $\rho_{xy}(k) = \rho_{yx}(-k)$. Portanto, a função de correlação cruzada, FCC, mede não somente a intensidade de uma associação mas também sua direção. Para se ver a figura inteira da relação entre as séries X_t e Y_t , é importante examinar a FCC, $\rho_{xy}(k)$, para ambos os lags positivos e negativos, $k > 0$ e $k < 0$. O gráfico da FCC é conhecido como correlograma cruzado.

3.10.2 Relação entre Função de Correlação Cruzada e Função de Transferência

O Modelo de função de tranferência em (3.27), supondo um tempo $t+k$, pode ser escrito como:

$$Y_{t+k} = v_0 X_{t+k} + v_1 X_{t+k-1} + v_2 X_{t+k-2} + \dots + N_{t+k} \quad (3.34)$$

Sem perda de generalidades pode-se assumir que $\mu_x = 0$ e $\mu_y = 0$. Multiplicando ambos os termos em (3.32) por X_t , e calculando a esperança tem-se:

$$\gamma_{xy}(k) = v_0 \gamma_{xx}(k) + v_1 \gamma_{xx}(k-1) + v_2 \gamma_{xx}(k-2) + \dots \quad (3.35)$$

Lembre-se que ao adotar $\gamma_{xx}(k) = 0$, para todo K , tem-se:

$$\rho_{xy}(k) = \frac{\sigma_x}{\rho_y} [v_0 \rho_x(k) + v_1 \rho_x(k-1) + v_2 \rho_x(k-2) + \dots]. \quad (3.36)$$

A relação entre a função de correlação cruzada $\rho_{xy}(k)$ e a função resposta de impulso v_j é claramente influenciada pela autocorrelação existente nas séries de entrada X_t .

Mesmo se $r=0$ em (3.29) e a função de tranferência v_B conter um número finito de pesos de resposta do impulso, então o uso da equação (3.36) na construção de um sistema de equações para resolver v_j em função de $\rho_{xy}(k)$ e $\rho_x(k)$ é difícil.

A variância e a covariância das estimativas da amostra de $\rho_{xy}(k)$ são claramente influenciadas pela autocorrelação existente na série de saída x_t . Como mostra a equação (3.36), é difícil realizar os testes de $\rho_{xy}(k)$ e de $v(B)$.

A variância e covariância da estimativa amostral de $\rho_{xy}(k)$ são claramente contaminadas pela estrutura de auto-correlação das séries de entrada x_t como mostra a equação (3.36), tornando os testes de ambos $\rho_{xy}(k)$ e v_k difíceis. Entretanto, se a série de entrada for um ruído branco, isto é, $\rho_x(k) = 0$ para $k \neq 0$, a equação em (3.38), reduz-se a:

$$v_k = \frac{\sigma_y}{\sigma_x} \cdot \rho_{xy}(k) \quad (3.37)$$

Portanto, a função de resposta de impulso v_k é diretamente proporcional a função de correlação cruzada $\rho_{xy}(k)$.

Algumas observações importantes:

1. A função de correlação cruzada $\rho_{XY}(K)$ é definida somente quando X_t e Y_t são processos bivariados conjuntamente estacionários. Caso o processo não seja do tipo estacionário, deve-se aplicar um processo de diferenciações para se obter a estacionariedade desejada, ou aplicar transformações para se estabilizar a variância. Deste modo, a menos que seja mencionado, os processos X_t e Y_t serão assumidos conjuntamente estacionários.
2. No modelo geral de função de transferência

$$X_t = v(B)X_t + N_t, \quad (3.38)$$

Pode-se assumir que a série de entrada X_t segue um processo ARMA, ou seja:

$$\phi_x(B)X_t = \theta_x(B)\alpha_t \quad (3.39)$$

onde α_t é ruído branco. A série α_t , definida como do tipo,

$$\alpha_t = \frac{\phi_x(B)}{\theta_x(B)} \cdot X_t \quad (3.40)$$

é chamada freqüentemente de série de entrada pré ajustada (prewhitened). E aplicando à mesma um processo de transformação de pré ajustamento na série de saída Y_t , obtém-se a série de saída filtrada.

$$\beta_t = \frac{\phi_x(B)}{\theta_x(B)} \cdot Y_t \quad (3.41)$$

fazendo, $\epsilon_t = \theta_x^{-1}(B) \cdot \phi_x(B)n_t$, o modelo de função de transferência em (3.39) (substituindo (3.40) e (3.41) em (3.42), torna-se :

$$\beta_t = v(B)\alpha_t + \epsilon_t, \quad (3.42)$$

Daí, os pesos de resposta de impulso v_j , para a função de transferência podem, conseqüentemente ser encontrados como (ver eq. (3.37)),

$$v_k = \frac{\sigma_\beta}{\sigma_\alpha} \rho_{\alpha\beta}(k) \quad (3.43)$$

onde a partir de então, as etapas fundamentais de identificação dos modelos de função de transferência podem ser conduzidas.

3.11 Construção do Modelo de Função de Transferência

3.11.1 Função de Correlação Cruzada Amostral

Para um conjunto de dados das séries X_t e Y_t , $1 \leq t \leq n$, a função de correlação cruzada;

$$\rho_{xy}(k) = \frac{\gamma_{xy}(k)}{\sigma_x \sigma_y}, \quad k = 0, \pm 1, \pm 2, \dots, \quad (3.44)$$

Passa a ser estimada pela seguinte função de correlação cruzada amostral,

$$\hat{\rho}_{xy}(k) = \frac{\hat{\gamma}_{xy}(k)}{S_x S_y}, \quad k = 0, \pm 1, \pm 2, \dots, \quad (3.45)$$

onde;

$$\hat{\gamma}_{xy}(k) = \begin{cases} \frac{1}{n} \sum_{t=1}^{n-k} (X_t - \bar{x})(Y_{t+k} - \bar{y}), & k \geq 0, \\ \frac{1}{n} \sum_{t=1-k}^n (X_t - \bar{x})(Y_{t+k} - \bar{y}), & k < 0 \end{cases} \quad (3.46)$$

com, $S_x = \sqrt{\hat{\gamma}_{xx}(0)}$ e $S_y = \sqrt{\hat{\gamma}_{yy}(0)}$, e \bar{x} e \bar{y} são as médias amostrais das séries X_t e Y_t respectivamente.

3.11.2 Identificação do Modelo de Função de Transferência

A partir de então, a Função de Transferência $v(B)$ pode ser obtida procedendo da seguinte forma:

1. Pré ajustando as séries de entrada,

$$\phi_x(B)X_t = \theta_x(B)\alpha_t, \quad (3.47)$$

isto é,

$$\alpha_t = \frac{\phi_x(B)}{\theta_x(B)}X_t, \quad (3.48)$$

onde, α_t é uma série de ruído branco, com média zero e variância δ_a^2

2. Calcular a série de saída filtrada. Isto é, transformar a série de saída Y_t utilizando o modelo pré ajustado, no item 1, para gerar a série:

$$\beta_t = \frac{\phi_x(B)}{\theta_x(B)}Y_t, \quad (3.49)$$

3. Calcular a função de autocorrelação cruzada amostral $\hat{\rho}_{\alpha\beta}(k)$ entre α e β para estimar v_k ,

$$\hat{v}_k = \frac{\hat{\sigma}_\beta}{\hat{\sigma}_\alpha} \hat{\rho}_{\alpha\beta}(k), \quad (3.50)$$

O significância da função de correlação cruzada é equivalente a \hat{v}_k , para esta verificação basta analisar seu desvio padrão $(n - k)^{-\frac{1}{2}}$.

4. Identificar b, $\delta_r(B) = (1 - \delta_1 B - \dots - \delta_r B^r)$, $eW_s(B) = (\omega_0 - \omega_1(B) - \dots - \omega_s B^s)$ pelo ajuste padrão de \hat{v}_k com o de v_k , discutidos em 3.11.1 e 3.11.2. Uma vez escolhidos b, r e s, as estimativas preliminares $\hat{\omega}_j$ e $\hat{\delta}_j$ relacionadas com v_k podem ser estimadas, como mostrado em (3.29). Portanto, tem-se uma estimativa da função de transferência $v(B)$ como:

$$\hat{v}_t = \frac{\hat{\omega}_s(B)}{\hat{\delta}_r(B)} B^b. \quad (3.51)$$

3.11.3 Identificação do Modelo de Ruído

Uma vez obtido o modelo preliminar da função de transferência, pode-se calcular a série de ruídos estimada,

$$\hat{N}_t = Y_t - \hat{v}(B)X_t$$

$$\hat{N}_t = Y_t - \frac{\hat{w}(B)}{\hat{\delta}(B)}B^b X_t \quad (3.52)$$

O modelo apropriado para o ruído pode ser identificado examinando-se sua *FAC* e *FACP* ou por outra ferramenta de identificação de séries univariadas de tempo. Então,

$$\phi(B)n_t = \theta(B)a_t \quad (3.53)$$

- **Observações sobre a identificação do Modelo de Função de Transferência.**

Combinando (3.52) e (3.53) temos o seguinte modelo de função de transferência:

$$Y_t = \frac{\omega(B)}{\delta(B)}X_{t-b} + \frac{\theta(B)}{\phi(B)}a_t \quad (3.54)$$

- **Alguns pontos importantes na construção do modelo de função de transferência:**

1. Na construção do modelo é assumido que as variáveis X_t , Y_t e N_t são todas estacionárias. Portanto, para series não estacionárias alguma estabilização da variância e diferenciação podem ser necessárias.
2. No processo de identificação acima, da função de transferência $v(B)$, a serie de entrada deve ser pré-ajustada. O modelo pré-ajustado é usado para filtrar a serie de saída, mas não necessariamente para ajustá-lo. Este é um método normal e simples para a construção de um modelo de função de transferência causal. Entretanto, para a construção de um possível sistema não causal com fenômeno de feedback, onde Y_t é influenciado por X_t e X_t também é influenciado por Y_t , ambas as séries de entrada e saída deveriam ser pré-ajustadas antes

de se examinar sua FCC. Isso é freqüentemente referido como um duplo pré-ajustamento. Entretanto, é geralmente mais vantajoso modelar um sistema não causal por meio de um processo de vetor (VAR).

3.11.4 Estimativas dos Parâmetros do Modelo de Função de Transferência

Após identificação do modelo de função de transferência preliminar,

$$Y_t = \frac{\omega(B)}{\delta(B)} X_{t-b} + \frac{\theta(B)}{\phi(B)} a_t \quad (3.55)$$

é necessário a realização da estimação dos parâmetros $\phi = (\phi_1, \dots, \phi_p)'$, $\theta = (\theta_1, \dots, \theta_q)'$ e σ_a^2 . Pode-se escrever a equação em (3.55) da seguinte forma:

$$\delta(B)\phi(B)Y_t = \phi(B)\omega(B)X_{t-b} + \delta(B)\theta(B)a_t \quad (3.56)$$

ou ainda, de forma equivalente;

$$c(B)y_t = d(B)x_{t-b} + e(B)a_t \quad (3.57)$$

onde:

$$\begin{aligned} c(B) &= \delta(B)\phi(B) = (1 - \delta_1 B - \dots - \theta_r B^r)(1 - \phi_1 B - \dots - \phi_p B^p) \\ &= (1 - c_1 B - c_2 B^2 - \dots - c_{p+r} B^{p+r}), \end{aligned}$$

$$\begin{aligned} d(B) &= \phi(B)\omega(B) = (1 - \phi_1 B - \dots - \phi_p B^p)(\omega_0 - \omega_1 B - \dots - \omega_s B^s) \\ &= d_0 - d_1 B - d_2 B^2 - \dots - d_{p+s} B^{p+s}), \end{aligned}$$

$$\begin{aligned} e(B) &= \delta(B)\theta(B) = (1 - \delta_1 B - \dots - \theta_r B^r)(1 - \theta_1 B - \dots - \theta_q B^q) \\ &= (1 - e_1 B - e_2 B^2 - \dots - e_{r+q} B^{r+q}). \end{aligned}$$

Assim,

$$\begin{aligned} a_t &= y_t - c_1 y_{t-1} - \dots - c_{p+r} y_{t-p-r} - d_0 x_{t-b} + d_1 x_{t-b-1} \\ &\quad + \dots + d_{p+s} x_{t-b-p-s} + e_1 a_{t-1} + \dots + e_{r+q} a_{t-r-q} \end{aligned} \quad (3.58)$$

onde, c_i , d_j e e_k , são funções de δ_i , ω_j , ϕ_k e θ_l . Sob a suposição de que os a_t são ruídos da série com as seguintes características $N(0, \delta_k^2)$. Daí, tem-se a seguinte função condicional de verossimilhança:

$$L(\delta, \omega, \phi, \theta, \sigma_a^2 | b, x, y, x_0, y_0, a_0) = (2\pi\sigma_a^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma_a^2} \sum_{t=1}^n a_t^2 \right] \quad (3.59)$$

onde, x_0 , y_0 , a_0 são valores apropriados para começar a computar os resíduos a_t em (3.64) e se comporta similar aos valores necessários para estimar os modelos **ARIMA**, descritos no Capítulo 2.

3.11.5 Diagnóstico dos Modelos de Função de Transferência

A partir do modelo identificado e seus parâmetros estimados, é necessário realizar a verificação da adequação do mesmo, pois só a partir desta adequação é que o mesmo pode ser utilizado para gerar previsões, controle, assim como, outras finalidades.

No modelo de função de transferência é assumido que a_t são ruído branco e independente da serie de entrada X_t e, portanto, também independente da série de entrada pré-ajustada α_t . Deste modo, a checagem do diagnóstico de um modelo de função de transferência deve-se examinar os resíduos \hat{a}_t do modelo de ruído assim como os resíduos α_t do modelo de entrada pré-ajustado para se ver melhor se as considerações se mantêm.

No modelo de função de transferência, supõe-se que são ruídos brancos independentes da série de entrada, assim como, os pesos α_t . Assim, quando se realiza diagnóstico de um modelo de função de tranferência, tem-se antes que examinar os resíduos \hat{a}_t do modelo realizado dos resíduos α_t .

Para verificar se as suposições são satisfeitas deve-se:

1. **Checagem da Correlação Cruzada:** Deve-se verificar se os resíduos a_t de uma série de entrada são independentes. Para que o modelo seja adequado, a função de autocorrelação amostral $\rho_{\hat{a}\hat{a}}(k)$, entre \hat{a}_t e α_t do modelo de entrada não deve apresentar nenhuma correlação fora do limite de $2(n-k)^{-\frac{1}{2}}$. Usa-se o seguinte teste de probabilidade.

$$Q_0 = m(m+2) \sum_{j=0}^k (m-j)^{-1} \hat{\rho}_{\alpha\hat{a}}(j), \quad (3.60)$$

Com uma distribuição aproximada χ^2 com $(K+1)-m$ graus de liberdade, onde $m = (n - t_0 + 1)$, que é o número de resíduos \hat{a}_t calculados e m é o número de parâmetros estimados da função de transferência $v(B) = \frac{\omega(B)}{\delta(B)}$.

2. **Checagem da Auto-correlação:** Para verificar se o modelo dos resíduos encontra-se adequado basta analisar a *FAC* e a *FACP* de \hat{a}_t , caso o modelo seja adequado, a *FAC* e a *FACP* da amostra de \hat{a}_t não deve mostrar nenhum padrão.

$$Q_1 = m(m+2) \sum_{j=1}^k (m-j)^{-1} \hat{\rho}_a(j). \quad (3.61)$$

Q_1 é uma estatística aproximada baseada na distribuição χ^2 com $(k-p-q)$ graus de liberdade. Isto dependerá do número de observações utilizadas para a construção do modelo de resíduos.

Caso o modelo seja inadequado, deve-se levantar duas hipóteses:

- A função de transferência $v(B)$ é incorreta, neste caso, de qualquer maneira se o modelo proposto para o ruído é adequado, tem-se $\rho_{aa}(k) \neq 0$ e $\rho_a(k) \neq 0$ para algum k . Para se observar isto, basta supor que o modelo correto é:

$$Y_t = v(B)X_t + \psi(B)a_t, \quad (3.62)$$

Se o termo da função de transferência $v(B)$ for escolhido de forma incorreta, então seremos conduzido a um erro do termo b_t .

$$Y_t = v_0(B)X_t + \psi_0(B)a_t \quad (3.63)$$

Então,

$$b_t = \psi_0^{-1}(B)[v(B) - v_0(B)]X_t + \psi_0^{-1}(B)\psi(B)a_t \quad (3.64)$$

Assim, b_t será não somente uma função de correlação cruzada de X_t com a_t , mas ambas serão autocorrelacionadas. Isto é verdade, mesmo quando $\psi_0(B) = \psi(B)$, isto é, quando o modelo do ruído está correto. Então, a solução é primeiramente reidentificar a função de transferência $v(B)$ para só assim, modificar o modelo de ruído.

- A função de transferência está correta, e apenas o modelo do ruído é inadequado. Neste caso, $\rho_{\alpha\alpha}(k) = 0$ para algum k . O teste padrão para $\hat{\rho}_{\hat{a}}(k) = 0$ pode ser usado para modificar o modelo de ruído $[\theta(B)/\phi(B)]a_t$.

Em resumo, para que um modelo de função de transferência seja adequado ambos $\hat{\rho}_{\hat{a}}(k) = 0$ e $\hat{\rho}_{\hat{a}}(k) = 0$ devem ser estatisticamente não significantes e não mostrar nenhum padrão, uma vez que o modelo de ruído seria contaminado por uma função de transferência incorreta, na checagem do diagnóstico de modelos de função de transferência. É prudente realizar primeiro o cheque da correlação cruzada. Em essência, é similar a um modelo de análise de variância com um termo de interação onde o teste de um efeito de interação deveria ser realizado antes do teste dos efeitos principais.

3.11.6 Estimação do Modelo de Função de Transferência

Para a estimação dos modelos de função de transferência, pode-se utilizar as seguintes expressões:

$$(1 - B)Y_t = \frac{w_0}{(1 - \delta_1 B)}(1 - B)X_t + \frac{1}{(1 - \phi_1 B)}a_t \quad (3.65)$$

3.12 Validação de Modelos de Função de Transferência

Os modelos de função de Transferência são apropriados para melhorar as previsões da série de saída y_t utilizando sua história passada, associada à série de entrada x_t . Isto é válido se a série de entrada for o indicador principal.

3.12.1 Erro Quadrático Médio de Séries de Entrada e Saída Estacionárias

Dadas duas séries X_t e Y_t estacionárias e relacionadas no seguinte modelo de função de transferência.

$$Y_t = \frac{\omega(B)}{\delta(B)}B^b X_t + \frac{\theta(B)}{\phi(B)}a_t, \quad (3.66)$$

e

$$\phi_x(B)X_t = \theta_x(B)\alpha_t, \quad (3.67)$$

onde, $\omega(B)$, $\delta(B)$, $\theta(B)$, $\phi(B)$, $\phi_x(B)$ e $\theta_x(B)$ são polinômios de B de ordem finita; as raízes de $\omega(B) = 0$, $\delta(B) = 0$, $\theta(B) = 0$, $\phi(B) = 0$, $\phi_x(B) = 0$ e $\theta_x(B) = 0$ estão

fora do círculo unitário. E a_t e α_t são resíduos independentes com média zero e variância σ_a^2 e σ_α^2 , respectivamente.

$$u(B) = \frac{\omega(B)B^b\theta_x(B)}{\theta(B)\phi_x(B)} = u_0 + u_1 + u_2B^2 + \dots, \quad (3.68)$$

e

$$\psi(B) = \frac{\theta(B)}{\phi(B)} = 1 + \psi_1B + \psi_2B^2 \quad (3.69)$$

A qual pode ser escrita, ainda da seguinte forma:

$$y_t = u(B)\alpha_t + \psi(B)a_t$$

$$y_t = \sum_{j=0}^{\infty} u_j\alpha_{t-j} + \sum_{j=0}^{\infty} \psi_j a_{t-j} \quad (3.70)$$

onde, $\psi_0 = 1$. Assim, $Y_{t+l} = \sum_{j=0}^{\infty} u_j \cdot \alpha_{t+l-j} + \sum_{j=0}^{\infty} \psi_j \cdot a_{t+l-j}$. Deixando, $\hat{Y}_t(l) = \sum_{j=0}^{\infty} u_{l+j}^* \cdot \alpha_{t-j} + \sum_{j=0}^{\infty} \psi_{l+j}^* \cdot a_{t-j}$, realizar previsões um passo a frente de Y_{t+l} . Então o erro de previsão torna-se:

$$\begin{aligned} Y_{t+l} - \hat{Y}_t(l) &= \sum_{j=0}^{l-1} [u_j \cdot \alpha_{t+l-j} + \psi_{t+l-j}] \\ &- \sum_{j=0}^{\infty} [u_{l+j}^* - u_{l+j}] \cdot \alpha_{t-j} - \sum_{j=0}^{\infty} (\psi_{l+j}^* - \psi_{l+j}) \cdot a_{t-j}, \end{aligned} \quad (3.71)$$

O erro quadrático médio **EQM** das previsões $E[Y_{t+l} - \hat{Y}_t(l)]^2$ é dado por:

$$E[Y_{t+l} - \hat{Y}_t(l)]^2 = \sum_{j=0}^{l-1} (\sigma_\alpha^2 u_j^2 + \sigma_a^2 \psi_j^2) + \sum_{j=0}^{\infty} \sigma_\alpha^2 (u_{l+j}^* - u_{l+j})^2 + \sum_{j=0}^{\infty} \sigma_a^2 (\psi_{l+j}^* - \psi_{l+j})^2 \quad (3.72)$$

Observa-se que este erro é minimizado quando, $u_{l+j}^* = u_{l+j}$ e $\psi_{l+j}^* = \psi_{l+j}$. Ou seja, o erro quadrático médio de previsão mínimo entre \hat{Y}_t e Y_{t+l} no tempo t é dado pela esperança condicional de Y_{t+l} no tempo t . Isto deve-se pelo fato, que $E[Y_{t+l} - \hat{Y}_t(l)] = 0$, e a variância de previsão ser dada por:

$$V(l) = E[Y_{t+l} - \hat{Y}_t(l)]^2 = \sigma_\alpha^2 \sum_{j=0}^{l-1} u_j^2 + \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2. \quad (3.73)$$

3.12.2 Resumo

Com o objetivo de se identificar o modelo e as componentes pertencentes da função de transferência, é necessário a utilização das Funções de Auto-Correlações Cruzadas FCC, as quais proporcionam a identificação de existência de lags negativos.

Para modelagem e previsão válidas a partir de modelos de funções de transferência, é importante observar que nenhum lag negativo deve ter spikes significativos. Caso isto aconteça, então ocorre um chamado modelo com feedback. Neste caso, modelos do tipo Autotregressivo Vetorial (**VAR**) são mais adequados. Em seguida, deve-se checar os seguintes itens:

- A quantidade de defasagens no modelo
- Se há spike no lag zero o modelo então não apresenta defasagem;
- O modelo terá um período de defasagens n se não há spikes no gráfico de correlação cruzada no lag 0;
- Depois do período de defasagem deve haver um spike no gráfico de correlação cruzada
- O número de spikes depois do primeiro spike corresponde ao número de parâmetros do numerador na função de transferência do modelo
- Se não existem spikes depois do primeiro spike, então o processo é provavelmente ruído branco;
- Seguindo os spikes, o gráfico de correlação cruzada pode ter um dos seguintes padrões:
 1. Ele pode decair exponencialmente, o que indica a necessidade de um parâmetro no denominador na função de transferência do modelo;
 2. Decair exponencialmente com padrão senoidal, que indica a necessidade de dois ou mais parâmetros no denominador na função de transferência do modelo;
 3. Cair imediatamente para zero, que indica que não é necessário nenhum parâmetro no denominador no modelo.

Capítulo 4

Sistemas de Implementações Inteligentes - Redes Neurais

4.1 Introdução

As Redes Neurais Artificiais estão embasadas diretamente nos estudos realizados sobre a estrutura fisiológica do cérebro humano para tentar comportar-se como um programa de forma inteligente no processamento de informações. Alguns estudos realizados consideram que a riqueza computacional do cérebro humano está associada ao grande número de neurônios, interconectados por uma rede complexa de sinapses. Estima-se que a quantidade de neurônios existentes no cérebro humano esteja na casa dos bilhões. Contudo, a velocidade de processamento destes componentes é relativamente baixa, quando comparada aos computadores tradicionais. Esta deficiência na velocidade de processamento dos neurônios é superada pela imensa quantidade de neurônios existentes operando de forma paralela (PORTUGAL e FERNANDES, 1995).

Estudos demonstram que nem sempre quando se trabalha com alguma ferramenta ou técnica da Inteligência Artificial - IA de forma individualizada, obtém-se os resultados desejados. Muitos pesquisadores recentemente vêm utilizando as mais variadas estratégias e técnicas referentes à Inteligência Computacional para *implementação de modelos* mais eficazes, ou seja, modelos que apresentem melhores ajustes às variáveis em questão e assim, representem de forma mais adequada a ocorrência do fenômeno estudado. Esta implementação conjugada de diferentes técnicas são de grandes importância, principalmente quando se trata de funções complexas, pois tendem a proporcionar ao pesquisador melhores resultados em comparação com aqueles obtidos quando aplicado metodologias de forma individualizadas.

Atualmente busca-se implementações utilizando mais de uma técnica do grupo da **I.A** para a melhoria de resultados, como por exemplo, a utilização de forma conjugada de **Redes Neurais e Algoritmos Genéticos** na busca da otimização de um mesmo problema, assim como o uso de **Redes Neurais e Lógica Fuzzy** para ajustar modelos mais robustos. Essas conjugações são conhecidas como **Soluções Híbridas**.

Muitas são as técnicas empregadas como estratégias para formulação de implementações; as mais utilizadas são:

- Redes Neurais
- Lógica Fuzzy
- Computação Evolutiva
- Metodologias Estatísticas em Geral, etc....

4.2 Redes Neurais Artificiais - RNA's

As pesquisas sobre redes neurais iniciaram em 1943, quando *Warrem McCulloch* e *Walter Pitts* estabeleceram as bases da neurocomputação, concebendo procedimentos matemáticos análogos ao funcionamento do cérebro biológico. Porém estes procedimentos jamais foram testados pelos seus criadores em situações práticas. No entanto, o passo mais importante quando se trata de estudar redes neurais, foi dado por **Donald Hebb** em 1949, quando ele propôs um modo de condicionar as redes neurais à capacidade de aprendizado. A partir de então, cada vez mais a comunidade científica vem implementando estruturas de RNA's mais eficazes. Este fato, se deve à evolução da tecnologia computacional que contribuiu com computadores cada vez mais modernos e velozes em processamento.

Segundo Dahmer(1998), quando se trata das RNA's, inúmeras características podem ser citadas sobre sua utilização no meio científico:

- Controle altamente paralelo e distribuído;
- Capacidade de aprendizagem através da observação de um conjunto de exemplos, sendo estes com ou sem informação de resposta desejadas a cada estímulo;

- Robustez, isto é, uma queda de desempenho gradual no eventual funcionamento irregular ou até inutilização de parte da rede;
- Manutenção do desempenho na presença de ruído e capacidade de lidar com dados incompletos.

Segundo Hernandez(2005), as Redes Neurais ou Redes de Neurônios Artificiais são sistemas de computação adaptativos inspirados nas características de processamento de informação encontradas nos neurônios biológicos e nas características de suas interconexões, (Fig. 4.1).

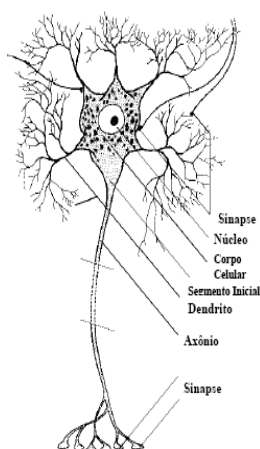


Figura 4.1 *Estrutura de Neurônio Biológico*

O neurônio biológico pode ser visto como o dispositivo computacional elementar do sistema nervoso, composto de muitas entradas e uma saída. As entradas são formadas através das conexões sinápticas que conectam os dendritos aos axônios de outras células nervosas. Os sinais que chegam por estes axônios são pulsos elétricos conhecidos como impulsos nervosos ou potenciais de ação e, constituem a informação que o neurônio processa para produzir como saída um impulso nervoso no seu axônio (Kovacs, 1996).

Estes sistemas, podem ser vistos como sistemas massivamente paralelos que podem ser implementados tanto em Hardware quanto em Software, sendo que os elementos de processamento individualmente têm capacidade relativamente limitados. Tais elementos de processamento básicos são os chamados neurônios artificiais, (Fig. 4.2).

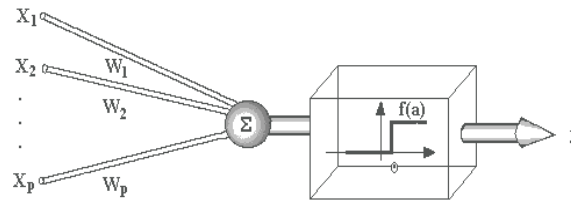


Figura 4.2 *Neurônios Artificial apresentado por McCulloch*

As redes neurais apresentam estratégias na pretensão de igualar ou até mesmo superar alguns aspectos da inteligência humana. Baseia-se no entendimento das estruturas de armazenamento e de processamento de informações realizados pelo sistema nervoso. A forma específica na qual a atividade neural é representada no modelo contextualizado de forma matemática pode assumir diversas simplificações. Todo este processo dependerá do grau de refinamento na representação de fenômeno biológico que se tem em mente, e varia conforme a ênfase do trabalho a ser realizado.

4.2.1 Funções de ativação

As funções de transferência (ou de ativação), f , utilizadas na construção de redes neurais artificiais têm sido, entre outras: funções degrau e funções lineares do tipo (Fig. 4.3 e 4.4):

- **Função de ativação linear**

$$f(x) = a.x \quad (4.1)$$

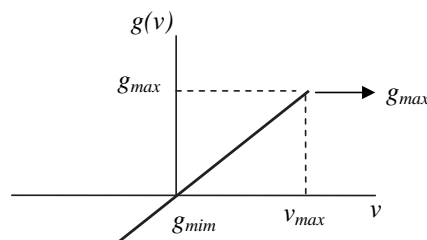


Figura 4.3 *Função de ativação linear*

- Função de ativação degrau

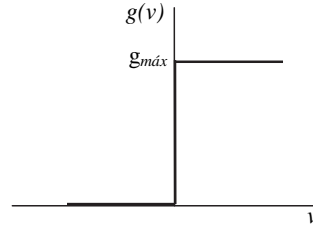


Figura 4.4 Função de ativação degrau

$$g(v) = \begin{cases} a, & v \geq \theta \\ -a, & v < \theta \end{cases} \quad (4.2)$$

- Função de ativação rampa

São funções do tipo anterior mas com um contradomínio igual a um intervalo fechado (Fig. 4.5)

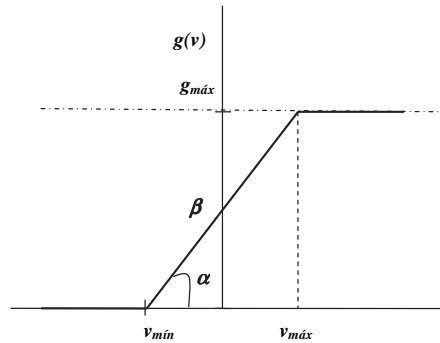


Figura 4.5 Função de ativação rampa

$$g(v) = \begin{cases} a, & v > \theta \\ a, & v \geq \theta \\ -a, & v < \theta \end{cases} \quad (4.3)$$

Também são muito utilizadas as funções crescentes de classe C^∞ cujo contradomínio se encontra contido num certo intervalo aberto e limitado, essas funções de transferência quando crescente e cujo contradomínio se encontra no intervalo $[0, 1]$ ou $[-1, 1]$ e tais que os limites são da forma:

$$\lim_{x \rightarrow +\infty} f(x) = 1 \quad (4.4)$$

$$\lim_{x \rightarrow -\infty} f(x) = 0 \quad (4.5)$$

$$\lim_{x \rightarrow -\infty} f(x) = -1 \quad (4.6)$$

são conhecidas como funções **sigmodais**, dessa classe, as mais utilizadas são as funções **tangente hiperbólica** e as funções de **Fermi** caracterizada por:

$$Fermi \longrightarrow f(x) = \frac{1}{1 + e^{-x}} \quad (4.7)$$

Dentre os vários modelos de redes neurais artificiais que serão apresentados, cumpre definir a função de ativação do neurônio, **$g(v)$** que, tem sido aproximada por meio de inúmeras funções monotônicas que procuram preservar essas característica mediante a constância de propósitos. A partir daí procede-se das seguintes formulações matemáticas:

- Função 1: Integração espaço temporal das atividades de entrada de uma árvore dentrital.

$$f_T = g\left[\int_T \cdot \sum_i \alpha_i(t) x_i(t) dt\right] \quad (4.8)$$

em que os $\alpha_i(t)$ representam os ganhos sinápticos, e os $x_i(t)$ as entradas dos neurônios.

Por outro lado, a função **$g(v)$** possui as conexões sinápticas excitatória e inibitória.

- Função 2: Modelo matemático do perceptron de Seebert.

$$g(v) = \frac{v^k}{v^k + \alpha} \quad (4.9)$$

A aproximação de **Seebert** permite, ajustar o ganho médio nessa faixa em que apareceu o parâmetro **K** . Normalmente, as funções de ativação mais utilizados são a sigmóide e a tangente hiperbólica, abaixo representados com as suas respectivas derivadas.

- Função 3: Função matemática sigmóide utilizada em redes neurais, Fig (4.6).

$$g(v) = [1 + \exp(-2\beta)]^{-1} = g(v) = \frac{1}{1 + e^{(av)}} \quad (4.10)$$

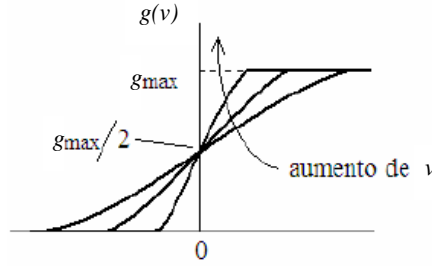


Figura 4.6 Função de ativação sigmóide -(logsig)

- Função 4: Função Tangente hiperbólica utilizada em redes neurais, Fig. (4.7).

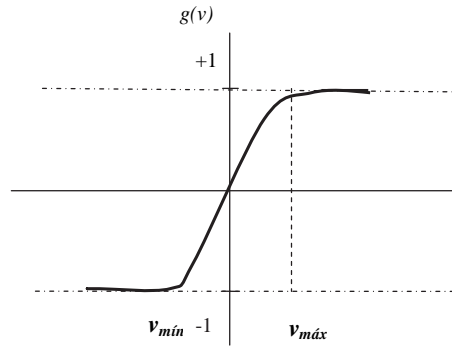


Figura 4.7 Função de ativação tangente hiperbólica -(tansig)

$$g(v) = \tanh(\beta v) \quad (4.11)$$

- Função 5: Exemplo de utilização de derivadas nas equações de redes neurais.

$$\frac{d}{dv} = ((1 + \exp(-2\beta v))^{-1} \quad (4.12)$$

$$\frac{d}{dv} = 2\beta(1 - g(v)) \cdot g(v) \quad (4.13)$$

- Função 6: A derivada da função 5 em relação a v fornece,

$$\frac{d}{dv} = (\tanh(\beta v)) = \beta(1 - g(v))^2 \quad (4.14)$$

A partir dessas equações pode-se implementar alguns importantes algoritmos ligados às redes neurais. Estes algoritmos estão abaixo relacionados:

1. Algoritmo Booleano de Warren McCulloch e Pitts
2. Algoritmo de Hebb
3. Algoritmo de Widrow Stanford
4. Algoritmo de Multicamadas de Rummelhart, Hinton e Williams

4.3 Algumas Teorias Referentes a Algoritmos de Redes Neurais

4.3.1 Modelos de Redes Neurais de Warren McCulloch e Pitts

Uma das grandes contribuições referentes ao estudo de redes neurais, deu-se com estudos realizados por **Warren McCulloch**. Nesses estudos procurou-se implementar algumas funções do tipo Booleanas.

Segundo (**McCulloch e Pitts**) em publicação de seu primeiro artigo referente a redes neurais, "a inteligência é equivalente a cálculo proporcional, que por sua vez pode ser implementado por funções booleanas. Por outro lado, o sistema nervoso é composto de rede de neurônios que, com as devidas simplificações, apresentam a capacidade básica de implementar essas funções". Esses autores foram os primeiros a proporem um modelo computacional para o neurônio biológico, o qual abriu caminho para a definição correta da inteligência artificial e a modulação de redes neurais, abrangendo o conceito de taxa de aprendizado, de retenção e de treinamento, (Haikin, 2001).

A partir de uma minimização dos erros e de uma maximização das taxas de retenção através de modelos de redes neurais, talvez se chegue a uma das possíveis soluções para o treinamento e desenvolvimento integrado, de forma a fornecer uma melhor visão sistêmica que venha lançar as bases do futuro conhecimento e toda uma cadeia integrada de suprimentos.

A taxa de aprendizado reflete a taxa com que os ganhos se alteram em consequência dos erros. Trata-se de uma regra local que visa a atualização dos parâmetros do sistema, ou seja, este algoritmo nos mostra a taxa de retenção dos conhecimentos e minimização dos erros, com uma consequente maximização dos resultados obtidos.

4.3.2 Neurônios de Remmelhart, Hilton e Williams - RHW

O algoritmo de retropropagação para redes de neurônios de múltiplas camadas conhecido por **Remmelhart, Hilton e Williams** (1986), Figura (4.8), permitiu superar uma das grandes limitações fundamentais para o treinamento em redes complexas, usando funções de ativação sempre lineares, consistindo num diagrama de blocos e neurônios conhecidos como **RHW**, composto basicamente de um **Adaline de Windrow**, seguido de uma função de ativação $g(v)$, (HAIKIN, 2001).

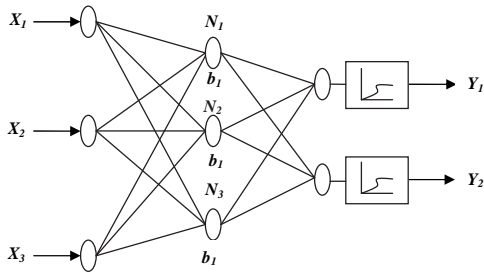


Figura 4.8 Rede com neurônio (RHW)

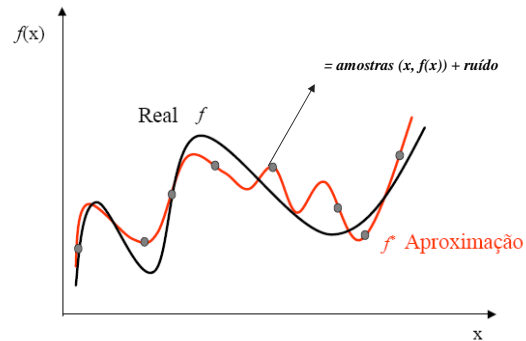


Figura 4.9 Erro de aproximação do (RHW)

Este neurônio gera uma função saída que depende de uma série de entradas convergentes no ponto conhecido como mínimo de erro, a fim de que se obtenha os resultados esperados. Consequentemente, a ocorrência será a minimização das distâncias entre o previsto e o real da função de treinamento, ver Figura (4.9).

4.3.3 Modelos de Saída Binária - Binário

Esta classe de modelos, também conhecida como **Adaline** assume que modelos neurais com saída binária são suficientemente poderosos para representar uma classe bastante ampla de problemas computacionais, devido ser incorporada ao modelo neural uma função que permitirá a viabilização das funcionalidades não linear. Esta função é comumente conhecida como **função de transferência**. Estes modelos apresentam uma grande vantagem em relação aos demais, pois apresentam uma simplicidade estrutural e com isso permitem a formalização matemática em um grau não possível para os modelos mais complexos, (Haikin, 2001).

A forma mais simples desta função de transferência f pode adotar valores (binários) discretos. Esta adoção permite que o modelo apresente saída binária, permitindo assim, a representação funcional de característica lógica realizadas pelos neurônios biológicos.

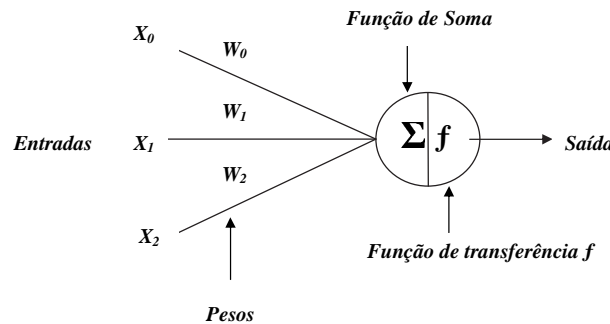


Figura 4.10 Neurônio artificial do tipo (Adaline)

Todos os tipos de redes neurais apresentam basicamente a mesma unidade de processamento, caracterizando-se desta forma como um neurônio artificial e simulando, deste modo, um neurônio biológico com uma estrutura com varias entradas, que correspondem às conexões sinápticas com outras unidades similares a ele próprio, e uma saída cujo valor é influenciado diretamente pela soma ponderada de todas as saídas dos demais neurônios a ele conectado. No cálculo de integração dos estímulos das várias entradas, a ponderação é dada por pesos que cada conexão possui e sua saída corresponde a soma, como:

$$\sum = w_1x_1 + w_2x_2 + \dots + w_nx_n + b = \sum_{i=1}^n w_ix_i + b \quad (4.15)$$

quando;

Este modelo de redes neurais multicamadas pode também ser considerado como *Adaline de Widrow de uma função de ativação do tipo $g(v)$* . Sempre que se trabalha com este tipo de algoritmo deve-se escolher um conjunto de parâmetros iniciais e uma taxa de aprendizado, estabelecendo assim um critério de parada que representa o número máximo de interações, (Haikin, 2001).

4.3.5 Propagação do erro

Depois de apresentado o padrão de entrada para a rede, a resposta de uma unidade é propagada como entrada para as unidades na camada seguinte (ocultas), até a camada de saída, onde é obtida a resposta da rede, e onde, o erro é calculado. Como o objetivo principal é a busca do erro mínimo, caso este, não seja o ótimo a rede retropropaga o erro, e o processo recomeça até que o erro mínimo seja obtido Fig.(4.12), (Haikin, 2001).

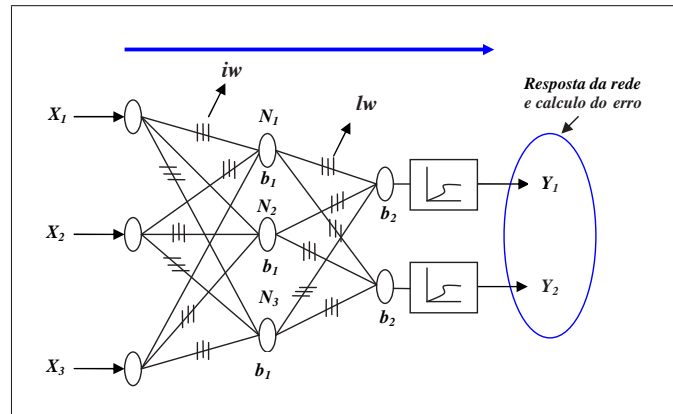


Figura 4.12 Propagação do erro na rede

$X_i \longrightarrow$ variáveis de entrada apresentada a rede;

$iw \longrightarrow$ pesos de entrada;

$lw \longrightarrow$ pesos de saída;

$N's$ neurônios de composição da rede;

$B's$ bias;

$Y_1 e Y_2 \longrightarrow$ resposta na saída da rede.

4.3.6 Retropropagação do erro

O "Backpropagation" é um algoritmo criado com a finalidade de obter treinamento de redes Multi-Camadas, e é comumente o mais utilizado no meio acadêmico. Este algoritmo baseia-se no Aprendizado Supervisionado por Correção de Erros gerados pela saída da rede, ou seja, este algoritmo calcula os erros locais, para cada unidade, desde a camada de saída até a de entrada. Também referenciado de Computação para Trás ou (Retropropagação) Fig (4.13):

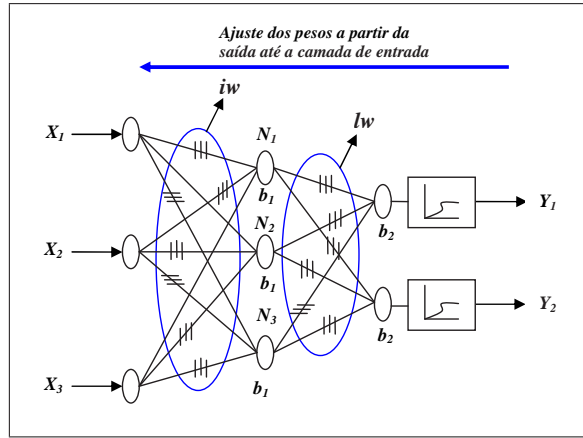


Figura 4.13 Retropropagação do erro na rede

$\delta_j(n) = e_j(n)O_j(n)(1 - O_j(n)) \rightarrow$ erro local para a unidade da camada de saída.

$\delta_j(n) = O_j(n)(1 - O_j(n)) \sum \delta_k w_{jk} \rightarrow$ erro local para as unidades das demais camadas.

$O_j(1 - O_j) \rightarrow$ função de ativação diferenciada em função do argumento -(valor de ativação).

$\delta_k \rightarrow$ é o erro das unidades da camada anterior conectadas a unidade j .

$W_{jk} \rightarrow$ são os pesos das conexões com a camada anterior.

4.3.7 Processo de Aprendizado por Correção de Erros

O Processo de Aprendizado por Correção de Erros é inteiramente baseado no processo de aprendizado por supervisão (**Aprendizado Supervisionado**). Para que determinada Rede Neural Artificial possa fornecer resultados convenientes, é necessário que passe por

uma fase de treinamento, onde seus pesos são ajustados de forma que ela se adapte aos diferentes estímulos de entrada pois, no decorrer deste processo de treinamento é que ocorre o seu aprendizado.

$$e = d_j - O_j \quad (4.16)$$

O erro de uma Rede Neural pode ser calculado como a diferença entre a saída desejada e a saída real gerada pela rede, fornecida em um ensino supervisionado.

Para um estímulo k , tem-se:

e - sinal de erro;

d_j - saída desejada apresentada durante o treinamento;

O_j - saída real da rede após a apresentação do estímulo de entrada.

Durante o aprendizado supervisionado, os erros vão sendo calculados sucessivamente, até que cheguem a um valor satisfatório, definido a priori. Sendo assim, surge uma curva de erros, a qual está diretamente relacionada à natureza do modelo de neurônio utilizado. Se a rede é formada por unidades lineares, como no modelo de **McCulloch e Pits**, na superfície de erro será encontrado um único valor mínimo. Por outro lado, se a rede é constituída por unidades não-lineares, podem ser encontrados diversos valores mínimos chamados de mínimos locais, além do mínimo global.

O processo de Aprendizado por Correção de Erros utiliza algoritmos para caminhar sobre a curva de erros, com o intuito de alcançar o menor valor de erro possível, o mínimo global, ver Figura 4.14. Muitas vezes, o algoritmo não alcança este mínimo global, atingindo o que chamamos de mínimo local. Caso este erro alcançado seja desfavorável, é necessário recomençar processo de aprendizado, (Haikin, 2001).

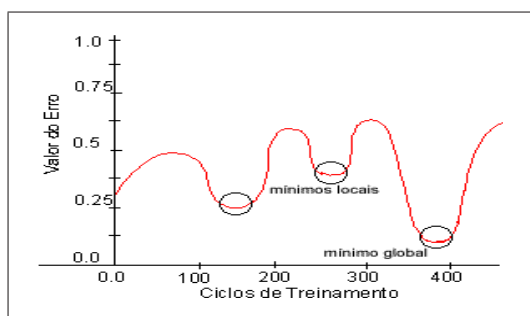


Figura 4.14 Superfície de erro com mínimos locais e o mínimo global

Para a correção do erro, os pesos da rede devem ser ajustados, de forma a aproximar a saída real à desejada. Tal ajuste dependerá do próprio erro calculado; do valor do estímulo de entrada que é "transmitido" pelo peso a ser ajustado; e também da taxa de aprendizado, a qual relaciona-se à cautela com que a curva de erros é percorrida. Para um dado estímulo k , no passo de treinamento n :

$$\Delta W_{ij} = \eta e_k(n) x_j(n) \quad (4.17)$$

$\Delta w(n)$ - valor de ajuste a ser acrescentado ao peso w_{ij} ;

η - taxa de aprendizado;

$e(n)$ - valor do erro;

$x_j(n)$ - valor do estímulo.

O valor atualizado do peso será dado pela expressão:

$$W(n+1) = W(n) + \Delta W_{kj}(n) \quad (4.18)$$

Portanto, podemos utilizar a *Regra Delta* para corrigir os valores dos pesos, minimizando a função de erro ε , também conhecida como "*função de custo*":

$$\varepsilon(n) = \frac{1}{2} e^2(n) \quad (4.19)$$

Onde:

$e(n)$ - erro da rede no passo n do treinamento;

$\varepsilon(n)$ - valor da função de custo no passo n do treinamento.

4.3.8 Redes Neurais: Pontos Fortes e Limitações

Conforme Simões e Shaw(1999), as redes neurais oferecem características inteligentes muito interessantes, tais como: aprendizagem, adaptação, tolerância à falhas e generalização. Assim, o ponto forte de aplicação de redes neurais está no reconhecimento de padrões, reconhecimento de caracteres e formas, estimação de funções não lineares, previsões financeiras, e controle de processos. O controle de processos também pode ser considerado uma tarefa de reconhecimento de padrões quando a rede neural puder reconstruir modelos de

conjuntos de dados. Na ausência de modelos matemáticos de processos, as redes neurais podem utilizar um histórico de dados para se construir modelos preditivos. Depois de treinada, a rede neural pode se tornar um equipamento eletrônico equivalente, ou seja, um modelo que pode então prever as reações do processo às novas condições. Entretanto as redes neurais têm certas limitações que devem ser consideradas:

- Não é possível traçar a maneira pela qual uma rede neural chegou à uma determinada solução.
- É impossível se interpretar as causas de um comportamento particular, uma vez que não se pode olhar para o sentido dos pesos e conexões de uma rede neural; além de ser impossível alterar manualmente a estrutura de uma rede neural para se obter o resultado desejado.
- A maioria dos métodos de treinamento não são completamente compreendidos, pois além de depender de um enfoque de tentativa de erro, há poucas regras de projeto a serem seguidas. Por exemplo, a definição do número de neurônios, ou o estabelecimento de uma determinada faixa de convergência depende de experimentação com os dados.
- O tempo de treinamento não é previsível, podendo em alguns casos ser muito longo.
- O tempo de execução depende do número de conexões, sendo aproximadamente proporcional ao quadrado do número de neurônios utilizados, onde a introdução de alguns nós a mais na rede pode aumentar consideravelmente o tempo de execução.

4.4 A Estatística do Processo de Aprendizagem

Nesta seção, retrata-se os aspectos estatísticos de aprendizagem. Segundo (Haikin, 2001), neste processo o interesse não está somente na evolução do vetor de pesos \mathbf{w} enquanto a rede neural passa por um algoritmo de aprendizagem, e sim no desvio entre uma função alvo $f(\mathbf{x})$ e a função real $F(\mathbf{x}, \mathbf{w})$ realizada pela rede neural, sendo que o vetor \mathbf{x} representa o sinal de entrada e o desvio é expresso em termos estatísticos. Lembre-se, que uma rede neural é meramente uma forma pela qual um conjunto de medidas que caracterizam determinado fenômeno pode ser explicado, denominado de *conhecimento empírico* e

este fenômeno físico ou ambiente de interesse pode ser codificado através de treinamento, um fenômeno estocástico descrito por um vetor aleatório \mathbf{X} formado por um conjunto de variáveis independentes, e um escalar D na qual caracteriza uma variável dependente, onde o vetor aleatório \mathbf{X} pode apresentar os mais diversos significados físicos particulares, diferentes entre si.

Nesta situação adota-se que a variável dependente D é do tipo escalar e que foi formulada simplesmente para facilitar a exposição, sem ocasionar perda de generalidade. Portanto, caso se adote um vetor aleatório constituído de N realizações, uma amostra de treinamento pode ser representada da seguinte forma:

$$\tau = (x_i, d_i)_{i=1}^N \quad (4.20)$$

onde;

$(x_i)_{i=1}^N$ representação do vetor aleatório X com N realizações;

$(d_i)_{i=1}^N$ conjunto correspondente de realizações do escalar aleatório D .

Observe que normalmente não se conhece a relação funcional exata entre X e D , por este fato normalmente, é utilizado o seguinte modelo estatístico proposto por (White, 1989);

$$D = f(X) + \epsilon \quad (4.21)$$

com;

$f(\cdot)$ é uma função determinística do seu argumento vetorial;

ϵ é um erro de expectativa aleatório que representa a nossa falta de conhecimento sobre a dependência de D e \mathbf{X} .

O modelo estatístico descrito por White representado na função (4.21) é denominado como um *modelo regressivo* com o erro ϵ caracterizado como variável aleatória com média zero e probabilidade de ocorrência positiva, cuja representação é mostrada na Fig. 4.15.

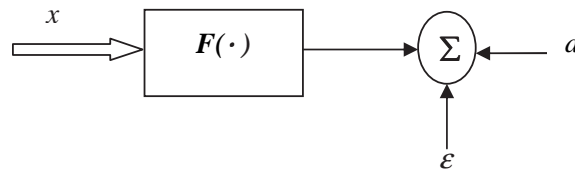


Figura 4.15 *Modelo Matemático Regressivo*

4.4.1 O Erro de Expectativa (ϵ)

O valor médio do erro de expectativa (ϵ), dada qualquer realização \mathbf{x} , é zero, isto é,

$$E = [\epsilon/x] = 0 \quad (4.22)$$

onde \mathbf{E} é o operador estatístico do valor esperado (esperança matemática). Deste modo, podemos afirmar que a função de regressão $f(X)$ é a média condicional da saída do modelo D , dado que a entrada $\mathbf{X}=\mathbf{x}$, que segue diretamente da equação (4.21), considerando a esperança E na equação (4.22) assim;

$$f(x) = E[D|x] \quad (4.23)$$

É muito importante observar que o erro de expectativa (ϵ) não é correlacionado com a função de regressão $f(X)$, isto é;

$$E[\epsilon f(X)] = 0 \quad (4.24)$$

O modelo regressivo da Fig. 4.15 é uma descrição "matemática" de um ambiente estocástico. O seu propósito é utilizar o vetor \mathbf{X} para explicar ou prever a variável dependente \mathbf{D} . A Figura (4.16) é o modelo "físico" correspondente do ambiente.

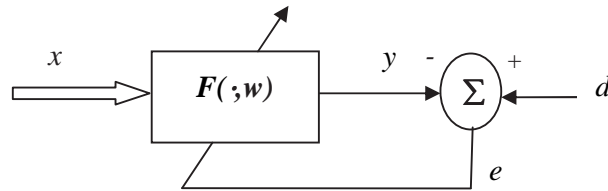


Figura 4.16 *Modelo Físico de Rede Neural*

O propósito deste segundo modelo é baseado em uma rede neural, e tem por característica codificar o conhecimento empírico, representado pela amostra de treinamento \mathfrak{S} em um conjunto correspondente de vetores de pesos sinápticos, \mathbf{w} , como mostrado por:

$$\tau \rightarrow \mathbf{w} \quad (4.25)$$

Na verdade, a rede neural fornece uma "aproximação" para o modelo regressivo da (Fig. 4.16). Suponha que a resposta real da rede neural, produzida em resposta ao vetor de entrada \mathbf{x} , seja representada pela variável aleatória do tipo;

$$Y = F(X, w) \quad (4.26)$$

onde, $\mathbf{F}(\cdot, \mathbf{w})$ é a função de entrada-saída realizada pela rede neural \mathfrak{S} . Observa-se que préviamente conhecidos os dados de treinamento o vetor de peso \mathbf{w} é obtido pela minimização da função de custo dado por;

$$\xi(w) = \frac{1}{2} \sum_{i=1}^N (d_i - F(x_i, w))^2 \quad (4.27)$$

onde $1/2$ é um fator de consistência;

$\xi(w)$ é a diferença quadrática entre a resposta desejada \mathbf{d} e a resposta real \mathbf{y} da rede neural, calculada como a média sobre todo o conjunto de dados de treinamento \mathfrak{S} . Observe que caso adote-se um operador média $E_{\mathfrak{S}}$ tomado sobre todo o conjunto de treinamento \mathfrak{S} é representado por x e d , onde os pares formados (\mathbf{x}, d) caracterizam uma amostra de treinamento de \mathfrak{S} . Por outro lado, o operador estatístico do valor esperado \mathbf{E} atua sobre todo o algoritmo de variáveis aleatórias \mathbf{X} e \mathbf{D} , incluindo desta forma \mathfrak{S} como um subconjunto.

Considerando-se a transformação descrita pela Eq. (4.27), podemos usar alternativamente $F(\mathbf{x}, \mathbf{w})$ e $F(\mathbf{x}, \mathfrak{S})$ e assim rescrever:

$$\xi(w) = \frac{1}{2} E_{\mathfrak{S}}[(d - F(x, \mathfrak{S}))^2] \quad (4.28)$$

Adicionando e subtraindo $f(\mathbf{x})$ ao argumento $(d - F(\mathbf{x}, \mathfrak{S}))$ e então utilizando a Eq. (4.28), podemos escrever;

$$d - F(x, \mathfrak{S}) = (d - f(x)) + (f(x) - F(x, \mathfrak{S})) = \epsilon + (f(x) - F(x, \mathfrak{S})) \quad (4.29)$$

Observe que esta expressão pode ser reescrita quando substituída na Eq. (4.28) e então expandindo os termos, pode-se reformular a função de custo $\xi(w)$ na forma equivalente;

$$\xi(w) = \frac{1}{2} E_{\mathfrak{S}}[\epsilon^2] + \frac{1}{2} E_{\mathfrak{S}}[f(x) - F(x, \mathfrak{S})^2] + E_{\mathfrak{S}}[\epsilon(f(x) - F(x, \mathfrak{S}))] \quad (4.30)$$

Observe, entretanto, que o último termo do valor esperado no lado direito da Eq. (4.30) é zero por duas razões:

- O erro de expectativa é *não-correlacionado* com a função de regressão $f(x)$ devido à Eq.(4.40), interpretada em termos do operador $E_{\mathfrak{S}}$.
- O erro de expectativa ϵ é relativo ao modelo de regressão da Fig.(4.19), enquanto que a função aproximativa $\mathbf{F}(\mathbf{x}, \mathbf{w})$ é relativa ao modelo de rede neural da Fig. (4.20). Conseqüentemente, a Eq. (4.49) se reduz a:

$$\xi(w) = \frac{1}{2} E_{\mathfrak{S}}[\epsilon^2] + \frac{1}{2} E_{\mathfrak{S}}[f(x) - F(x, \mathfrak{S})]^2 \quad (4.31)$$

O primeiro termo no lado direito da Eq. (4.31) é a variância do erro de expectativa (*do modelo regressivo*) ϵ , calculado sobre o conjunto de treinamento \mathfrak{S} . Este termo representa o *erro intrínseco*, porque ele é *independente* do vetor de pesos \mathbf{w} . Ele pode ser ignorado, na medida em que seja considerada a minimização da função de custo $\xi(w)$ em relação a \mathbf{w} . Assim, o valor particular do vetor de pesos \mathbf{w}^* que minimiza a função de custo $\xi(w)$ também irá minimizar a média de "*ensemble*" da distância quadrática entre a função de regressão $f(x)$ e a função aproximativa $\mathbf{F}(\mathbf{x}, \mathbf{w})$. Em outras palavras, a medida natural da eficiência de $\mathbf{F}(\mathbf{x}, \mathbf{w})$ em prever a resposta desejada d é definida por:

$$L_{med} = (f(x), F(x, w)) = E_{\mathfrak{S}}[(f(x) - F(x, \mathfrak{S}))^2] \quad (4.32)$$

4.4.2 Verificação dos Termos "Bias" e Variância

Fazendo uso da Eq.(4.32), pode-se redefinir a distância quadrática entre $f(x)$ e $\mathbf{F}(\mathbf{x}, \mathbf{w})$ como:

$$L_{med} = (f(x), F(x, w)) = E_{\mathfrak{S}}[(E[D|X = x] - F(x, \mathfrak{S}))^2] \quad (4.33)$$

Esta expressão pode também ser vista como o valor médio do erro estimativo entre a função de regressão $f(x) = E[D|X=x]$ e a função aproximativa $F(x, w)$, calculada sobre toda a amostra de treinamento \mathfrak{S} . Observa-se que a média condicional $E[D|X=x]$ apresenta

um valor esperado constante em relação a todo o conjunto de dados de treinamento \mathfrak{S} . Devido este fator pode-se constatar que:

$$E[D|X = x] - F(x, \mathfrak{S}) = (E[D|X = x] - E_{\mathfrak{S}}[F(x, \mathfrak{S})]) + (E_{\mathfrak{S}}[F(x, \mathfrak{S})] - F(x, \mathfrak{S})) \quad (4.34)$$

onde simplesmente adicionamos e subtraímos a média $E_{\mathfrak{S}}[F(x, \mathfrak{S})]$. A Equação (4.34) pode ser reformulada como soma de dois termos:

$$L_{med} = (f(x), F(x, \mathfrak{S})) = B^2(w) + V(w) \quad (4.35)$$

onde $B(w)$ e $V(w)$ são, definidos da seguinte forma:

$$B(w) = E_{\mathfrak{S}}[F(x, \mathfrak{S})] - E[D|X = x] \quad (4.36)$$

sendo:

$B(w)$ o termo representativo do "bias" do valor médio da função aproximativa $F(x, \mathfrak{S})$, medido em relação à função de regressão $f(x) = E[D|X=x]$. O fator mais importante referente a este termo é que o mesmo representa a incapacidade da rede neural definida pela função $F(x, w)$ de aproximar com precisão a função de regressão $f(x) = E[D|X=x]$. Portanto, pode-se analisar o bias $B(w)$ como um erro aproximativo.

$$V(w) = E_{\mathfrak{S}}[(F(x, \mathfrak{S}) - E_{\mathfrak{S}}[F(x, \mathfrak{S})])^2] \quad (4.37)$$

$V(w)$ é o termo representativo da variância da função aproximativa $F(x, w)$, medida sobre toda a amostra de treinamento \mathfrak{S} . Este referido termo representa a não-adequação da informação contida na amostra de treinamento \mathfrak{S} acerca da função de regressão $f(x)$. Pode-se, portanto, avaliar a variância $V(w)$ como a manifestação de um *erro estimativo*.

4.5 Algoritmo de Levenberg Marquardt - LMA

Algoritmo Levenberg Marquardt (LM) se basea no método dos mínimos quadrados para ajustar funções com comportamento não linear, procurando encontrar o melhor ajustamento para um conjunto de informações na busca de minimizar a soma dos quadrados das diferenças entre a curva ajustada, tais diferenças são caracterizadas de resíduos. O LM é uma interpolação entre o Algoritmo de Gauss-Newton (GN) e o método do Gradiente Descendente, o LM é mais eficiente na maioria dos casos por atingir com maior rapidez a convergência (objetivo). Para um conjunto de dados cujas as variáveis dependentes e independentes (x_i, y_i) , onde o objetivo é otimizar o parâmetro β da função $f(x, \beta)$, de forma que seus resíduos (erros), $e(w)$ sejam mínimos.

$$e(w) = \frac{1}{n} \cdot \sum_{i=1}^m [(y - y_i)]^2 \quad (4.38)$$

E com base em função dos pesos, podemos reescreve-la da seguinte forma:

$$E(W) = \sum_{i=1}^m [(f(W^T, \beta)) - y_i], \quad (4.39)$$

e com a utilização da matriz J jacobiana da função f em β , cuja matriz é dada por:

$$J = \frac{\vartheta e(w)}{\vartheta w}, \quad (4.40)$$

utilizada para se obter a aproximação da matriz H hessiana :

$$H = \frac{\vartheta^2 E_R(W)}{\vartheta(w)^2}, \quad (4.41)$$

Observe que agora para a atualização de uma camada de pesos W_{k+1} na rede:

$$W_{k+1} = W_k + \delta_{wk}, \quad (4.42)$$

com:

$$\delta_{wk} = -H^{-1} \cdot \rho_{(k)} \quad (4.43)$$

onde, $\rho_{(k)}$ é definido como:

$$\rho_{(k)} = 2J^T(w).E(w) \quad (4.44)$$

E com as respectivas atualizações dos pesos no processo de retropropagação da rede, tem-se:

$$W_{k+1} = W_k - [J^T(w).J(w) + \lambda I]^{-1}.J^T(w).E(w) \quad (4.45)$$

sendo J^T o jacobiano de $\rho_{(k)}$, gerada a partir da primeira derivada de $\rho_{(k)}$, I a matriz identidade e λ é a constante do método de Levenberg-Marquardt. Ou ainda:

$$\rho_{(k)} = -[J_k^T \cdot \rho_{(W_k)}].[J_k^T J_k + \lambda I]^{-1} \quad (4.46)$$

4.6 Normalização dos Dados de Entrada

Segundo Palit e Popovic (2005), o processo de normalização de dados é um processo de preparação final dos dados antes que os mesmos sejam apresentados ao processo de treinamento da rede neural. Ele inclui a normalização dos dados processados e de seus intervalos naturais, a fim de que, os dados estejam normalizados de forma adequada a satisfazer as exigências da camada de entrada da rede e são adaptados à não linearidade dos neurônios, de modo que seus resultados não devam ultrapassar o limite de saturação.

4.6.1 Normalização por Mínimo e Máximo

A normalização obtida por mínimo e máximo são de grande utilidade para normalizar dados de uma forma não linear, utilizando a função tangente hiperbólica descrita no espaço $(-1,1)$ ou no caso da função Tansig descrita no espaço $(0,1)$. Na prática, a normalização mais simples é dada por:

$$x_{ni} = \frac{x_i}{x_{max}} \quad (4.47)$$

E a normalização linear, se dá da seguinte forma:

$$x_{ni} = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (4.48)$$

4.6.2 Normalização por Desvio Padrão

Widrow (1962), demonstrou que se os padrões de treino apresentarem as características baseadas em uma distribuição normal padrão, ou seja, apresentarem média zero e variância unitária $x'_k \sim N(0, 1)$. Portanto, cada entrada x_k é transformada em x'_k por:

$$x'_k = \frac{x_k - \bar{x}}{\sigma_x} \quad (4.49)$$

onde \bar{x} e σ representam respectivamente o valor médio das entradas x_k e o desvio padrão associado ; e k é o índice de padrão.

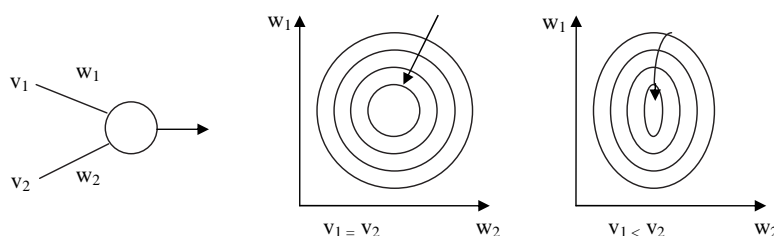


Figura 4.17 Efeito da normalização dos dados de entrada

Através da Figura 4.17, nota-se que quando as variáveis são $v_1 = v_2$, existe uma constância na direção do gradiente, enquanto que, para um comportamento $v_1 \neq v_2$, esta direção é alterada em cada passo.

Estas normalizações são as mais freqüentemente utilizadas. Além da utilização da normalização **linear** e **logarítmica** para um melhor dimensionamento dos sinais de entrada da rede, outros tipos de normalização são utilizadas para moderar possíveis problemas de não linearidade durante a fase de treinamento, são elas: **transformações logarítmicas** utilizadas para comprimir as regiões onde são observadas grandes variações nos valores observados, e **transformações exponenciais**, a qual pode ampliar a escala na região de pequenos valores observados, etc. No entanto, existe falhas a serem encontradas neste processo, p.e, pode-se na preparação destes dados se ter uma eventual perda de informações nos novos dados adquiridos.

4.7 Desenvolvimento e Aplicações de Redes Neurais

Para Kovács (1998), quando se utiliza redes neurais no estudo dos mais variados tipos de problemas, deve-se sempre seguir alguns passos para o desenvolvimento de aplicações dessas redes.

- Coleta de Dados
- Separação em Conjuntos
- Configuração de Rede
- Treinamento
- Teste
- Integração

1. *Coleta de Dados e Separação em Conjuntos*

Esta tarefa requer uma análise cuidadosa sobre o problema em si, para haver uma minimização de ambiguidades e erros nos dados. Outro fator muitíssimo importante é que os dados coletados devem apresentar de forma bastante significativa e cobrir amplamente o domínio do problema; não devendo cobrir somente as operações normais ou rotineiras, mas também as exceções e as condições nos limites do domínio do problema.

Normalmente, os dados coletados são separados em duas categorias : ***Dados de Treinamento***, que serão utilizados para o treinamento da rede e ***Dados de Teste***, os quais serão utilizados para verificar a robustez sob a condição real de utilização da rede. Além do mais, pode-se usar também uma subdivisão do conjunto de treinamento, criando um conjunto de validação. Este, será utilizado para verificar a eficiência da rede quanto a sua capacidade de generalização durante o treinamento, podendo ser empregado como critério de parada do treinamento.

Após a ocorrência do treinamento desses conjuntos, eles geralmente são dispostos em ordem de forma aleatória para prevenção de tendências associadas à ordem de apresentação dos dados. Em muitas situações pode ser necessário pré-processar estes dados, através de normalizações, escalonamento e conversões de formato para torná-los mais apropriados à sua utilização na rede.

2. *Configuração de Rede*

Este passo é caracterizado pela definição da configuração da rede, encontra-se dividido em três etapas:

- Seleção do paradigma neural apropriado à aplicação
- Determinação da topologia da rede a ser utilizada, assim como, o número de camadas, o número de unidades em cada camada, etc.
- Determinação de parâmetros do algoritmo de treinamento e funções de ativação. Esta etapa apresenta um grande impacto na eficácia do sistema resultante. Além dessas etapas não se deve deixar de resaltar que as definições de configuração de redes neurais exige muito da capacidade e experiências anteriores dos seus implementadores.

3. *Treinamento*

Após a escolha do algoritmo de treinamento, deve-se ajustar os pesos das conexões. Observa-se que, nesta etapa é muito importante considerar alguns aspectos tais como: **a inicialização da rede, o modo adotado para o treinamento e o tempo de treinamento**. Esta etapa é considerada uma das mais importantes, pois uma excelente escolha dos valores iniciais dos pesos da rede pode ocasionar uma expressiva diminuição do tempo de treinamento. Esses pesos iniciais apresentam-se na forma de uma distribuição uniforme, ou seja, os mesmos tem características de números aleatórios uniformemente distribuídos. Quando refere-se ao modo de treinamento o mais adotado é o modo padrão, pois este apresenta um menor armazenamento de dados, além de apresentar menos problemas de mínimos locais, motivado pela utilização de processos estocásticos.

O outro método utilizado é o de **Batch**, que também apresenta uma grande eficiência das estimativas do vetor gradiente, tornando o treinamento mais estável. A escolha do modo a ser utilizado dependerá em muito do problema que está se tratando. O treinamento deve ser interrompido sempre que a rede encontrar-se com uma boa capacidade de generalização e quando a taxa de erro for bastante pequena, encontrando assim, um ponto ótimo de parada com erro mínimo e capacidade de generalização máxima.

4. *Teste ou Validação*

Nesta tarefa o pesquisador fará testes para verificação da confiabilidade e adequação da rede aos dados que foram anteriormente utilizados, verificando assim, a capacidade da rede trabalhar em situações reais do próprio problema que está sendo verificado. Alguns outros testes podem ser adotados para a verificação da existência de valores muito pequenos, visto que, caso estes valores existam, as conexões associadas podem ser consideradas insignificantes e assim ser eliminadas (**prunning**), e por outro lado, caso haja, valores muito grandes que os demais é um indicativo de que houve um **over-training** da rede. No entanto, é importante ressaltar que testes utilizados em modelos lineares, como por exemplo, em modelos de Box e Jenkins, onde utiliza-se testes de Ljung-Box-Pierce para validação e escolha do melhor modelo entre os ajustados, o mesmo não é válido quando se trata de modelos não lineares (LIMÃO, 1999).

5. Integração

Já com a rede treinada e avaliada, pode-se integrá-la em um sistema do ambiente operacional da aplicação para melhor facilidade de aquisições de dados. Esta integração deve, de preferência, ser aplicada em sistemas que apresentem facilidade de utilização, não esquecendo de realizar um processo de treinamento dos usuários para que assim, haja um sucesso imediato do processo.

Uma estrutura significativa de redes neurais apresenta uma enorme capacidade de adaptação e aprendizagem. Esta capacidade de adaptação e aprendizagem pelo ambiente significa que modelos de redes neurais podem lidar com dados imprecisos e situações não totalmente definidas, ou seja, caso seja apresentado a determinada rede entradas que ainda não lhe foram apresentadas, ela tem a capacidade de generalizar com os dados existentes. As Redes Neurais apresentam uma grande capacidade de aproximar qualquer função contínua do tipo não linear de um grau de correção desejado.

4.8 Redes Neurais em Sistemas de Controle

Segundo Kovács (2002), os sistemas de controle são baseados em estruturas físicas claramente definidas, e por isso, constituem uma área que permite uma aplicação eficiente das técnicas de redes neurais. Os sistemas de controle podem ser distinguidos em duas categorias distintas: os *sistemas de controle dinâmicos* e os *sistemas de controle de processos* que, embora exista uma relação entre ambos, podem ser aplicados em classes diferentes de problemas e com abordagens próprias.

Os sistemas de controle dinâmicos sempre envolvem o controle de uma *planta* cujas variáveis evoluem de acordo com um sistema de equações diferenciais especificadas pelas leis físicas que as governam. Os objetivos do controle dinâmico neste caso, é disciplinar a evolução destas variáveis, conforme critérios previamente descritos, ou seja, propor estabilização ou, até mesmo, impor trajetória de referência que expliquem o mesmo.

4.8.1 Sistemas de Controle Dinâmicos

Com estudos no decorrer dos últimos anos, os sistemas de controle dinâmicos, particularmente os que envolvem modelos lineares, estão sendo aplicados em vários conjuntos de problemas, assim como, as suas respectivas análises, (KOVÁCS, 2002).

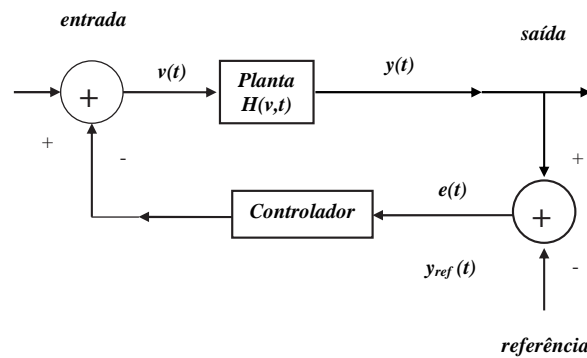


Figura 4.18 Esquema de um sistema de controle dinâmico

A representação do sistema de controle dinâmico, descrita na Figura 4.18, é definida da seguinte forma:

- $u(t) \rightarrow$ caracteriza a excitação, ou entrada do sistema, a qual é composta pelos sinais recebidos pela planta;
- $y(t) \rightarrow$ caracteriza a saída, caracterizada pela resposta, ou projeção que serve de base para a comparação com a trajetória de referência;
- $e(t) \rightarrow$ é o erro produzido pelo afastamento da saída em relação a trajetória de referência. Este se caracteriza por definir como uma importante medida utilizada pelo controlador para determinar uma ação de controle o mais eficiente possível que, ao ser aplicada à entrada da planta, produza um menor erro possível, ou até mesmo o elimine.
- $H(v,t) \rightarrow$ operador ou função de transferência da planta, responsável por sintetizar a relação entre a entrada e a saída.

Observa-se que, uma vez especificado o objetivo de controle e caracterizada a dinâmica da planta, considerando-a na forma de equações diferenciais ou funções de transferência, é necessário determinar um controlador que atenda as especificações prévias definidas.

É muito importante observar que a dinâmica de um determinado sistema pode ser descrita por um sistema multivariável de equações diferenciais lineares de parâmetros fixos, e para um outro sistema pode ser descrito por função de transferência linear variante no tempo. Consequentemente, as abordagens adotadas para o projeto de controladores em ambos os casos são completamente diferentes.

4.8.2 Identificação de Sistemas Dinâmicos

Levando em consideração uma planta de tempo discreto que apresente como entrada e saída as seguintes sequências $u(k)$ e $y(k)$ com dinâmica desconhecida. E, assumindo como primeiro passo a definição de uma classe parametrizada $P(W)$ de modelos de identificação e, dentro desta classe mediante algum algoritmo apropriado, procurar identificar a parametrização que de forma mais adequada represente a dinâmica da planta (Fig 4.19).

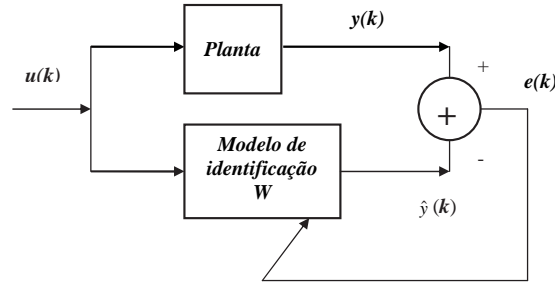


Figura 4.19 Identificação de um sistema dinâmico

A medida de adequação que melhor ajuste o sistema é definida através da minimização de erro $e(k)$ de reconstrução da saída, também utilizado para reajustar os parâmetros do modelo de identificação.

$$\|y(k) - \hat{y}(k)\| = \|y(k) - P(\hat{W}; u(k))\| < \varepsilon \quad (4.50)$$

para $k \geq K$ e para todo $u(k) \in U$.

4.8.3 Retropropagação Dinâmica (Dynamic Back Propagation)

Como já anteriormente definido na equação 4.21, em que o algoritmo de retropropagação estática, cujo resultado de um método do gradiente para minimizar uma função de erro quadrático pelo qual um parâmetro arbitrário w é ajustado, apresenta a seguinte característica.

$$W(K+1) = W(k) - \eta \text{grad}(E(W(k))) \quad (4.51)$$

onde k é o índice dos exemplos da tabela de treinamento.

A partir de então, considera-se um modelo de identificação não linear parametrizado por $W = [w_1, w_2, \dots, w_n]$:

$$x(K+1) = \phi(x(k), u(k), W) \quad (4.52)$$

$$\hat{y}(k) = \psi(x(k))$$

onde o erro quadrático $E(W)$ é definido como:

$$E(W) = \|y(k) - \hat{y}(k)\| \quad (4.53)$$

e como gradiente do erro $E(W)$ em relacionado ao parâmentro w_i , implica em:

$$\frac{\delta E(W)}{\delta w_i} = -2(y(k) - \hat{y}(k)) \cdot \frac{\delta \hat{y}(k)}{\delta w_i} = -2\partial_y(k) \cdot \hat{y}_{wi}(k) \quad (4.54)$$

4.8.4 Sistemas de Controle de Processos

Com utilização nas mais variadas especialidades, entre elas a Engenharia Elétrica, Engenharia Mecânica, Biomedicina, Engenharia de Trafêgo, os sistemas de controle de processos em qualquer situação apresentam funções essenciais voltadas ao processo em si: a temporização, o sequenciamento, a sincronização e o intertravamento de eventos. E os voltados especialmente ao operador: monitoração, supervisão, detecção de falhas e processamento de alarme.

Diagnóstico de falhas na operação de sistemas pode ser reduzido a um problema de detecção de padrões, no caso padrões anômalos, e já se sabe que as redes neurais são eficientes detetores de padrões (KOVÁCS, 2002).

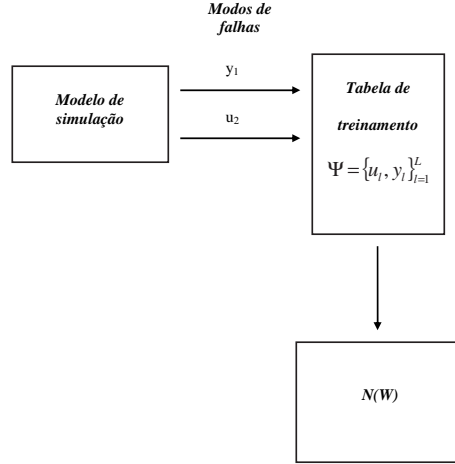


Figura 4.20 *Treinamento de uma rede neural para diagnóstico de falhas*

4.8.5 Análise de Dados

Embora na função objetivo do modelo de Redes Neurais Artificiais (RNA) tenha sido adotado a função de minimização do erro quadrático médio, o desempenho dos modelos de previsão será avaliado através de um conjunto maior de indicadores estatísticos. Assim, além do erro quadrático médio (**EQM**), serão calculados e analisados o erro absoluto médio (**EAM**), erro relativo percentual médio (**ERPM**) o erro relativo percentual máximo (**ERPmax**). Todos estes testes de avaliação dos erros para projeções futuras são descritos abaixo:

$$EQM = \frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2 \quad (4.55)$$

$$EAM = \frac{1}{n} \sum_{j=1}^n |y_j - \hat{y}_j| \quad (4.56)$$

$$ERPM = \frac{100}{n} \sum_{j=1}^n \frac{|y_j - \hat{y}_j|}{y_j} \quad (4.57)$$

$$ERP_{max} = \max \cdot 100 \left| \frac{y_j - \hat{y}_j}{y_j} \right| \quad (4.58)$$

onde: $n = 1, \dots, N$, com N o número de padrões analisados; $Y_j =$ são caracterizados como os valores reais observados; \hat{y}_j são caracterizados como os valores previstos, ou seja, valores obtidos após modelagem, também conhecidos como previsões, ver Figura (4.21).

4.8.6 Previsões com Redes Neurais

Devido a grande complexidade dinâmica de determinados processos ou eventos não lineares no decorrer de um intervalo de tempo t , torna-se muito difícil a obtenção de previsões com grau de confiabilidade significativos. Com isso, em determinadas situações, os modelos baseados em inferência estatística não são adequados para este fim.

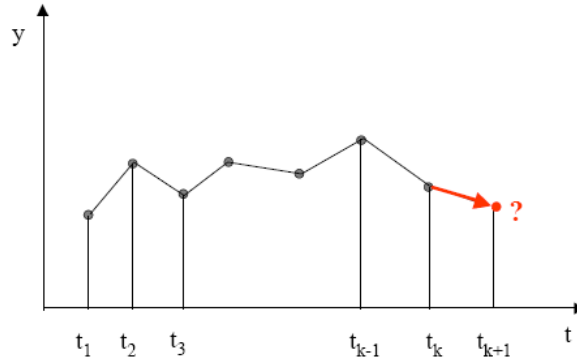


Figura 4.21 Valores Previstos

$$y(t_1), y(t_j), \dots, y(t_z) \rightarrow y(t_{k+1}) \quad (4.59)$$

$$y(t_i), y(t_{i+1}), \dots, y(t_k) \rightarrow y(t_{k+1}) \rightarrow \text{um passo a frente.} \quad (4.60)$$

$$y(t_i), y(t_j), \dots, y(t_z) \rightarrow y(t_{k+h}) \rightarrow k \text{ passo a frente.} \quad (4.61)$$

Portanto, nestes casos, modelos baseados em redes neurais suprem estas necessidades com excelentes resultados. Como um exemplo de processo que apresenta esta característica pode-se citar, o mercado financeiro onde a movimentação das ações por apresentar algumas relações não lineares, não é possível modelar utilizando modelos estruturais, regressão e ARIMA.

Capítulo 5

Análise de Dados

5.1 Processo de filtração dos cristais de hidrato

Para a produção de alumina, a bauxita após minerada é transportada para a fábrica onde sofre a intervenção das principais fases do processo de sua obtenção: a moagem, digestão, **filtração/evaporação**, precipitação e calcinação.

A bauxita após moída com solução cáustica e cal em moinhos conjugados, forma uma polpa de cristais de hidrato. A polpa em solução pré-aquecida de soda cáustica é transferida aos **digestores**, onde é realizado a primeira separação de impurezas como o **Ferro** e o **Sílica** pois, estes são fatores que são prejudiciais à qualidade final do alumínio. A partir de então, o hidrato de alumina é dissolvido, formando o licor de aluminato de sódio e outras impurezas. A solução que contém o aluminato de sódio e os resíduos de bauxita em suspensão é transferida aos espessadores, onde, com auxílio de agentes floclulantes, os resíduos são separados por sedimentação, formando uma lama densa, a “**lama vermelha**”, que é lavada e filtrada para recuperar o residual de soda, sendo transferida com baixo teor de umidade para o depósito de rejeitos, (<http://www.serteq.com.br/>) Fig.(5.1).

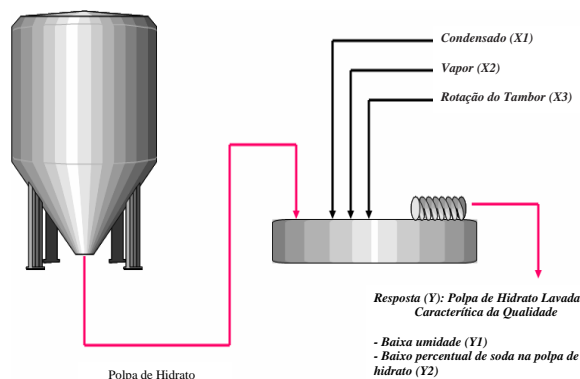
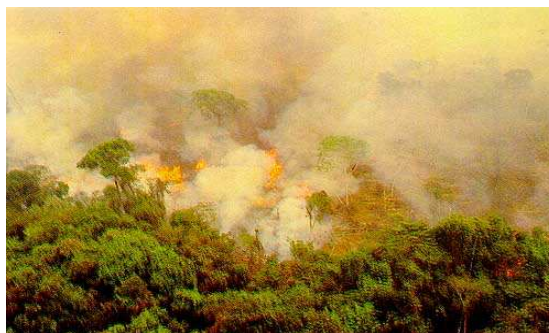


Figura 5.1 Processo de lavagem e filtração da polpa de hidrato

A polpa de cristais de hidrato é bombeada dos tanques de **precipitação** para filtros rotativos de discos ou de tambor horizontal, cuja finalidade é a máxima remoção da umidade e dos resíduos de soda presentes na torta filtrada.

No processo é utilizado **condensado**, cuja finalidade é lavar a torta formada sobre o tambor removendo resíduos de soda cáustica, e após esta lavagem com a ajuda de **calcinadores**, movidos a queima de óleo **BTF** (Baixo Teor de Fluidéz). O óleo **BTF** tem como função produzir calor e é utilizado em equipamentos destinados à geração de energia térmica. Os calcinadores trabalham em temperaturas equivalentes a 600 e 700 graus centígrados. É aplicado um **vapor** que tem como finalidade a aceleração da secagem da torta, removendo assim os vestígios da água não combinada (livre) molecularmente com a alumina. Deste modo é obtido o primeiro produto do processo de produção de alumínio: A **alumina** com características de um pó branco e refinado de consistência semelhante ao açúcar, (SANTOS, 2007 e <http://www.abralatas.org.br>).

Os altos números de ocorrências de queimadas e a utilização pela indústria de produtos derivados de petróleo como combustível, entre eles o óleo **BTF**, são os grandes responsáveis pelos elevados índices de emissões de gases na atmosfera contribuindo com o aumento do efeito estufa e ocorrência de chuvas ácidas. Pois com a combustão do óleo **BTF**, devido a alta concentração de enxofre em sua composição, são liberados na atmosfera dióxidos de enxofre (SO_2) e ionizado na atmosfera forma o (SO_4^{-2}). Portanto, na atmosfera estes dióxidos reagem com as gotículas de água em suspensão resultando em H_2SO_4 , ácido sulfúrico, provocando por consequência chuva ácida, (www.sbirt.ibict.br). Deste modo, a queima excessiva de óleo **BTF**, representa uma ameaça ao meio ambiente e consequentemente à saúde da população.

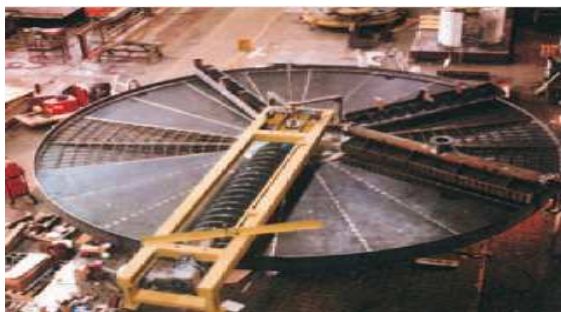
Figura 5.2 *Emissões de Gases por Indústrias*Figura 5.3 *Emissões de Gases por Queimadas*

Responsáveis pela emissões no período de 1971 e 2004, de cerca de 36% do principal gás causador do efeito estufa o CO_2 , as indústrias mundiais dentre elas a de produção e beneficiamento de alumínio, apresentou um aumento de 61% no consumo de energia, principalmente em alguns países emergentes importantes, o que aumenta ainda mais a quantidade de CO_2 lançada à atmosfera, (<http://noticias.terra.com.br/>).

5.1.1 Filtros de discos rotativos

Um fator de grande importância na cadeia produtiva do alumínio é a energia elétrica onde somente no Brasil, segundo Dr. Célio Berman 50% de toda energia produzida é consumida por indústrias, sendo que 30% se restringe a seis setores: Cimento, Aço, Alumínio, Ferro-Ligas, Petroquímica e Papel e Celulose. Onde 8% particularmente é consumida na indústria de produção de alumínio em movimentação de filtros rotativos, fornos etc. (www.socioambiental.org)

Os filtros de discos rotativos (Fig 5.4), pertencem ao grupo com alimentação lateral e já estão disponíveis há muitos anos. São usados geralmente em aplicações resistentes tais como enxugamento do taconito do minério de ferro, da hematita, do hidrato de carvão, **de alumínio**, do concentrado de cobre, dos concentrados da flutuação de pirita e de outros processos de beneficiamento. Pode-se destacar que uma das maiores finalidades dos filtros rotativos no processo é a de filtração de impurezas e seleção de materias.

Figura 5.4 *Filtro de disco rotativo*Figura 5.5 *Substituição do Tecido Filtrante*

5.1.2 Critérios de Eventuais Paradas no Processo

No decorrer da filtração da polpa de cristais de hidrato, alguns fatores contribuem para a ocorrência de eventuais e indispensáveis interrupções no processo. Estes fatores influenciam diretamente na qualidade do produto e conseqüentemente na não otimização dos resultados. Dentre estes fatores, destaca-se a substituição do **tecido filtrante** (Fig 5.5) dos filtros rotativos. Esta substituição é realizada a cada 1.500h de trabalho, e a finalidade é evitar a não eficiência do enxugamento da polpa e a não retenção correta das impurezas. Para uma eficiente lavagem da polpa é necessário a realização da **limpeza cáustica**, intervenção necessária à cada 200 horas ininterruptas do processo.

5.1.3 Eficiência do Processo

Para a máxima eficiência no processo de lavagem e secagem da polpa de cristais de hidrato é extremamente necessário que as variáveis vapor e condensado, sejam combinadas e controladas com a rotação do tambor, proporcionando assim, o tempo de residência adequado para que a lavagem e a secagem possam agir na máxima remoção da água livre e da soda residual.

5.2 Análise dos Dados Através de Modelo de Rede Neural

Para o desenvolvimento do ajuste e controle do processo por sistema inteligente de redes neurais artificiais, são utilizadas as informações coletadas a cada 8 horas do processo ininterrupto. No entanto, não existem as informações ocorridas nos finais de semana e, em determinadas situações as amostras das variáveis foram retiradas em apenas um horário ou em apenas uma semana de eventuais meses. As variáveis utilizadas foram: Condensado, rotação do tambor e vapor. No entanto, essas informações de coleta de dados são consideradas irrelevantes quando se trabalha com modelagens via sistemas de aprendizagem inteligentes. Pois para o treinamento da rede, as informações devem ser selecionadas de forma que se torne a mais representativa possível do evento a ser estudado, dispensando assim, a ordem de ocorrência das observações em cada variável.

As amostras são referentes ao período de 12 de julho de 2005 a 22 de junho de 2006, correspondendo um total de 498 informações. No entanto, na fase de treinamento e validação da rede foram utilizados somente um conjunto de 350 observações, das quais 57% foram para treinamento o correspondente a 200 observações e para validação utilizou-se o correspondente a 43% dessa amostra o que equivale a 150 observações (Fig 5.6).

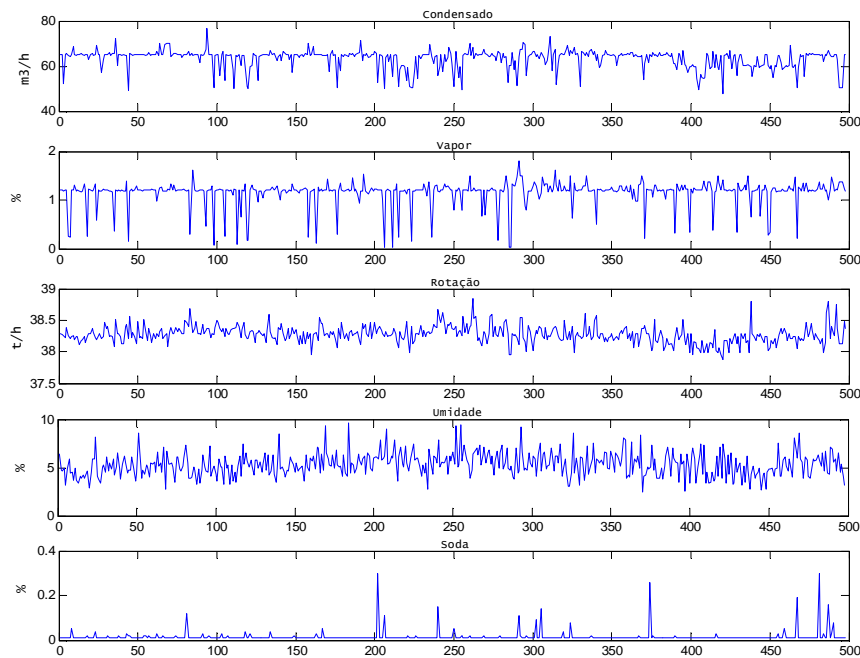


Figura 5.6 Comportamentos das Variáveis do processo

Na busca de uma modelagem de aproximação, ou seja, de uma modelagem que caracterize da forma mais significativa possível o processo estudado, que seja capaz de acompanhar o padrão gerado, realizou-se vários ensaios de simulações utilizando alguns fatores, como por exemplo: normalização dos dados de entrada da rede por Desvio Padrão e normalização no espaço $(-1, 1)$, além de alteração na estrutura da própria rede, através de mudanças nos seguintes parâmetros: número de camadas escondidas de neurônios, alterações das taxas de aprendizagem e de momento e os tipos de funções de ativação para cada camada.

No entanto, após a realização de vários ensaios, foram selecionados três situações, onde utilizou-se como método de treinamento o algoritmo alternativo ao Backpropagation padrão porém bastante eficiente, ou seja, uma variação do algoritmo Backpropagation com o método ***Levenberg Marquardt***, também conhecido como ***Levenberg Marquardt Backpropagation***. Na escolha deste algoritmo levou-se em consideração o fato, do mesmo ser considerado como um dos mais rápidos e eficiente para se obter a convergência em treinamento (Hagan e Menhaj, 1994).

5.3 Estruturação e Treinamento da Rede - 1º Ensaio

Neste primeiro ensaio de simulação, a fase de treinamento foi realizada em 10.000 iterações, com as variáveis normalizadas por ***desvio padrão***, ver Função 4.39. O parâmetro de atualização de tela adotado foi de 50 em 50 épocas, com uma taxa de aprendizagem de 0,07, taxa de momento de 0,8 e um erro final de treinamento equivalente $E(W)=0$.

Topologia * da Rede:

- 1ª Camada: Camada de entrada da rede, representada pelas variáveis independentes (condensado, rotação e vapor).
- 2ª Camada: Formada por 14 neurônios com função de ativação identidade, significando neurônios estáticos e com a saída sendo representada por uma função do tipo Logística (logsig).

* Na matemática, a topologia é a área em que se estudam os espaços topológicos. Em engenharias, está associada à disposição lógica de elementos. Em informática, **topologia de rede** é a forma por meio da qual ela se apresenta fisicamente, ou seja, como os elementos de rede (network nodes) estão dispostos. Obtido em "<http://pt.wikipedia.org/wiki/Topologia>"

- 3ª Camada: Formada por 12 neurônios com função de ativação identidade, significando neurônios estáticos e com a saída sendo representada por uma função do tipo Logística (logsig).
- 4ª Camada: Camada de saída da rede, formada por apenas dois neurônio, cuja entrada apresenta uma função de ativação identidade, com neurônios estáticos e com as saídas representadas por uma função do tipo linear (purelin).

Com esta configuração de rede foi obtido o respectivo erro de treinamento $MSE = 2 \times 10^{-2}$, Figura 5.7.

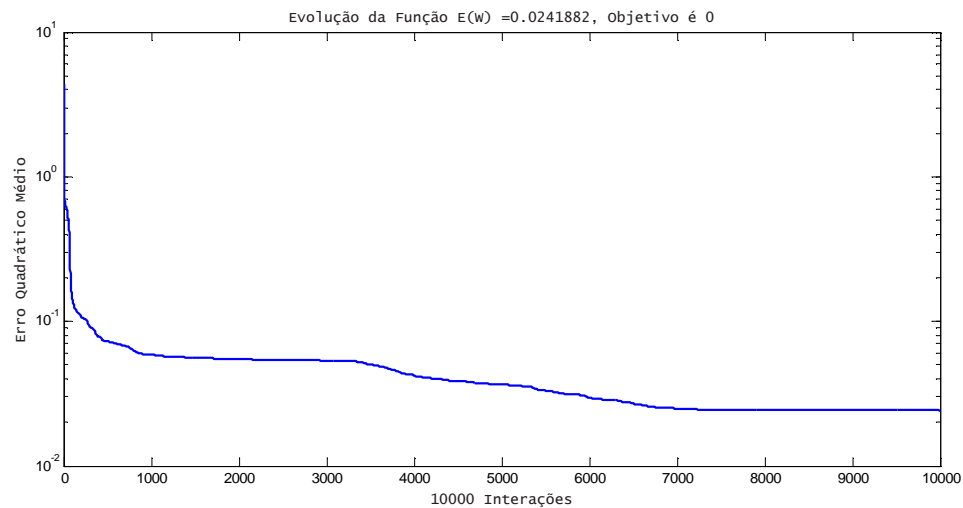
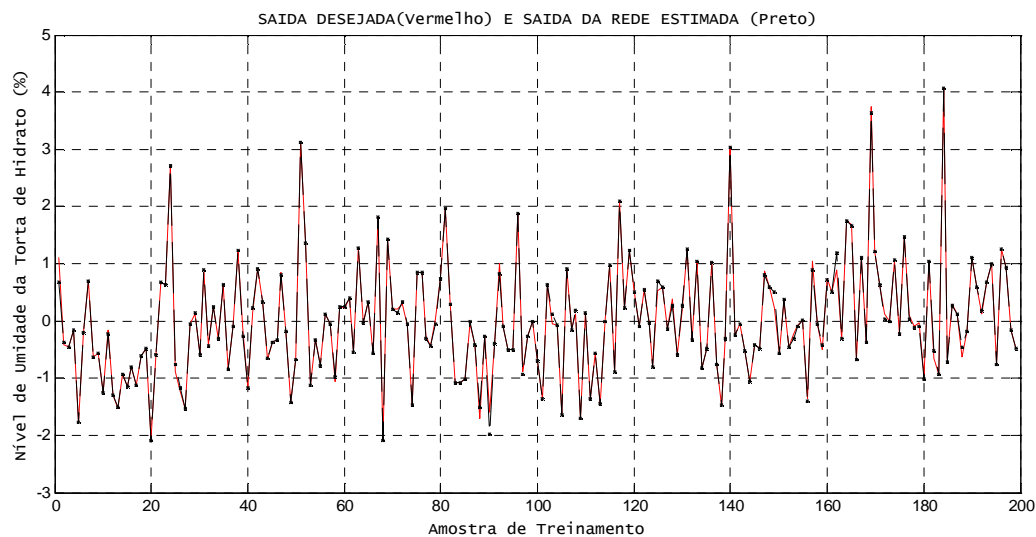
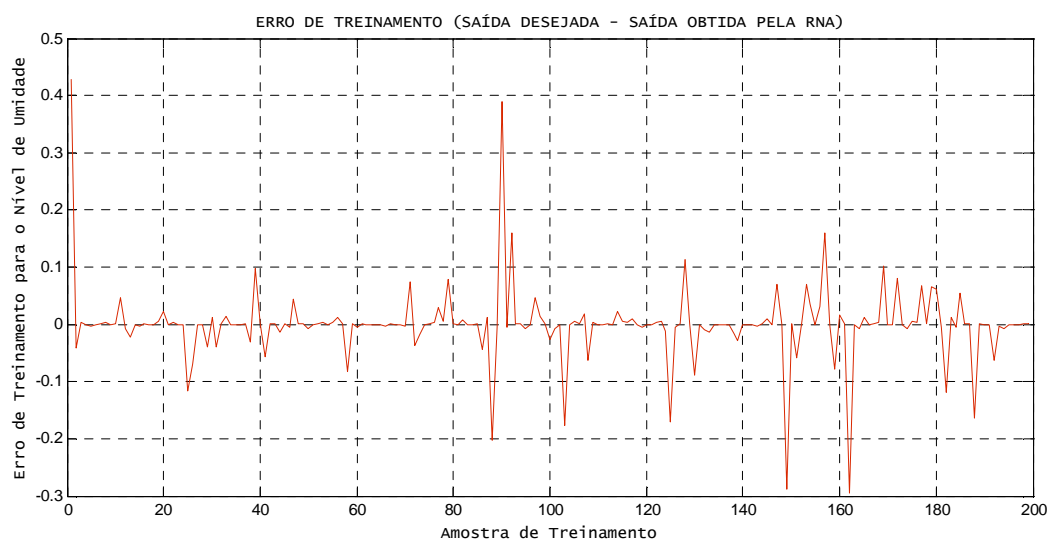


Figura 5.7 *Evolução da Função Erro $E(W)$*

5.3.1 Simulação de Treinamento da Variável Nível de Umidade- 1º Ensaio

Figura 5.8 *Treinamento da Variável Umidade*Figura 5.9 *Erro de Treinamento da Variável Umidade*

Pode-se observar na Figura 5.8, que o ajuste de treinamento para a variável nível de umidade, obtido através da simulação do modelo de RNA com apenas duas camadas foi de razoável significância, pois um grande número de observações não acompanhou o padrão do processo. Esta afirmação pode ser comprovada através das Figura 5.9, gerada a partir dos erros quadráticos médios de treinamento para a variável dependente em questão.

5.3.2 Simulação de Treinamento da Variável Nível de Soda- 1º Ensaio

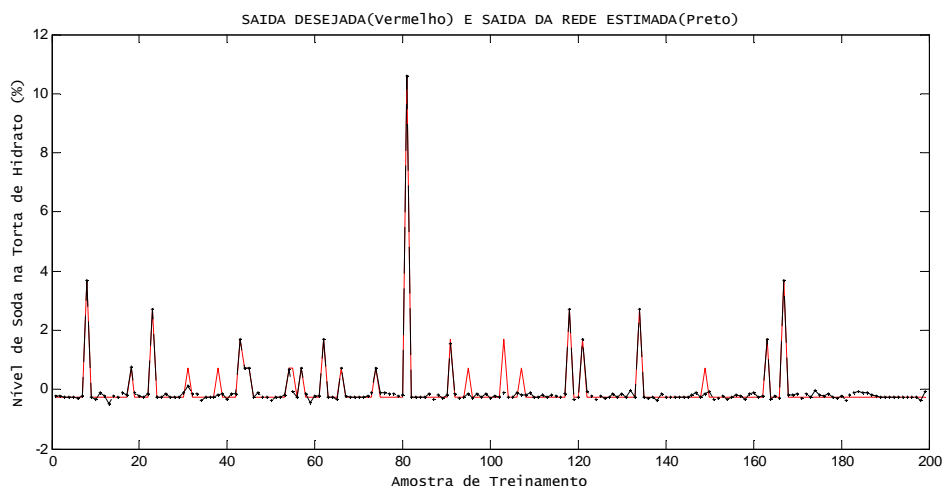


Figura 5.10 Treinamento da Variável Soda

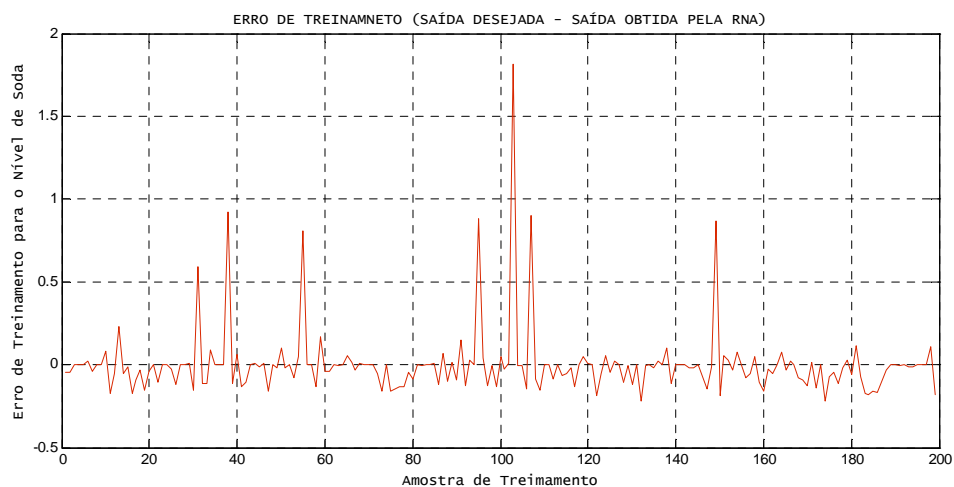


Figura 5.11 Erro de Treinamento da Variável Soda

Pode-se observar através da Figura 5.10, que o ajuste de treinamento para a variável nível de Soda presente na Polpa de cristais de hidrato, semelhante ao treinamento da umidade, também se caracterizou como de razoável significância, devido uma grande parte das observações não ter acompanhado o padrão do processo. Esta afirmação pode ser comprovada através da Figura 5.11, correspondente aos erros quadráticos médios de treinamento desta variável.

5.3.3 Simulação de Validação da Variável Nível de Umidade- 1º Ensaio

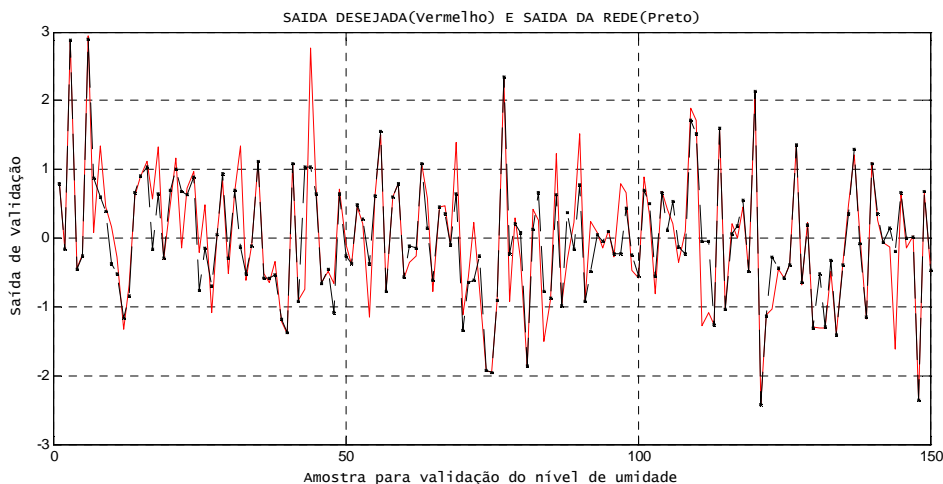


Figura 5.12 Validação da Variável Nível de Umidade

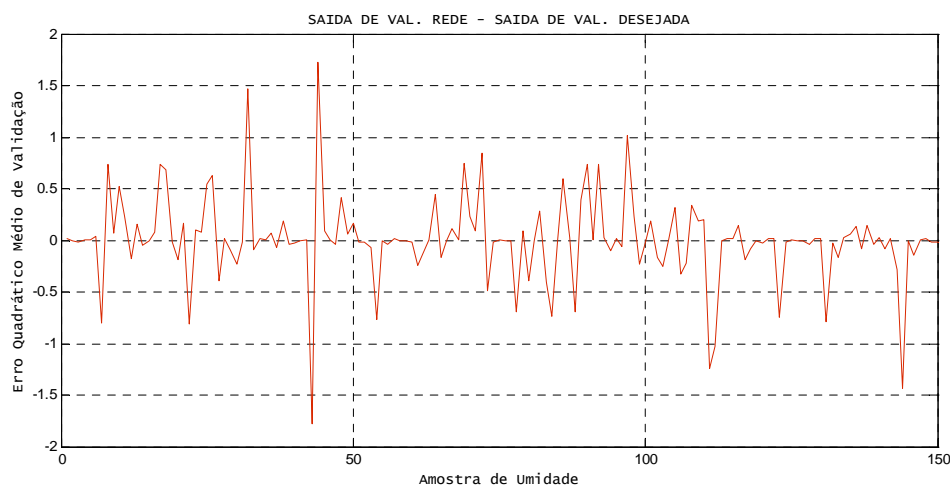


Figura 5.13 Erro de Validação da Variável Nível de Umidade

Pode-se observar através da Figura 5.12, que quando passou-se para a fase de ajuste de validação da variável nível de umidade, para o modelo de RNA com apenas duas camadas, as respostas obtidas pela rede não apresentaram respostas significativas. Pode-se observar uma grande variação entre os valores reais e os valores gerados pela rede. Para esta verificação é suficiente a análise da Figura 5.13, gerada a partir dos erros quadráticos médios $MSE = 5 \times 10^{-2}$ de validação desta variável.

5.3.4 Simulação de Validação da Variável Nível de Soda- 1º Ensaio

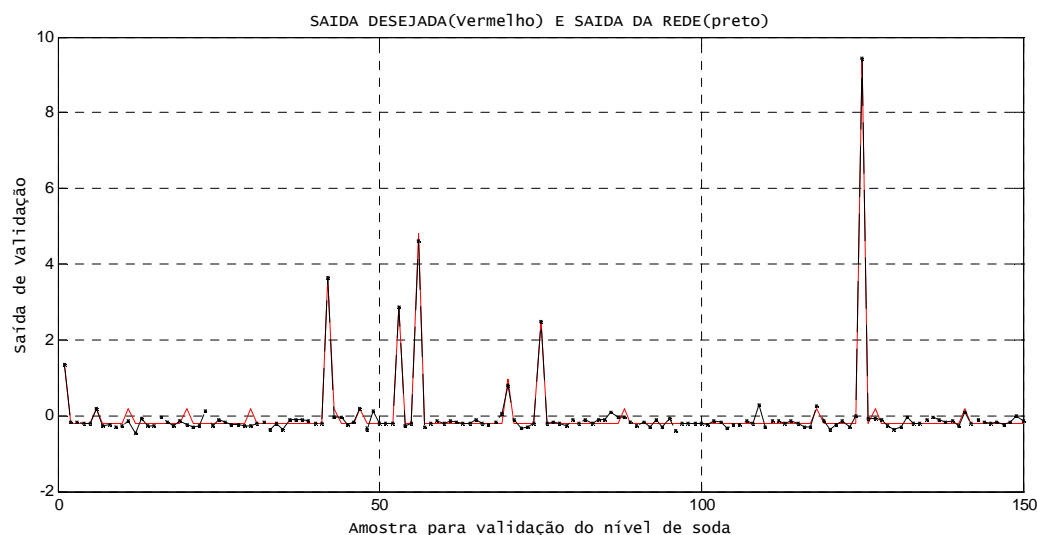


Figura 5.14 Validação da Variável Nível de Soda

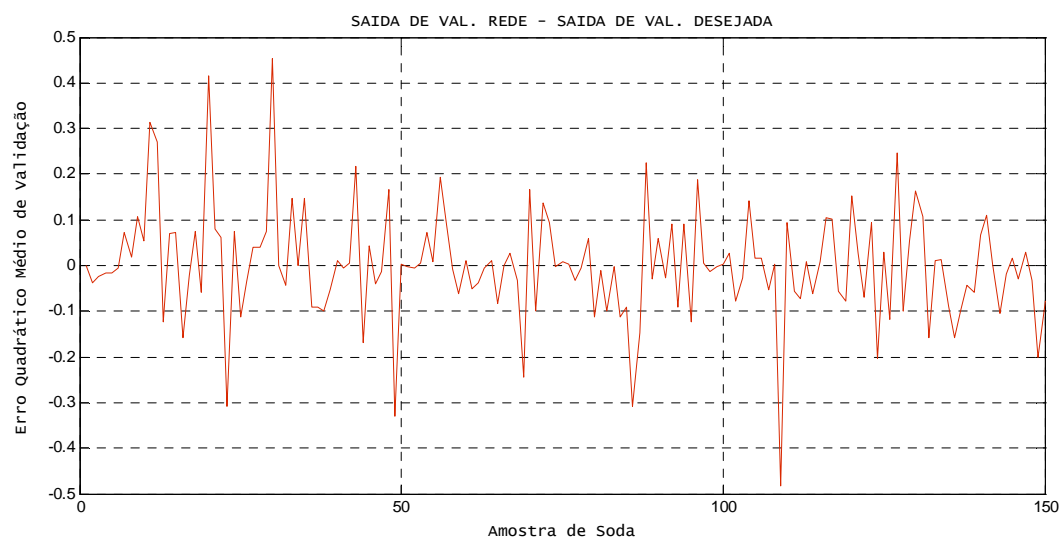


Figura 5.15 Erro de Validação da Variável Nível de Soda

Observa-se através da Figura 5.14, que o mesmo ocorreu quando se validou a variável nível de soda, ou seja, a rede não conseguiu de forma robusta, acompanhar o padrão gerado pelo processo. E, através da Figura 5.15, formada pelos erros quadráticos médios fornecidos pelo processo, verifica-se com mais clareza a grande variabilidade entre o padrão do processo real e o padrão gerado pela rede neural.

5.3.5 Simulação das Validações Desnormalizadas- 1º Ensaio

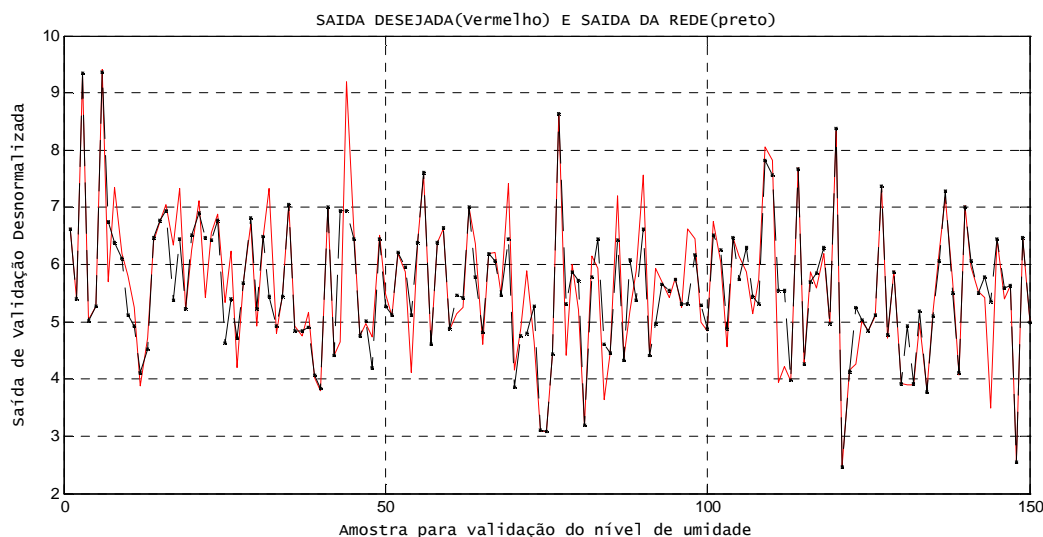


Figura 5.16 Validação Desnormalizada da Variável Umidade

Para uma melhor visualização das absrvações geradas pela rede em relação as ob-sevações originais do processo, ver Figuras 5.16 e 5.17, que foram construídas a partir do processo de desnormalização dos resultados. Através destas, pode-se verificar que a rede proposta neste ensaio não obteve resultados satisfatórios, ou seja, a mesma não conseguiu monitorar de forma desejada o padrão gerador do processo.

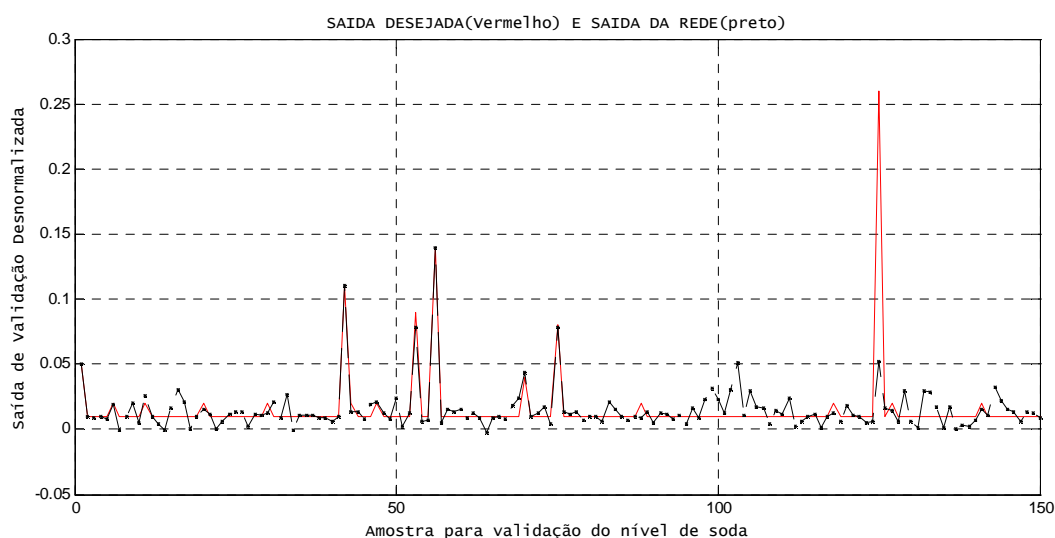


Figura 5.17 Validação Desnormalizada Variável Soda

5.4 Estruturação e Treinamento da Rede - 2º Ensaio

Na busca de melhores resultados, houve a necessidade de num segundo momento ajustar nova rede com duas camadas internas. Porém, para esta nova rede aplicou-se aos dados uma normalização no espaço $(-1, 1)$. Esta escolha se justifica pela definição da função de ativação Tangente Hiperbólica (Tansig), definida na Função 4.11. Para esta RNA convergir, foi suficiente a utilização de apenas 2 camadas de neurônios, com as seguintes configurações:

Topologia da Rede:

- 1ª Camada: Camada de entrada da rede, representada pelas variáveis independentes (condensado, rotação e vapor).
- 2ª Camada: Formada por 12 neurônios com função de ativação identidade, significando neurônios estáticos e com a saída sendo representada por uma função do tangente hiperbólica (tansig).
- 3ª Camada: Formada por 10 neurônios com função de ativação identidade, significando neurônios estáticos e com a saída sendo representada por uma função do tangente hiperbólica (tansig).
- 4ª Camada: Camada de saída da rede, formada por apenas dois neurônios, cuja entrada apresenta uma função de ativação identidade com neurônios estáticos e com as saídas representadas por duas funções do tipo linear (purelin).

5.4.1 Simulação de Treinamento da Rede - 2º Ensaio

A fase do treinamento foi realizada em 15.000 interações (= épocas), sendo utilizado um parâmetro de atualização da tela de 50 em 50 épocas. Adotou-se uma taxa de aprendizagem de 0,03, taxa de momento de 0,7 e pré determinou-se um erro final de treinamento equivalente $E(W)=10e^{-4}$.

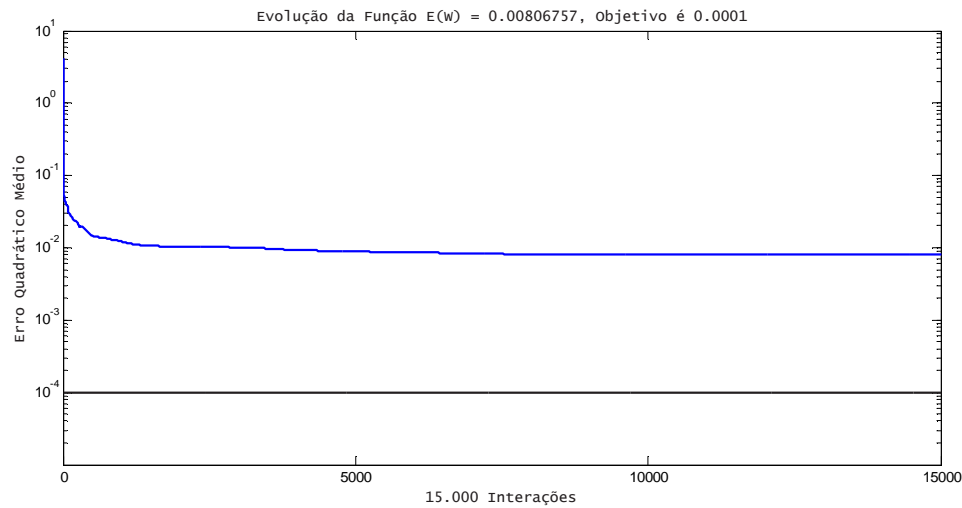
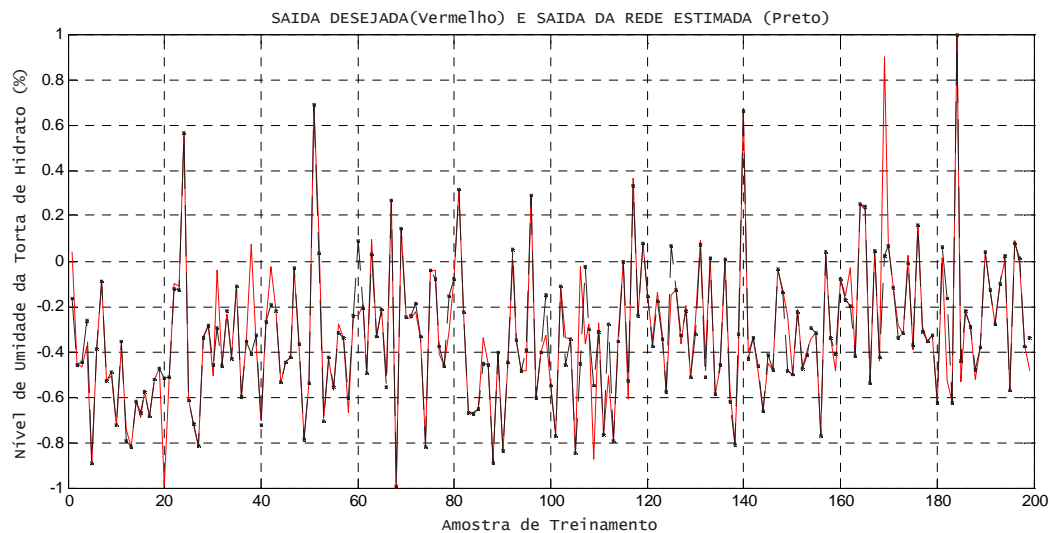
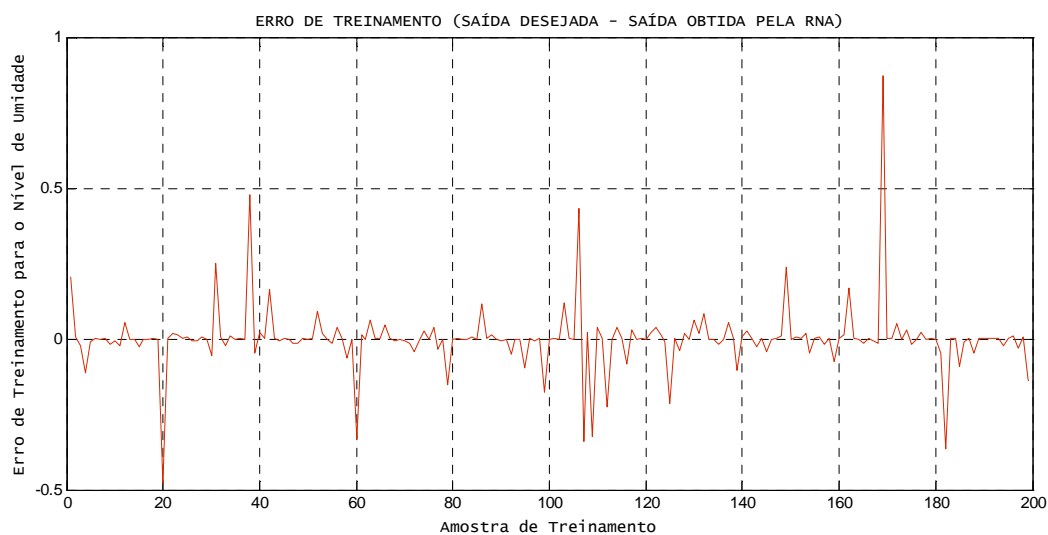


Figura 5.18 *Evolução da Função Erro $E(W)$*

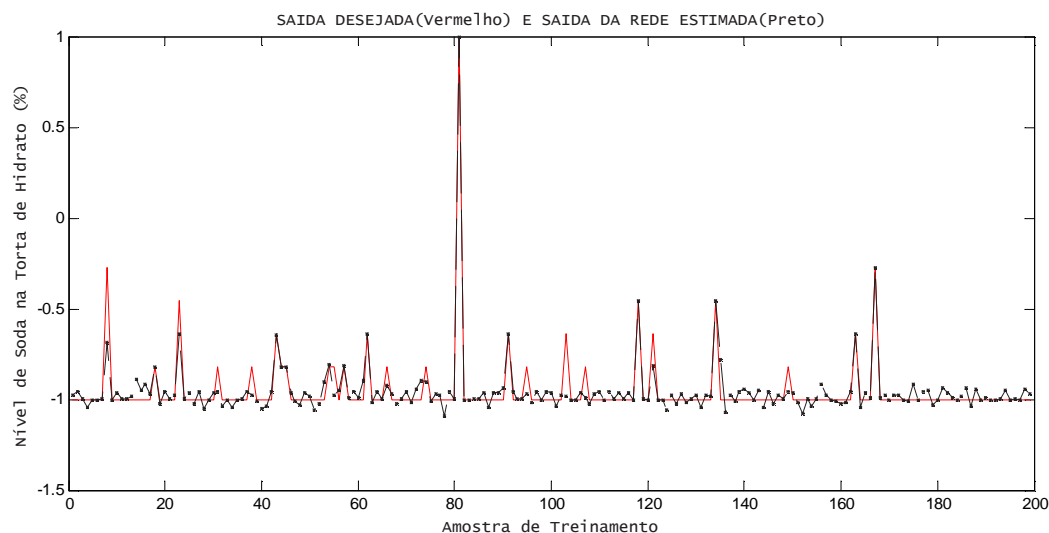
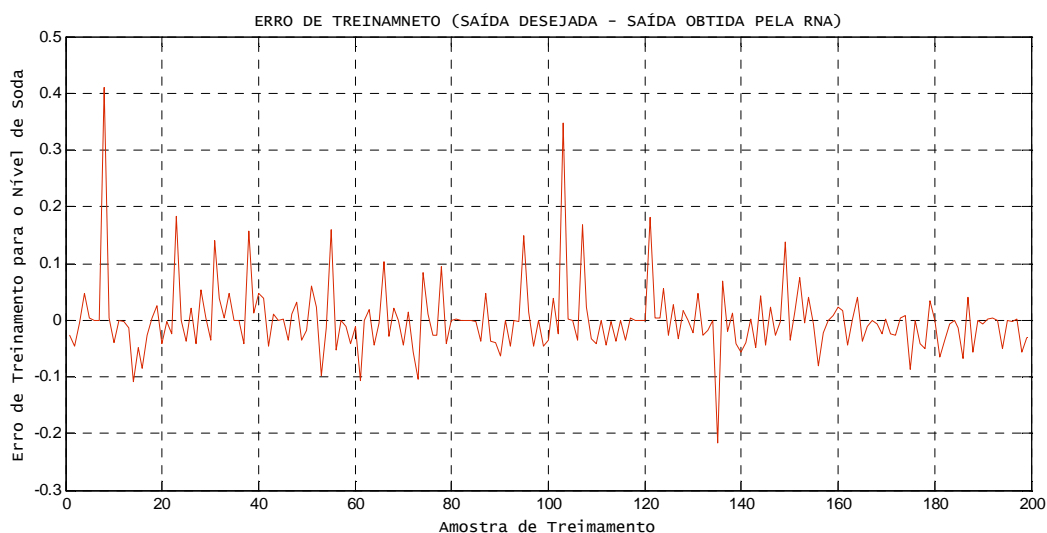
Com esta configuração de rede neste ensaio se obteve um erro de treinamento equivalente a $MSE = 8 \times 10^{-3}$, ver Figura (5.18), na qual o erro pode ser considerar bastante eficiente e um tempo de convergência significativo.

5.4.2 Simulação de Treinamento da Variável Nível de Umidade - 2º Ensaio

Figura 5.19 *Treinamento da Variável Umidade*Figura 5.20 *Erro de Treinamento da Variável Umidade*

Pode-se observar através da Figura 5.19, que o ajuste de treinamento para o modelo de RNA com apenas duas camadas, apresentou boa significância. Entretanto, algumas observações não acompanharam o padrão gerado pelo processo. Isto é confirmado através da Figura, 5.20 dos erros quadráticos médios do treinamento para esta variável.

5.4.3 Simulação de Treinamento da Variável Nível de Soda - 2º Ensaio

Figura 5.21 *Treinamento da Variável Soda*Figura 5.22 *Erro de Treinamento da Variável Soda*

A Figura 5.21, mostra que a simulação do treinamento para a variável nível de soda, através do modelo de RNA com apenas duas camadas, apresentou boa robustez. No entanto, ainda se detecta que algumas observações não acompanham o padrão do processo, isto é confirmado através da Figura 5.22 dos erros quadráticos médios de treinamento para a variável em análise.

5.4.4 Simulação de Validação da Variável Nível de Umidade - 2º Ensaio

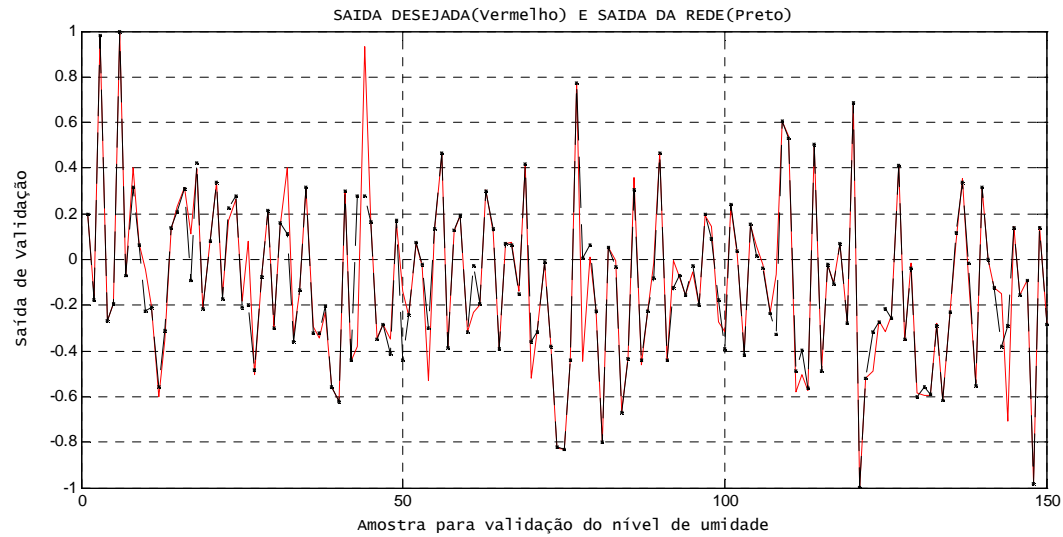


Figura 5.23 Validação da Variável Nível de Umidade

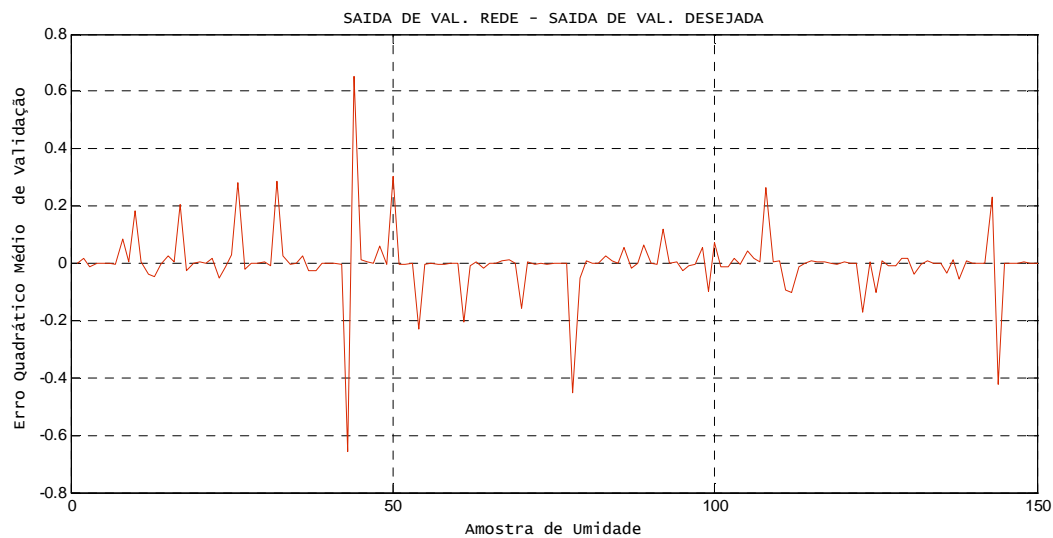


Figura 5.24 Erro de Validação da Variável Nível de Umidade

Ao analisar a Figura 5.23, observa-se que o ajuste de validação para o modelo da RNA com apenas duas camadas para a variável níveis de umidade, apresentou significativa melhora. No entanto, os resultados obtidos não acompanharam em maioria o padrão gerado pelo processo. Isto é confirmado através das Figura 5.24, gerada pelos erros quadráticos médios de treinamento, onde se observou um erro de validação equivalente a $MSE = 7 \times 10^{-2}$.

5.4.5 Simulação de Validação da Variável Nível de Soda - 2º Ensaio

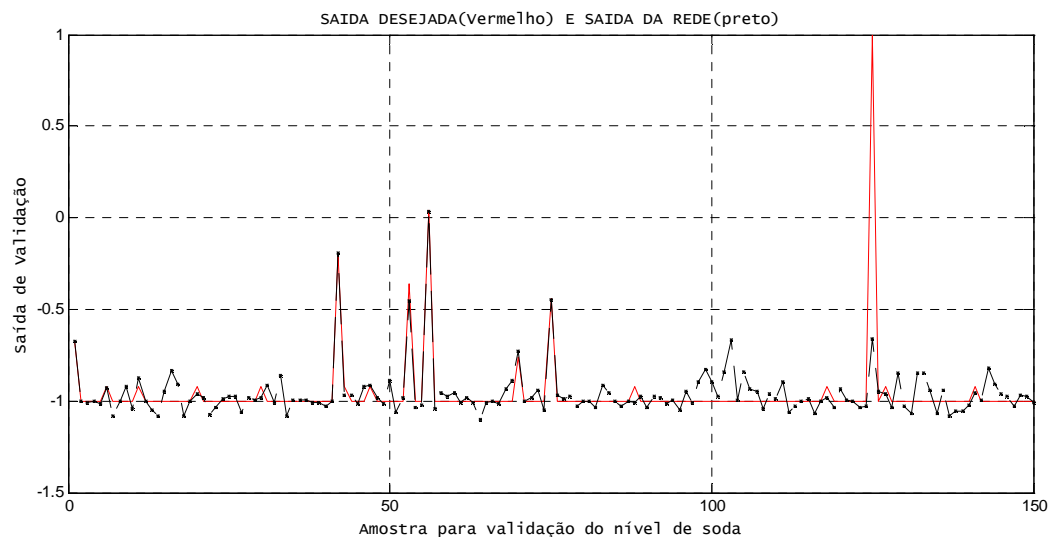


Figura 5.25 Validação da Variável Nível de Soda

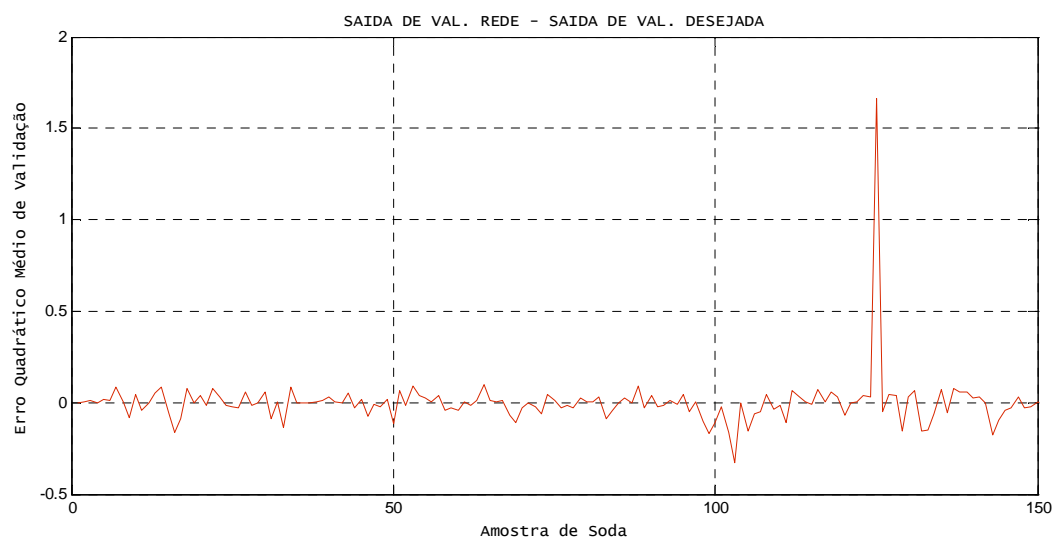


Figura 5.26 Erro de Validação da Variável Nível de Soda

Pode-se observar através da Figura 5.25, que o ajuste do modelo de RNA para a validação da variável níveis de soda foi bastante significativo. No entanto, com as mesmas características do ajuste da variável níveis de umidade, algumas observações não conseguiram acompanhar o padrão do processo. Isto é confirmado através da Figura 5.26, correspondente aos erros quadráticos médios de validação para esta variável.

5.4.6 Simulação das Validações Desnormalizadas - 2º Ensaio

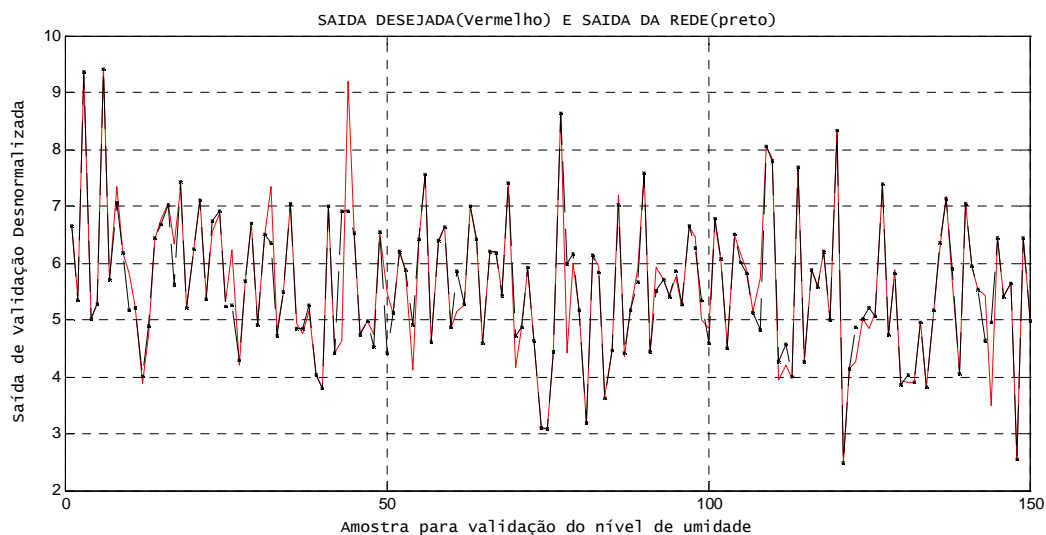


Figura 5.27 Validação Desnormalizada Variável Umidade

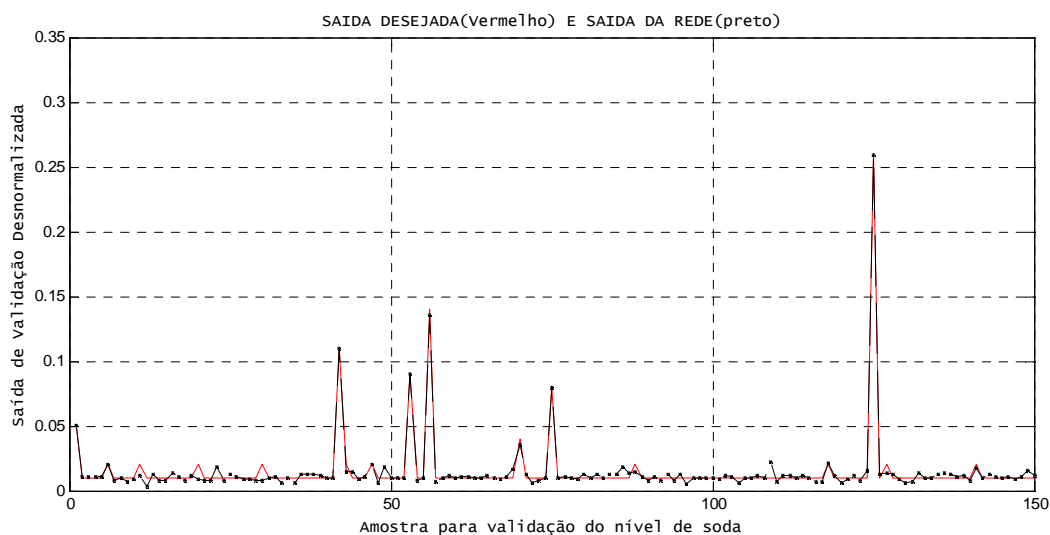


Figura 5.28 Validação Desnormalizada Variável Soda

Para uma melhor visualização das absevações geradas pela rede em relação as obsevações originais do processo, ver Figuras 5.27 e 5.28, que foram construídas a partir do processo de desnormalização dos resultados.

5.5 Estruturação e Treinamento da Rede - 3º Ensaio

Na busca por melhores aproximações, num terceiro momento, a rede foi treinada, com as variáveis de entrada normalizadas no espaço $(-1, 1)$ e, com três camadas internas de neurônios, com as seguintes características:

Topologia da Rede:

- 1ª Camada: Camada de entrada da rede, representada pelas variáveis independentes (condensado, rotação e vapor).
- 2ª Camada: Formada por 12 neurônios com função de ativação identidade e com saída apresentando uma função do tipo tangente hiperbólica (tansig).
- 3ª Camada: Formada por 10 neurônios com função de ativação identidade e com saída apresentando uma função do tipo tangente hiperbólica (tansig).
- 4ª Camada: Formada por 8 neurônios com função de ativação identidade e com saída apresentando uma função do tipo tangente hiperbólica (tansig).
- 5ª Camada: Definida como camada de saída da rede, formada por apenas dois neurônio, cuja a entrada apresenta uma função de ativação identidade e com a de saídas do tipo linear (purelin).

5.5.1 Treinamento da Rede - 3º Ensaio

A Figura 5.29, mostra a evolução da função $E(W)$, que se caracteriza por definir o valor do Erro Quadrático Médio (EQM) em cada interação de ajuste dos parâmetros do modelo de rede neural proposto.

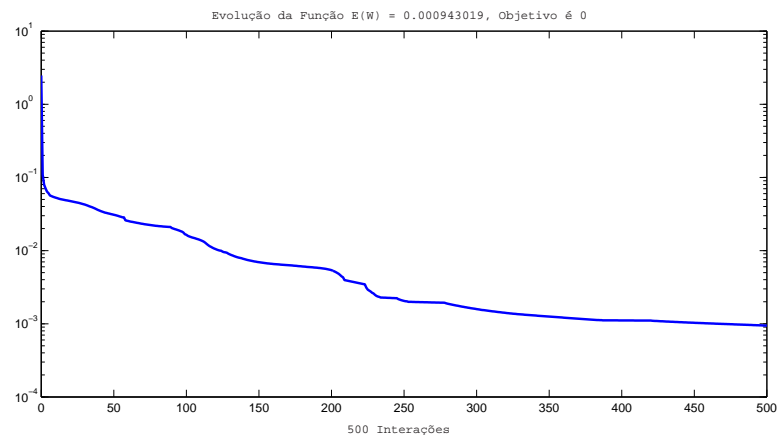


Figura 5.29 *Evlução da Função Erro $E(W)$*

Para esta rede foi estipulado inicialmente um número pré definido de interações, equivalente a 2000 interações.

No entanto, a mesma obteve convergência para um erro mínimo no treinamento em apenas 500 interações. Os parâmetros utilizados foram os seguintes: atualização da tela foi adotado de 50 em 50 épocas, taxa de aprendizagem de 0,07 e taxa de momento de 0,8 e pré determinou-se um erro final de treinamento equivalente $E(W)=0$ (zero). Esta configuração de rede, proporcionou um erro de treinamento bstante significativo, equivalente a $MSE = 9 \times 10^{-4}$.

5.5.2 Simulação de Treinamento da Variável Umidade - 3º Ensaio

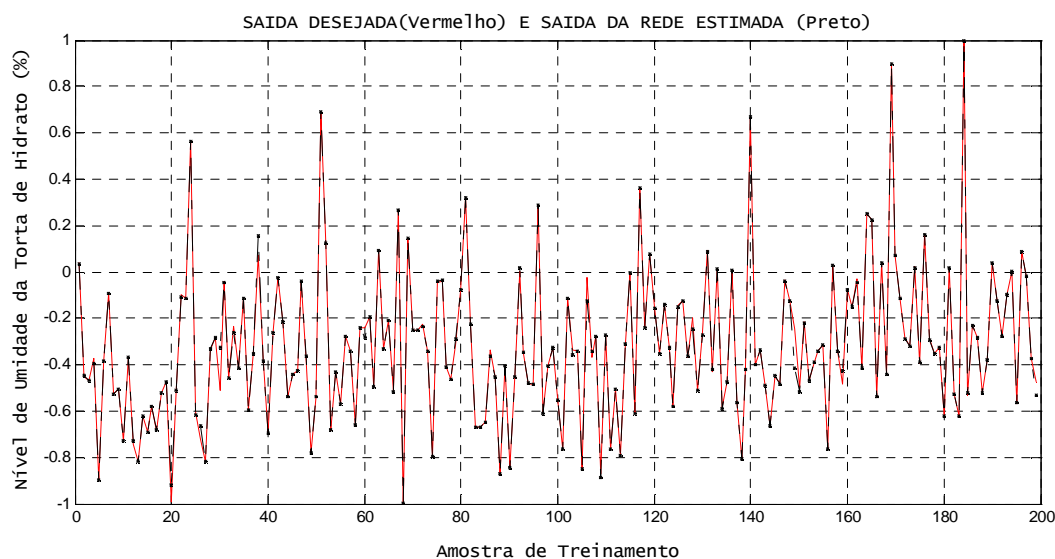


Figura 5.30 Treinamento da Variável Umidade

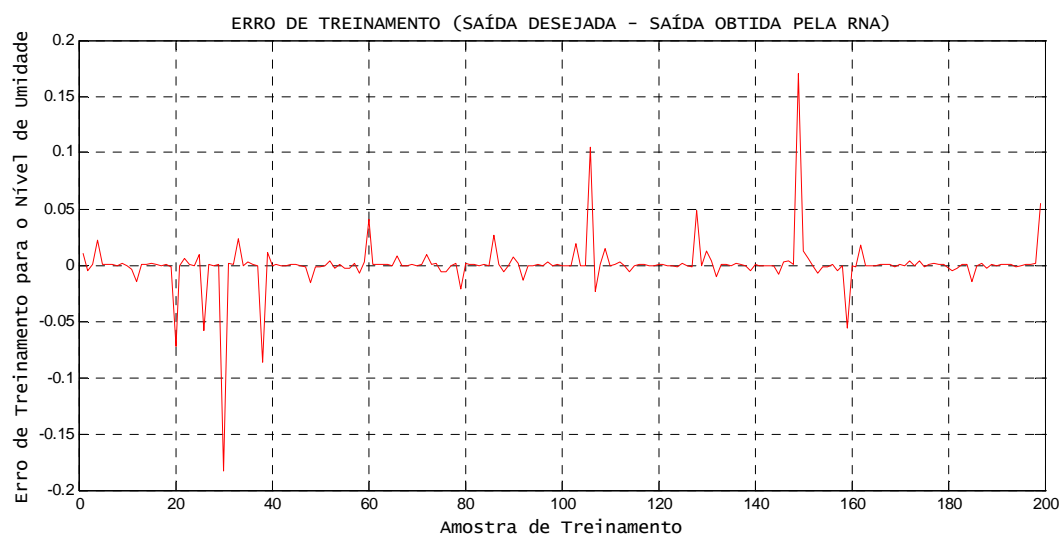
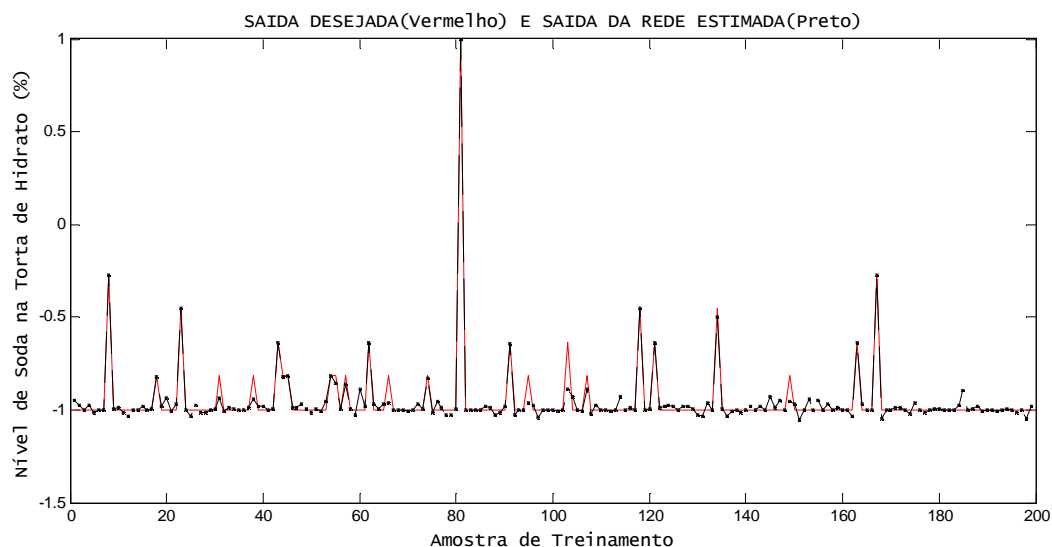
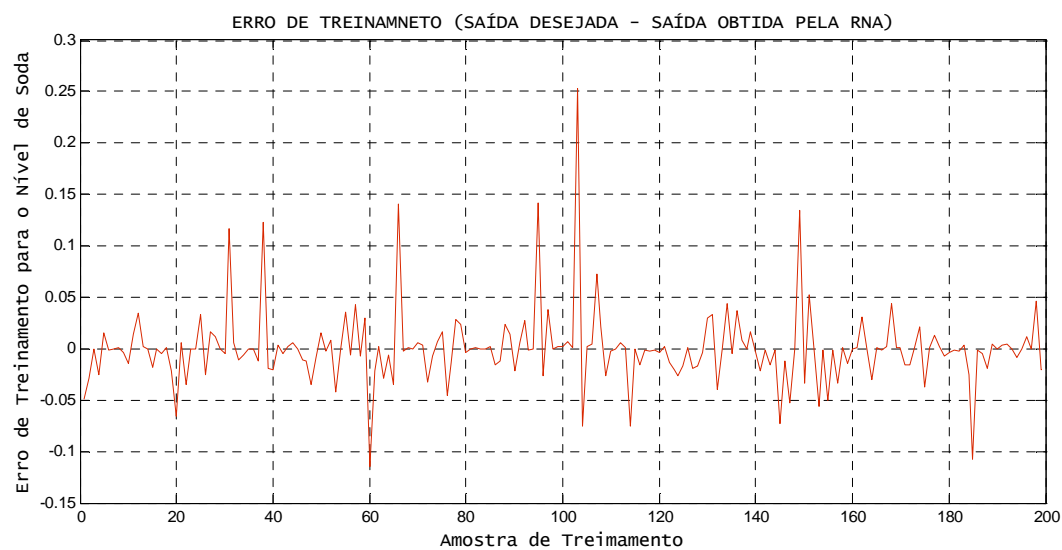


Figura 5.31 Erro de Treinamento da Variável Umidade

Pode-se observar através da Figura 5.30, que o ajuste do modelo de RNA para o treinamento da variável níveis de umidade foi bastante significativo, conseguindo acompanhar o comportamento do processo na grande maioria das observação da amostra selecionada. Isto é confirmado através das Figura 5.31, gerada a partir dos erros de treinamento cometidos para esta variável.

5.5.3 Simulação de Treinamento da Variável Soda - 3º Ensaio

Figura 5.32 *Treinamento da Variável Nível de Soda*Figura 5.33 *Erro de Treinamento da Variável Soda*

Pode-se observar através da Figura 5.32, que o ajuste de treinamento para o modelo de RNA conseguiu acompanhar o comportamento do processo na grande maioria das observação da amostra da variável Soda selecionada para treinamento. Isto é confirmado através da Figura 5.33, gerada a partir dos erros de treinamento fornecidos por esta variável.

5.5.4 Simulação de Validação da Variável Umidade - 3º Ensaio

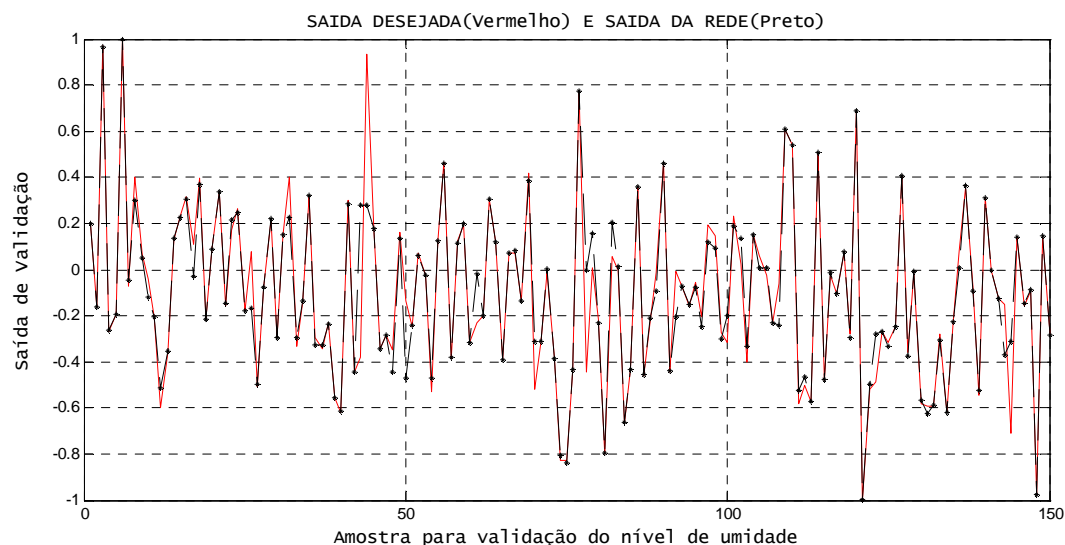


Figura 5.34 Validação da Variável Nível de Umidade

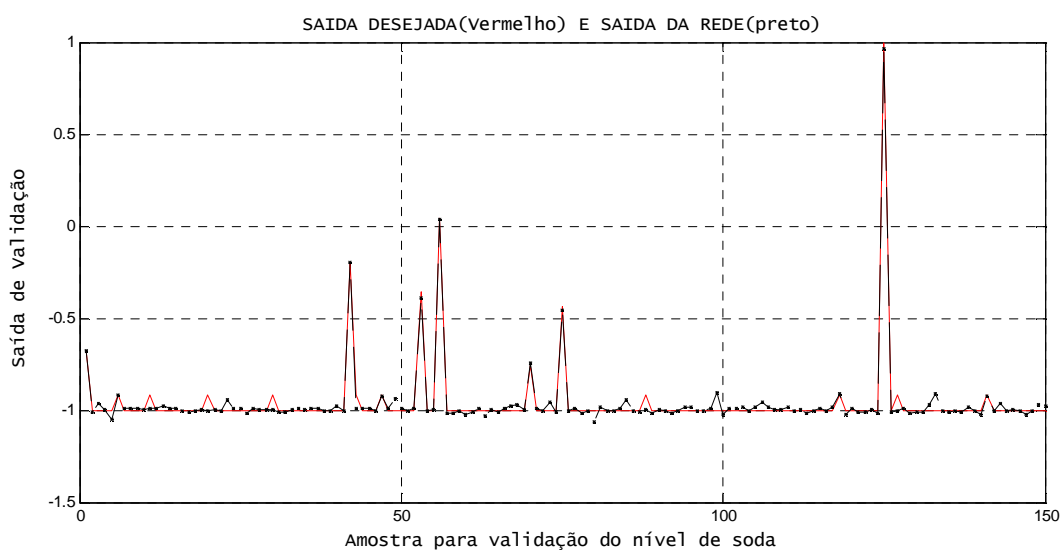


Figura 5.35 Validação da Variável Nível de Soda

Pode-se observar através das Figuras 5.34 e 5.35 que o ajuste de validação para o modelo de RNA proposto quando comparado com as informações das variáveis níveis de umidade e níveis de soda do processo real respectivamente, apresentam um comportamento preciso na grande maioria das observação da amostra de validação.

5.5.5 Simulação de Validação da Variável Soda - 3º Ensaio

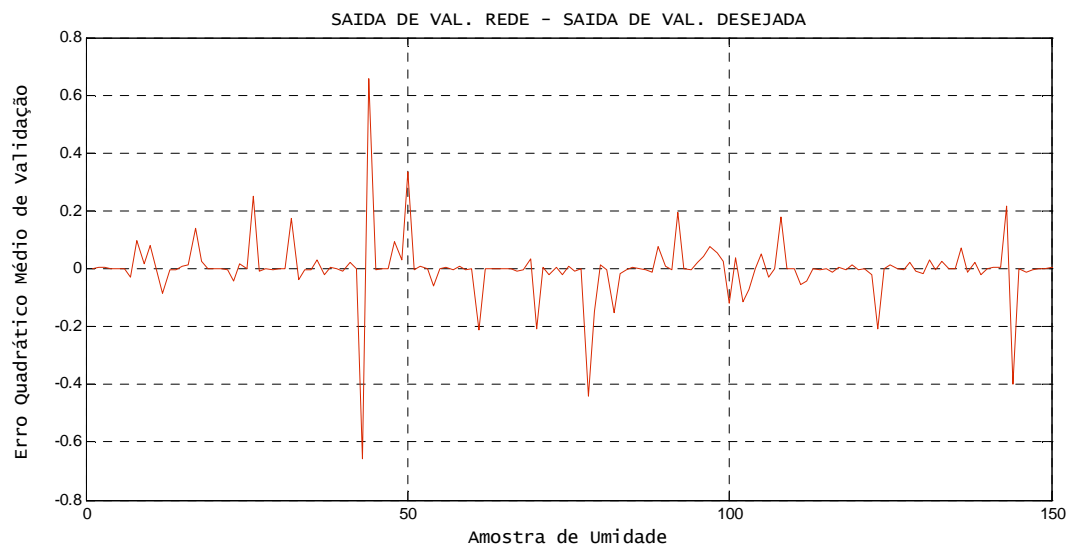


Figura 5.36 Erro de Validação da Variável Nível de Umidade

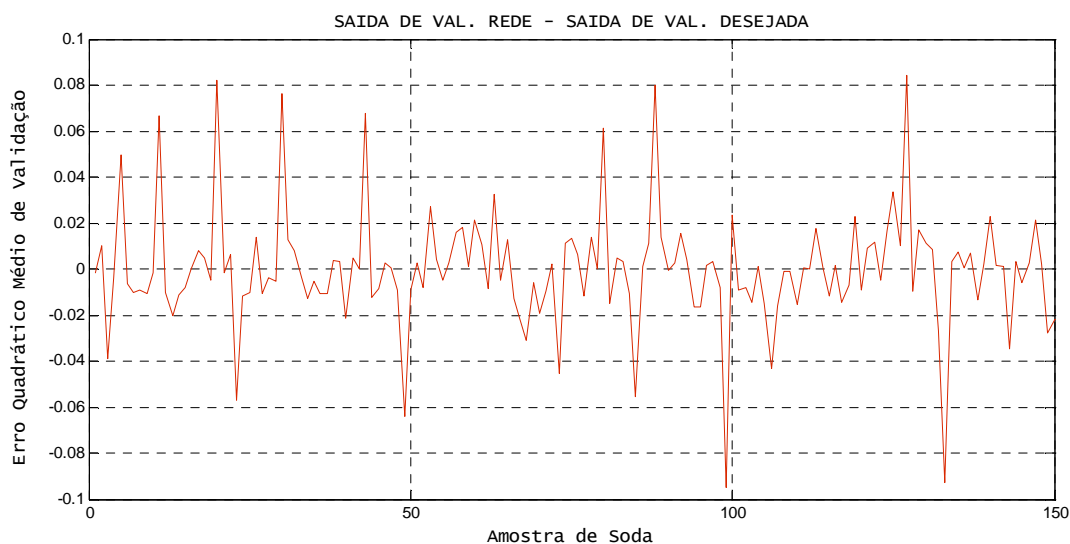


Figura 5.37 Erro de Validação da Variável Nível de Soda

A afirmação anterior, é de fácil comprovação ao se analisar as Figuras, 5.36 e 5.37, correspondentes ao erro de validação do nível de umidade e ao erro de validação do nível de soda presente na polpa. Estes gráficos apresentam erros bastante significativos em torno de $MSE = 6 \times 10^{-3}$ e conseqüentemente, espera-se que esta rede após instalada e adequada à planta, venha a fornecer excelentes resultados de mapeamento e controle do processo.

5.5.6 Simulação das Validações Desnormalizadas - 3º Ensaio

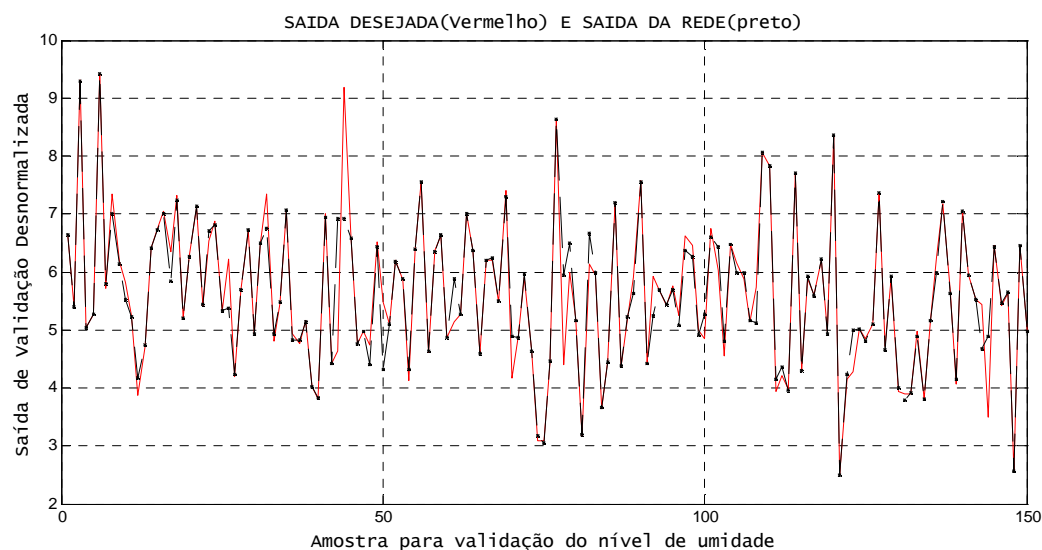


Figura 5.38 Validação Desnormalizada Variável Umidade

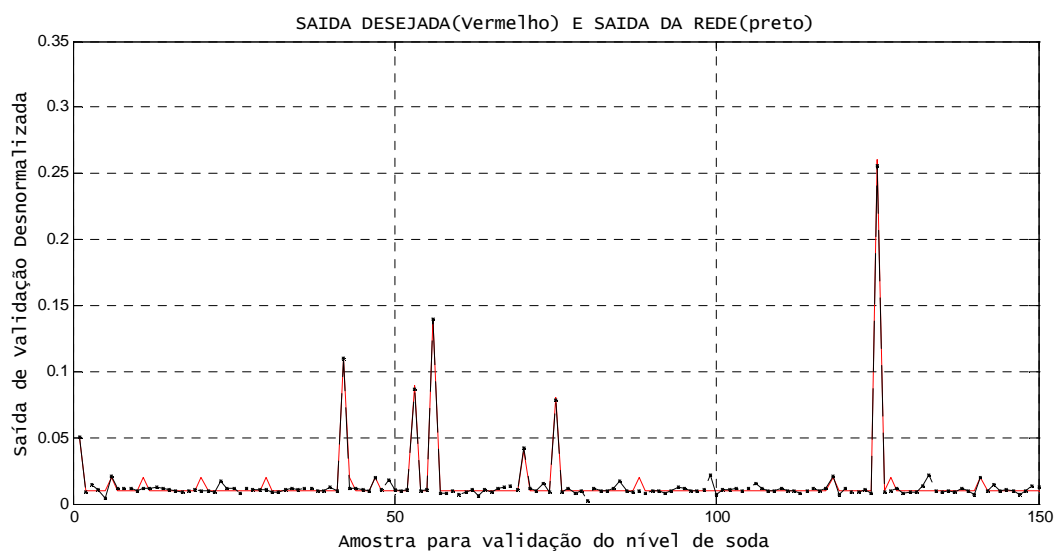


Figura 5.39 Validação Desnormalizada Variável Soda

Através das análises gráficas, realizadas com auxílio das Figuras 5.38 e 5.39, pode-se verificar que esta estrutura de rede foi a que melhor conseguiu se aproximar da verdadeira ocorrência do processo.

5.5.7 Sínteses dos Modelos de RNA Ajustados

Com o objetivo de melhor visualização e entendimento dos momentos de ajuste do modelo de rede neural, proposto neste trabalho, na Tabela 5.1 apresentamos todas as configurações adotadas nos três momentos dos ensaios realizados. Pelos dados, fica evidente os motivos que levaram à escolha da rede estruturada no terceiro momento.

Tabela 5.1 *Comparativo para Decisão entre as Três Redes Neurais Propostas*

Ensaio	Neurônios	Cam. Inter.	1ª Camada	2ª Camada	3ª Camada	Normalização	Algoritmo	MSE-t	MSE-v.
1º	(14-12-2)	2	logsig	logsig	—	Desvio Padrão	Trainlm	2×10^{-2}	5×10^{-2}
2º	(12-10-2)	2	tansig	tansig	—	(-1, 1)	Trainlm	8×10^{-3}	7×10^{-2}
3º	(12-10-8-2)	3	tansig	tansig	tansig	(-1, 1)	Trainlm	9×10^{-4}	2×10^{-3}

Os sistemas inteligentes de Redes Neurais Artificiais mostram-se bastante eficientes quando são utilizadas com o objetivo de monitoramento de processos industriais complexos. Especificamente neste trabalho, onde a preocupação foi estudar o padrão do processo estabelecendo a relação de influência das variáveis de entrada (*variáveis independentes*) nas variáveis de saída (*variáveis dependentes*), permitindo assim, um melhor controle e compreensão do processo produtivo e, conseqüentemente sua otimização.

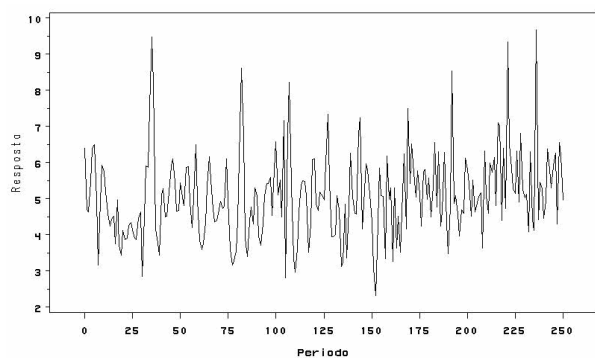
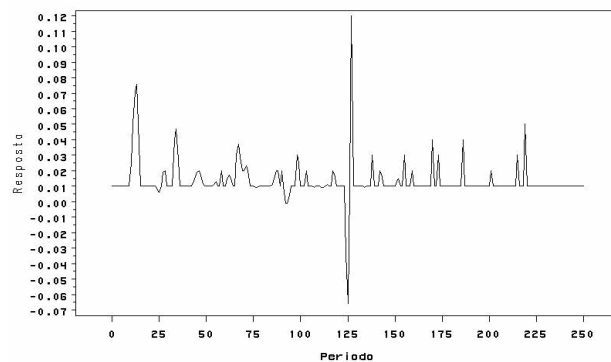
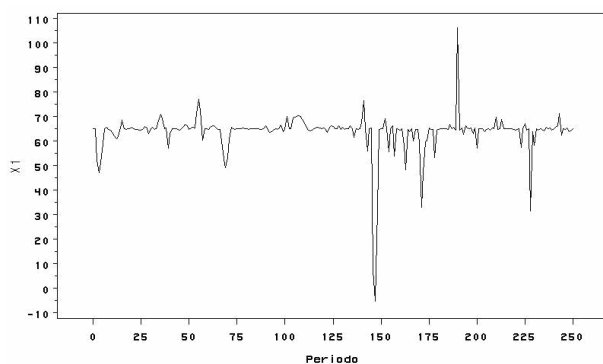
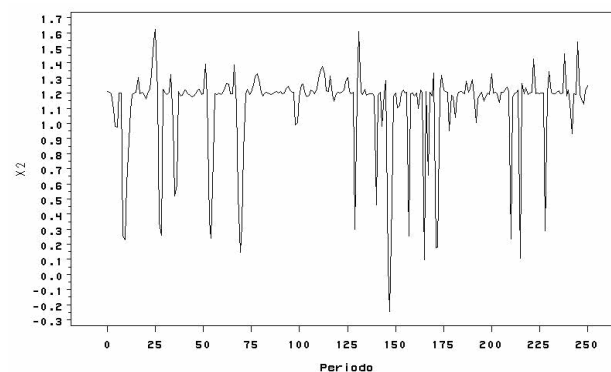
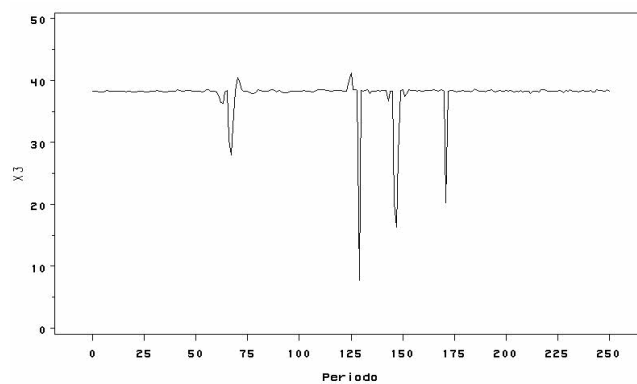
Portanto, o modelo de RNA descrito no terceiro momento, é o que apresenta melhor capacidade de acompanhar o padrão do processo de monitoramento da filtração e secagem da polpa de cristais de hidrato. Pois apresentou um excelente ajuste no treinamento e na validação, mostrando-se bastante eficiente. Assim, considera-se o objetivo específico relacionado a esta metodologia de estruturar uma Rede Neural Artificial que reconhecesse de forma robusta o padrão gerador do processo e que auxilie na avaliação comparativa entre as metodologias estudadas, alcançado.

5.6 Modelagem Multivariada com Função de Transferência

Sabe-se que qualquer série que descreva algum fenômeno físico, em geral, pode apresentar pelo menos duas componentes: uma relacionada ao sinal ou informação desejada daquele fenômeno e outra relacionada a um ruído que é caracterizado por apresentar os dados independentes e identicamente distribuídos (iid). Se uma série carrega alguma informação na metodologia de séries temporais, isso é caracterizado pela presença de um padrão não aleatório na série. Esse padrão não aleatório é, em geral, associado à ocorrência de correlação entre as observações da(s) série(s) em tempos diferentes. A determinação de quais são as observações que contribuem para a formação do padrão da série é fornecida pela função de auto-correlação, auto-correlação parcial e a correlação cruzada, no caso das séries temporais multivariadas. Deste modo, o objetivo da construção de um modelo em séries temporais é, em resumo, separar a informação do ruído, ou seja, capturar a informação presente na série.

Para a modelagem multivariada com função de transferência, foram utilizadas 200 observações de cada variável e 53 observações foram separadas para a validação dos modelos. Inicialmente para a modelagem foram consideradas as variáveis umidade na torta de hidrato como resposta, vazão de condensado, vazão de vapor e rotação do tambor como variáveis de controle. Em um segundo momento considerou-se as variáveis níveis de soda cáustica como a variável resposta, vazão de condensado, vazão de vapor e rotação do tambor novamente como variáveis de controle.

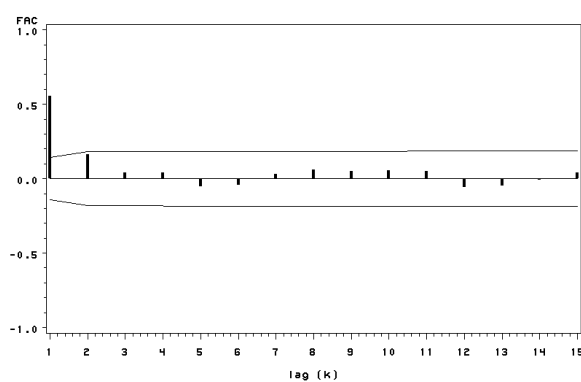
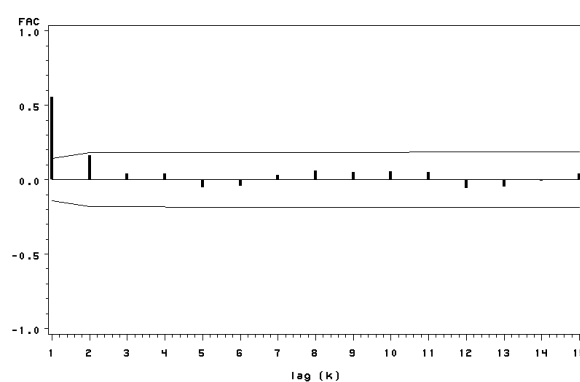
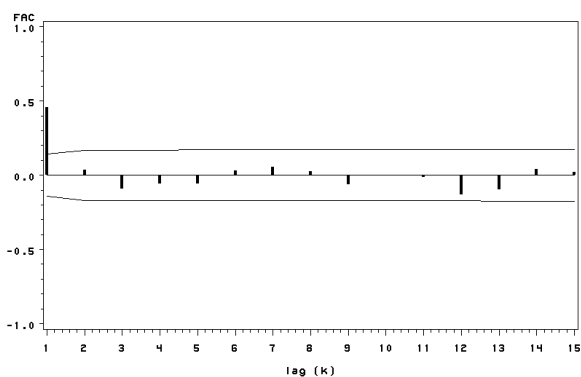
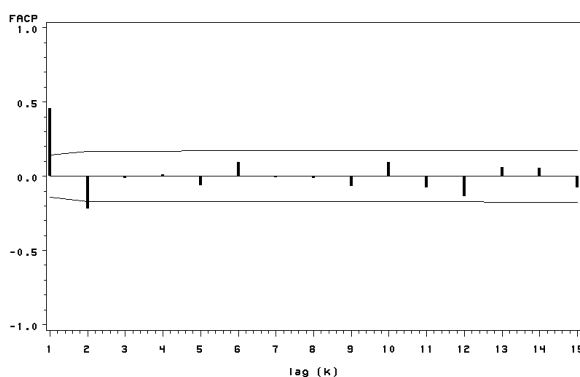
Os gráficos das séries são apresentados nas Figuras 5.40 a 5.44, onde pode-se observar que todas as séries são estacionárias e não há indícios de sazonalidade, não sendo assim, necessário diferenciações ou transformações. Entretanto, observa-se uma acentuada variabilidade relativa presente em todas elas.

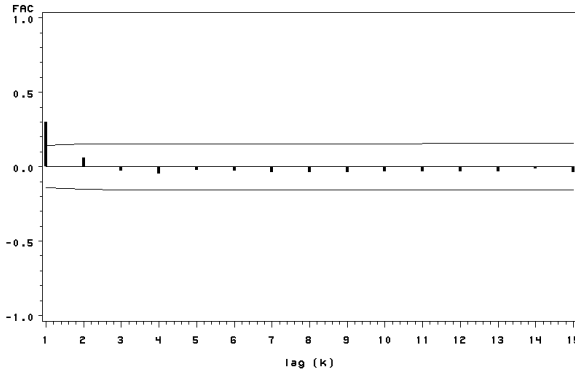
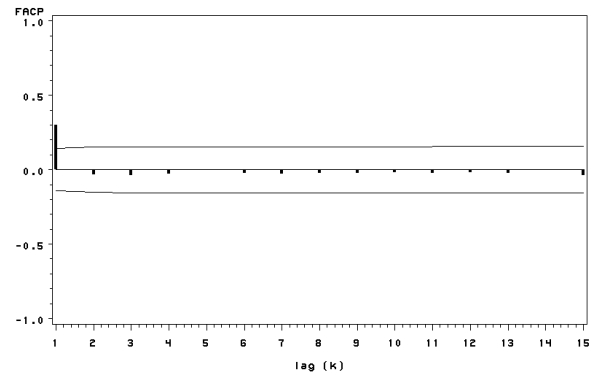
Figura 5.40 *Níveis de Umidade na Torta Y_1* Figura 5.41 *Níveis de Soda na Torta Y_2* Figura 5.42 *Vazão de Condensado X_1* Figura 5.43 *Vazão de Vapor X_2* Figura 5.44 *Rotação do Tambor X_3*

5.6.1 Modelagem para a Variável Nível de Umidade

5.6.2 Análise das *FAC* e *FACP*

Na tentativa de se ajustar modelos mais significativos para as variáveis explicativas, diversos modelos *ARIMA* são ajustados até que se encontre aquele que se ajuste mais adequadamente aos dados. A escolha dos modelos a serem ajustados são baseados na análise das *funções de auto-correlação (FAC)* e *auto-correlação parcial (FACP)*, Figuras 5.45 a 5.50.

Figura 5.45 *FAC para a Variável X_1* Figura 5.46 *FACP para a Variável X_1* Figura 5.47 *FAC para a Variável X_2* Figura 5.48 *FACP para a Variável X_2*

Figura 5.49 *FAC para a Variável X_3* Figura 5.50 *FACP para a Variável X_3*

Pode-se verificar analisando os gráficos das Funções de Autocorrelações e de Auto-correlações Parciais das variáveis de controle que, os intervalos de confiança para os lag's ficaram entre ± 2 desvio padrão. E, para este intervalo observou-se que em todas as séries há evidências de autocorrelações até o lag 2, com as autocorrelações parciais descrevendo seu comportamento como senóides amortecidas, implicando em possíveis modelos com parâmetros autorregressivos e médias móveis.

Portanto, as função de auto-correlação e função de auto-correlação parcial das variáveis de controle conduziram a três modelos para as variáveis condensado (X_1), vazão de vapor (X_2) e rotação (X_3):

- Um modelo auto-regressivo de ordem 2 (**AR**(2));
- um modelo de médias móveis de ordem 2 (**MA**(2));
- E um modelo de médias móveis de ordem 1 (**MA**(1))

$$X_{1t} = \phi_1 X_{1t-1} - \phi_2 X_{1t-2} + a_t \quad (5.1)$$

$$X_{2t} = \theta_1 a_{t-1} - \theta_2 a_{t-2} + a_t \quad (5.2)$$

$$X_{3t} = \theta_1 a_{t-1} + a_t \quad (5.3)$$

5.7 Ajuste de Modelos ARIMA

Para se ajustar modelos de função de transferência, deve-se previamente ajustar as séries modelos **ARIMA** do tipo descrito em 2.22, pois estes terão como finalidade sua utilização como filtros de pré-filtragens.

5.7.1 Ajuste de Modelos ARIMA para a Variável X1

Tabela 5.2 *Cálculo e Avaliação dos Parâmetros da Variável X1*

Parâmetros	Estimativas	Erro Padrão	t-valor	$\text{approx}Pr > t $	Lag
MU	64.607	0.929	68.43	< .0001	0
AR1,1	0.669	0.069	9.62	< .0001	1
AR1,2	-0.208	0.069	-2.99	< .0028	2

Tabela 5.3 *Estudo das Auto correlações dos Resíduos*

Lag	χ^2	Grau de Liberdade	$Pr > \chi^2$	Auto correlações					
6	6.71	4	0.1520	0.018	-0.031	-0.004	0.146	-0.090	-0.044
12	13.99	10	0.1733	0.051	0.040	-0.009	0.011	0.119	-0.125
18	16.88	16	0.3932	0.006	-0.013	0.043	-0.016	0.092	-0.049
24	31.04	22	0.0953	-0.142	0.092	0.025	-0.069	-0.043	0.163

A partir das avaliações dos parâmetros na Tabela 5.2, verificou-se que as estimativas dos parâmetros estimados foram significativos a um nível descritivo de 5%. E, com auxílio da Tabela 5.3, verificou-se que os resíduos apresentaram auto-correlações não significativas, o que caracteriza ajuste aceitável, ou seja, o modelo conseguiu capturar o padrão da série. Portanto, tem-se um modelo proposto **ARIMA(2,0,0)** ou **AR(2)** da seguinte forma:

$$X_{1t} = 0,667X_{1t-1} - 0,208X_{1t-2} + a_t \quad (5.4)$$

5.7.2 Ajuste de Modelos ARIMA para a Variável X2

Tabela 5.4 *Cálculo e Avaliação dos Parâmetros da Variável X2*

Parâmetros	Estimativas	Erro Padrão	t-valor	$\text{approx}Pr > t $	Lag
MU	1.104	0.032	34.39	< .0001	0
MA1,1	-0.574	0.070	-8.16	< .0001	1
MA1,2	-0.152	0.070	-2.17	.030	2

Tabela 5.5 *Estudo das Auto correlações dos Resíduos*

Lag	χ^2	Grau de Liberdade	$Pr > \chi^2$	Auto correlações					
6	3.01	4	0.5567	-0.011	-0.040	-0.065	0.014	-0.079	0.046
12	10.81	10	0.3729	0.026	0.057	-0.116	0.042	0.049	-0.122
18	23.31	16	0.1058	-0.090	0.097	0.030	-0.128	0.115	0.095
24	29.18	22	0.1397	0.128	-0.057	-0.003	-0.058	-0.048	-0.028

Para a variável X2 ajustou-se um modelo, onde os resultados das estimativas dos parâmetros estão dispostos na Tabela 5.4, onde pode-se observar que todos os parâmetros estimados foram significativos a um nível descritivo de 5%. As análises dos resíduos descritos na Tabela 5.5, apresentaram auto-correlações não significativas, ou seja, os mesmos não são autocorrelacionados o que caracteriza ajuste aceitável do modelo. Portanto, tem-se um modelo **ARIMA(0,0,2)** ou **MA(2)**, com a seguinte estrutura funcional:

$$X_{2t} = 0.574a_{t-1} - 0.152a_{t-2} + a_t \quad (5.5)$$

5.7.3 Ajuste de Modelos ARIMA para a Variável X3

Tabela 5.6 *Cálculo e Avaliação dos Parâmetros da Variável X3*

Parâmetros	Estimativas	Erro Padrão	t-valor	$\text{approx}Pr > t $	Lag
MU	37.673	0.298	126.28	< .0001	0
MA1,1	-0.286	0.068	-4.20	< .0001	1

Tabela 5.7 *Estudo das Auto correlações dos Resíduos*

Lag	χ^2	Grau de Liberdade	$Pr > \chi^2$	Auto correlações					
6	1.51	5	0.912	0.021	0.062	-0.035	-0.038	-0.008	-0.018
12	2.23	11	0.997	-0.026	-0.025	-0.027	-0.021	-0.024	-0.018
18	17.71	17	0.407	-0.028	-0.005	-0.015	-0.079	0.170	0.184
24	22.46	23	0.492	0.059	-0.044	-0.042	-0.061	0.013	0.099

Para a variável X3 ajustou-se um modelo preliminar, onde os resultado da estimativa do parâmetro, estão dispostos na Tabela 5.6, onde observa-se que o parâmetro estimado de média móvel foi significativo a um nível descritivo de 5%. E, no estudo realizado nos resíduos do modelo verificou-se que os mesmos, apresentaram auto-correlações não significativas, ou seja, os mesmos não são autocorrelacionados o que caracteriza ajuste aceitável do modelo. Portanto, tem-se um modelo $MA(0,0,1)$ ou $MA(1)$, o qual será utilizado como filtro para pré-filtragem para a modelagem de função de transferência, apresentando a seguinte estrutura funcional:

$$X_{3t} = 0.286a_{t-1} + a_t \quad (5.6)$$

5.8 Correlações Cruzadas

Com o objetivo de se verificar a relação de influência das variáveis explicativas nas variáveis dependentes, foram construídas e analisadas as funções de auto-correlações cruzadas entre estas variáveis. As mesmas, são apresentadas a seguir.

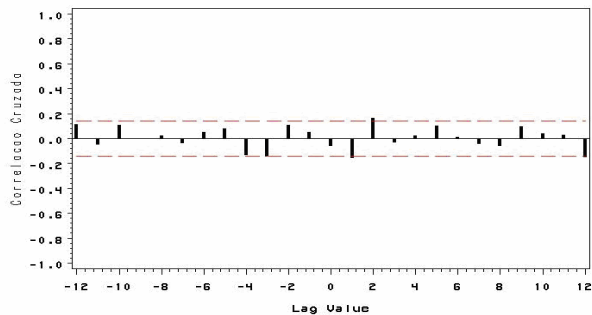


Figura 5.51 *Função de Correlação Cruzada Y_1 vs X_1*

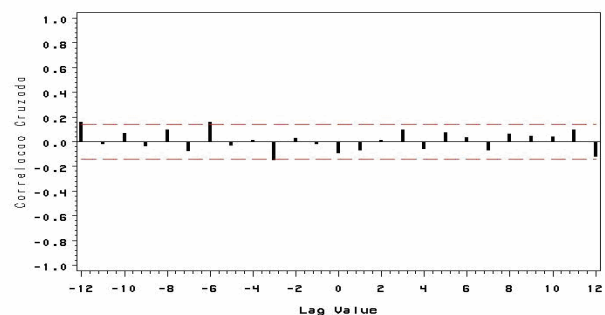


Figura 5.52 *Função de Correlação Cruzada Y_1 vs X_2*

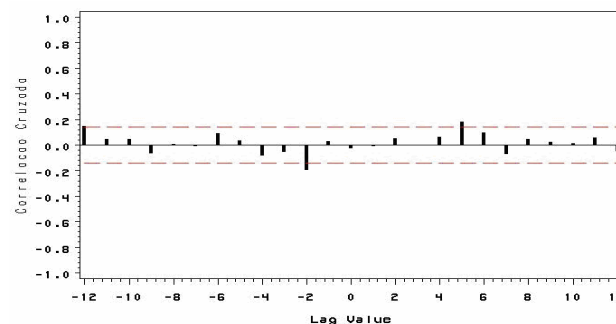


Figura 5.53 *Função de Correlação Cruzada Y_1 vs X_3*

5.8.1 Análise da Correlação Cruzada entre Y_1 vs X_1

Observa-se na Figura 5.51, correspondente a Correlação Cruzada entre Y e X_1 que não existe spike no lag 0. Logo o modelo pode apresentar defasagem de 1 lag, percebe-se também que os lag's 1 e 2 apresentam correlação significativa e após este spike o gráfico apresenta um decaimento imediatamente para zero, indicando que não há necessidade de parâmetro no denominador para o modelo de função de transferência.

5.8.2 Análise da Correlação Cruzada entre Y_1 vs X_2

Observa-se na Figura 5.52, correspondente a Correlação Cruzada entre Y e X2 que não há presença de spike significativo no **lag 0**, porém observa-se que existe correlação nos lag's -3 e -6 caracterizando um modelo com **feedback**. Portanto, será necessário pré-ajustar também a série de saída Y1.

5.8.3 Análise da Correlação Cruzada entre Y_1 vs X_3

A Correlação cruzada entre Y e X3 representada na Figura 5.53, mostra que não há spike no lag 0, nem nos lag's positivos, e verifica-se que há presença de correlação significativa no lag -2. Este comportamento é um indicativo de modelo com **feedback**. Sendo assim, é necessário pré-ajustar também a série de entrada Y1. A FCC apresenta um decaimento exponencial a partir do lag 5, indicando possivelmente a necessidade de parâmetros no denominador do modelo de função de transferência.

5.8.4 Ajuste dos Modelos de Função de Transferência

A partir das análises feitas sobre as funções de correlação cruzada, testou-se vários modelos de função de transferência, sendo que, em todos os modelos houve a necessidade de se fazer o ajuste da componente de erro ou ruído no modelo final da função de transferência. Alguns desses modelos, apresentados na forma de operador atraso B , cujos os respectivos resultados e análises, estão apresentados a seguir:

$$\text{Modelo1 : } Y_{1t} = (\varpi_{01} - \varpi_{11}B)X_{1t-1} + \varpi_{02}X_{2t} + (\varpi_{03} - \varpi_{13}B^5)X_{3t} + (1 - \theta B)a_t \quad (5.7)$$

$$\text{Modelo 2 : } Y_{1t} = (\varpi_{01} - \varpi_{11}B)X_{1t-1} + (\varpi_{03} - \varpi_{13}B^5)X_{3t} + (1 - \theta B)a_t \quad (5.8)$$

$$\text{Modelo3 : } Y_{1t} = \frac{(\varpi_{01} - \varpi_{11}B)}{(1 - \delta_{11}B)}X_{1t} + (\varpi_{03} - \varpi_{13}B^5)X_{3t} + (1 - \theta B)a_t \quad (5.9)$$

5.8.5 Estimativas para a Estruturação do Modelo 1 em 5.7

Tabela 5.8 *Cálculo dos Parâmetros Estimados para o Modelo de F.T.*

Parâmetros	Estimativas	E.P	t-valor	$\text{aproxPr} > t $	Lag	Variável	Deslocamento
MU	3.026	1.412	2.14	0.032	0	Y1	0
MA1,1 (θ)	-0.363	0.070	-5.21	< 0.000	1	Y1	0
NUM1 (ϖ_{01})	-0.024	0.011	-2.17	0.029	0	X1	1
NUM1,1 (ϖ_{11})	-0.023	0.011	-2.18	0.029	1	X1	1
NUM2 (ϖ_{02})	-0.411	0.316	-1.30	0.194	0	X2	0
NUM3 (ϖ_{03})	0.013	0.026	0.50	0.618	0	X3	0
NUM1,1 (ϖ_{13})	-0.052	0.023	-2.28	0.023	5	X3	0

Através das Tabela 5.8, pode-se observar que as estimativas dos parâmetros ϖ_{01} e ϖ_{11} , correspondente à variável X1, e à estimativa ϖ_{13} correspondente à variável X3 e a estimativa do parâmetro média móvel θ apresentam estatísticas t significativas a 5%. E, portanto serão utilizadas para a formulação do modelo de função de transferência Equação 5.10.

$$Y_{1t} = (-0,024 + 0,023B)X_{1t-1} - 1,30X_{2t} + (0,013 + 0,015B^5)X_{3t} + (1 + 0,362B)a_t \quad (5.10)$$

Os números em **destaque** na tabela 5.8, referem-se à estatística t . Onde, para valores de $|t| > 2,00$ as estimativas dos parâmetros para o modelo são consideradas significativas ao nível de 5

5.8.6 Análise Residual para o Modelo de F.T em 5.7

Tabela 5.9 *Estudo das Auto correlações dos Resíduos*

Lag	χ^2	Grau de Liberdade	$Pr > \chi^2$	Auto correlações					
6	7.70	5	0.173	0.035	0.048	-0.162	-0.065	-0.025	-0.063
12	14.64	11	0.199	0.134	-0.004	0.019	-0.114	0.030	-0.039
18	16.91	17	0.461	-0.064	-0.026	-0.056	0.025	-0.013	0.044
24	20.24	23	0.627	-0.013	-0.010	-0.041	-0.089	0.065	0.030

Tabela 5.10 *Estudo das Correlação Cruzada dos Resíduos vs X1*

Lag	χ^2	Grau de Liberdade	$Pr > \chi^2$	Correlação Cruzada					
5	0.90	5	0.970	0.009	0.023	-0.032	-0.009	0.054	-0.007
11	7.89	11	0.723	-0.003	-0.066	0.072	0.055	0.072	-0.133
17	11.95	17	0.803	-0.022	-0.057	0.068	0.000	0.047	0.102
23	18.25	23	0.744	0.101	-0.065	-0.054	-0.010	-0.120	-0.023

Tabela 5.11 *Estudo das Correlação Cruzada dos Resíduos vs X2*

Lag	χ^2	Grau de Liberdade	$Pr > \chi^2$	Correlação Cruzada					
5	3.31	6	0.769	0.013	-0.014	-0.078	0.080	-0.064	-0.007
11	7.05	12	0.855	0.023	-0.015	0.057	0.011	0.076	0.097
17	13.49	18	0.761	-0.097	-0.020	-0.067	0.082	0.040	0.102
23	20.46	24	0.670	0.010	0.102	-0.115	0.004	0.094	-0.056

Tabela 5.12 *Estudo das Correlação Cruzada dos Resíduos vs X3*

Lag	χ^2	Grau de Liberdade	$Pr > \chi^2$	Correlação Cruzada					
5	2.01	5	0.848	0.004	0.072	-0.011	-0.038	0.059	0.010
11	8.04	11	0.709	0.093	-0.094	0.081	0.006	0.021	0.081
17	13.30	17	0.716	-0.060	0.073	-0.094	-0.075	0.051	0.033
23	19.85	23	0.651	0.132	0.040	-0.071	-0.086	0.022	-0.041

Para o modelo 1 proposto em 5.10, observa-se que a avaliação das auto-correlações dos resíduos Tabela 5.9, têm estatísticas Q não significativas o que indica que os resíduos desse modelo preliminar de função de transferência são ruídos brancos. Deste modo, o modelo se ajustou adequadamente aos dados. Além do mais, o cheque da correlação cruzada dos resíduos com as variáveis X1, X2 e X3 na entrada Tabela 5.10 a 5.11, apresentam correlações cruzadas dos resíduos não significativas, evidenciando que as funções de transferência relacionadas a essas variáveis são adequadas.

5.8.7 Estimativas para a Estruturação do Modelo 2

Tabela 5.13 *Cálculo dos Parâmetros Estimados para o Modelo de F.T.*

Parâmetros	Estimativas	E.P	t-valor	$\text{aproxPr} > t $	Lag	Variável	Deslocamento
MU	3.257	1.408	2.31	0.021	0	Y1	0
MA1,1 (θ)	-0.372	0.069	-5.40	< 0.0001	1	Y1	0
NUM1 (ϖ_{01})	-0.026	0.011	-2.40	0.016	0	X1	1
NUM1,1 (ϖ_{11})	-0.023	0.011	-2.19	0.029	1	X1	1
NUM3 (ϖ_{03})	-0.002	0.023	-0.07	0.947	0	X3	0
NUM1,1 (ϖ_{13})	-0.052	0.023	-2.27	0.024	5	X3	0

Através das Tabela 5.13, pode-se observar que as estimativas dos parâmetros ϖ_{01} , ϖ_{11} , correspondente à variável X1, e à estimativa ϖ_{13} correspondente à variável X3 e à estimativa do parâmetro média móvel θ apresentam estatísticas t significativas a 5%. E, portanto serão utilizadas para a formulação do modelo de função de transferência (Eq. 5.11).

$$Y_{1t} = (-0,026 + 0,023B)X_{1t-1} + (0,002 + 0,052B^5)X_{3t} + (1 + 0,372B)a_t \quad (5.11)$$

5.8.8 Análise Residual para o Modelo de F.T em 5.8

Tabela 5.14 *Estudo das Auto correlações dos Resíduos*

Lag	χ^2	Grau de Liberdade	$Pr > \chi^2$	Auto correlações					
6	7.45	5	0.189	0.037	0.053	-0.154	-0.059	-0.027	-0.071
12	15.16	11	0.175	0.132	-0.018	0.017	-0.132	0.025	-0.039
18	18.39	17	0.365	-0.055	-0.025	-0.081	0.021	-0.022	0.062
24	22.08	23	0.516	-0.010	-0.008	-0.040	-0.096	0.070	0.026

Para o modelo proposto em 5.11, observa-se que a avaliação das auto-correlações dos resíduos Tabela 5.14, têm estatísticas Q não significativas o que indica que os resíduos desse modelo de função de transferência são ruídos brancos. Deste modo, o modelo se ajustou adequadamente aos dados.

Tabela 5.15 *Estudo das Correlação Cruzada dos Resíduos vs X1*

Lag	χ^2	Grau de Liberdade	$Pr > \chi^2$	Correlação Cruzada					
5	0.93	5	0.968	0.009	0.023	-0.031	-0.013	0.055	-0.006
11	6.95	11	0.803	-0.003	-0.064	0.073	0.046	0.056	-0.125
17	10.50	17	0.881	-0.017	-0.055	0.062	0.010	0.049	0.092
23	16.19	23	0.847	0.088	-0.061	-0.061	-0.007	-0.116	-0.022

Tabela 5.16 *Estudo das Correlação Cruzada dos Resíduos vs X3*

Lag	χ^2	Grau de Liberdade	$Pr > \chi^2$	Correlação Cruzada					
5	1.82	5	0.873	0.003	-0.058	-0.020	-0.045	0.058	0.011
11	7.59	11	0.749	0.093	-0.096	0.087	0.002	0.019	0.062
17	12.62	17	0.762	-0.049	0.073	-0.096	-0.069	0.059	0.023
23	18.65	23	0.721	0.121	0.033	-0.070	-0.089	0.028	-0.039

Através do cheque da correlação cruzada dos resíduos com as variáveis X1, e X3 na entrada Tabelas 5.15 e 5.16, apresentam correlações cruzadas dos resíduos não significativas, evidenciando que as funções de transferência relacionadas a essas variáveis são adequadas.

5.8.9 Estimativas para a Estruturação do Modelo 3 em 5.9

Tabela 5.17 *Cálculo dos Parâmetros Estimados para o Modelo de F.T.*

Parâmetros	Estimativas	E.P	t-valor	$approx Pr > t $	Lag	Variável	Deslocamento
MU	4.176	1.316	3.17	0.002	0	Y1	0
MA1,1 (θ)	-0.372	0.069	-5.38	< 0.0001	1	Y1	0
NUM1 (ϖ_{01})	-0.003	0.012	-0.21	0.834	0	X1	1
NUM1,1 (ϖ_{11})	-0.027	0.011	2.34	0.019	1	X1	1
DEM1,1 (δ_{11})	-0.0525	0.296	-1.77	0.077	1	X1	1
NUM3 (ϖ_{03})	0.0004	0.026	0.01	0.988	0	X3	0
NUM1,1 (ϖ_{13})	-0.053	0.023	-2.31	0.021	5	X3	0

Através das Tabela 5.17, pode-se observar que as estimativas dos parâmetros ϖ_{11} , δ_{11} , correspondente à variável X1, e à estimativa ϖ_{13} correspondente à variável X3 e à estimativa do parâmetro média móvel θ apresentam estatísticas t significativas a 5%. E, portanto serão utilizadas para a formulação do modelo de função de transferência (Eq. 5.12).

$$Y_{1t} = \frac{(-0,003 - 0,027B)}{(1 - 0,525B)}X_{1t} + (0,0004 + 0,053B^5)X_{3t} + (1 + 0,372B)a_t \quad (5.12)$$

5.8.10 Análise Residual para o Modelo 3 de F.T em 5.9

Tabela 5.18 *Estudo das Auto correlações dos Resíduos*

Lag	χ^2	Grau de Liberdade	$Pr > \chi^2$	Auto correlações					
6	8.09	5	0.152	0.040	0.059	-0.151	-0.057	-0.040	-0.086
12	16.22	11	0.133	0.137	-0.003	0.016	-0.133	0.022	-0.048
18	18.33	17	0.369	-0.041	-0.015	-0.060	0.034	-0.012	0.054
24	22.39	23	0.497	-0.005	-0.014	-0.045	-0.105	0.066	0.027

Tabela 5.19 *Estudo das Correlação Cruzada dos Resíduos vs X1*

Lag	χ^2	Grau de Liberdade	$Pr > \chi^2$	Correlação Cruzada					
5	2.53	4	0.640	-0.008	-0.009	0.088	0.036	-0.016	0.059
11	5.98	10	0.817	-0.006	-0.017	-0.075	0.075	0.050	0.061
17	10.82	16	0.820	-0.126	-0.007	-0.056	0.055	0.011	0.050
23	18.11	22	0.699	0.095	0.090	-0.055	-0.073	-0.010	-0.109

Tabela 5.20 *Estudo das Correlação Cruzada dos Resíduos vs X3*

Lag	χ^2	Grau de Liberdade	$Pr > \chi^2$	Correlação Cruzada					
5	2.90	5	0.715	0.004	0.079	0.065	-0.005	0.065	0.012
11	8.50	11	0.668	0.093	-0.096	0.083	0.008	0.023	0.058
17	14.04	17	0.664	-0.048	0.074	-0.101	-0.074	0.064	0.028
23	19.56	23	0.668	0.124	0.049	-0.033	-0.088	0.022	-0.037

Para o modelo proposto em 5.12, observa-se que a avaliação das auto-correlações dos resíduos Tabela 5.18, apresenta estatísticas Q não significativas, o que indica que os resíduos desse modelo de função de transferência são ruídos brancos. Deste modo, o modelo se ajustou adequadamente aos dados. Além do mais, o cheque da correlação cruzada dos resíduos com as variáveis X1, e X3 na entrada Tabela 5.19 e 5.20, apresentam correlações cruzadas dos resíduos não significativas evidenciando que as funções de transferência relacionadas a essas variáveis são adequadas.

5.8.11 Seleção do Melhor Modelo para o Nível de Umidade

Ao comparar os ajustes fornecidos pelos três modelos Figuras 5.54 a 5.56, verifica-se que a Figura 5.54, gerada a partir do teste de ajuste, foi o que produziu menor erro quadrático médio equivalente à $MSE = 1,049$.

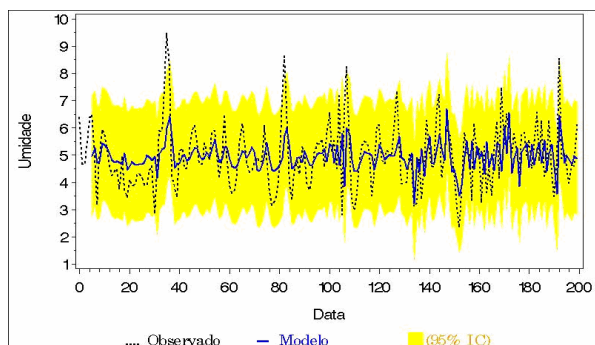


Figura 5.54 Ajuste do Modelo 5.10 de Função de Transferência

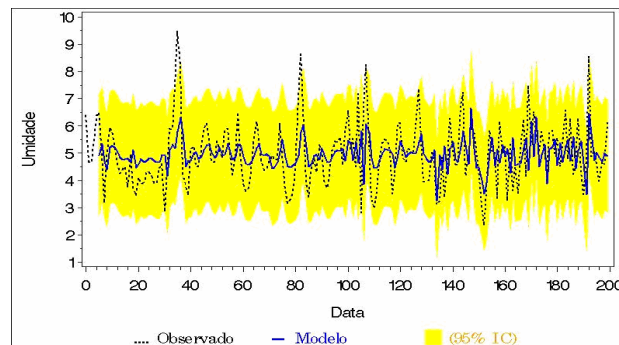


Figura 5.55 Ajuste do Modelo 5.11 de Função de Transferência

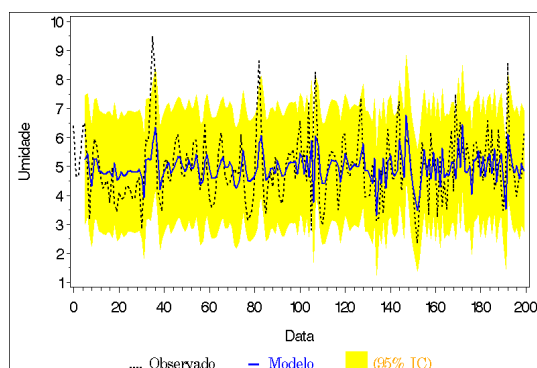


Figura 5.56 Ajuste do Modelo 5.12 de Função de Transferência

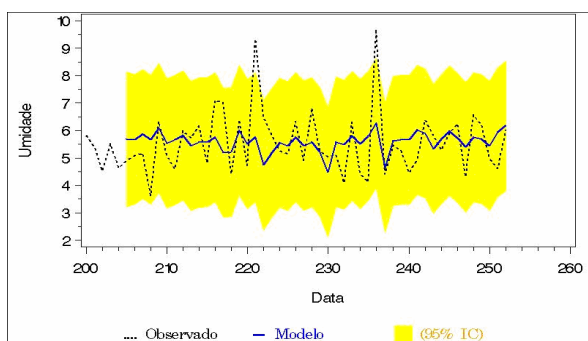
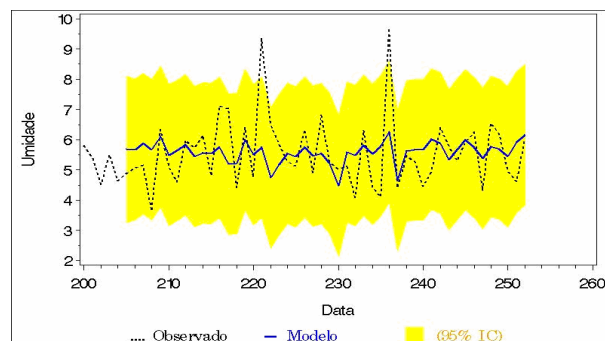
Um outro método de comparação para se verificar e selecionar o modelo que melhor se ajustou aos dados é utilizando os critérios de **Akaike** e **SBC**, onde o modelo que apresentar menores valores nessas estatísticas é considerado o que apresenta melhor ajuste. Portanto, através da Tabela 5.21, pode-se verificar que o segundo modelo descrito em 5.11 é o que apresenta melhores resultados ao descrever o padrão do processo, ou seja, é o que melhor explica a dinâmica da série gerada pela variável nível de umidade do processo.

Tabela 5.21 *Cr terios para Sele  o do Melhor Modelo em Ajuste Ajustado*

Cr�terio	Modelo 1	Modelo 2	Modelo 3
AIC	580.482	580.223	583.692
SBC	603.393	599.861	606.603
MSE	1.049	1.063	1.073
RMSE	1.024	1.036	1.036

5.8.12 Teste de Valida  o para os Modelos do N vel de Umidade

Esses modelos foram avaliados e validados com a aplica  o de um conjunto de 53 observa  es separadas do banco de dados. Em todos os casos as fun  es de transfer  ncia relacionadas foram adequadas, isto  , os res duos apresentaram-se n o correlacionados e as correla  es cruzadas entre os res duos e as vari veis de controle foram n o significativas. Nas Figuras 5.57, 5.58 e 5.59 s o apresentados os resultados de cada modelo e, na Tabela 5.22, s o apresentados os cr terios de compara  o.

Figura 5.57 *Teste de Valida  o do 1  Modelo*Figura 5.58 *Teste de Valida  o do 2  Modelo*

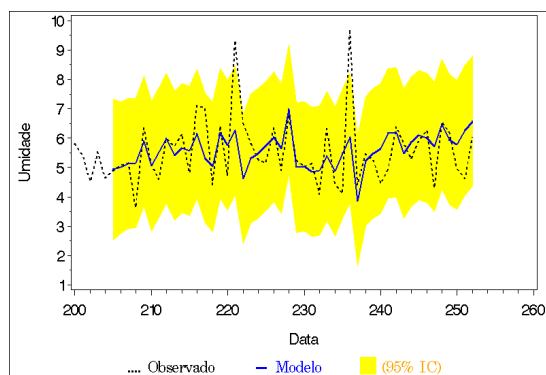


Figura 5.59 *Teste de Validação do 3º Modelo*

Tabela 5.22 *Crítérios para Seleção do Melhor Modelo em Validação*

Crítério	Modelo 1	Modelo 2	Modelo 3
AIC	160.524	158.528	153.057
SBC	173.622	169.755	166.155
MSE	1.292	1.292	1.074
RMSE	1.136	1.136	1.036

Para o ajuste em validação pode-se observar que o modelo descrito na equação 5.12 e representado na Figura (5.59) apresentou melhores resultados. Portanto, este modelo será utilizado na comparação entre as duas classes de modelos dinâmicos.

5.9 Modelagem para a Variável Nível de Soda

5.9.1 Análise das *FAC* e *FACP*

As correlações cruzadas entre a variável Y_2 (soda) e as variáveis de controle são apresentadas nas Figuras 5.60 a 5.62.

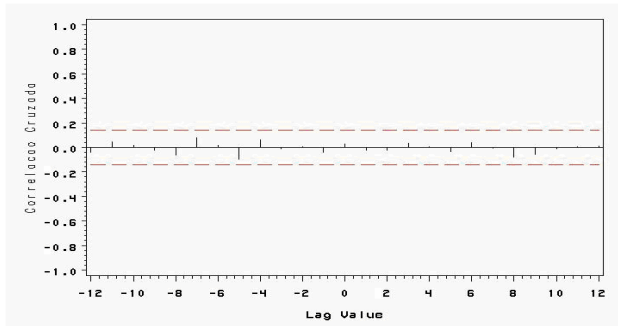


Figura 5.60 Função de Correlação Cruzada Y_2 vs X_1

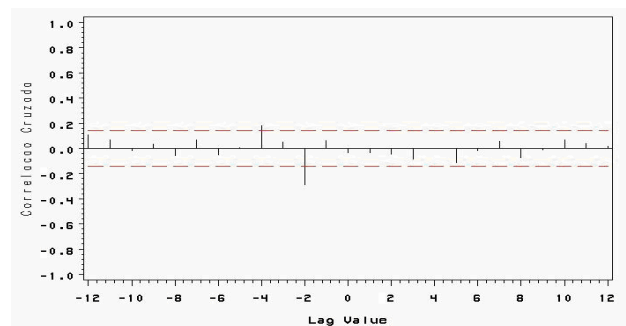


Figura 5.61 Função de Correlação Cruzada Y_2 vs X_2

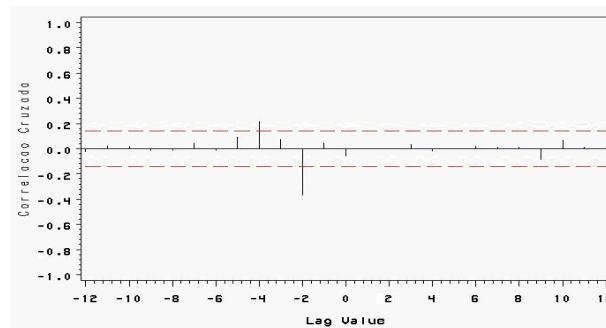


Figura 5.62 Função de Correlação Cruzada Y_2 vs X_3

Através de análise nas Funções de autocorrelação Cruzada, verificou-se os seguintes fatores:

- As variáveis resposta Y_2 (soda) e X_1 (condensado) não apresentam correlação entre si. Deste modo, essa variável pode ser descartada do modelo.
- Quanto às variáveis X_2 (vapor) e X_3 (rotação), nota-se que não apresentam correlação significativa no lag zero, indicando talvez a ocorrência de defasagem e a possível presença de modelo com **feedback** devido a presença de lags negativos com correlações significativas.

5.9.2 Ajuste do Modelo de Função de Transferência - Soda

Sabe-se que, devido os fatores observados em 5.10.1, torna-se difícil encontrar modelos de funções de transferência. Porém, foram testados vários modelos de função de transferência, sendo que, em todos os modelos houve a necessidade de se fazer o ajuste da componente de erro ou ruído no modelo final da função de transferência. Esses modelos foram apresentados na forma de operador atraso B , com as respectivas estruturas funcionais.

$$\text{Modelo 1 : } Y_{2t} = \frac{\varpi_{01}}{(1 - \delta_{11}B)} X_{2t-2} + \varpi_{03} X_{3t} + (1 - \theta B) a_t \quad (5.13)$$

$$\text{Modelo 2 : } Y_{2t} = (\varpi_{01} - \varpi_{11}B) X_{2t-1} + (\varpi_{03} - \varpi_{13}B^5) X_{3t} + (1 - \theta B) a_t \quad (5.14)$$

5.9.3 Estimativas para a Estruturação do Modelo 1 em 5.13

Tabela 5.23 Cálculo dos Parâmetros Estimados para o Modelo de F.T.

Parâmetros	Estimativas	E.P	t-valor	$\text{aprox} Pr > t $	Lag	Variável	Deslocamento
MU	0.050	0.018	2.82	0.0047	0	Y2	0
MA1,1 (θ)	-0.551	0.064	-8.65	< 0.0001	1	Y2	0
NUM1 (ϖ_{01})	-0.004	0.003	-1.50	0.1339	0	X2	2
NUM1,1 (ϖ_{11})	0.756	0.233	3.24	0.0012	1	X2	2
NUM2 (ϖ_{02})	-0.0005	0.0003	-1.84	0.0664	0	X3	0

Através das Tabela 5.23, pode-se observar que as estimativas dos parâmetros ϖ_{11} , δ_{11} , correspondente à variável X2, e à estimativa ϖ_{13} correspondente à variável X3 e à estimativa do parâmetro média móvel θ apresentam estatísticas t significativas a 5%. E, portanto, serão utilizadas para a formulação do modelo 1 de função de transferência em 5.13. Portanto, as seguintes estimativas foram obtidas para o modelo 1:

$$\text{Modelo 1 : } Y_{2t} = \frac{(-0,004)}{(1 - 0,756B)} X_{2t-2} - 0,005 X_{3t} + (1 + 0,551B) a_t \quad (5.15)$$

5.9.4 Análise Residual para o Modelo de F.T em 5.15

Tabela 5.24 *Estudo das Autocorrelação dos Resíduos do Modelo*

Lag	χ^2	Grau de Liberdade	$Pr > \chi^2$	Autocorrelações					
6	3.93	5	0.560	-0.024	-0.103	-0.068	0.032	-0.047	-0.023
12	7.26	11	0.778	-0.034	-0.048	-0.015	-0.049	0.076	-0.063
18	8.41	17	0.957	0.020	-0.026	0.048	0.038	-0.020	0.010
24	10.89	23	0.984	-0.021	-0.041	0.070	0.046	0.004	0.044

Tabela 5.25 *Estudo das Correlação Cruzada dos Resíduos vs X2*

Lag	χ^2	Grau de Liberdade	$Pr > \chi^2$	Correlação Cruzada					
5	4.18	5	0.523	0.018	-0.036	0.058	-0.117	0.036	0.039
11	7.33	11	0.772	-0.049	-0.007	0.062	0.035	0.021	-0.091
17	9.66	17	0.917	0.032	-0.063	0.022	-0.005	0.004	-0.080
23	12.37	23	0.964	0.084	-0.009	0.068	-0.041	0.010	0.020

Tabela 5.26 *Estudo das Correlação Cruzada dos Resíduos vs X3*

Lag	χ^2	Grau de Liberdade	$Pr > \chi^2$	Correlação Cruzada					
5	1.57	6	0.955	0.006	0.019	0.038	0.071	-0.006	0.033
11	5.14	12	0.953	0.027	0.017	0.013	-0.092	0.088	-0.026
17	7.30	18	0.987	0.029	-0.024	0.043	-0.059	0.057	0.031
23	8.35	24	0.999	0.044	0.018	0.024	0.032	-0.014	-0.036

Para o modelo proposto em 5.15, observa-se que a avaliação das auto-correlações dos resíduos apresentam estatísticas Q não significativas (Tabela 5.24), o que indica que os resíduos desse modelo de função de transferência são ruídos brancos. Deste modo, o modelo se ajustou adequadamente aos dados. Além do mais, o cheque da correlação cruzada dos resíduos com as variáveis X1, e X3 na entrada Tabelas 5.25 e 5.26, apresentam correlações cruzadas dos resíduos não significativas, evidenciando que as funções de transferência relacionadas a essas variáveis são adequadas.

5.9.5 Estimativas para a Estruturação do Modelo 2 em 5.14

Tabela 5.27 Cálculo dos Parâmetros Estimados para o Modelo de F.T.

Parâmetros	Estimativas	E.P	t-valor	$\text{aproxPr} > t $	Lag	Variável	Deslocamento
MU	0.044	0.012	3.69	0.0002	0	Y2	0
MA1,1 (θ)	-0.5663	0.064	-8.86	< 0.0001	1	Y2	0
NUM1 (ϖ_{01})	-0.0024	0.003	-0.73	0.4633	0	X2	2
NUM1,1 (ϖ_{11})	0.0054	0.003	1.64	0.1014	1	X2	2
NUM2 (ϖ_{02})	-0.0006	0.0003	-2.08	0.0377	0	X3	0

Através das Tabela 5.27, pode-se observar que as estimativas do parâmetro ϖ_{11} da variável X2 é significativo a 10% e a estimativa ϖ_{13} correspondente à variável X3 e à estimativa do parâmetro média móvel θ apresentam estatísticas t significativas a menos 5% e menos de 0,001%, respectivamente. E, portanto serão utilizadas para a formulação do modelo 2 de função de tranferência, descrito em 5.14. Portanto, as seguintes estimativas foram obtidas para o modelo 2:

$$\text{Modelo 2 : } Y_{2t} = (-0,0024 + 0,0054B)X_{2t-1} - 0,0006X_{3t} + (1 + 0,5663B)a_t \quad (5.16)$$

5.9.6 Análise Residual para o Modelo de F.T em 5.16

Tabela 5.28 *Estudo das Autocorrelação dos Resíduos do Modelo*

Lag	χ^2	Grau de Liberdade	$Pr > \chi^2$	Autocorrelações					
6	5.00	5	0.416	-0.022	-0.110	-0.063	0.0574	-0.072	-0.011
12	8.22	11	0.694	-0.019	-0.045	0.008	-0.049	0.078	-0.067
18	9.17	17	0.935	0.021	-0.026	0.054	0.014	-0.013	0.008
24	12.12	23	0.969	-0.020	0.024	0.088	0.041	-0.005	0.052

Tabela 5.29 *Estudo das Correlação Cruzada dos Resíduos vs X2*

Lag	χ^2	Grau de Liberdade	$Pr > \chi^2$	Correlação Cruzada					
5	4.35	5	0.500	-0.012	0.020	0.012	-0.145	0.002	0.027
11	7.71	11	0.739	-0.070	-0.010	0.051	0.036	0.005	-0.091
17	10.72	17	0.871	0.039	-0.073	0.017	0.003	0.012	-0.091
23	14.19	23	0.921	0.093	-0.014	0.074	-0.058	0.010	0.012

Tabela 5.30 *Estudo das Correlação Cruzada dos Resíduos vs X3*

Lag	χ^2	Grau de Liberdade	$Pr > \chi^2$	Correlação Cruzada					
5	2.10	6	0.910	0.012	0.022	0.023	0.094	-0.024	0.012
11	5.84	12	0.924	0.014	0.003	0.002	-0.100	0.087	-0.035
17	8.7	18	0.960	0.031	-0.033	0.056	-0.079	0.058	0.033
23	10.50	24	0.992	0.049	0.008	0.034	0.032	-0.020	-0.052

Para o modelo proposto na (Eq. 5.16), observa-se que a avaliação das auto-correlações dos resíduos apresentam estatísticas Q não significativas o que indica que os resíduos desse modelo de função de transferência são ruídos brancos Tabela 5.28. Deste modo, o modelo se ajustou adequadamente aos dados. Além do mais, o cheque da correlação cruzada dos resíduos com as variáveis X1, e X3 na entrada Tabela 5.29 e 5.30, apresentam correlações cruzadas dos resíduos não significativas evidenciando que as funções de transferências relacionadas a essas variáveis são adequadas.

5.9.7 Seleção do Melhor Modelo para o Nível de Soda

Ao comparar os ajustes fornecidos pelos dois modelos Figuras 5.63 e 5.64, verifica-se que praticamente não existe distinção entre os mesmos, ou seja, ambos apresentam a mesma capacidade de explicar a dinâmica da variável nível de soda no processo.

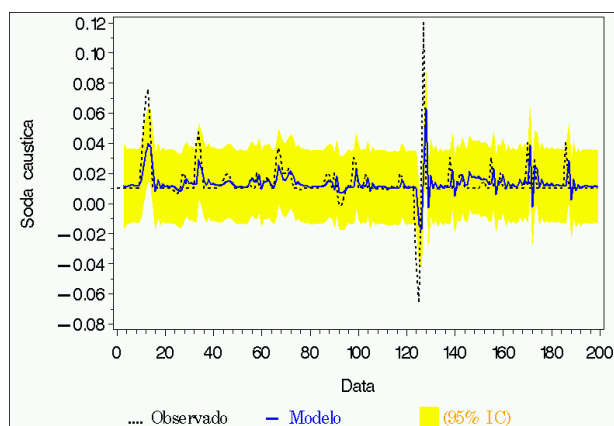


Figura 5.63 Ajuste do Modelo 5.15 de Função de Transferência

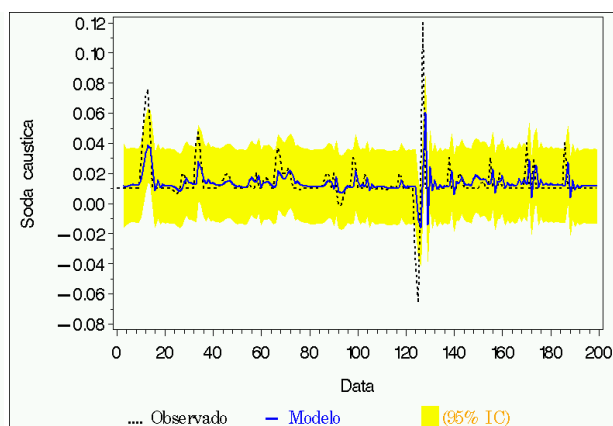


Figura 5.64 Ajuste do Modelo 5.16 de Função de Transferência

A mesma conclusão é verificada quando compara-se os critérios de **Akaike** e **SBC**, apresentados na Tabela 5.31.

Tabela 5.31 Critérios para Seleção do Melhor Modelo Ajustado

Critério	Modelo 1	Modelo 2
AIC	-1165.60	-1165.46
SBC	-1149.18	-1149.05
MSE	0.00015	0.00015
RMSE	0.01221	0.01221

5.9.8 Teste de Validação para os Modelos do Nível de Soda

Para a avaliação desses modelos, ou seja, para a validação, foi aplicado os mesmos critérios utilizados no processo de validação do modelo para a variável nível de umidade. Constatou-se que em todos os casos as funções de transferências relacionadas foram adequadas, isto é, os resíduos apresentaram-se não correlacionados e as correlações cruzadas entre os resíduos e as variáveis de controle foram não significativas. Nas Figuras 5.65 e 5.66, são apresentados os resultados de cada modelo e na Tabela 5.32 estão disponíveis os critérios utilizados na seleção.

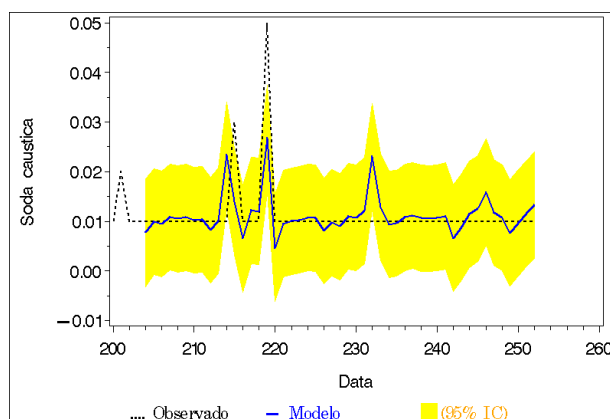
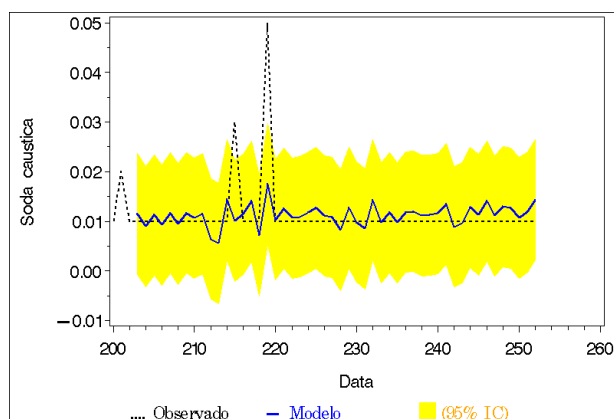


Figura 5.65 Teste de Validação do 1º Modelo

Figura 5.66 Teste de Validação do 2º Modelo

Tabela 5.32 Critérios para Seleção do Melhor Modelo Validado

Critério	Modelo 1	Modelo 2
AIC	-360.878	-366.857
SBC	-349.406	-357.397
MSE	0.00003	0.000027
RMSE	0.00550	0.005192

Para o ajuste em validação, pode-se observar que o modelo 2 descrito em 5.16, apresentou melhores resultados.

5.10 Conclusões

Levando em consideração todos os testes de ajuste e validação realizados para os modelos, assim como, os critérios de seleção, pode-se concluir em relação aos modelos de função de transferência ajustados que:

- O modelo que melhor representa o real comportamento do padrão da variável nível de umidade presente na polpa de cristais de hidrato no processo em estudo é o modelo descrito na (Eq. 5.12).

$$Y_{1t} = \frac{(-0,003 - 0,027B)}{(1 - 0,525B)}X_{1t} + (0,0004 + 0,053B^5)X_{3t} + (1 + 0,372B)a_t \quad (5.17)$$

- Já para a variável controle do nível de soda cáustica presente na polpa de cristais de hidrato, constatou-se que qualquer modelo dentre os ajustados pode representar o comportamento desta variável dentro do processo, pois devido os critérios AIC e BIC ambos conseguiram capturar o padrão gerador do processo. São eles:

$$Y_{2t} = \frac{-0,004}{(1 - 0,756B)}X_{2t-2} - 0,005X_{3t} + (1 + 0,551B)a_t \quad (5.18)$$

$$Y_{2t} = (-0,0024 + 0,0054B)X_{2t-1} - 0,0006X_{3t} + (1 + 0,5663B)a_t \quad (5.19)$$

No entanto, o modelo descrito na (Eq. 5.19) apresentou melhores respostas em validação. Assim, este modelo será utilizado na avaliação comparativa de resultados entre as classes de modelos dinâmicos dentro das metodologias estudadas.

Portanto, baseado nos ajustes de validação dos modelos estudados, conclui-se que o modelo dinâmico multivariado que melhor capturou o padrão gerador do processo de filtração e secagem da polpa de cristais de hidrato foi o modelo de Rede Neural Multicamada Direta com algoritmo de treinamento Backpropagation. No entanto, é importante resaltar que ambos os modelos foram significativos em relação ao objetivo do trabalho, e para a futura implementação de um desses modelos no controle do processo na planta, é de muitíssima importância levar em consideração o estudo de viabilidade de ambas as situações.

Capítulo 6

Considerações Finais

O referente trabalho cumpriu com o propósito inicial especificado em seu objetivo, ou seja, propor modelos dinâmicos multivariados relacionando a sua viabilidade no monitoramento da filtração e secagem da polpa de cristais de hidrato no processo de fabricação da alumina e dentre estes, selecionar o mais significativo. Para tal, foram avaliadas as aproximações entre as respostas originais do processo e as fornecidas através da utilização das metodologias de modelagem: Rede Neural Artificial e Modelos ARIMA Multivariado com auxílio de Função de Transferência.

É muito importante ressaltar que as metodologias de séries temporais e redes neurais são bem diferentes, no sentido de que, em redes neurais, procura-se recuperar a série em estudo sem a preocupação de separar o sinal do ruído. Assim, em redes neurais praticamente toda a série, incluindo a informação e o ruído, é recuperada enquanto na metodologia de séries temporais procura-se recuperar somente o sinal ou a informação. Neste sentido, o erro quadrático médio pode não ser adequado para comparação dos métodos, uma vez que ele é resultado da diferença média entre o modelo ajustado e a série observada.

Além deste fato, em Modelos de Séries Temporais com Função de Transferência, deve-se levar em consideração a elevada variabilidade presente nas séries. Nesse caso, para trabalhos futuros talvez fosse interessante a aplicação de modelos que capturem, além do nível da série, também a sua variabilidade, ou seja, recomenda-se a aplicação de modelos da classe dos ***ARCH***, ou seja, modelos ***Auto-regressivos com heterocedasticidade condicional generalizado GARCH multivariados***. E, devido a ocorrência de modelos tipo ***feedback***, talvez sejam obtidos melhores resultados com a aplicação aos dados de modelos da classe dos modelos ***Auto-regressivo vetorial VAR***. No entanto, não se trabalhou estes tipos de modelos devido à proposta inicial do trabalho, em explorar modelos com funções de transferência.

Finalizando, é muito importante deixar claro que, isso não implica que a estrutura de rede neural proposta, assim como, o modelo de função de transferência ajustados são os melhores e mais adequados para controle deste processo pois, pode-se para os mesmos dados, ajustar redes com estruturas ainda mais significativas porém com outros padrões e parâmetros de treinamento e validação, ou ainda com outras dinâmicas, não se encaixando mais a este tipo de Rede Neural.

Apêndice A

ANEXO

A.1 Introdução

Neste apêndice, será apresentado, além das tabelas referentes ao banco de dados os algoritmos implementados para a obtenção das redes neurais e do modelo de séries temporais multivariado.

A.1.1 Algoritmo para Obtenção de Rede Neural - MATLAB 7.0

4

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%  
NORMALIZAÇÃO DAS VARIÁVEIS POR DESVIO PADRÃO  
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%  
  
condnorm=(cond-mean(cond))/(std(cond));  
  
vapnorm=(vap-mean(vap))/(std(vap));  
  
rotnorm=(rot-mean(rot))/(std(rot));  
  
umidnorm=(umid-mean(umid))/(std(umid));  
  
sodnorm=(sod-mean(sod))/(std(sod));  
  
condnormv=(condv-mean(condv))/(std(condv));  
  
vapnormv=(vapv-mean(vapv))/(std(vapv));  
  
rotnormv=(rotv-mean(rotv))/(std(rotv));  
  
umidnormv=(umidv-mean(umidv))/(std(umidv));  
  
sodnormv=(sodv-mean(sodv))/(std(sodv));  
  
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%  
DESNORMALIZAÇÃO DO DESVIO  
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%  
  
umid_desnorpad=(Umidade_sai_rede_teste*std(umidv))+ mean(umidv);  
sod_desnorpad= (soda_sai_rede_teste*std(sodv))+ mean(sodv);
```

Figura A.1 Gera a Normalização por Desvio Padrão

```

clc
clear
load filtro_mod
pa1=now;
N=498;
Inicio=datestr(now)
obj = findobj(gcf, 'Tag', 'epochs');
epochs = str2double(get(obj,'String'));
obj = findobj(gcf, 'Tag', 'goal');
goal = str2double(get(obj,'String'));
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
ESCOLHA DOS PADROES APRESENTADOS A REDE
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

obj = findobj(gcf, 'Tag', 'nap');
nap = str2double(get(obj,'String'));

obj = findobj(gcf, 'Tag', 'ndp');
ndp = str2double(get(obj,'String'));

obj = findobj(gcf, 'Tag', 'nav');
nav = str2double(get(obj,'String'));

obj = findobj(gcf, 'Tag', 'ndv');
ndv = str2double(get(obj,'String'));

space = 1; % índice que regula os valores que serão apresentados no treinamento na rede

i = [1:space:(ndp-nap)];
iv = [1:space:(ndv-nav)];
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
VARIÁVEIS DO PROCESSO
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
cond=condensado(nap:ndp);
vap=vapor(nap:ndp);
rot=rotacao(nap:ndp);
umid=umidade(nap:ndp);
sod=soda(nap:ndp);
[m,n] = size(cond);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%ENTRADAS PARA VALIDAÇÃO.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
condv=condensado(nav:ndv);
vapv=vapor(nav:ndv);
rotv=rotacao(nav:ndv);
umidv=umidade(nav:ndv);
sodv=soda(nav:ndv);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
TESTE PARA VERIFICAR SE O TREINO ESTA SENDO EXECUTADO
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

cond = cond(i);
vap = vap(i);
rot = rot(i);
umid = umid(i);
sod = sod(i);
condv = condv(iv);
vapv = vapv(iv);
rotv = rotv(iv);
umidv = umidv(iv);
sodv = sodv(iv);

```

Figura A.2 *Padrões apresentados a rede*

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
VETORES A SEREM APRESENTADOS A REDE
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

P = [condnorm;vapnorm;rotnorm];
T = [umidnorm;sodnorm];
v.P = [condnormv;vapnormv;rotnormv];
v.T = [umidnormv;sodnormv];

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
FUNÇÃO DE TREINAMENTO DA REDE
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
net = newff([-1 1;-1 1;-1 1],[14 12 2],{'logsig' 'logsig' 'purelin'},'trainlm'); % 1ª Rede 1º momento
net = newff([-1 1;-1 1;-1 1],[12 10 2],{'tansig' 'tansig' 'purelin'},'trainlm'); % 2ª Rede 2º momento
net = newff([-1 1;-1 1;-1 1],[12 10 8 2],{'tansig' 'tansig' 'tansig' 'purelin'},'trainlm'); % 3ª Rede 3º momento

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
PARAMETROS DE TREINAMENTO - 3ª REDE
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

net.trainParam.show = 50; % Resposta da rede de 50 em 50 épocas
net.trainParam.epochs = epochs; % N° de épocas para o treinamento e validação
net.trainParam.goal = goal; % Objetivo do erro a ser alcançado pela rede
net.trainParam.lr = 0.07; % Taxa de aprendizagem (0.01 a 1.5)
net.trainParam.mc = 0.8; % Taxa de momentum (0.7 a 0.9)

% net.trainParam.lr_inc = 1.05; % Taxa de Incremento da lr
% net.trainParam.lr_dec = 0.7; % Taxa de decremento lr
% net.trainParam.max_perf_inc = 1.05; % Incremento máximo do erro
% net.performFcn = 'mse'; %
% net.performParam.ratio = 0.5; %

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
EXECUTA O TREINAMENTO
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

[net,tr]=train(net,p,t,[ ],v);
[net,tr]=train(net,p,t);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
SALVA OS PESOS
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

peso_entrada = net.IW;
peso_intermed = net.LW;
peso_bias = net.b;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
EXECUTA TESTE E VALIDAÇÃO
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

sai_rede=sim(net,p);
sai_rede_testev=sim(net,v.P);

```

Figura A.3 *Função e parâmetros de treinamento apresentados a rede*

3

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
NORMALIZAÇÃO DAS VARIÁVEIS NO ESPAÇO (-1, 1)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

ymax = 1;
ymin = -1;

% variável condensado
condmax=max(cond);
condmin=min(cond);
condmaxv=max(condv);
condminv=min(condv);
condnorm=((ymax-ymin)*(cond-condmin)/(condmax-condmin))+ymin;
condnormv=((ymax-ymin)*(condv-condminv)/(condmaxv-condminv))+ymin;

% variável vapor
vapmax=max(vap);
vapmin=min(vap);
vapmaxv=max(vapv);
vapminv=min(vapv);
vapnorm=((ymax-ymin)*(vap-vapmin)/(vapmax-vapmin))+ymin;
vapnormv=((ymax-ymin)*(vapv-vapminv)/(vapmaxv-vapminv))+ymin;

% variável rotação
rotmax=max(rot);
rotmin=min(rot);
rotmaxv=max(rotv);
rotminv=min(rotv);
rotnorm=((ymax-ymin)*(rot-rotmin)/(rotmax-rotmin))+ymin;
rotnormv=((ymax-ymin)*(rotv-rotminv)/(rotmaxv-rotminv))+ymin;

% variável umidade
umidmax=max(umid);
umidmin=min(umid);
umidmaxv=max(umidv);
umidminv=min(umidv);
umidnorm=((ymax-ymin)*(umid-umidmin)/(umidmax-umidmin))+ymin;
umidnormv=((ymax-ymin)*(umidv-umidminv)/(umidmaxv-umidminv))+ymin;

% variável soda
sodmax=max(sod);
sodmin=min(sod);
sodmaxv=max(sodv);
sodminv=min(sodv);
sodnorm=((ymax-ymin)*(sod-sodmin)/(sodmax-sodmin))+ymin;
sodnormv=((ymax-ymin)*(sodv-sodminv)/(sodmaxv-sodminv))+ymin;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
DESNORMALIZAÇÃO DO NO ESPAÇO (-1, 1)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

umid1=((Umidade_sai_rede_teste + 1)*(umidmax-umidmin))+2*umidmin)/2;
sod1=((soda_sai_rede_teste + 1)*(sodmax-sodmin))+2*sodmin)/2;

```

Figura A.4 Gera a Normalização no Espaço (-1, 1)

5

Figura A.5 *Gera a Análise Gráfica de Treinamento*

6

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
AGORA COMECA A ANÁLISE GRÁFICA DE VALIDAÇÃO
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

figure(5)
    Umidade_desejv=(umidnormv);
    plot(Umidade_desejv,'r')
    hold on
    Umidade_sai_rede_testev=(sai_rede_testev(1,:));
    plot(Umidade_sai_rede_testev,'k')
    xlabel('Amostra para validacao do nível de umidade')
    ylabel('Saída de Validação')
    title('SAIDA DESEJADA(Vermelho) E SAIDA DA REDE(Preto)')
    grid on
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

figure(6)
    erro_saida_umidadev=(Umidade_desejv-Umidade_sai_rede_testev);
    plot(erro_saida_umidadev,'m')
    grid on
    title('SAIDA DE VAL. REDE - SAIDA DE VAL. DESEJADA')
    ylabel('Erro Quadrático Médio de Validação ')
    xlabel('Amostra de Umidade')
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

figure(7)

    soda_desejv=(sodnormv);
    plot(soda_desejv,'r')
    hold on
    soda_sai_rede_testev=(sai_rede_testev(2,:));
    plot(soda_sai_rede_testev,'k')
    xlabel('Amostra para validacao do nível de soda')
    ylabel('Saída de Validação')
    title('SAIDA DESEJADA(Vermelho) E SAIDA DA REDE(preto)')
    grid on
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

figure(8)

    erro_saida_soda2=(soda_desejv-soda_sai_rede_testev);
    plot(erro_saida_soda2,'m')
    grid on
    title('SAIDA DE VAL. REDE - SAIDA DE VAL. DESEJADA')
    ylabel('Erro de Validação E(W)')
    xlabel('Amostra de Soda')
    grid on

```

Figura A.6 Gera a Análise Gráfica de Validação

7

%%
AGORA COMECA A ANÁLISE GRÁFICA DE VALIDAÇÃO DESNORMALIZADA
 %%%

figure (9)

```

plot(umid,'r')
hold on
plot(umid1,'k');
xlabel('Amostra para validacao do nível de umidade')
ylabel('Saída de Validação Desnormalizada')
title('SAIDA DESEJADA(Vermelho) E SAIDA DA REDE(preto)')
grid on
  
```

figure (10)

```

plot(sod,'r')
hold on
plot(sod1,'k');
xlabel('Amostra para validacao do nível de soda')
ylabel('Saída de Validação Desnormalizada')
title('SAIDA DESEJADA(Vermelho) E SAIDA DA REDE(preto)')
grid on
  
```

%%
GERA OS RESULTADOS DE VALIDAÇÃO (Nível de Umidade e Nível de Soda)
 %%%

```

a = umid1
b = sod1
  
```

Figura A.7 *Gera a Análise Gráfica da Validação Desnormalizada*


```

function varargout = indexcaustico(varargin)
gui_Singleton = 1;
gui_State = struct('gui_Name',    mfilename, ...
    'gui_Singleton',  gui_Singleton, ...
    'gui_OpeningFcn', @indexcaustico_OpeningFcn, ...
    'gui_OutputFcn',  @indexcaustico_OutputFcn, ...
    'gui_LayoutFcn',  [ ], ...
    'gui_Callback',   [ ]);
if nargin & isstr(varargin{1})
    gui_State.gui_Callback = str2func(varargin{1});
end
if nargout
    [varargout{1:nargout}] = gui_mainfcn(gui_State, varargin{:});
else
    gui_mainfcn(gui_State, varargin{:});
end
function indexcaustico_OpeningFcn(hObject, eventdata, handles, varargin)
handles.output = hObject;
guidata(hObject, handles);

function varargout = indexcaustico_OutputFcn(hObject, eventdata, handles)
varargout{1} = handles.output;
function epochs_CreateFcn(hObject, eventdata, handles)
if ispc
    set(hObject,'BackgroundColor','white');
else
    set(hObject,'BackgroundColor',get(0,'defaultUicontrolBackgroundColor'));
end
function goal_CreateFcn(hObject, eventdata, handles)
if ispc
    set(hObject,'BackgroundColor','white');
else
    set(hObject,'BackgroundColor',get(0,'defaultUicontrolBackgroundColor'));
end

function nap_CreateFcn(hObject, eventdata, handles)
if ispc
    set(hObject,'BackgroundColor','white');
else
    set(hObject,'BackgroundColor',get(0,'defaultUicontrolBackgroundColor'));
end
function ndp_CreateFcn(hObject, eventdata, handles)
if ispc
    set(hObject,'BackgroundColor','white');
else
    set(hObject,'BackgroundColor',get(0,'defaultUicontrolBackgroundColor'));
end
function nav_CreateFcn(hObject, eventdata, handles)
if ispc
    set(hObject,'BackgroundColor','white');
else
    set(hObject,'BackgroundColor',get(0,'defaultUicontrolBackgroundColor'));
end
function ndv_CreateFcn(hObject, eventdata, handles)
if ispc
    set(hObject,'BackgroundColor','white');
else
    set(hObject,'BackgroundColor',get(0,'defaultUicontrolBackgroundColor'));
end

```

Figura A.8 *Algoritmo para iniciar o treinamento e validação*

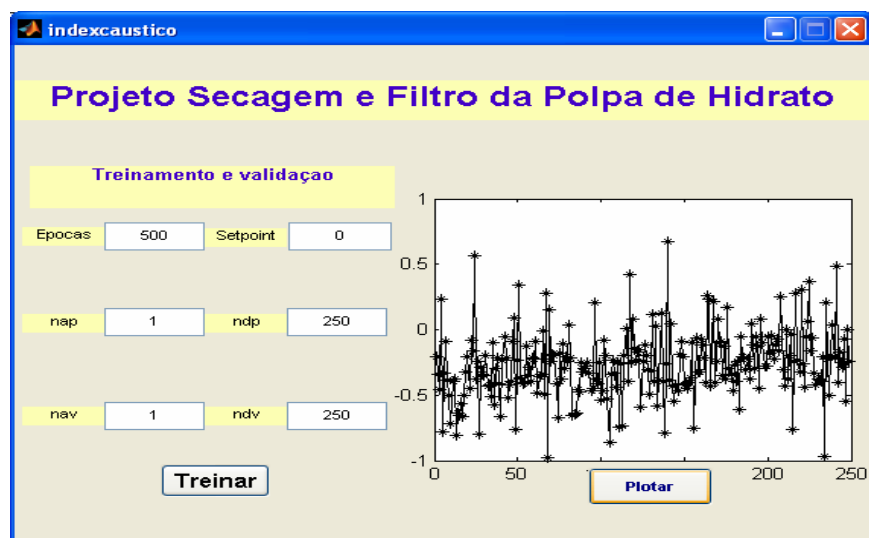


Figura A.9 Projeto para seleção dos dados de treinamento e validação

A.1.2 Algoritmo para Obtenção do Modelo de Séries Temporais - SAS-9

```
data albras;
  infile "e:\felix\albras100\hidrato100.txt" expandtabs;
  input t y x1 x2 x3 x31;
  *   format date date7.;
  datalines;
ods html file='albras';
ods graphics on;
proc arima ;
  i var=x1 nlag=12 outcov=x1out;
  e q=1 p=1 method=ml plot;
  i var=x2 nlag=12 outcov=x2out;
  e q=1 p=1 method=ml plot;
  i var=x31 nlag=12 outcov=x3out;
  e q=1 p=(1)(3) method=ml plot;
run;
ods html close;
quit;
```

Figura A.10 *Gera as Séries das Variáveis*

```

data albras;
infile "g:\felix\hidsod\hidsod200.txt"
expandtabs;
input t y1 y2 x1 x2 x3 @@ ;
label t
      y1      ='Umidade'
      y2      ='soda'
      x1      ='condensado'
      x2      ='vapor'
      x3      ='rotacao';

run;
ods html file='albras';
ods graphics on;
proc arima ;
  i var=y1 nlag=15 outcov=yresp;
  i var=x1 nlag=15 outcov=x1out;
  i var=x2 nlag=15 outcov=x2out;
  i var=x3 nlag=15 outcov=x3out;
run;
data acf1;
  set yresp;
  lcl=-2*stderr;      ucl=2*stderr;
run;
proc gplot data=acf1;
plot corr*lag=1
      lcl*lag=2
      ucl*lag=2/overlay vminor=4 vaxis=axis1
haxis=axis2 ;
axis1 order=(-1.0 to 1.0 by 0.5)
label=(angle=90 'FAC Umidade') ;
axis2 order=(1.0 to 15 by 1.0) label=('lag
(k)');
symbol1 color=black interpol=needle
value=None width=4;
symbol2 color=black interpol=join value=None
width=1;
*title h=2 "FAC da resposta umidade";
run;
proc gplot data=acf1;
plot partcorr*lag=1
      lcl*lag=2
      ucl*lag=2/overlay vminor=4 vaxis=axis1
haxis=axis2 ;
axis1 order=(-1.0 to 1.0 by 0.5)
label=(angle=90 'FACP Umidade') ;
axis2 order=(1.0 to 15 by 1.0) label=('lag
(k)');
symbol1 color=black interpol=needle
value=None width=4;
symbol2 color=black interpol=join value=None
width=1;
*title h=2 "FACP da resposta umidade";
run;
data acf2;
  set x1out;
  lcl=-2*stderr;      ucl=2*stderr;
run;
proc gplot data=acf2;
plot corr*lag=1
      lcl*lag=2
      ucl*lag=2/overlay vminor=4 vaxis=axis1
haxis=axis2 ;
axis1 order=(-1.0 to 1.0 by 0.5)
label=(angle=90 'FAC X1') ;
axis2 order=(1.0 to 15 by 1.0) label=('lag
(k)');
symbol1 color=black interpol=needle
value=None width=4;
symbol2 color=black interpol=join value=None
width=1;
*title h=2 "FAC de X1";
run;
proc gplot data=acf2;
plot partcorr*lag=1
      lcl*lag=2
      ucl*lag=2/overlay vminor=4 vaxis=axis1
haxis=axis2 ;
axis1 order=(-1.0 to 1.0 by 0.5)
label=(angle=90 'FACP X1') ;
axis2 order=(1.0 to 15 by 1.0) label=('lag
(k)');
symbol1 color=black interpol=needle
value=None width=4;
symbol2 color=black interpol=join value=None
width=1;
*title h=2 "FACP de X1";
run;
ods html close;
quit;

symbol1 color=black interpol=needle
value=None width=4;
symbol2 color=black interpol=join value=None
width=1;
*title h=2 "FACP de X1";
run;
data acf3;
  set x2out;
  lcl=-2*stderr;      ucl=2*stderr;
run;
proc gplot data=acf3;
plot corr*lag=1
      lcl*lag=2
      ucl*lag=2/overlay vminor=4 vaxis=axis1
haxis=axis2 ;
axis1 order=(-1.0 to 1.0 by 0.5)
label=(angle=90 'FAC X2') ;
axis2 order=(1.0 to 15 by 1.0) label=('lag
(k)');
symbol1 color=black interpol=needle
value=None width=4;
symbol2 color=black interpol=join value=None
width=1;
*title h=2 "FAC de X2";
run;
proc gplot data=acf3;
plot partcorr*lag=1
      lcl*lag=2
      ucl*lag=2/overlay vminor=4 vaxis=axis1
haxis=axis2 ;
axis1 order=(-1.0 to 1.0 by 0.5)
label=(angle=90 'FACP X2') ;
axis2 order=(1.0 to 15 by 1.0) label=('lag
(k)');
symbol1 color=black interpol=needle
value=None width=4;
symbol2 color=black interpol=join value=None
width=1;
*title h=2 "FACP de X2";
run;
data acf4;
  set x3out;
  lcl=-2*stderr;      ucl=2*stderr;
run;
proc gplot data=acf4;
plot corr*lag=1
      lcl*lag=2
      ucl*lag=2/overlay vminor=4 vaxis=axis1
haxis=axis2 ;
axis1 order=(-1.0 to 1.0 by 0.5)
label=(angle=90 'FAC X3') ;
axis2 order=(1.0 to 15 by 1.0) label=('lag
(k)');
symbol1 color=black interpol=needle
value=None width=4;
symbol2 color=black interpol=join value=None
width=1;
*title h=2 "FAC de X3";
run;
proc gplot data=acf4;
plot partcorr*lag=1
      lcl*lag=2
      ucl*lag=2/overlay vminor=4 vaxis=axis1
haxis=axis2 ;
axis1 order=(-1.0 to 1.0 by 0.5)
label=(angle=90 'FACP X3') ;
axis2 order=(1.0 to 15 by 1.0) label=('lag
(k)');
symbol1 color=black interpol=needle
value=None width=4;
symbol2 color=black interpol=join value=None
width=1;
*title h=2 "FACP de X3";
run;
ods html close;
quit;

```

Figura A.11 Gera a FAC e FACP das Séries

```

data albras;
infile "c:\pendrive\felix\hidsod\hidsod200.txt" expandtabs;
input t y1 y2 x1 x2 x3 @@ ;
label t      ='tempo em horas'
      y1      ='Umidade'
      y2      ='soda'
      x1      ='condensado'
      x2      ='vapor'
      x3      ='rotacao';
datalines;
run;
ods html file='albras2';
ods graphics on;
goptions cback=white colors=(black) border reset=(axis symbol);
axis1 offset=(1 cm)
      label=('Período') minor=none;
axis2 label=(angle=90 'Umidade');
symbol1 i=join;
proc gplot ;
  plot y1*t / haxis=axis1
        vaxis=axis2
        vminor=1;
  title 'Umidade na torta de hidrato - Y1';
axis2 label=(angle=90 'Resposta');
run;
proc gplot ;
  plot y2*t / haxis=axis1
        vaxis=axis2
        vminor=1;
  title 'Soda caustica - Y2';
axis2 label=(angle=90 'Resposta');
run;
proc gplot ;
  * format date date7.;
  plot x1*t / haxis=axis1
        vaxis=axis2
        vminor=1;
  title 'Vazao de condensado - X1';
axis2 label=(angle=90 'X1');
run;
proc gplot ;
  * format date date7.;
  plot x2*t / haxis=axis1
        vaxis=axis2
        vminor=1;
  title 'Vazao de vapor - X2';
axis2 label=(angle=90 'X2');
run;
proc gplot ;
  * format date date7.;
  plot x3*t / haxis=axis1
        vaxis=axis2
        vminor=1;
  title 'Rotacao do tambor - X3';
axis2 label=(angle=90 'X3');
run;
ods html close;
quit;

```

Figura A.12 Gera os Modelos ARIMA

```

data albras;
infile "c:\pendrive\felix\hidsod\hidsod052.txt" expandtabs;
input t y1 y2 x1 x2 x3 @@ ;
label t      ='tempo em horas'
      y1     ='Umidade'
      y2     ='soda'
      x1     ='condensado'
      x2     ='vapor'
      x3     ='rotacao';
ods html file='albras';
ods graphics on;
proc arima ;
* i var=x1 outcov=x1out ;
* e p=2 method=ml noprint ;
i var=x2 outcov=x2out ;
e q=2 method=ml noprint;
i var=x3 outcov=x3out ;
e q=1 method=ml noprint;
run;
i var=y2 crosscor=(x2 x3) nlag=12 ;
run;
e input= ( 2$(1) x2 2$ x3) q=1
method=ml ;
forecast lead=0 out=saida;
run;
proc forecast data=saida outest=estim;
run;
proc print data=estim;
title 'parametros de saida';
run;
data b;
merge albras saida;
keep t x1 x2 x3 y2 forecast 195 u95;
run;
proc print data=b;
options reset=global gunit=pct border cback=white
        colors=(white yellow) ctext=black
        ftitle=swissb ftext=swiss htitle=4 htext=4;
proc gplot data=b;
plot 195*t      =3
      u95*t     =3
      y2*t      =1
      forecast*t=2/overlay
                                haxis=axis1
                                hminor=4
                                vaxis=axis2
                                vminor=1
                                caxis=black
                                areas=2;

axis1 label=('Data') ;
axis2 label=(angle=90 'Soda caustica');
symbol1 i=2 w=2.5 color=black i=join ;
symbol2 w=2 i=join c=blue;
symbol3 l=20 i=join c=green;

footnotel c=black f=centx      ' .... Observado'
          c=b f=centx          ' - Modelo'
          c=orange f=centx     '(95% IC)';
run;
ods html close;
quit;

```

Figura A.13 Gera os Testes dos Modelos ARIMA

```

data albras;
infile "c:\pendrive\felix\hidsod\hidsod200.txt" expandtabs;
input t y1 y2 x1 x2 x3 @@ ;
label t      ='tempo em horas'
      y1      ='Umidade'
      y2      ='soda'
      x1      ='condensado'
      x2      ='vapor'
      x3      ='rotacao';
ods html file='albras';
ods graphics on;
proc arima ;
  i var=y1;
  e q=1 method=ml;
  i var=x1;
  e p=2 method=ml;
  i var=x2;
  e q=2 method=ml;
  i var=x3;
  e q=1 method=ml;
run;
  i var=y1 crosscor=(x1) nlag=12 outcov=cross1;
  i var=y1 crosscor=(x2) nlag=12 outcov=cross2;
  i var=y1 crosscor=(x3) nlag=12 outcov=cross3;
run;
data crosscorr1;
  set cross1;
keep obs crossvar stderr corr partcorr lag;
if invcorr=.;
lcl=-1.96*stderr;
ucl=1.96*stderr;
keep lcl ucl ;
run;
proc gplot;
  axis1 label=(angle=90 'Correlacao Cruzada') order=-1 to 1 by 0.2 length=3 in;
  axis2 length=5 in order=-12 to 12 by 2;
  plot (corr ucl lcl)*lag / overlay vaxis=axis1 haxis=axis2 vref=0;
  symbol1 i=needle color=black;
  symbol2 v=none l=21 i=join color=red;
  symbol3 v=none l=21 i=join color=red;
  title 'Correlacao Cruzada Y1 vs. X1';
run;
data crosscorr2;
  set cross2;
keep obs crossvar stderr corr partcorr lag;
if invcorr=.;
lcl=-1.96*stderr;
ucl=1.96*stderr;
keep lcl ucl ;
run;
proc gplot;
  axis1 label=(angle=90 'Correlacao Cruzada') order=-1 to 1 by 0.2 length=3 in;
  axis2 length=5 in order=-12 to 12 by 2;
  plot (corr ucl lcl)*lag / overlay vaxis=axis1 haxis=axis2 vref=0;
  symbol1 i=needle color=black;
  symbol2 v=none l=21 i=join color=red;
  symbol3 v=none l=21 i=join color=red;
  title 'Correlacao Cruzada Y1 vs. X2';
run;
data crosscorr3;
  set cross3;
keep obs crossvar stderr corr partcorr lag;
if invcorr=.;
lcl=-1.96*stderr;
ucl=1.96*stderr;
keep lcl ucl ;
run;
proc gplot;
  axis1 label=(angle=90 'Correlacao Cruzada') order=-1 to 1 by 0.2 length=3 in;
  axis2 length=5 in order=-12 to 12 by 2;
  plot (corr ucl lcl)*lag / overlay vaxis=axis1 haxis=axis2 vref=0;
  symbol1 i=needle color=black;
  symbol2 v=none l=21 i=join color=red;
  symbol3 v=none l=21 i=join color=red;
  title 'Correlacao Cruzada Y1 vs. X3';
run;
ods html close;
quit;

```

Figura A.14 Gera as Funções de Correlação Cruzada

```

data albras;
infile "c:\pendrive\felix\hidsod\hidsod200.txt" expandtabs;
input t y1 y2 x1 x2 x3 @@ ;
label t      ='tempo em horas'
      y1     ='Umidade'
      y2     ='soda'
      x1     ='condensado'
      x2     ='vapor'
      x3     ='rotacao';
ods html file='albras';
ods graphics on;
proc arima ;
  i var=y1 outcov=y1out ;
  e q=1 method=ml noprint ;
  i var=x1 outcov=x1out ;
  e p=2 method=ml noprint ;
  i var=x2 outcov=x2out ;
  e q=2 method=ml noprint;
  i var=x3 outcov=x3out ;
  e q=1 method=ml noprint;
run;
  i var=y1 crosscor=(x1 x2 x3) nlag=12 ;
run;
  e input=(1$(1) x1 x2 (5) x3) q=1
  method=ml plot;
  forecast lead=0 out=saida;
run;
proc forecast data=saida outest=estim;
run;
proc print data=estim;
  title 'parametros de saida';
run;
data b;
  merge albras saida;
  keep t x1 x2 x3 y1 forecast 195 u95;
run;
proc print data=b;
options reset=global gunit=pct border cback=white
        colors=(white yellow) ctext=black
        ftitle=swissb ftext=swiss htitle=4 htext=4;
proc gplot data=b;
plot    195*t      =3
        u95*t      =3
        y1*t       =1
        forecast*t=2/overlay
                                haxis=axis1
                                hminor=4
                                vaxis=axis2
                                vminor=1
                                caxis=black
                                areas=2;

axis1 label=('Data') ;
axis2 label=(angle=90 'Umidade');
symbol1 l=2 w=2.5 color=black i=join ;
symbol2 w=2 i=join c=blue;
symbol3 l=20 i=join c=green;

footnotel c=black f=centx      ' .... Observado'
          c=b f=centx         ' - Modelo'
          c=orange f=centx     '(95% IC)';
run;
ods html close;
quit;

```

Figura A.15 *Teste do Modelo 1 de F.T - Umidade*


```

data albras;
infile "g:\felix\hidsod\hidsod200.txt" expandtabs;
input t y1 y2 x1 x2 x3 @@ ;
label t          ='tempo em horas'
      y1         ='Umidade'
      y2         ='soda'
      x1         ='condensado'
      x2         ='vapor'
      x3         ='rotacao';
ods html file='albras';
ods graphics on;
proc arima ;
  i var=x1 outcov=x1out ;
  e p=2 method=ml noprint ;
  i var=x2 outcov=x2out ;
  e q=2 method=ml noprint;
  i var=x3 outcov=x3out ;
  e q=1 method=ml noprint;
run;
  i var=y1 crosscor=(x1 x3) nlag=12 ;
run;
  e input=((1)/(1) x1 (5) x3) q=1
  method=ml plot;
  forecast lead=0 out=saida;
run;
proc forecast data=saida outest=estim;
run;
proc print data=estim;
  title 'parametros de saida';
run;
data b;
  merge albras saida;
  keep t x1 x2 x3 y1 forecast l95 u95;
run;
proc print data=b;
goptions reset=global gunit=pct border cback=white
  colors=(white yellow) ctext=black
  ftitle=swissb ftext=swiss htitle=4 htext=4;
proc gplot data=b;
plot
  195*t      =3
  u95*t      =3
  y1*t       =1
  forecast*t=2/overlay
                                haxis=axis1
                                hminor=4
                                vaxis=axis2
                                vminor=1
                                caxis=black
                                areas=2;

*title 'Albras Data';
axis1 label=('Data') ;
axis2 label=(angle=90 'Umidade');
symbol1 l=2 w=2.5 color=red i=join ;
symbol2 w=2 i=join c=blue;
symbol3 l=20 i=join c=green;

footnotel c=r f=centx      ' ... Observado'
          c=b f=centx      ' - Previsao'
          c=yellow f=centx  '(95% IC)';
run;
ods html close;
quit;

```

Figura A.16 *Teste do Modelo 2 de F.T - Umidade*

```

data albras;
infile "g:\felix\hidsod\hidsod200.txt" expandtabs;
input t y1 y2 x1 x2 x3 @@ ;
label t          ='tempo em horas'
      y1          ='Umidade'
      y2          ='soda'
      x1          ='condensado'
      x2          ='vapor'
      x3          ='rotacao';
ods html file='albras';
ods graphics on;
proc arima ;
  i var=y1 outcov=y1out ;
  e q=1 method=ml noprint ;
  i var=x1 outcov=x1out ;
  e p=2 method=ml noprint ;
  i var=x2 outcov=x2out ;
  e q=2 method=ml noprint;
  i var=x3 outcov=x3out ;
  e q=1 method=ml noprint;
run;
  i var=y1 crosscor=(x1 x3) nlag=12 ;
run;
  e input=(1$(1) x1 (5) x3) q=1
  method=ml plot;
  forecast lead=0 out=saida;
run;
proc forecast data=saida outest=estim;
run;
proc print data=estim;
  title 'parametros de saida';
run;
data b;
  merge albras saida;
  keep t x1 x2 x3 y1 forecast 195 u95;
run;
proc print data=b;
goptions reset=global gunit=pct border cback=white
  colors=(white yellow) ctext=black
  ftitle=swissb ftext=swiss htitle=4 htext=4;
proc gplot data=b;
plot 195*t      =3
      u95*t     =3
      y1*t      =1
      forecast*t=2/overlay
                                haxis=axis1
                                hminor=4
                                vaxis=axis2
                                vminor=1
                                caxis=black
                                areas=2;

axis1 label=('Data') ;
axis2 label=(angle=90 'Umidade');
symbol1 l=2 w=2.5 color=black i=join ;
symbol2 w=2 i=join c=blue;
symbol3 l=20 i=join c=green;

footnotel c=black f=centx      ' ... Observado'
          c=b f=centx         ' - Modelo'
          c=orange f=centx     '(95% IC)';
run;
ods html close;
quit;

```

Figura A.17 *Teste do Modelo 3 de F.T - Umidade*

```

data albras;
infile "c:\pendrive\felix\hidsod\hidsod200.txt" expandtabs;
input t y1 y2 x1 x2 x3 @@ ;
label t      ='tempo em horas'
      y1      ='Umidade'
      y2      ='soda'
      x1      ='condensado'
      x2      ='vapor'
      x3      ='rotacao';
ods html file='albras';
ods graphics on;
proc arima ;
  i var=y2 outcov=x1out ;
  e q=1 method=ml noprint ;
  i var=x2 outcov=x2out ;
  e q=2 method=ml noprint;
  i var=x3 outcov=x3out ;
  e q=1 method=ml noprint;
run;
  i var=y2 crosscor=(x2 x3) nlag=12 ;
run;
  e input=( 2$(1) x2 1$ x3) q=1
  method=ml plot;
  forecast lead=0 out=saida;
run;
proc forecast data=saida outest=estim;
run;
proc print data=estim;
  title 'parametros de saida';
run;
data b;
  merge albras saida;
  keep t x1 x2 x3 y2 forecast 195 u95;
run;
proc print data=b;
goptions reset=global gunit=pct border cback=white
  colors=(white yellow) ctext=black
  ftitle=swissb ftext=swiss htitle=4 htext=4;
proc gplot data=b;
plot 195*t      =3
      u95*t      =3
      y2*t       =1
      forecast*t=2/overlay
                                haxis=axis1
                                hminor=4
                                vaxis=axis2
                                vminor=1
                                caxis=black
                                areas=2;
axis1 label=('Data') ;
axis2 label=(angle=90 'Soda caustica');
symbol1 l=2 w=2.5 color=black i=join ;
symbol2 w=2 i=join c=blue;
symbol3 l=20 i=join c=green;

footnotel c=black f=centx      ' .... Observado'
          c=b f=centx         ' - Modelo'
          c=orange f=centx     '(95% IC)';
run;
ods html close;
quit;

```

Figura A.18 Gera Modelo de F.T - Soda

```

data albras;
infile "c:\pendrive\felix\hidsod\hidsod200.txt" expandtabs;
input t y1 y2 x1 x2 x3 @@ ;
label t      ='tempo em horas'
      y1      ='Umidade'
      y2      ='soda'
      x1      ='condensado'
      x2      ='vapor'
      x3      ='rotacao';
ods html file='albras';
ods graphics on;
proc arima ;
  i var=y2 outcov=Y2out ;
  e Q=1 method=ml noprint ;
  i var=x2 outcov=x2out ;
  e q=2 method=ml noprint;
  i var=x3 outcov=x3out ;
  e q=1 method=ml noprint;
run;
  i var=y2 crosscor=(x2 x3) nlag=12 ;
run;
  e input=(2$ (2) x2 x3) q=1
  method=ml ;
  forecast lead=0 out=saida;
run;
proc forecast data=saida outest=estim;
run;
proc print data=estim;
  title 'parametros de saida';
run;
data b;
  merge albras saida;
  keep t x1 x2 x3 y2 forecast l95 u95;
run;
proc print data=b;
goptions reset=global gunit=pct border cback=white
  colors=(white yellow) ctext=black
  ftitle=swissb ftext=swiss htitle=4 htext=4;
proc gplot data=b;
plot
  l95*t      =3
  u95*t      =3
  y2*t       =1
  forecast*t =2/overlay
                                haxis=axis1
                                hminor=4
                                vaxis=axis2
                                vminor=1
                                caxis=black
                                areas=2;
axis1 label=('Data') ;
axis2 label=(angle=90 'Soda caustica');
symbol1 l=2 w=2.5 color=black i=join ;
symbol2 w=2 i=join c=blue;
symbol3 l=20 i=join c=green;

footnotel c=black f=centx      ' .... Observado'
          c=b f=centx         ' - Modelo'
          c=orange f=centx     ' (95% IC)';
run;
ods html close;
quit;

```

Figura A.19 *Teste do Modelo 1 de F.T - Soda*

```

data albras;
infile "c:\pendrive\felix\hidsod\hidsod052.txt" expandtabs;
input t y1 y2 x1 x2 x3 @@ ;
label t          ='tempo em horas'
      y1         ='Umidade'
      y2         ='soda'
      x1         ='condensado'
      x2         ='vapor'
      x3         ='rotacao';
ods html file='albras';
ods graphics on;
proc arima ;
  i var=Y2 outcov=Y2out ;
  e Q=1 method=ml noprint ;
  i var=x2 outcov=x2out ;
  e q=2 method=ml noprint;
  i var=x3 outcov=x3out ;
  e q=1 method=ml noprint;
run;
  i var=y2 crosscor=(x2 x3) nlag=12 ;
run;
  e input=(2$ (3) x2 x3) q=1
  method=ml ;
  forecast lead=0 out=saida;
run;
proc forecast data=saida outest=estim;
run;
proc print data=estim;
  title 'parametros de saida';
run;
data b;
  merge albras saida;
  keep t x1 x2 x3 y2 forecast 195 u95;
run;
proc print data=b;
goptions reset=global gunit=pct border cback=white
  colors=(white yellow) ctext=black
  ftitle=swissb ftext=swiss htitle=4 htext=4;
proc gplot data=b;
plot 195*t      =3
      u95*t     =3
      y2*t      =1
      forecast*t=2/overlay
                                haxis=axis1
                                hminor=4
                                vaxis=axis2
                                vminor=1
                                caxis=black
                                areas=2;

axis1 label=('Data') ;
axis2 label=(angle=90 'Soda caustica');
symbol1 l=2 w=2.5 color=black i=join ;
symbol2 w=2 i=join c=blue;
symbol3 l=20 i=join c=green;

footnotel c=black f=centx      ' .... Observado'
          c=b f=centx         ' - Modelo'
          c=orange f=centx'    (95% IC)';
run;
ods html close;
quit;

```

Figura A.20 *Teste do Modelo 2 de F.T - Soda*

A.1.3 Tabela do Banco de Dados

Estes dados foram adquiridos junta a empresa de fabricação de alumínio do Norte do País. Esta empresa destaca-se por ser a maior beneficiadora de minério de Bauxita extraído no município de Paragominas, este também pertencente ao Estado do Pará.

Tabela A.1 *Valores Fornecidos pelo Processo*

Y_1	Y_2	X_1	X_2	X_3
Nível de Umidade (%)	Soda (%)	Vazão de Condensado (m^3/h)	Vazão de Vapor (t/h)	Rotação (%)
6,390	0,010	65,052	1,210	38,282
4,690	0,010	65,169	1,208	38,268
4,630	0,010	52,032	1,195	38,259
4,960	0,010	64,925	1,200	38,232
3,160	0,010	65,697	1,200	38,390
4,920	0,010	64,708	0,256	38,252
5,930	0,010	64,418	0,228	38,218
4,440	0,050	65,277	1,201	38,249
4,510	0,010	68,478	1,204	38,263
3,740	0,010	65,119	1,303	38,190
4,960	0,010	64,730	1,195	38,249
3,700	0,010	65,025	1,204	38,249
3,430	0,010	65,167	1,195	38,115
4,110	0,010	65,758	1,163	38,176
3,870	0,010	65,146	1,195	38,190
4,260	0,010	64,793	1,328	38,241
3,900	0,010	64,921	1,209	38,153
4,460	0,020	65,203	0,260	38,249
4,620	0,010	62,973	1,224	38,282
2,840	0,010	65,094	1,203	38,282
4,490	0,010	65,622	1,192	38,365
5,900	0,010	64,727	1,206	38,301
5,860	0,040	65,052	1,326	38,232
8,180	0,010	69,184	0,593	38,195
4,160	0,010	64,985	1,212	38,263
3,770	0,010	65,215	1,182	38,282
3,430	0,010	57,065	1,187	38,209
5,100	0,010	62,998	1,223	38,249
5,270	0,010	65,122	1,215	38,467
4,500	0,010	64,963	1,197	38,365
6,100	0,020	64,995	1,198	38,398
4,670	0,010	66,153	1,190	38,249
5,430	0,010	64,724	1,195	38,256
4,810	0,010	65,535	1,173	38,240
5,860	0,010	65,340	0,358	38,085
4,200	0,010	72,268	1,197	38,511
5,030	0,010	60,431	1,189	38,282
6,490	0,020	64,995	1,203	38,282
4,970	0,010	64,936	1,191	38,249
3,850	0,010	64,769	1,201	38,136
5,340	0,010	65,296	1,195	38,136
6,160	0,010	64,746	1,195	38,398
5,510	0,030	64,878	1,389	30,438
4,400	0,020	49,254	0,147	38,212
4,720	0,020	65,614	1,195	38,558
4,780	0,010	64,936	1,227	38,224
6,100	0,010	64,756	1,190	38,310
4,940	0,010	64,919	1,206	38,234
3,560	0,010	65,190	1,289	38,136
4,390	0,010	64,927	1,204	38,482
8,620	0,010	64,949	1,203	38,224
6,680	0,010	65,132	1,206	38,249

Tabela A.2 *Continuação da Tabela A.1*

Y_1	Y_2	X_1	X_2	X_3
Nível de Umidade (%)	Soda (%)	Vazão de Condensado (m^3/h)	Vazão de Vapor (t/h)	Rotação (%)
3,900	0,010	65,082	1,194	38,217
4,770	0,020	64,890	1,205	38,518
4,290	0,020	64,622	1,209	38,325
5,290	0,010	65,329	1,200	38,176
5,080	0,020	66,112	1,209	38,365
3,960	0,010	64,646	1,196	38,249
5,420	0,010	64,994	1,223	38,151
5,410	0,010	64,904	1,205	38,282
5,590	0,010	65,281	1,208	38,209
4,540	0,030	66,557	0,988	38,282
6,570	0,010	64,754	1,178	38,241
5,110	0,010	70,094	1,253	38,318
5,520	0,010	64,949	1,260	38,375
4,500	0,020	65,024	1,193	38,209
7,170	0,010	69,465	1,176	38,252
2,810	0,010	69,893	1,187	38,291
6,740	0,010	69,999	1,219	38,086
5,380	0,010	70,039	1,197	38,325
5,380	0,010	64,082	1,327	38,392
5,470	0,010	65,190	1,211	38,282
5,070	0,010	65,445	1,313	38,282
3,510	0,020	65,545	1,192	38,384
6,090	0,010	65,177	1,198	38,325
6,100	0,010	65,207	1,206	38,371
4,840	0,010	65,134	1,202	38,317
4,660	0,010	63,654	1,207	38,282
5,180	0,010	65,119	1,220	38,282
5,980	0,010	64,963	1,209	38,492
7,340	0,120	64,905	1,194	38,465
5,480	0,010	66,259	1,205	38,398
3,950	0,010	64,861	0,297	7,684
3,950	0,010	65,741	1,194	38,421
4,010	0,010	64,815	1,610	38,302
5,090	0,010	64,949	1,206	38,372
4,690	0,010	64,965	1,191	38,489
3,240	0,010	64,923	1,187	38,282
4,850	0,010	61,610	1,198	38,282
3,360	0,010	64,935	1,197	38,276
4,700	0,030	64,825	1,198	38,434
6,260	0,010	65,039	1,183	38,282
5,050	0,010	67,398	0,462	38,342
4,590	0,010	76,500	1,185	38,365
4,580	0,020	65,087	1,208	38,372
7,240	0,010	65,081	1,198	38,232
4,160	0,010	65,265	1,285	38,282
4,860	0,010	7,260	0,068	20,280
5,130	0,010	64,711	1,180	38,392
4,350	0,010	65,039	1,200	38,503
3,620	0,010	65,193	1,203	38,472
5,850	0,010	55,623	1,220	38,365
5,090	0,030	65,007	1,202	38,328
5,070	0,010	66,235	1,205	38,365
3,330	0,010	53,755	0,253	38,249
6,170	0,010	64,936	1,197	38,365
4,990	0,020	64,840	1,202	38,282
5,290	0,010	64,579	1,179	38,279
3,250	0,010	64,963	1,202	38,358
5,300	0,010	58,401	1,100	38,368
3,620	0,010	48,163	1,221	38,372
4,510	0,010	64,637	1,204	38,392
3,520	0,010	64,020	0,099	38,459
5,160	0,010	65,088	1,211	38,398
6,240	0,010	60,015	0,658	38,220
4,150	0,010	64,739	1,205	38,352
7,490	0,010	64,878	1,183	38,472
5,410	0,040	59,608	1,332	38,272
6,510	0,010	32,878	0,175	20,206
5,710	0,010	47,037	0,182	38,209
5,040	0,030	59,875	1,209	38,328
5,770	0,010	60,179	1,317	38,333

Tabela A.3 *Continuação da Tabela A.2*

Y_1	Y_2	X_1	X_2	X_3
Nível de Umidade (%)	Soda (%)	Vazão de Condensado (m^3/h)	Vazão de Vapor (t/h)	Rotação (%)
5,130	0,010	65,370	1,223	38,317
4,240	0,010	64,646	1,211	38,151
5,740	0,010	64,715	1,206	38,282
5,810	0,010	53,393	0,953	38,306
4,990	0,010	64,423	1,186	38,335
5,570	0,010	64,784	1,176	38,282
4,490	0,010	64,971	1,038	38,255
5,350	0,010	64,949	1,173	38,282
6,560	0,010	64,910	1,206	38,143
4,770	0,010	65,161	1,211	38,342
6,290	0,010	64,815	1,205	38,594
4,230	0,040	66,708	1,197	38,328
4,620	0,010	65,043	1,278	38,226
6,280	0,010	65,569	1,204	38,209
4,310	0,010	64,548	1,231	38,231
3,480	0,010	106,356	1,287	38,115
4,780	0,010	64,324	1,193	38,384
8,540	0,010	65,319	1,003	38,348
4,870	0,010	62,625	1,162	38,457
5,090	0,010	66,317	1,198	38,291
4,570	0,010	65,417	1,205	38,176
3,970	0,010	64,920	1,146	38,205
4,690	0,010	65,213	1,177	38,375
4,600	0,010	63,009	1,201	38,209
6,120	0,010	65,326	1,193	38,282
5,810	0,010	57,120	1,330	38,392
5,400	0,020	65,084	1,199	38,282
4,520	0,010	65,192	1,204	38,348
5,510	0,010	65,088	1,194	38,249
4,630	0,010	63,874	1,138	38,121
4,890	0,010	64,742	1,205	38,255
5,070	0,010	64,955	1,204	38,136
5,160	0,010	64,414	1,222	38,325
3,620	0,010	65,342	1,238	38,127
6,320	0,010	64,967	1,209	38,303
5,080	0,010	69,789	0,234	38,217
4,590	0,010	64,800	1,172	38,282
5,980	0,010	65,079	1,194	37,950
5,720	0,010	68,854	1,206	38,274
6,150	0,010	65,060	1,221	38,282
4,810	0,030	65,013	0,106	38,302
7,100	0,010	65,027	1,266	38,176
7,020	0,010	64,969	1,195	38,545
4,400	0,010	65,028	1,235	38,452
6,390	0,050	64,676	1,190	38,457
4,730	0,010	64,578	1,202	38,260
9,330	0,010	64,719	1,204	38,282
6,490	0,010	65,035	1,429	38,282
5,850	0,010	57,403	1,195	38,220
5,270	0,010	65,139	1,200	38,314
5,140	0,010	67,107	1,195	38,273
6,320	0,010	64,566	1,201	38,325
4,900	0,010	65,108	1,199	38,209
6,800	0,010	31,500	0,291	38,063
5,250	0,010	63,790	1,195	38,365
5,030	0,010	58,216	1,341	38,209
5,120	0,010	65,192	1,204	38,472
4,090	0,010	64,441	1,195	38,282
6,300	0,010	65,119	1,195	38,209
4,430	0,010	64,666	1,204	38,282
4,110	0,010	64,285	1,217	38,176
9,680	0,010	65,602	1,194	38,308
4,410	0,010	64,991	1,201	38,330
5,460	0,010	65,465	1,460	38,260
5,270	0,010	64,432	1,179	38,282
4,450	0,010	64,963	1,222	38,282
4,940	0,010	65,719	1,131	38,325

Tabela A.4 *Continuação da Tabela A.3*

Y_1	Y_2	X_1	X_2	X_3
Nível de Umidade (%)	Soda (%)	Vazão de Condensado (m^3/h)	Vazão de Vapor (t/h)	Rotação (%)
6,000	0,010	64,604	1,303	38,393
5,480	0,150	55,811	1,190	38,672
7,910	0,010	59,889	1,169	38,511
5,010	0,010	65,180	1,196	38,472
5,610	0,010	57,350	1,204	38,624
5,450	0,010	65,170	1,320	38,438
5,590	0,010	62,143	1,272	38,295
6,220	0,010	63,662	1,296	38,438
4,280	0,010	68,176	1,276	38,421
6,150	0,010	65,507	1,202	38,472
5,690	0,010	60,966	1,199	40,366
6,650	0,050	51,049	0,790	40,352
5,400	0,010	66,432	1,213	38,399
9,310	0,010	62,475	1,195	38,245
5,030	0,010	54,858	1,257	38,324
5,280	0,010	59,946	1,195	38,344
9,420	0,020	49,604	0,786	40,258
5,710	0,010	64,666	1,204	40,556
7,350	0,010	65,381	1,211	40,309
6,190	0,010	65,038	1,200	40,444
5,810	0,010	65,230	1,193	40,294
5,240	0,020	69,913	1,499	40,405
3,880	0,010	64,671	1,195	40,425
4,720	0,010	68,363	1,207	38,844
6,410	0,010	63,314	1,257	38,528
6,760	0,010	65,560	1,179	38,560
7,050	0,010	64,592	1,199	38,171
6,340	0,010	65,330	1,206	38,230
7,330	0,010	64,967	1,206	38,125
5,210	0,010	65,054	0,664	38,097
6,260	0,020	67,462	1,202	38,317
7,120	0,010	65,598	0,681	38,282
5,420	0,010	64,890	1,202	38,136
6,550	0,010	64,370	1,186	38,282
6,870	0,010	64,266	1,210	38,563
5,330	0,010	65,686	1,250	38,588
6,230	0,010	64,846	1,199	38,249
4,200	0,010	64,212	1,204	38,398
5,690	0,010	62,608	1,186	38,272
6,700	0,010	66,417	0,178	38,238
4,930	0,020	65,477	1,210	38,398
6,480	0,010	61,602	1,198	38,368
7,340	0,010	65,416	1,172	38,249
4,800	0,010	63,768	1,211	38,465
5,470	0,010	64,285	1,497	38,513
7,050	0,010	52,664	1,213	38,282
4,930	0,010	0,133	0,023	37,950
4,760	0,010	0,133	0,023	37,950
5,160	0,010	65,890	1,322	38,282
4,030	0,010	58,402	1,284	38,263
3,800	0,010	62,645	1,294	38,278
7,000	0,010	51,349	1,474	38,231
4,420	0,110	64,274	1,803	38,115
4,640	0,020	67,367	1,502	38,545
9,190	0,010	67,367	1,502	38,545
6,570	0,010	70,406	1,309	38,230
4,760	0,010	69,762	0,886	38,166
4,970	0,020	55,557	0,801	38,009
4,740	0,010	64,273	1,189	38,422
6,520	0,010	65,143	1,176	38,325
5,480	0,010	65,165	1,179	38,219
5,100	0,010	66,271	1,292	38,026
6,200	0,010	66,547	1,362	38,196
5,870	0,090	64,246	1,284	38,332
4,120	0,010	64,991	1,214	38,282
6,390	0,010	63,421	1,023	38,398
7,560	0,140	65,829	1,143	38,369
4,610	0,010	65,212	1,484	38,261

Tabela A.5 *Continuação da Tabela A.4*

Y_1	Y_2	X_1	X_2	X_3
Nível de Umidade (%)	Soda (%)	Vazão de Condensado (m^3/h)	Vazão de Vapor (t/h)	Rotação (%)
6,380	0,010	63,514	1,287	38,209
6,630	0,010	60,801	1,320	38,381
4,860	0,010	62,568	1,323	38,188
5,150	0,010	65,089	1,175	38,252
5,260	0,010	72,932	1,280	38,249
7,010	0,010	57,791	1,175	38,329
6,370	0,010	65,011	1,226	38,372
4,600	0,010	67,268	1,616	38,398
6,200	0,010	51,745	1,300	38,391
6,210	0,010	60,960	1,194	38,260
5,480	0,010	64,730	1,402	38,208
7,410	0,010	65,665	1,189	38,184
4,160	0,040	65,309	1,202	38,198
4,880	0,010	65,235	1,296	38,498
5,900	0,010	67,225	1,195	38,249
4,640	0,010	68,969	1,204	38,187
3,090	0,010	64,840	1,187	38,325
3,080	0,080	66,742	1,502	38,200
4,430	0,010	65,378	0,626	38,397
8,630	0,010	68,291	1,344	37,983
4,410	0,010	65,141	1,206	38,249
5,990	0,010	65,671	1,195	37,997
5,200	0,010	64,350	1,196	38,257
3,180	0,010	50,947	1,199	38,249
6,140	0,010	66,251	1,201	38,209
5,940	0,010	66,021	1,296	38,236
3,650	0,010	64,771	1,291	38,613
4,460	0,010	66,025	1,503	38,236
7,200	0,010	64,347	1,195	38,282
4,360	0,010	66,531	1,193	38,398
5,180	0,020	67,157	1,190	38,198
5,900	0,010	64,975	1,200	38,249
7,570	0,010	66,340	1,199	30,502
4,420	0,010	62,569	0,502	30,584
5,930	0,010	65,263	1,226	30,256
5,700	0,010	59,164	1,199	38,207
5,420	0,010	60,412	1,218	38,282
5,760	0,010	64,744	1,176	38,282
5,240	0,010	65,205	1,179	38,164
6,630	0,010	66,752	1,198	38,221
6,460	0,010	65,169	1,209	38,167
5,000	0,010	65,064	1,215	38,230
4,850	0,010	64,753	1,235	38,249
6,750	0,010	64,493	1,261	38,282
6,040	0,010	65,208	1,194	38,176
4,560	0,010	64,332	1,195	38,220
6,480	0,010	60,176	1,161	38,249
6,160	0,010	64,588	1,200	38,200
5,880	0,010	64,660	1,292	38,282
5,150	0,010	65,392	1,289	38,335
5,740	0,010	64,873	1,204	38,282
8,060	0,010	65,165	1,190	38,365
7,830	0,010	65,137	1,282	38,408
3,940	0,010	65,466	1,183	38,209
4,210	0,010	65,135	1,156	38,282
3,970	0,010	66,132	1,315	38,176
7,690	0,010	66,002	1,015	38,327
4,290	0,010	64,183	1,234	38,063
5,880	0,010	65,192	0,970	38,282
5,600	0,010	64,441	0,986	38,282
6,200	0,020	64,410	1,305	38,080
4,980	0,010	64,326	1,332	38,249
8,340	0,010	64,627	1,502	38,365
2,490	0,010	54,180	1,341	38,367
4,150	0,010	61,016	0,212	37,420
4,270	0,010	64,986	1,217	38,289
5,020	0,010	61,695	1,213	38,282
4,850	0,260	65,099	1,212	38,209
5,100	0,010	64,400	1,196	38,014
7,360	0,020	65,807	1,061	38,063
4,720	0,010	64,551	1,314	38,511
5,890	0,010	64,930	1,303	38,176
3,930	0,010	64,192	1,190	38,141
3,900	0,010	65,802	1,199	38,281
3,900	0,010	64,675	1,195	38,345

Tabela A.6 *Continuação da Tabela A.5*

Y_1	Y_2	X_1	X_2	X_3
Nível de Umidade (%)	Soda (%)	Vazão de Condensado (m^3/h)	Vazão de Vapor (t/h)	Rotação (%)
4,980	0,010	64,840	1,204	38,233
3,810	0,010	65,101	1,204	38,022
5,170	0,010	64,903	1,026	38,099
6,240	0,010	65,949	1,208	38,203
7,180	0,010	64,730	1,205	38,116
5,700	0,010	64,298	1,219	38,136
4,070	0,010	65,453	1,204	38,136
7,040	0,010	60,978	1,201	38,017
5,950	0,020	60,896	0,312	38,218
5,530	0,010	58,298	1,320	38,379
5,430	0,010	65,497	1,196	38,201
3,500	0,010	65,222	1,184	38,223
6,450	0,010	60,614	1,300	38,277
5,410	0,010	67,048	1,197	38,514
5,640	0,010	60,628	1,318	38,176
2,570	0,010	60,032	1,225	38,120
6,450	0,010	57,116	1,234	38,282
4,980	0,010	59,561	0,339	37,983
3,400	0,010	59,777	1,225	38,282
6,430	0,010	60,476	1,152	37,950
7,040	0,010	59,949	1,238	38,168
5,870	0,010	59,836	1,196	38,148
3,340	0,010	54,300	1,204	38,063
5,060	0,010	49,408	1,210	38,043
6,410	0,010	56,111	1,204	37,983
7,400	0,010	55,027	1,203	38,176
4,370	0,010	54,540	1,396	37,988
7,390	0,010	54,672	1,296	38,026
3,420	0,010	65,675	1,177	38,136
4,500	0,010	64,726	1,378	37,993
3,360	0,010	58,847	1,369	38,142
3,450	0,010	59,861	1,195	38,136
4,530	0,010	66,604	0,390	38,064
3,630	0,010	59,229	1,168	38,356
3,780	0,030	64,963	1,195	38,026
6,150	0,010	59,504	1,497	37,983
4,070	0,010	58,664	1,355	38,063
7,150	0,010	64,666	1,209	37,950
3,910	0,010	47,558	1,413	37,875
7,500	0,010	64,832	1,208	38,176
5,900	0,010	64,955	1,204	38,175
3,400	0,010	65,406	1,191	38,180
6,330	0,010	63,444	1,199	37,983
4,300	0,010	66,684	1,213	38,209
6,500	0,010	58,684	1,190	38,137
3,290	0,010	64,520	1,236	38,249
3,390	0,010	64,960	1,204	38,282
6,030	0,010	55,952	0,335	38,136
4,500	0,010	61,659	1,360	38,249
5,250	0,010	64,931	1,209	37,983
5,500	0,010	64,964	1,177	38,282
3,140	0,010	60,019	1,226	38,209
5,200	0,010	60,086	1,314	38,282
3,530	0,010	60,387	1,208	38,142
6,250	0,010	60,683	1,492	38,176
4,770	0,010	60,320	1,195	37,983
2,810	0,010	59,869	0,655	38,802
4,300	0,010	55,385	1,195	38,282
5,190	0,010	60,301	1,190	38,249
3,860	0,010	59,571	1,204	38,176
4,270	0,010	64,343	1,204	38,279
4,560	0,010	59,149	1,300	38,249
4,980	0,010	60,000	0,677	38,282
2,700	0,010	60,088	1,213	38,209
4,200	0,010	60,682	1,208	38,141
4,430	0,010	64,910	1,309	38,249

Tabela A.7 *Continuação da Tabela A.6*

Y_1	Y_2	X_1	X_2	X_3
Nível de Umidade (%)	Soda (%)	Vazão de Condensado (m^3/h)	Vazão de Vapor (t/h)	Rotação (%)
2,760	0,010	65,119	1,190	38,249
4,530	0,010	60,395	0,275	38,026
4,880	0,010	59,017	0,333	38,065
4,870	0,010	60,383	1,172	38,173
4,970	0,010	54,759	1,254	38,136
4,140	0,010	60,239	1,254	38,209
4,860	0,010	59,941	1,172	38,063
4,960	0,030	64,739	1,186	38,209
7,590	0,010	64,472	1,182	38,267
4,080	0,010	61,365	1,176	38,365
4,080	0,010	58,325	1,218	38,282
5,770	0,050	59,415	1,195	38,365
6,120	0,010	60,239	1,176	38,282
5,190	0,010	59,941	1,195	38,176
4,670	0,010	60,353	1,296	38,282
5,860	0,010	69,019	1,227	38,209
4,740	0,010	59,746	1,221	38,206
4,000	0,010	59,264	1,176	38,249
8,110	0,010	60,541	1,479	38,249
6,670	0,190	14,532	0,218	38,046
6,040	0,010	59,360	1,276	38,249
8,600	0,010	65,053	1,227	38,171
5,840	0,010	64,324	1,182	38,218
5,580	0,010	64,817	1,350	38,176
3,940	0,010	55,364	1,190	38,209
6,050	0,010	65,343	1,341	38,472
5,810	0,010	64,792	1,172	38,249
4,090	0,010	65,389	1,204	38,295
3,880	0,010	65,228	1,189	39,448
4,610	0,010	65,490	1,209	39,335
5,850	0,010	65,325	1,190	39,189
5,670	0,010	64,922	1,188	39,290
5,730	0,010	65,157	1,223	38,219
6,350	0,300	65,119	1,373	38,176
5,760	0,010	64,592	1,218	38,136
5,120	0,010	64,963	1,227	38,209
6,560	0,030	61,814	1,332	37,983
4,630	0,010	64,925	1,178	38,141
6,690	0,010	64,739	1,204	38,624
4,380	0,160	66,709	1,006	38,795
4,340	0,010	64,817	1,218	38,209
7,180	0,030	64,331	1,474	38,398
6,910	0,080	64,212	1,314	38,282
5,600	0,010	65,202	1,176	39,262
5,910	0,010	65,192	1,222	39,750
4,630	0,010	63,065	1,211	38,365
6,470	0,010	29,435	1,373	38,136
4,540	0,010	29,435	1,373	38,136
5,410	0,010	29,435	1,373	38,136
4,070	0,010	65,179	1,222	38,484
3,130	0,010	64,890	1,190	39,372
5,850	0,010	65,325	1,190	39,189
5,670	0,010	64,922	1,188	39,290
5,730	0,010	65,157	1,223	38,219
6,350	0,300	65,119	1,373	38,176
5,760	0,010	64,592	1,218	38,136
5,120	0,010	64,963	1,227	38,209
6,560	0,030	61,814	1,332	37,983
4,630	0,010	64,925	1,178	38,141
6,690	0,010	64,739	1,204	38,624
4,380	0,160	66,709	1,006	38,795
4,340	0,010	64,817	1,218	38,209
7,180	0,030	64,331	1,474	38,398
6,910	0,080	64,212	1,314	38,282
5,600	0,010	65,202	1,176	39,262
5,910	0,010	65,192	1,222	39,750
4,630	0,010	63,065	1,211	38,365
6,470	0,010	29,435	1,373	38,136
4,540	0,010	29,435	1,373	38,136
5,410	0,010	29,435	1,373	38,136
4,070	0,010	65,179	1,222	38,484
3,130	0,010	64,890	1,190	39,372

Bibliografia

- ABRALATAS. *Fluxo de Produção de Alunínio*. Disponível em [http :
//www.abralatas.org.br/flash/Fluxo_prod_aluminio.swf](http://www.abralatas.org.br/flash/Fluxo_prod_aluminio.swf). Acesso em: 15 de Março de 2007.
- AKAIKE, H. *Maximum Likelihood Identification of Gaussian Autoregressive Moving Average Models*. Biometrika, 1973, 60, 255-265
- AKAIKE, H. *A New Look at The Statistical Model Identification*. IEEE, Transactions an Automatic Control, 1974, AC-19, 716-723
- AKAIKE, H. *On Entropy Maximization Principle. In Applications of Statistics*.(P.R. Krishnariah, Ed.), 27-41. Amsterdam: North-Holland, 1977.
- BERMAM, C. *Consumo de Energia Elétrica*. Disponível em [http :
//www.socioambiental.org/manchetes/index_h.html](http://www.socioambiental.org/manchetes/index_h.html). Acesso em: 7 de Agosto de 2007.
- BERTASI, D. e SIQUEIRA, M. L. *Reconhecimento de Caracteres Utilizando Redes Neurais Artificiais*, Porto Algre, Dezembro 1999
- BOX, G. E. P. and JENKINS, G.M. *Time Series Analysis: Forecasting and Control Third Edition*. Englewood Cliffs: Prentice Hall, 1994.
- BROCKLEBANK, J. C. and DICKEY, D. A. *SAS System for Forecasting Time Series*. SAS Institute Inc, Cary, NC, USA, 1986.
- CERQUEIRA, E. O. ANDRADE, J. C. e POPPI, R, J *Redes Neurais e suas Aplicações em Calibração Multivariada*, São Paulo, 2001.
- DAHMER, A. *Segmentação de Imagens Ecocardiográficas Utilizando Redes Neurais e Medidas de Textura*, Porto Alegre: CPGCC, 1998. Dissertação de Mestrado.
- Estudo Sobre Inteligência Artificial*, 2006.Disponivel em: [www.citi.pt/educacao_final/trab_final
_inteligencia_artificial/ia.html](http://www.citi.pt/educacao_final/trab_final_inteligencia_artificial/ia.html). Acesso em: 17 de Outubro de 2006
- Efeito Estufa e Aquecimento Global*. Disponivel em: [http :
//paginas.terra.com.br/lazer/staruck/efeitoestufa.htm](http://paginas.terra.com.br/lazer/staruck/efeitoestufa.htm). Acesso em: 22 de Novembro de 2007.
- FERNANDES, L. G. L. PORTUGAL, M. S. e NAVAUX, P. O. *A Previsão de Séries de Tempo: Redes Neurais Artificiais e Modelos Estruturais*, Rio Grande do Sul, 1995

- FERNANDES, L. G. L. *Utilização de Redes Neurais na Análise de Séries Temporais*, Porto Alegre: CPGCC, 1995. Dissertação de Mestrado.
- GUIMARÃES, B. B. e NETO, F. R. *Redes Neurais para a Previsão de Índices do Mercado de Ações: uma análise do trabalho de Amol S. Kulkarni*, Artigo, UFBA - Salvador - Bahia.
- HECHT - NIELSEN, R. *Neurocomputing: Addison Wesley Publ. Co.*, New York, 1990.
- HAGAN, Martin T. e MENHAJ, Mohammed B. (1994). *Training Feedforward Networks with the Marquardt Algorithm*. IEEE Transactions on Neural Networks, v.5, n.6.
- HAIKIN, S. *Redes Neurais - Princípios e Prática*, Porto Alegre: Editora Bookman, 2001.
- HAMILTON, J. D. *Time Series Analysis* Princeton University Press, Princeton New Jersey, 1994.
- HERNANDEZ, E. D. L. *Inteligência Computacional e Redes Neurais em Engenharia Elétrica*, São Paulo: PSI-EPUSP, 2005.
- HINES, J. W. *MATLAB Supplement to Fuzzy and Neural Approaches in Engineering*
- LIMÃO, R. C. *Rede Neural com Dinâmica Interna Aplicada a Problemas de Identificação e Controle Não-Linear*, Santa Catarina: UFSC, 1999.
- MORETTIN, P. A. e TOLOI, C. M. C. *Análise de Séries Temporais*, Edgard Blücher, São Paulo, 2004.
- MORETTIN, P. A. e TOLOI, C. M. C. *Previsão de Séries Temporais*. São Paulo: Editora Atual. 1987.
- PALIT, A. K. and POPOVIC, D. *Computational Intelligence in Time Series Forecasting Theory and Engineering Applications - Advances in Industrial Control*. Springer-Verlag, 20053, p.1-142
- PINHEIRO, J. R. *Funções de Transferências* - Notas - Disponível em <http://www.ufsm.br/gepoc/renes/Templates/arquivos/elc1031/ELC1031.L2.4.pdf> - acesso em 27 de Setembro de 2007.
- PORTUGAL, M. S. e FERNANDES, L. G. L. *Redes Neurais Artificiais e Previsão de Séries Econômicas: Uma Introdução*, p. 22, 1994.
- PLANTULLO, V. L. *Teoria Geral da Administração: De Taylor às Redes Neurais*, p.172, Rio de Janeiro, 2001.
- RISSANEM, J. *Modelling by Shortest data Description. Automatic*, 14, 465-471, 1978.
- RICK, E. e KNIGHT, K. *Inteligência Artificial*, São Paulo - 2ª edição: Makron Books do Brasil Editora Ltda, 2003.
- SANTOS, J. *O Alumínio*. 21-23 de maio de 2007. Disponível em <http://w3.ualg.pt/~ealmeida/Semin%C3%A1rios/U.algarve-%20EST-EXTRUSAL.pdf>. Acesso em: 22 de

Fevereiro de 2007.

SERTEC. *Alumínio - o que é?*. Disponível em http://www.serteq.com.br/noticia.asp?codigo=146&COD_MENU=50. Acesso em: 02 de Março de 2007.

SCHWARTZ, G. *Estimating the Dimension of a Model*, Annals of Statistics, 6, 461-464, 1978.

SIMÕES, M. G. e SHAW, I. S. *Controle e Modelagem Fuzzy*, p. 165, São Paulo, 1999.

SOUZA R. C. *Modelos Estruturais para Previsão de Séries Temporais*, IMPA-CNPQ - Rio de Janeiro, 1989, 17º Colóquio Brasileiro de Matemática.

SOUZA, A. N. SILVA, I. N e BORDON, M. E. *Identificação de Processos de Ensaio de Alta Tensão Através de Redes Neurais Artificiais*, São Paulo, 1999, IV Congresso Brasileiro de Redes Neurais, pag. 072-077.

WEI, W. W. S *Time Series Analysis - Univariate and Multivariate Methods*, Department of Statistics Temple University, 1989

WILOMAŃSKA, A. *Multivariate Time Series*, Hugo Steinhaus Center Wroclaw University of Technology. Berlin, 2005.

WYLOMANSKA, A. *Multivariate Time Series*, Hugo Steinhaus Center, Wroclaw University of Technology, Berlin, July 2005

KOIVO, H. N. *Neural Networks: Basics Using MATLAB Neural Network Toolbox*, February 2000

KOVÁCS, K. L. *Redes Neurais Artificiais - Fundamentos e aplicações*, Editora Acadêmica. São Paulo, 1996.

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](#)