

Universidade Estadual de Campinas
Faculdade de Engenharia Química
Área de Concentração
Desenvolvimento de Processos Químicos

AVALIAÇÃO DA UTILIZAÇÃO DE REDES NEURONAIS APLICADAS A PROCESSOS QUÍMICOS

Autor: Carlos Alberto Ubirajara Gontarski

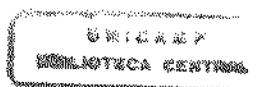
Orientador: Prof. Dr. Milton Mori

Tese de Doutorado apresentada à Faculdade de Engenharia Química como parte dos requisitos exigidos para a obtenção do título de Doutor em Engenharia Química.

Campinas – São Paulo

Dezembro de 2000

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE



x 30
 MARCA:
 UNICAMP
 5588a
 Ex.
 60/43837
 16-392101
 D
 1100
 2010-2101

-00153319-1

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

G588a Gontarski, Carlos Alberto Ubirajara
 Avaliação da utilização de redes neuronais aplicadas a
 processos químicos / Carlos Alberto Ubirajara
 Gontarski.--Campinas, SP: [s.n.], 2000.

Orientador: Milton Mori.

Tese (doutorado) - Universidade Estadual de
Campinas, Faculdade de Engenharia Química

1. Polpação alcalina por sulfito. 2. Madeira -
 Classificação. 3. Águas residuais - Purificação -
 Tratamento biológico. 4. Papel - Indústria. 5. Celulose.
 6. Redes neurais (Computação). I. Mori, Milton. II.
 Universidade Estadual de Campinas. Faculdade de
 Engenharia Química. III. Título.

QUÍMICA

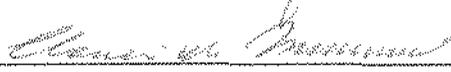
UNIVERSIDADE ESTADUAL DE CAMPINAS

2010

Tese de Doutorado defendida por Carlos Alberto Ubirajara Gontarski e aprovada em 06 de dezembro de 2000 pela banca examinadora constituída pelos doutores:



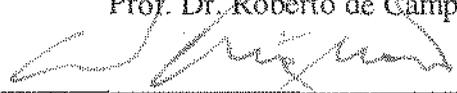
Prof. Dr. - Orientador Milton Mori



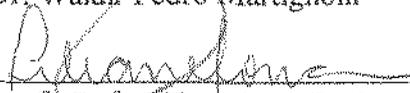
Prof. Dr. Cláudio Augusto Oller do Nascimento



Prof. Dr. Roberto de Campos Giordano



Dr. Waldir Pedro Martignoni



Prof.^a. Dr.^a. Liliane Lona Batista

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

Folha de versão final

Este exemplar corresponde à versão final da Tese de Doutorado em
Engenharia Química.



Prof. Dr. Milton Mori

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

Dedico este trabalho à Suzana, Amanda, Nicole e Rosi, minha família que sempre me dá o apoio que preciso, e em especial a meu querido e falecido pai Alberto Michal Gontarski, engenheiro químico, grande homem ao qual procuro ter como exemplo.

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

Agradecimentos

Ao grande amigo, Prof. Dr. Milton Mori, pela orientação deste trabalho e pela orientação na minha vida profissional.

Ao programa PICDT da CAPES pelo apoio financeiro no desenvolvimento deste trabalho.

À Ripasa Celulose e Papel SA pela disposição em realizar um trabalho em conjunto, fornecendo dados e recursos para que o mesmo tivesse sucesso. Em especial aos engenheiros Manoel Telhada e Fabiano Periotto pelas proveitosas discussões a respeito do trabalho.

À Rhodiaco Ltda. pelo trabalho realizado em conjunto, fornecendo dados e recursos para que o mesmo tivesse sucesso. Em especial aos engenheiros Leo Prenen, Wellington Bonifácio e Paulo Rodrigues pelas diversas reuniões, visitas e discussões a respeito do trabalho.

Aos professores do Departamento de Processos Químicos da Unicamp, Prof^o. Dr^a. Liliane Lona, pelos valiosos comentários nas apresentações sobre o trabalho, ao Prof. Dr. Martin Aznar e a Prof^a Dr^a Aline Costa pelas discussões, apoio nesta fase final da tese e principalmente pela grande amizade.

Aos colegas do Departamento de Engenharia Química da UFPR, em especial aos amigos Luiz Fernando Luz Jr, Georges Kaskantzis e Carlos Yamamoto, com quem sempre pude discutir assuntos afetos a este trabalho entre outros.

Aos colegas do Laboratório de Modelagem e Simulação de Processos Químicos da UNICAMP, em especial aos amigos Henry Meier, Jailson e Alexandre Peres pelos auxílios nos momentos difíceis e debates sobre o trabalho.

A todos que não mencionei aqui e que contribuíram direta ou indiretamente para que o trabalho fosse realizado.

Resumo

Este trabalho apresenta alguns aspectos importantes na aplicação de redes neurais a simulação de processos químicos. As redes são utilizadas em dois processos industriais distintos, fornecendo uma forma de comparar as dificuldades e potencialidades da aplicação desta técnica.

Na Ripasa S/A, uma das maiores indústrias de papel e celulose do Brasil, a classificação da matéria-prima para o processo de polpação que tem sido usada é baseada em uma das propriedades da madeira, a sua densidade básica. Na Ripasa se utiliza a madeira de árvores do gênero *Eucalyptus*. O capítulo 3 apresenta uma alternativa mais rigorosa para o critério de classificação da matéria-prima do processo, levando em consideração outros dados tais como propriedades físicas, composição química da madeira e informações relativas à origem das árvores. Acredita-se que uma boa classificação deve se basear nas propriedades finais da polpa obtida, que são o rendimento depurado, o número Kappa, a viscosidade e a alcalinidade residual. A predição destas propriedades é um dos objetivos deste trabalho, desde que as suas determinações experimentais demandam custo e tempo de obtenção consideráveis. Utiliza-se uma técnica estatística para estimar os erros de amostragem e medida, para serem comparados com os erros gerados pelas redes neurais. Redes neurais foram treinadas a partir dos dados analisados da coleta de 165 amostras de madeira oriundas de oito regiões diferentes.

No segundo estudo (capítulo 4) se apresenta uma forma de prever as propriedades ambientais da corrente de saída da unidade de tratamento de efluentes líquidos na Rhodiaco Ltda, uma das plantas químicas mais importantes no Brasil. A indústria produz o ácido tereftálico e gera águas residuárias que necessitam ser tratadas em um sistema de lodo ativado. Redes neurais do tipo "backpropagation" são utilizadas para prever a eliminação do carbono orgânico total (TOC) pela unidade de tratamento, usando o algoritmo delta-bar-delta para ajustar os pesos e a função sigmoideal como função de transferência nos neurônios. A influência das variáveis de entrada é analisada e resultados satisfatórios são obtidos na predição de algumas situações analisadas.

A principal conclusão destes trabalhos é que o uso de redes neurais pode auxiliar a estabelecer melhores condições operacionais para processos químicos industriais.

As redes neuronais se apresentam como uma possível ferramenta de auxílio a operação, de modo a prever situações de pico, e propiciando ações preventivas que minimizam as flutuações na produção. No futuro, vários trabalhos deverão surgir para prever as condições do efluente industrial baseados em dados de operação reais.

Palavras-chave : Polpação, Digestores de polpa, classificação de madeira, número Kappa, predição de propriedades ambientais, tratamento de águas residuárias, aplicações de redes neuronais.

Abstract

This work presents some important aspects when applying neural networks for chemical process simulations. Neural networks are used in two different industrial processes providing a way of comparison of difficulties and potentialities for the application of this technique.

Currently at Ripasa S/A, one of the largest paper industries in Brazil, the raw material classification for the pulp digester has been done using density as the property of the wood. At Ripasa S/A, the wood comes from trees of the genus *Eucalyptus*. The chapter 3 presents a more rigorous alternative for the criteria of raw material classification by taking into account other data such as physical properties, chemical composition of the wood and origin of the trees. A good classification should be based on the final properties of the process, which are yield, Kappa number, viscosity and residual alkalinity. Prediction of these properties by neural networks was one of the objectives of this work since experimental measurement is time-consuming and expensive.

A statistical technique was used to estimate sampling and measurement errors. These values were compared with the errors generated by the neural network. Different back-propagation networks were created to evaluate a suitable set of network parameters, such as number of nodes on the hidden layer, learning parameters and input range, among others. The networks were trained by data recorded from the analysis of 165 wood samples from eight different regions.

The second work (chapter 4) presents a way to predict the environmental properties of the output stream from the wastewater treatment plant at Rhodiaco Ltda, one of the major chemical plants in Brazil. The industrial plant produces purified terephthalic acid and generates wastewater that should be treated in an activated sludge system. Back-propagation neural networks are used to predict the elimination of total organic carbon (TOC) in the treatment plant, using the delta-bar-delta algorithm for estimation of weights and the sigmoid function as the neuron transfer function. The influence of input variables is analyzed, and satisfactory predicted results are obtained for some situations analyzed.

The main conclusion of this work is that the neural networks can be used to establish a better operating condition for industrial chemical processes. Neural networks represent a possible aid to operations in order to predict upsets and proactively act to minimize output fluctuations. In the future, some work will be done to predict effluent conditions based on the actual operation data set.

Keywords: Pulp digesters, pulp raw material classification, Kappa number, Prediction of environmental properties, Wastewater treatment, Neural network applications.

Sumário

RESUMO	vii
ABSTRACT	ix
SUMÁRIO	xi
LISTA DE FIGURAS	xvi
LISTA DE TABELAS	xxi
NOMENCLATURA	xxiv
CAPÍTULO 1. - INTRODUÇÃO	1
1.1. JUSTIFICATIVA DA ESCOLHA DO TEMA	1
1.2. PROPOSTA E OBJETIVOS DO TRABALHO:	2
CAPÍTULO 2. REVISÃO BIBLIOGRÁFICA	5
2.1. REDES NEURONAIS:	5
2.1.1. FILOSOFIA E HISTÓRICO	5
2.1.2. CLASSIFICAÇÃO DE TIPOS	7
2.1.3. ESTRUTURA BÁSICA	8
2.1.4. REDES NEURONAIS "LVQ – LEARNING VECTOR QUANTIZATION"	13
2.1.5. TREINAMENTO E PROGRAMAÇÃO	15
2.1.6. REDES NEURONAIS E OUTRAS TECNOLOGIAS	17
2.1.7. APLICAÇÕES PARA REDES NEURONAIS.	18
2.1.8. APLICAÇÕES DE REDES NEURONAIS PARA A ÁREA DE ENGENHARIA QUÍMICA	18
2.1.9. PUBLICAÇÕES SOBRE A APLICAÇÃO DE REDES NEURONAIS A PROCESSOS DE POLPAÇÃO E TRATAMENTO DE EFLUENTES.	21
2.2. O PROCESSO DE FABRICAÇÃO DA CELULOSE E DO PAPEL.	21
2.3. TRATAMENTO DE EFLUENTES LÍQUIDOS	26
2.3.1. PRINCIPAIS VARIÁVEIS MONITORADAS	36
2.3.1.1. Demanda bioquímica de oxigênio	36

2.3.1.2.	Demanda química de oxigênio – (DQO)	37
2.3.1.3.	Carbono Orgânico Total (TOC)	37
2.3.1.4.	pH	37
2.3.1.5.	Sólidos suspensos	37
2.4.	USO DA ESTATÍSTICA PARA ANÁLISE DOS DADOS.	38
2.4.1.	CONCEITOS BÁSICOS	38
2.4.2.	TESTES DE HIPÓTESES:	38
2.4.3.	ANÁLISE DA VARIÂNCIA	39
2.4.4.	DESENHOS HIERÁRQUICOS	39
2.4.5.	PLANEJAMENTO DE EXPERIMENTOS	40

CAPÍTULO 3. ESTUDO DE CLASSIFICAÇÃO DA MATÉRIA-PRIMA PARA A POLPAÇÃO

3.1.	CARACTERIZAÇÃO DO PROBLEMA A SER ESTUDADO	41
3.2.	OBJETIVOS PROPOSTOS DESTE ESTUDO	42
3.3.	VARIÁVEIS E PROPRIEDADES ENVOLVIDAS	43
3.4.	DADOS FORNECIDOS PARA REALIZAR O ESTUDO	44
3.4.1.	BANCO DE DADOS COM AS PROPRIEDADES PRINCIPAIS DO LOTE	44
3.4.2.	CLASSIFICAÇÃO OBTIDA COM OS ESPECIALISTAS EM MATÉRIA-PRIMA	44
3.5.	METODOLOGIA ADOTADA NA REALIZAÇÃO DO ESTUDO	44
3.5.1.	ALGORITMOS DE TREINAMENTO UTILIZADOS NESTE ESTUDO	45
3.5.2.	CÁLCULO DAS MÉDIAS	45
3.5.3.	PADRONIZAÇÃO DE TALHÃO, FAZENDA (PARQUE FLORESTAL) E ESPÉCIE	45
3.5.4.	MANIPULAÇÃO DOS DADOS	45
3.5.5.	ALEATORIZAÇÃO DOS CONJUNTOS DE TREINO E DE TESTE	46
3.5.6.	ALEATORIZAÇÃO NO TREINAMENTO DA REDE	46
3.5.7.	TREINAMENTO E TESTE . PROCEDIMENTO “SAVE BEST”	46
3.5.8.	INSTRUMENTOS UTILIZADOS DO SOFTWARE NEURALWORKS PRO II	47
3.5.8.1.	Erro RMS (“RMS error”)	47
3.5.8.2.	Matriz de Resultados (“Confusion matrix”)	47

3.5.8.3.	Índice de correlação (“Correlation”)	48
3.5.8.4.	Somatório das entradas (“Summation”)	48
3.5.9.	INSTRUMENTOS PARA ANÁLISE EXTERNOS AO SOFTWARE NEURALWORKS	48
3.5.9.1.	Estatísticas	48
3.5.9.2.	Gráficos	48
3.5.9.3.	“Flash code”	49
3.6.	RESULTADOS OBTIDOS	49
3.6.1.	ANÁLISE INICIAL DOS DADOS	49
3.6.2.	ESTIMATIVA DO ERRO DE AMOSTRAGEM E MEDIDA DAS VARIÁVEIS	53
3.6.3.	AVALIAÇÃO INICIAL DO NÚMERO DE CAMADAS, NÚMERO DE NEURÔNIOS, FUNÇÃO DE ATIVAÇÃO E ALGORITMO DE TREINAMENTO	57
3.6.4.	DEPURAÇÃO DO RESULTADO	59
3.6.5.	CLASSIFICAÇÃO	64
3.7.	PROGRAMAS APLICATIVOS GERADOS	68
3.8.	ANÁLISE DE SENSIBILIDADE DAS VARIÁVEIS DE ENTRADA	72
3.8.1.	ANÁLISE PARAMÉTRICA - NÍVEIS ESTIPULADOS EM 10% ABAIXO E ACIMA DO VALOR MÉDIO	74
3.8.1.1.	Planejamento dos experimentos	76
3.8.1.2.	Resultados e análises	77
3.8.1.3.	Análise dos fatores principais somente:	78
3.8.1.4.	Análise considerando as interações até de terceiro grau	78
3.8.2.	ANÁLISE PARAMÉTRICA - NÍVEIS ESTIPULADOS POR UM DESVIO PADRÃO ABAIXO E ACIMA DA MÉDIA	78
3.8.2.1.	Planejamento dos experimentos	80
3.8.2.2.	Resultados e análises (interações até terceira ordem)	80
3.8.2.3.	Outras análises	81
3.8.3.	CONCLUSÕES ESPECÍFICAS DESTE ESTUDO	81
3.9.	CONCLUSÃO SOBRE O ESTUDO DA CLASSIFICAÇÃO DA MATÉRIA-PRIMA PARA A POLPAÇÃO DA MADEIRA	83

<u>CAPÍTULO 4. PREDIÇÃO DE COMPORTAMENTO DO SISTEMA DE</u>	
<u>TRATAMENTO DE EFLUENTES USANDO A TECNOLOGIA</u>	
<u>DE REDES NEURONAIS</u> _____ 85	
4.1. CARACTERIZAÇÃO DO PROBLEMA A SER ESTUDADO _____	85
4.2. OBJETIVOS DO ESTUDO REALIZADO _____	87
4.3. TRATAMENTO INICIAL DOS DADOS _____	87
4.4. TREINAMENTO DAS ESTRUTURAS DE REDES NEURONAIS PARA PREDIÇÃO DO TOC	
FINAL DA UNIDADE _____	91
4.4.1. PRIMEIRO CONJUNTO DE DADOS - JUN/98 A OUT/98: _____	91
4.4.1.1. Treinando cada reator com dados de uma série e a outra como teste: _____	92
4.4.1.2. Treinando para cada reator individualmente: _____	98
4.4.2. SEGUNDO CONJUNTO DE DADOS - NOV/98 + MAR/99 A JUN/99: _____	104
4.4.3. CONJUNTO DIFERENCIADO DE DADOS _____	108
4.5. ESTUDO DE CASO COM AS REDES TREINADAS _____	115
4.6. ELABORAÇÃO DO PROGRAMA PARA PREDIÇÃO DO TOC _____	115
4.6.1. CONCLUSÕES FINAIS SOBRE O ESTUDO NA UNIDADE DE TRATAMENTO DE	
EFLUENTES _____	126
 <u>CAPÍTULO 5. DISCUSSÃO GERAL DOS RESULTADOS, CONCLUSÕES E</u>	
<u>SUGESTÕES</u> _____ 129	
5.1. DISCUSSÃO GERAL DOS RESULTADOS _____	129
5.1.1. SIMULADOR PARA UNIDADES DE TRATAMENTO DE EFLUENTES _____	131
5.1.2. PROCESSO DE TRATAMENTO DE EFLUENTES DA RIPASA SA _____	131
5.1.3. PROCESSO DE POLPAÇÃO EM UNIDADE PILOTO DA MADEIRA (RIPASA SA) _____	132
5.1.4. PROCESSO DE TRATAMENTO DE EFLUENTES DA RHODIACO LTDA. _____	132
5.2. CONCLUSÕES _____	132
5.3. SUGESTÕES PARA PROSSEGUIMENTO DO TRABALHO: _____	134
REFERÊNCIAS BIBLIOGRÁFICAS _____	137
ANEXOS _____	143

Lista de Figuras

Figura 1 - Esquema básico de uma rede neuronal de retropropagação. _____	8
Figura 2 – Esquema das operações envolvidas em um elemento de processamento (neurônio) típico para redes neuronais. _____	9
Figura 3 – Exemplo de uma rede neuronal LVQ – Learning Vector Quantization. (BULSARI, 1995) _____	15
Figura 4 - Fluxograma simplificado do processo de fabricação do papel e celulose _____	22
Figura 5 – Sistema Reator (CSTR) e Sedimentador com reciclo de material celular (METCALF & EDDY, 1991). _____	31
Figura 6 - Degradação da matéria orgânica (PESSSOA & JORDÃO, 1982) _____	32
Figura 7 - Histograma da idade em anos _____	50
Figura 8 - Histograma da produtividade (em esteres/hectare.ano) _____	50
Figura 9 - Histograma da densidade básica (em kg matéria seca/m ³) _____	50
Figura 10 - Histograma de extrativos (em %massa) _____	50
Figura 11 - Histograma de lignina solúvel (em %massa) _____	50
Figura 12 - Histograma de lignina insolúvel (em %massa) _____	50
Figura 13 - Histograma de holocelulose (em %massa) _____	51
Figura 14 - Histograma de rendimento depurado (em %massa) _____	51
Figura 15 - Histograma de teor de rejeitos (em %massa) _____	51
Figura 16 - Histograma do índice kappa _____	51
Figura 17 - Histograma de viscosidade (em cP) _____	51
Figura 18 - Histograma de alcalinidade residual (em g/l) _____	51
Figura 19 - Histograma do código de fazenda _____	52

Figura 20 - Comparação dos erros RMS para o conjunto de teste e de treinamento, ao longo dos ciclos de treinamento para uma das redes treinadas. _____	58
Figura 21 - Comparação dos erros RMS para o conjunto de teste e de treinamento, mudando-se o número de neurônios na camada interna. _____	59
Figura 22 – Tela do programa de treinamento para a rede de predição de propriedades. Ponto onde o erro do conjunto de teste é o menor encontrado. _____	60
Figura 23 – Tela do programa de treinamento para a rede de predição de propriedades. Prosseguindo o treinamento além do ponto onde o erro do conjunto de teste é o menor encontrado observa-se a elevação deste erro. _____	61
Figura 24 – Comparação entre o valor obtido pela rede neuronal e o valor desejado para o rendimento depurado nos conjuntos de treino e de teste. Indica-se os desvios de 1% e 5% no valor da ordenada da linha $y=x$. _____	62
Figura 25 – Comparação entre o valor obtido pela rede neuronal e o valor desejado para o número kappa nos conjuntos de treino e de teste. Indica-se os desvios de 1% e 5%. _____	62
Figura 26 – Comparação entre o valor obtido pela rede neuronal e o valor desejado para a viscosidade nos conjuntos de treino e de teste. Indica-se os desvios de 1% e 5% no valor da ordenada da linha $y=x$. _____	63
Figura 27 – Comparação entre o valor obtido pela rede neuronal e o valor desejado para o álcali ativo residual nos conjuntos de treino e de teste. Indica-se os desvios de 1% e 5%. _____	63
Figura 28 – Tela do programa de treinamento da rede LVQ para a classificação segundo o especialista Trovó. _____	65
Figura 29 – Tela do programa de treinamento da rede LVQ para a classificação segundo o especialista Gilberto. _____	66
Figura 30 – PREDPROP programa para a predição das propriedades após a polpação _____	69
Figura 31 – PREVDADO programa para fornecer a classificação dos lotes, sem usar a predição. _____	70

Figura 32 – ANTOTAL programa que gera um arquivo com todas as informações disponíveis sobre os lotes de matéria-prima, usando a predição das propriedades.	71
Figura 33– Fluxograma simplificado da unidade de tratamento de efluentes	85
Figura 34- Fluxograma simplificado da unidade de tratamento de efluentes, constando o código das amostras coletadas.	86
Figura 35 – Esquema simplificado de uma rede neuronal do tipo “back-propagation”, para o caso estudado.	86
Figura 36 - REATOR 1 – Valor predito do TOC (em mg/l) da amostra 203 em relação ao valor encontrado.	94
Figura 37 - REATOR 1 – Valor predito do TOC da amostra 203 em relação ao valor encontrado. Escala ampliada para valores inferiores a 400 mg/l.	95
Figura 38 - REATOR 2 – Valor predito do TOC (em mg/l) da amostra 204 em relação ao valor encontrado.	95
Figura 39 - REATOR 2 – Valor predito do TOC da amostra 204 em relação ao valor encontrado. Escala ampliada para valores inferiores a 400 mg/l.	96
Figura 40- REATOR 3 – Valor predito do TOC (em mg/l) da amostra 205 em relação ao valor encontrado.	96
Figura 41 - REATOR 3 – Valor predito do TOC da amostra 205 em relação ao valor encontrado. Escala ampliada para valores inferiores a 400 mg/l.	97
Figura 42 - LAGOA – Valor predito do TOC (em mg/l) da amostra 209 em relação ao valor encontrado.	97
Figura 43 - REATOR 1A – Valor predito do TOC (em mg/l) da amostra 203A em relação ao valor encontrado.	98
Figura 44 - REATOR 1A – Valor predito do TOC da amostra 203A em relação ao valor encontrado. Valores inferiores a 400 mg/l.	99
Figura 45 - REATOR 1A – Valor predito do TOC (em mg/l) da amostra 203A em relação ao valor encontrado. Treinada até obter erro de treino muito pequeno.	99

Figura 46 - REATOR 1B – Valor predito do TOC (em mg/l) da amostra 203B em relação ao valor encontrado. _____	100
Figura 47 - REATOR 1B – Valor predito do TOC da amostra 203B em relação ao valor encontrado. Valores inferiores a 400 mg/l. _____	101
Figura 48 - REATOR 1A – Valor predito do TOC (em mg/l) da amostra 203A em relação ao valor encontrado. Utiliza-se como variável de entrada a carga do reator. _____	102
Figura 49 - REATOR 1A – Valor predito do TOC da amostra 203A contra o valor observado. Utiliza-se como variável de entrada a carga do reator. Valores inferiores a 400 mg/l. _____	102
Figura 50 – Acompanhamento da vazão de TOC em kg/h ao longo dos dias do período entre jun/98 e out/98. _____	103
Figura 51 - REATOR 1A – Valor predito do TOC da amostra 203A em relação ao valor encontrado. Rede treinada para o segundo conjunto de dados sem usar conjunto de teste, e com todas as variáveis propostas. _____	107
Figura 52 - REATOR 1A – Valor predito do TOC da amostra 203A em relação ao valor encontrado para o conjunto de avaliação . Rede treinada para o segundo conjunto de dados sem usar conjunto de teste, e com todas as variáveis propostas. _____	107
Figura 53 – Tela do programa para treinar a rede do reator 1A _____	110
Figura 54 – Tela do programa para treinar a rede do reator 1B _____	110
Figura 55 – Tela do programa para treinar a rede do reator 3A _____	111
Figura 56 – Tela do programa para treinar a rede do reator 3B _____	111
Figura 57 – Tela do programa para treinar a rede da lagoa _____	112
Figura 58– Predição do TOC para o reator 1A _____	112
Figura 59– Predição do TOC para o reator 1B _____	113
Figura 60- Predição do TOC para o reator 2A _____	113
Figura 61- Predição do TOC para o reator 2B _____	113

Figura 62- Predição do TOC para o reator 3A	114
Figura 63- Predição do TOC para o reator 3B	114
Figura 64- Predição do TOC para a lagoa de polimento	114
Figura 65– Influência do pH no TOC residual do reator 1A predito pela rede	115
Figura 66 – O Menu principal e as opções de entrada de dados	121
Figura 67 – As opções para a predição e saída de resultados	121
Figura 68 - Entrada de dados para o reator 1A referente ao dia 0.	122
Figura 69 - Entrada de dados para o reator 1B referente ao dia 0.	122
Figura 70 - Entrada de dados para o reator 2A referente ao dia 0. Para a entrada do dia 1 a tela de diálogo é a mesma, mudando somente o campo correspondente.	123
Figura 71 - Entrada de dados para o reator 2B referente ao dia 0. Para a entrada do dia 1 a tela de diálogo é a mesma, mudando somente o campo correspondente.	123
Figura 72 - Entrada de dados para o reator 3A referente ao dia 0. Para a entrada dos dias 1 e 2 a tela de diálogo é a mesma, mudando somente o campo correspondente.	124
Figura 73 - Entrada de dados para o reator 3B referente ao dia 0. Para a entrada dos dias 1 e 2 a tela de diálogo é a mesma, mudando somente o campo correspondente.	125
Figura 74 - Entrada de dados para a lagoa referente ao dia 0. Para a entrada dos dias 1, 2 e 3 a tela de diálogo é a mesma, mudando somente o campo correspondente.	125
Figura 75 - Resultado para a predição de quatro dias. As predições para 1, 2 e 3 dias geram janela idêntica.	126

Lista de Tabelas

Tabela 1 - Contaminantes importantes no tratamento de águas residuárias. _____	28
Tabela 2 - Características físicas, químicas e biológicas das águas residuárias e suas fontes. _____	29
Tabela 3 - Distribuição das ocorrências dos testes segundo a fazenda e espécie. _____	53
Tabela 4 - Resultados para a análise da variância do conjunto de dados, mostrando a variabilidade do <u>rendimento depurado</u> entre as amostras/análises para cada talhão. _____	54
Tabela 5 - Resultados para a análise da variância do conjunto de dados, mostrando a variabilidade do <u>índice Kappa</u> entre as amostras/análises para cada talhão. _____	55
Tabela 6 - Resultados para a análise da variância do conjunto de dados, mostrando a variabilidade da <u>viscosidade</u> entre as amostras/análises para cada talhão. _____	55
Tabela 7 - Resultados para a análise da variância do conjunto de dados, mostrando a variabilidade da <u>alcalinidade residual</u> entre as amostras/análises para cada talhão. _____	56
Tabela 8 - Resultados para a análise da variância do conjunto de dados, mostrando a variabilidade do <u>teor de rejeitos</u> entre as amostras/análises para cada talhão. _____	56
Tabela 9- Estatística dos desvios dos resultados apresentados pela rede neuronal _____	64
Tabela 10 – Valores dos escores adotados na Classificação proposta neste trabalho, para cada variável e faixa de valores. _____	67
Tabela 11 - Fazendas e espécies selecionadas para os experimentos, e o talhão correspondente _____	75
Tabela 12 - Níveis inferior e superior dos fatores a testar _____	76
Tabela 13 - Estatística Descritiva e níveis inferior e superior dos fatores. _____	79
Tabela 14 - Estatística Descritiva das variáveis resposta. _____	79

Tabela 15 - Extrato da saída do programa MINITAB para análise do experimento incluindo interações de até terceira ordem. _____	80
Tabela 16 - Quadro resumo com os fatores de menor importância. Os locais assinalados significam que o fator não tem muita influência sobre aquela variável saída da rede. _____	82
Tabela 17 - Quadro resumo com os fatores de maior importância. Os locais assinalados significam que o fator tem muita influência sobre aquela variável saída da rede. _____	83
Tabela 18 – Resultados obtidos com dados de Jun/98 a Out/98, para as variáveis pH e TOC. Para o primeiro reator o TOC é o da amostra 203, para o segundo o da amostra 204, para o terceiro o da amostra 205, e o da lagoa é o TOC da amostra 209. _____	94
Tabela 19– Resultados obtidos no treinamento da rede para predizer o TOC (em mg/l) do sistema de tratamento de efluentes. _____	109
Tabela 20 – Limites e média dos valores utilizados no treinamento das redes, para o reator 1-A. _____	117
Tabela 21 – Limites e média dos valores utilizados no treinamento das redes, para o reator 1-B. _____	118
Tabela 22 – Limites e média dos valores utilizados no treinamento das redes, para o reator 2-A. _____	118
Tabela 23 – Limites e média dos valores utilizados no treinamento das redes, para o reator 2-B. _____	119
Tabela 24 – Limites e média dos valores utilizados no treinamento das redes, para o reator 3-A. _____	119
Tabela 25 – Limites e média dos valores utilizados no treinamento das redes, para o reator 3-B. _____	120
Tabela 26 – Limites e média dos valores utilizados no treinamento das redes, para a lagoa. _____	120
Tabela 27 – Valores medidos e preditos pela rede neuronal para o TOC da amostra 209 com 4 dias de antecedência. _____	127

Tabela 28 – Dados fornecidos pela RIPASA SA, para o estudo da classificação da matéria-prima	144
Tabela 29 - Tabela gerada pelo programa ANTOTAL, utilizando-se o conjunto dos pontos com as médias por talhão.	151
Tabela 30 – Tabela gerada pelo programa ANTOTAL, utilizando-se o conjunto de todos os pontos (e não as médias). Observe-se que pela rede de predição ter sido treinada para as médias apenas estes resultados não são tão bons, mas a classificação obtida desta forma não fica tão prejudicada.	153

Nomenclatura

Latinas

d_k valor desejado da variável de saída para o dado k .

$e_k^{(o)}$ erro gerado para o dado k para a camada de saída (output).

E erro global.

$E[k]$ valor do erro em um instante k .

$f()$ função que estabelece uma relação entre as entradas e a saída de cada neurônio. Usualmente ela é chamada de função de ativação do neurônio.

$I_j^{[s]}$ somatório ponderado das entradas do j -ésimo neurônio na camada s .

K_S constante de meia velocidade, é a concentração do substrato quando se atinge a metade da taxa de crescimento máxima.

k taxa máxima de utilização do substrato por unidade de massa de substrato.

k_d coeficiente de decaimento endógeno, em unidades de tempo⁻¹.

$lcoef$ é o coeficiente de treinamento.

o_k valor obtido da variável de saída para o dado k .

Q vazão do influente.

Q_w vazão da corrente que descarta o lodo em excesso que não é reciclado.

R coeficiente de correlação linear.

(r'_g) taxa líquida de crescimento dos microorganismos.

(r_{S_0}) taxa de utilização do substrato.

S_0 concentração inicial do substrato limitante do crescimento no líquido.

S concentração do substrato limitante do crescimento no líquido após um tempo θ .

V volume do reator.

V_s volume do reator acrescido do valor do volume do sedimentador.

$w_{ji}^{[k]}$ peso da conexão unindo o i -ésimo neurônio na camada $(s-1)$ com o j -ésimo neurônio na camada s .

$w[k]$ valor do peso no instante k .

X concentração de microorganismos no reator, de onde é descartada a quantidade Q_W .

X_e concentração de microorganismos no efluente da unidade de sedimentação.

X_0 concentração de microorganismos na corrente de entrada da unidade.

$x_j^{[s]}$ valor corrente de saída do j -ésimo neurônio na camada s .

Y coeficiente de máxima produção, definido como a razão entre a massa de células formadas e a massa de substrato consumido.

Gregas:

$\alpha[k]$ taxa de treinamento para a conexão no instante k .

$\alpha[0]$ taxa de treinamento inicial.

$\alpha[bound]$ valor limite a que pode chegar a taxa de treinamento.

$\Delta\alpha[k]$ variação da taxa de treinamento da conexão no instante k .

$\Delta w[k]$ variação no peso w no instante k .

$\delta[k]$ termo do gradiente na variação do peso.

$\bar{\delta}[k]$ média ponderada do termo do gradiente na variação do peso.

κ razão de incremento na taxa de treinamento.

φ razão de decréscimo na taxa de treinamento.

θ fator para transformação dos pesos (convex).

μ_1 média do tratamento 1.

μ_2 média do tratamento 2.

SIGLAS

AAR	Alcali ativo residual.
Adj MS	é a média dos quadrados ajustada.
Adj SS	representa a soma dos quadrados ajustada.
ANOVA	Análise da Variância.
COBEQ	Congresso Brasileiro de Engenharia Química
DBD	algoritmo delta-bar-delta .
DBO	demanda bioquímica de oxigênio.
DF	representa os graus de liberdade.
DQO	demanda química de oxigênio.
EDVAC	Electronic Discrete Variable Automatic Computer.
ENIAC	Electronic Numerical Integrator And Computer.
Epoch	número de pontos para cada ciclo de atualização dos pesos.
LVQ	Learning Vector Quantization.
MÉS	fração de sólidos voláteis suspensos.
PE	elementos de processamento, neurônios da rede.
PID	Controle Proporcional-Integral-Derivativo.
Quickprop	Algoritmo de treinamento para redes (FAHLMAN, 1988).
RMS	refere-se ao erro médio quadrático como definido no capítulo 3.
RNA	Redes Neurais Artificiais
SDCD	Sistema Digital de Controle Distribuído.
TALHAOC	se refere ao código do talhão correspondente.
TOC	carbono orgânico total.

CAPÍTULO 1. - INTRODUÇÃO

1.1. JUSTIFICATIVA DA ESCOLHA DO TEMA

As redes neuronais têm sido recentemente alvo de muita atenção, em grande parte por causa da sua enorme aplicabilidade e de sua facilidade de resolver problemas complexos e não-lineares. Na década de 80 as pesquisas, envolvendo principalmente simulação de processos, mostravam um relativo esforço em utilizar técnicas como a colocação ortogonal e outras que conseguiam reduzir a carga computacional para resolução numérica de modelos determinísticos através da discretização em número reduzido de pontos dentro do intervalo desejado. Porém, na década de 90, mesmo com o avanço tecnológico dos equipamentos computacionais, surgiram pesquisas comprovando a viabilidade de se utilizar a técnica das redes neuronais aos diversos campos do conhecimento. No campo da Engenharia Química esta tecnologia encontra grandes possibilidades e deverá em breve representar um impacto significativo na abordagem dos problemas (BHAT,1990). Uma constatação deste fato se faz observando o que está ocorrendo com os principais criadores de programas de simulação: estão se associando ou mesmo comprando as empresas que criam programas que usam redes neuronais. Não será surpresa quando, em pouco tempo, estes simuladores já vierem com módulos próprios para redes neuronais. Diante desta perspectiva, e afirmando que não se deve querer aplicar esta tecnologia indiscriminadamente, deseja-se estabelecer alguns critérios para que se possa avaliar melhor quando podem ser obtidos resultados satisfatórios com as redes neuronais. Basicamente, indica-se o seu emprego quando os modelos determinísticos ou mesmo os empíricos apresentam dificuldades de serem implementados ou desenvolvidos (LEE, 1994).

Nesta tese pretende-se demonstrar duas aplicações de redes neuronais a processos distintos que foram realizados. Uma breve introdução ao assunto de interesse é apresentada, envolvendo os aspectos que vão ser abordados. O primeiro estudo refere-se a uma classificação da matéria-prima da polpação da madeira em indústria de papel e celulose. Neste caso houve a colaboração da Ripasa Papel e Celulose SA, indústria química situada em Limeira, São Paulo. Como os critérios propostos para classificação exigia algumas variáveis obtidas ao final do processo, de determinação experimental laboriosa, utiliza-se uma rede para prever estes dados em função de outras propriedades da madeira que são

mais fáceis de obter. Outra rede foi usada para processar tais informações e classificar enfim as amostras de madeira. Uma análise de sensibilidade foi realizada para avaliar as variáveis mais importantes a serem medidas e as que poderiam ser excluídas do estudo. Este trabalho está no Capítulo 3.

No outro processo contamos com a parceria da Rhodiaco Ltda., indústria química que situa-se em Paulínia, São Paulo, que forneceu dados e informações valiosas para a realização do estudo. Neste caso, deseja-se prever as condições do efluente final de acordo com as condições atuais de entrada e as outras condições operacionais. Foram criadas várias redes, individuais para cada um dos equipamentos da estação de tratamento. Com isto foi necessária a criação de um programa para utilizar tais redes treinadas. Uma análise da influência das entradas da rede também foi realizada aqui, mas de uma forma diferente daquela do processo anterior. Este trabalho é abordado no Capítulo 4.

1.2. PROPOSTA E OBJETIVOS DO TRABALHO:

Destacam-se os principais objetivos globais do trabalho desenvolvido com sendo:

- aprender a utilizar a técnica das redes neuronais.
- estabelecer uma visão crítica sobre a aplicação de redes neuronais em casos diversos.
- tentar integrar a utilização de redes neuronais a outras técnicas de simulação e análise de processos.
- avaliar de maneira crítica a utilização de redes neuronais a processos químicos, indicando possíveis soluções aos problemas encontrados.

Para isto a proposta de trabalho envolve os seguintes itens:

A. **Técnicas a utilizar:** redes neuronais e estatística.

B. **Processos:**

- Processo de polpação da madeira na Indústria de Papel e Celulose, Ripasa SA.
- Reatores de lodo ativado, Processo do tratamento de efluentes da Rhodiaco Ind. Quim. Ltda.

C. Objetivos específicos:

- Otimizar a produção e a qualidade do produto.
- Diminuir sistematicamente a carga da unidade de tratamento de efluentes.
- Aumentar o nível de controle sobre os processos.

D. Dados:

- Utilizar dados históricos, fornecidos pelas indústrias.

E. Programas (“Softwares”) a utilizar:

- Neuralworks Pro II Plus, com os Pacotes: (Designer Pack e UDND), da Neuralware; para o treinamento e geração da rede neuronal.
- Datasculptor, da Neuralware, para preparação dos dados e análise de resultados.
- Microsoft Excel, também para preparação e análise de dados.
- Microcal Origin, tanto para a análise de dados como para a construção dos gráficos.
- Compiladores Fortran e C, para construção do programa para utilizar a rede treinada.
- Outros pacotes estatísticos: Minitab, Statistica,...

CAPÍTULO 2. REVISÃO BIBLIOGRÁFICA

2.1. REDES NEURONAIS:

2.1.1. Filosofia e histórico

Computadores são mais eficientes fazendo determinados tipos de trabalho que seres humanos. Porém em outras tarefas ocorre o contrário. Os melhores computadores, em algumas tarefas, não conseguem nem chegar perto da performance humana, como por exemplo o reconhecimento de faces. Por outro lado, podem desenvolver outras tarefas de forma mais rápida e confiável do que um conjunto de humanos, como o controle de contas bancárias. Pode-se entender a diferença básica entre os dois sistemas de processamento. Um computador geralmente consiste em um único processador operando, executando instruções, escritas por um programador, uma de cada vez em sequência. Os sistemas nervosos humanos consistem em bilhões de células nervosas (neurônios), intensamente conectadas entre si, realizando operações relativamente simples e sem o benefício de um programador.

As redes neuronais são tentativas humanas de simular e entender o que se passa em um sistema nervoso, com a esperança de que se possam capturar algumas das potencialidades desses sistemas biológicos.

A capacidade do cérebro aprender a partir de estímulos externos é alvo de estudos desde a antiguidade. Com o início do entendimento da estrutura celular cerebral quase no início do século (CAJAL, 1894) e da construção de dispositivos automáticos para cálculo, os pesquisadores começaram a desejar uma máquina que pudesse realizar tarefas e pensar como um ser humano. Em 1943, MCCULLOCH e PITTS descreveram um modelo de cálculo lógico de uma rede neuronal. Além de orientar todo o desenvolvimento posterior de redes neuronais, este trabalho também ajudou os pesquisadores da Universidade da Pensilvânia a construir o EDVAC (Eletronic Discrete Variable Automatic Computer), e o seu sucessor o ENIAC (Eletronic Numerical Integrator And Computer), entre 1943 e 1946. O trabalho de McCulloch e Pitts também inspirou outros pesquisadores como Marvin Minsky na área de inteligência artificial e Norbert Wiener na área de cibernética.

Em 1958, ROSEMBLAT publicou um trabalho onde apresenta-se a estrutura do perceptron, uma unidade de reconhecimento de padrões que é a base do neurônio de uma rede neuronal. As técnicas de treinamento foram se aperfeiçoando e novas aplicações foram propostas, tendo um período de grande avanço até meados de 1969. Nesta época foi publicado o livro *Perceptrons* de Minsky e Pappert (1969), que enfatizaram em demasia as limitações dos perceptrons isolados, o que desencorajou pesquisadores e agências de financiamento a investirem em redes neurais. Na década de 70 os custos de computação eram elevados, sendo um período em que muito pouco se desenvolveu na área, como no caso de reconhecimento de imagens, restando apenas aplicações de reconhecida eficácia (YAMAMOTO, 1998).

Após a publicação do livro “*Parallel Distributed Processing: Exploration in Microstructures of Cognition*”, de RUMELHART e MCCLELLAND (1986), resurgiu o desenvolvimento das redes neurais com trabalhos importantes em diversas áreas e com a disseminação da técnica pelo mundo. Um histórico sobre redes neurais também pode ser encontrado em NELSON e ILLINGWORTH (1991), HAYKIN (1994), DAYHOFF (1990), KOVÁCS (1996) e HUNT et al. (1992).

O futuro da aplicação de redes neurais é promissor pelas suas características principais. Como são funções que estimam relações do tipo entrada-saída de um sistema, constituem-se em mais uma ferramenta possível para a modelagem de processos. Diz-se que elas “aprendem” essas relações a partir de um processo de “treinamento” semelhante ao aprendizado de um cérebro humano, sem necessitar da idealização, elaboração e validação de modelos matemáticos complexos.

Uma outra característica é sua estrutura de distribuição paralela de informações. Concebida de forma semelhante ao cérebro, esse paralelismo vem atraindo a atenção de inúmeros pesquisadores de todos os setores do conhecimento, em particular os da área de informática. Já se sabe que os computadores só conseguirão aumentar sua capacidade e velocidade de cálculo adotando-se uma estrutura essencialmente paralela de fluxo de informações, como a que ocorre no tecido nervoso cerebral. Hoje já existem os “chips” neurais e em breve, computadores pessoais baseados nesses “chips” começarão a surgir no mercado (YAMAMOTO, 1998).

2.1.2. Classificação de tipos

Em uma tentativa de estabelecer uma classificação dos tipos existentes de redes neuronais duas características básicas são levadas em consideração:

- A forma como a informação flui pela rede. Em algumas ela flui sempre em um sentido único, da entrada para a saída. Em outras a informação flui independente de sentido ou direção.
- A forma como os resultados são interpretados, quando a rede neuronal compara as suas respostas com uma desejada durante o seu treinamento, ou quando a rede neuronal por si própria estabelece a melhor forma de encontrar uma resposta.

As redes onde as informações avançam em apenas uma direção são denominadas de diretas (“feedforward”). São bastante comuns por sua relativa simplicidade e estabilidade. As redes conhecidas como “backpropagation” são exemplos deste tipo.

Aquelas redes onde as conexões permitem que as informações fluam em qualquer sentido são chamadas de recorrentes (“feedback”). Em geral, as redes deste tipo são capazes de representar sistemas mais complexos, mas isto pode ser às custas de complicações indesejáveis para o treinamento da rede.

A forma como são treinadas as redes também é uma característica usada em sua classificação. O método mais comum para treinar uma rede é alimentá-la com dados de entrada e com os dados de saída desejados. Os primeiros resultados são produzidos pela rede e comparados com os resultados desejados gerando os erros. Estes são usados para ajustar os pesos na rede de forma que a próxima vez que o mesmo dado for usado, a rede se aproximará da resposta desejada. Esta forma de treinamento é chamada de supervisionada, e é o caso das redes diretas.

Em algumas redes do caso inverso, não se fornece o resultado desejado, mas a ela é permitido organizar os dados numa forma que ela entenda estarem ajustados. Redes deste tipo são chamadas de não supervisionadas, ou ainda de auto organizadas. Algumas redes têm uma fase inicial auto organizada seguida de uma fase supervisionada.

2.1.3. Estrutura básica

Redes neurais são constituídas por neurônios. Outro termo freqüentemente usado para neurônio é nó, ou ainda pela terminologia preferida de alguns programas aplicativos, elementos de processo (PE). Estes são arranjados em camadas, e os neurônios de uma camada podem estar conectados a muitos outros em diversas outras camadas, como pode ser observado na Figura 1.

Em alguns casos, os neurônios podem estar conectados a outros da mesma camada, ou até mesmo com ele próprio. Cada neurônio processa as entradas que ele recebe por estas conexões, e fornece um valor analógico contínuo a outros neurônios através de suas conexões de saída. Como em sistemas biológicos, a força dessas conexões são variadas. Criar uma rede neuronal consiste em decidir como os neurônios são conectados, como estes processam as informações e como as intensidades das conexões devem ser modificadas.

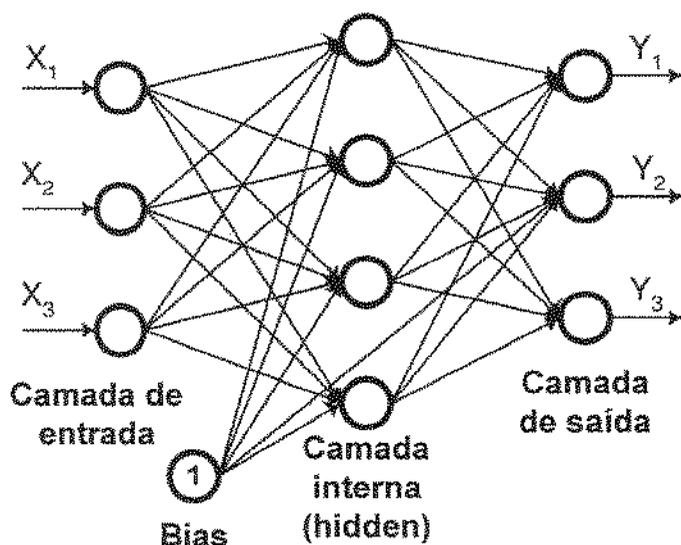


Figura 1 - Esquema básico de uma rede neuronal direta.

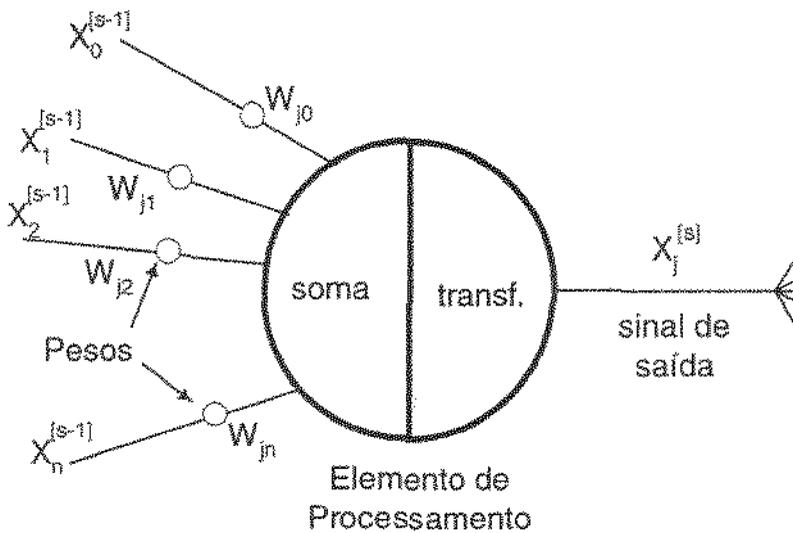


Figura 2 – Esquema das operações envolvidas em um elemento de processamento (neurônio) típico para redes neurais.

Na Figura 1 X_i e Y_i correspondem as entradas e saídas da RN, respectivamente. Os círculos representam os neurônios, que são normalmente distribuídos em camadas consecutivas. Cada conexão entre os neurônios tem um valor associado (denominado como peso). Estes pesos são ajustados durante o treinamento da rede neuronal, definindo o conhecimento acumulado por ela.

Um elemento na rede direta, ilustrado na Figura 2, transfere suas entradas com as seguintes operações:

$$x_j^{[s]} = f\left(\sum_i (w_{ji}^{[s]} \cdot x_i^{[s-1]})\right) \quad (1)$$

ou na forma simplificada

$$x_j^{[s]} = f(I_j^{[s]}) \quad (2)$$

onde

$x_j^{[s]}$ = valor corrente de saída do j -ésimo neurônio na camada s ,

$w_{ji}^{[s]}$ = peso da conexão unindo o i -ésimo neurônio na camada $(s-1)$ com o j -ésimo neurônio na camada s ,

$I_j^{[s]}$ = somatório ponderado das entradas do j -ésimo neurônio na camada s e

$f()$ = função que estabelece uma relação entre as entradas e a saída de cada neurônio. Usualmente ela é chamada de função de ativação do neurônio.

Na notação aqui utilizada, o superscrito entre colchetes indica a camada da rede que está sendo considerada. A função f mais intensamente empregada é a função sigmoideal, mas pode ser qualquer função que seja diferenciável. A função sigmoideal é definida como

$$f(z) = (1 + e^{-z})^{-1} \quad (3)$$

e outras funções também empregadas são:

Linear $f(z) = Az + B$ (4)

Tangente hiperbólica $f(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$ (5)

Limite ("threshold") $f(z) = +1$ se $z > 0$ senão $= 0$ (6)

Gaussiana $f(z) = e^{\left(\frac{-z^2}{\sigma^2}\right)}$ (7)

Função de base radial $f(z) = e^{\left(\frac{-z^2}{\beta^2}\right)}$ (8)

A propagação do erro local para as camadas internas segue a seguinte relação:

$$e_j^{[s]} = f'(I_j^{[s]}) \cdot \sum_k (e_k^{[s+1]} \cdot w_{kj}^{[s+1]}) \quad (9)$$

Note-se que a equação 9 não pode ser usada para as camadas de saída. O erro local na camada externa seria simplesmente a diferença entre o valor desejado para a variável e o valor obtido, contudo aplica-se um fator de escala, $f'(I_k)$, obtendo o erro a ser utilizado,

$$e_k^{(o)} = (d_k - o_k) \cdot f'(I_k) \quad (10)$$

onde $e_k^{(o)}$ = erro gerado para o dado k para a camada de saída (output),

d_k = valor desejado da variável de saída para o dado k ,

o_k = valor obtido da variável de saída para o dado k .

O objetivo do processo de treinamento é minimizar o erro global, E , do sistema pelo ajuste dos pesos. Para isto, na verdade se utilizará dos erros locais em cada um dos elementos (neurônios). Dado um conjunto de pesos $w_{ji}^{[s]}$, necessita-se determinar como

incrementá-los ou decrementá-los de forma a reduzir o erro global. Uma das formas de estimar esse incremento é usar a regra do gradiente descendente:

$$\Delta w_{ji}^{[s]} = -lcoef \left(\frac{\partial E}{\partial w_{ji}^{[s]}} \right) \quad (11)$$

onde *lcoef* é o coeficiente de treinamento. Em outras palavras, cada peso muda de acordo com a intensidade e direção do gradiente negativo na superfície do erro. A derivada parcial pode ser decomposta e colocada em termos do erro local, $e_j^{[s]}$:

$$\left(\frac{\partial E}{\partial w_{ji}^{[s]}} \right) = -e_j^{[s]} \cdot x_i^{[s-1]} \quad (12)$$

ou ainda combinando as duas últimas,

$$\Delta w_{ji}^{[s]} = lcoef \cdot e_j^{[s]} \cdot x_i^{[s-1]} \quad (13)$$

A correção dos valores dos pesos pode ser feita após a apresentação de cada um dos pares de dados à rede, ou ainda acumula-se tais correções para um conjunto de *n* dados. A esse conjunto de dados denomina-se de “epoch”.

A regra de treinamento chamada de delta-bar-delta (DBD) é um aprimoramento da regra convencional que foi sugerido inicialmente por JACOBS (1988), e se utiliza uma notação simplificada com algumas características peculiares, assim:

$E[k]$ = valor do erro em um instante *k*

$w[k]$ = valor do peso no instante *k*

$\Delta w[k]$ = variação no peso *w* no instante *k*

$\alpha[k]$ = taxa de treinamento para a conexão no instante *k*

$\Delta \alpha[k]$ = variação da taxa de treinamento da conexão no instante *k*

$\delta[k]$ = termo do gradiente na variação do peso

$\bar{\delta}[k]$ = média ponderada do termo do gradiente na variação do peso

θ = fator para transformação dos pesos (convex)

κ = razão de incremento na taxa de treinamento

φ = razão de decréscimo na taxa de treinamento.

Como no método convencional da retropropagação, o componente do gradiente é definido como:

$$\delta[k] = \frac{\partial E[k]}{\partial w[k]} \quad (14)$$

A equação convencional para atualização dos pesos pela regra Delta é:

$$w[k+1] = w[k] + \alpha \delta[k] \quad (15)$$

Neste caso, α seria uma constante fixa. No caso da regra DBD a taxa de treinamento é definida para cada conexão, e a equação de atualização dos pesos torna-se:

$$w[k+1] = w[k] + \alpha[k] \cdot \delta[k] \quad (16)$$

Para implementar a heurística que deverá incrementar e decrementar as taxas de treinamento para cada conexão, Jacobs empregou uma média ponderada dos componentes do gradiente tal como,

$$\bar{\delta}[k] = (1 - \theta) \delta[k] + \theta \delta[k - 1] \quad (17)$$

e ainda:

$$\Delta \alpha[k] = \begin{cases} \kappa & \bar{\delta}[k - 1] \delta[k] > 0 \\ -\varphi \alpha[k] & \bar{\delta}[k - 1] \delta[k] < 0 \\ 0 & \text{para } = 0 \end{cases} \quad (18)$$

onde κ e φ são constantes. Esta regra aumenta a taxa de treinamento de forma linear, mas o decréscimo é geométrico. Assim, quando deseja-se apressar a convergência deve-se modificar o $\alpha[0]$, o $\alpha[bound]$ (que são os valores inicial e limite para $\alpha[k]$, respectivamente) e o κ para valores mais elevados. Pode-se ainda experimentar baixar o valor de φ , mas este parâmetro controla sensivelmente as oscilações no método. Portanto, para eliminar oscilações no erro deve-se reduzir os valores de $\alpha[0]$, $\alpha[bound]$, κ e de φ .

Em síntese, o processo de treinamento da rede se constitui basicamente em um problema de otimização e exige a apresentação da base de dados à rede por diversas vezes, até que o conjunto de pesos encontrados satisfaça a condição desejada para o erro apurado.

Desta forma, as mesmas dificuldades associadas à otimização de processos também

ocorrem na fase de aprendizado da rede, tais como problemas de convergência, existência de mínimos locais e tempo de computação indeterminado.

Portanto, a modelagem com redes neuronais envolve as seguintes etapas principais:

- análise do problema, configuração das variáveis de entrada e de saída,
- coleta e seleção de dados a serem empregados,
- definição da estrutura da RN mais adequada,
- divisão dos dados em grupos para treinamento, teste e validação,
- treinamento da RN utilizando os dados que representam o processo,
- validação da rede treinada,
- utilização da rede treinada no simulador.

Uma importante fase para obter-se um resultado satisfatório quando se trabalha com redes neuronais é a definição do número de camadas e do número de neurônios em cada uma delas. Não existe ainda nenhuma regra que esclareça satisfatoriamente a forma correta de configurar esses parâmetros. Contudo, pode-se afirmar que quanto mais complexa a rede criada, maior sua capacidade de ajustar-se aos dados, e maior também é o tempo para treiná-la. Mas um número elevado de neurônios, ou mesmo a adição de mais camadas podem não significar um melhor resultado. Isto porque desta forma, com mais pesos a serem treinados, é bem possível que venha a ocorrer o chamado “overfitting”, onde a rede representa muito bem os dados usados no treinamento, mas perde a capacidade de generalizar para outras situações. Assim, aliado a outros cuidados, recomenda-se sempre a estrutura mais simples possível.

2.1.4. Redes Neuronais “LVQ – Learning Vector Quantization”

Este tipo de rede neuronal é utilizada para a classificação da madeira a partir das propriedades estimadas da polpa no estudo referente a Ripasa S/A. Originalmente sugerida por KOHONEN (1988), esta rede contém uma camada interna, que também é referida como camada Kohonen, onde a classificação é apreendida e realizada. Neste tipo de rede, cada parâmetro a utilizar corresponde a um neurônio da camada de entrada, e na camada de

saída cada neurônio corresponde a uma classe. A Figura 3 mostra a estrutura de uma rede LVQ (BULSARI, 1995). A rede é totalmente interligada entre a camada de entrada e a camada interna, e parcialmente interligada entre a camada interna e a camada de saída, com cada neurônio de saída conectado em distintos grupos de neurônios da camada interna. Os pesos das conexões entre a camada interna e a de saída são todos iguais a 1. Os pesos das conexões entre a camada de entrada e a interna formam as componentes do vetor de referência. Um vetor de referência é associado a cada um dos neurônios da camada interna. Estes se modificam ao longo do treinamento da rede. As saídas dos neurônios da camadas interna e de saída são binárias. Quando um padrão de entrada é fornecido à rede, o neurônio da camada interna que apresentar o vetor de referência mais próximo deste padrão é dito como como neurônio “vencedor”, e sua saída é a única a produzir o valor 1. Todos os outros neurônios da camada intermediária são forçados a produzir o valor 0. Da mesma forma, o neurônio de saída que estiver conectado ao neurônio vencedor será o único a emitir o valor unitário. Este neurônio de saída indica, portanto, a classe para um determinado padrão de entrada. Cada um dos neurônios de saída refere-se exclusivamente a uma outra classe diferente.

Um procedimento de treinamento simples para uma LVQ seria:

- (i) inicializar os pesos dos vetores de referência,
- (ii) apresentar um padrão de entrada para o treinamento da rede,
- (iii) calcular a distância Euclideana entre o padrão de entrada e cada um dos vetores de referência,
- (iv) atualizar os pesos do vetor de referência mais próximo do padrão de entrada, que é o vetor de referência do neurônio vencedor da camada interna. Se este neurônio pertencer ao grupo ligado a classe correta (neurônio correto da camada de saída), então o vetor de referência pode ser aproximado ao padrão de entrada. Caso contrário, ele é afastado do padrão,
- (v) retornar a (ii) com um novo padrão de entrada para treinamento e repetir os procedimentos até que todos os padrões sejam classificados corretamente, ou quando outro critério de parada seja satisfeito.

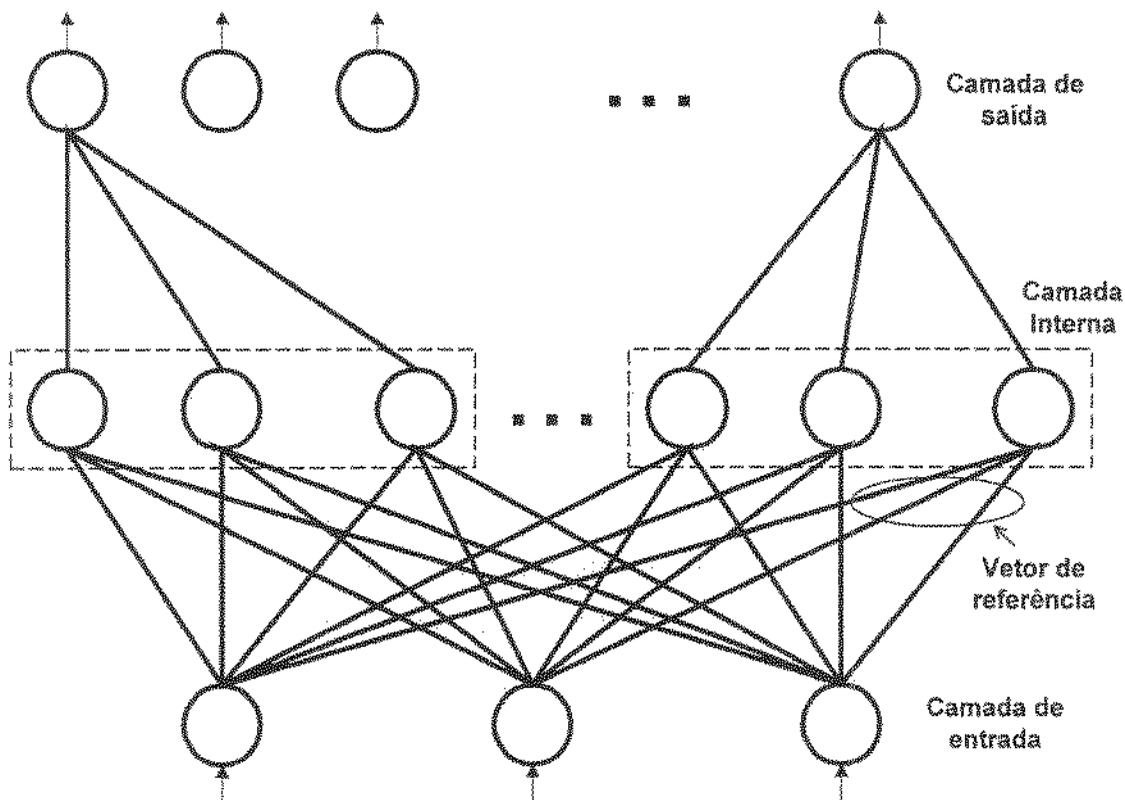


Figura 3 – Exemplo de uma rede neuronal LVQ – Learning Vector Quantization.
(BULSARI, 1995)

O cálculo da distância Euclidiana, d_i , na camada de saída entre cada vetor de treinamento, x_j , e o vetor dos pesos, w_{ij} , é feito pela equação

$$d_i = \left\{ \sum_{j=1}^N (w_{ij} - x_j)^2 \right\}^{1/2} \quad (19)$$

O neurônio que apresentar o menor valor de d_i será considerado como “vencedor”, e os pesos deste serão ajustados segundo a heurística abaixo, usando um fator de atração, α , e uma taxa de repulsão, γ .

$$w' = \begin{cases} w + \alpha(x - w) & \text{se o neurônio vencedor está na classe correta} \\ w - \gamma(x - w) & \text{se o neurônio vencedor não está na classe correta} \end{cases} \quad (20)$$

2.1.5. Treinamento e programação

As técnicas convencionais de programação exigem que alguém crie um algoritmo. Existem casos onde esta tarefa se torna muito complexa, como por exemplo o

reconhecimento de uma face humana. Isto pela diversidade de variações possíveis que tem de ser levadas em consideração, tais como se a pessoa está sorrindo ou séria.

As redes neuronais, ao contrário de serem programadas, são treinadas. Isto significa que exemplos lhes são apresentados, e elas se ajustam a estes através de um método de aprendizado. Com certeza, há necessariamente que haver um número representativo de exemplos para se alimentar à rede.

As redes neuronais de aprendizagem supervisionada possuem uma grande capacidade para inferir sobre as relações entre as variáveis de entrada e saída, mesmo quando o número de neurônios é pequeno. Isto facilita para que os cálculos possam ser efetuados a baixo custo até em microcomputadores de pequeno porte. Mesmo sendo simples, a rede “feedforward” mostra ser uma ferramenta muito interessante para a modelagem de processos estacionários não-lineares. Uma característica importante das redes é a possibilidade de se incorporar variáveis subjetivas, assim como quente, frio, morno, bom, ruim, rejeitado e aprovado.

Muitos trabalhos enfocam a necessidade do desenvolvimento de técnicas para melhorar a capacidade preditiva da rede neuronal, já que se constata que na etapa de treinamento podem surgir respostas que não correspondem ao treinamento ótimo. Uma das maneiras de se melhorar a convergência para uma condição de mínimo erro global é a injeção de ruído no treinamento a fim de gerar perturbações nos pesos (HAYAKAWA et al., 1995; GRANDVALET e CANU, 1995; BISHOP, 1995; YAMAZAKI e OGAWA, 1993; ALBEVERIO et al., 1995).

Para modelagem de processos em regime transiente tem sido utilizada a rede do tipo recorrente com relativo sucesso, apesar dos poucos trabalhos encontrados que utilizam esta estrutura (PINEDA, 1987; PEARLMUTTER, 1989; PEARLMUTTER, 1990). Este tipo de rede exige um tempo computacional maior na fase de treinamento do que aquele para as estruturas “feedforward”, mas consegue assimilar melhor a dinâmica dos processos em regime transiente.

A discussão e o equacionamento de diversos tipos de estruturas para redes neuronais podem ser encontrados em HAYKIN (1994), BAUGHMAN e KOVÁCS (1996).

2.1.6. Redes neuronais e outras tecnologias

O campo da inteligência artificial reúne elementos tais como as redes neuronais, sistemas especialistas, algoritmos genéticos, lógica “fuzzy” e outros. Contudo, eles podem ser significativamente diferentes entre si, como o são para a programação tradicional.

Os sistemas especialistas são registrados como sendo a tecnologia mais antiga de inteligência artificial. Eles diferem da programação convencional por separar a base de um conhecimento da forma como ele é processado. Isto permite que conhecimentos adicionais possam ser incluídos ao sistema sem necessidade de reprogramação.

Esta técnica exige que o conhecimento numa determinada área esteja desenvolvido e seja disponível, de forma que se possam estabelecer algumas regras. Um exemplo de regra com um fator de confiança seria: se existe um brinco preso na orelha, a pessoa é do sexo feminino - confiança 70%.

Quando se desenvolve uma rede neuronal como uma solução para um problema nem o conhecimento específico, nem as regras explícitas de processamento deste são levadas em conta pelo programador. Ao invés disso, a rede procura aprender intrinsecamente estas regras para processar o conhecimento.

Pode se visualizar as redes neuronais como um modelo extremamente generalizado que é parametrizado por pesos ajustáveis. Assim, não é necessário ter acesso ao conhecimento profundo do problema para desenvolver uma solução deste. Em contrapartida, nos sistemas especialistas e na programação convencional se conhece como se chegou a uma determinada resposta, o que não é possível com uma rede neuronal. Por este motivo, pode acontecer que a solução do problema por redes neuronais não seja a mais indicada.

O uso da estatística prevê a formulação de algumas hipóteses sobre os dados e deve-se, muitas vezes, limitar a análise em um certo número de interações possíveis. Em contraste, sob o ponto de vista prático, as redes neuronais são basicamente não-paramétricas, mesmo que se pense nelas como sendo parametrizadas pelos seus pesos. Incluindo cada vez mais dados para serem analisados, mais as interações complexas entre as entradas são possíveis de serem consideradas. De maneira geral, a estatística pode ser

útil para o conhecimento dos dados, o que pode levar a desenvolver um modelo de rede neuronal melhor.

2.1.7. Aplicações para redes neuronais.

Redes neuronais têm sido aplicadas a centenas de áreas tais como:

- ⊗ Bancária – com a leitura de cheques e documentos, avaliação de investimentos e sistemas de identificação,
- ⊗ Telecomunicações – com a análise e processamento de sinais, compressão de dados e imagens, eliminação de ruídos,
- ⊗ Produção - controle de processos produtivos, classificação e otimização de processos, análise e projeto de produtos, processos e máquinas,
- ⊗ Robótica – sistemas de visão, controle de trajetórias, automatização, reconhecimento de voz
- ⊗ Informática - reconhecimento ótico de caracteres, reconhecimento de comandos vocais,
- ⊗ Médica – Análise e identificação de células cancerosas, auxílio em exames complexos, projeto de próteses, especialista médico para diagnóstico de problemas cardíacos, entre outros exemplos.

2.1.8. Aplicações de redes neuronais para a área de engenharia química

O número de trabalhos publicados relacionados com redes neuronais cresce rapidamente, como se pode observar em uma simples busca bibliográfica. Uma revisão compreensiva de tudo que surge relacionado nesta área seria muito extensa para ser apresentada neste trabalho. Foram selecionadas algumas para serem mencionadas aqui, cujo conteúdo auxiliaram no processo de obtenção das redes treinadas deste trabalho.

- ⊗ Em 1990, BHAT e McAVOY utilizaram uma rede “feedforward” para propor um algoritmo de controle do pH em um reator CSTR.

- ⊗ No mesmo ano, YDSTIE (1990) publica uma análise da aplicação de redes neurais em diversas técnicas de controle de processos.
- ⊗ Utilizando uma rede neuronal, THIBAUT et al. (1990) conseguiram prever variáveis de um processo de fermentação com vistas ao controle do processo.
- ⊗ Em 1992, BHAT e McAVOY apresentaram um algoritmo para determinar uma estrutura de rede neuronal mínima pela eliminação de neurônios da rede.
- ⊗ Para modelar a produção de penicilina de uma unidade industrial MASSIMO et al., (1992) se utilizaram de uma rede “feedforward”.
- ⊗ HUNT et al. (1992) publicam uma revisão sobre estruturas, algoritmos e possibilidades de aplicação de redes neurais na modelagem e controle de processos químicos.
- ⊗ Ainda em 1992, WILLIS et al. analisaram o potencial de uso de redes na estimação de parâmetros e em controle de processos.
- ⊗ HOSKINS e HIMMELBLAU (1992) utilizaram uma rede neuronal para controlar um CSTR, comparando o desempenho com um controlador PID existente.
- ⊗ Em 1993, WALT et al. utilizaram uma rede “feedforward” para prever a viscosidade cinemática de óleo cru e suas frações, conseguindo uma precisão de 98,74 % na previsão para dados fora do conjunto de treinamento.
- ⊗ NASCIMENTO e YAMAMOTO (1993) apresentaram resultados de aplicação de redes “feedforward” na modelagem da síntese de amônia, polimerização de olefinas, de reator fotoquímico e uma comparação de modelo estocástico com modelo de rede neuronal.
- ⊗ No mesmo ano ainda, TANABE et al. (1993) modelaram a distribuição de potência em reator com banco de geradores de vapor, com auxílio de redes neurais.

- ⊗ Para diagnosticar falhas numa planta com várias colunas de destilação BECRAFT e LEE (1993) integraram uma rede neuronal e um sistema especialista.
- ⊗ VEAUX et al. (1993) comparam o desempenho de uma rede “feedforward” com uma regressão multilinear adaptativa.
- ⊗ Em 1994, CHENG e TITTERINGTON apresentam uma revisão sobre redes neuronais sobre a ótica da estatística.
- ⊗ MORRIS et al. (1994) revisaram as principais estruturas de redes neuronais que podem ser aplicadas na modelagem e controle de processos.
- ⊗ Em 1995, BARATTI et al. desenvolveram uma rede para inferir o desempenho de duas colunas de destilação a partir de reduzido número de variáveis de estado.
- ⊗ Utilizando uma rede com recorrência externa para modelar uma coluna de destilação recheada, MAC MURRAY e HIMMELBLAU (1995) mostraram melhores resultados do que utilizando modelos lineares utilizados em controle preditivo de processos.
- ⊗ Para modelar e simular o processo de refino de cavacos de madeira QIAN e TESSIER (1995) utilizaram-se de uma rede neuronal.
- ⊗ TETKO et al. (1995) publicaram um trabalho analisando o problema de “overfitting” e “overtraining” comparando o desempenho de uma rede “feedforward” com uma regressão multilinear. Contrariando o trabalho de VEAUX et al (1993), neste trabalho a rede ofereceu vantagens.
- ⊗ Em 1996, HIMMELBAU e KARJALA utilizaram uma rede recorrente para retificação de valores experimentais medidos num tanque com descarga.

2.1.9. Publicações sobre a aplicação de redes neurais a processos de polpação e tratamento de efluentes.

Devido aos trabalhos desenvolvidos no Laboratório de Modelagem e Simulação de Processos Químicos (LMSPQ) do Departamento de Processos Químicos, algumas publicações foram possíveis. Sobre a utilização de redes neurais para predição da DBO e DQO de saída das lagoas de tratamento de efluentes da RIPASA S/A, tem-se um trabalho realizado em conjunto com alunos de iniciação científica do laboratório (GONTARSKI et al., 98c).

Relativas a polpação da madeira no processo de fabricação de papel e celulose, trabalho realizado também em parceria com a RIPASA S/A, surgiram duas publicações em congresso. Uma em congresso nacional, o COBEQ (GONTARSKI et al., 1998a), e outra em um congresso internacional (GONTARSKI et al., 1998b).

Em parceria com a Rhodiaco surgiram diversas publicações. Em congressos internacionais tem-se (GONTARSKI et al., 1999a), (GONTARSKI et al., 1999b), (GONTARSKI et al., 2000a) e (GONTARSKI et al., 2000b). O primeiro da série salienta os primeiros resultados, com o conjunto inicial de dados. No segundo, aprimora-se os resultados para um novo conjunto de dados. Os dois últimos tratam de análises realizadas com as redes treinadas, mostrando a influência das principais variáveis no processo. Este último trabalho em congresso ainda foi aceito para publicação em revista indexada, a *Computers and Chemical Engineering* (GONTARSKI et al., 2000c).

Além dessas publicações, participou-se do EPFEQ, evento de pesquisa da Faculdade de Engenharia Química da Unicamp, com resumo publicado (GONTARSKI et al., 1999c), e mais recentemente de congresso nacional, (GONTARSKI et al., 2000d).

2.2. O PROCESSO DE FABRICAÇÃO DA CELULOSE E DO PAPEL.

Um dos processos onde se aplicam redes neurais neste trabalho é um dos mais importantes na indústria de papel e celulose. Trata-se da polpação da madeira. Contudo, como muitas variáveis analisadas no trabalho tem importância para o restante do processo, uma revisão mais abrangente sobre o mesmo se faz necessária. Após a apresentação de um fluxograma simplificado do processo segue uma descrição sobre as principais fases envolvidas.

- Linhas Gerais da fabricação de celulose e papel

Sendo a madeira uma fonte abundante de celulose, constitui a principal matéria prima utilizada na manufatura de polpa, que é o material celulósico fibroso empregado na produção do papel e de outros produtos derivados da celulose.

A confecção de papel pode ser dividida em três fases: a polpação da madeira, a descoloração da polpa e a obtenção do produto final, o papel. A matéria prima é, em geral, reduzida a fibras que são consecutivamente refinadas, com lavagens sucessivas. A fabricação do papel pode acontecer de forma integrada com a polpação de forma que a polpa recebe enchimento, acabamento e o produto final é transformado em folhas de papel. Os enchimentos geralmente utilizados são a argila, o talco e o gesso.

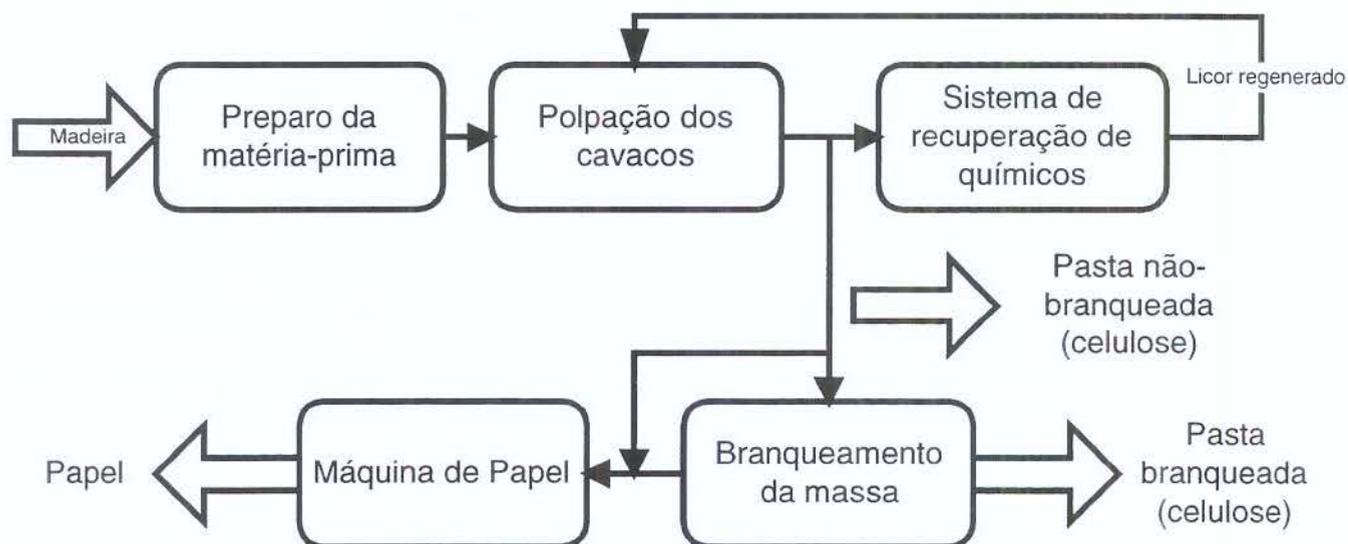


Figura 4 - Fluxograma simplificado do processo de fabricação do papel e celulose

- Preparo da madeira e o processo de Polpação (cozimento)

Em locais de clima quente e úmido, a madeira é entregue à fábrica o mais rápido possível após ter sido cortada, uma vez que o armazenamento nestas condições ocasiona perda de fibras e manchas devido à presença de fungos. Já em climas frios ela pode ser armazenada por diversos anos sem perda de sua qualidade. Na fase de preparação da madeira, os seguintes procedimentos são feitos:

☒ corte,

- ⊗ armazenamento,
- ⊗ descasque e lavagem dos troncos,
- ⊗ corte em cavacos para as condições do cozimento,
- ⊗ peneiramento dos cavacos, seleção e armazenamento dos mesmos e
- ⊗ transporte para os digestores.

A polpação é o processo pelo qual a madeira é transformada em uma massa fibrosa. Para que isto aconteça, as fortes ligações existentes no interior da madeira devem ser rompidas, existindo várias maneiras para tal, por meios mecânicos, térmicos e químicos.

O processo Kraft é caracterizado pelo cozimento dos cavacos em um solução formada por hidróxido de sódio (NaOH) e sulfeto de sódio (Na₂S), chamada licor de cozimento. O álcali ataca a molécula de lignina quebrando-a em pequenos fragmentos na forma de sais de sódio. As polpas Kraft geralmente produzem papéis resistentes, contudo, são mais difíceis de branquear do que polpas produzidas segundo outros processos. Algumas vantagens: grande flexibilidade com relação às várias espécies de madeira, polpas obtidas com alta resistência, recuperação eficiente dos químicos do cozimento, entre outras.

Nomenclatura empregada no processo KRAFT:

LICOR BRANCO: composto pelos químicos do cozimento (NaOH e Na₂S);

LICOR PRETO: contém os produtos da reação de deslignificação. Este licor é concentrado e queimado no forno de recuperação produzindo um fundido inorgânico de carbonato de sódio (Na₂CO₃) e sulfeto de sódio (Na₂S);

LICOR VERDE: originado pela dissolução do fundido do licor preto, que reage com cal virgem (CaO) convertendo o carbonato de sódio em hidróxido de sódio, e assim regenerando o licor branco.

No processo de cozimento em batelada, o aquecimento é feito de acordo com um programa pré determinado, no qual a temperatura é elevada, gradualmente, durante 50 a 90 minutos até atingir um valor, geralmente próximo de 170°C, e é mantida durante um certo período de tempo.

Após a reação, a mistura segue normalmente para estágios de peneiramento, com a separação dos cavacos mal cozidos. Posteriormente têm-se estágios de lavagem, onde o licor e os orgânicos solubilizados são separados das fibras.

Os digestores descontínuos permitem maior flexibilidade operacional, menores quedas na produção e tempos menores para manutenção em relação aos de operação contínua.

Segundo a literatura, o ataque químico à lignina durante o processo se deve à ação conjunta dos íons hidroxila (OH⁻) e hidro-sulfeto (SH⁻), presentes no licor de cozimento. Carbohidratos, especialmente hemiceluloses (50%, em média) e, em menor extensão, celulose (10%, em média), são também atacados e removidos dos cavacos. As principais reações envolvendo os carbohidratos estão dispostas a seguir:

- ⊗ Inchamento alcalino: provoca alterações físicas na parede da fibra, facilitando a penetração do licor de cozimento e a difusão dos reagentes, influenciando na uniformidade do cozimento. Esta reação ocorre em um pH aproximadamente igual a 13;
- ⊗ Solubilização alcalina: ocorre sempre quando alguma nova região da madeira se torna acessível ou na presença de cadeias moleculares suficientemente pequenas, que possibilitem a solubilização;
- ⊗ Hidrólise alcalina dos grupos acetila da hemicelulose: geralmente ocorre no início do cozimento provocando a formação de íons acetato que, em meio aquoso, forma o ácido acético, consumindo assim uma parte do licor de cozimento;
- ⊗ Despolimerização terminal (reação de desfolhamento): provoca o consumo da maior parte do álcali do licor de cozimento (60 a 65% do álcali efetivo) explicando o baixo rendimento dos processos alcalinos (45 a 50% de polpa seca em relação a massa de madeira seca). Este tipo de reação ocorre em ciclos, onde a unidade terminal redutora é atacada e retirada dos polissacarídeos, aparecendo um novo grupo terminal semelhante ao original, iniciando deste modo um novo ciclo;
- ⊗ Hidrólise alcalina das ligações β -glucosídicas: promove o fracionamento dos polissacarídeos, expondo uma nova unidade terminal redutora, sensível a

reações de desfolhamento. Ocorre principalmente a temperatura máxima de cozimento (170°C);

- ⊗ Reprecipitação e adsorção: ocorre com as moléculas de carboidratos em solução que, com a queda do pH, sofrem reprecipitação, sendo adsorvidas pelas regiões cristalinas das microfibrilas da celulose.

Os polissacarídeos retirados da madeira pelo processo Kraft não continuam no licor em sua forma original, sendo transformados em componentes ácidos (geralmente ácidos orgânicos), elevando o consumo de álcali.

Analisando as principais variáveis que influenciam o processo de cozimento, podemos classificá-las em duas categorias:

- ⊗ variáveis associadas à madeira: espécie, densidade, composição química, fatores de crescimento, estocagem e dimensões do cavaco;
- ⊗ variáveis associadas à operação de cozimento: relação entre as massa dos reagentes e a massa de madeira seca, concentração dos reagentes no licor de cozimento, sulfidez do licor, temperatura e tempo de cozimento.

A espécie de madeira empregada no cozimento é responsável por diferenças apreciáveis na qualidade e nas propriedades da polpa obtida. O comprimento das fibras diferencia as espécies folhosas (fibras curtas, com comprimento entre 1.0mm e 1.5mm) das coníferas (fibras longas, com comprimento entre 3.0mm e 5.0mm). Além das composições químicas diferentes entre elas, as folhosas deslignificam mais rapidamente que as coníferas, por apresentarem um menor teor de lignina, ocorrência preferencial da lignina na lamela média, menor tendência à reprecipitação e condensação e uma estrutura celular mais favorável à penetração do licor (através dos elementos de vasos).

Devido às características anatômicas das madeiras, a celulose derivada das coníferas apresenta resistência mecânica maior do que a similar das folhosas, viabilizando assim seu emprego em papéis de resistência muito elevada (por exemplo, sacos de cimento).

O ponto de desfibramento do material é atingido quando a dissolução de substâncias da madeira for suficiente para permitir a separação das fibras por diferencial de pressão.

Os reagentes do licor de cozimento penetram nos cavacos através dos espaços vazios e pela difusão dos íons na água já presente nas fibras. A penetração é importante no estágio inicial do cozimento para a retirada do ar ocluso dos cavacos, durante o período de aquecimento. Como o álcali é consumido rapidamente nas reações com carboidratos no estágio inicial, o álcali introduzido pela penetração é rapidamente gasto, ocorrendo necessidade de se introduzir quantidades adicionais de reagente para promover a deslignificação do cavaco já impregnado pelo licor. Logo, os reagentes empregados nas reações de deslignificação propriamente ditas são transportadas basicamente por difusão. Os produtos resultantes deixam os cavacos também por difusão no sentido oposto ao da entrada dos reagentes. A concentração dos íons hidroxila diminui no sentido do seu movimento para o interior dos cavacos, por serem neutralizados na reação com os componentes da madeira. Quando a deslignificação não for suficiente para atingir o ponto de desfibramento, haverá aumento da quantidade de rejeitos. Pode-se dizer que a espessura do cavaco e a temperatura do processo controlam a uniformidade da reação, pois, com um acréscimo de temperatura, o aumento ocorrido na difusão é muito menor do que o sofrido pela taxa de reação, ocasionando uma falta de uniformidade da reação dentro do cavaco; o mesmo se aplica quando a espessura do cavaco está acima do ideal.

A quantidade de álcali ativo necessário para se atingir o ponto de desfibramento depende da espécie da madeira, dimensões dos cavacos, tempo e temperatura do cozimento e da sulfidez.

A concentração de álcali empregada deve manter em solução os compostos dissolvidos da madeira, pois, ocorrendo queda de pH, fragmentos de lignina e hemicelulose poderão ser reprecipitados sobre as fibras, elevando o número Kappa, exigindo maiores quantidades de químicos em estágios de tratamento posteriores.

2.3. TRATAMENTO DE EFLUENTES LÍQUIDOS

Em muitos processos químicos são geradas correntes líquidas contendo substâncias indesejáveis no processo e que precisam ser descartadas. Normalmente o que se faz é retornar essas correntes ao rio ou até mesmo ao oceano, que são os chamados corpos receptores do despejo. O lançamento de compostos orgânicos em corpos receptores pode causar conseqüências muito graves, comprometendo o equilíbrio ecológico. Em condições normais, um rio é capaz de receber uma carga apreciável, eliminando-a gradativamente

mediante ações naturais que se processam ao longo de vários quilômetros do seu percurso, entretanto quando o curso de água fica sem oxigênio dissolvido suficiente ocorre a morte dos organismos aeróbios e a aplicação desta água fica praticamente impossibilitada para diversos usos e finalidades.

No território brasileiro as águas doces, salobras e salinas são protegidas pela legislação brasileira, de acordo com a portaria 36/GM de 18 de junho de 1986 (CONSELHO NACIONAL DO MEIO AMBIENTE, 1990). No âmbito do estado de São Paulo se estabelecem padrões para os efluentes de qualquer fonte poluidora quanto a certas características tais como pH, temperatura, materiais sedimentáveis, DBO₅ e concentrações de substâncias tóxicas.

Com isto, obrigam-se as indústrias e quaisquer geradores de águas residuárias a possuírem sistemas de tratamento de efluentes líquidos, de forma a atender às exigências legais. O método de tratamento está relacionado intimamente com as características do efluente. METCALF & EDDY (1991) denominam como operações unitárias aqueles métodos de tratamento onde predomina a aplicação de forças físicas, e como processos unitários aqueles métodos que removem os contaminantes através de reações químicas ou biológicas. Os principais contaminantes em tratamento de águas residuárias, segundo METCALF & EDDY (1991), estão apresentados na Tabela 1 e as principais características físicas, químicas e biológicas das águas residuárias, e respectivas fontes estão na Tabela 2.

Nas duas unidades industriais onde foram realizados os estudos apresentados neste trabalho as características do efluente dependem muito da sua origem no processo.

Na Rhodiaco, indústria que fabrica em larga escala um produto químico utilizado na indústria de polímeros, as águas residuárias vindas do processo se misturam com o esgoto doméstico. Na Ripasa SA, praticamente todo o resíduo proveniente da fabricação do papel e celulose é gerado nos seguintes processos:

- máquina de papel;
- branqueamento da celulose, e
- fábrica de recuperação (YANG, 1996).

Tabela I - Contaminantes importantes no tratamento de águas residuárias.

Contaminate	Motivo da importância
Sólidos suspensos	Sólidos suspensos podem ser a causa de depósitos de lodo e de desenvolvimento de condições anaeróbias quando o efluente líquido não tratado é lançado no ambiente aquático.
Orgânicos biodegradáveis	Compostos principalmente por proteínas, carboidratos e gorduras, os orgânicos biodegradáveis são medidos frequentemente em termos de DBO (demanda bioquímica de oxigênio) e DQO (demanda química de oxigênio). Se lançado sem tratamento ao ambiente, sua estabilização biológica pode levar à exaustão da reserva de oxigênio natural presente e ao desenvolvimento de condições sépticas.
Patogênicos	Algumas doenças podem ser transmitidas por organismos patogênicos em águas residuárias.
Nutrientes	Tanto nitrogênio quanto fósforo, juntos ao carbono, são nutrientes essenciais para o crescimento. Quando lançados no ambiente aquático, estes nutrientes podem levar ao crescimento de colônias de seres aquáticos indesejáveis. Quando lançados em excessivas quantidades sobre a Terra, também podem poluir águas subterrâneas.
Poluentes perigosos	Compostos orgânicos e inorgânicos selecionados com base no conhecimento de apresentarem carcinogenicidade, mutagenicidade, teratogenicidade ou toxicidade. Muitos destes compostos são encontrados em águas residuárias.
Orgânicos refratários	Estes orgânicos tendem a resistir a métodos convencionais de tratamento de efluentes líquidos. Exemplos típicos incluem surfactantes, fenóis e pesticidas agrícolas.
Metais pesados	Metais pesados em águas residuárias são provenientes geralmente de atividades comercial e industrial e devem ser removidos quando o efluente for reutilizado.
Inorgânicos dissolvidos	Constituintes inorgânicos como cálcio, sódio e sulfato são adicionados à água de abastecimento doméstico e devem ser removidos se o efluente for reutilizado.

Tabela 2 - Características físicas, químicas e biológicas das águas residuárias e suas fontes.

CARACTERÍSTICAS	FONTE
PROPRIEDADES FÍSICAS	
• Cor	Resíduos domésticos e industriais, degradação natural de materiais orgânicos
• Odor	Águas residuárias em decomposição e resíduos industriais
• Sólidos	Abastecimento de água potável, resíduos domésticos e industriais, erosão de solos, infiltração
• Temperatura	Resíduos domésticos e industriais
CONSTITUINTES QUÍMICOS	
• Orgânicos	
Carboidratos	Resíduos domésticos, comerciais e industriais
Gorduras, óleos e graxas	Resíduos domésticos, comerciais e industriais
Pesticidas	Resíduos agrícolas
Fenóis	Resíduos industriais
Proteínas	Resíduos domésticos, comerciais e industriais
Poluente perigosos	Resíduos domésticos, comerciais e industriais
Surfactantes	Resíduos domésticos, comerciais e industriais
Compostos orgânicos voláteis	Resíduos domésticos, comerciais e industriais
Outros	Degradação natural de matérias orgânicos
• Inorgânicos	
Alcalinidade	Resíduos domésticos, abastecimento de água potável, infiltração de água subterrânea
Cloretos	Resíduos domésticos, abastecimento de água potável, infiltração de água subterrânea
Metais pesados	Resíduos industriais
Nitrogênio	Resíduos domésticos e agrícolas
PH	Resíduos domésticos, comerciais e industriais
Fósforo	Resíduos domésticos, comerciais e industriais, drenagem natural de água
Poluente perigosos e tóxicos	Resíduos domésticos, comerciais e industriais
Enxofre	Abastecimento de água potável, resíduos domésticos, comerciais e industriais
• Gases	
Gás sulfídrico	Decomposição de resíduos domésticos
Metano	Decomposição de resíduos domésticos
Oxigênio	Abastecimento de água potável, infiltração com a superfície da água
CONSTITUINTES BIOLÓGICOS	
• Animais	Cursos d'água abertos e plantas de tratamento de efluentes
• Plantas	Cursos d'água abertos e plantas de tratamento de efluentes
• Protistas	Resíduos domésticos, infiltração com a superfície da água e plantas de tratamento de efluentes
• Vírus	Resíduos domésticos

Fonte: METCALF & EDDY (1991)

Este resíduo é composto por produtos químicos usados durante o processo como compostos organo-clorados, peróxidos e hipocloritos, sendo portanto bastante ácido (pH aproximadamente 3) e é produzido em grande quantidade.

Para preservar a qualidade dos cursos de água naturais, a Legislação Federal Brasileira divide-os em classes de acordo com suas características. As legislações estaduais

devem impor restrições aos parâmetros dos efluentes lançados a estes cursos de água de forma a preservar sua classificação (OLIVEIRA, 2000).

Quando o efluente não se enquadra nas condições permitidas, ele deverá ser tratado, e será necessária a implantação de um sistema de tratamento adequado às características desse efluente. Os efluentes inorgânicos são tratados preferencialmente por processos físico-químicos. Os orgânicos podem receber tratamento por processo físico-químico ou biológico. Os processos biológicos são adequados quando o efluente orgânico apresenta características biodegradáveis. Basicamente, no chamado tratamento primário são empregadas operações unitárias como peneiramento, sedimentação e filtração para a remoção de sólidos sobrenadantes ou sedimentáveis normalmente encontrados em efluentes líquidos. No tratamento secundário são usados processos químicos ou biológicos para eliminar grande parte da carga de materiais orgânicos. No tratamento terciário usam-se combinações de operações e processos unitários para remover outros constituintes que não foram removidos no tratamento secundário, tais como nitrogênio e fósforo (METCALF & EDDY, 1991).

Os processos biológicos que ocorrem na presença de oxigênio livre, denominados aeróbios, são bem conhecidos e os mais utilizados. A implantação de sistemas com processos aeróbios oferece, portanto, poucos riscos de investimento, pois têm custo baixo de construção e apresentam operação e manutenção simples e econômicas. No entanto, são sistemas que ocupam áreas consideráveis, os equipamentos necessários como os aeradores demandam energia e o lodo, gerado como subproduto, ainda é passível de decomposição por ação biológica. O objetivo do tratamento biológico de efluentes industriais ou domésticos é coagular e remover sólidos coloidais que não sedimentam e ainda estabilizar a matéria orgânica. Esta redução da concentração da matéria orgânica é usualmente medida como DBO (demanda bioquímica de oxigênio), TOC (carbono orgânico total), ou DQO (demanda química de oxigênio) (METCALF & EDDY, 1991).

Outros processos de decomposição biológica ocorrem na ausência de oxigênio livre e se denominam como processos anaeróbios. Quando o tratamento é realizado em lagoas anaeróbias pode-se ter problemas estéticos devido a formação de espuma na superfície da água; já quando realizado em bio-reatores fechados, o processo pode apresentar as vantagens de produção do gás metano como subproduto que é utilizado como

fonte de energia, menor área de ocupação, e ainda produção de lodo em menor quantidade e já estabilizado (OLIVEIRA, 2000).

Neste trabalho foram estudados reatores aeróbios, sendo que em um dos casos se trata de lagoas de grande volume, e no outro tem-se um conjunto de reatores de aeração. No caso das lagoas não há retorno do lodo, enquanto que no segundo caso tem-se o retorno, configurando o tratamento com lodo ativado. METCALF & EDDY (1991) descrevem este tipo de sistema, como ilustrado na Figura 5. O material orgânico é introduzido no reator onde uma cultura bacteriana aeróbia é mantida em suspensão. No reator, as bactérias se encarregam de converter grande parte dos orgânicos em basicamente água, gás carbônico e moléculas orgânicas de menor cadeia. A condição aeróbia é alcançada pela aeração através de difusores ou outros sistemas mecânicos, os quais também ajudam a manter uma mistura mais eficaz dentro do reator.

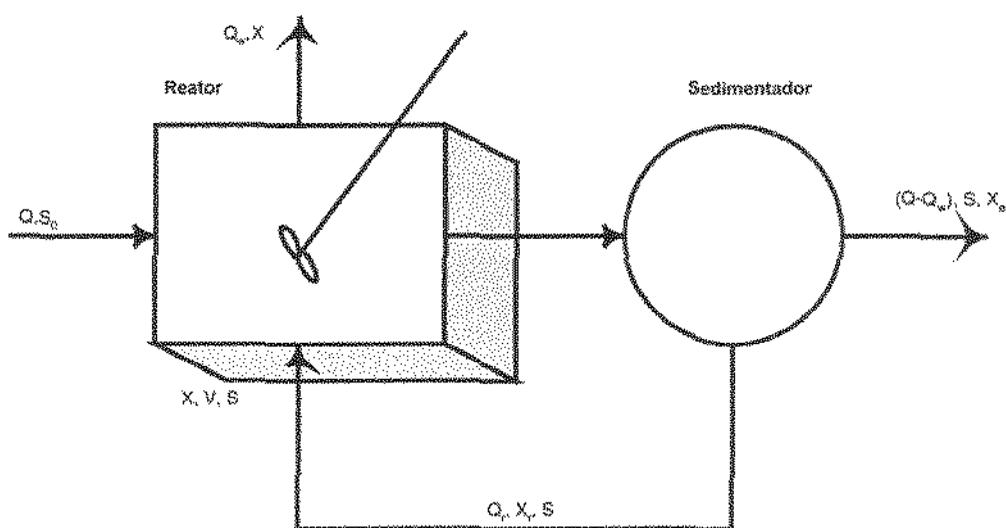


Figura 5 – Sistema Reator (CSTR) e Sedimentador com reciclo de material celular (METCALF & EDDY, 1991).

Depois de um certo período de tempo, a mistura contendo células novas e velhas segue a um sedimentador, onde as células são isoladas do efluente tratado. Uma parte deste material celular é reciclado para o reator com a finalidade de manter a concentração desejada de organismos no reator, e a outra parte é descartada. O nível em que a massa biológica deve ser mantida no reator depende da eficiência desejada do tratamento e de outras considerações relacionadas à cinética de crescimento microbiano. Na natureza, a função chave das bactérias no ecossistema é a de decompor a matéria orgânica produzida

por outros organismos vivos. No processo com lodo ativado, as bactérias são os microorganismos mais importantes porque eles são responsáveis pela decomposição do material orgânico na corrente influente. No reator, uma parte do despejo orgânico é usada pelas bactérias aeróbias e facultativas para obter energia para a síntese do material orgânico restante em novas células. A Figura 6 ilustra exatamente isto. Somente uma pequena parte do despejo original é realmente oxidada a compostos de baixa energia, tais como NO_3 , SO_4 e CO_2 , a grande porção é mesmo sintetizada como material celular.

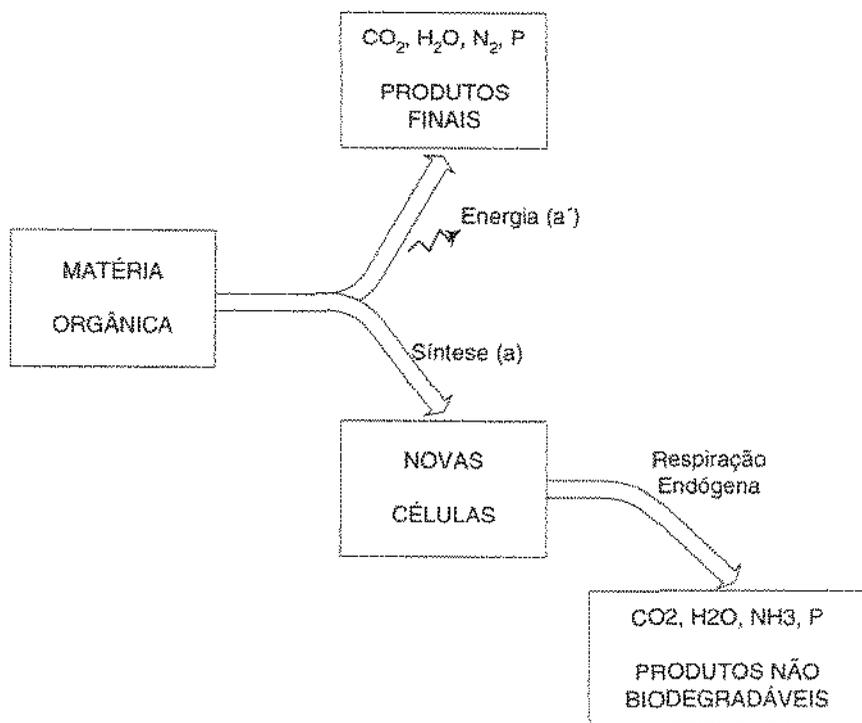


Figura 6 - Degradação da matéria orgânica (PESSSOA & JORDÃO, 1982)

Assim como é importante que as bactérias decomponham o despejo orgânico o mais rápido possível, também é essencial que elas formem um floco satisfatório, o qual é um pré-requisito para uma separação efetiva na unidade de sedimentação do material biológico. Quando o tempo de residência médio do lodo no sistema aumenta se observa que as suas características de decantação melhoram bastante. Segundo METCALF & EDDY (1991), esse tempo de residência do lodo ideal pode ficar entre 5 a 15 dias para o sistema da Figura 5, mas no caso de tratar exclusivamente esgoto doméstico, este tempo fica entre 3 a 4 dias. No caso das lagoas, como as da Ripasa SA, os tempos de residência variam de 2 a

10 dias, não sendo comum o desenvolvimento de algas. Em termos de redução de DBO, a eficiência das lagoas aeradas pode ser superior a 95%.

Algumas relações importantes definem alguns dos parâmetros essenciais ao processo, e certas considerações são necessárias para estas serem obtidas. A primeira destas é que não se considera nenhuma presença de lodo na corrente influente do despejo industrial. Outra é que a estabilização do efluente pelos microorganismos ocorre somente dentro do reator, e ainda considera-se que o volume usado no cálculo do tempo médio de residência do lodo para o sistema inclui somente o volume dos reatores. Desta forma, o tempo médio de residência hidráulico para o sistema, θ_s é definido como

$$\theta_s = \frac{V_s}{Q} \quad (21)$$

onde V_s = volume do reator acrescido do volume do sedimentador,

Q = vazão do influente.

O tempo de residência para o reator será apenas

$$\theta = \frac{V}{Q} \quad (22)$$

onde V = volume do reator.

Para o sistema da Figura 5, o tempo de residência médio para o lodo, θ_c , torna-se então:

$$\theta_c = \frac{VX}{Q_w X + (Q - Q_w) X_e} \quad (23)$$

onde Q_w = vazão da corrente que descarta o lodo em excesso que não é reciclado,

X = concentração de microorganismos no reator, de onde é descartada a quantidade Q_w ,

X_e = concentração de microorganismos no efluente da unidade de sedimentação.

Em um sistema com uma unidade de sedimentação operando em condições apropriadas, a quantidade de material celular no efluente seria muito pequena, e a equação 23 poderia ser simplificada como

$$\theta_c \equiv \frac{V}{Q_w} \quad (24)$$

Um balanço de massa para os microorganismos no sistema pode ser escrito como:

$$\left[\begin{array}{c} \text{Taxa de} \\ \text{acúmulo de} \\ \text{microorganismos} \\ \text{dentro do sistema} \end{array} \right] = \left[\begin{array}{c} \text{Vazão de} \\ \text{entrada de} \\ \text{microorganismos} \\ \text{no sistema} \end{array} \right] - \left[\begin{array}{c} \text{Vazão de} \\ \text{saída de} \\ \text{microorganismos} \\ \text{do sistema} \end{array} \right] + \left[\begin{array}{c} \text{Taxa de} \\ \text{crescimento de} \\ \text{microorganismos} \\ \text{dentro do sistema} \end{array} \right] \quad (25)$$

ou ainda

$$\frac{dX}{dt} V = QX_0 - [Q_w X + (Q - Q_w) X_c] + V(r'_g) \quad (26)$$

onde X_0 = concentração de microorganismos na corrente de entrada da unidade,

(r'_g) = taxa líquida de crescimento dos microorganismos.

A expressão da taxa líquida de crescimento pode ser substituída, e ainda admitindo que a concentração de microorganismos na corrente de entrada da unidade é zero e ainda o estado estacionário, chega-se a seguinte equação:

$$\frac{[Q_w X + (Q - Q_w) X_c]}{VX} = -Y \frac{(r_{sw})}{X} - k_d \quad (27)$$

onde Y = coeficiente de máxima produção, definido como a razão entre a massa de células formadas e a massa de substrato consumido,

k_d = coeficiente de decaimento endógeno, em unidades de tempo⁻¹,

(r_{sw}) = taxa de utilização do substrato, que pode ser escrita como

$$(r_{sw}) = -\frac{kXS}{K_S + S} = \frac{S_0 - S}{\theta} \quad (28)$$

onde k = taxa máxima de utilização do substrato por unidade de massa de substrato,

K_S = constante de meia velocidade, é a concentração do substrato quando se atinge a metade da taxa de crescimento máxima,

S_0 = concentração inicial do substrato limitante do crescimento no líquido,

S = concentração do substrato limitante do crescimento no líquido após um tempo θ .

De forma a propiciar um dimensionamento adequado da unidade de tratamento, se faz necessário determinar as constantes cinéticas, basicamente K_S , k e k_d , através de experimento em laboratório em aparato específico.

Dentre os principais fatores que influenciam a atividade de lagoas e tanques de aeração têm-se:

a) Fatores controláveis:

- temperatura;
- agitação (ventos);
- evaporação, e
- chuvas.

b) Fatores parcialmente controláveis:

- permeabilidade do fundo da lagoa;
- nutrientes, e
- carga poluidora.

c) Fatores relacionados ao projeto de construção

- localização;
- número de unidades;
- disposição das unidades;
- dimensões;
- período de detenção, e
- detalhes construtivos (OLIVEIRA, 2000).

Dependendo da situação, os fatores realmente controláveis se reduzem a poucas condições operacionais. No caso da Rípasa SA, por exemplo, procura-se alterar o grau de aeração nas lagoas e no caso da Rhodiaco procura-se estabelecer uma distribuição das cargas pelos diversos reatores. Em alguns casos, só se consegue o objetivo no final do tratamento com a redução deliberada da atividade produtiva da unidade industrial, o que,

com certeza, gera custos indesejáveis. De forma a monitorar a unidade, são feitas frequentemente algumas medições que serão definidas a seguir.

2.3.1. Principais variáveis monitoradas

A preocupação com as condições do despejo industrial final faz com que as principais variáveis sejam avaliadas e comparadas com determinados parâmetros de controle. Estas informações coletadas devem ser suficientes para informar sobre a realidade do comportamento da biomassa e como esta reage a variações das características da água residuária sob tratamento. Devido às características dos efluentes nos casos estudados, a principal análise é a da quantificação de matéria orgânica no efluente. Durante anos se tem procurado desenvolver um teste que melhor indique esta propriedade. Inicialmente, acreditava-se que a medida da fração de sólidos voláteis para os sólidos totais seria suficiente. Contudo, ela é sujeita a diversos erros e, agora, raramente usada (METCALF & EDDY, 1991). Os métodos de laboratório atualmente mais empregados são o da demanda bioquímica de oxigênio (DBO, ou BOD), o da demanda química de oxigênio (DQO, ou COD) e o de carbono orgânico total (COT, ou TOC). Assim, os principais parâmetros de monitoração são definidos a seguir:

2.3.1.1. Demanda bioquímica de oxigênio

Segundo diversos autores, este é o parâmetro mais empregado para poluição orgânica aplicada a águas residuárias ou mesmo as naturais. Na verdade o que se usa é a chamada DBO₅, que é referida a um período de 5 dias. A determinação implica na medida do oxigênio dissolvido utilizado pelos microorganismos na oxidação bioquímica da matéria orgânica. Contudo, sabe-se que tal determinação possui limitações. Mas esta informação é importante porque (1) determina a quantidade aproximada de oxigênio que será necessário para estabilizar biologicamente a matéria orgânica presente, (2) dimensiona os equipamentos da unidade de tratamento e (3) mede a eficiência de determinado processo de tratamento (METCALF & EDDY, 1991). Também pode-se interpretar esta medida como sendo a concentração de matéria orgânica existente na água residual que pode ser biodegradada (OLIVEIRA, 2000).

Pela legislação é necessário reduzir, pelo tratamento de efluentes, a carga orgânica em termos de DBO de 60 a 80%.

2.3.1.2. *Demanda química de oxigênio – (DQO)*

Esta propriedade corresponde à quantidade de oxigênio necessária para oxidar a fração orgânica de uma amostra que seja oxidável pelo permanganato ou dicromato de potássio em solução ácida (PESSOA & JORDÃO, 1982). Este processo oxida quase todos os compostos orgânicos existentes na amostra analisada, liberando o gás carbônico e a água, cuja reação tem, geralmente, uma eficácia de mais de 95% de oxidação de todos os compostos existentes na água.

Apesar desta análise ser obtida rapidamente, em 3 horas aproximadamente, ela não especifica a proporção de matéria orgânica existente na água residual que poderia ser oxidada por bactérias. Porém, do valor encontrado para a DQO uma fração deste corresponde a DBO.

2.3.1.3. *Carbono Orgânico Total (TOC)*

Indicado para casos onde a concentração de matéria orgânica não é tão elevada, este teste é realizado injetando uma quantidade conhecida da amostra em um forno de alta temperatura. O carbono orgânico é oxidado a dióxido de carbono na presença de um catalisador adequado. O CO₂ produzido é então medido quantitativamente por meio de um analisador de infravermelho. Como este teste fornece resultado quase imediatos, ele tem se tornado cada vez mais empregado (METCALF & EDDY, 1991).

2.3.1.4. *pH*

O pH é um fator importante no crescimento das bactérias, sendo que a maioria delas não tolera pH acima de 9,5 ou abaixo de 4,0. METCALF & EDDY (1991) citam a faixa ótima de pH para o crescimento de bactérias entre 6,5 e 7,5.

2.3.1.5. *Sólidos suspensos*

Por definição, a matéria sólida total em uma água residuária corresponde a toda matéria que permanece como resíduo após evaporação a uma temperatura de 103 °C. Se este resíduo é calcinado a 600 °C, as substâncias orgânicas se volatilizam e as minerais permanecem sob a forma de cinzas. Desta maneira se determina o conteúdo de matéria sólida volátil e a matéria fixa (PESSOA & JORDÃO, 1982). A análise de sólidos voláteis é empregada de forma mais comum para medir a estabilidade biológica de lodos de esgotos.

A matéria sólida total pode ainda ser classificada em matéria em suspensão e dissolvida. Os sólidos suspensos são aqueles que passam por um filtro que retém as partículas de diâmetro superior a 1 micron (METCALF & EDDY, 1991).

2.4. USO DA ESTATÍSTICA PARA ANÁLISE DOS DADOS.

2.4.1. Conceitos básicos

Neste trabalho serão apresentadas algumas análises das respostas obtidas que dependem de certos conceitos e princípios de estatística (MONTGOMERY, 1991). Portanto, alguns dos principais aspectos sobre o assunto estão reunidos nesta revisão.

A replicação de um experimento é a repetição de um experimento respeitando-se algumas características. Permite conhecer aproximadamente o erro experimental. A avaliação se dois sistemas são estatisticamente diferentes baseia-se na comparação da média e da variância dos conjuntos de medidas.

A aleatorização de um experimento é um dos cuidados essenciais a serem tomados na execução do experimento. Os métodos estatísticos exigem que as observações (ou os erros) sejam variáveis aleatórias independentemente distribuídas. A aleatorização normalmente torna esta consideração válida. Também é com ela que efeitos casuais tendem a ser eliminados.

A blocagem é uma técnica para aumentar a precisão de um experimento. Blocos são porções homogêneas do material experimental. Significa fazer comparações entre condições de interesse dentro de cada bloco (COLACIOPPO, 1997).

A distribuição normal é uma das distribuições de probabilidade das mais importantes e mais utilizadas. É caracterizada por 2 parâmetros: a média m e a variância s^2 . Muitas vezes utiliza-se a notação $N(m, s^2)$ para representá-la. Assim, a expressão $N(10, 16)$ identifica a distribuição normal com média $m=10$ e variância $s^2 = 16$ (desvio padrão $s=4$).

2.4.2. Testes de hipóteses:

Uma hipótese estatística é uma afirmação sobre os parâmetros de uma distribuição de probabilidade (COLACIOPPO, 1997). Assim, por exemplo:

$H_0: \mu_1 = \mu_2$ (hipótese nula)

$H_1: \mu_1 \neq \mu_2$ (hipótese alternativa).

O objetivo é comparar a resposta de um processo (ou sua variância) entre duas condições experimentais, ou com um dado valor específico. No caso exemplificado está se testando se a média da resposta de um processo em que foi aplicado o tratamento 1 (μ_1) é igual ou diferente da média quando aplicado o tratamento 2 (μ_2).

2.4.3. Análise da variância

Quando se deseja comparar entre si mais de dois tratamentos, testando a igualdade de várias médias, o procedimento adequado é a Análise da Variância (ANOVA). Ao testar a igualdade de vários tratamentos de um fator, as hipóteses formuladas numa análise deste tipo são:

$H_0: \mu_1 = \mu_2 = \dots = \mu_n$ (hipótese nula: todos os tratamentos têm a mesma média)

H_1 : (hipótese alternativa: pelo menos a média de um tratamento difere das demais).

Os cálculos da ANOVA levam a obter uma estatística (F_0) para o teste acima, que quando comparado com a distribuição F de referência, obtém-se o p-valor do teste. O p-valor do teste deve ser interpretado como a probabilidade de ocorrer eventos como aquele que foi observado no experimento, sendo a hipótese nula verdadeira (todas as médias dos tratamentos iguais). Sendo o p-valor um valor baixo, por exemplo 0.01, isso indica que, na hipótese das médias serem iguais, eventos como aquele observado no experimento ocorrem 1% das vezes. Com isso o evento pode ser considerado raro ou então duvidamos da hipótese nula. No caso, pode-se rejeitar a hipótese nula ao nível de 1% de significância.

2.4.4. Desenhos hierárquicos

O objetivo pretendido aqui na utilização de modelos hierárquicos é deduzir os valores da variabilidade que não podem ser medidos diretamente. Assim, num conjunto de medidas de uma variável de processo estão envolvidos nos desvios calculados os desvios no procedimento de amostragem, no procedimento de análise e também devidos ao processo propriamente dito. Portanto, com auxílio dos desenhos hierárquicos (desenho

experimental especialmente planejado) é possível obter uma estimativa desses desvios (COLACIOPPO, 1997).

2.4.5. Planejamento de experimentos

Por planejamento estatístico de experimentos entende-se ao processo de planejar o experimento de maneira apropriada para que os dados coletados possam ser analisados por métodos estatísticos, resultando em conclusões objetivas e válidas (MONTGOMERY, 1991 e BARROS NETO et al., 1996).

CAPÍTULO 3. ESTUDO DE CLASSIFICAÇÃO DA MATÉRIA-PRIMA PARA A POLPAÇÃO

3.1. CARACTERIZAÇÃO DO PROBLEMA A SER ESTUDADO

A classificação da matéria-prima vegetal (troncos de Eucaliptus) para a indústria de papel e celulose, normalmente, vem sendo feita por uma de suas propriedades que pode ser facilmente obtida (que é a densidade básica). Existe um estudo com a finalidade de se estender a análise considerando também outras informações do lote da madeira. Assim, dados da procedência, bem como as propriedades básicas, podem ser utilizadas para este fim. A classificação mais adequada depende de uma definição das informações a serem utilizadas. Há a tendência de se tentar utilizar variáveis como o rendimento depurado, o índice Kappa, e outras ligadas ao resultado da operação de polpação da madeira, como definidas na seção 3.3. Como estes dados são de análise dispendiosa e que requerem um tempo relativamente elevado, pretende-se obter tais propriedades pela sua relação com as outras variáveis que podem ser medidas mais facilmente. Para este fim, utilizar-se-á Redes Neurais. Este estudo foi realizado em cooperação com a Empresa Ripasa S.A. Papel e Celulose, que necessitava de avaliar novas propostas para classificar os lotes de matéria-prima para usá-los de maneira mais adequada no processo.

Numa primeira etapa, foi feita uma análise estatística nos dados que foram fornecidos pela empresa. Em seguida, os dados foram colocados no formato exigido para se introduzir no programa de treinamento de redes neurais. No processo de treinamento das redes foram observadas diversas das condições necessárias para se obter resultados satisfatórios. Assim, realizou-se o treinamento de várias configurações e topologias de redes para definir as que ofereceriam resultados promissores. A partir daí, definiu-se como objetivo a redução do erro RMS, como definido na seção 3.5.8.1, no conjunto de teste. Conseguiu-se atingir um resultado que oferece, tanto para o conjunto de treino como para o de teste, um erro RMS compatível com a qualidade dos dados coletados.

O passo seguinte foi obter uma regra que fosse consagrada para classificar a matéria-prima. A Ripasa teve a iniciativa de utilizar a experiência dos especialistas em matéria-prima da empresa para serem aplicadas a esse estudo. Assim, recebeu-se uma

classificação proposta por dois dos especialistas para cada lote do banco de dados. Também foi perguntado a esses especialistas quais os principais critérios adotados. De posse destas informações foi possível elaborar redes neuronais que, de acordo com as variáveis principais indicadas, forneciam a classe de cada lote.

Finalmente, com as redes de melhor resultado, foram criados os aplicativos que permitem a análise completa dos resultados obtidos, e os que podem ser empregados para classificar os novos lotes.

A análise da sensibilidade paramétrica foi realizada partindo da escolha da rede com o melhor ajuste com o conjunto de dados. Inicialmente, realizou-se a análise utilizando como níveis o padrão da média descontando-se e somando-se à ela 10% de seu valor, para ter-se, respectivamente, os níveis inferior e superior de cada variável. Contudo, a análise dos resultados foi complicada por este padrão ser altamente dependente do valor absoluto das variáveis e das unidades utilizadas. Assim, procurou-se uma maneira de evitar este problema. A proposta seria a de utilizar o desvio padrão de cada variável. Acredita-se que desta forma os resultados obtidos tornaram-se mais coerentes.

3.2. OBJETIVOS PROPOSTOS DESTE ESTUDO

No intuito de propiciar um processo global de investigação sobre o processo produtivo, foi realizado um estudo sobre a primeira etapa do processo. Assim, enfocando-se a classificação da madeira - que é a matéria-prima essencial do processo - e a sua relação direta com a operação seguinte, a polpação, pretende-se:

1. Indicar ações para melhorar o processo, aumentando o rendimento em produtos e a qualidade dos mesmos.
2. definir, com auxílio da técnica de redes neuronais, um novo procedimento para a utilização dos diversos lotes de madeira, visando uma produção maximizada com as melhores características possíveis para o produto. Pode ser considerado ainda uma situação de menor volume dos despejos.
3. Prever alguns dos efeitos da implantação da nova classificação no processo de cozimento.
4. Gerar, em tempo real, valores para variáveis de processo não mensuráveis *on-line*.

5. Implementar a técnica das redes neuronais na indústria, tornando-a uma ferramenta de fácil acesso aos engenheiros de processo.

3.3. VARIÁVEIS E PROPRIEDADES ENVOLVIDAS

As variáveis que compõe o banco de dados existente atualmente são as seguintes:

Parque Florestal – procedência da madeira.

Espécie - espécie da madeira utilizada.

Talhão – identificação da área da fazenda onde se desenvolveu determinada madeira.

Idade - refere-se a informação do tempo decorrido entre o plantio e o corte da madeira, em meses.

Produtividade – incremento médio anual do volume da madeira em unidades de esteres ♦/hectare.ano.

Densidade básica - propriedade da madeira, determinada em laboratório, e expressa em kg de matéria seca /m³.

Extrativos totais – constituinte da madeira que corresponde a quantidade de material facilmente extraído em solventes orgânicos e água quente. Em % massa em base seca.

Lignina solúvel – constituinte da madeira que é solubilizada no ácido sulfúrico após a retirada dos extrativos totais. Expressa em % massa em base seca.

Lignina insolúvel - constituinte da madeira que não é solubilizada no ácido sulfúrico após a retirada dos extrativos totais. Expressa em % massa em base seca.

Holocelulose - principal constituinte da madeira para o processo. Em % massa em base seca.

♦ Unidade equivalente a volume. (1 estere ~ 0,7 m³)

Rendimento depurado – é a relação entre a quantidade de massa seca após a polpação e depuração e a quantidade de madeira seca utilizada, avaliada em experimento piloto no laboratório. Em % massa em base seca.

Teor de rejeitos - massa seca detida no depurador após polpação em laboratório. Em % massa em base seca.

Número kappa – indica, de forma indireta, o teor de lignina residual na polpa.

Viscosidade – indica, de forma indireta, o grau de degradação da polpa (em cP).

Alcali ativo residual (AAR) - quantidade residual de reagente presente no licor após a polpação em laboratório. Expressa em g/l.

3.4. DADOS FORNECIDOS PARA REALIZAR O ESTUDO

Foi fornecido para realizar o estudo um conjunto de dados obtidos pela Ripasa em seus laboratórios contendo os valores das variáveis acima para 165 amostras totais, tendo-se 3 amostras por talhão, sendo, portanto 55 talhões.

3.4.1. Banco de dados com as propriedades principais do lote

A Tabela 28 que se encontra no Anexo (pag. 144) traz os dados utilizados no treinamento das redes para a predição das propriedades da polpa após o processo de polpação em laboratório.

3.4.2. Classificação obtida com os especialistas em matéria-prima

Na mesma Tabela 28 (Anexo, pag. 144), as duas últimas colunas correspondem às classificações arbitradas por dois especialistas da Ripasa e que serão utilizadas no treinamento de redes com o objetivo de determinar a classe de novos lotes.

3.5. METODOLOGIA ADOTADA NA REALIZAÇÃO DO ESTUDO

Utilizando um software específico para redes neuronais, o Neuralworks Professional II Plus, foram testadas várias configurações de números de neurônios

distribuídos em uma ou duas camadas internas. Além disto foram utilizadas outras funções de transferência além da sigmoidal e ainda algumas técnicas de treinamento.

3.5.1. Algoritmos de treinamento utilizados neste estudo

Os principais algoritmos de treinamento que foram usados neste estudo foram: o “Backpropagation” padrão, o “Quickprop” (FAHLMAN, 1988) e o “Delta-bar-delta” (Neuralware, 1995). Para considerar a classificação final da matéria-prima foi treinada uma rede neuronal do tipo “Learning Vector Quantization (LVQ)”.

3.5.2. Cálculo das médias

Para um mesmo talhão existem 3 amostras que são analisadas em laboratório independentemente. O cálculo da média desses valores confere aos dados uma maior confiabilidade na representação do lote oriundo deste talhão. As médias foram obtidas através de planilha eletrônica, pelo software excel.

3.5.3. Padronização de talhão, fazenda (parque florestal) e espécie

A variável talhão foi utilizada no treinamento das redes iniciais, e fez-se uma nova numeração sequenciada de 1 a 55. Posteriormente, quando se decidiu retirar esta variável como entrada da rede, deixou-se de renumerar esta variável. As variáveis parque florestal e espécie precisam ser codificadas de forma a fornecer à rede estas informações adequadamente. Neste caso, criam-se tantas variáveis novas quanto é o número de classes. Cada uma destas variáveis (faz_cod0, faz_cod1, ..., faz_cod9) assume valor zero ou um, identificando assim se um dado se refere a um determinado parque florestal ou não. O mesmo procedimento ocorre com a espécie. A ordem das espécies e fazendas é estipulada pela frequência em que cada uma aparece, tendo, por exemplo, como faz_cod0=1 a de menor número de casos.

3.5.4. Manipulação dos dados

O tratamento dos dados e dos resultados obtidos é feito por intermédio do Datasculptor da Neuralware, que também integra o pacote do programa de treinamento de

redes neuronais. Alguns valores obtidos para avaliação de resultados podem ser retirados do próprio Neuralworks, como o erro médio quadrático (RMS) e os coeficientes de correlação lineares (R) entre o valor desejado e o obtido das variáveis para o conjunto de treinamento e para o de teste.

Este software permite uma fácil modificação na estrutura dos dados, através da aplicação de determinadas operações organizadas seqüencialmente. Assim, os dados brutos foram filtrados para as variáveis desejadas, na ordem requerida, e as variáveis espécie e fazenda foram ainda codificadas, possibilitando a criação do arquivo de entrada para o treinamento das redes. Depois que as redes são treinadas no Neuralworks, pode-se utilizar o Datasculptor para aplicar o “flash code” que é gerado no treinamento. Este procedimento será melhor explicado no item 3.5.9.3.

3.5.5. Aleatorização dos conjuntos de treino e de teste

Inicialmente trabalhou-se com a totalidade do conjunto de pontos para uma primeira avaliação das possibilidades, e posteriormente trabalhou-se com um conjunto de pontos aleatórios para treinamento, distinto do complementar, para teste (CAUDILL, M., 1991). O tamanho do conjunto de teste adotado foi de 20% do tamanho do conjunto total. Para este fim utilizou-se o Datasculptor.

3.5.6. Aleatorização no treinamento da rede

Um ponto importante no treinamento das redes é a aleatorização da seqüência dos pontos a serem apresentados na rede. No caso, o próprio Neuralworks faz isto na primeira vez que lê o arquivo de treinamento.

3.5.7. Treinamento e Teste . Procedimento “Save best”

Como já foi dito anteriormente, a rede mais adequada não é aquela onde se observa o menor erro para o conjunto usado no treinamento, mas sim aquela que tem um erro para o conjunto de teste no valor mínimo. O procedimento “save best” do Neuralworks faz com que periodicamente a rede treinada seja usada com o conjunto de teste, e quando o erro diminui salva a rede. Assim, ao final do processo, resgata-se o último arquivo da rede salvo para se ter a melhor rede obtida na sessão de treinamento executado.

3.5.8. Instrumentos utilizados do software Neuralworks Pro II

O Neuralworks Pro II possui alguns instrumentos para auxiliar a análise do processo de treinamento de redes. Dentre eles, os que foram mais utilizados foram:

3.5.8.1. Erro RMS ("RMS error")

É o parâmetro mais utilizado para avaliar como anda o treinamento. O programa apresenta um valor que é ponderado nas variáveis e para todo o conjunto de dados:

$$RMS_{\text{ponto}_m} = \sqrt{\frac{\sum_1^n [(d_i - o_i)^2]}{n}} \quad (29)$$

$$RMS_{\text{epoch}} = \sqrt{\frac{\sum_1^m \left\{ \frac{\sum_1^n [(d_i - o_i)^2]}{n} \right\}}{m}} \quad (30)$$

onde d_i são os valores desejados para a variável de saída da rede; o_i são os valores obtidos para a variável de saída; m é o número de pontos para cada ciclo de atualização dos pesos (epoch); n é o número de variáveis de saída; RMS_{ponto_m} é o erro para cada um dos pontos; e RMS_{epoch} é o erro para todas as variáveis de saída e computados todos os pontos do ciclo (epoch).

3.5.8.2. Matriz de Resultados ("Confusion matrix")

Para cada uma das variáveis de saída pode-se usar esse instrumento que cria um gráfico da distribuição dos pontos em várias faixas do valor desejado da variável. A altura da barra em cada posição da matriz corresponde ao número de ocorrências naquela faixa. Assim, o ideal seria que todos os pontos caíssem na diagonal principal. O número que aparece ao lado corresponde ao coeficiente de correlação entre o valor obtido e o desejado.

3.5.8.3. *Índice de correlação (“Correlation”)*

Na nova versão do software surgiu uma nova opção que é o gráfico do índice de correlação. Assim é possível acompanhar o índice ao longo do processo, para cada uma das variáveis de saída.

3.5.8.4. *Somatório das entradas (“Summation”)*

Um instrumento auxiliar é o gráfico do somatório das entradas. Ele é utilizado para se ter uma idéia do valor do somatório ao longo do “epoch” (número de casos usados para reajustar os pesos), valor este que é alimentado na função de transferência. É importante monitorar este valor lembrando que a função sigmoideal, apesar de não apresentar problemas com o seu domínio, retorna valores muito próximos à unidade quando o valor introduzido está acima de +6 ou abaixo de -6. Assim, se o somatório variar acima desta faixa pouco treinamento está sendo executado, já que o valor da saída não sofre mudanças. Por outro lado, se os somatórios forem valores pequenos, o processo é mais demorado.

3.5.9. Instrumentos para análise externos ao software Neuralworks

Algumas análises de resultado foram realizadas em outros programas que não o Neuralworks, as principais são:

3.5.9.1. *Estatísticas*

Utilizando um software de estatística algumas análises foram realizadas com o conjunto de dados. Basicamente se utilizou a estatística descritiva.

3.5.9.2. *Gráficos*

Talvez o recurso mais interessante e que melhor indica como foi o treinamento seja o gráfico do valor obtido pelo valor desejado de cada variável. Para facilitar ainda mais a análise foram colocadas duas curvas auxiliares indicando os desvios em nível de 1% e de 5%. Deve-se lembrar que esse desvio é na ordenada do gráfico, correspondendo às funções $y=0,99x$; $y=1,01x$; $y=0,95x$ e $y=1,05x$, para 1% e 5% de desvio respectivamente.

3.5.9.3. "Flash code"

Tendo-se a rede treinada é desejável poder utilizá-la em outras condições fora do conjunto de treino ou de teste. A forma mais fácil de se fazer isto é através do "flash code", que é o arquivo fonte gerado em linguagem C pelo próprio Neuralworks. Assim, para utilizar este recurso necessita-se de um compilador C e de um programa que gerencie a chamada da subrotina. Contudo, no programa DATASCULPTOR, existe a possibilidade de se importar este arquivo e utilizá-lo como um módulo do programa. Infelizmente, a Neuralware, fornecedora do Datasculptor, não atualizou mais o programa e segundo, informações por e-mail, a Aspentech não irá distribuir as versões antigas do Datasculptor. Além disso, algumas redes geradas pela nova versão do Neuralware (backpropagation é uma delas!) não são diretamente lidas pelo Datasculptor, exigindo algumas mudanças no código para serem utilizadas. A saída que encontramos foi a de obter o "flash code" com a versão antiga do Neuralworks. Para isto, é necessário salvar os pesos da rede treinada, criar uma rede idêntica na versão antiga e aí importar os pesos e criar o código fonte. Já para as redes LVQ usadas na fase de classificação propriamente dita, não foi necessário este procedimento, exigindo algumas poucas mudanças no código gerado.

3.6. RESULTADOS OBTIDOS

Neste capítulo os principais resultados do estudo para classificação da matéria-prima do processo de polpação da indústria de papel e celulose são mostrados e comentados.

3.6.1. Análise inicial dos dados

Com o objetivo de demonstrar a distribuição dos dados para cada uma das variáveis que serão utilizadas no treinamento da rede, apresenta-se o conjunto de histogramas construídos em 10 níveis (exceto para fazenda e idade). O histograma da idade refere-se ao conjunto de dados inicialmente fornecidos em anos.

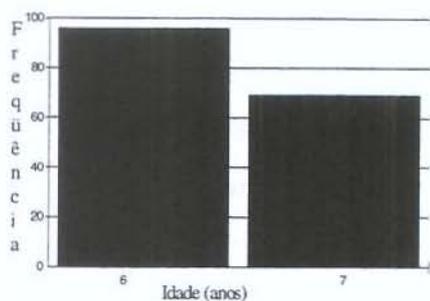


Figura 7 - Histograma da idade em anos

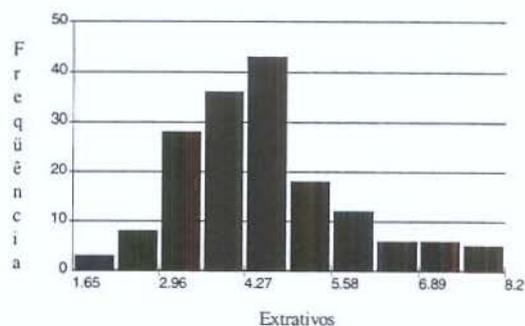


Figura 10 - Histograma de extrativos (em %massa)

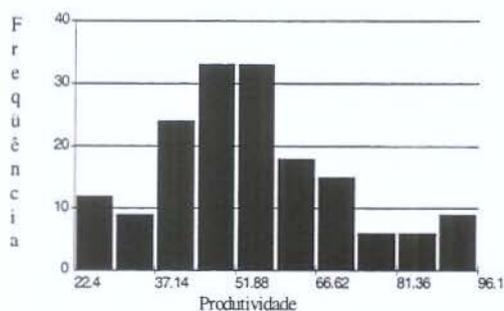


Figura 8 - Histograma da produtividade (em esteres/hectare.ano)

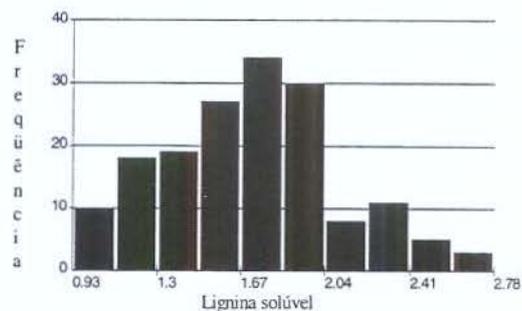


Figura 11 - Histograma de lignina solúvel (em %massa)

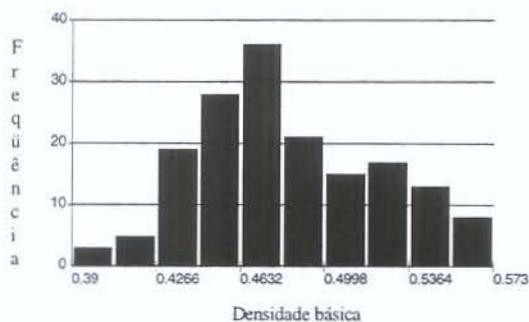


Figura 9 - Histograma da densidade básica (em kg matéria seca/m³)

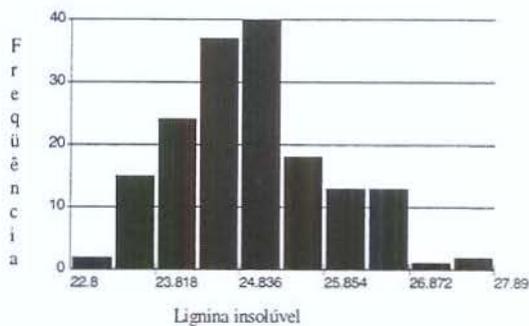


Figura 12 - Histograma de lignina insolúvel (em %massa)

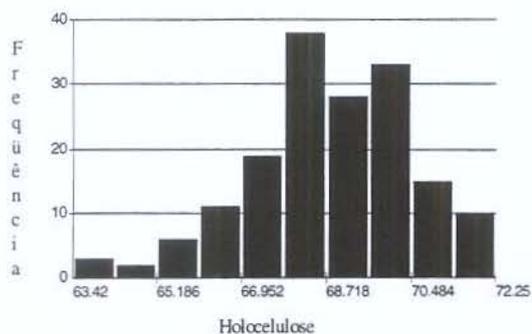


Figura 13 - Histograma de holocelulose (em %massa)

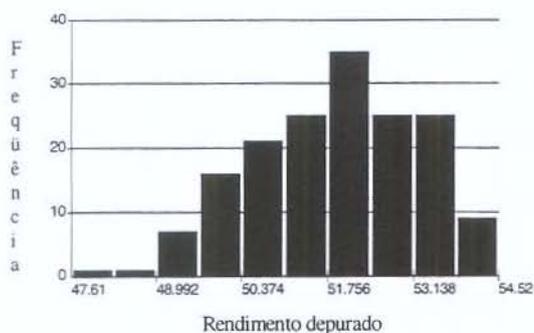


Figura 14 - Histograma de rendimento depurado (em %massa)

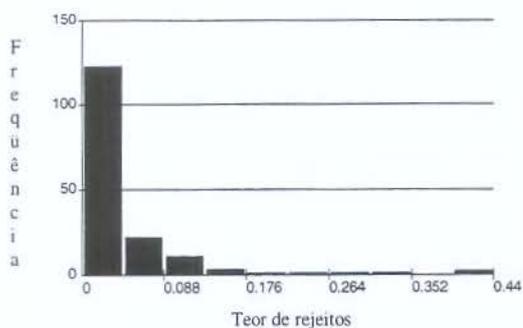


Figura 15 - Histograma de teor de rejeitos (em %massa)

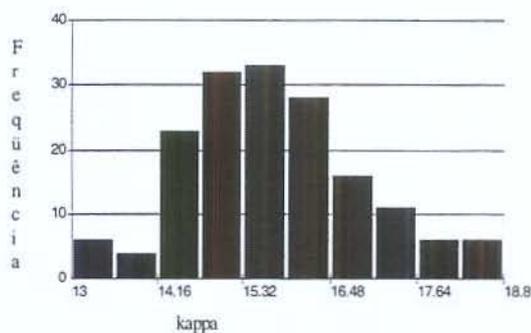


Figura 16 - Histograma do índice kappa

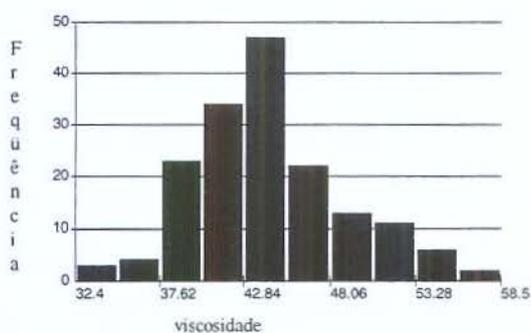


Figura 17 - Histograma de viscosidade (em cP)

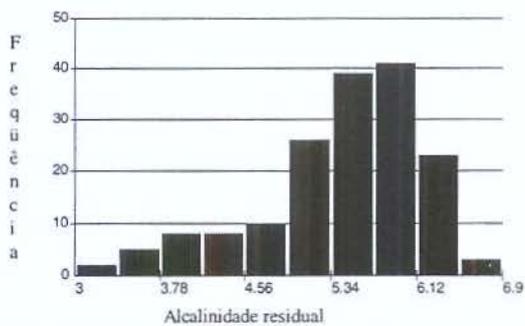


Figura 18 - Histograma de alcalinidade residual (em g/l)

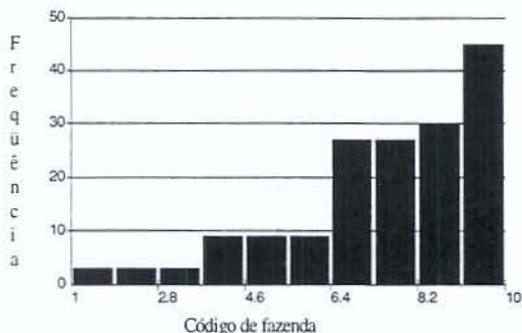


Figura 19 - Histograma do código de fazenda

Observe-se que a maioria das variáveis tem uma distribuição muito parecida com a distribuição normal. As variáveis de codificação como fazenda e espécie obviamente não seguem este padrão devido ao método usado de codificação para as mesmas. O programa DATASCULPTOR faz a conversão dos nomes de fazenda e espécie por um número. O padrão adotado segue o número de apresentações de cada uma das fazendas e espécies. Por isso nota-se que a medida as maiores quantidades sempre estão codificadas com o maior número (“ranking”). Contudo, esta forma de codificação foi posteriormente modificada para ser usada no treinamento, como foi explicado no item 3.5.3.

Outra observação importante é sobre o comportamento da variável teor de rejeitos. Observa-se que a maioria dos dados possui baixo valor do teor. Surgem valores mais elevados, mas isolados, o que faz com que note-se uma extensa cauda. Mas, isto não deve ser motivo para um tratamento especial para esta variável, a não ser lembrar que, segundo a fonte dos dados, o erro experimental da determinação desta propriedade é elevado.

A seguir apresenta-se uma grade da distribuição dos casos estudados mostrando a ocorrência dos experimentos segundo a espécie e segundo a fazenda (parque florestal). Observa-se a exiguidade de experimentos em determinadas espécies e fazendas. Há a preocupação de se averiguar o que a rede treinada pode fornecer como resultado nestes casos. Verificamos na Tabela 3 a contagem para a ocorrência de cada caso possível entre os fatores Fazenda e Espécie.

Tabela 3 - Distribuição das ocorrências dos testes segundo a fazenda e espécie.

Código da Fazenda	Código referente à espécie plantada					
	1	2	3	4	5	ALL
1	0	0	0	3	0	3
2	0	0	0	0	3	3
3	0	0	0	0	3	3
4	0	0	0	0	9	9
5	0	0	0	0	9	9
6	0	0	0	0	9	9
7	3	0	0	9	15	27
8	0	0	9	0	18	27
9	0	0	0	15	15	30
10	0	15	15	0	15	45
ALL	3	15	24	27	96	165

3.6.2. Estimativa do erro de amostragem e medida das variáveis

No processo de treinamento da rede obtém-se alguns parâmetros para comparar o resultado esperado com o obtido pela rede em cada iteração. Até o momento, o critério de parada para o processo obedecia a valores arbitrariamente estabelecidos. Pretende-se utilizar a análise da variância para se ter uma boa estimativa do erro envolvido no processo de amostragem e de análise das amostras. Estes valores estimados, portanto, serviriam para avaliar o momento de encerrar o procedimento de treinamento, já que além deste ponto não existe certeza estatística no valor experimental utilizado. A forma utilizada aqui foi comparar a variância dos dados experimentais com a variância do erro entre o valor desejado e o obtido pela rede, para cada uma das variáveis de saída da rede.

Os dados utilizados provém de banco de dados existente, e acredita-se que, na sua elaboração, os devidos cuidados com a aleatorização dos amostradores e técnicos de laboratório tenham sido tomados. De outra forma, os resultados aqui analisados podem ser utilizados, mas sem a certeza de representatividade da realidade do processo.

Os dados foram dispostos considerando que retirou-se três amostras do lote de um determinado talhão, e estas foram analisadas (com o cuidado da aleatorização na seqüência, codificação de amostras, etc.). Pode-se esperar que as análises dessem resultados iguais, mas devido a variabilidade inerente ao processo de amostragem, de análise e por outros fatores não detectáveis, os valores apresentam diferenças.

A organização dos dados seguiu um desenho hierárquico, onde para cada talhão se obtém três replicações para as propriedades em estudo. Como não há replicação para cada amostra, engloba-se toda a variabilidade dos processos de amostragem e medida no valor que será obtido desta forma (COLACIOPO, 1997). Em outras palavras, a variabilidade destes processos estaria incluída na variabilidade em talhão. O programa MINITAB prevê a possibilidade de trabalhar com fatores aninhados, e a média dos quadrados (Adj. MS) é reconhecida como uma estimativa para a variância relativa ao fator. Executando o programa MINITAB para o conjunto de dados em estudo, obteve-se os resultados mostrados nas Tabelas 4 a 8, para cada uma das variáveis medidas. Nestas tabelas, que se referem ao resultado do programa, a sigla DF representa os graus de liberdade, Adj SS representa a soma dos quadrados ajustada, Adj MS é a média dos quadrados ajustada e TALHAOC se refere ao código do talhão correspondente.

Rendimento Depurado

Tabela 4 - Resultados para a análise da variância do conjunto de dados, mostrando a variabilidade do *rendimento depurado* entre as amostras/análises para cada talhão.

Analysis of Variance for rendimento depurado			
Source	DF	Adj SS	Adj MS
TALHAOC	54	256.5399	4.7507
Amostra(TALHAOC)	110	42.0094	0.3819
Error	0	0.0000	0.0000
Total	164		

Índice Kappa

Tabela 5 - Resultados para a análise da variância do conjunto de dados, mostrando a variabilidade do *índice Kappa* entre as amostras/análises para cada talhão.

Analysis of Variance for índice kappa			
Source	DF	Adj SS	Adj MS
TALHAOC	54	178.3245	3.3023
Amostra(TALHAOC)	110	46.4200	0.4220
Error	0	0.0000	
Total	164		

Viscosidade

Tabela 6 - Resultados para a análise da variância do conjunto de dados, mostrando a variabilidade da *viscosidade* entre as amostras/análises para cada talhão.

Analysis of Variance for Viscosidade			
Source	DF	Adj SS	Adj MS
TALHAOC	54	2464.296	45.635
Amostra(TALHAOC)	110	988.940	8.990
Error	0	0.0000	
Total	164		

Alcalinidade Residual

Tabela 7 - Resultados para a análise da variância do conjunto de dados, mostrando a variabilidade da *alcalinidade residual* entre as amostras/análises para cada talhão.

Analysis of Variance for Alcalinidade residual			
Source	DF	Adj SS	Adj MS
TALHAOC	54	84.96909	1.5735
Amostra(TALHAOC)	110	13.06000	0.1187
Error	0	0.0000	0.0000
Total	164		

Teor de Rejeitos

Tabela 8 - Resultados para a análise da variância do conjunto de dados, mostrando a variabilidade do *teor de rejeitos* entre as amostras/análises para cada talhão.

Analysis of Variance for Teor de rejeitos			
Source	DF	Adj SS	Adj MS
TALHAOC	54	0.5022036	0.0093001
Amostra(TALHAOC)	110	0.1762667	0.0016024
Error	0	0.0000	0.0000
Total	164		

Esses resultados de variância: 0,3819 para o rendimento depurado, 0,422 para o Kappa, 8,99 para a viscosidade, 0,1187 para a alcalinidade residual e 0,0016024 para o teor de rejeitos, serão utilizados posteriormente quando avaliarmos o resultado da rede treinada.

3.6.3. Avaliação inicial do número de camadas, número de neurônios, função de ativação e algoritmo de treinamento

Avaliações preliminares sobre estes parâmetros foram realizadas com o banco de dados, a fim de identificar em primeira instância as alternativas mais promissoras. Inicialmente a idéia foi de trabalhar com o conjunto total dos pontos (165 pontos), considerando-os individualmente. Avaliou-se a questão de utilizar o conjunto de dados ordenado contra o de organização aleatória, e realmente os resultados do aleatório foram melhores, principalmente quando se trabalha com um “epoch” (tamanho do conjunto de dados para atualizar os pesos no treinamento) menor que o total. Treinou-se várias redes com números de neurônios na camada interna entre 1 e 10. Utilizou-se também a opção “Cascade Learning” onde o treinamento é sucessivo para um número crescente de neurônios, encontrando o melhor número de neurônios para uma camada. O número que foi indicado foi o de 10 neurônios. Até aqui, a média do RMS (“round mean square”) encontrado foi de 0,07, chegando no mínimo a 0,06. Ainda foram usadas duas camadas, a opção de conectar os neurônios com todos os anteriores, outras funções de ativação e outras regras de treinamento.

Os resultados inicialmente obtidos apontaram para uma rede de 1 camada, com 8 a 12 neurônios, conectados com todos os anteriores, usando a função sigmoideal, com epoch total e usando “quickprop” (FAHLMAN, 1988) como regra de treinamento. Mesmo assim, a barreira do RMS de 0,058 não conseguia ser ultrapassada. Descobriu-se, então um meio de melhorar a convergência, usando a faixa de entrada padrão das variáveis. Acontece que o programa normaliza os valores de todas as variáveis automaticamente. A faixa utilizada pode ser ajustada, podendo ser de (-1 , 1), de (0 , 1) ou ainda (-1,5 , 1,5). Observou-se por um instrumento específico que a somatória das entradas em cada um dos elementos (neurônios) estava muito aquém do valor máximo permitido para a função sigmoideal. Aumentando, porém, a faixa de entrada (“input range”) esses somatórios se tornam melhores distribuídos e o aprendizado ocorre mais rápido. O resultado alcançado em 34.E6 iterações foi de RMS= 0,0475.

A partir deste ponto se começou a trabalhar com a média por talhão. Apesar de incluir-se nas entradas, variáveis que poderiam explicar a diferença entre as amostras de mesmo talhão, considerou-se que a maior parte desta variabilidade não estaria sendo

explicada à rede para que esta pudesse aprender as diferenças. Com isso, foram testadas novas possibilidades, retirando entradas tais como a indicação de talhão, e aquelas que exigiam análise em laboratório (extrativos, lignina solúvel e insolúvel e holocelulose no caso).

Treinando uma rede com 4 neurônios de saída até 99 E6 iterações, com 10 neurônios em uma camada, usando faixa de (-2, 2), "quickprop", com todas as entradas, chegou-se ao resultados seguinte: RMS=0,0281. O índice de correlação foi de 0,9863, 0,9755, 0,9743, 0,9771, 0,9620 para o rendimento depurado, o kappa, a viscosidade, a alcalinidade residual e o teor de rejeitos respectivamente.

O conjunto dos pesos das redes treinadas considerados como os melhores são aqueles que apresentam o mínimo erro RMS para o conjunto de teste. Ao observar a Figura 20, nota-se que o erro para o conjunto de treinamento tende a diminuir indefinidamente, enquanto que o erro para o conjunto de teste apresenta uma curva com um ponto de mínimo.

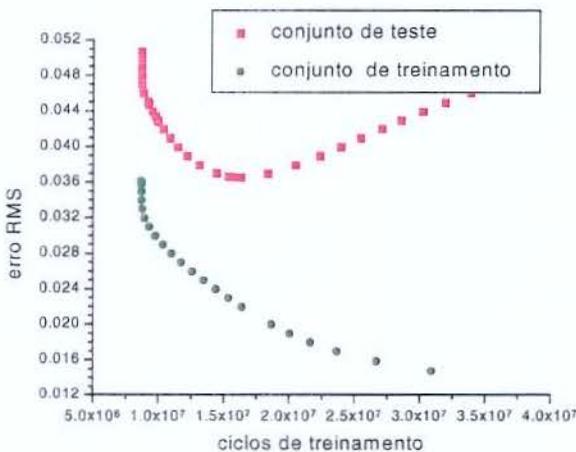


Figura 20 - Comparação dos erros RMS para o conjunto de teste e de treinamento, ao longo dos ciclos de treinamento para uma das redes treinadas.

Reavaliando, então, o número ideal de neurônios na camada interna, pela Figura 21 nota-se que a rede com menor erro RMS do conjunto de teste é aquela que tem 8 neurônios na camada interna. O erro RMS para o conjunto de treinamento oscila entre os pontos, isto devido principalmente à convergência no processo de treinamento da rede. Mas, sem

dúvida, existe uma tendência clara e previsível que este deva diminuir a medida que se aumente o número de neurônios na camada interna.

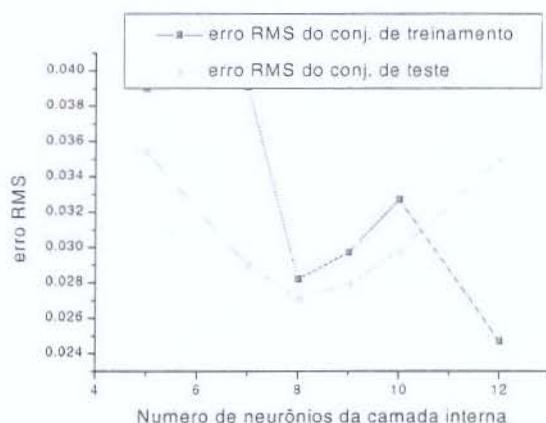


Figura 21 - Comparação dos erros RMS para o conjunto de teste e de treinamento, mudando-se o número de neurônios na camada interna.

Modificando a regra de treinamento para a Delta-bar-Delta, usando a conexão dos elementos com todos os anteriores, as duas configurações anteriores apresentaram melhores resultados. Mas descobriu-se, então, que os dados fornecidos pela RIPASA continham alguns poucos erros de digitação, e decidiu-se treinar novamente a rede com os dados corrigidos.

3.6.4. Depuração do resultado

Os resultados anteriores foram obtidos com o banco de dados inicialmente fornecido pela Ripasa. Oportunamente, observou-se que haviam erros de digitação nos dados, bem como um erro de interpretação no número de parques florestais. Acontece que haviam três nomes diferentes para o mesmo parque florestal, e por isso foi interpretado inicialmente como se fossem 10 parques florestais (fazendas) distintos no total, e na realidade são 8. Ao realizar as mudanças no banco de dados, reiniciou-se o processo de treinamento da rede. Estranhamente os erros foram muito acima daqueles que haviam sido obtidos antes. Chegou-se a conclusão que isto seria devido ao fato de que os pesos encontrados anteriormente estarem otimizados. Assim, partiu-se da melhor rede encontrada com 10 neurônios na camada interna e foram treinadas redes com pequenas modificações

de cada vez. Primeiramente mudou-se o código de fazenda para aqueles dados com problemas. Com a rede treinada para a nova condição, retirou-se as entradas a mais e novos pesos foram obtidos. Assim, sucessivamente para todas as modificações necessárias tais como a eliminação do teor de rejeitos da camada de saída. A rede obtida desta forma apresentou um erro RMS para o conjunto de treino de 0,0307, com índice de correlação de 0,9762; para o conjunto de teste de 0,0296, com índice de correlação de 0,9843. Individualmente para cada variável o índice de correlação encontrado foi de $R_{rendep}=0.978$, $R_{kappa}=0.988$, $R_{visco}=0.960$, $R_{alc_re}=0.978$; e para o conjunto de teste foi de $R_{rendep}=0.988$, $R_{kappa}=0.989$, $R_{visco}=0.970$, $R_{alc_re}=0.990$.

A impressão da tela para o ponto ótimo encontrado pode ser vista adiante na Figura 22. A Figura 23 demonstra o que acontece ao deixar treinando após o ponto ótimo. Para facilitar a visualização, o gráfico após o ponto ótimo só registra os pontos de 50 em 50. Repare-se que ocorre o esperado, ou seja, o erro do conjunto de treino decresce enquanto que o erro do conjunto de teste começa a se elevar gradualmente.

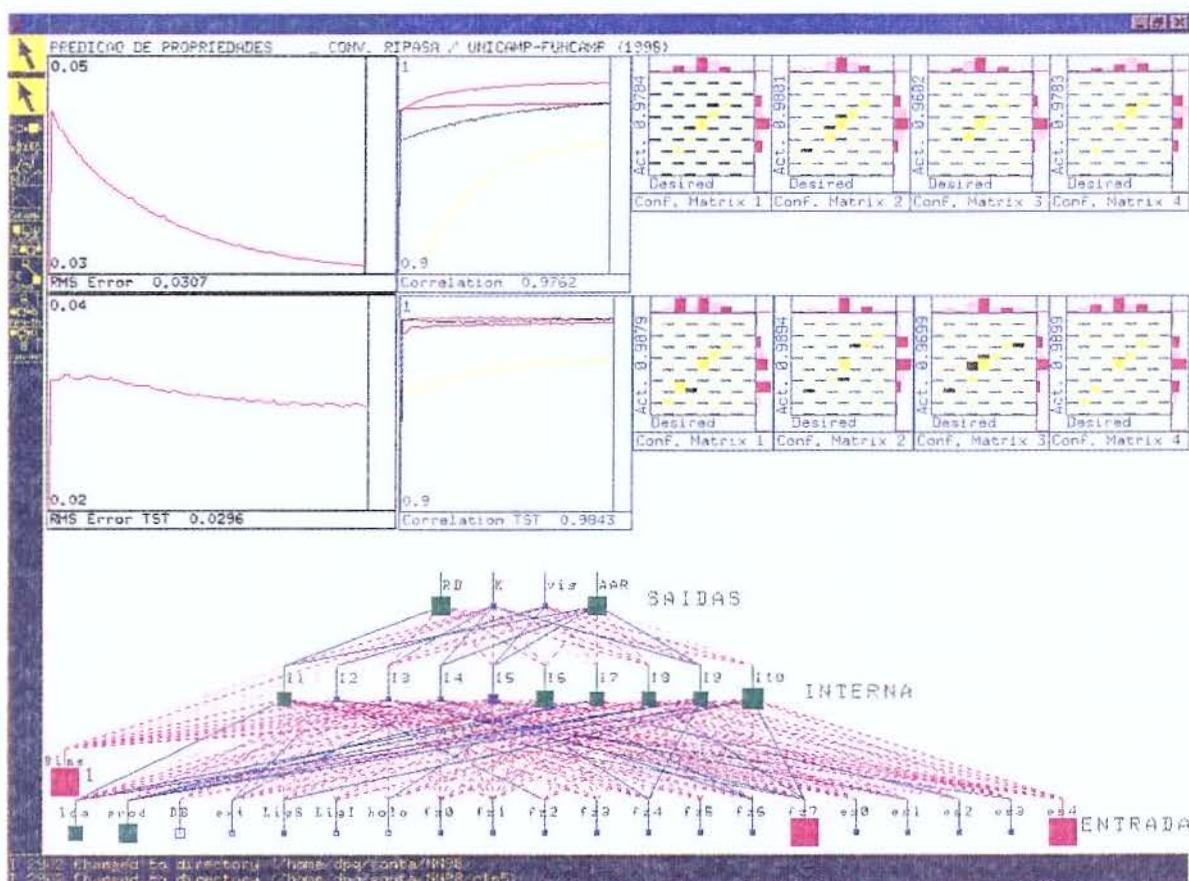


Figura 22 – Tela do programa de treinamento para a rede de predição de propriedades. Ponto onde o erro do conjunto de teste é o menor encontrado.

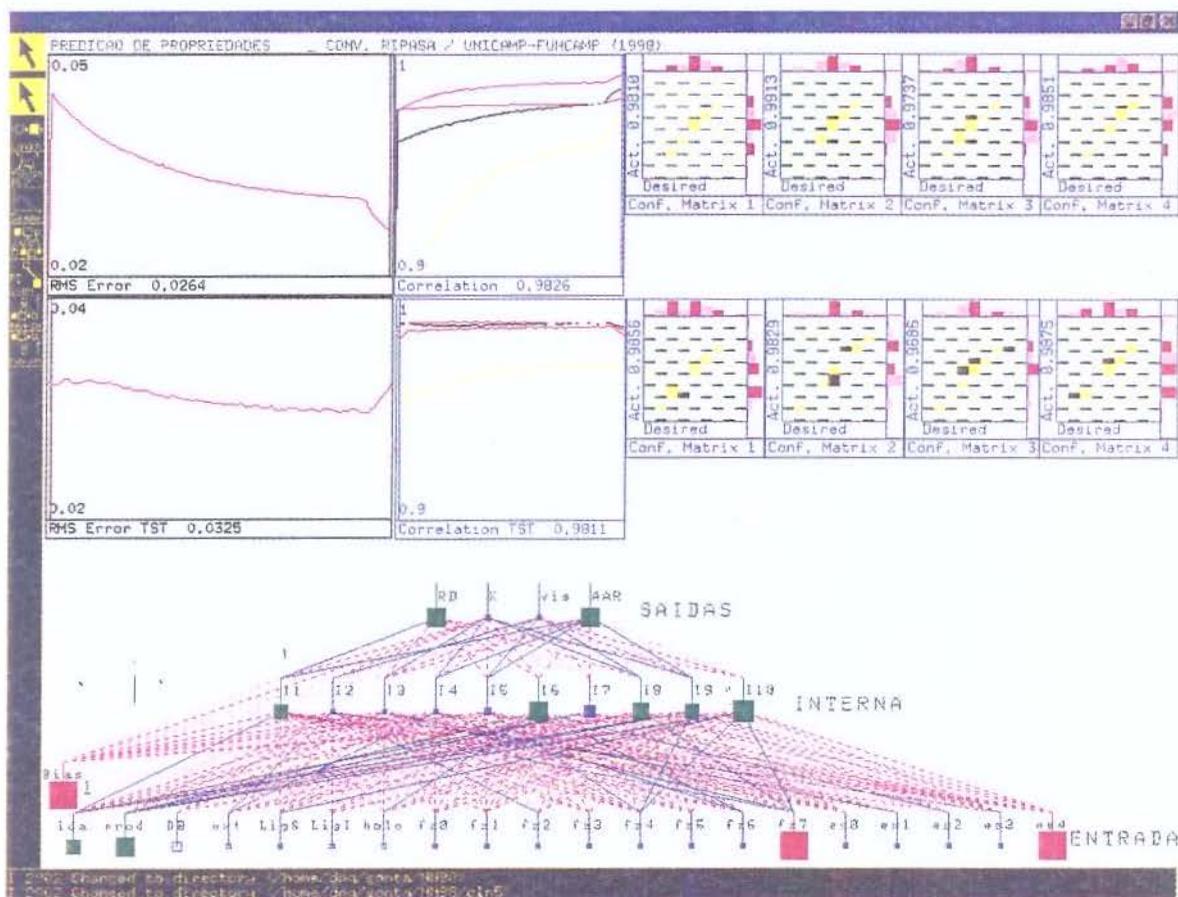


Figura 23 – Tela do programa de treinamento para a rede de predição de propriedades. Prosseguindo o treinamento além do ponto onde o erro do conjunto de teste é o menor encontrado observa-se a elevação deste erro.

A Figura 24 demonstra o erro na predição da propriedade rendimento depurado. Observa-se que os pontos usados como teste, não utilizados no processo de treinamento portanto, estão com erros compatíveis com aqueles dos pontos do conjunto de treino. Nesta variável observa-se que quase todos os pontos caem dentro da faixa de 1% de desvio.

A Figura 25 demonstra o erro na predição da propriedade do número kappa. Observa-se que os pontos usados como teste também estão com erros compatíveis com aqueles dos pontos do conjunto de treino. Nesta variável poucos pontos caem fora da faixa de 1% de desvio, mas todos estão dentro da faixa de 5% de desvio.

A Figura 26 demonstra o erro na predição da viscosidade. Observa-se que os pontos usados como teste também estão com erros compatíveis com aqueles dos pontos do conjunto de treino. Nesta variável observa-se que vários pontos caem fora da faixa de 1% de desvio, e poucos estão fora da faixa de 5% de desvio. Como esta variável não é importante para a classificação, que é o nosso objetivo, estes erros são aceitáveis.

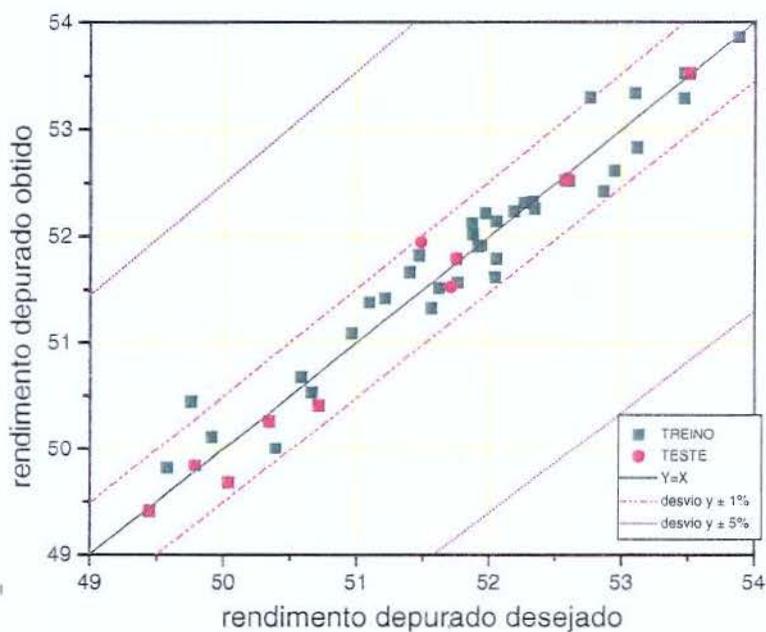


Figura 24 – Comparação entre o valor obtido pela rede neuronal e o valor desejado para o rendimento depurado nos conjuntos de treino e de teste. Indica-se os desvios de 1% e 5% no valor da ordenada da linha $y=x$.

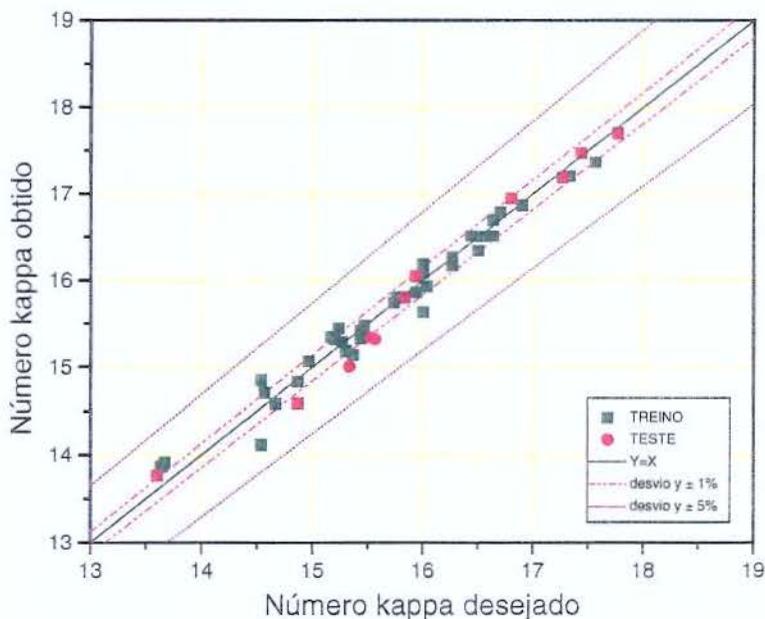


Figura 25 – Comparação entre o valor obtido pela rede neuronal e o valor desejado para o número kappa nos conjuntos de treino e de teste. Indica-se os desvios de 1% e 5%.

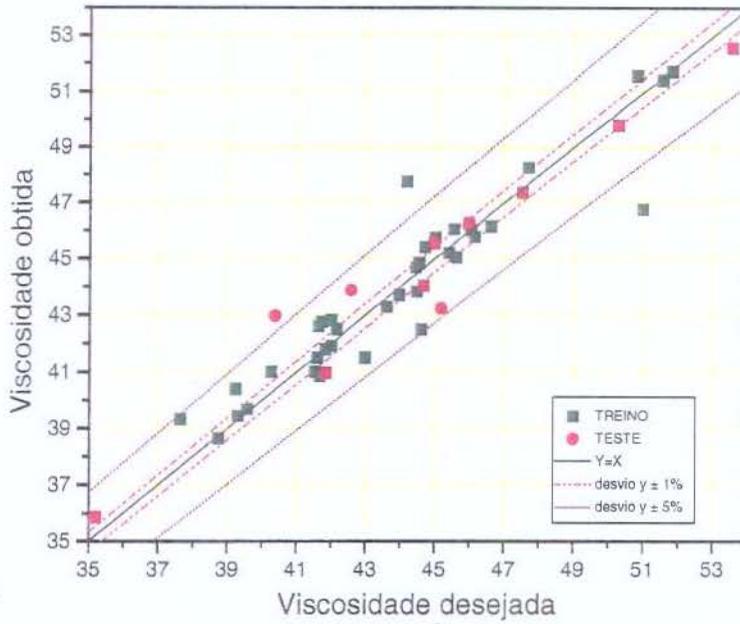


Figura 26 – Comparação entre o valor obtido pela rede neuronal e o valor desejado para a viscosidade nos conjuntos de treino e de teste. Indica-se os desvios de 1% e 5% no valor da ordenada da linha $y=x$.

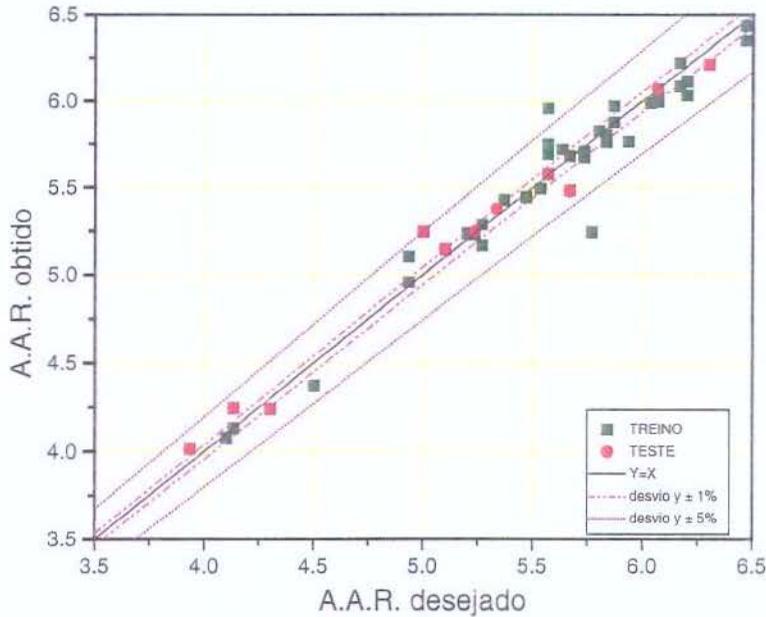


Figura 27 – Comparação entre o valor obtido pela rede neuronal e o valor desejado para o álcali ativo residual nos conjuntos de treino e de teste. Indica-se os desvios de 1% e 5%.

A Figura 27 demonstra o erro na predição da propriedade álcali ativo residual. Observa-se que os pontos usados como teste também estão com erros compatíveis com aqueles dos pontos do conjunto de treino. Nesta variável alguns pontos caem fora da faixa de 1% de desvio, mas quase todos, com exceção de dois, estão dentro da faixa de 5% de desvio.

Tabela 9- Estatística dos desvios dos resultados apresentados pela rede neuronal

	Resíduo do rendimento depurado	Resíduo do número kappa	Resíduo da viscosidade	Resíduo da alcalinidade residual
Máximo	0.60778	0.99107	2.6448	0.49944
Mínimo	-0.72538	-0.93003	-3.91161	-0.47067
Média	0.00137	-0.00418	0.00416	0.00153
Variância	0.06738	0.11652	1.8434	0.04637
D. Padrão	0.25959	0.34135	1.35772	0.21533

Comparando as variâncias dos erros, mostrados acima na Tabela 9 com as variâncias de amostragem e medida das variáveis, que foram calculadas anteriormente, pode-se dizer que os resultados apresentados pela rede são satisfatórios. Depurar ainda mais o treinamento não teria nenhuma vantagem a partir deste ponto, já que o erro gerado na predição da rede já é inferior aos erros naturais associados à amostragem, medida e leitura das propriedades avaliadas.

3.6.5. Classificação

Com a questão da predição das propriedades resolvida, a classificação dos lotes de madeira pode agora ser mais abrangente, envolvendo outros critérios, com outras variáveis importantes. O procedimento adotado pela Ripasa foi de utilizar-se da experiência profissional de dois especialistas no assunto, que de posse das informações do banco de dados, atribuíram um conceito a cada um dos 165 casos estudados. Os conceitos atribuídos vão de A a D, identificando, respectivamente, as madeiras de melhor a pior desempenho no

processo. As variáveis que foram utilizadas como base para a classificação variaram entre os dois especialistas. O especialista A baseou-se na densidade básica, no número kappa e no rendimento depurado; enquanto que o especialista B utilizou-se da densidade básica, do número kappa, do rendimento depurado, da quantidade de holocelulose e do álcali ativo residual. O especialista B colocou conceito para cada uma das 165 amostras, mas o especialista A o fez para cada um dos 55 talhões. Para padronizar o tratamento, posteriormente, o especialista A esclareceu seus critérios a ponto de que se pudesse completar a classificação para todas as amostras.

Com as duas classificações foi possível treinar redes neuronais do tipo LVQ (“Learning Vector Quantization”) usando como entradas as variáveis base dos especialistas, e a saída como o conceito arbitrado. As Figuras 28 e 29 demonstram a tela no instante onde se encontrou a melhor rede. Foram treinadas redes com diversos números de neurônios na camada chamada de “Kohonen”, e as de melhor resultado foram escolhidas.

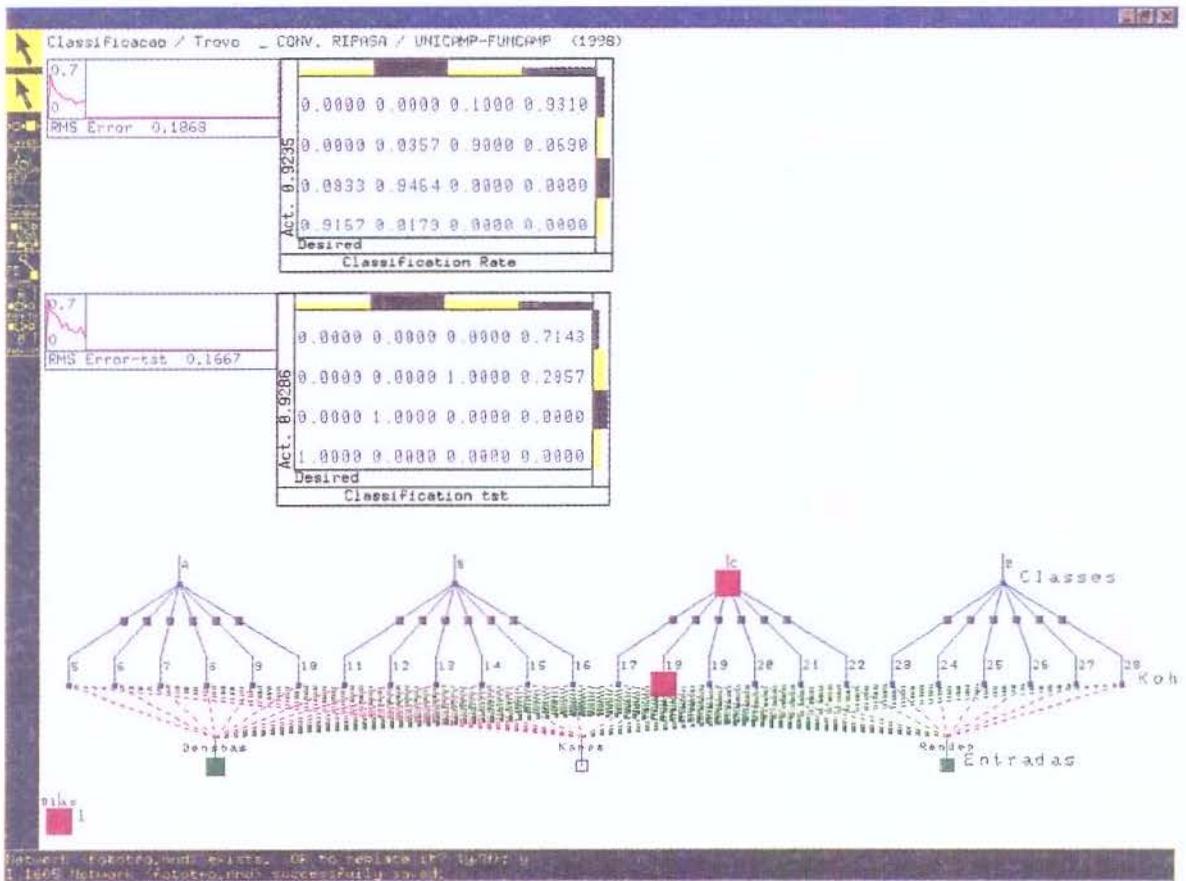


Figura 28 – Tela do programa de treinamento da rede LVQ para a classificação segundo o especialista A.

Para a classificação do especialista A (Figura 28), com 24 nós na camada "Kohonen", encontrou-se um erro RMS para o conjunto de treino de 0,1868, e para o conjunto de teste de 0,1667. O índice de correlação também é mostrado e igual a 0,9235 para o conjunto de treino e 0,9286 para o de teste. Os números que aparecem na matriz correspondem a fração de cada classe que apresentam resultado obtido igual ao desejado (diagonal principal) e as que o resultado obtido não confere com o desejado (diagonais secundárias). Assim, por exemplo, na matriz do conjunto de teste observa-se que na maioria dos casos, a classificação obtida confere com a desejada, exceto em 28,6% dos casos desejados da classe "D", onde a rede fornece a classe "C".

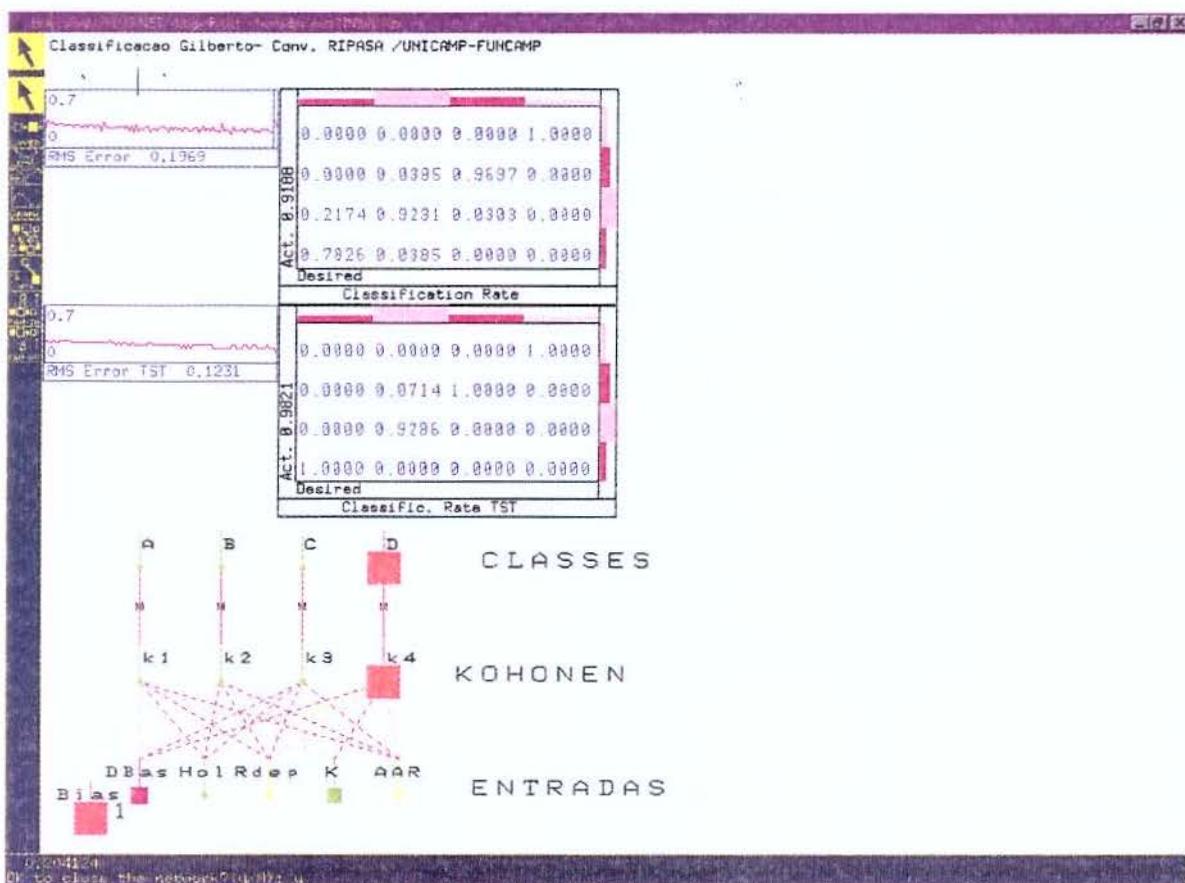


Figura 29 – Tela do programa de treinamento da rede LVQ para a classificação segundo o especialista B.

Para a classificação do especialista B (Figura 29), com apenas 4 nós na camada Kohonen, encontrou-se um erro RMS para o conjunto de treino de 0,1969, e para o conjunto de teste de 0,1231. O índice de correlação é igual a 0,9180 para o conjunto de treino e 0,9821 para o de teste.

Uma terceira classificação é proposta neste trabalho, e envolve os critérios inicialmente utilizados pelos especialistas, mas utiliza-se de um sistema de escores ao invés de uma rede neuronal para determinar a classe final do lote. Não se pretende desmerecer a avaliação dos especialistas, mas apenas criar um critério fixo e bem conhecido para avaliar a classificação. O sistema adotado é o de acumular o escore de acordo com cada uma das variáveis, e de acordo com a sua somatória dar o conceito final. A tabela abaixo demonstra mais claramente o peso diferenciado adotado entre as variáveis e o padrão adotado. Recorde-se que tais padrões podem ser facilmente ajustados conforme se observem modificações mais rígidas na classificação.

Tabela 10 – Valores dos escores adotados na Classificação proposta neste trabalho, para cada variável e faixa de valores.

Variável	Classe	A	B	C	D
Densidade Básica	Faixa	0,4 – 0,46	0,461 – 0,48	0,481 – 0,5	> 0,5
	Escore	250	170	80	0
Número Kappa	Faixa	≤ 15,0	15,1 – 16,0	16,1 – 17,0	> 17,0
	Escore	250	170	80	0
Rendimento Depurado	Faixa	> 52,0	51,1 – 52,0	50,0 – 51,0	< 50,0
	Escore	200	130	70	0
Álcali Ativo Residual	Faixa	≥ 6,3	5,5 – 6,29	4,5 – 5,4	≤ 4,5
	Escore	150	100	50	0
Holocelulose	Faixa	≥ 70,51	68,81 – 70,50	67,12 – 68,80	≤ 67,11
	Escore	150	100	50	0
FINAL	ESCORE	850 – 1000	610 – 840	250 - 600	0 – 240

Os resultados completos em forma de tabela estão no anexo (pág. 151). A decisão final de qual a classificação que será empregada caberá aos engenheiros da Ripasa.

3.7. PROGRAMAS APLICATIVOS GERADOS

Foram criados 4 programas aplicativos auxiliares, com o objetivo de utilizar-se as redes treinadas pelo Neuralworks. O primeiro deles, chamado de PREDPROP, é exclusivo para se fazer a predição das propriedades após a polpação, e está ilustrado na Figura 30. Como dados de entrada, delimitados entre si por tabulações no arquivo COMPLETO.TXT, deve-se introduzir :

Parque florestal / Talhão / Espécie / Idade (meses) / Produtividade / Densidade Básica ... / ...Extrativos / Lignina solúvel / Lignina insolúvel / Holocelulose ;

Retornando, no arquivo de saída PREDPROP.TXT, as seguintes informações:

Talhão / Rendimento depurado / Número Kappa / Viscosidade / Álcali ativo residual.

Exemplo de arquivo de entrada:

!# Arquivo de dados para DATASCULPTOR - Arquivo: Completo.dat Descriptor: Complet.ddd										
!#										
!# Fazenda	Talhão	Espécie	Idade	Produtividade	Db	ET	LS	LI	H	
flecha azul	144	E. hibrido	84.0000	55.7	0.5360	5.2767	1.9533	24.0467	68.7233	
flecha azul	139	E. grandis	80.5000	62.6	0.4533	6.3500	1.7933	24.3767	67.4800	
flecha azul	134	E. grandis	84.0000	47.3	0.4917	4.8533	1.5433	24.9433	68.6600	

Exemplo do arquivo de saída:

"flecha azul"	144	"E. hibrido"	84	55.7	0.536	5.277	1.953	24.05	68.72	49.85	17.48	47.36	5.147
"flecha azul"	139	"E. grandis"	80.5	62.6	0.4533	6.35	1.793	24.38	67.48	51.42	15.75	44.82	5.764
"flecha azul"	134	"E. grandis"	84	47.3	0.4917	4.853	1.543	24.94	68.66	51.8	16.05	44.04	6.215
"flecha azul"	267	"E. grandis"	75	61.5	0.4917	4.1	1.903	25.22	68.78	51.92	16.17	41.82	6.086

O segundo programa, chamado de PREVDADO, ilustrado na Figura 31, destina-se a fornecer a classificação dos lotes sendo dadas todas as informações. Como dados de entrada (delimitados entre si por tabulações no arquivo CLASSIMP.TXT) deve-se introduzir :

Talhão / Densidade Básica / Holocelulose / Rendimento depurado / Número Kappa / Álcali ativo residual ;

retornando no arquivo de saída PREVDADO.TXT:

Talhão / Classificação REDE especialista A / Classificação REDE especialista B /
Classificação Proposta

Exemplo de arquivo de entrada:

```

!# Arquivo de dados para Previsao das Classes c/ dados da massa
!#      Arquivo: clasimp.txt      Descriptor: clasimp.ddd
!# Talhão Db      H      RD      NK      A.A.R.
144      0.5360  68.7233  49.7833  17.4333  5.1000
139      0.4533  67.4800  51.2133  15.7333  5.9333
134      0.4917  68.6600  51.7500  15.9333  6.3000
  
```

Exemplo do arquivo de saída:

```

144      "D"      "D"      "D"
139      "B"      "B"      "B"
134      "C"      "B"      "B"
  
```

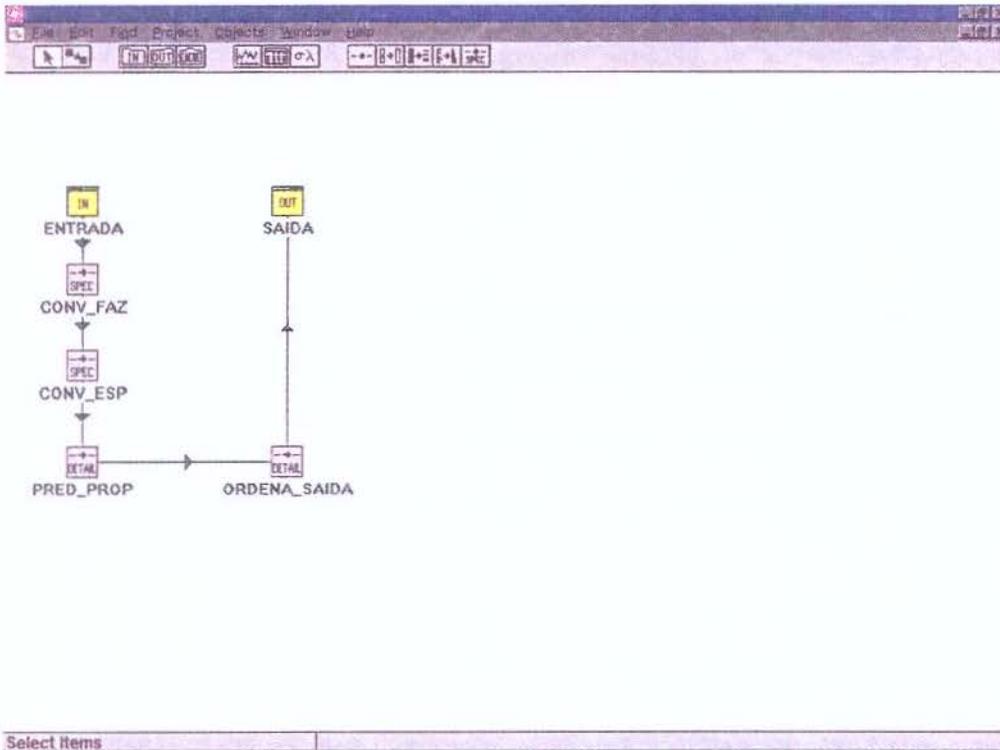


Figura 30 – PREDPROP programa para a predição das propriedades após a polpação

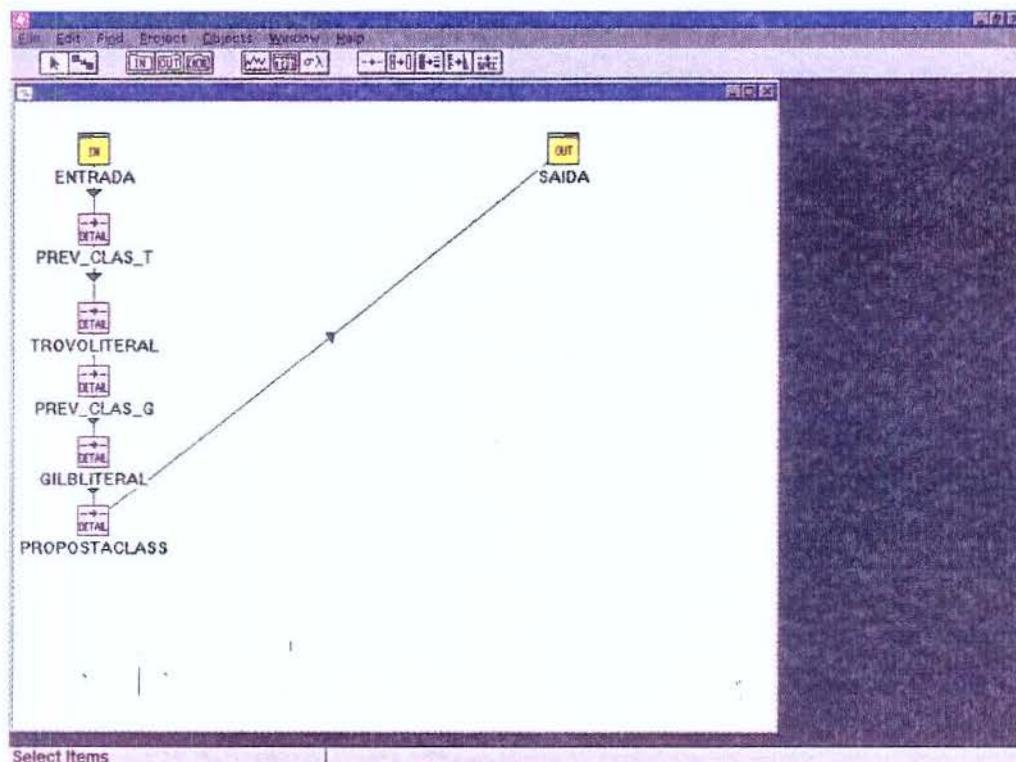


Figura 31 – PREVDADO programa para fornecer a classificação dos lotes, sem usar a predição.

O terceiro programa, chamado de PREVPRED, destina-se a fornecer a classificação dos lotes através da predição das propriedades após a polpação. Como dados de entrada (delimitados entre si por tabulações no arquivo COMPLETO.TXT) deve-se introduzir :

Parque florestal / Talhão / Espécie / Idade (meses) / Produtividade / Densidade Básica ... / ...Extrativos / Lignina solúvel / Lignina insolúvel / Holocelulose ;

retornando no arquivo de saída PREVPRED.TXT:

Talhão / Classificação REDE especialista A / Classificação REDE especialista B / Classificação Proposta

Exemplo de arquivo de entrada:

!# Arquivo de dados para Datasculptor/Análise de Resultados ANTOTAL - VALORES MEDIOS

!# Arquivo: MEDIOST.TXT Descriptor: MEDIOST.DDD

!# Fazenda	Talhão	Espécie	Idade	Prod	Db	ET	LS	LI	H	RD	NK	V	A.A.R.	TrGil
flecha azul	144	E. hibrido	84.0000	55.7	0.5360	5.2767	1.9533	24.0467	68.7233	49.7833	17.4333	47.5000	5.1000	D D
flecha azul	139	E. grandis	80.5000	62.6	0.4533	6.3500	1.7933	24.3767	67.4800	51.2133	15.7333	44.5333	5.9333	B B
flecha azul	134	E. grandis	84.0000	47.3	0.4917	4.8533	1.5433	24.9433	68.6600	51.7500	15.9333	44.6667	6.3000	C B

Exemplo do arquivo de saída:

144	"D"	"D"	"D"
139	"C"	"C"	"C"
134	"C"	"B"	"C"

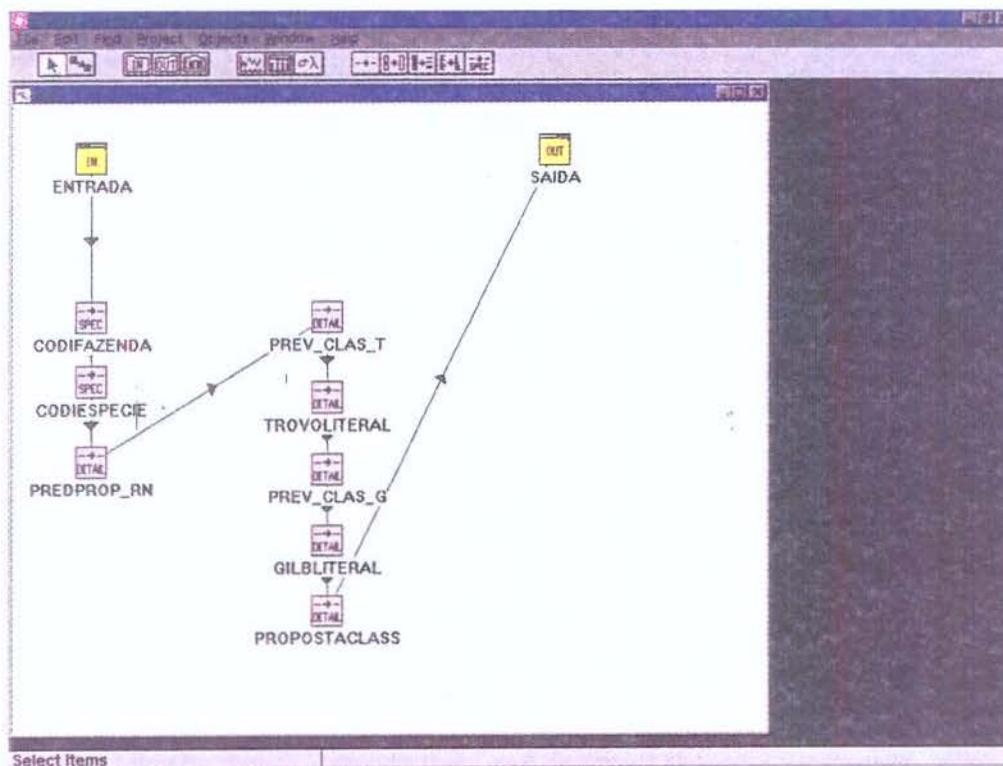


Figura 32 – ANTOTAL programa que gera um arquivo com todas as informações disponíveis sobre os lotes de matéria-prima, usando a predição das propriedades.

O último programa, chamado de ANTOTAL, destina-se a gerar um arquivo com todas as informações disponíveis sobre os lotes de matéria-prima. O seu fluxograma está ilustrado na Figura 32. A classificação dos lotes é realizada pela predição das propriedades após a polpação. Como dados de entrada (delimitados entre si por tabulações no arquivo COMPLETO.TXT) deve-se introduzir :

Parque florestal / Talhão / Espécie / Idade (meses) / Produtividade / Densidade Básica ... / ...Extrativos / Lignina solúvel / Lignina insolúvel / Holocelulose / Rendimento depurado / Número kappa / Viscosidade / Álcali ativo residual / Classificação por especialista A / Classificação por especialista B;

retornando no arquivo de saída ANTOTAL.TXT:

Parque florestal / Talhão / Espécie / Idade (meses) / Produtividade / Densidade Básica ... / ...Extrativos / Lignina solúvel / Lignina insolúvel / Holocelulose / Rendimento depurado / Número kappa / Viscosidade / Álcali ativo residual / Classificação por Especialista A / Classificação por Especialista B / Rendimento depurado (predito) / Número kappa (predito) / Viscosidade (predito) / Álcali ativo residual (predito) / Classificação REDE Especialista A / Classificação REDE Especialista B / Classificação Proposta

As figuras acima mostram a ordem das operações realizadas em cada um dos programas aplicativos gerados. A figura correspondente ao programa PREVPRED não é mostrada porque é idêntica à do programa ANTOTAL. O que diferencia um programa do outro, são as configurações dos arquivos de entrada e saída apenas.

As Tabelas 29 e 30, que constam no anexo (pag. 151 e 153), mostram os resultados gerados pelas redes treinadas. A primeira é obtida com os valores médios para talhão, e que por isso é a que tem a credibilidade compatível com os procedimentos adotados. A segunda é apenas um teste para avaliar as redes treinadas, onde se usa, ao invés dos dados médios, os dados de cada amostra coletada. Os resultados, principalmente no que tange a predição das propriedades devem ser olhados com a devida atenção ao detalhe que a rede não foi treinada para estas condições.

3.8. ANÁLISE DE SENSIBILIDADE DAS VARIÁVEIS DE ENTRADA

Para realizar este estudo, considera-se que o processo de treinamento da rede foi suficientemente realizado a ponto de se considerar que a rede pode dar bons resultados, mesmo para condições não utilizadas no treinamento. Pretende-se utilizar o Planejamento de experimentos e a análise da variância correspondente, para se ter uma idéia da influência de cada fator, tentando identificar se existiria algum deles que possa ser eliminado das análises por não exercer influência nos resultados.

A função da rede neuronal seria de prever algumas propriedades da matéria-prima frente a condições padrão de tratamento (polpação em laboratório). Este ensaio demora muito tempo e exige vigilância constante, o que justifica a aplicação do método. As propriedades que devem ser previstas são:

- rendimento depurado (RENDEP)
- índice Kappa (KAPPA)
- a viscosidade (VISCO)
- a alcalinidade residual (ALCRES)

As variáveis que se pretende usar para prever estas quatro, por vezes chamadas de fatores, são as seguintes:

- a idade de corte (IDADE)
- a produtividade obtida no talhão (PROD)
- parque florestal de origem da madeira (FAZ)
- a espécie da árvore (ESP)
- a densidade básica da amostra coletada (DENSB)
- a identificação do talhão de origem (TALHAO)
- nível de extrativos da amostra (EXTRAT)
- a quantidade relativa de lignina solúvel (LIGSOL)
- a quantidade relativa de lignina insolúvel (LIGINS)
- a quantidade relativa de holocelulose (HOLO)

Esta análise foi realizada antes da depuração do treinamento (vide item 3.6.4), assim a rede treinada escolhida foi a que apresentou o menor erro médio quadrático (RMS) correspondente a 0,0095, para todos os pontos e para todas as variáveis de saída. Os índices de correlação entre os valores desejados e obtidos que foram apresentados pelo instrumento específico do NEURALWORKS, chamado de “confidence matrix”, são os seguintes:

$$R(\text{RENDEP}) = 0.9977$$

$$R(\text{KAPPA}) = 0.9993$$

$$R(\text{VISCO}) = 0.9985$$

$$R(\text{ALCRES}) = 0.9952$$

Tais valores evidenciam um alto grau de ajuste entre o valor desejado e os preditos. Lembrando que aqui não avaliou-se o erro do conjunto de teste, sendo todos os pontos pertencentes ao conjunto de treinamento. Deseja-se, contudo, saber se alguma destas variáveis poderia ser desconsiderada, em virtude da baixa influência que exerceria sobre o resultado. Para isso, aplicou-se o planejamento fatorial de experimentos (BARROS NETO et al., 1996 e MONTGOMERY, 1991), utilizando-se da rede treinada.

Um dos procedimentos fundamentais para o sucesso da análise refere-se a escolha dos fatores a analisar e os respectivos níveis de variação. Quanto aos fatores, neste caso, procurou-se avaliar todos eles. Talvez, com exceção da variável talhão, onde existe uma certa dificuldade de estabelecer-se um padrão de identificação adequado para o treinamento da rede. Neste estudo, considerou-se que as variáveis fazenda e espécie definem o talhão, ou seja existe uma dependência desta variável às outras duas. Isto impede a aplicação direta do planejamento fatorial. Em certo ponto, este problema pode ser encarado como um planejamento em blocos. Tal análise foi feita, mas os seus resultados não foram diferentes de quando a variável talhão foi simplesmente ignorada. Para avaliar a influência do talhão no processo utiliza-se uma análise especial que será detalhada adiante.

3.8.1. Análise paramétrica - níveis estipulados em 10% abaixo e acima do valor médio

Como já fora mencionado, a decisão a respeito dos níveis a serem testados é de grande influência no resultado final da análise. Assim, procedeu-se a escolha em função da análise das variáveis, e depois pela formulação de dois padrões diferentes.

Avaliando as variáveis, percebe-se que estas podem ser classificadas em variáveis discretas e em contínuas. Os fatores que tem valores somente inteiros e dentro de faixa bem definida, que são as variáveis discretas, são:

A **idade** de corte, que assume somente dois valores, 6 ou 7 anos. Neste caso, a coleta de dados arredondou os meses para enquadrar a idade em números inteiros. Inicialmente, essa poderia ser a primeira preocupação da análise a ser realizada. Ou seja, se for constatado grande influência desta variável dentro deste intervalo, fica caracterizada a necessidade de aumentar a resolução da variável, colocando-a talvez em meses e não mais em anos.

A **fazenda** de origem, que no caso do banco de dados utilizados pode ser identificada por 10 valores diferentes. Como o número de amostras por fazenda não é homogêneo, ocorrendo casos de apenas uma amostra como casos de várias amostras da mesma fazenda, há de se tomar um cuidado especial. Assim, pesquisou-se para escolher as fazendas onde se tivesse o maior número de ocorrências. As fazendas escolhidas foram a “ibiti” e “guarujá”.

A **espécie** da madeira plantada. Da mesma forma que a fazenda, são cinco tipos de espécies envolvidas no estudo. A escolha das duas espécies para estudo também seguiu o mesmo raciocínio que para a fazenda. Na verdade, procurou-se a combinação onde se tinha o maior número de casos. Assim as espécies escolhidas foram a *E. grandis* e a *E. saligna*.

O **talhão** de origem dentro da fazenda. Como existe uma dependência muito forte deste fator com os dois anteriores, decidiu-se aleatorizar um talhão escolhido entre os possíveis. Ou seja, por exemplo, na fazenda guarujá, tem-se amostras de madeira de *E. saligna* dos talhões numero 177, 179 e 283. Por sorteio, o talhão escolhido foi o numero 177. Resumindo em uma tabela, temos os seguintes talhões.

Tabela 11 - Fazendas e espécies selecionadas para os experimentos, e o talhão correspondente

Fazenda	Espécie	Talhão sorteado
Guarujá	<i>E. saligna</i>	177
Guarujá	<i>E. grandis</i>	211
Ibiti	<i>E. saligna</i>	308
Ibiti	<i>E. grandis</i>	222

Com relação as outras variáveis que são contínuas, adotou-se como critério inicialmente o valor da média descontado e acrescido em 10 %, para se ter os níveis inferior e superior respectivamente. A tabela abaixo mostra os valores mínimo e máximo encontrados e

a média dos valores de cada variável do banco utilizado para o treinamento. O nível inferior (nível -), corresponde a 90% do valor da média. O nível superior (nível +) corresponde a 110% do valor da média. Os valores mínimo e máximo são apresentados a fim de detectar se o nível de variação (no caso 10%) sobrepuja os limites utilizados no treinamento da rede. Neste caso, aparece na variável lignina insolúvel e em holocelulose uma extrapolação da faixa, que pode ser considerada suficientemente pequena.

Tabela 12 - Níveis inferior e superior dos fatores a testar

Fatores	Mínimo	Máximo	Média	Nível (-)	Nível (+)
Produtividade	22,4	96,1	55,193	49,6737	60,7123
Densidade Básica	0,39	0,573	0,483	0,4347	0,5313
Extrativos	1,65	8,2	4,504	4,0536	4,9544
Lignina solúvel	0,93	2,78	1,72	1,548	1,892
Lignina insolúvel	22,8	27,89	24,975	22,4775	27,4725
Holocelulose	63,4	72,25	68,801	61,9209	75,6811

3.8.1.1. Planejamento dos experimentos

Definidos os fatores, os níveis a utilizar, o modelo e as respostas, executa-se o planejamento dos experimentos a serem realizados com auxílio do programa MINITAB. A primeira providência é criar o conjunto das possibilidades do teste. Como são 9 fatores estudados a dois níveis, teríamos 2^9 experimentos para se ter o fatorial completo. O programa MINITAB não prevê a possibilidade da execução de tal número de experimentos, aceitando como 128 o número máximo de experimentos. Assim, o que se realizará será um experimento fatorial fracionado, conhecido como 2^{9-2} .

Lembre-se que este experimento tem resolução VI, o que quer dizer que podemos avaliar algumas das associações entre interações de terceiro grau apenas, tendo-se todos os efeitos principais e secundários.

3.8.1.2. *Resultados e análises*

Com o recurso da utilização de um programa auxiliar que faz parte do pacote do software adquirido da Neuralware pelo Laboratório de Modelagem e Simulação de processos Químicos, chamado de DATASCULPTOR, a execução do experimento proposto foi bastante facilitada. Existe uma integração entre o software que cria e treina a rede neuronal e este outro que ajuda a colocar os dados na forma desejada. Depois de treinada a rede neuronal, o NEURALWORKS PRO II pode criar uma subrotina em C que representa a rede neuronal treinada. O programa DATASCULPTOR tem uma ferramenta que interpreta essa subrotina e fornece os valores de saída da rede para qualquer conjunto de entrada estipulado. Assim, depois de importados os arquivos com os valores dos fatores (que são as entradas da rede) e a subrotina da rede neuronal treinada para o DATASCULPTOR, o arquivo gerado com as saídas correspondentes é transportado novamente para a planilha do MINITAB. A análise do experimento será realizada por este programa estatístico. Alguns dos resultados desta análise são mostrados, e se percebe que parte da informação emitida pelo programa tem sentido diferente daquele usualmente empregado. Lembre-se de que o modelo utilizado não dá a variabilidade esperada num processo real. Portanto o resultado para um conjunto de fatores vai ser sempre único, o que não é observado nos processos reais. Assim, o que aqui será admitido como resíduo, na verdade é a conjugação de efeitos considerados pequenos por serem de elevada ordem. São os efeitos de interação entre 3, 4 ou mais fatores, que não são incluídos expressamente no modelo de ajuste do MINITAB. O p-valor que aparece então nas listagens de saída do programa refere-se usualmente a uma probabilidade de ocorrer o evento da hipótese nula. A hipótese testada é a de que todos os efeitos dos fatores são iguais e de valor nulo, em base estatística, o que atribuiria a causa das variabilidades nas respostas ao resíduo (ou ruído do sistema). Assim, quando um fator, ou interações entre eles, apresenta um efeito que se sobressalta ao sinal dos resíduos (ruído), há de se avaliar a sua variabilidade frente à dos resíduos. Daí surge o p-valor, que define a que nível de significância determinado efeito pode ser considerado diferente de zero (rejeitando a hipótese nula). Poderia ainda se dizer de outra forma: Considerando que a hipótese nula é verdadeira, a ocorrência de um determinado evento segue a probabilidade estipulada por p-valor. Assim, quando um efeito apresenta o valor 4, por exemplo, e o p-valor seja de 0.01, considerando que a hipótese de que na verdade este número seja zero, a probabilidade deste evento ocorrer seria de uma em 100 ocorrências. No caso, onde verificamos que tal interpretação não se aplica diretamente, qual seria então o

significado deste p-valor?. Seria um valor de comparação entre um efeito (principal, secundário ou ternário) e o conjunto de efeitos de maior grau, que sempre são pequenos o bastante para serem negligenciados. Um baixo valor do p-valor equivaleria a dizer que o fator apresenta um efeito que não pode ser confundido com os aqueles de maior grau (perfeitamente negligenciáveis).

3.8.1.3. *Análise dos fatores principais somente:*

O programa MINITAB fornece uma análise mostrando um relatório que contém os comandos utilizados, os efeitos estimados, a análise da variância, as ocorrências incomuns e o gráfico dos efeitos para cada uma das variáveis de saída. Neste caso incluiu-se todos os efeitos secundários e de maior grau como resíduo.

3.8.1.4. *Análise considerando as interações até de terceiro grau*

A inclusão das interações de até terceiro grau na análise altera (diminui) o p-valor para os termos já que a variabilidade do resíduo se torna menor relativamente, aumentando o valor da coluna de t-valor que determina a probabilidade.

Observa-se em alguns casos que tem-se alguns casos de interações de 3ª ordem com efeitos relevantes. Fato que indica que algo pode não estar coerente. Avaliando melhor estes casos pode-se dizer que a maioria destes casos apresenta efeitos de valores baixos, considerando o valor do desvio padrão para cada variável resposta. Isto sugere que a faixa de variação adotada para os fatores, ou simplesmente para algum deles, é muito pequena em relação aos outros fatores.

3.8.2. Análise Paramétrica - níveis estipulados por um desvio padrão abaixo e acima da média

A escolha dos níveis anteriormente feita apresenta problemas na análise, já que depende do valor absoluto da média e não existe correspondência entre os fatores. Assim, os efeitos apresentados não permitem definir quais os fatores mais importantes. Pensando em uma forma de poder analisar e comparar diretamente os valores dos efeitos, surgiram algumas idéias. Poderia ser adotado os valores mínimo e máximo encontrados, os valores do 1o. e 3o. quartis, ou ainda usar o desvio padrão para afastar os níveis da média. Essa

última pareceu ser aquela mais representativa da variabilidade normal de cada variável, o que sugere que os efeitos podem ser comparados entre si.

O cálculo dos desvios padrão foi feito com o banco de dados utilizados no treinamento da rede neuronal, com auxílio do programa DATASCULPTOR. A Tabela abaixo mostra a estatística dos fatores, bem como os níveis adotados e também a estatística das variáveis resposta.

Tabela 13 - Estatística Descritiva e níveis inferior e superior dos fatores.

	Produtividade	densidade básica	Extrativos	lignina solúvel	Libnina insol.	Holocelulose
Maximum	96.1	0.564	7.19	2.657	27.147	71.87
Minimum	22.4	0.4	2.657	0.963	23.207	65.037
Count	55	55	55	55	55	55
Mean	55.1927	0.4834	4.50395	1.71989	24.9748	68.8013
Variance	296.98	0.0014242	1.22323	0.137413	0.700618	2.40557
S.Deviation	17.2331	0.0377386	1.106	0.370692	0.837029	1.55099
Nível (-)	37.960	0.4456	3.398	0.9785	23.300	67.250
Nível (+)	72.426	0.5211	6.716	2.4613	26.649	70.352

Tabela 14 - Estatística Descritiva das variáveis resposta.

	Rendimento depurado	Índice KAPPA	viscosidade	alcalinidade residual	teor de rejeitos
Maximum	54.193	18.433	53.533	6.467	0.393
Minimum	49.443	13.6	35.2	3.4	0.0
Count	55	55	55	55	55
Mean	51.8424	15.743	44.2909	5.44545	0.0422546
Variance	1.55348	1.08063	14.935	0.515108	0.00303975
S.Deviation	1.24639	1.03953	3.86458	0.71771	0.055134

3.8.2.1. *Planejamento dos experimentos*

A mudança dos valores dos níveis não muda significativamente o planejamento dos experimentos. Assim, temos a mesma saída do programa obtida anteriormente. Logicamente, a diferença está no arquivo usado pelo DATASCULPTOR para processar o conjunto de experimentos.

Na análise dos resultados, com auxílio do programa MINITAB, estudou-se a variação na forma de estipular os termos do modelo que o programa ajusta. Isto permite salientar alguns fatos que serão explorados posteriormente. Foram realizadas análises com interações de até terceira ordem, com interações de até segunda ordem e só com os efeitos principais; e ainda bloqueando a variável talhão.

3.8.2.2. *Resultados e análises (interações até terceira ordem)*

A saída do programa MINITAB para a análise proposta está resumida abaixo:

Tabela 15 - Extrato da saída do programa MINITAB para análise do experimento incluindo interações de até terceira ordem.

Rendep		
Term	Effect	P
A=prod	0.1023	0.075
B=densb	-0.1411	0.017
C=extrat	-0.1280	0.028
D=ligsol	-0.6645	0.000
E=ligins	0.1267	0.030
F=holo	0.1748	0.004
G=idade	0.1733	0.004
H=faz	-0.5895	0.000
I=esp	1.5383	0.000
Kappa		
Term	Effect	P
A=prod	-0.4170	0.000
B=densb	0.6527	0.000
C=extrat	0.0461	0.559
D=ligsol	0.9752	0.000
E=ligins	0.5533	0.000
F=holo	-0.1077	0.179
G=idade	-0.7073	0.000
H=faz	0.1567	0.054
I=esp	0.2320	0.006

	Visco	
Term	Effect	P
A=prod	-2.135	0.
B=densb	-3.930	0.
C=extrat	-2.890	0.
D=ligsol	-1.502	0.
E=ligins	1.792	0.
F=holo	0.610	0.
G=idade	2.572	0.
H=faz	-3.178	0.
I=esp	-1.256	0.

As tabelas acima exibem o efeito dos fatores principais e o p-valor correspondente.

3.8.2.3. Outras análises

Considerando somente os efeitos de até segundo grau, as interações entre 3 fatores ou mais são consideradas como resíduo.

A análise para somente os efeitos principais foi realizado no MINITAB apenas para melhorar a visualização desses efeitos, lembrando que neste caso todas as interações são admitidas como ruído.

A variável talhão, em última análise, se caracteriza como um bloco. Os resultados da análise foram obtidos pelo MINITAB bloqueando na variável talhão. Note-se que desta forma as variáveis fazenda e espécie não mais entram como fatores, já que ao considerar talhão como bloco estas perdem o sentido na análise.

3.8.3. Conclusões específicas deste estudo

Avaliando então as análises acima, com o intuito de identificar se alguma das variáveis pode ser desconsiderada para o treinamento da rede, chega-se ao quadro resumo abaixo, onde mostra-se as variáveis com pouca influência na resposta da rede para cada uma das saídas.

Tabela 16 - Quadro resumo com os fatores de menor importância. Os locais assinalados significam que o fator não tem muita influência sobre aquela variável saída da rede.

Fator	Saídas da rede neuronal			
	Rend. Depurado	Kappa	Viscosidade	Alcal. Resid.
Prod	☒			
Densb	☒			
Extrat	☒	☒		
Ligsol				☒
Ligins	☒			
Holo	☒	☒	☒	
Idade	☒			
Faz		☒		☒
Esp		☒	☒	

Ao observar o quadro acima, e lembrando que a viscosidade é a variável de saída de menor importância para o estudo da classificação da madeira, pode-se chegar à conclusão de que o fator holocelulose poderia ser retirado do treinamento da rede, já que tem alguma influência apenas na alcalinidade residual.

Ainda, a quantidade de extrativos têm influência média apenas na viscosidade, podendo ser excluída, dependendo da conveniência (econômica ou operacional), mas se prever a alcalinidade residual for importante, deve-se mantê-la pois aí exerce grande influência.

A definição sobre a espécie é de extrema importância para a determinação do rendimento depurado, razão pela qual não é recomendável retirá-la do modelo.

O quadro resumo onde se apresentam as maiores influências está exposto na Tabela 17.

Para efeito de decisão, com relação ao grau de importância no resultado alcançado, baseando-se nas tabelas acima e nos valores dos efeitos, pode-se estabelecer uma escala da importância das variáveis de entrada. Assim, classificam-se como:

Fatores de elevada importância:

Quantidade de lignina solúvel

Densidade básica

Fazenda

Espécie

Tabela 17 - Quadro resumo com os fatores de maior importância. Os locais assinalados significam que o fator tem muita influência sobre aquela variável saída da rede.

Fator	Saídas da rede neuronal			
	Rend. Depurado	Kappa	Viscosidade	Alcal. Resid.
Prod		<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>
Densb		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
Extrat				<input checked="" type="checkbox"/>
Ligsol	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		
Ligins		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
Holo				
Idade		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Faz	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	
Esp	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>

Fatores de média importância:

Idade

Lignina insolúvel

Extrativos

Fatores de baixa importância:

Produtividade

Quantidade de holocelulose

3.9. CONCLUSÃO SOBRE O ESTUDO DA CLASSIFICAÇÃO DA MATÉRIA-PRIMA PARA A POLPAÇÃO DA MADEIRA

Neste capítulo apresentou-se os principais procedimentos adotados para se conseguir um resultado satisfatório na predição das propriedades da matéria-prima após a sua polpação com a finalidade de estipular uma classificação mais abrangente do que a atual. Primeiro, apresenta-se uma análise estatística dos dados. Depois mostram-se os primeiros resultados obtidos, e a avaliação de algumas características testadas. Ainda, demonstra-se o uso dos programas aplicativos gerados. Por fim, uma análise de sensibilidade indica as variáveis mais importantes para a predição das propriedades.

Algumas redes foram treinadas para várias condições e configurações diferentes, onde algumas dessas apresentam um ótimo resultado, mostrando o bom ajuste alcançado. Isto viabiliza, portanto, a classificação da matéria-prima em função das propriedades previstas.

CAPÍTULO 4. PREDIÇÃO DE COMPORTAMENTO DO SISTEMA DE TRATAMENTO DE EFLUENTES USANDO A TECNOLOGIA DE REDES NEURONAIS

4.1. CARACTERIZAÇÃO DO PROBLEMA A SER ESTUDADO

Neste projeto propõe-se a utilização de técnicas de simulação com auxílio de redes neurais artificiais com a finalidade de fornecer subsídios para se proceder um melhor controle sobre o processo de tratamento de efluentes da unidade da Rhodiaco Ind. Químicas Ltda. de Paulínia.

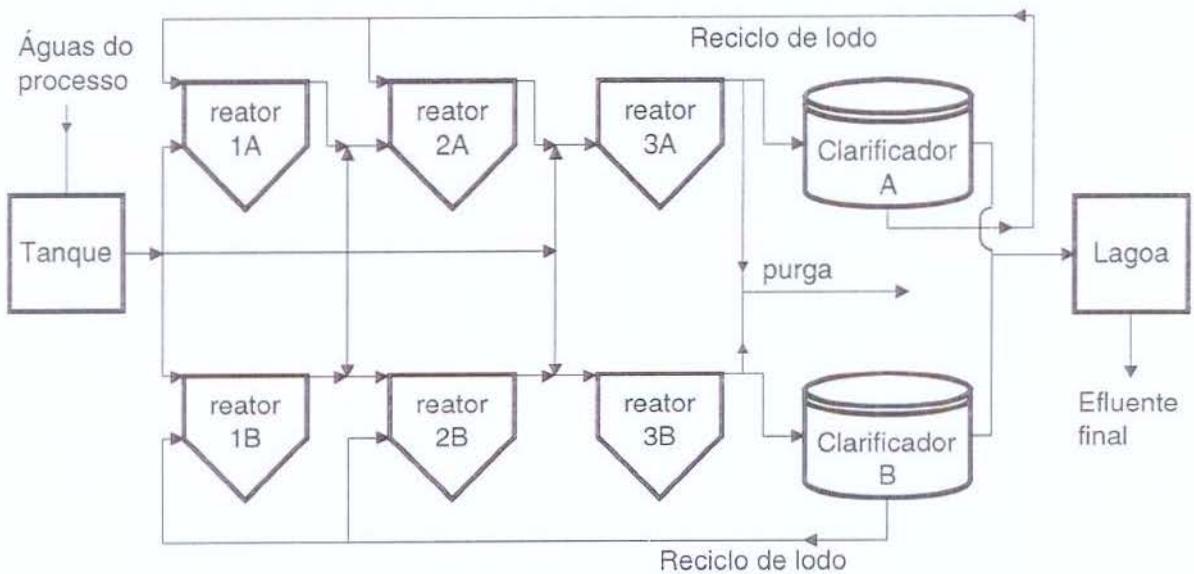


Figura 33– Fluxograma simplificado da unidade de tratamento de efluentes

O sistema é composto basicamente de dois conjuntos de reatores de aerção que funcionam paralelamente, e ainda os seus clarificadores. São três reatores em série, que recebem também fluxos do efluente que vêm do tanque reservatório e parte do lodo que recicla. Os clarificadores separam o lodo, que retorna ao processo, do efluente líquido que irá para uma lagoa de estabilização antes de ser despejado no receptor. Para controlar a idade do lodo, e manter a relação entre as vazões de lodo e de efluente dentro de limites operacionais estipulados, realiza-se a purga de sólidos retirando uma quantidade do efluente do último reator. Este passa por um espessador, sendo que a parte líquida segue para a

lagoa e os sólidos são levados a lugar apropriado de descarte. A Figura 34 mostra o fluxograma simplificado do sistema.

Os dados disponíveis envolvem informações provenientes tanto do laboratório como do Sistema Digital de Controle Distribuído (SDCD). Tem-se, disponíveis, os registros de Junho a Novembro de 1998 e de Março a Agosto de 1999. Lembrando que a idéia inicial sempre foi utilizar os dados existentes para o estudo, procurou-se encontrar o melhor resultado possível com os dados colocados à nossa disposição.

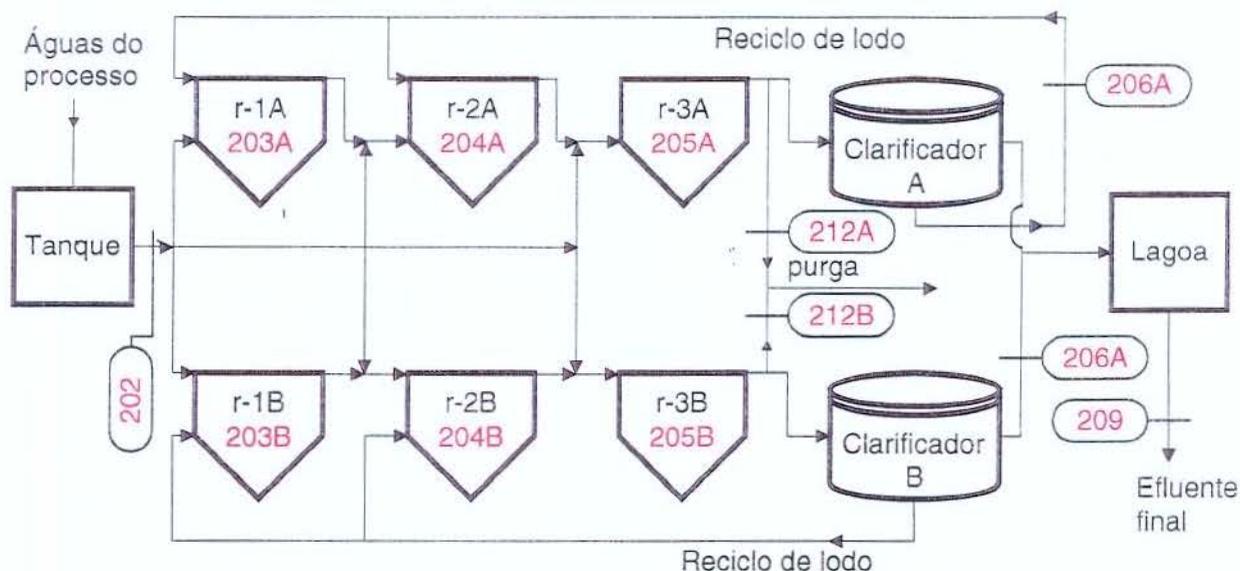


Figura 34- Fluxograma simplificado da unidade de tratamento de efluentes, constando o código das amostras coletadas.

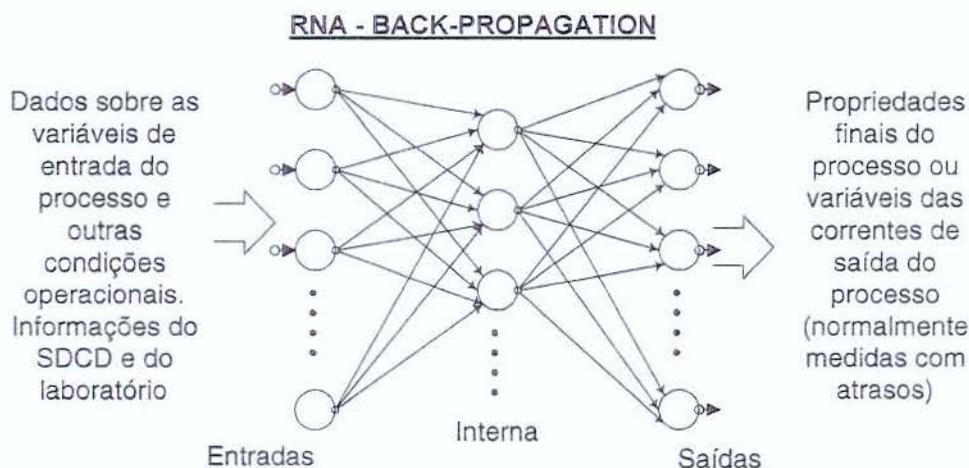


Figura 35 – Esquema simplificado de uma rede neuronal do tipo “back-propagation”, para o caso estudado.

4.2. OBJETIVOS DO ESTUDO REALIZADO

- Geral:

Indicar ações para melhorar o processo na unidade de tratamento de efluentes, aumentando o grau de controle sobre os despejos enviados ao receptor.

- Específicos:

Deseja-se iniciar a investigação sobre os fatores que influenciam a eficiência de remoção da carga poluidora (medida pela concentração de carbono orgânico total ou TOC) dos efluentes da unidade industrial. Assim, considerando dados obtidos pelo SDCD (sistema digital de controle distribuído) e do laboratório, pretende-se:

- obter, com auxílio da técnica de redes neurais, um novo instrumento para a investigação de melhores condições operacionais, visando a suficiente depuração da carga poluidora.
- Gerar valores preditos para as variáveis de processo não mensuráveis *on-line*, no caso os valores de TOC residual nos reatores e no descarte para o receptor.
- Implementar a técnica das redes neurais dentro da Rhodiaco, com a finalidade de fornecer subsídios para que os seus engenheiros de processo possam avaliar uma possível aplicação em outras unidades.

Observa-se que com o programa simulador criado será possível estudar diferentes formas de atuar no processo frente a uma situação a ser enfrentada. Contudo, deve-se respeitar rigorosamente os intervalos de validade para cada uma das variáveis, porque as redes criadas não extrapolam valores de maneira confiável.

4.3. TRATAMENTO INICIAL DOS DADOS

Com auxílio de tabelas e gráficos, procurou-se definir melhor a relação entre as variáveis existentes no banco de dados. Um dos principais resultados desta fase foi estabelecer um tempo de residência médio para relacionar os valores de entrada e saída da unidade. Outra decisão importante foi optar por simular cada reator separadamente. Isto devido às condições encontradas no banco de dados, onde não haveriam dados suficientes para se tentar treinar uma rede única para a unidade.

A partir de várias planilhas envolvendo os meses de Junho/98 a Agosto/98 foram criados os arquivos necessários ao treinamento das redes neuronais. Cada mês possuía dados em planilhas distintas para as informações do SDCD e do laboratório. Assim, procurou-se unir todos os meses e planilhas destes em uma só. Feito isto, foram eliminados todos os dados onde faltava alguma das informações importantes do laboratório. Com este grupo de dados criaram-se os arquivos respeitando as exigências do programa de treinamento de redes. Os problemas encontrados com relação à estrutura dos dados existentes referem-se ao fato de que:

- a) Os dados do laboratório são amostras simples e em horários ímpares. Já os dados do SDCD são valores médios por turno, onde os turnos terminam em horas pares. Admitiu-se que o sistema estaria em regime estacionário a ponto de não haver grandes diferenças durante este tempo.
- b) A periodicidade das amostras está voltada para uma adequação operacional. Assim, seria o correto ter as amostragens sincronizadas com os tempos de residência em cada equipamento, considerando portanto a vazão.
- c) Algumas análises são realizadas com uma periodicidade diferente das outras. Isto provocou a eliminação de diversos dados, pois na falta de uma única variável o ponto não pode ser utilizado.

As variáveis que inicialmente pretendia-se considerar na avaliação são as seguintes:

- 1) TOC do efluente de entrada em cada reator
- 2) Vazão do efluente na entrada de cada reator.
- 3) Concentração de sólidos (MES) do lodo recirculado
- 4) Concentração de sólidos (MES) observado em cada reator
- 5) Idade média do lodo utilizado.
- 6) Vazão de lodo que recircula.
- 7) Concentração de oxigênio (% O_2) dissolvido nos reatores.
- 8) pH de entrada e nos reatores.
- 9) TOC residual em cada reator e na saída da lagoa de polimento.
- 10) Vazão de purga do lodo.
- 11) Concentração de Manganês.

De acordo com as condições impostas para operação do sistema de tratamento de efluentes, tais como as vazões, TOC de entrada, taxa de reciclo do lodo, idade do lodo, entre outras, haverá em consequência, um determinado grau de eliminação do TOC. As avaliações sobre a condição final do despejo que são realizadas, atualmente, se baseiam em diversas variáveis que fornecem indícios do andamento do processo, e na análise final do despejo que vai ao receptor. Além do fato de que esta medida só é realizada no final da operação, quando já se passaram aproximadamente de 3 a 5 dias de quando o efluente entrou na unidade, não se dispõe de nenhum instrumento capaz de fornecer uma boa idéia de como estabelecer condições operacionais estabilizadas. Isto gera uma dificuldade para o controle do processo, já que a sua principal variável de avaliação demora muito a ser verificada. O que se propõe é, através de redes neurais devidamente treinadas, obter uma predição do TOC residual imediatamente em função das condições operacionais a serem utilizadas no momento. Isto permitirá um grau de atuação sobre o processo mais eficiente, aumentando possivelmente a eficiência da unidade como um todo.

Neste intuito, ficou decidido criar um simulador para a unidade usando redes neurais treinadas para cada um dos reatores. Assim, existe uma rede específica para o primeiro reator, e outras duas para os outros. Isto se explica pelo fato de que se ter um maior número de informações para o primeiro reator do que para os seguintes. Então foi considerado inicialmente o maior número possível de variáveis relativas ao primeiro reator de cada série. Surgiu então a questão sobre a medida da concentração de oxigênio nos reatores. Os medidores de oxigênio estiveram por um tempo desativados ou fornecendo valores não confiáveis. Se estes dados fossem eliminados, o número restante de pontos disponíveis para usar nos treinamentos seria muito reduzido. As redes foram treinadas então sem esta variável. Existem diversas formas para abordar o problema da diferenciação entre os primeiros reatores das séries. Inicialmente optou-se por usar os dados da série B no treinamento propriamente dito e os dados da série A no teste das redes treinadas, para avaliar a característica de generalização delas.

Assim, para o primeiro reator das duas plantas preparou-se arquivos contendo como variáveis de saída:

- a) o TOC da amostra 203, no dia seguinte.
- b) o pH dentro do reator no dia seguinte

E como variáveis de entrada:

- a) o pH da amostra 202
- b) o TOC da amostra 202
- c) o pH da amostra 203, no tempo corrente.
- d) o TOC da amostra 203, no tempo corrente. Note-se que na saída da rede o TOC seria do dia seguinte.
- e) a concentração de sólidos (MES) da amostra 203
- f) a concentração de sólidos (MES) da amostra 206
- g) a vazão de entrada de efluente no reator (FI-6302)
- h) a vazão de entrada de lodo (G-6413)
- i) a vazão da corrente de purga de sólidos da unidade (FI-6303). O intuito de incluir este dado é de indicar uma relação com a idade do lodo.

Nestas condições e eliminando os pontos onde faltasse alguma destas variáveis o banco de dados se reduziu para 58 pontos para a trancha B (treino) e 62 para a trancha A (teste). Lembrando que nesta ocasião o banco de dados incluía dados até outubro/98.

Para estudar o problema da predição da concentração de manganês foram realizadas várias reuniões com os engenheiros da indústria, para se definir modificações na periodicidade e propostas de criação de análises de laboratório que seriam necessárias. As informações à disposição não seriam suficientes para a predição da variável, principalmente porque só se fazia a análise na saída do sistema. Apesar do esforço de formar um banco de dados compatível para o estudo, o conjunto coletado de dados ainda era insuficiente para acumular as informações necessárias para treinar a rede. Como a análise do manganês não é realizada continuamente, parando nos fins de semana e feriados, isto acaba por diminuir a quantidade de dados quando se considera o tempo de retenção hidráulica no sistema.

Com os dados fornecidos foram selecionados apenas 12 pontos que continham as informações necessárias. O número de variáveis a serem utilizadas como as vazões, o pH, a concentração de sólidos (MES), e as concentrações de entrada e saída do manganês é maior do que o número de pontos. Além disto, foi informado que nem todos os dados deveriam ser usados porque havia sido feita uma modificação no tratamento para reduzir a

concentração no manganês. Com esta modificação, deveria ser treinada uma rede específica e assim, foi sugerido que se aguardasse a formação de um volume suficiente de dados, seguindo o cronograma de análises atual, ou ainda que se intensifique a quantidade de amostragens e análises para o manganês.

4.4. TREINAMENTO DAS ESTRUTURAS DE REDES NEURONAIS PARA PREDIÇÃO DO TOC FINAL DA UNIDADE

A medida que os dados foram sendo fornecidos, as redes foram treinadas para prever o TOC. A esta altura do estudo, novos dados (meses) já estavam disponíveis. Ao tentar treinar novas redes com estes, a qualidade do ajuste caiu sensivelmente. Avaliando melhor o conjunto de dados, identifica-se pelo comportamento do primeiro reator, que ocorreram duas fases distintas na operação do sistema de tratamento. Assim decidiu-se por trabalhar estas fases independentemente. O 1º conjunto refere-se a dados colhidos entre Junho/98 e Outubro/98; e o 2º conjunto refere-se a Novembro/98 e ainda ao período de Março/99 a Junho/99. Nos meses de Dezembro/98 a Fevereiro/99 a empresa realizou modificações em seu sistema de armazenamento de dados, sendo que eles não foram disponibilizados.

4.4.1. Primeiro conjunto de dados - JUN/98 a OUT/98:

Tratamento inicial dos dados. Basicamente o que se fez foi juntar os meses numa única planilha, estabelecer a relação entrada/saída de acordo com o tempo de residência médio do líquido e eliminar linhas onde faltavam dados.

Modelagem para cada reator. Percebe-se, pelas planilhas de dados, que existem diferenças no número de dados disponíveis para cada reator. Os dois primeiros reatores possuem mais dados do que o último. Assim para poder analisar melhor estes reatores, onde se acredita que é onde grande parte do TOC é consumido, decidiu-se por manter a estrutura de uma rede para cada reator.

Variável de saída: Deseja-se prever o valor do TOC na saída de cada reator, usando dados de entrada com 24 h de diferença. Este é o tempo de retenção hidráulica médio na faixa de controle da vazão de operação. Lembrando que a predição do pH do reator, que até então era uma variável de saída da rede, foi deixada para estudos futuros.

Como os valores pouco variam e como o erro de leitura nestas condições é crítico, as redes criadas geravam somatórias dos erros nas predições muito elevadas. Ainda porque existe o problema do arredondamento na primeira casa decimal que não é realizado pela rede.

Variáveis de entrada: Muitas combinações foram sugeridas e testadas, e resolveu-se mostrar aqui os resultados que oferecem melhores condições de avaliação. Primeiro se tentou treinar a rede com dados de uma única série e usar os dados da outra somente para teste. Posteriormente, abandonou-se esta forma para treinar redes distintas para cada série. Além disso, devido às diferenças existentes, e mesmo pelas diferenças na periodicidade das análises, as variáveis de entrada variam de acordo com cada reator dentro da mesma série. No item seguinte especificam-se as variáveis de entrada considerada em cada um dos casos.

4.4.1.1. *Treinando cada reator com dados de uma série e usando a outra como teste:*

Como se tratam de duas séries (A e B) de reatores com dados de projeto idênticos, poderia se esperar que o desempenho entre uma série e a outra fosse bastante similar. Assim, espera-se que a rede captando as principais relações para o reator de uma das séries possa prever as condições da saída para o reator correspondente da outra série. Para cada um dos reatores as variáveis de entrada consideradas são:

- para o 1º Reator:

pH (202), TOC (202), pH e TOC (203, inicial), MES (203 e 206), vazão de entrada de líquido no 1º reator da série, vazão de purga, vazão de entrada de lodo no 1º reator da trancha. A série B constitui o conjunto de treino e a série A o conjunto de teste, para todos os reatores. O banco de dados se reduziu para 58 pontos para a série B e 62 para a série A.

- para o 2º Reator:

pH (202), TOC (202), pH (203), TOC (203), pH e TOC (204, inicial), MES (203, 204, 206), vazões de entrada de líquido no 1º e 2º reatores da série, vazão de purga, vazões de entrada de lodo no 1º e 2º reatores da série. O banco de dados se reduziu para 59 pontos para a série B e 61 para a série A.

- para o 3º Reator:

pH (202), TOC (202), pH (204), TOC (204), vazões de entrada de líquido nos reatores da série, vazão de purga, vazões de entrada de lodo nos reatores da série. O banco de dados se reduziu para 71 pontos para a série B e 72 para a série A.

- para a Lagoa:

Vazões de saída do último reator de cada série, TOC (205A e 205B), kg de TOC/h correspondente (descarga), Vazão total no espessador. O banco de dados se reduziu para 12 pontos para o conjunto de treino e 6 para o conjunto de teste.

O tempo de residência, utilizado para relacionar as entradas com as saídas do processo, foi de um dia.

Infelizmente, algumas variáveis consideradas importantes não puderam ser usadas. No caso da concentração de oxigênio, preferiu-se por não usá-la devido aos problemas no medidor. As concentrações de N e P também foram desconsideradas porque acreditava-se que o sistema de controle sobre sua alimentação poderia levar perturbações ao sistema. A idade do lodo não foi calculada diretamente, mas a vazão de purga foi sempre utilizada para representar indiretamente a influência desta condição de operação.

Resultados:

Os resultados alcançados utilizando estas variáveis podem ser visualizados na Tabela 18. Note-se que os resultados obtidos para o pH também constam. Em vista da dificuldade em obter uma boa predição para esta variável com o conjunto de informações disponíveis, a predição do pH foi colocada em segundo plano. Na tabela estão mostrados os valores dos erros RMS para o conjunto de treino e de teste, e os índices de correlação correspondentes. O valor indicado como *final* refere-se ao encontrado quando a rede foi treinada até que o erro de treino fosse muito pequeno. Já o valor indicado como *melhor* se refere ao ponto de menor erro no conjunto de teste.

Para observar melhor o ajuste destas redes pode-se observar os gráficos das Figuras 36 a 42, os quais relacionam o valor predito com o desejado.

Tabela 18 – Resultados obtidos com dados de Jun/98 a Out/98, para as variáveis pH e TOC. Para o primeiro reator o TOC é o da amostra 203, para o segundo o da amostra 204, para o terceiro o da amostra 205, e o da lagoa é o TOC da amostra 209.

Variável Final/Melhor	RMS conj. de treino	RMS Conj. de teste	Índice de correlação conj. de treino	Índice de correlação conj. de teste
TOC – Final	0.0001	0.0506	1.0000	0.9433
TOC – melhor	0.0353	0.0327	0.9630	0.9698
PH – Final	0.0001	0.2147	1.0000	0.2094
PH – melhor	0.1427	0.1022	0.0091	0.0111
REATOR2	0.0157	0.0371	0.9926	0.9281
REATOR3	0.0347	0.0355	0.9412	0.8624
LAGOA	0.0664	0.0321	0.9394	0.9552

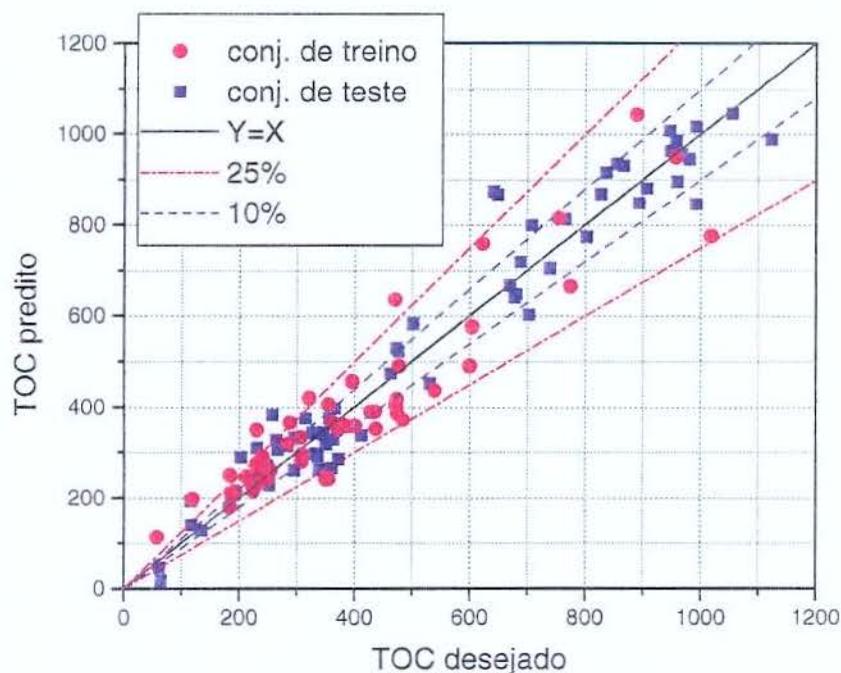


Figura 36 - REATOR 1 – Valor predito do TOC (em mg/l) da amostra 203 em relação ao valor encontrado.

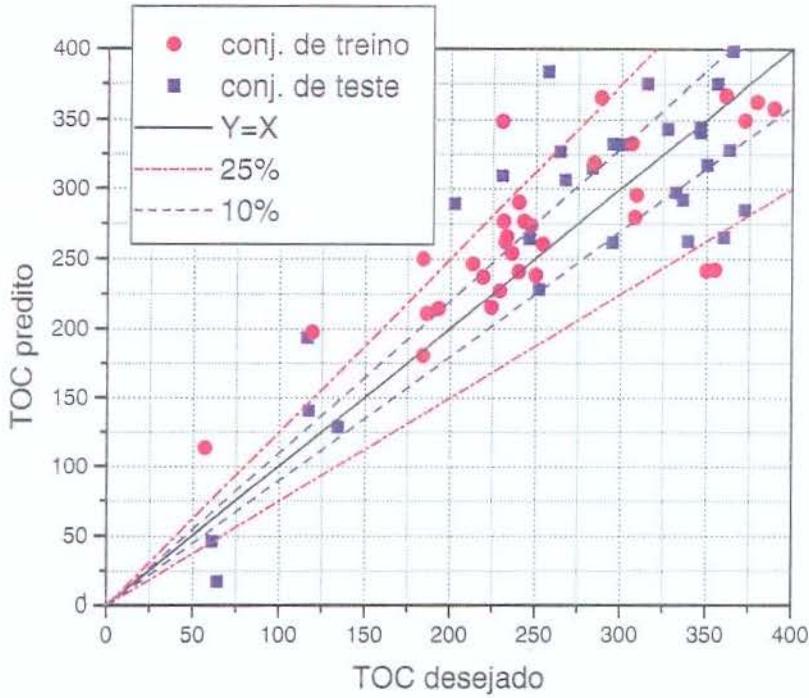


Figura 37 - REATOR 1 – Valor predito do TOC da amostra 203 em relação ao valor encontrado. Escala ampliada para valores inferiores a 400 mg/l.

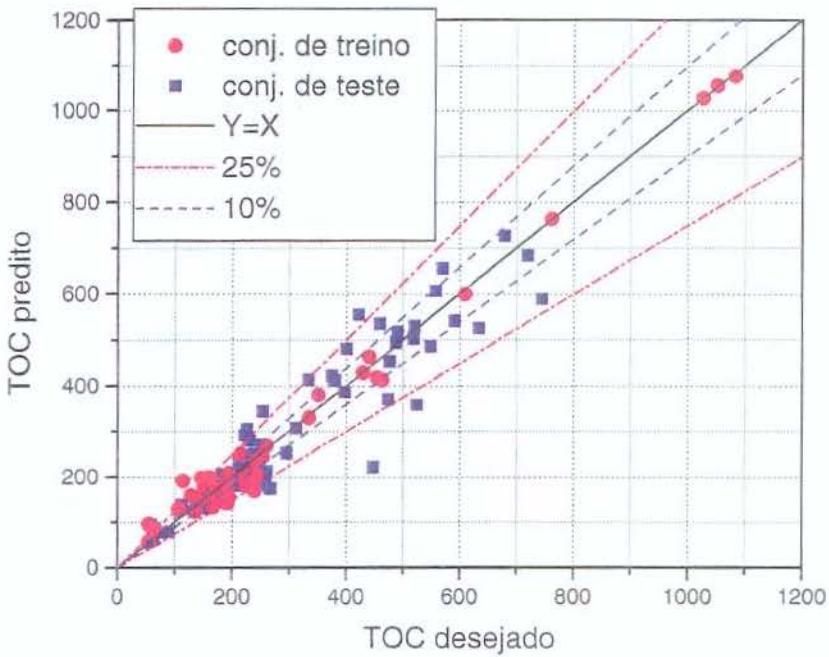


Figura 38 - REATOR 2 – Valor predito do TOC (em mg/l) da amostra 204 em relação ao valor encontrado.

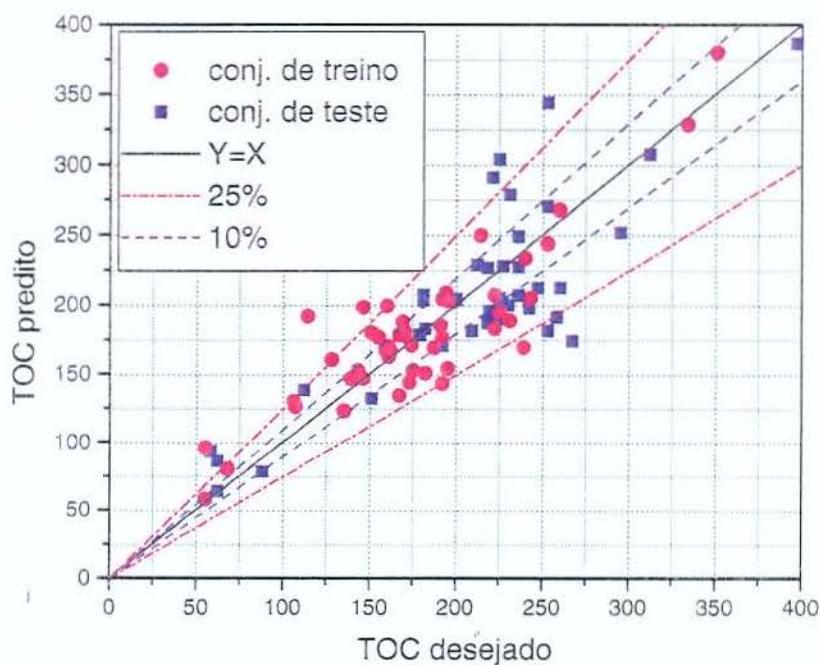


Figura 39 - REATOR 2 – Valor predito do TOC da amostra 204 em relação ao valor encontrado. Escala ampliada para valores inferiores a 400 mg/l.

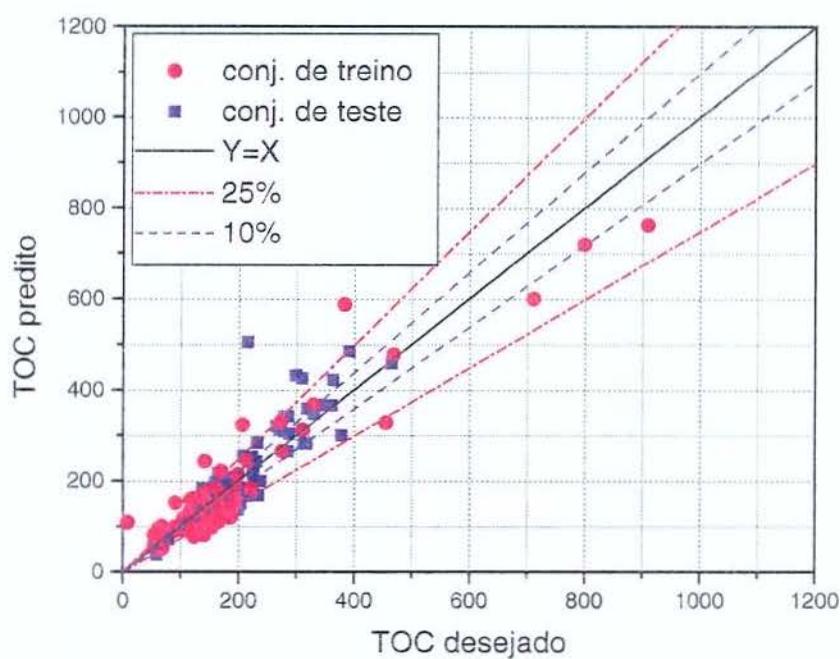


Figura 40- REATOR 3 – Valor predito do TOC (em mg/l) da amostra 205 em relação ao valor encontrado.

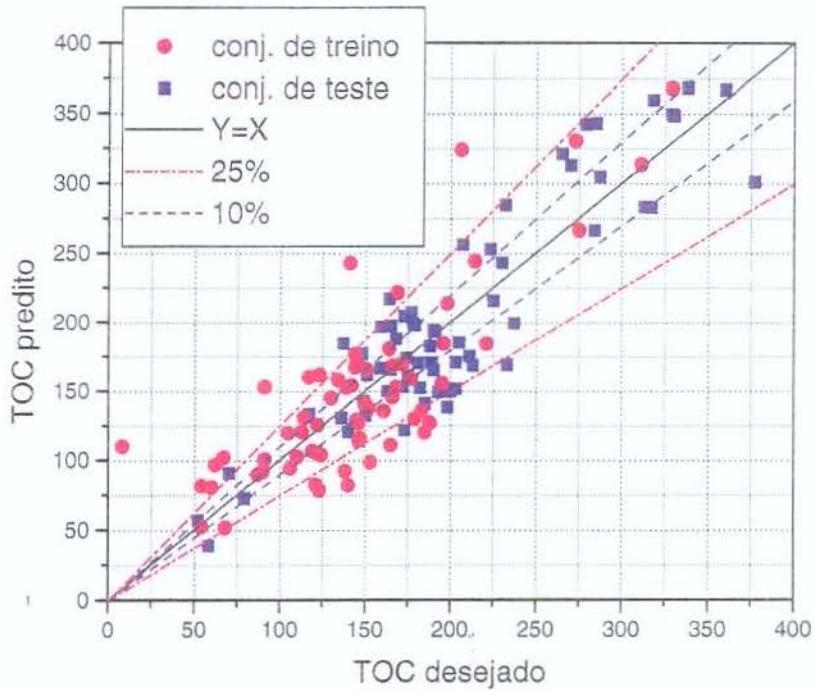
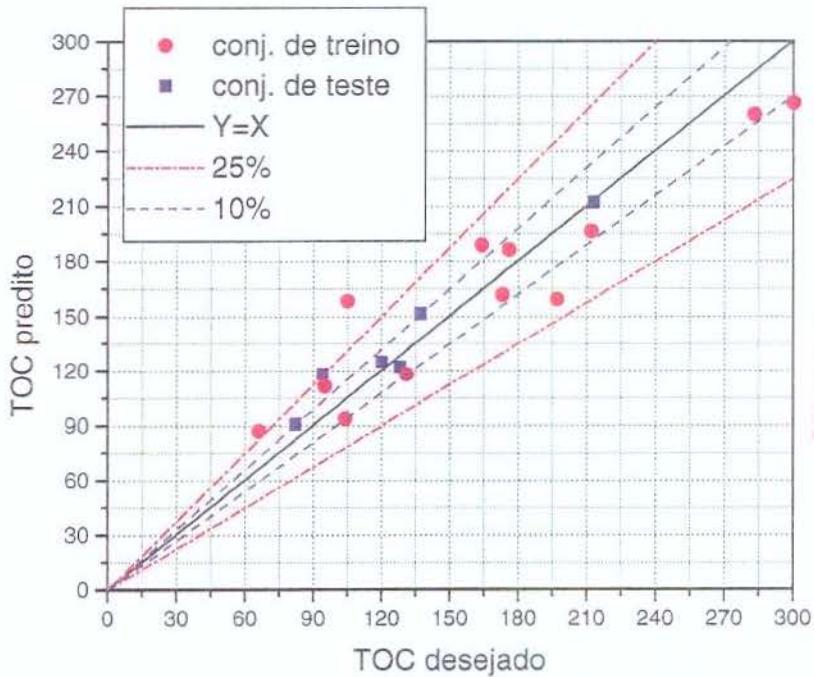


Figura 41 - REATOR 3 – Valor predito do TOC da amostra 205 em relação ao valor encontrado. Escala ampliada para valores inferiores a 400 mg/l.



UNICAMP
 BIBLIOTECA CENTRAL
 SEÇÃO CIRCULANTE

Figura 42 - LAGOA – Valor predito do TOC (em mg/l) da amostra 209 em relação ao valor encontrado.

4.4.1.2. *Treinando para cada reator individualmente:*

Os reatores da série B foram construídos de forma a duplicar os reatores da série A. Mesmo assim existem algumas diferenças entre eles, e investigou-se qual seria o resultado se o treinamento fosse feito para os reatores em separado. A seguir os erros, os índices de correlação e os gráficos correspondentes são apresentados.

REATOR 1 - Série A:

$RMS_{(TREINO)}$	$RMS_{(TESTE)}$	$R_{(TREINO)}$	$R_{(TESTE)}$
0.0329	0.0377	0.9802	0.9774

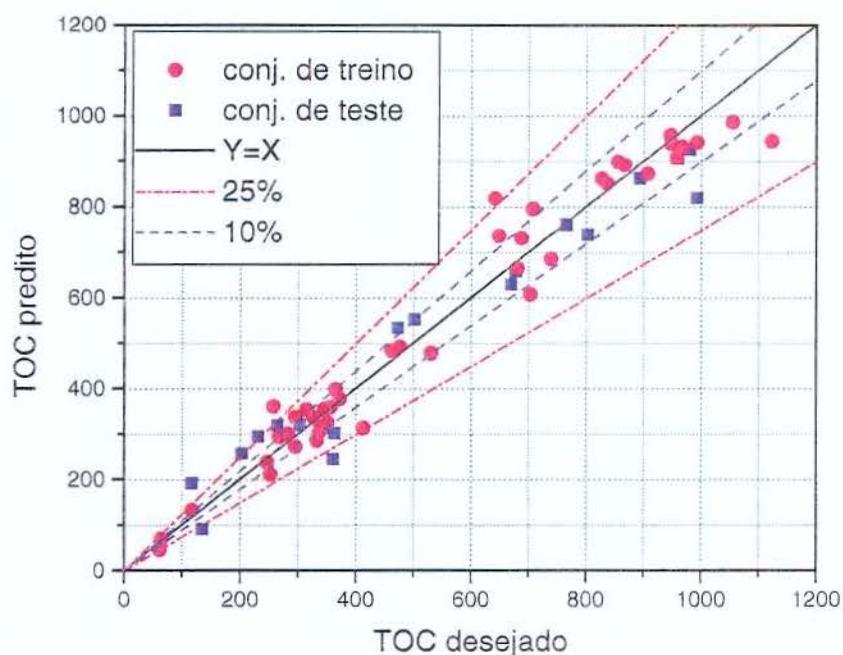


Figura 43 - REATOR 1A – Valor predito do TOC (em mg/l) da amostra 203A em relação ao valor encontrado.

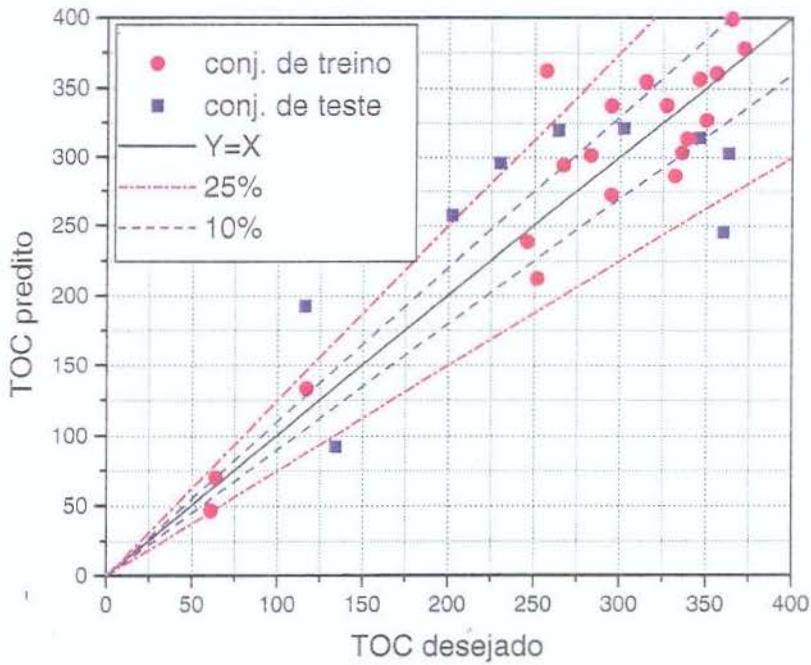


Figura 44 - REATOR 1A – Valor predito do TOC da amostra 203A em relação ao valor encontrado. Valores inferiores a 400 mg/l.

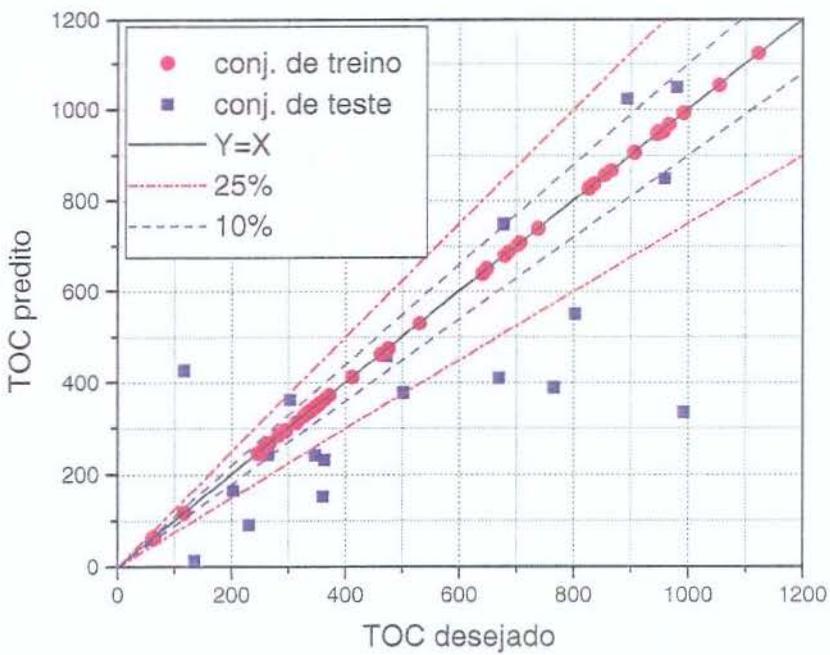


Figura 45 - REATOR 1A – Valor predito do TOC (em mg/l) da amostra 203A em relação ao valor encontrado. Treinada até obter erro de treino muito pequeno.

Demonstrando o que acontece ao se deixar a rede treinar indefinidamente, observa-se no gráfico da Figura 45 que os erros para o conjunto de treino estão muito próximos de zero, enquanto que os erros do conjunto de teste aumentaram consideravelmente. Assim, diz-se que a rede da Figura 43 tem melhores características para a generalização do que a rede da Figura 45. Foi por esta razão que se optou pela rede a qual apresenta um erro RMS mínimo para o conjunto de teste como a rede a ser utilizada no simulador.

REATOR 1 - Série B:

$RMS_{(TREINO)}$	$RMS_{(TESTE)}$	$R_{(TREINO)}$	$R_{(TESTE)}$
0.0254	0.0524	0.9611	0.9459

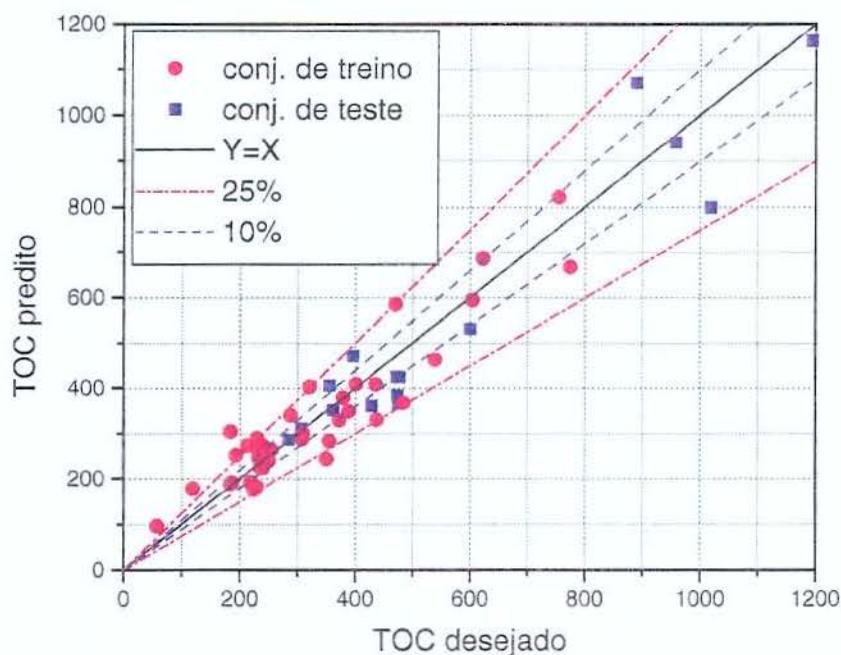


Figura 46 - REATOR 1B – Valor predito do TOC (em mg/l) da amostra 203B em relação ao valor encontrado.

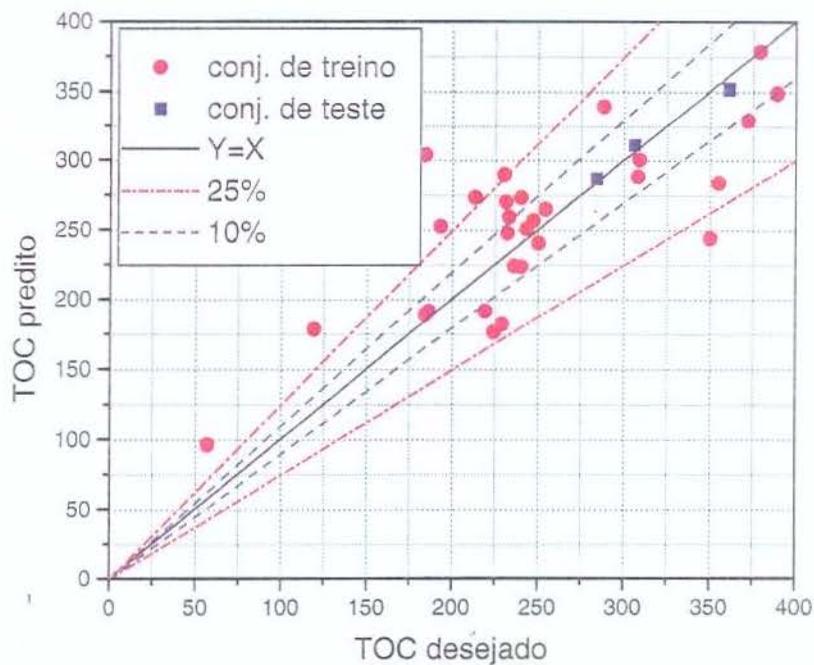


Figura 47 - REATOR 1B – Valor predito do TOC da amostra 203B em relação ao valor encontrado. Valores inferiores a 400 mg/l.

Como tentativa de melhorar a predição treinou-se a rede introduzindo também como variável de entrada a carga do reator em termos de kg/h de TOC. Os resultados em relação ao primeiro reator estão demonstrados nas Figuras 48 a 50.

REATOR 1 – Série A - COM AS CARGAS COMO ENTRADAS

$RMS_{(TREINO)}$	$RMS_{(TESTE)}$	$R_{(TREINO)}$	$R_{(TESTE)}$
0.0332	0.0415	0.9799	0.9801

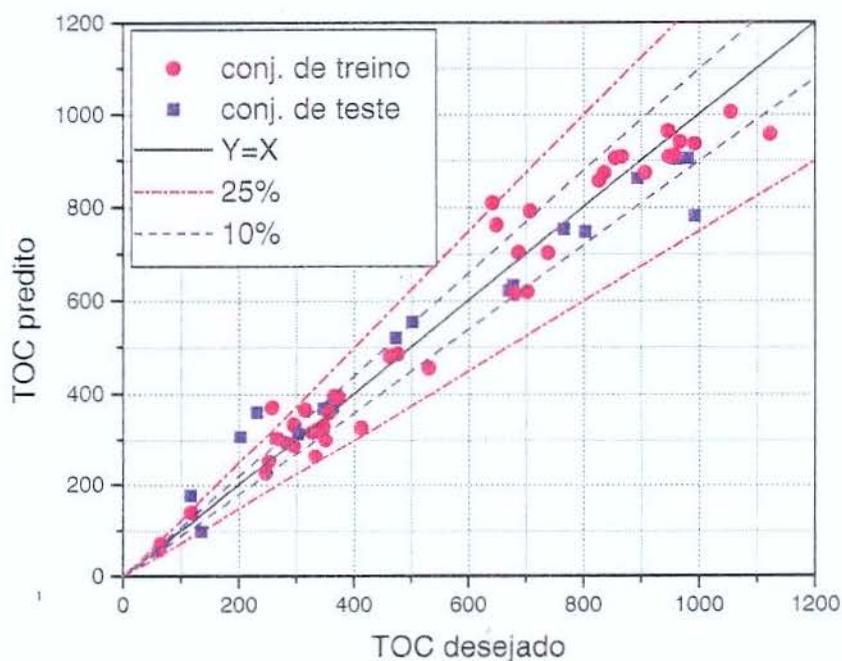


Figura 48 - REATOR 1A – Valor predito do TOC (em mg/l) da amostra 203A em relação ao valor encontrado. Utiliza-se como variável de entrada a carga do reator.

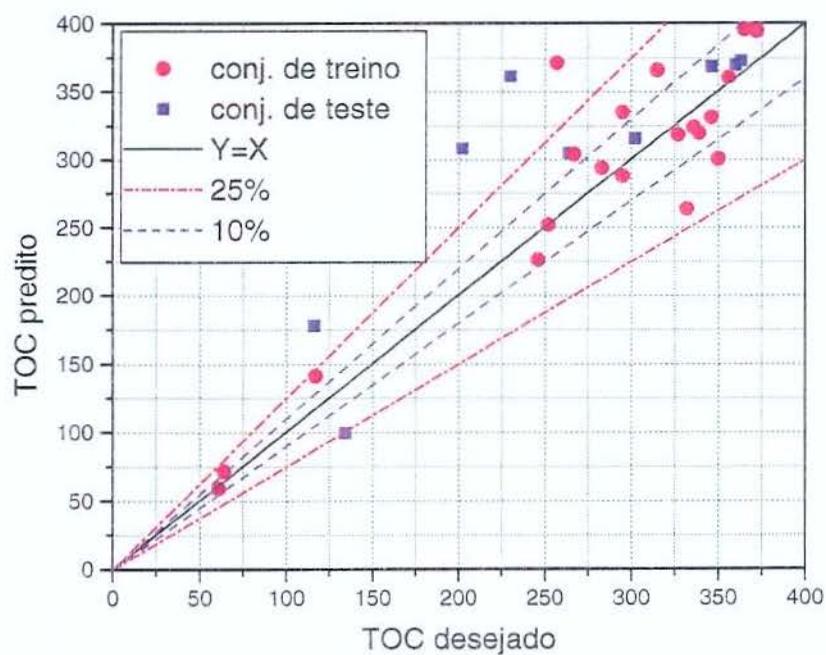


Figura 49 - REATOR 1A – Valor predito do TOC da amostra 203A contra o valor observado. Utiliza-se como variável de entrada a carga do reator. Valores inferiores a 400 mg/l.

O gráfico da Figura 50 mostra a vazão calculada de saída de em kg/h de TOC para o primeiro reator. O período ainda corresponde ao da primeira fase dos dados. Dentro deste período se observa que o sistema trabalhou em dois patamares distintos. O primeiro onde a descarga de TOC gira em torno de 35 kg/h de TOC; e o segundo onde a descarga gira em torno de 13 kg/h de TOC. No final do gráfico ainda observa-se o início de um terceiro período com descargas abaixo de 10 kg/h. Posteriormente observou-se que a unidade passou a operar dentro desta última faixa de valores, e por isso o estudo foi realizado em separado. O conjunto de dados total foi dividido neste ponto, treinando-se redes em separado para cada período.

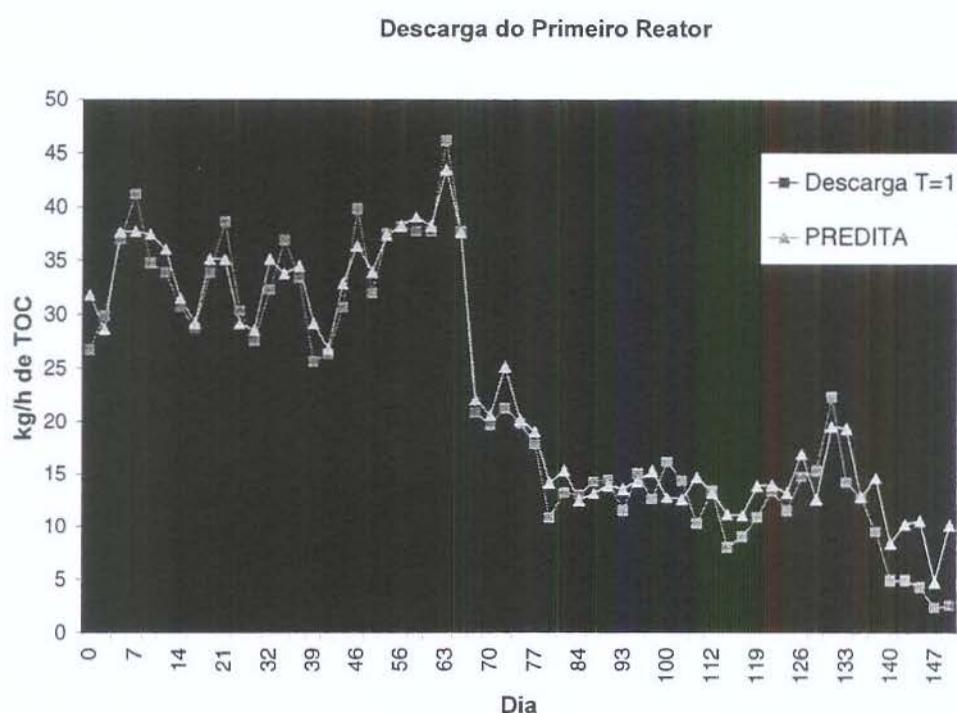


Figura 50 – Acompanhamento da vazão de TOC em kg/h ao longo dos dias do período entre jun/98 e out/98.

Conclusões parciais:

1. Trabalhar com as variáveis de saída (TOC, pH) em redes separadas. A predição do pH apresentou problemas, pois a faixa é muito estreita e a falta de arredondamento para a primeira casa decimal gera erros elevados. A predição do pH foi abandonada até que se evidencie a sua necessidade.

2. Otimizar o número de neurônios da camada intermediária.
3. Avaliar a inclusão de novas variáveis, principalmente a carga da unidade.
4. Treinar isoladamente os reatores de cada série.
5. Eliminar dados inconsistentes. Para isto deve-se estabelecer alguns critérios que possam detectar quais os pontos em que houve algum problema na planta, provocando por exemplo flutuações muito grandes.

4.4.2. Segundo conjunto de dados - NOV/98 + MAR/99 a JUN/99:

Tratamento inicial dos dados. Identificadas algumas possíveis melhorias possíveis, avaliou-se nesta fase do trabalho a utilização de algumas idéias que haviam surgido na fase anterior. Devido a dois problemas principais intrínsecos na simulação de unidades industriais:

- a) Estabilidade do sistema (a simulação dinâmica exige uma maior complexidade na rede e uma frequência de coleta de dados mais adequada).
- b) Erros de amostragem, análise, leitura e digitação; deve haver uma seleção dos dados coletados para que não se introduza no treinamento esses ruídos para a rede.

Assim, procurou-se elaborar um critério para eliminar pontos inconsistentes para o treinamento das redes e utilizar todas as variáveis possíveis e as combinações que possam relacionar uma nova informação. Cita-se como exemplo, a idade do lodo que é função de outras variáveis que já constavam no modelo. As variáveis estudadas foram as seguintes:

Amostra 202 – TOC

Amostra 203 – pH, TOC, Nitrogênio, Fósforo, MES

Amostra 206 – MES

Vazão de líquido e de lodo no 1º reator

Oxigênio dissolvido

TOC médio (202)

Carga (202i e 202média, 203 t=0)

Idade do Lodo (trancha e reator)

K_{mesx} , K_{mesy} , K_{subx} , K_{suby}

Estes últimos parâmetros referem-se aos termos das retas ao colocar em gráfico a cinética simplificada para o lodo:

$$\left[\left(\frac{\Delta X}{\Delta t} + \frac{24Q_R X_R}{V} \right) \frac{1}{X_e} + 24 \frac{Q}{V} \right] = Y \left[\frac{1}{X} \left(\frac{24QS_0}{V} - \frac{24(Q+Q_R)S}{V} - \frac{\Delta S}{\Delta t} \right) \right] - k_d \quad (31)$$

e para o substrato:

$$\left[\frac{1}{X} \left(\frac{24QS_0}{V} - \frac{24(Q+Q_R)S}{V} - \frac{\Delta S}{\Delta t} \right) \right]^{-1} = \frac{K_S}{K} \left(\frac{1}{S} \right) + \frac{1}{K} \quad (32)$$

Estas equações derivam da equação 27 e 28, que estão na pág. 34. Os parâmetros mencionados são portanto:

$$K_{SUB-Y} = \left[\left(\frac{\Delta X}{\Delta t} + \frac{24Q_R X_R}{V} \right) \frac{1}{X_e} + 24 \frac{Q}{V} \right] \quad (33)$$

$$K_{SUB-X} = \left[\frac{1}{X} \left(\frac{24QS_0}{V} - \frac{24(Q+Q_R)S}{V} - \frac{\Delta S}{\Delta t} \right) \right] \quad (34)$$

$$K_{MES-Y} = \left[\frac{1}{X} \left(\frac{24QS_0}{V} - \frac{24(Q+Q_R)S}{V} - \frac{\Delta S}{\Delta t} \right) \right]^{-1} \quad (35)$$

$$K_{MES-X} = \left(\frac{1}{S} \right) \quad (36)$$

Os resultados alcançados ao se considerar estas condições não correspondeu às expectativas. Analisando o problema percebeu-se, principalmente, que quanto mais variáveis procurava-se incluir menor se tornava o conjunto de dados. Contudo, mesmo com estas dificuldades, resolveu-se utilizar estes procedimentos para uma análise primária das variáveis mais importantes, procurando então eliminar as que tivessem menor influência. Neste caso foram eliminadas da relação inicialmente proposta:

Amostra 203 – Nitrogênio, MES

Amostra 206 – MES

Vazão de lodo no 1º reator

Carga (202 média)

Idade do Lodo (trancha e reator)

K_{mesx} , K_{mesy} , K_{suby} .

Outro avanço alcançado com estes estudos foi a obtenção de critérios para eliminar os pontos de maior flutuação. Esses critérios foram sugeridos ao se analisar os gráficos dos parâmetros da cinética resumida em comparação com a série temporal de cada uma das variáveis importantes. Mesmo de forma subjetiva, escolhiam-se os pontos que aparentavam não pertencer ao conjunto dos outros pontos. Avaliava-se então as outras variáveis e sempre se encontrava alguma que sugeria a anomalia.

Avaliação dos resultados: Trabalhando com 29 pontos totais, usando 15 destes para o treinamento, o resultado não foi aceitável, onde o erro RMS ficou acima de 0.1 e o índice de correlação abaixo de 0.7. Treinou-se a rede com os 29 pontos, sem conjunto de teste portanto. A tentativa seria de com isso poder eliminar variáveis menos importantes e com isso o número de pontos poderia ser maior para poder se treinar as redes com as variáveis restantes. Os resultados alcançados utilizando-se os 29 pontos podem ser avaliados nos gráficos seguintes. Na Figura 51 mostra-se o ajuste para o conjunto de dados que foi utilizado no treinamento e o gráfico da Figura 52 avalia a predição justamente para aqueles dados que haviam sido eliminados pelos critérios mencionados acima, tais dados foram reunidos em um conjunto denominado de avaliação. Como, neste caso, não se utilizou um conjunto de teste para controlar o treinamento haveria uma dúvida sobre a generalidade da rede. Contudo, mesmo utilizando o conjunto de avaliação para testar a rede, o resultado alcançado foi bastante aceitável para o propósito deste estudo. Acredita-se que se a rede forneceu uma predição razoável para estes pontos porque com o aumento no número de dados ela pode ter se tornado mais genérica do que a anterior com menos pontos para treinar.

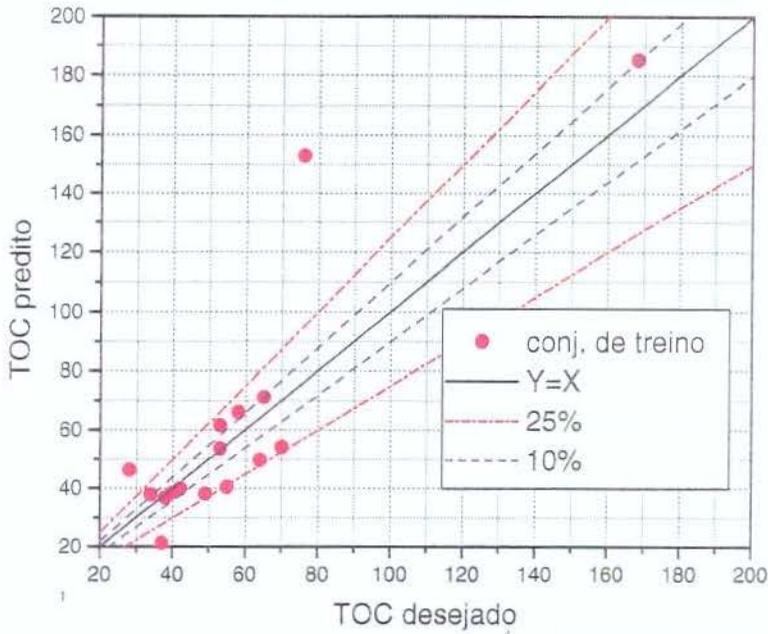


Figura 51 - REATOR 1A – Valor predito do TOC da amostra 203A em relação ao valor encontrado. Rede treinada para o segundo conjunto de dados sem usar conjunto de teste, e com todas as variáveis propostas.

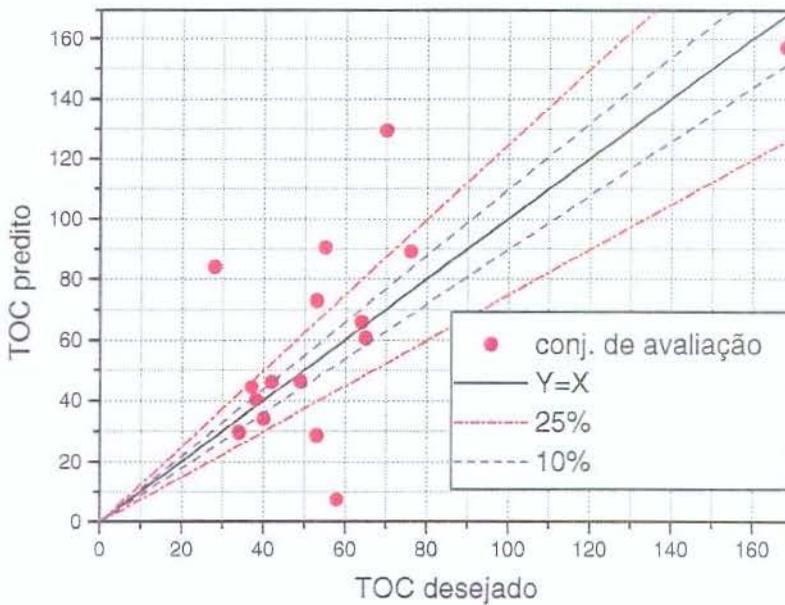


Figura 52 - REATOR 1A – Valor predito do TOC da amostra 203A em relação ao valor encontrado para o **conjunto de avaliação**. Rede treinada para o segundo conjunto de dados sem usar conjunto de teste, e com todas as variáveis propostas.

Com as redes treinadas, procedeu-se uma avaliação da influência de cada variável. Eliminando-se as variáveis de menor influência e retreinando a rede houve uma degradação considerável nos resultados. Muitos pontos ainda continuavam dentro de uma faixa de erro razoável. Contudo, concluiu-se que estas redes não poderiam ser usadas desta forma. Talvez o resultado pudesse ser melhor ao aumentar-se ainda mais o número de pontos. Para resolver este problema poderia ser feito uma eliminação ainda mais drástica nas variáveis a utilizar, de forma a aumentar o número de dados. Desta forma resolveu-se utilizar a mesma configuração de variáveis da primeira fase deste trabalho. Com isto o número de dados foi satisfatório e os resultados também.

Conclusões parciais:

- 1) Existe uma necessidade de aumentar o número de dados. Considerou-se que adicionar os pontos da primeira fase não iria resolver o problema pois a diferença no desempenho da estação é muito grande e a tendência é que a situação atual perdesse, inviabilizando o uso daqueles dados.
- 2) Foram estabelecidos alguns critérios para eliminar os pontos inconsistentes, ligados a possíveis problemas de instabilidade ou algum problema de amostragem, por exemplo. Os mais importantes são os seguintes:

$$\Delta\text{TOC}_{202} > 400 \text{ (mg/l)/dia} \quad (\text{média } 1900)$$

$$\Delta\text{TOC}_{203} > 20 \text{ (mg/l)/dia} \quad (\text{média } 40)$$

$$|\text{TOC}_{202i} - \text{TOC}_{202\text{méd}}| > 200 \text{ (mg/l)/dia}$$

$$\Delta\text{Carga}_{202} > 3 \text{ (kg/dia)/dia} \quad (\text{média } 3)$$

$$\Delta\text{MES}_{203} > 800 \text{ (mg/l)/dia} \quad (\text{média } 3500)$$

$$\Delta\text{Idade Lodo} > 100\% / \text{dia}$$

4.4.3. Conjunto diferenciado de dados

Trata-se do 2º conjunto - Novembro/98 e Março/99 a Junho/99, adicionando os dados de baixo TOC de saída do primeiro reator do conjunto anterior; e ainda usando como variáveis de entrada as mesmas usadas para o 1º conjunto.

Pretendendo melhorar o grau de predição, decidiu-se por treinar as redes para o 2º conjunto da mesma maneira como se tinha feito para o 1º conjunto, utilizando as mesmas variáveis de entrada. As Figuras 58 a 64 mostram a relação entre o resultado predito e o observado. Observa-se, também pela Tabela 19, que de uma maneira geral, os índices alcançados com estes dados são próximos aos registrados no treinamento do 1º conjunto. Nos casos onde houve um resultado pior comparado com os anteriores é possível encontrar as razões que explicam a maior dificuldade encontrada. Como é o caso do segundo reator da série A. Analisando os dados, percebe-se que a vazão de purga nesse período esteve bem acima dos valores anteriores. Nos casos onde os índices de correlação foram inferiores aos observados anteriormente percebe-se que isto é devido a um menor número de informações, ocorrendo casos com poucas variáveis ou com poucos pontos experimentais, como é o caso das redes do terceiro reator e da lagoa, respectivamente. Mesmo assim, a predição global não é prejudicada porque grande parte da carga poluidora é eliminada principalmente nos dois primeiros reatores.

Tabela 19– Resultados obtidos no treinamento da rede para predizer o TOC (em mg/l) do sistema de tratamento de efluentes.

Rede Equipamento	Erro RMS conjunto de treino	Erro RMS conjunto de teste	Índice de correlação do conj. de treino	Índice de correlação do conj. de teste
REATOR 1 A	0.0155	0.0265	0.9928	0.9779
REATOR 2 A	0.0445	0.0463	0.9709	0.9615
REATOR 3 A	0.0244	0.0307	0.9825	0.9724
REATOR 1 B	0.0187	0.0229	0.9879	0.9807
REATOR 2 B	0.0285	0.0339	0.9825	0.9699
REATOR 3 B	0.0309	0.0311	0.9769	0.9812
LAGOA	0.0396	0.0401	0.9730	0.9684

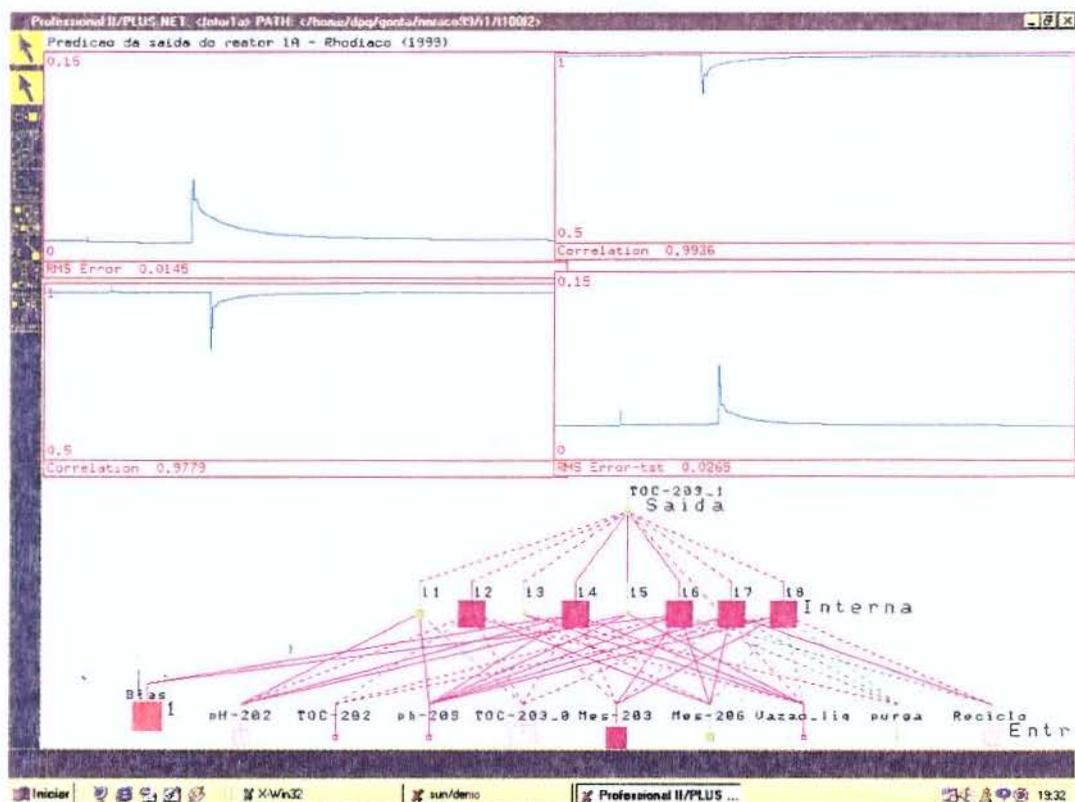


Figura 53 – Tela do programa para treinar a rede do reator 1A

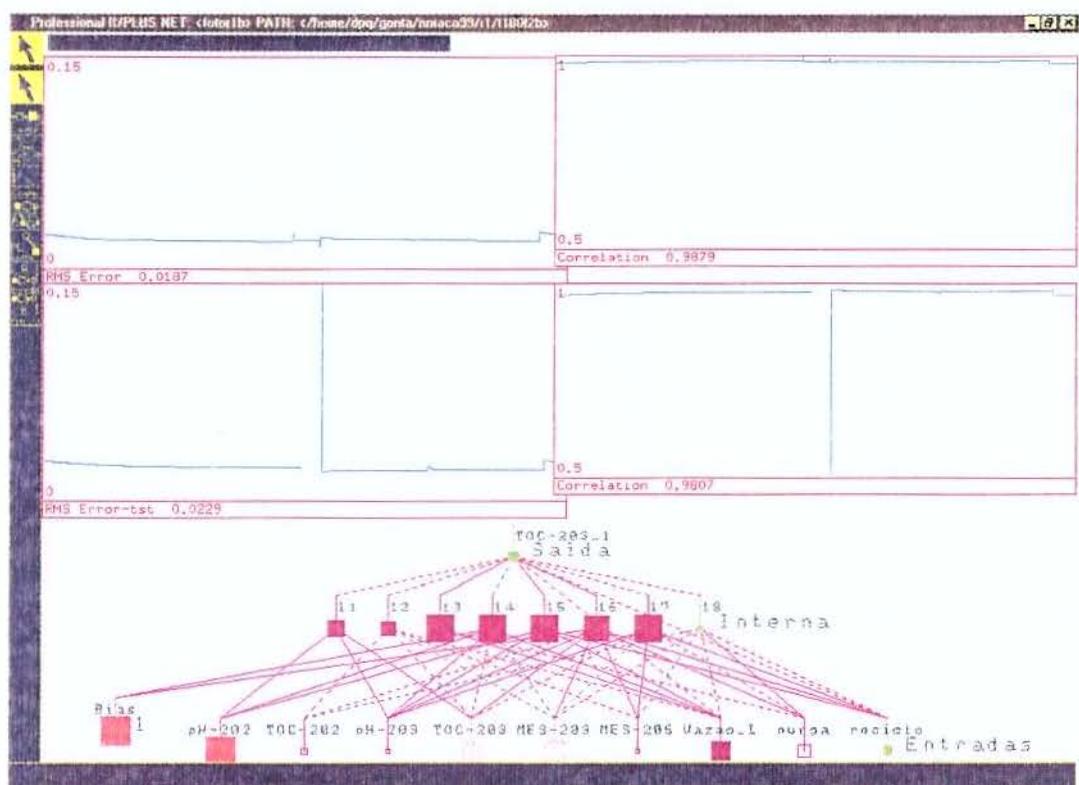


Figura 54 – Tela do programa para treinar a rede do reator 1B

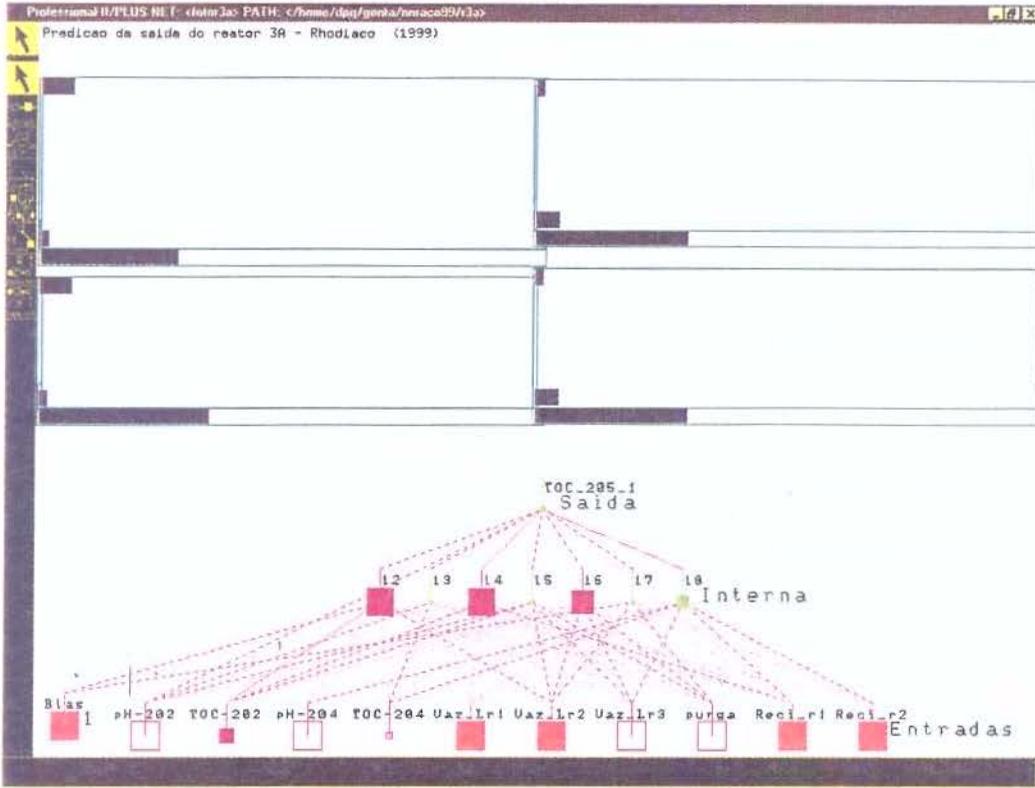


Figura 55 – Tela do programa para treinar a rede do reator 3A

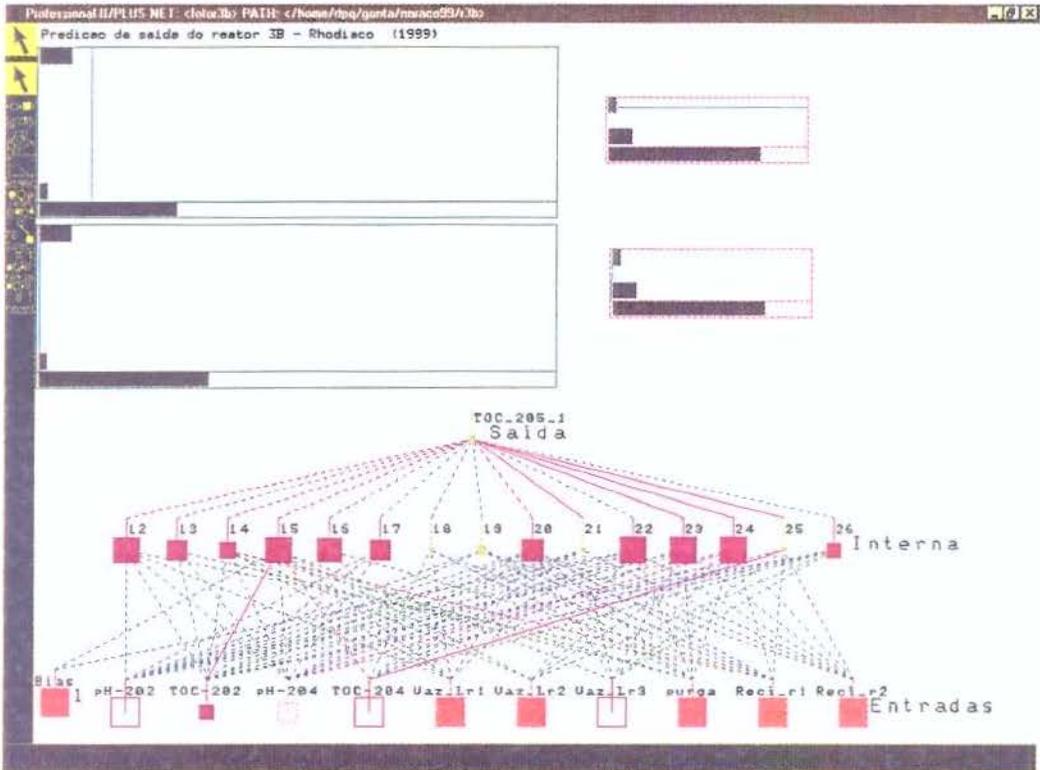


Figura 56 – Tela do programa para treinar a rede do reator 3B

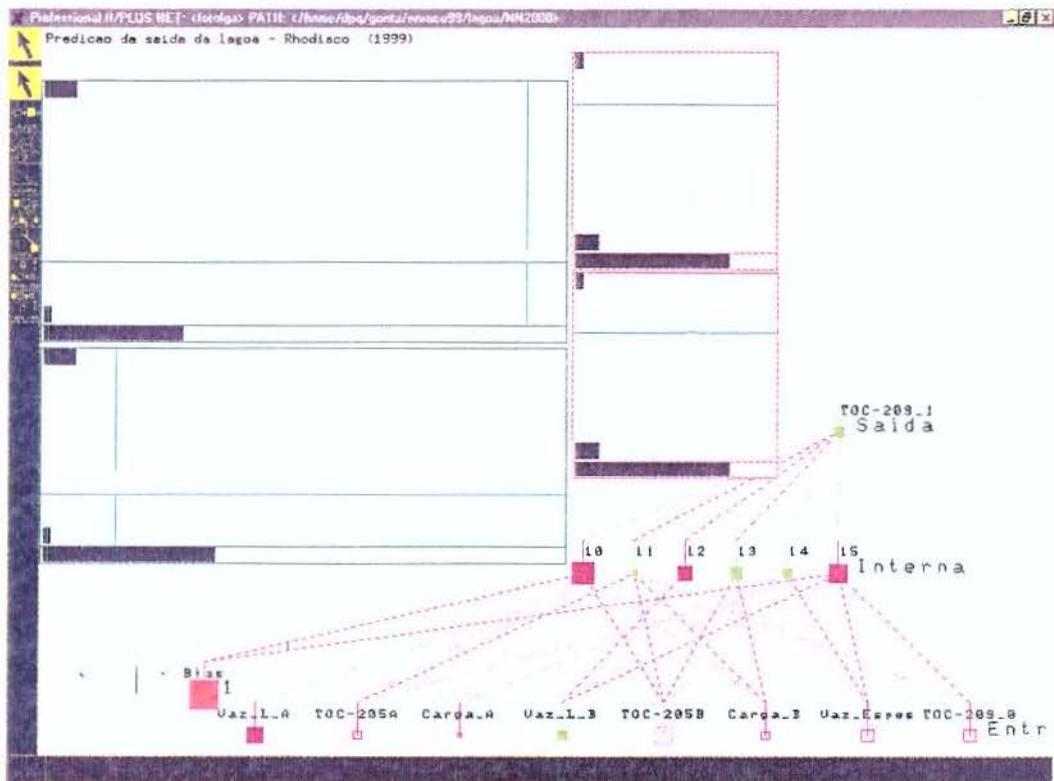


Figura 57 – Tela do programa para treinar a rede da lagoa

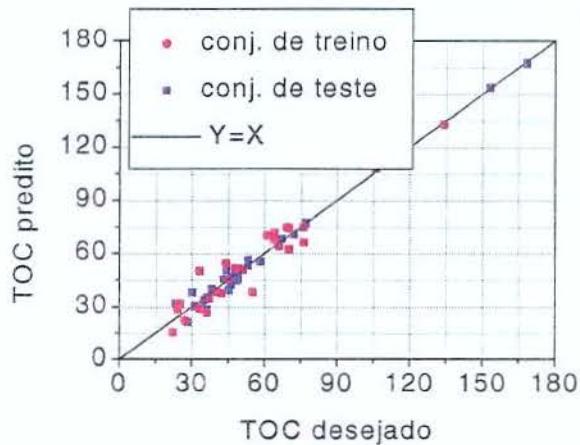


Figura 58– Predição do TOC para o reator 1A

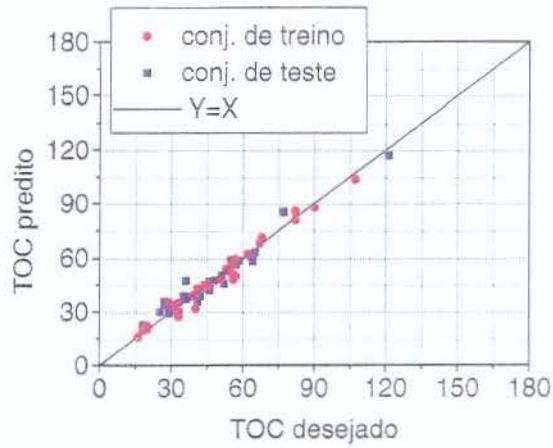


Figura 59- Predição do TOC para o reator 1B

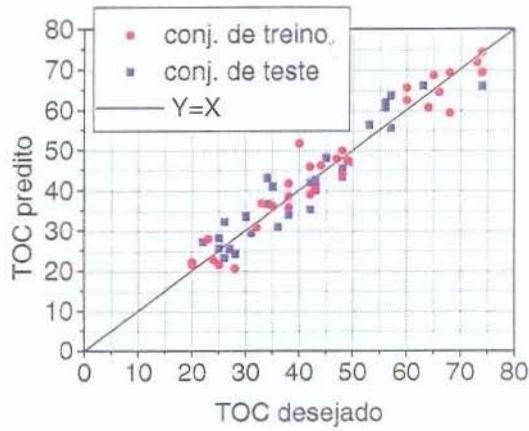


Figura 60- Predição do TOC para o reator 2A

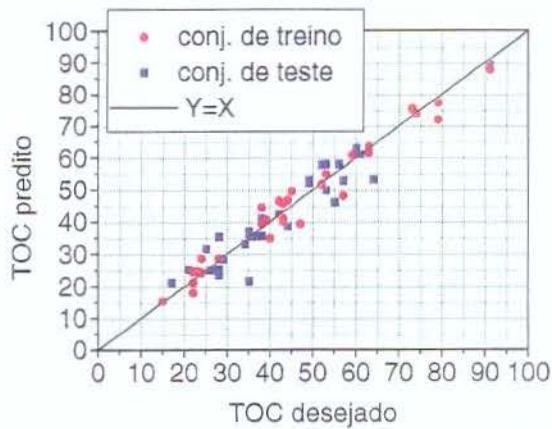


Figura 61- Predição do TOC para o reator 2B

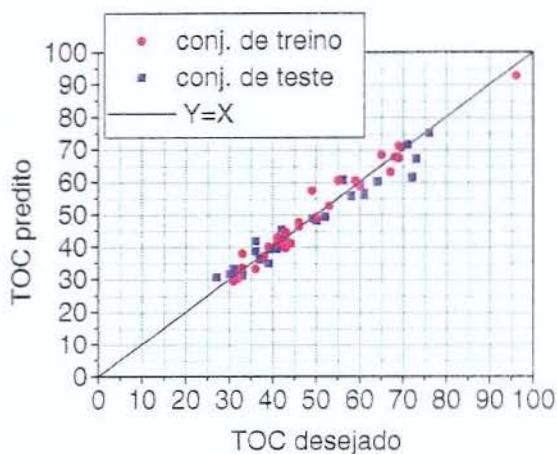


Figura 62- Predição do TOC para o reator 3A

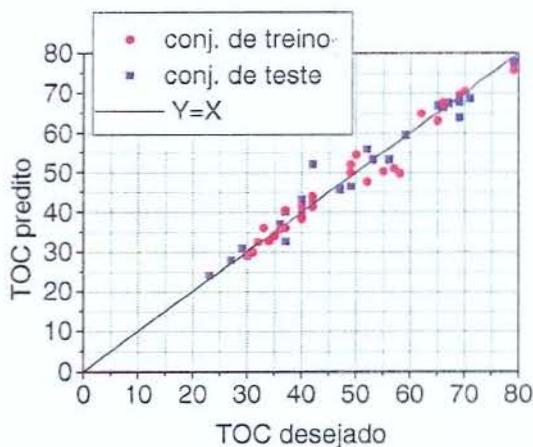


Figura 63- Predição do TOC para o reator 3B

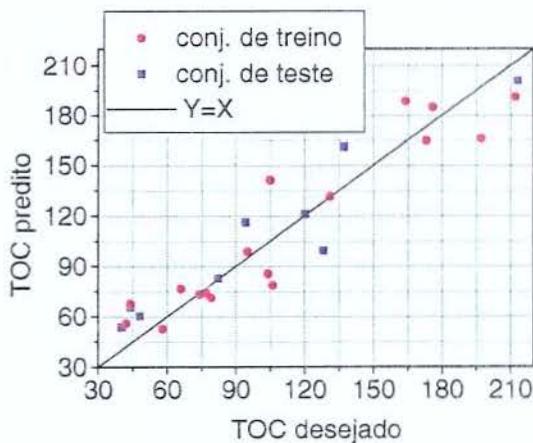


Figura 64- Predição do TOC para a lagoa de polimento

4.5. ESTUDO DE CASO COM AS REDES TREINADAS

Um dos aspectos que se desejava conhecer seria o impacto sobre o TOC residual por uma alteração de pH na entrada do primeiro reator. Através de uma aplicação simples da rede treinada pode-se obter um resultado que indica a influência na concentração final do substrato, como se observa na Figura 65. Com a rede neuronal obteve-se os valores preditos da concentração de TOC variando-se o pH dentro da faixa utilizada para o treinamento e utilizando-se dos valores médios para as outras variáveis.



Figura 65– Influência do pH no TOC residual do reator 1A predito pela rede

Contudo, uma decisão sobre o valor ideal do pH na entrada deve considerar, além de diversas faixas de trabalho para as outras variáveis, também a sua influência sobre o lodo. Saliente-se que o que se está verificando é que a rede interpreta que na maioria das vezes que o pH de entrada foi baixo os valores de TOC na saída foram menores. Esperava-se que em algum ponto houvesse um valor mínimo para o TOC, pois se sabe que valores baixos do pH não são desejáveis. O que poderia prever melhor este efeito seria uma análise sobre a eficiência do lodo após exposição contínua em pH baixo.

4.6. ELABORAÇÃO DO PROGRAMA PARA PREDIÇÃO DO TOC

A fim de propiciar a utilização das redes treinadas foi criado um programa para prever os valores do TOC da amostra 209. O nome escolhido para ele foi SIMNEURAL. Na versão atual, o programa faz basicamente:

- 1) Leitura de dados através de arquivo.

- 2) Entrada dos dados através do teclado utilizando diálogos próprios. A estrutura utilizada prevê que os dados devem ser introduzidos, de acordo com a necessidade, para um mesmo reator em diversos dias. Observe-se pela Figura 68, por exemplo, que no cabeçalho está a designação de qual equipamento se refere (série, reator ou lagoa), e ainda o dia a que se refere, no campo específico.

Admite-se que o usuário, na maioria das vezes, irá querer uma previsão a partir dos dados de hoje, por exemplo. Assim, designou-se como “data atual” o dia a partir do qual se tem os dados. No diálogo sempre se refere a “data atual” pela abreviação D.A. Por exemplo, se a data atual seria hoje, quando se tem D.A.+1 está se referindo ao dia de amanhã e assim por diante. Para facilitar a introdução do dado correto, optou-se por utilizar o nome da variável como foi apresentado nas planilhas com os dados.

Para verificar os valores que constam na memória deve-se utilizar os mesmos diálogos da entrada de dados, pois quando ele é mostrado já vem com os valores atuais das variáveis. Quando a variável é predita aparece então o valor predito. Esta é uma das formas de se ver todos os valores utilizados e calculados. A outra forma seria a seguinte.

- 3) Gravação dos dados em arquivo para registro ou estudos repetidos com os mesmos. É possível portanto introduzir os dados pelos diálogos e gravá-los no formato apropriado para posterior leitura. O nome do arquivo criado neste caso é de “TEMPOR.DAT”. Assim, ao sair do programa ele pode ser renomeado e ser utilizado como arquivo de entrada de dados. O outro arquivo que é criado chama-se “RESULTS.TXT”.

Ele contém resumidamente todos os valores que podem ser preditos no programa, ou seja, na ordem:

TOC203A(1), TOC203B(1), TOC204A(1), TOC204A(2), TOC204B(1), TOC204B(2), TOC205A(1), TOC205A(2), TOC205A(3), TOC205B(1), TOC205B(2), TOC205B(3), TOC209(1), TOC209(2), TOC209(3), TOC209(4).

Os valores entre parênteses correspondem ao dia de referência, começando pela D.A: com o valor 1.

- 4) Comandos para a predição do TOC da amostra 209, (aquela retirada antes de ir para o receptor) desde DA+1 ate DA+4. Neste caso, se os dados ainda não

foram introduzidos, ou se foram introduzidos através do arquivo o programa reconhece esta situação e vai mostrando os diálogos necessários a predição pedida para que se introduzam os dados, ou se faça a conferência dos valores (caso da leitura do arquivo). Ao usar uma vez este diálogo o programa reconhece quando todos os dados foram digitados e poderá executar as rotinas das redes neuronais para a predição. O mesmo diálogo não é automaticamente chamado mais de uma vez. Se for necessário mudar algum dado que já tenha sido introduzido deve-se chamar o diálogo correspondente pela opção “Entrada de Dados” do menu, e depois executar o preditor. Quando se executa o preditor após a conferência individual de todos os diálogos da opção do menu “Entrada de Dados”, observe-se que ele mostra imediatamente o resultado predito.

Outro aspecto muito importante na obtenção de resultados confiáveis é obedecer aos limites das variáveis utilizadas no treinamento. Os limites empregados no treinamento estão mostrados nas Tabelas 20 a 26.

Tabela 20 – Limites e média dos valores utilizados no treinamento das redes, para o reator 1-A.

Variável	Máximo	Mínimo	Média
PH-202	9.7000	4.2000	6.2129
TOC-202	3651.0000	1256.0000	2095.6453
PH-203a	8.8000	7.7000	8.2048
TOC-203a(TO)	237.0000	19.0000	57.7903
MES-203a	6600.0000	2780.0000	4934.9194
MES-206a	12980.0000	4555.0000	9017.1611
FI-6302	46.4530	21.9270	38.5270
G-6413a	46.7740	5.7800	33.0051
FI-6303	45.0090	10.9670	25.6940
TOC-203a	168.0000	22.0000	53.1129

Tabela 21 – Limites e média dos valores utilizados no treinamento das redes, para o reator 1-B.

Variável	Máximo	Mínimo	Média
PH-202	9.7000	4.2000	6.2328
TOC-202	3651.0000	1256.0000	2098.868
PH-203b	8.7000	7.8000	8.2607
TOC-203b(T0)	105.0000	19.0000	46.2459
MES-203b	7590.0000	2650.0000	4351.393
MES-206b	12270.0000	3940.0000	7829.672
FI-6302	45.5800	9.3980	20.8969
G-6413b	25.0000	4.7007	10.1749
FI-6303	39.3500	4.7969	14.7316
TOC-203b	121.0000	16.0000	47.6557

Tabela 22 – Limites e média dos valores utilizados no treinamento das redes, para o reator 2-A.

Variável	Máximo	Mínimo	Média
PH-202	9.7000	4.2000	6.2129
TOC-202a	3651.0000	1256.0000	2095.645
PH-203a	8.8000	7.7000	8.2048
TOC-203a	237.0000	19.0000	57.7903
MES-203a	6600.0000	2780.0000	4934.919
PH-204a	8.9000	7.9000	8.3323
TOC-204a(T0)	86.0000	20.0000	43.9355
MES-204a	6970.0000	2140.0000	4854.241
MES-206a	12980.0000	4555.0000	9017.161
FI-6302	46.4530	21.9270	38.5270
FI-6305	19.9050	9.3911	16.6153
G-6413a	46.7740	5.7800	33.0051
FI-6303	45.0090	10.9670	25.6940
FI-6304	25.0120	4.7016	11.4593
TOC-204a	74.0000	20.0000	43.5968

Tabela 23 – Limites e média dos valores utilizados no treinamento das redes, para o reator 2-B.

Variável	Máximo	Mínimo	Média
PH-202	9.7000	4.2000	6.2328
TOC-202b	3651.0000	1256.0000	2098.868
PH-203b	8.7000	7.8000	8.2607
TOC-203b	105.0000	19.0000	46.2459
MES-203b	7590.0000	2650.0000	4351.393
PH-204b	8.7000	7.6000	8.3311
TOC-204b(T0)	73.0000	18.0000	42.1311
MES-204b	6090.0000	2645.0000	4273.934
MES-206b	12270.0000	3940.0000	7829.672
FI-6307	45.5800	9.3980	20.8969
FI-6308	19.5400	0.0102	3.3308
G-6413b	25.0000	4.7007	10.1749
FI-6310	39.3500	4.7969	14.7316
FI-6311	16.8700	4.6846	10.5020
TOC-204b	91.0000	15.0000	44.3443

Tabela 24 – Limites e média dos valores utilizados no treinamento das redes, para o reator 3-A.

Variável	Máximo	Mínimo	Média
PH-202	9.7000	4.8000	6.3364
TOC-202a	3090.0000	1090.0000	2110.890
PH-204a	8.8000	8.0000	8.3455
TOC-204a	74.0000	5.0000	48.0727
FI-6302	47.2540	22.7870	37.7998
FI-6305	20.2120	9.7482	16.2305
FI-6306	3.0103	0.0194	0.1384
G-6413a	47.1510	0.0400	28.1575
FI-6303	38.9600	11.3860	22.7155
FI-6304	16.7000	4.8807	9.7575
TOC-204a	96.0000	27.0000	49.1273

Tabela 25 – Limites e média dos valores utilizados no treinamento das redes, para o reator 3-B.

Variável	Máximo	Mínimo	Média
PH-202	9.7000	4.8000	6.3364
TOC-202b	3090.0000	1090.0000	2110.890
PH-204b	8.7000	8.0000	8.3564
TOC-204b	91.0000	24.0000	49.5091
FI-6307	49.0100	9.7607	24.2524
FI-6308	21.0000	0.0129	5.8556
FI-6309	23.1920	0.0100	12.3518
G-6413b	11.5300	0.0400	8.2643
FI-6310	39.3800	8.8085	17.8793
FI-6311	16.8600	7.4283	10.8533
TOC-205b	79.0000	23.0000	48.0727

Tabela 26 – Limites e média dos valores utilizados no treinamento das redes, para a lagoa.

Variável	Máximo	Mínimo	Média
vzd6311	53.6400	4.2817	35.0519
toc205a	64.0000	33.0000	44.7000
Cargaa	2291.2600	274.0315	1445.492
vzd6314	57.3300	22.4199	41.6627
toc205b	65.0000	30.0000	45.9000
Cargab	2964.0000	1098.5764	1880.314
vzesp	51.4607	11.5700	30.2103
toc209	106.0000	33.0000	59.8000
tocd1	106.0000	40.0000	61.2000

Na Tabela 26 as variáveis de entrada para a rede neuronal que representa a lagoa ao final do tratamento são: a vazão de saída do reator 3A (vzd6311), o TOC do reator 3A (toc205A), a carga de TOC correspondente a série A, a vazão de saída do reator 3B (vzd6314), o TOC do reator 3B (toc205B), a carga de TOC correspondente a série B, a vazão do espessador (vzesp) e o TOC da amostra 209 um dia antes da previsão (toc209). A variável de saída desta rede é o TOC da amostra 209 previsto (tocd1). Apresentam-se nas Figuras 66 a 75, as telas do programa para que se possa entender melhor alguns detalhes de como opera o programa criado:



Figura 66 – O Menu principal e as opções de entrada de dados

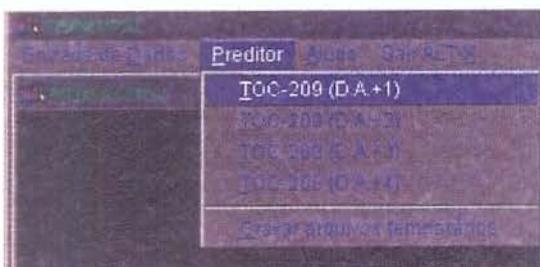


Figura 67 – As opções para a predição e saída de resultados

Entrada de dados para o reator 1A referente ao dia 0

Varíavel	Valor
1) pH-202	4.800
2) TOC-202	2268.000
3) pH-203-A	8.200
4) TOC-203-A	19.000
5) MES-203-A	5200.000
6) MES-203-A	7520.000
7) FI-6302	34.674
8) FI-6401	34.701
9) FI-6303	44.991

D.A. -

Figura 68 - Entrada de dados para o reator 1A referente ao dia 0.

Entrada de dados para o reator 1B referente ao dia 0

Varíavel	Valor
1) pH-202	4.800
2) TOC-202	2268.000
3) pH-203-B	7.900
4) TOC-203-B	19.000
5) MES-203-B	4585.000
6) MES-206-B	7800.000
7) FI-6307	14.793
8) FI-6402	17.831
9) FI-6310	8.990

D.A. -

Figura 69 - Entrada de dados para o reator 1B referente ao dia 0.

Varíavel	Valor	Varíavel	Valor
1) pH-202	4.800	8) MES-204-A	5070.000
2) TOC-202	2268.000	9) MES-205-A	7520.000
3) pH-203-A	8.200	10) FI-5302	34.674
4) TOC-203-A	19.000	11) FI-5305	14.851
5) MES-203-A	5200.000	12) FI-5401	34.701
6) pH-204-A	8.300	13) FI-5303	44.991
7) TOC-204-A	20.000	14) FI-5304	22.289

Figura 70 - Entrada de dados para o reator 2A referente ao dia 0. Para a entrada do dia 1 a tela de diálogo é a mesma, mudando somente o campo correspondente.

Varíavel	Valor	Varíavel	Valor
1) pH-202	4.800	8) MES-204-B	4750.000
2) TOC-202	2268.000	9) MES-205-B	7800.000
3) pH-203-B	7.900	10) FI-5307	14.793
4) TOC-203-B	19.000	11) FI-5306	0.011
5) MES-203-B	4585.000	12) FI-5402	17.831
6) pH-204-B	7.900	13) FI-5310	8.990
7) TOC-204-B	22.000	14) FI-5311	8.987

Figura 71 - Entrada de dados para o reator 2B referente ao dia 0. Para a entrada do dia 1 a tela de diálogo é a mesma, mudando somente o campo correspondente.

Modelo para o reator 3A - Entrada de dados - 11/05/2011

Varivel	Valor	Dia	D.A.
1) pH-202	4.800	0	0
2) T.DC-202	2268.000		
3) pH-204A	8.300		
4) T.DC-204A	20.000		
5) FI-6302	34.674		
6) FI-6305	14.851		
7) FI-6308	0.015		
8) FI-640	34.701		
9) FI-6303	44.991		
10) FI-6304	22.289		

OK

Cancel

Figura 72 - Entrada de dados para o reator 3A referente ao dia 0. Para a entrada dos dias 1 e 2 a tela de diálogo é a mesma, mudando somente o campo correspondente.

Forma de dados simulados da Lagoa

Varável	Valor
1) pH-202	4.800
3) TQC-202	2268.000
3) pH-204 B	7.900
4) TQC-204 B	22.000
8) FI-6307	14.793
8) FI-6308	0.011
7) FI-6309	41.608
8) FI-6402	17.831
9) FI-6310	8.990
10) FI-6311	8.987

Dia =

Figura 73 - Entrada de dados para o reator 3B referente ao dia 0. Para a entrada dos dias 1 e 2 a tela de diálogo é a mesma, mudando somente o campo correspondente.

Forma de dados simulados da Lagoa

Varável	Valor	Varável	Valor
1) Velocidade de Saída do Decantador da Trança A	14.839	FI-6302	34.674
2) TQC-205A	0.000	FI-6305	14.851
3) Carga de TQC da Trança A	0.000	FI-6306	0.015
4) Velocidade de Saída do Decantador da Trança B	38.581	FI-6401	34.701
5) TQC-205B	0.000	FI-6307	14.793
6) Carga de TQC da Trança B	0.000	FI-6308	0.011
7) Velocidade para escoamento	52.582	FI-6309	41.608
8) TQC-203	23.000	FI-6402	17.831

Dia =

Figura 74 - Entrada de dados para a lagoa referente ao dia 0. Para a entrada dos dias 1, 2 e 3 a tela de diálogo é a mesma, mudando somente o campo correspondente.

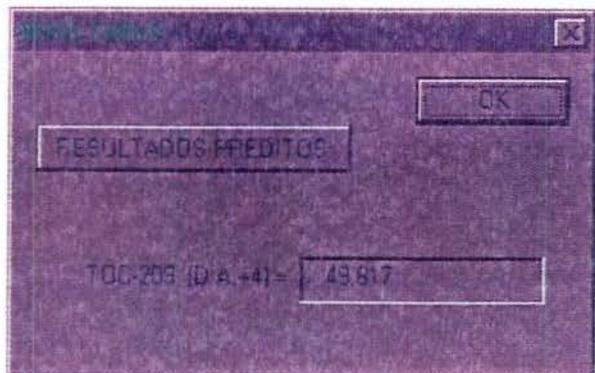


Figura 75 - Resultado para a predição de quatro dias. As predições para 1, 2 e 3 dias geram janela idêntica.

Para testar o programa, selecionou-se do banco de dados os conjuntos que continham um número mínimo de informações. Assim foram criados 11 arquivos contendo os casos onde não faltavam as principais informações. Esses arquivos estão no formato próprio para serem lidos pelo SIMNEURAL. Em diversos destes casos ainda faltava algum tipo de dado, contudo decidiu-se por estimar os valores que faltam. De fato, somente em um destes casos é que se tem todas as informações necessárias ao programa, o referente ao dia 161. Os dados referentes aos dias 402 e 416 são de certa forma muito importantes, já que são dados não utilizados anteriormente no treinamento ou no teste das redes. Na maioria dos casos faltavam dados a respeito do TOC da amostra 205 ou da 209 no dia da D.A., ou ainda da concentração de sólidos (MES) de algum dos reatores. Os resultados encontrados para os 11 casos, para o TOC da amostra 209 com um intervalo de 4 dias contam da Tabela 27.

4.6.1. Conclusões Finais sobre o Estudo na Unidade de Tratamento de Efluentes

Neste trabalho demonstra-se a capacidade que uma rede treinada com os devidos cuidados pode oferecer para simular uma unidade através de dados reais coletados. Considera-se que o grau de predição dos valores pelas redes treinadas satisfaz aos propósitos deste trabalho, e estas redes são instrumentos que permitem uma análise de situações que podem ocorrer na realidade. Podem ainda serem usadas para o treinamento de operadores, aumentar o grau de conhecimento sobre o processo, ou indicar ações para a otimização da planta.

Tabela 27 – Valores medidos e preditos pela rede neuronal para o TOC da amostra 209 com 4 dias de antecedência.

Ponto/ dia:	Valor predito:	Valor esperado:
159	92,09	84,00
160	81,60	106,00
161	77,23	79,00
164	71,77	60,00
166	71,98	55,00
167	79,21	77,00
174	67,06	44,00
307	57,80	40,00
335	55,14	42,00
402	63,35	68,00
416	62,74	79,00

Contudo, analisando o que ocorreu nos dois períodos estudados, para aumentar o grau de confiança na predição é recomendável se fazer o treinamento das redes periodicamente, ou quando se perceber que as condições na operação revelam padrões novos a serem considerados.

Cabe ressaltar ainda as principais dificuldades enfrentadas que acabaram por influir sobre muitas das decisões sobre a abordagem do problema foram:

- a) Os dados do laboratório são amostras simples e em horários ímpares. Já os dados do SDCD são valores médios por turno, onde os turnos terminam em horas pares. Admitiu-se que o sistema estaria em regime estacionário a ponto de não haver grandes diferenças durante este tempo.

- b) A periodicidade das amostras está voltada para uma adequação operacional. Assim, seria o correto ter as amostragens sincronizadas com os tempos de residência em cada equipamento, considerando portanto a vazão.
- c) Algumas análises são realizadas com uma periodicidade diferente das outras. Isto provocou a eliminação de diversos dados, pois na falta de uma única variável o ponto não pode ser utilizado.
- d) Definir quando as variações que ocorrem não influem no resultado esperado.
- e) Tratamento dos dados pelo excell. Tarefas repetidas e sujeitas a erros. Conferência freqüente.

A principal sugestão seria de que as medidas de laboratório, tanto para a unidade de tratamento de efluentes como para qualquer outra, deveriam ser planejadas não somente observado o caráter do controle operacional, mas também poderia existir uma relação causa e efeito entre os dados que fossem coletados. Isto seria muito útil na aplicação futura de estudos no gênero deste, bem como da implantação de sistemas inteligentes mais complexos de controle.

CAPÍTULO 5. DISCUSSÃO GERAL DOS RESULTADOS, CONCLUSÕES E SUGESTÕES

5.1. DISCUSSÃO GERAL DOS RESULTADOS

Para avaliar a utilização de redes neurais a processos da engenharia química se procurou fazer parcerias com indústrias que tivessem um interesse pelo assunto e que oferecessem condições para a realização do mesmo.

Retomando o histórico das atividades desenvolvidas, na época da escolha do tema de tese, alguns contatos já haviam sido feitos com a Ripasa SA, que estava interessada na potencialidade do uso de redes neurais. A princípio, a idéia era de procurar resolver um problema ligado à lagoa do tratamento de efluentes. Uma avaliação inicial do problema apontava para a necessidade de obter um número maior de informações experimentais. A empresa, no entanto, não poderia fornecer todos os dados solicitados e se decidiu passar este trabalho para alunos de iniciação científica que, inclusive, fariam as análises que iriam faltar. Enfim, com mais dados, o trabalho dos alunos apresentou redes que apresentaram um resultado razoável, e que posteriormente foi melhorado através de um trabalho de tese de mestrado (OLIVEIRA, 2000).

Paralelamente ao estudo na lagoa de tratamento, surgiu o interesse em utilizar redes para classificação da matéria-prima da fabricação de papel e celulose, a qual se entendia estar sendo feita de uma maneira muito simplificada. Para sofisticar esta classificação a empresa ofereceu dados sobre um estudo que havia sido feito com a obtenção em unidade piloto das propriedades da massa após a polpação. Em virtude das dificuldades inerentes a obtenção destas propriedades importantes para a classificação propôs-se o uso de redes neurais para a sua predição a partir de dados que fossem facilmente obtidos. Desta forma, resolveu-se o problema da classificação em duas etapas, a primeira usando uma rede do tipo “feedforward” com o algoritmo “backpropagation” que estima o valor de 4 variáveis que caracterizam as propriedades da massa polpada, a partir de 9 variáveis de entrada. E na segunda etapa utilizou-se uma rede do tipo “Linear Vector Quantization” (LVQ) para estabelecer a classe do lote da madeira. Os resultados obtidos

foram bastante satisfatórios, já que o erro nas redes era inferior aos erros de amostragem e medição.

Devido a estes resultados alcançados para o caso da Ripasa SA, havia a possibilidade de testar a técnica em outra situação. Outros processos dentro da própria Ripasa foram considerados, mas concluiu-se que os dados disponíveis para estes não seriam suficientes para o estudo com redes. Assim, nova parceria surgiu com a Rhodiaco Ltda. que também estava bastante interessada em começar a usar esta ferramenta para solucionar alguns problemas. A diferença básica entre as duas empresas era de que muitos dos dados que nos colocaram a disposição vinham de um Sistema Digital de Controle Distribuído (SDCD) e também de análises laboratoriais. No primeiro contato feito, ficou claro que o interesse era avaliar esta grande quantidade de dados de forma a concluir se a unidade estaria trabalhando em condições otimizadas.

Optou-se, então, por aplicar redes neurais à unidade de tratamento de efluentes. Após algumas adequações nos dados, percebeu-se que o número de dados fornecidos não era suficiente pois a menor periodicidade de algumas análises laboratoriais acabava comprometendo um bom número dos dados armazenados no SDCD. Desta maneira, as redes foram sendo treinadas e aprimoradas a medida que novos dados surgiam. Percebeu-se, principalmente no primeiro reator, que a unidade teve uma mudança bastante acentuada em determinado momento, passando a eliminar muito mais a carga orgânica. O treinamento com todos os dados provou ser muito ineficiente e mesmo uma análise grosseira dos dados de entrada permitia prever que a rede não teria como explicar porque os valores das saídas reduziram tanto. Assim, decidiu-se treinar uma rede para cada um desses períodos, de TOC de saída alto e de TOC de saída baixo. Treinando as redes desta forma o resultado foi mais satisfatório, porém evidencia-se que seria necessário usar mais dados em que ocorressem estas mudanças na unidade e envolver outras variáveis que pudessem estar relacionadas com os efeitos destas mudanças para obter um modelo ainda mais confiável. Com isto, não foi possível aprofundar o estudo para a influência do lodo no processo, nem prever a atividade do lodo frente a outras variáveis tais como concentração de nutrientes e de oxigênio.

Algumas propostas de abordagem envolvendo mais variáveis tiveram que ser abandonadas porque ao incluí-las o número de dados para o treinamento das redes reduzia muito. Em função desta quantidade de dados disponíveis para cada um dos reatores é que se

decidiu treinar redes separadamente para cada um deles. Com isto, foram treinadas 7 redes, uma para cada reator e outra que prediz o TOC final em função do efluente de cada uma das séries. Uma análise de sensibilidade permitiu escolher quais variáveis deviam ser usadas, mas também foi considerado o número correspondente de dados que se poderia ter em cada caso. O resultado encontrado nestas condições foi bastante satisfatório, mostrando novamente o potencial do emprego das redes neuronais na simulação desta unidade.

Observa-se que em todos os casos estudados o número de dados foi citado como um fator importante para se obter um bom resultado com redes neuronais. Obviamente não é somente a quantidade de informações que põe em risco o sucesso do emprego desta técnica, o grau de confiabilidade da rede treinada depende essencialmente da qualidade das informações processadas. Para demonstrar este fato, pode-se avaliar os quatro casos que foram estudados e fazer uma correlação direta entre o resultado obtido e a qualidade dos dados utilizados. Assim, sempre obteve-se melhores resultados quando os dados eram mais representativos do processo. Os casos estudados e as principais características e resultados referem-se ao treinamento de redes neuronais para dados obtidos do:

5.1.1. Simulador para unidades de tratamento de efluentes

Primeiramente, executou-se um teste, uma primeira experiência com redes neuronais, onde se gerou um bom número de dados com o simulador Superpro II, para um sistema de tratamento de efluentes assemelhado ao da Ripasa SA. Os dados utilizados, portanto, não possuíam nenhum ruído, e ao treinar redes para este caso obteve-se um índice de correlação da ordem de 0,99.

5.1.2. Processo de Tratamento de Efluentes da Ripasa SA

Depois trabalhou-se com os dados da lagoa da Ripasa, os quais envolviam fatores complicantes, tais como o tempo de residência indefinido e o padrão complexo de fluxo na lagoa que sugere a retenção de parte do material por um tempo maior que o médio. As amostras neste caso são compostas, o que torna os dados um pouco mais representativos, mas em algumas análises de laboratório podem ser observados erros de leitura consideráveis. Trabalhando estes dados conseguiu-se obter uma rede onde o índice de correlação ficava abaixo de 0,9.

5.1.3. Processo de polpação em unidade piloto da madeira (Ripasa SA)

Partindo para o estudo da matéria prima para o processo de fabricação de papel e celulose, abordado com maiores detalhes no capítulo 3, verificou-se que os dados utilizados são bastante confiáveis. Além do fato de ser possível estimar os erros envolvidos na amostragem e leitura através da replicação dos experimentos devidamente realizada, percebe-se que houve um planejamento experimental para a coleta das amostras. Com isto, os resultados obtidos giraram em torno de 0,98.

5.1.4. Processo de tratamento de efluentes da Rhodiaco Ltda.

E, finalmente, para o caso deste processo, tratado com maiores detalhes no capítulo 4, pode se considerar que os dados armazenados pelo SDCD são bastante confiáveis, com exceção da concentração de oxigênio, e os recursos para análises laboratoriais são bastante apropriados. Mesmo assim, devido ao problema na periodicidade das análises já mencionado, alguns efeitos sobre o lodo não puderam ser levados em consideração, e o índice de correlação encontrado para as redes treinadas variou entre 0,97 e 0,99.

5.2. CONCLUSÕES

A aplicação de redes neuronais ao controle de variáveis de processo tem se tornado bastante evidente e tudo indica que em um futuro próximo ela estará bem difundida, mas são poucos trabalhos que a utilizam para a simulação de processos. Ao escolher este tema para desenvolver a tese de doutorado pretendia-se mostrar se a utilização de redes neuronais pode ser uma alternativa para simular alguns processos da engenharia química. A intenção é a de recorrer a uma ferramenta que permite obter resultados a partir de um bom conjunto de dados, mas somente para aqueles casos onde um modelo teórico, com base fenomenológica, não pode ser resolvido com as técnicas e recursos computacionais atuais. Existem ainda os casos onde o modelo exige dados de composição e de propriedades físicas, coeficientes cinéticos e de transferência que são muito difíceis ou onerosos de serem obtidos.

No caso estudado no Capítulo 3, a classificação da matéria prima para o processo de fabricação de papel e celulose, a aplicação de redes neuronais se justificou pelos resultados encontrados. Ela ajudou a aprimorar o sistema de classificação, que depois de devidamente implementado na operação de mistura dos cavacos que alimentam o processo, poderá diminuir as variações nas propriedades da polpa, trazendo vantagens para a qualidade do produto final obtido.

Para o sistema de tratamento de efluentes da Rhodiaco Ltda, que foi estudado no Capítulo 4, foi necessário treinar um número maior de redes de forma a aproveitar melhor as informações disponíveis e isto garantiu um resultado considerado como satisfatório. Contudo, devido a complexidade no sistema, seria necessário aprimorar ainda mais o banco de dados, tanto no que se refere à periodicidade das análises como à inclusão de variáveis a serem consideradas.

Conclui-se que o número de dados e o seu grau de representabilidade do processo são fatores essenciais para se obter um bom modelo com redes neuronais. Assim, sempre obteve-se os melhores resultados quando os dados estavam adequadamente medidos. Esta relação é importante para quem pretende aplicar redes neuronais a processos reais.

Conclui-se, que de forma geral, a utilização de redes neuronais a processos químicos é bastante satisfatória se forem observadas considerações como as que foram expostas. Ainda assim, seria muito apropriado se a adequação da periodicidade das análises de forma a poder se correlacionar os efeitos observados com as suas respectivas causas fosse feito desde já nas indústrias, obviamente sem prejudicar a preocupação habitual do monitoramento do processo. Com certeza isto facilitaria sobremaneira trabalhos como este, onde se deseja utilizar de dados históricos para se tirar conclusões sobre o processo.

Enfim, espera-se que este trabalho possa contribuir para esclarecer as potencialidades e os problemas na utilização desta técnica. Acredita-se que em breve se poderá empregar modelos neuronais de maneira acoplada aos principais simuladores comerciais que são utilizados na área da engenharia química. A preocupação é de alertar para que o uso destes modelos com redes neuronais não seja feita de forma indiscriminada e descuidada, diante da facilidade que se colocará nas mãos do usuário.

Algumas das técnicas usadas com sucesso para aprimorar o treinamento das redes no caso estudado no Capítulo 3, não surtiram o mesmo efeito no caso seguinte, estudado no

Capítulo 4. Isto demonstra que não existe uma receita única a seguir para se encontrar um modelo neuronal perfeito. O assunto, portanto, deve render ainda muitos estudos neste sentido. Mesmo assim, conclui-se que desde que o usuário tenha conhecimentos sobre o processo e saiba como aprimorar o treinamento da rede neuronal, o uso de redes, nos casos mais críticos é uma alternativa viável.

De maneira a poder automatizar o processo de geração de redes neuronais para qualquer processo será necessário um maior número de pesquisas sobre o assunto. As redes ainda necessitam da supervisão de alguém com experiência para que o treinamento alcance bons resultados. O processo permanente de validação e atualização das redes treinadas com a inclusão dos dados mais recentes ainda é incipiente e carece de algoritmos genéricos que permitam a geração segura de novas redes. A partir do conjunto de pontos e definidas as entradas e saídas desejadas, o caminho para se obter uma rede satisfatória por vezes oferece alguma dificuldade.

5.3. SUGESTÕES PARA PROSSEGUIMENTO DO TRABALHO:

Como sugestões ao prosseguimento deste trabalho sugere-se avaliar outros processos com características diferentes dos já estudados a fim de confirmar estas impressões sobre a aplicação de redes. Também é importante que se faça o acompanhamento do nível da predição para os dados mais recentes a fim de verificar a validade da rede anteriormente treinada, frente a possíveis modificações que possam ter ocorrido.

Para os processos estudados neste trabalho seria interessante reestruturar a periodicidade das análises de forma a garantir um maior aproveitamento das análises e garantindo a correlação entre causa e efeito no processo.

Considerando o tratamento de efluentes da Rhodiaco Ltda, para aumentar o grau de confiabilidade nas predições é necessário ainda evoluir para um modelo que considere adicionalmente efeitos sobre o lodo. Para isto deve se considerar um conjunto maior de variáveis que não puderam ser utilizadas neste trabalho por diversos motivos, tais como o problema na periodicidade das medidas e problemas nos sensores de oxigênio dissolvido. É possível que desta forma o modelo neuronal seja mais robusto frente a modificações

bruscas que afetem o desempenho do lodo, desde que se possa analisar ocorrências destas modificações.

Relativamente à dificuldade em obter bons resultados quando se tentou treinar a rede incluindo dados novos do sistema de tratamento de efluentes (Capítulo 4) que continham valores de saída bem inferiores aos anteriores. Percebeu-se que a rede não conseguia ajustar-se adequadamente a estes pontos por causa da forma como o erro gerado é minimizado. Os valores mais altos da variável de saída levam vantagem pois o erro nestes pontos gera um erro global maior enquanto que nos menores valores o erro absoluto é pequeno frente aos anteriores, mas o erro relativo nestes casos torna-se relevante. Assim sugere-se estudar algoritmos que minimizem o erro relativo ao invés de considerar somente o erro absoluto entre o valor desejado e o predito pela rede.

Referências Bibliográficas

- ALBEVERIO, S., FEN, G., QIAN, M. Role of noises in neural networks. **Physical Review E**, v. 52, n. 6, p. 6593-6606, 1995.
- BARATTI, R.F., VACCA, G., SERVIDA, A., Neural Network Modeling of Distillation Columns, **Hydrocarbon Processing**, 35-38, June 1995.
- BARROS NETO, B., SCARMINO, I. S., BRUNS, R.E. **Planejamento e Otimização de Experimentos**. 2a. ed. Editora da Unicamp, 1996.
- BAUGHMAN, D.R., LIU, Y.A., **Neural networks in bioprocessing and chemical engineering**, Academic Press, 1995.
- BECRAFT, W.R., LEE, P.L., An integrated neural network / expert system approach for fault diagnosis, **Computers & Chemical Engineering**, 17(10), 1001-1014, 1993.
- BHAGAT,P. An Introduction to Neural Nets. **Chem. Eng. Progress**, Aug. 1990, pag. 55-60.
- BHAT, N.V., McAVOY, T., Use of neural nets for dynamic modeling and control of chemical process systems, **Computers & Chemical Engineering**, v.14(4/5), p. 573-583, 1990.
- BHAT, N.V., McAVOY, T., Determining model structure for neural models by network stripping, **Computers & Chemical Engineering**, v.16(4), p. 271-281, 1992.
- BISHOP, C.M., Training with noise is equivalent to Tikhonov regularization, **Neural Computation**, 7, 108-116, 1995.
- BULSARI, A.B. *Neural Networks for Chemical Engineers*. Amsterdam: Elsevier, 1995, 680 p.
- CAJAL, S.R., **Die Retine der Wirbelthiere**, Weisbaden, J.F. Gergman, 1894.
- CAUDILL, M. Neural Network Training Tips and Techniques. **AI Expert** . Jan, 1991.
- CHENG, B., TITTERINGTON, D.M., Neural Networks: a Review from a Statistical Perspective, **Statistical Science**, vol. 9, No. 1, p. 2-54, 1994.

- COLLASCIOPO, R. Apostila do Curso de Planejamento de Experimentos Avançado. Unicamp, 1997.
- DAYHOFF, J.E., *Neural Network Architectures – An Introduction*, Van Nostrand Reinhold, New York, 1990.
- GONTARSKI, C.A., MORI, M., PEREIRA, A.M., PERIOTO, F.R., TELHADA, M.A.L. Utilização de Redes Neurais na Avaliação do Processo de Polpação da Madeira para Fabricação de Papel. Anais do 12^o Congresso Brasileiro de Engenharia Química, trab. n^o 451, 1998, Porto Alegre. **Proceedings ABEQ**, 1998a. (CDROM)
- GONTARSKI, C.A., MORI, M., PEREIRA, A.M., PERIOTO, F.R., TELHADA, M.A.L. The Use of Neural Networks for the Prediction of Wood Properties for Pulp Digesters, In International Symposium on the Use of Computers in Engineering and the Quality of the Software, Rio de Janeiro, 9/nov/1998. **Proceedings**. 1998b.
- GONTARSKI, C.A., CASTRO, C., SOUZA, C., FERRARI, L., VITTI, J., MORI, M., PEREIRA, A.M. Avaliação do processo de Tratamento de Resíduos Industriais de Papel e Celulose com Redes Neurais. In XVIII Congresso Interamericano de Ingeniería Química, San Juan, Porto Rico, 07/dez/98. **Proceedings**. 1998c.
- GONTARSKI, C. A., MORI, M., BONIFÁCIO, W. Predição de Comportamento de um Sistema de Tratamento de Efluentes Industriais Utilizando Redes Neurais. In II Congresso de Engenharia de Processos do **Mercosul**, Florianópolis-SC, 02/set/99. **Proceedings**. 1999a.
- GONTARSKI, C. A., MORI, M., BONIFÁCIO, W. Predição das Propriedades do Efluente Líquido de uma Unidade de Tratamento de Esgotos Industriais Utilizando Redes Neurais. In IV Congresso Interamericano de Computación Aplicada a la Indústria de Procesos. San José, Costa Rica, 02/nov/99. **Proceedings**. 1999b.
- GONTARSKI, C.A. , BONIFÁCIO, W., MORI, M., Avaliação do Comportamento de um Sistema de Tratamento de Efluentes usando Redes Neurais Artificiais. In 1^o Encontro EPEFEQ. UNICAMP, Campinas, Brasil, 1999. **Abstracts**. 1999c.
- GONTARSKI, C.A. , RODRIGUES, P.R., MORI, M., PRENEM, L.F. Liquid Effluent Properties Prediction from an Industrial Wastewater Treatment Plant Using Artificial Neural Networks. In Escape 10, Florença, Itália, 2000. **Proceedings**. 2000a.

- GONTARSKI, C.A. , RODRIGUES, P.R., MORI, M., PRENEM, L.F. Simulation of an Industrial Wastewater Treatment Plant Using Artificial Neural Networks. PSE2000, Colorado, USA, 2000. Proceedings. 2000b.
- GONTARSKI, C.A. , RODRIGUES, P.R., MORI, M., PRENEM, L.F. Simulation of an Industrial Wastewater Treatment Plant Using Artificial Neural Networks. Computers and Chemical Engineering, v.24, p. 1719-1723. July 2000c.
- GONTARSKI, C.A. , RODRIGUES, P.R., MORI, M., PRENEM, L.F. Avaliação das Condições de Operação de um Sistema de Tratamento de Efluentes com auxílio de Redes Neurais. In 13^o Congresso Brasileiro de Engenharia Química. São Pedro, Brasil, 2000. Proceedings. 2000d.
- GRANDVALET, Y., CANU, S., Comments on "noise injection int inputs in backpropagation learning", *IEEE Transactions on Systems, Man and Cybernetics*, 25(4), 678-681, 1995.
- HAYAKAWA, Y., MARUMOTO, A., SAWADA, Y., Effects of the caotic noise on the performance of a neural network model for optimization problems, *Physical Review E.*, v.51, n.4, 51, april 1995.
- HAYKIN, S., *Neural Networks*, Macmillan College Publishing Company, New York, 1994.
- HIMMELBLAU, D.M., KARJALA, T.W., Rectification of Data in a Dynamic Process using Artificial Neural Networks, *Computers & Chemical Engineering*, v.20(6/7), p. 805-812, 1996.
- HOSKINS, J.C., HIMMELBLAU, D.M., Process Control via Artificial Neural Networks and Reinforced Learning, *Computers & Chemical Engineering*, 16, 4, 241-251, 1992.
- HUNT, K.J., SBARBARO, D., ZBIKOWSKI, R., GAWTHROP, P.J., Neural Networks for Control Systems- A Survey, *Automatica*, vol. 28, n. 6, p. 1083-1112, Great Britain, 1992.
- JACOBS, R.A. Increased Rates of Convergence Through Learning Rate Adaptation. *Neural Networks*, v.1, p. 295-307, 1988.

- KOHONEN, T, et al., Statistical Pattern Recognition with Neural Networks: Benchmark Studies. Proceedings of the Second Annual IEEE International Conference on Neural Networks. V. 1, 1988.
- KOVÁCS, Z.L. **Redes Neurais Artificiais: Fundamentos e Aplicações**. São Paulo: Edição Acadêmica, 1996.
- LEE, J.H., DATTA, A.K. Nonlinear Inferential Control of Pulp Digesters. **AICHE Journal** , v.40 n. 1, Jan. 1994.
- MacMURRAY, J.C., Himmelblau, D.M., Modeling and Control of a Packed Distillation Column Using Artificial Neural Networks, **Computers Chem. Engng.**, 19(10), 1077-1088, 1995.
- MASSIMO, C.D., MONTAGUE, G.A., WILLIS, M.J., THAM, M.T., MORRIS, A.J., Towards improved penicilin fermentation via artificial neural networks, **Computers Chem. Engng.**, v.16(4), 283-291, 1992.
- McCULLOCH, W., PITTS, W., A Logical Calculus of the Ideas Immanent in Nervous Activity, **Bulletin of Mathematical Biophysics**, 1943.
- METCALF & EDDY, Inc., **Wastewater Engineering: treatment, disposal and reuse**. 3rd ed. New York, McGraw-Hill, 1991.
- MINSKY, M.L. and PAPERT, S.A. **Perceptrons**, MIT Press, 1969.
- MONTGOMERY, D.C. **Design and Analysis of Experiments**, 3rd ed. Wiley International Edition, 1991.
- MORRIS, A.J., MONTAGUE, G.A., WILLIS, M.J., Artificial Neural Networks: Studies in Process Modelling and Control, **Trans. I. Chem. E.**, vol. 72, part A, 3-19, 1994.
- NASCIMENTO, C.A.O., YAMAMOTO, C.I. Modelagem de Processos Químicos via Redes Neurais. Em Anais – Workshop em Aplicações de Redes Neurais, 29-30 nov., São Paulo. **Proceedings**. 1993. p. 1-13.
- NELSON, M.M., ILLINGWORTH, W.T., **A practical Guide to Neural Nets**, Addison-Wesley Publishing Company, 1991.
- NEURALWARE, Inc. **Neural Computing - A Technology Handbook for Professional II/PLUS**. Technical Publications Group. USA. 1995.

- OLIVEIRA, K.P.S. **Aplicação da técnica de Redes Neurais e de Análise de Componentes Principais na Modelagem de uma Lagoa Aerada da RIPASA S/A.** Campinas, Faculdade de Engenharia Química, Unicamp, Fevereiro 2000, 64 p. Tese (Mestrado).
- PEARLMUTTER, B.A., **Dynamic Recurrent Neural Networks**, Carnegie Mellon University, 1990.
- PEARLMUTTER, B.A., Learning State Space Trajectories in Recurrent Neural Networks, **Neural Computation**, 1, 263-269, 1989.
- PESSÔA, C. A., e JORDÃO, E. P., **Tratamento de esgotos domésticos**. 2 ed., v.1, Rio de Janeiro, Associação Brasileira de Engenharia Sanitária e Ambiental, 1982.
- PINEDA, F.J., Generalization of Backpropagation to Recurrent Neural Networks, **Physical Review Letters**, vol. 50(19), 2229-2232, 1987.
- QIAN, Y., TESSIER, P.J.C., Modelling of a Woodchip Refiner Using Artificial Neural Network, **Chem. Eng. Technol.**, v.18, 337-342, 1995.
- RIPASA, Manual de procedimentos - Polpação da madeira - RIPASA SA
- ROSENBLATT, R., **Principles of Neurodynamics**, Spartan Books, New York, 1962.
- RUMELHART, D.E., McCLELLAND, J.L., editors, **Parallel Distributed Processing: Explorations in the Microstructure of Cognition**. V. 1, Foundations. Mit Press, 1986.
- TANABE, A., YAMAMOTO, T., SHINFUKU, K., NAKAMAE, T., Development of Neural Network for Analysis of Local Power Distributions in BWR Fuel Bundles, **Journal of Nuclear Sci. and Tech.**, 30(8), 804-812, 1993.
- TETKO, I.V., LIVINGSTONE, D.J., LUIK, A.I., Neural networks studies. 1. Comparison of overfitting and overtraining, **J. Chem. Inf. Comput. Sci.**, 35, 826-833, 1995.
- THIBAUT, J., BREUSEGEM, V.V., CHÉURY, A., On-line prediction of fermentation variables using neural networks, **Biotechnology and Bioengineering**, 36, 1041-1048, 1990.

UNICAMP

BIBLIOTECA CENTRAL

SEÇÃO CIRCULANTE

- VEAUX, R.D.D., PSICHOGIOS, D.C., UNGAR, L.H., A comparison of two nonparametric estimation schemes: MARS and neural networks, **Computers Chem. Engng.** v.17(8), 819-837, 1993.
- WALT, T.J., DEVENTER, J.S.J., BARNARD, E., The Estimation of Kinematic Viscosity of Petroleum Crude Oils and Fractions with a Neural Net, **The Chem. Engng. Journal**, 51, 151-158, 1993.
- WILLIS, M.J., MONTAGUE, G.A., MASSIMO, C., THAM, M.T., MORRIS, A.J., Artificial Neural Networks in Process Estimation Control, **Automatica**, v.28, n.6, p. 1181-1187, 1992.
- YAMAMOTO, C.I. **Modelagem Matemática e Otimização do Processo Industrial de Síntese de Amônia Utilizando Redes Neurais.** São Paulo, Escola Politécnica da Universidade de São Paulo, 1998, 230 p. Tese (Doutorado).
- YAMASAKI, K., OGAWA, H., A theory of over-learning in the presence of noise, **IEEE**, 266-271, 1993.
- YANG, G.B.O., Managing secondary treatment systems, **TAPPI Journal**, v. 52-54, 1996.
- YDSTIE, B.E., Forecasting and Control Using Adaptative Connectionist Networks, **Computers Chem. Engng.**, v.14, n. 4/5, 583-599, 1990.

Anexos

Tabela 28 – Dados fornecidos pela RIPASA SA, para o estudo da classificação da matéria-prima

Fazenda	Talhão	Espécie	Idade	Produtividade	Db	ET	LS	LI	H	RD	NK	V	A.A.R.	Classe Esp. A	Classe Esp. B
flecha azul	144	E. híbrido	84.0	55.7	0.536	6.97	1.93	23.66	67.44	50.21	16.5	44.2	5.30	D	C
flecha azul	144	E. híbrido	84.0	55.7	0.522	4.21	1.95	23.87	69.97	49.77	17.2	44.5	5.30	D	D
flecha azul	144	E. híbrido	84.0	55.7	0.550	4.65	1.98	24.61	68.76	49.37	18.6	53.8	4.70	D	D
flecha azul	139	E. grandis	80.5	62.6	0.439	5.74	1.77	24.95	67.54	51.37	15.3	42.8	6.60	B	B
flecha azul	139	E. grandis	80.5	62.6	0.463	7.38	1.75	24.60	66.27	50.67	15.5	42.6	5.60	B	B
flecha azul	139	E. grandis	80.5	62.6	0.458	5.93	1.86	23.58	68.63	51.60	16.4	48.2	5.60	B	B
flecha azul	134	E. grandis	84.0	47.3	0.504	4.29	1.65	24.34	69.72	51.95	15.4	42.3	6.80	B	B
flecha azul	134	E. grandis	84.0	47.3	0.462	5.08	1.50	25.32	68.10	51.25	15.5	48.5	6.30	B	B
flecha azul	134	E. grandis	84.0	47.3	0.509	5.19	1.48	25.17	68.16	52.05	16.9	43.2	5.80	C	C
flecha azul	267	E. grandis	75.0	61.5	0.499	1.65	1.78	26.44	70.13	50.71	16.3	40.9	6.30	C	C
flecha azul	267	E. grandis	75.0	61.5	0.468	5.52	2.02	24.72	67.74	52.53	16.3	43.4	6.20	B	B
flecha azul	267	E. grandis	75.0	61.5	0.508	5.13	1.91	24.50	68.46	52.57	16.2	41.2	6.00	C	C
flecha azul	273	E. grandis	71.5	54.7	0.499	3.71	1.75	25.62	68.92	51.84	16.4	41.4	5.90	C	C
flecha azul	273	E. grandis	67.5	54.7	0.493	4.58	1.80	25.62	68.00	52.67	16.2	40.9	5.90	C	C
flecha azul	273	E. grandis	71.7	54.7	0.488	3.76	1.87	26.07	68.30	51.65	17.0	42.6	5.60	C	C
flecha azul	270	E. grandis	72.4	48.8	0.499	4.34	1.75	26.43	67.48	51.17	16.4	44.3	5.50	C	C
flecha azul	270	E. grandis	72.4	48.8	0.461	4.84	1.84	26.48	66.84	51.97	16.3	44.2	5.60	B	B
flecha azul	270	E. grandis	72.5	48.8	0.468	4.74	1.71	26.71	66.84	52.60	16.1	43.4	5.80	B	B
flecha azul	280	E. urophylla	73.5	45.2	0.529	6.65	1.60	26.59	65.16	48.67	18.3	50.7	3.40	D	D
flecha azul	280	E. urophylla	73.5	45.2	0.541	7.70	1.64	26.49	64.17	50.59	18.3	51.3	3.30	D	D
flecha azul	280	E. urophylla	73.6	45.2	0.537	7.22	1.65	25.34	65.78	50.52	18.7	50.9	3.50	D	D
flecha azul	271	E. urophylla	71.3	50.6	0.522	7.56	1.54	23.97	66.93	49.25	17.3	52.5	4.00	D	D
flecha azul	271	E. urophylla	71.3	50.6	0.506	6.58	1.48	23.83	68.11	49.57	16.9	53.7	4.10	D	D

flecha azul	271	<i>E. urophylla</i>	71.3	50.6	0.508	7.35	1.48	23.52	67.65	49.90	15.7	46.2	4.20	D	C
flecha azul	142	<i>E. urophylla</i>	60.8	43.6	0.557	6.14	1.91	23.86	68.09	50.40	15.8	44.0	4.00	C	C
flecha azul	142	<i>E. urophylla</i>	60.8	43.6	0.538	5.96	1.50	23.90	68.64	49.90	17.2	53.5	4.50	D	D
Fazenda	Talhão	Espécie	Idade	Produtividade	Db	ET	LS	LI	H	RD	NK	V	A.A.R.	Classe Esp. A	Classe Esp. B
flecha azul	142	<i>E. urophylla</i>	60.8	43.6	0.520	6.17	1.64	24.01	68.18	49.80	18.8	53.3	3.90	D	D
Fortaleza	32	<i>E. grandis</i>	73.0	40.8	0.499	3.69	1.55	23.66	71.10	50.80	13.4	45.6	5.60	C	B
Fortaleza	32	<i>E. grandis</i>	73.0	40.8	0.494	4.23	1.48	24.23	70.06	51.15	15.6	40.4	5.80	C	B
Fortaleza	32	<i>E. grandis</i>	73.0	40.8	0.526	4.19	1.48	24.39	69.94	51.32	16.5	39.2	5.80	C	C
Fortaleza	25	<i>E. grandis</i>	71.3	59.0	0.461	3.03	1.62	25.19	70.16	52.48	14.6	42.0	5.60	B	A
Fortaleza	25	<i>E. grandis</i>	71.3	59.0	0.481	2.61	1.53	24.98	70.88	52.61	14.3	43.0	5.50	B	B
Fortaleza	25	<i>E. grandis</i>	72.4	59.0	0.470	3.27	1.54	25.14	70.05	51.70	15.1	41.0	5.60	B	B
Fortaleza	24	<i>E. grandis</i>	71.3	49.0	0.473	5.75	2.02	24.57	67.66	51.47	15.7	41.6	5.90	B	B
Fortaleza	24	<i>E. grandis</i>	71.3	49.0	0.474	5.46	1.91	25.48	67.15	51.26	16.8	43.3	5.20	C	C
Fortaleza	24	<i>E. grandis</i>	71.3	49.0	0.471	5.34	1.98	24.68	68.00	52.13	15.3	40.1	5.30	B	B
Campinas	24	<i>E. grandis</i>	82.1	68.4	0.429	4.55	2.52	26.18	66.75	52.27	16.3	46.1	5.80	B	B
Campinas	24	<i>E. grandis</i>	82.1	68.4	0.443	4.59	2.31	25.81	67.29	50.60	16.4	45.6	5.40	B	C
Campinas	24	<i>E. grandis</i>	82.1	68.4	0.456	5.24	2.18	26.05	66.53	52.75	15.3	40.8	5.80	B	B
Fortaleza	90	<i>E. grandis</i>	92.6	38.3	0.495	2.55	1.46	25.79	70.20	51.98	14.5	41.4	5.80	B	B
Fortaleza	90	<i>E. grandis</i>	93.3	38.3	0.490	3.96	1.41	25.58	69.05	52.01	15.7	49.7	5.10	C	B
Fortaleza	90	<i>E. grandis</i>	93.3	38.3	0.519	4.16	1.26	26.11	68.47	50.46	16.5	44.4	5.10	C	C
Fortaleza	253	<i>E. grandis</i>	89.9	46.6	0.469	4.40	1.82	25.26	68.52	51.94	16.0	45.0	5.50	B	B
Fortaleza	253	<i>E. grandis</i>	89.9	46.6	0.478	3.95	1.75	25.79	68.51	52.19	15.7	45.6	5.20	B	B
Fortaleza	253	<i>E. grandis</i>	89.9	46.6	0.457	3.62	1.84	26.64	67.90	51.15	16.3	42.8	5.10	B	B
Fortaleza	258	<i>E. urophylla</i>	78.7	58.5	0.515	6.93	2.56	24.55	65.96	49.59	17.6	51.4	3.40	D	D
Fortaleza	258	<i>E. urophylla</i>	78.7	58.5	0.509	6.87	2.67	24.21	66.25	49.47	16.7	50.1	3.40	D	D
Fortaleza	258	<i>E. urophylla</i>	78.7	58.5	0.488	6.06	2.74	25.03	66.17	49.50	17.3	48.3	3.50	D	D

Fortaleza	270	<i>E. urophylla</i>	77.6	42.7	0.521	5.96	1.27	25.50	67.27	51.45	16.7	54.6	4.10	D	C
Fortaleza	270	<i>E. urophylla</i>	77.6	42.7	0.517	4.53	1.35	26.86	67.26	50.52	16.5	49.3	4.50	D	C
Fortaleza	270	<i>E. urophylla</i>	77.6	42.7	0.527	5.79	1.28	26.51	66.42	50.01	16.9	50.7	3.80	D	D
Campinas	33	<i>E. grandis</i>	92.5	54.4	0.446	2.79	1.63	25.33	70.25	53.20	14.8	38.7	6.10	A	A
Fazenda	Talhão	Espécie	Idade	Produtividade	Db	ET	LS	LI	H	RD	NK	V	A.A.R.	Classe Esp. A	Classe Esp. B
Campinas	33	<i>E. grandis</i>	92.5	54.4	0.444	4.03	1.61	24.99	69.37	51.81	15.7	46.9	5.40	B	B
Campinas	33	<i>E. grandis</i>	92.5	54.4	0.468	4.52	1.65	25.24	68.59	51.15	17.5	50.6	5.20	B	C
Fortaleza	252	<i>E. urophylla</i>	89.2	33.4	0.542	4.57	1.42	24.66	69.35	50.62	17.7	49.8	4.70	D	D
Fortaleza	252	<i>E. urophylla</i>	89.8	33.4	0.557	8.20	1.43	23.40	66.97	50.00	17.2	53.8	3.90	D	D
Fortaleza	252	<i>E. urophylla</i>	89.8	33.4	0.563	6.91	1.50	23.98	67.61	49.10	17.1	51.8	4.90	D	D
Fortaleza	262	<i>E. urophylla</i>	82.1	22.4	0.557	3.88	1.20	27.19	67.73	50.72	17.4	45.5	4.40	D	D
Fortaleza	262	<i>E. urophylla</i>	82.4	22.4	0.570	5.15	1.21	26.71	66.93	50.00	17.8	46.4	4.20	D	D
Fortaleza	262	<i>E. urophylla</i>	82.4	22.4	0.566	7.80	2.39	26.39	63.42	47.61	18.1	43.0	4.30	D	D
Fortaleza	35	<i>E. urophylla</i>	92.5	51.8	0.520	6.40	2.40	25.19	66.01	49.73	14.7	50.5	4.70	D	C
Fortaleza	35	<i>E. urophylla</i>	92.5	51.8	0.542	6.47	2.27	25.37	65.89	50.95	16.4	51.8	4.10	D	D
Fortaleza	35	<i>E. urophylla</i>	92.5	51.8	0.544	7.72	2.41	26.13	63.74	50.34	16.4	58.3	3.00	D	D
sta geneveva	50	<i>E. grandis</i>	68.5	36.5	0.449	3.28	1.55	24.61	70.56	52.60	15.8	45.3	5.20	B	A
sta geneveva	50	<i>E. grandis</i>	68.5	36.5	0.456	3.76	1.57	24.73	69.94	51.19	16.3	45.3	6.20	B	B
sta geneveva	50	<i>E. grandis</i>	68.5	36.5	0.442	4.01	1.55	24.81	69.63	51.81	15.9	52.4	6.10	B	B
sta geneveva	5	<i>E. grandis</i>	79.3	25.6	0.472	5.18	2.35	27.89	64.58	51.49	17.1	45.2	5.70	C	C
sta geneveva	5	<i>E. grandis</i>	79.4	25.6	0.489	5.64	2.26	26.68	65.42	51.45	16.0	47.5	5.80	C	B
sta geneveva	5	<i>E. grandis</i>	79.4	25.6	0.502	4.80	2.00	26.87	66.33	51.25	16.2	47.1	5.50	C	C
sta geneveva	72	<i>E. grandis</i>	71.2	24.8	0.482	5.37	0.95	24.47	69.21	52.41	16.5	43.9	5.60	C	C
sta geneveva	72	<i>E. grandis</i>	71.2	24.8	0.481	4.33	1.15	24.66	69.86	51.93	16.3	45.3	5.60	C	C
sta geneveva	72	<i>E. grandis</i>	71.4	24.8	0.488	4.77	0.93	26.08	68.22	51.78	16.7	44.1	5.40	C	C
Ibill	308	<i>E. saligna</i>	69.3	49.1	0.485	3.48	1.48	25.25	69.79	50.72	16.2	43.5	5.40	C	C

Ibiti	308	E. saligna	69.3	49.1	0.523	3.07	1.65	24.52	70.76	50.59	16.0	40.6	5.30	C	C
Ibiti	308	E. saligna	69.3	49.1	0.527	3.14	1.65	27.76	67.45	50.43	15.9	44.8	5.00	C	C
Ibiti	310	E. saligna	67.9	59.1	0.492	3.15	1.89	25.32	69.64	50.93	15.0	37.9	5.20	B	C
Ibiti	310	E. saligna	67.9	59.1	0.500	4.87	1.80	25.32	68.01	52.12	14.7	41.3	5.30	B	B
Ibiti	310	E. saligna	67.9	59.1	0.508	4.88	1.84	25.39	67.89	51.64	15.2	39.5	5.10	B	C
Fazenda	Talhão	Espécie	Idade	Produtividade	Db	ET	LS	LI	H	RD	NK	V	A.A.R.	Classe Esp. A	Classe Esp. B
Ibiti	205	E. saligna	88.5	93.7	0.450	2.64	1.66	24.01	71.69	52.44	15.5	39.9	6.30	B	A
Ibiti	205	E. saligna	88.5	93.7	0.471	2.55	1.66	23.54	72.25	51.79	14.4	39.0	5.90	B	B
Ibiti	205	E. saligna	88.5	93.7	0.445	3.11	1.71	23.51	71.67	51.68	15.7	41.9	5.90	B	B
Ibiti	237	E. saligna	83.3	69.9	0.471	2.05	1.68	24.30	71.97	52.07	15.0	41.4	6.30	B	A
Ibiti	237	E. saligna	83.3	69.9	0.464	2.67	1.57	24.37	71.39	52.49	15.2	44.4	6.20	B	B
Ibiti	237	E. saligna	83.3	69.9	0.466	3.25	1.71	24.08	70.96	52.42	16.1	40.7	5.70	B	B
Ibiti	230	E. saligna	84.7	46.9	0.487	4.36	1.95	25.33	68.36	51.86	14.9	36.6	5.60	B	B
Ibiti	230	E. saligna	84.7	46.9	0.501	4.39	1.89	25.72	68.00	50.95	15.7	41.2	5.40	C	C
Ibiti	230	E. saligna	84.7	46.9	0.489	4.46	1.80	25.03	68.71	52.32	15.4	43.3	4.70	B	C
Ibiti	250	E. dunnii	75.1	43.9	0.537	2.97	2.02	25.11	69.90	51.11	16.9	39.6	5.30	D	C
Ibiti	250	E. dunnii	75.1	43.9	0.542	4.11	2.14	23.73	70.02	50.27	16.6	41.3	5.00	D	C
Ibiti	250	E. dunnii	75.1	43.9	0.514	2.55	2.21	23.90	71.34	50.76	16.9	44.6	4.70	D	C
Ibiti	253	E. dunnii	78.1	54.4	0.531	3.66	2.04	25.02	69.28	49.72	17.8	43.0	4.90	D	D
Ibiti	253	E. dunnii	78.1	54.4	0.573	3.18	2.14	23.96	70.72	49.83	15.7	32.8	6.30	D	C
Ibiti	253	E. dunnii	78.1	54.4	0.538	3.12	2.21	24.04	70.63	49.71	16.4	37.1	6.10	D	C
Ibiti	210	E. dunnii	85.2	57.3	0.531	4.60	1.34	25.25	68.81	51.11	15.6	40.9	5.20	C	C
Ibiti	210	E. dunnii	85.2	57.3	0.529	5.62	1.52	24.98	67.88	51.04	17.2	43.5	4.80	D	D
Ibiti	210	E. dunnii	85.2	57.3	0.515	5.05	1.45	25.74	67.76	50.73	17.9	40.2	4.80	D	D
Ibiti	212	E. dunnii	85.2	45.7	0.535	6.43	2.23	23.90	67.44	49.82	17.9	39.8	4.90	D	D
Ibiti	212	E. dunnii	85.2	45.7	0.532	5.00	2.16	24.03	68.81	51.26	16.0	37.3	5.40	C	C

Ibiti	212	E. dunnii	85.2	45.7	0.538	4.74	2.32	23.59	69.35	50.08	18.8	39.1	4.50	D	D
Ibiti	196	E. dunnii	94.6	72.9	0.554	4.94	2.22	24.39	68.45	52.16	13.4	41.7	6.10	C	C
Ibiti	196	E. dunnii	94.6	72.9	0.561	3.92	1.44	23.57	71.07	52.60	13.0	43.0	6.00	C	C
Ibiti	196	E. dunnii	94.6	72.9	0.549	4.11	1.35	23.77	70.77	53.04	14.5	40.1	5.50	C	C
Ibiti	433	E. grandis	59.2	44.5	0.466	5.41	1.87	23.92	68.80	52.88	15.7	39.2	5.80	B	B
Ibiti	433	E. grandis	59.2	44.5	0.472	5.38	1.89	24.94	67.79	52.01	15.7	40.1	5.70	B	B
Fazenda	Talhão	Espécie	Idade	Produtividade	Db	ET	LS	LI	H	RD	NK	V	A.A.R.	Classe Esp. A	Classe Esp. B
Ibiti	433	E. grandis	59.2	44.5	0.460	4.80	1.88	24.62	68.70	53.63	14.9	38.0	5.50	B	A
Ibiti	305	E. grandis	68.9	63.7	0.439	4.27	1.84	25.91	67.98	53.15	15.2	43.8	5.70	A	B
Ibiti	305	E. grandis	68.9	63.7	0.455	5.47	1.75	25.82	66.96	52.77	15.9	43.3	5.30	A	B
Ibiti	305	E. grandis	68.9	63.7	0.453	5.12	1.84	24.81	68.23	53.37	15.3	43.7	5.70	A	B
Ibiti	412	E. grandis	69.8	95.4	0.414	4.06	1.84	25.38	68.72	54.40	14.3	39.2	6.90	A	A
Ibiti	412	E. grandis	69.8	95.4	0.396	2.66	2.03	25.16	70.15	54.52	14.4	41.5	6.30	A	A
Ibiti	412	E. grandis	69.8	95.4	0.390	4.08	1.98	25.05	68.89	53.66	14.9	45.3	6.20	A	A
Ibiti	244	E. grandis	85.5	61.3	0.448	3.91	1.79	24.93	69.37	53.86	13.9	39.9	6.40	A	A
Ibiti	244	E. grandis	85.5	61.3	0.444	4.36	1.74	23.99	69.91	54.42	13.4	37.8	6.50	A	A
Ibiti	244	E. grandis	85.5	61.3	0.447	4.54	1.76	24.78	68.92	53.36	13.7	40.2	6.50	A	A
Ibiti	222	E. grandis	88.1	86.1	0.458	4.55	1.76	24.96	68.73	53.32	13.9	40.1	6.50	A	A
Ibiti	222	E. grandis	88.1	86.1	0.455	4.84	1.86	24.87	68.43	53.99	13.4	33.1	6.10	A	A
Ibiti	222	E. grandis	88.1	86.1	0.443	4.44	1.89	25.72	67.95	53.22	13.5	32.4	5.60	A	A
Guaruja	280	E. grandis	71.3	73.5	0.464	3.73	1.23	24.42	70.62	54.09	15.2	40.7	6.00	A	B
Guaruja	280	E. grandis	71.3	73.5	0.423	3.61	1.25	25.32	69.82	53.27	15.5	40.8	5.70	A	B
Guaruja	280	E. grandis	71.3	73.5	0.440	3.81	1.21	24.59	70.39	53.67	14.7	39.6	5.70	A	A
Guaruja	211	E. grandis	72.8	96.1	0.435	3.88	1.38	23.27	71.47	53.35	14.2	45.6	5.90	A	A
Guaruja	211	E. grandis	72.8	96.1	0.429	3.73	1.29	24.82	70.16	52.39	15.3	47.9	6.00	A	B
Guaruja	211	E. grandis	72.8	96.1	0.443	1.79	1.20	24.78	72.23	52.88	15.4	44.6	5.90	A	B

Guaruja	70	E. grandis	75.4	64.2	0.474	3.55	1.54	25.40	69.51	53.81	15.4	44.8	6.10	B	B
Guaruja	70	E. grandis	75.4	64.2	0.467	3.34	1.79	24.53	70.34	53.70	15.2	43.1	6.40	B	B
Guaruja	70	E. grandis	75.4	64.2	0.451	3.54	1.73	24.70	70.03	52.89	15.5	46.2	6.10	B	B
Guaruja	204	E. grandis	85.1	42.7	0.492	3.32	1.16	26.26	69.27	53.68	14.3	42.6	6.30	B	B
Guaruja	204	E. grandis	85.1	42.7	0.474	3.71	1.07	26.15	69.07	53.32	15.0	44.1	6.10	B	B
Guaruja	204	E. grandis	85.1	42.7	0.479	4.02	1.05	26.00	68.93	53.41	15.3	51.4	6.20	B	B
Guaruja	203	E. grandis	85.3	58.3	0.494	4.60	2.25	24.60	68.55	53.13	15.4	45.6	6.00	B	B
Fazenda	Talhão	Espécie	Idade	Produtividade	Db	ET	LS	LI	H	RD	NK	V	A.A.R.	Classe Esp. A	Classe Esp. B
Guaruja	203	E. grandis	85.3	58.3	0.478	4.29	2.18	24.26	69.27	52.89	15.7	44.8	5.60	B	B
Guaruja	203	E. grandis	85.3	58.3	0.467	5.32	1.90	24.25	68.53	53.32	14.8	43.6	5.90	B	B
Guaruja	184	E. grandis	80.6	79.2	0.445	4.44	1.97	25.25	68.34	52.42	14.8	41.0	5.90	A	A
Guaruja	184	E. grandis	80.6	79.2	0.437	4.41	2.25	25.12	68.22	53.29	14.7	40.2	5.90	A	A
Guaruja	184	E. grandis	80.6	79.2	0.458	3.73	2.24	24.60	69.43	52.57	14.3	36.5	5.80	A	A
Guaruja	283	E. saligna	75.6	49.4	0.481	3.58	1.47	26.35	68.60	51.67	15.1	42.7	5.70	B	B
Guaruja	283	E. saligna	75.6	49.4	0.471	3.38	1.23	25.18	70.21	51.89	15.2	46.9	5.20	B	B
Guaruja	283	E. saligna	75.6	49.4	0.501	3.76	1.33	26.28	68.63	50.84	15.4	47.0	5.20	B	C
Guaruja	177	E. saligna	83.9	67.0	0.480	3.98	1.75	24.79	69.48	52.71	14.8	47.1	6.00	B	B
Guaruja	177	E. saligna	83.9	67.0	0.473	3.54	1.92	25.27	69.27	52.07	14.8	44.8	5.30	B	B
Guaruja	177	E. saligna	83.9	67.0	0.447	4.57	1.80	25.25	68.38	52.93	15.0	46.0	5.40	B	A
Guaruja	179	E. saligna	84.6	60.6	0.473	3.57	1.72	25.36	69.35	52.34	15.2	45.2	5.30	B	B
Guaruja	179	E. saligna	84.6	60.6	0.497	3.55	1.81	25.10	69.53	52.97	14.0	44.2	6.00	B	B
Guaruja	179	E. saligna	84.6	60.6	0.471	3.34	1.86	25.05	69.75	53.27	14.5	49.0	5.90	B	B
Saligna	58	E. grandis	71.4	57.2	0.443	3.99	1.25	24.58	70.18	51.97	15.2	47.3	5.80	B	B
Saligna	58	E. grandis	71.4	57.2	0.436	4.02	1.25	25.37	69.36	52.55	16.1	42.6	5.40	B	B
Saligna	58	E. grandis	71.4	57.2	0.461	3.61	1.23	24.54	70.62	53.25	15.3	37.8	5.80	B	B
Saligna	68	E. grandis	71.4	39.4	0.464	4.33	1.27	24.24	70.16	53.22	15.8	44.2	5.40	B	B

Saligna	68	E. grandis	71.4	39.4	0.455	4.81	1.27	24.17	69.75	51.57	15.7	44.2	5.20	B	B
Saligna	68	E. grandis	71.4	39.4	0.439	4.82	1.36	24.54	69.48	51.76	15.8	46.6	5.20	B	B
Saligna	65	E. grandis	71.3	30.4	0.456	4.54	2.41	24.51	68.54	53.38	15.2	44.3	4.30	B	B
Saligna	65	E. grandis	71.3	30.4	0.457	4.91	2.48	24.91	67.70	52.43	15.4	47.1	5.80	B	B
Saligna	65	E. grandis	71.3	30.4	0.464	5.79	2.78	26.06	65.37	51.21	15.7	45.4	5.60	B	B
Campinas	12	E. grandis	72.0	84.0	0.418	3.18	1.65	24.92	70.25	53.51	15.5	49.5	6.30	A	A
Campinas	12	E. grandis	72.0	84.0	0.426	3.24	1.74	25.15	69.87	52.53	15.8	58.5	5.80	B	B
Campinas	12	E. grandis	72.0	84.0	0.409	4.56	1.82	25.23	68.59	52.79	14.5	44.9	6.40	A	A
Fazenda	Talhão	Espécie	Idade	Produtividade	Db	ET	LS	LI	H	RD	NK	V	A.A.R.	Classe Esp. A	Classe Esp. B
Penha	187	E. grandis	62.7	81.0	0.434	3.47	1.84	24.50	70.19	53.84	14.6	45.5	6.40	A	A
Penha	187	E. grandis	62.7	81.0	0.431	3.69	1.88	23.81	70.62	54.44	14.7	49.2	6.50	A	A
Penha	187	E. grandis	62.7	81.0	0.403	3.54	1.96	24.37	70.13	53.60	14.7	41.3	6.30	A	A
Penha	96	E. grandis	73.9	29.6	0.472	4.47	0.99	24.55	69.99	51.98	15.2	41.6	5.40	B	B
Penha	96	E. grandis	73.9	29.6	0.493	4.58	1.01	23.63	70.78	52.01	15.2	45.1	5.20	B	C
Penha	96	E. grandis	73.9	29.6	0.469	4.27	0.98	24.12	70.63	52.37	15.2	44.3	5.40	B	B
Penha	94	E. grandis	73.9	56.9	0.453	3.83	0.98	23.39	71.80	52.89	14.6	48.8	5.40	A	A
Penha	94	E. grandis	73.9	56.9	0.467	4.11	0.94	22.80	72.15	53.32	14.3	41.7	6.40	A	A
Penha	94	E. grandis	73.9	56.9	0.436	4.18	0.97	23.43	71.42	53.87	14.2	51.0	6.00	A	A

Tabela 29 - Tabela gerada pelo programa ANTOTAL, utilizando-se o conjunto dos pontos com as médias por talhão.

talha	Fazenda	Espécie	idade	prod	DensB	extrat	lignS	lignIns	Holoc	RenD	K	Visc	aar	RenD	K	Visc	Aar	E A	E B	RT _r	RGI	Pro
144	"flecha azul"	"E. híbrido"	64	55.7	0.536	5.277	1.953	24.05	68.72	49.78	17.43	47.5	5.1	49.85	17.48	47.36	5.147	"D"	"D"	"D"	"D"	"D"
139	"flecha azul"	"E. grandis"	60.5	62.6	0.4533	6.35	1.793	24.38	67.48	51.21	15.73	44.53	5.933	51.42	15.75	44.82	5.764	"B"	"B"	"B"	"B"	"B"
134	"flecha azul"	"E. grandis"	64	47.3	0.4917	4.853	1.543	24.94	68.66	51.75	15.93	44.67	6.3	51.9	16.05	44.04	6.215	"C"	"B"	"C"	"C"	"C"
267	"flecha azul"	"E. grandis"	75	61.5	0.4917	4.1	1.903	25.22	68.78	51.94	16.27	41.83	6.167	51.92	16.17	41.82	6.086	"C"	"C"	"C"	"C"	"C"
273	"flecha azul"	"E. grandis"	70.23	54.7	0.4933	4.017	1.807	25.77	68.41	52.05	16.53	41.63	5.8	51.79	16.52	42.63	5.824	"C"	"C"	"C"	"C"	"C"
270	"flecha azul"	"E. grandis"	72.43	48.9	0.476	4.64	1.767	26.54	67.05	51.91	16.27	43.97	5.633	51.91	16.27	43.71	5.717	"C"	"B"	"C"	"B"	"C"
280	"flecha azul"	"E. urophylla"	73.53	45.2	0.5357	7.19	1.633	26.14	65.04	49.93	18.43	50.97	3.4	49.52	16.34	49.37	4.024	"D"	"D"	"D"	"D"	"D"
271	"flecha azul"	"E. urophylla"	71.3	50.6	0.512	7.163	1.5	23.77	67.66	49.57	16.63	50.8	4.1	49.82	16.7	51.58	4.078	"D"	"D"	"D"	"D"	"D"
142	"flecha azul"	"E. urophylla"	60.8	43.6	0.5383	6.09	1.683	23.92	68.3	50.03	17.27	50.27	4.133	49.68	17.19	49.78	4.247	"D"	"D"	"D"	"D"	"D"
32	"fortaleza"	"E. grandis"	73	40.8	0.5063	4.037	1.503	24.09	70.37	51.09	15.17	41.73	5.733	51.38	15.35	42.75	5.708	"C"	"B"	"C"	"C"	"C"
26	"fortaleza"	"E. grandis"	71.67	59	0.4707	2.97	1.663	25.1	70.36	52.26	14.67	42	5.567	52.32	14.59	42.84	5.749	"B"	"B"	"B"	"B"	"B"
24	"fortaleza"	"E. grandis"	71.3	49	0.4727	5.517	1.97	24.91	67.6	51.62	15.93	41.67	5.467	51.52	15.87	40.89	5.445	"C"	"B"	"B"	"B"	"B"
24	"campinas"	"E. grandis"	62.1	68.4	0.4427	4.793	2.337	26.01	66.88	51.87	16	44.17	5.667	52.03	16.19	47.76	5.485	"B"	"B"	"B"	"B"	"B"
90	"fortaleza"	"E. grandis"	93.07	38.3	0.5013	3.557	1.377	25.83	69.24	51.48	15.57	45.17	5.333	51.94	15.32	43.24	5.379	"C"	"B"	"B"	"C"	"C"
253	"fortaleza"	"E. grandis"	89.9	46.6	0.468	3.99	1.803	25.9	68.31	51.76	16	44.47	5.267	51.57	16.09	43.83	5.168	"B"	"B"	"B"	"C"	"C"
258	"fortaleza"	"E. urophylla"	78.7	58.5	0.504	6.62	2.657	24.6	66.13	49.52	17.2	49.93	3.433	51.15	14.84	49.14	4.206	"D"	"D"	"C"	"C"	"C"
270	"fortaleza"	"E. urophylla"	77.6	42.7	0.5217	5.427	1.3	26.29	68.98	50.66	16.7	51.53	4.133	50.54	16.79	51.45	4.133	"D"	"C"	"D"	"D"	"D"
33	"campinas"	"E. grandis"	92.5	54.4	0.4527	3.78	1.63	25.19	69.4	52.05	16	45.4	5.567	52.15	15.84	45.2	5.957	"B"	"B"	"B"	"B"	"B"
252	"fortaleza"	"E. urophylla"	89.6	33.4	0.554	6.56	1.45	24.01	67.98	49.91	17.33	51.8	4.5	50.12	17.21	51.74	4.374	"D"	"D"	"D"	"D"	"D"
262	"fortaleza"	"E. urophylla"	82.3	22.4	0.5643	5.61	1.6	26.76	66.03	49.44	17.77	44.97	4.3	49.41	17.71	45.52	4.244	"D"	"D"	"D"	"D"	"D"
35	"fortaleza"	"E. urophylla"	92.5	51.8	0.5353	6.863	2.36	25.56	65.21	50.34	15.83	53.53	3.933	50.27	15.81	52.58	4.019	"D"	"D"	"C"	"D"	"D"
talha	Fazenda	Espécie	idade	prod	DensB	extrat	lignS	lignIns	Holoc	RenD	K	Visc	aar	RenD	K	Visc	aar	Tr ₀	Gilb	RT _r	RGI	Pro
50	"sta geneveva"	"E. grandis"	68.5	36.5	0.449	3.683	1.557	24.72	70.04	51.87	16	47.67	5.933	52.13	16.16	48.27	5.761	"B"	"B"	"B"	"B"	"B"
5	"sta geneveva"	"E. grandis"	79.37	25.6	0.4877	5.207	2.203	27.15	65.44	51.4	16.43	46.6	5.667	51.67	16.52	46.16	5.68	"C"	"C"	"C"	"C"	"C"
72	"sta geneveva"	"E. grandis"	71.27	24.8	0.4937	4.823	1.01	25.07	69.1	52.04	16.5	44.43	5.533	51.62	16.35	44.7	5.496	"C"	"C"	"C"	"C"	"C"
308	"ibiti"	"E. saligna"	69.3	49.1	0.5117	3.23	1.593	25.84	69.33	50.58	16.03	42.97	5.233	50.69	15.93	41.54	5.234	"D"	"C"	"C"	"C"	"C"
310	"ibiti"	"E. saligna"	67.9	59.1	0.5	4.3	1.843	25.34	68.51	51.56	14.97	39.57	5.2	51.33	15.07	39.71	5.236	"B"	"C"	"B"	"C"	"C"
205	"ibiti"	"E. saligna"	88.5	93.7	0.4553	2.767	1.677	23.89	71.87	51.97	15.2	40.27	6.033	52.22	15.32	41.04	5.989	"B"	"B"	"B"	"B"	"A"

237	"ibiti"	"E. saligna"	83.3	69.9	0.467	2.057	1.653	24.25	71.44	52.33	15.43	42.16	6.067	52.33	15.33	42.53	5.999	"B"	"B"	"B"	"B"	"B"
230	"ibiti"	"E. saligna"	84.7	46.9	0.4923	4.403	1.86	25.36	68.36	51.71	15.33	40.37	5.233	51.53	15.01	42.96	5.246	"B"	"C"	"B"	"C"	"C"
250	"ibiti"	"E. dunni"	75.1	43.9	0.531	3.21	2.123	24.25	70.42	50.71	16.8	41.83	5	50.42	16.95	40.88	5.247	"D"	"C"	"D"	"C"	"C"
253	"ibiti"	"E. dunni"	78.1	54.4	0.5473	3.32	2.13	24.34	70.21	49.75	16.63	37.63	5.767	50.45	16.52	39.32	5.242	"D"	"C"	"D"	"C"	"C"
210	"ibiti"	"E. dunni"	85.2	57.3	0.525	5.09	1.437	25.32	68.15	50.96	16.9	41.53	4.933	51.09	16.87	41.03	4.96	"D"	"D"	"D"	"C"	"C"
212	"ibiti"	"E. dunni"	85.2	45.7	0.535	5.39	2.237	23.84	68.53	50.39	17.57	38.73	4.933	50.01	17.37	38.67	5.104	"D"	"D"	"D"	"D"	"D"
196	"ibiti"	"E. dunni"	94.6	72.9	0.5547	4.323	1.67	23.91	70.1	52.6	13.63	41.6	5.867	52.53	13.87	41.53	5.877	"C"	"C"	"C"	"C"	"B"
433	"ibiti"	"E. grandis"	59.2	44.5	0.466	5.197	1.88	24.49	68.43	52.84	15.43	39.1	5.667	51.95	16.46	40.98	5.456	"B"	"B"	"B"	"B"	"C"
305	"ibiti"	"E. grandis"	60.9	63.7	0.449	4.953	1.81	25.51	67.72	53.1	15.47	43.6	5.567	53.34	15.49	43.29	5.692	"A"	"B"	"A"	"B"	"B"
412	"ibiti"	"E. grandis"	69.8	95.4	0.4	3.6	1.95	25.2	69.25	54.19	14.53	42	6.467	53.92	14.12	41.94	6.436	"A"	"A"	"A"	"A"	"A"
244	"ibiti"	"E. grandis"	85.5	81.3	0.4463	4.27	1.763	24.57	69.4	53.88	13.67	39.3	6.467	53.86	13.92	39.45	6.349	"A"	"A"	"A"	"A"	"A"
222	"ibiti"	"E. grandis"	88.1	86.1	0.452	4.61	1.837	25.18	68.37	53.51	13.6	35.2	6.067	53.53	13.77	35.88	6.071	"A"	"A"	"A"	"A"	"A"
280	"guarujá"	"E. grandis"	71.3	73.5	0.4423	3.717	1.23	24.78	70.28	53.68	15.13	40.37	5.8	53.51	15.04	44.96	6.197	"A"	"B"	"A"	"A"	"B"
211	"guarujá"	"E. grandis"	72.8	96.1	0.4357	3.133	1.29	24.29	71.29	52.87	14.97	46.03	5.933	53.42	14.55	39.57	6.525	"A"	"B"	"A"	"A"	"A"
70	"guarujá"	"E. grandis"	75.4	64.2	0.464	3.477	1.887	24.88	69.96	53.47	15.37	44.7	6.2	53.3	15.14	45.41	6.114	"B"	"B"	"B"	"B"	"B"
204	"guarujá"	"E. grandis"	85.1	42.7	0.4817	3.683	1.09	26.14	69.09	53.47	14.87	46.03	6.2	53.53	14.84	46.03	6.036	"B"	"B"	"B"	"B"	"B"
203	"guarujá"	"E. grandis"	85.3	58.3	0.4797	4.737	2.11	24.37	68.78	53.11	15.3	44.6	5.833	52.84	15.19	42.51	5.802	"B"	"B"	"B"	"B"	"B"
184	"guarujá"	"E. grandis"	80.6	79.2	0.4467	4.193	2.153	24.99	68.66	52.76	14.53	39.23	5.867	53.31	14.86	40.42	5.971	"A"	"A"	"A"	"A"	"A"
283	"guarujá"	"E. saligna"	75.6	49.4	0.4843	3.573	1.343	25.94	68.15	51.47	15.23	45.53	5.367	51.82	15.45	46.03	5.427	"B"	"B"	"B"	"B"	"C"
177	"guarujá"	"E. saligna"	83.9	67	0.4667	4.03	1.823	25.1	69.04	52.57	14.87	45.97	5.567	52.53	14.6	46.26	5.58	"B"	"B"	"B"	"B"	"B"
179	"guarujá"	"E. saligna"	84.6	60.6	0.4803	3.49	1.797	25.17	69.54	52.86	14.57	46.13	5.733	52.42	14.71	45.76	5.673	"B"	"B"	"B"	"B"	"B"
talha	Fazenda	Espécie	idade	prod	DensB	extrat	lignS	lignins	Holoc	RenD	K	Visc	aar	RenD	K	Visc	aar	Tro	Glib	RTr	RGI	Pro
58	"saligna"	"E. grandis"	71.4	57.2	0.4467	3.873	1.243	24.83	70.05	52.59	15.53	42.57	5.667	52.55	15.35	43.89	5.482	"B"	"B"	"B"	"B"	"B"
68	"saligna"	"E. grandis"	71.4	39.4	0.4527	4.587	1.3	24.32	69.8	52.18	15.77	45	5.267	52.24	15.81	45.77	5.288	"B"	"B"	"B"	"B"	"B"
65	"saligna"	"E. grandis"	71.3	30.4	0.459	5.08	2.557	25.16	67.2	52.34	15.43	45.6	5.233	52.26	15.43	45.06	5.231	"B"	"B"	"B"	"B"	"B"
12	"campinas"	"E. grandis"	72	84	0.4177	3.66	1.737	25.1	69.5	52.94	15.27	50.97	6.167	52.62	15.3	46.76	6.223	"B"	"A"	"A"	"A"	"B"
187	"penha"	"E. grandis"	62.7	81	0.4227	3.567	1.893	24.23	70.31	53.96	14.67	45.33	6.4	53.5	13.9	42.7	6.298	"A"	"A"	"A"	"A"	"A"
96	"penha"	"E. grandis"	73.9	29.6	0.478	4.44	0.9933	24.1	70.47	52.12	15.2	43.67	5.933	52.1	15.93	45.31	5.655	"B"	"B"	"C"	"B"	"B"
94	"penha"	"E. grandis"	73.9	56.9	0.452	4.04	0.9633	23.21	71.79	53.36	14.37	47.17	5.833	52.83	14.38	43.35	6.283	"A"	"A"	"A"	"A"	"A"

Tabela 30 – Tabela gerada pelo programa ANTOTAL, utilizando-se o conjunto de todos os pontos (e não as médias). Observe-se que pela rede de predição ter sido treinada para as médias apenas estes resultados não são tão bons, mas a classificação obtida desta forma não fica tão prejudicada.

talha	Fazenda	espécie	idade	prod	DensB	extrat	lignS	lignins	Holoc	RenD	K	Visc	aar	RenD	K	Visc	aar	Tro	Glib	RTr	RGI	Pro
144	"flecha azul"	"E. híbrido"	84	55.7	0.536	6.97	1.93	23.66	67.44	50.21	16.5	44.2	5.3	49.58	17.15	50.54	4.608	"D"	"C"	"D"	"D"	"D"
144	"flecha azul"	"E. híbrido"	84	55.7	0.522	4.21	1.95	23.87	69.97	49.77	17.2	44.5	5.3	50.38	17.03	45.87	5.639	"D"	"D"	"D"	"C"	"C"
144	"flecha azul"	"E. híbrido"	84	55.7	0.55	4.66	1.98	24.61	68.76	49.37	18.6	53.8	4.7	49.79	17.68	46.06	5.176	"D"	"D"	"D"	"D"	"D"
139	"flecha azul"	"E. grandis"	80.5	62.6	0.439	5.74	1.77	24.95	67.54	51.37	15.3	42.8	6.6	51.82	15.7	44.68	5.869	"B"	"B"	"B"	"B"	"B"
139	"flecha azul"	"E. grandis"	80.5	62.6	0.463	7.38	1.75	24.6	66.27	50.67	15.5	42.6	5.6	51.12	15.46	47.42	5.376	"B"	"B"	"B"	"B"	"C"
139	"flecha azul"	"E. grandis"	80.5	62.6	0.458	5.93	1.86	23.58	68.63	51.6	16.4	48.2	5.6	51.41	15.63	42.17	6.098	"B"	"B"	"B"	"B"	"B"
134	"flecha azul"	"E. grandis"	84	47.3	0.504	4.29	1.65	24.34	69.72	51.95	15.4	42.3	6.8	51.71	16.91	43.08	6.375	"B"	"B"	"C"	"C"	"C"
134	"flecha azul"	"E. grandis"	84	47.3	0.462	5.08	1.5	25.32	66.1	51.25	15.5	48.5	6.3	52.1	15.8	44.61	6.195	"B"	"B"	"B"	"B"	"B"
134	"flecha azul"	"E. grandis"	84	47.3	0.509	5.19	1.48	25.17	68.16	52.05	16.9	43.2	5.8	51.83	16.12	44.3	6.098	"C"	"C"	"C"	"C"	"C"
267	"flecha azul"	"E. grandis"	75	61.5	0.499	1.65	1.78	26.44	70.13	50.71	16.3	40.9	6.3	52.08	16.16	42.42	6.367	"C"	"C"	"C"	"C"	"B"
267	"flecha azul"	"E. grandis"	75	61.5	0.468	5.52	2.02	24.72	67.74	52.53	16.3	43.4	6.2	51.74	15.94	44.11	5.898	"B"	"B"	"B"	"B"	"B"
267	"flecha azul"	"E. grandis"	75	61.5	0.508	5.13	1.91	24.5	68.46	52.57	16.2	41.2	6	51.51	16.23	41.98	6.027	"C"	"C"	"C"	"C"	"C"
273	"flecha azul"	"E. grandis"	71.5	54.7	0.499	3.71	1.75	25.62	68.92	51.84	16.4	41.4	5.9	51.77	16.53	43.04	5.937	"C"	"C"	"C"	"C"	"C"
273	"flecha azul"	"E. grandis"	67.5	54.7	0.493	4.58	1.8	25.62	68	52.67	16.2	40.9	5.9	51.85	16.51	42.8	5.72	"C"	"C"	"C"	"C"	"C"
273	"flecha azul"	"E. grandis"	71.7	54.7	0.488	3.76	1.87	26.07	68.3	51.66	17	42.6	5.6	51.93	16.45	42.38	5.844	"C"	"C"	"C"	"C"	"C"
talha	Fazenda	espécie	idade	prod	DensB	extrat	lignS	lignins	Holoc	RenD	K	Visc	aar	RenD	K	Visc	aar	Tro	Glib	RTr	RGI	Pro
270	"flecha azul"	"E. grandis"	72.4	48.8	0.499	4.34	1.75	26.43	67.48	51.17	16.4	44.3	5.5	51.7	16.57	43.29	5.745	"C"	"C"	"C"	"C"	"C"
270	"flecha azul"	"E. grandis"	72.4	48.8	0.461	4.84	1.84	26.48	66.84	51.97	16.3	44.2	5.6	51.99	16.08	44.33	5.687	"B"	"B"	"B"	"B"	"C"
270	"flecha azul"	"E. grandis"	72.5	48.8	0.468	4.74	1.71	26.71	66.84	52.6	16.1	43.4	5.8	52	16.11	44.15	5.703	"B"	"B"	"B"	"B"	"C"
280	"flecha azul"	"E. urophylla"	73.5	45.2	0.529	6.65	1.6	26.59	65.16	48.67	18.3	50.7	3.4	49.7	16.27	49.88	4.05	"D"	"D"	"D"	"D"	"D"
280	"flecha azul"	"E. urophylla"	73.5	45.2	0.541	7.7	1.64	26.49	64.17	50.59	18.3	51.3	3.3	49.44	16.22	48.77	4.012	"D"	"D"	"D"	"D"	"D"
280	"flecha azul"	"E. urophylla"	73.6	45.2	0.537	7.22	1.66	25.34	65.76	50.52	18.7	50.9	3.5	49.46	16.55	49.5	4.021	"D"	"D"	"D"	"D"	"D"
271	"flecha azul"	"E. urophylla"	71.3	50.6	0.522	7.56	1.54	23.97	66.93	49.25	17.3	52.5	4	49.66	16.61	50.76	4.035	"D"	"D"	"D"	"D"	"D"
271	"flecha azul"	"E. urophylla"	71.3	50.6	0.506	6.58	1.48	23.83	68.11	49.57	16.9	53.7	4.1	49.89	16.79	51.83	4.132	"D"	"D"	"D"	"D"	"D"
271	"flecha azul"	"E. urophylla"	71.3	50.6	0.508	7.35	1.49	23.52	67.65	49.9	15.7	46.2	4.2	49.84	16.71	51.91	4.084	"D"	"C"	"D"	"D"	"D"

142	"lecha azul"	"E. urophylla"	80.8	43.6	0.557	6.14	1.91	23.86	68.09	50.4	15.8	44	4	49.44	17.31	48.33	4.238	"C"	"C"	"D"	"D"	"D"
142	"lecha azul"	"E. urophylla"	80.8	43.6	0.538	5.96	1.5	23.9	68.64	49.9	17.2	53.5	4.5	49.76	17.24	49.43	4.294	"D"	"D"	"D"	"D"	"D"
142	"lecha azul"	"E. urophylla"	80.8	43.6	0.52	6.17	1.64	24.01	68.18	49.8	18.8	53.3	3.9	49.9	16.99	51.2	4.234	"D"	"D"	"D"	"D"	"D"
32	"fortaleza"	"E. grandis"	73	40.8	0.499	3.69	1.55	23.66	71.1	50.8	13.4	45.6	5.6	51.62	14.81	42.71	5.825	"C"	"B"	"B"	"B"	"B"
32	"fortaleza"	"E. grandis"	73	40.8	0.494	4.23	1.48	24.23	70.06	51.15	15.6	40.4	5.8	51.41	15.6	43.14	5.648	"C"	"B"	"C"	"C"	"B"
32	"fortaleza"	"E. grandis"	73	40.8	0.526	4.19	1.48	24.39	69.94	51.32	16.5	39.2	5.8	51.09	15.85	42.54	5.625	"C"	"C"	"C"	"C"	"C"
25	"fortaleza"	"E. grandis"	71.3	59	0.461	3.03	1.62	25.19	70.18	52.48	14.6	42	5.6	52.44	14.63	42.8	5.746	"B"	"A"	"B"	"B"	"B"
25	"fortaleza"	"E. grandis"	71.3	59	0.481	2.61	1.53	24.98	70.88	52.61	14.3	43	5.5	52.3	14.42	43.09	5.817	"B"	"B"	"B"	"B"	"B"
25	"fortaleza"	"E. grandis"	72.4	59	0.47	3.27	1.54	25.14	70.05	51.7	15.1	41	5.6	52.23	14.76	42.62	5.671	"B"	"B"	"B"	"B"	"B"
24	"fortaleza"	"E. grandis"	71.3	49	0.473	5.75	2.02	24.57	67.66	51.47	15.7	41.6	5.9	51.47	15.88	40.69	5.486	"B"	"B"	"B"	"B"	"B"
24	"fortaleza"	"E. grandis"	71.3	49	0.474	5.46	1.91	25.48	67.15	51.26	16.8	43.3	5.2	51.52	16.12	41.22	5.36	"C"	"C"	"B"	"C"	"C"
24	"fortaleza"	"E. grandis"	71.3	49	0.471	5.34	1.98	24.68	68	52.13	15.3	40.1	5.3	51.6	15.66	40.78	5.519	"B"	"B"	"B"	"B"	"B"
24	"campinas"	"E. grandis"	82.1	68.4	0.429	4.55	2.52	26.18	66.75	52.27	16.3	46.1	5.8	52.25	16.05	47.31	5.523	"B"	"B"	"B"	"B"	"B"
24	"campinas"	"E. grandis"	82.1	68.4	0.443	4.59	2.31	25.81	67.29	50.6	16.4	45.6	5.4	52.02	16.26	47.68	5.573	"B"	"C"	"B"	"B"	"B"
24	"campinas"	"E. grandis"	82.1	68.4	0.456	5.24	2.18	26.05	66.53	52.75	15.3	40.8	5.8	51.79	16.26	48.22	5.365	"B"	"B"	"B"	"B"	"C"
90	"fortaleza"	"E. grandis"	92.6	38.3	0.495	2.55	1.46	25.79	70.2	51.98	14.5	41.4	5.8	52.34	14.75	42.53	5.592	"B"	"B"	"B"	"B"	"B"
90	"fortaleza"	"E. grandis"	93.3	38.3	0.49	3.96	1.41	25.58	69.05	52.01	15.7	49.7	5.1	51.83	15.47	43.54	5.358	"C"	"B"	"B"	"C"	"C"
90	"fortaleza"	"E. grandis"	93.3	38.3	0.519	4.16	1.26	26.11	68.47	50.46	16.5	44.4	5.1	51.64	15.66	43.6	5.291	"C"	"C"	"C"	"C"	"C"
talha	Fazenda	espécie	idade	prod	DensB	extrat	lignS	lignIns	Holoc	RenD	K	Visc	aar	RenD	K	Visc	aar	Tro	Gilb	RTTr	RGi	Pro
253	"fortaleza"	"E. grandis"	89.9	46.6	0.469	4.4	1.82	25.26	68.52	51.94	16	45	5.5	51.39	16.04	43.64	5.231	"B"	"B"	"B"	"C"	"C"
253	"fortaleza"	"E. grandis"	89.9	46.6	0.478	3.95	1.75	25.78	68.51	52.19	15.7	45.6	5.2	51.52	16.04	43.78	5.2	"B"	"B"	"C"	"C"	"C"
253	"fortaleza"	"E. grandis"	89.9	46.6	0.457	3.62	1.84	26.64	67.9	51.15	16.3	42.8	5.1	51.83	16.09	44.14	5.138	"B"	"B"	"B"	"B"	"C"
258	"fortaleza"	"E. urophylla"	78.7	58.5	0.515	6.93	2.56	24.55	65.96	49.59	17.6	51.4	3.4	50.92	14.96	49.33	4.19	"D"	"D"	"C"	"C"	"C"
258	"fortaleza"	"E. urophylla"	78.7	58.5	0.509	6.87	2.67	24.21	66.25	49.47	16.7	50.1	3.4	51.07	14.75	49.23	4.255	"D"	"D"	"C"	"C"	"C"
258	"fortaleza"	"E. urophylla"	78.7	58.5	0.488	6.06	2.74	25.03	66.17	49.5	17.3	48.3	3.5	51.5	14.76	49.06	4.175	"D"	"D"	"B"	"C"	"C"
270	"fortaleza"	"E. urophylla"	77.6	42.7	0.521	5.96	1.27	25.5	67.27	51.45	16.7	54.6	4.1	50.46	16.81	52.85	4.164	"D"	"C"	"D"	"D"	"D"
270	"fortaleza"	"E. urophylla"	77.6	42.7	0.517	4.53	1.35	26.86	67.26	50.52	18.5	49.3	4.5	50.79	16.92	49.63	4.152	"D"	"C"	"D"	"D"	"D"
270	"fortaleza"	"E. urophylla"	77.6	42.7	0.527	5.79	1.28	26.51	66.42	50.01	16.9	50.7	3.8	50.36	16.65	51.77	4.112	"D"	"D"	"D"	"D"	"D"
33	"campinas"	"E. grandis"	92.5	54.4	0.446	2.79	1.63	25.33	70.25	53.2	14.8	38.7	6.1	52.5	15.32	44.75	6.309	"A"	"A"	"B"	"B"	"A"
33	"campinas"	"E. grandis"	92.5	54.4	0.444	4.03	1.61	24.99	69.37	51.81	15.7	46.9	5.4	52.13	15.68	45.17	5.941	"B"	"B"	"B"	"B"	"B"

33	"campinas"	"E. grandis"	92.5	54.4	0.468	4.52	1.65	25.24	68.59	51.15	17.5	50.6	5.2	51.78	15.91	46.82	5.642	"B"	"C"	"B"	"B"	"B"
252	"fortaleza"	"E. urophylla"	89.2	33.4	0.542	4.57	1.42	24.86	69.35	50.62	17.7	49.8	4.7	50.9	17.02	46.53	4.419	"D"	"D"	"D"	"D"	"D"
252	"fortaleza"	"E. urophylla"	89.8	33.4	0.557	8.2	1.43	23.4	68.97	50	17.2	53.8	3.9	49.83	16.83	54.87	4.232	"D"	"D"	"D"	"D"	"D"
252	"fortaleza"	"E. urophylla"	89.8	33.4	0.563	6.91	1.5	23.98	67.61	49.1	17.1	51.8	4.9	49.94	17.21	52.03	4.353	"D"	"D"	"D"	"D"	"D"
262	"fortaleza"	"E. urophylla"	82.1	22.4	0.557	3.88	1.2	27.19	67.73	50.72	17.4	45.5	4.4	49.9	17.7	43.09	4.176	"D"	"D"	"D"	"D"	"D"
262	"fortaleza"	"E. urophylla"	82.4	22.4	0.57	5.15	1.21	26.71	66.93	50	17.8	46.4	4.2	49.56	17.78	44.37	4.226	"D"	"D"	"D"	"D"	"D"
262	"fortaleza"	"E. urophylla"	82.4	22.4	0.566	7.8	2.39	26.39	63.42	47.61	18.1	43	4.3	49.27	16.66	47.31	4.149	"D"	"D"	"D"	"D"	"D"
35	"fortaleza"	"E. urophylla"	92.5	51.8	0.52	6.4	2.4	25.19	66.01	49.73	14.7	50.5	4.7	50.49	15.91	52.67	4.047	"D"	"C"	"C"	"D"	"D"
35	"fortaleza"	"E. urophylla"	92.5	51.8	0.542	6.47	2.27	25.37	65.89	50.95	16.4	51.8	4.1	50.26	16.15	52.59	4.086	"D"	"D"	"D"	"D"	"D"
35	"fortaleza"	"E. urophylla"	92.5	51.8	0.544	7.72	2.41	26.13	63.74	50.34	16.4	58.3	3	50.16	15.37	51.63	3.951	"D"	"D"	"C"	"D"	"D"
50	"sta genoveva"	"E. grandis"	68.5	36.5	0.449	3.28	1.55	24.61	70.56	52.6	15.8	45.3	5.2	52.23	15.88	48.24	5.873	"B"	"A"	"B"	"B"	"A"
50	"sta genoveva"	"E. grandis"	68.5	36.5	0.456	3.76	1.57	24.73	69.94	51.19	16.3	45.3	6.2	52.01	16.22	48	5.717	"B"	"B"	"B"	"B"	"B"
50	"sta genoveva"	"E. grandis"	68.5	36.5	0.442	4.01	1.55	24.81	69.63	51.81	15.9	52.4	6.1	52.14	16.32	48.49	5.705	"B"	"B"	"B"	"B"	"B"
5	"sta genoveva"	"E. grandis"	79.3	25.6	0.472	5.18	2.35	27.89	64.58	51.49	17.1	45.2	5.7	51.95	16.2	45.45	5.656	"C"	"C"	"B"	"B"	"C"
5	"sta genoveva"	"E. grandis"	79.4	25.6	0.489	5.64	2.26	26.68	65.42	51.45	16	47.5	5.8	51.48	16.63	45.72	5.674	"C"	"B"	"C"	"C"	"C"
5	"sta genoveva"	"E. grandis"	79.4	25.6	0.502	4.8	2	26.87	66.33	51.25	16.2	47.1	5.5	51.6	16.56	46.74	5.674	"C"	"C"	"C"	"C"	"C"
talha	Fazenda	espécie	idade	prod	DensB	extrat	lignS	lignins	Holoc	RenD	K	Visc	aar	RenD	K	Visc	aar	Tro	Glib	RTr	RGI	Pro
72	"sta genoveva"	"E. grandis"	71.2	24.8	0.482	5.37	0.95	24.47	69.21	52.41	16.5	43.9	5.6	51.54	16.37	44.03	5.501	"C"	"C"	"C"	"C"	"C"
72	"sta genoveva"	"E. grandis"	71.2	24.8	0.481	4.33	1.15	24.66	69.86	51.93	16.3	45.3	5.6	51.69	16.27	44.87	5.53	"C"	"C"	"C"	"C"	"C"
72	"sta genoveva"	"E. grandis"	71.4	24.8	0.488	4.77	0.93	26.08	68.22	51.78	16.7	44.1	5.4	51.66	16.28	45.32	5.523	"C"	"C"	"C"	"C"	"C"
308	"ibiti"	"E. saligna"	69.3	49.1	0.485	3.48	1.48	25.25	69.79	50.72	16.2	43.5	5.4	51.02	15.99	42.62	5.278	"C"	"C"	"C"	"C"	"C"
308	"ibiti"	"E. saligna"	69.3	49.1	0.523	3.07	1.65	24.52	70.76	50.59	16	40.6	5.3	50.57	16.13	40.99	5.198	"C"	"C"	"C"	"C"	"C"
308	"ibiti"	"E. saligna"	69.3	49.1	0.527	3.14	1.65	27.76	67.45	50.43	15.9	44.8	5	50.35	16.57	40.72	5.02	"C"	"C"	"D"	"C"	"D"
310	"ibiti"	"E. saligna"	67.9	59.1	0.492	3.15	1.89	25.92	69.64	50.93	15	37.9	5.2	51.44	15.52	41.12	5.427	"B"	"C"	"B"	"C"	"C"
310	"ibiti"	"E. saligna"	67.9	59.1	0.5	4.67	1.8	25.32	68.01	52.12	14.7	41.3	5.3	51.28	14.87	39.65	5.152	"B"	"B"	"B"	"C"	"C"
310	"ibiti"	"E. saligna"	67.9	59.1	0.508	4.88	1.84	25.38	67.89	51.54	15.2	39.5	5.1	51.15	14.87	38.99	5.145	"B"	"C"	"C"	"C"	"C"
205	"ibiti"	"E. saligna"	88.5	93.7	0.45	2.64	1.66	24.01	71.69	52.44	15.5	39.9	6.3	52.29	15.25	40.95	6.012	"B"	"A"	"B"	"B"	"A"
205	"ibiti"	"E. saligna"	88.5	93.7	0.471	2.55	1.66	23.54	72.25	51.79	14.4	39	5.9	52.17	15.4	41.09	6.024	"B"	"B"	"B"	"B"	"B"
205	"ibiti"	"E. saligna"	88.5	93.7	0.445	3.11	1.71	23.51	71.67	51.68	15.7	41.9	5.9	52.2	15.28	41.05	5.928	"B"	"B"	"B"	"B"	"A"
237	"ibiti"	"E. saligna"	83.3	69.9	0.471	2.05	1.68	24.3	71.97	52.07	15	41.4	6.3	52.29	15.48	42.1	6.128	"B"	"A"	"B"	"B"	"B"

237	"ibiti"	"E. saligna"	83.3	69.9	0.464	2.67	1.57	24.37	71.39	52.49	15.2	44.4	6.2	52.38	15.27	42.61	6.01	"B"	"B"	"B"	"B"	"B"
237	"ibiti"	"E. saligna"	83.3	69.9	0.466	3.25	1.71	24.08	70.96	52.42	16.1	40.7	6.7	52.32	15.23	42.88	5.851	"B"	"B"	"B"	"B"	"B"
230	"ibiti"	"E. saligna"	84.7	46.9	0.487	4.36	1.95	25.33	68.26	51.86	14.9	36.6	5.6	51.8	15.01	43.05	5.274	"B"	"B"	"B"	"B"	"C"
230	"ibiti"	"E. saligna"	84.7	46.9	0.501	4.39	1.89	25.72	68	50.95	15.7	41.2	5.4	51.36	15.07	42.51	5.179	"C"	"C"	"B"	"C"	"C"
230	"ibiti"	"E. saligna"	84.7	46.9	0.489	4.46	1.8	25.03	68.71	52.32	15.4	43.3	4.7	51.61	15	43.25	5.283	"B"	"C"	"B"	"C"	"C"
250	"ibiti"	"E. dunni"	75.1	43.9	0.537	2.97	2.02	25.11	69.9	51.11	16.9	39.6	5.3	50.34	17.4	41.73	5.101	"D"	"C"	"D"	"D"	"D"
250	"ibiti"	"E. dunni"	75.1	43.9	0.542	4.11	2.14	23.73	70.02	50.27	16.6	41.3	5	50.03	17.16	39.31	5.2	"D"	"C"	"D"	"D"	"D"
250	"ibiti"	"E. dunni"	75.1	43.9	0.514	2.55	2.21	23.9	71.34	50.76	16.9	44.8	4.7	50.8	16.54	42.05	5.441	"D"	"C"	"C"	"C"	"C"
253	"ibiti"	"E. dunni"	78.1	54.4	0.531	3.66	2.04	25.02	69.28	49.72	17.8	43	4.9	50.52	17.13	39.82	5.067	"D"	"D"	"D"	"C"	"D"
253	"ibiti"	"E. dunni"	78.1	54.4	0.573	3.18	2.14	23.96	70.72	49.83	15.7	32.8	6.3	50.18	16.32	38.54	5.256	"D"	"C"	"D"	"C"	"C"
253	"ibiti"	"E. dunni"	78.1	54.4	0.538	3.12	2.21	24.04	70.63	49.71	16.4	37.1	6.1	50.63	16.33	39.66	5.333	"D"	"C"	"C"	"C"	"C"
210	"ibiti"	"E. dunni"	85.2	57.3	0.531	4.6	1.34	25.25	68.81	51.11	15.6	40.9	5.2	51.29	16.5	40.77	5.063	"C"	"C"	"C"	"C"	"C"
210	"ibiti"	"E. dunni"	85.2	57.3	0.529	5.62	1.52	24.98	67.88	51.04	17.2	43.5	4.8	50.77	17.08	40.89	4.936	"D"	"D"	"D"	"C"	"D"
210	"ibiti"	"E. dunni"	85.2	57.3	0.515	5.05	1.45	25.74	67.76	50.73	17.9	40.2	4.8	51.23	16.93	41.32	4.903	"D"	"D"	"C"	"C"	"C"
talha	Fazenda	espécie	idade	prod	DensB	extrat	lignS	lignns	Holoc	RenD	K	Visc	aar	RenD	K	Visc	aar	Tro	Gilb	RTr	RGl	Pro
212	"ibiti"	"E. dunni"	85.2	45.7	0.535	6.43	2.23	23.9	67.44	49.82	17.9	39.8	4.9	49.66	17.85	38.47	4.872	"D"	"D"	"D"	"D"	"D"
212	"ibiti"	"E. dunni"	85.2	45.7	0.532	5	2.16	24.03	68.81	51.26	16	37.3	5.4	50.23	17.19	39.04	5.132	"C"	"C"	"D"	"D"	"D"
212	"ibiti"	"E. dunni"	85.2	45.7	0.538	4.74	2.32	23.59	69.35	50.08	18.8	39.1	4.5	50.22	16.8	38.58	5.282	"D"	"D"	"D"	"C"	"C"
196	"ibiti"	"E. dunni"	94.6	72.9	0.554	4.94	2.22	24.39	68.45	52.16	13.4	41.7	6.1	51.79	14.73	40.57	5.484	"C"	"C"	"C"	"C"	"C"
196	"ibiti"	"E. dunni"	94.6	72.9	0.561	3.92	1.44	23.57	71.07	52.6	13	43	6	52.72	13.65	40.97	6.023	"C"	"C"	"C"	"C"	"B"
196	"ibiti"	"E. dunni"	94.6	72.9	0.549	4.11	1.35	23.77	70.77	53.04	14.5	40.1	5.5	52.85	13.67	41.54	5.996	"C"	"C"	"C"	"C"	"B"
433	"ibiti"	"E. grandis"	59.2	44.5	0.466	5.41	1.87	23.92	68.8	52.88	15.7	39.2	5.8	52.03	16.08	40.83	5.591	"B"	"B"	"B"	"B"	"B"
433	"ibiti"	"E. grandis"	59.2	44.5	0.472	5.39	1.89	24.94	67.79	52.01	15.7	40.1	5.7	51.69	16.78	40.88	5.338	"B"	"B"	"C"	"C"	"C"
433	"ibiti"	"E. grandis"	59.2	44.5	0.46	4.8	1.88	24.62	68.7	53.63	14.9	38	5.5	52.11	16.4	41.17	5.472	"B"	"A"	"B"	"B"	"B"
305	"ibiti"	"E. grandis"	68.9	63.7	0.439	4.27	1.84	25.91	67.98	53.15	15.2	43.8	5.7	53.54	15.55	43.53	5.773	"A"	"B"	"A"	"B"	"B"
305	"ibiti"	"E. grandis"	68.9	63.7	0.455	5.47	1.75	25.82	66.96	52.77	15.9	43.3	5.3	53.11	15.68	43.83	5.54	"A"	"B"	"B"	"B"	"B"
305	"ibiti"	"E. grandis"	68.9	63.7	0.453	5.12	1.84	24.81	68.23	53.37	15.3	43.7	5.7	53.95	15.04	42.34	5.817	"A"	"B"	"B"	"B"	"B"
412	"ibiti"	"E. grandis"	69.8	95.4	0.414	4.06	1.84	25.38	68.72	54.4	14.3	39.2	6.9	53.85	14.09	42.56	6.255	"A"	"A"	"A"	"A"	"A"
412	"ibiti"	"E. grandis"	69.8	95.4	0.396	2.66	2.03	25.16	70.15	54.52	14.4	41.5	6.3	53.95	14.33	40.07	6.637	"A"	"A"	"A"	"A"	"A"
412	"ibiti"	"E. grandis"	69.8	95.4	0.39	4.08	1.98	25.05	68.89	53.66	14.9	45.3	6.2	53.89	13.99	43.6	6.344	"A"	"A"	"A"	"A"	"A"

244	"bifl"	"E. grandis"	85.5	61.3	0.448	3.91	1.79	24.93	68.37	53.86	13.9	39.9	6.4	53.91	13.99	39.85	6.376	"A"	"A"	"A"	"A"	"A"
244	"bifl"	"E. grandis"	85.5	61.3	0.444	4.36	1.74	23.99	69.91	54.42	13.4	37.8	6.5	53.85	13.7	39.09	6.453	"A"	"A"	"A"	"A"	"A"
244	"bifl"	"E. grandis"	85.5	61.3	0.447	4.54	1.76	24.76	68.82	53.36	13.7	40.2	6.5	53.8	14.14	39.52	6.201	"A"	"A"	"A"	"A"	"A"
222	"bifl"	"E. grandis"	88.1	86.1	0.458	4.55	1.76	24.96	68.73	53.32	13.9	40.1	6.5	53.53	13.63	35.54	6.146	"A"	"A"	"A"	"A"	"A"
222	"bifl"	"E. grandis"	88.1	86.1	0.455	4.84	1.86	24.87	68.43	53.99	13.4	33.1	6.1	53.44	13.7	35.66	6.051	"A"	"A"	"A"	"A"	"A"
222	"bifl"	"E. grandis"	88.1	86.1	0.443	4.44	1.89	25.72	67.95	53.22	13.5	32.4	5.6	53.61	14.04	36.73	6.017	"A"	"A"	"A"	"A"	"A"
280	"guaruja"	"E. grandis"	71.3	73.5	0.464	3.73	1.23	24.42	70.62	54.09	15.2	40.7	6	53.33	14.74	44.49	6.189	"A"	"B"	"B"	"B"	"A"
280	"guaruja"	"E. grandis"	71.3	73.5	0.423	3.61	1.25	25.32	69.82	53.27	15.5	40.8	5.7	53.65	15.47	45.49	6.169	"A"	"B"	"A"	"A"	"B"
280	"guaruja"	"E. grandis"	71.3	73.5	0.44	3.81	1.21	24.59	70.39	53.67	14.7	39.6	5.7	53.52	14.95	44.79	6.219	"A"	"A"	"A"	"A"	"A"
211	"guaruja"	"E. grandis"	72.8	96.1	0.435	3.88	1.38	23.27	71.47	53.35	14.2	45.6	5.9	53.26	14.56	38.54	6.402	"A"	"A"	"A"	"A"	"A"
211	"guaruja"	"E. grandis"	72.8	96.1	0.429	3.73	1.29	24.82	70.16	52.39	15.3	47.9	6	53.5	14.53	39.72	6.322	"A"	"B"	"A"	"A"	"A"
211	"guaruja"	"E. grandis"	72.8	96.1	0.443	1.79	1.2	24.78	72.23	52.88	15.4	44.6	5.9	53.46	14.67	41.32	6.757	"A"	"B"	"A"	"A"	"A"
latha	Fazenda	espécie	idade	prod	DensB	extrat	lignS	lignins	Holoc	RenD	K	Visc	aar	RenD	K	Visc	aar	Tro	Glib	RTr	RGI	Pro
70	"guaruja"	"E. grandis"	75.4	64.2	0.474	3.55	1.54	25.4	69.51	53.81	15.4	44.8	6.1	53.17	15.53	45.86	5.95	"B"	"B"	"B"	"B"	"B"
70	"guaruja"	"E. grandis"	75.4	64.2	0.467	3.34	1.79	24.53	70.34	53.7	15.2	43.1	6.4	53.3	14.81	45.07	6.216	"B"	"B"	"B"	"B"	"B"
70	"guaruja"	"E. grandis"	75.4	64.2	0.451	3.54	1.73	24.7	70.03	52.89	15.5	46.2	6.1	53.42	15.07	45.22	6.178	"B"	"B"	"A"	"B"	"B"
204	"guaruja"	"E. grandis"	85.1	42.7	0.492	3.32	1.15	26.26	69.27	53.68	14.3	42.6	6.3	53.49	14.82	46.13	6.066	"B"	"B"	"B"	"B"	"B"
204	"guaruja"	"E. grandis"	85.1	42.7	0.474	3.71	1.07	26.15	69.07	53.32	15	44.1	6.1	53.59	14.86	46.25	6.059	"B"	"B"	"B"	"B"	"B"
204	"guaruja"	"E. grandis"	85.1	42.7	0.479	4.02	1.06	26	68.93	53.41	15.3	51.4	6.2	53.51	14.84	45.76	5.961	"B"	"B"	"B"	"B"	"B"
203	"guaruja"	"E. grandis"	85.3	58.3	0.494	4.6	2.25	24.6	68.55	53.13	15.4	45.6	6	52.71	15.3	42.27	5.742	"B"	"B"	"B"	"B"	"B"
203	"guaruja"	"E. grandis"	85.3	58.3	0.478	4.29	2.18	24.26	69.27	52.89	15.7	44.6	5.6	52.97	14.9	42.72	5.98	"B"	"B"	"B"	"B"	"B"
203	"guaruja"	"E. grandis"	85.3	58.3	0.467	5.32	1.9	24.25	68.53	53.32	14.8	43.6	5.9	52.85	15.33	42.67	5.696	"B"	"B"	"B"	"B"	"B"
184	"guaruja"	"E. grandis"	80.6	79.2	0.445	4.44	1.97	25.25	68.34	52.42	14.6	41	5.9	53.26	15.16	40.94	5.81	"A"	"A"	"A"	"B"	"B"
184	"guaruja"	"E. grandis"	80.6	79.2	0.437	4.41	2.25	25.12	68.22	53.29	14.7	40.2	5.9	53.33	15.03	41.07	5.878	"A"	"A"	"A"	"A"	"B"
184	"guaruja"	"E. grandis"	80.6	79.2	0.458	3.73	2.24	24.6	69.43	52.57	14.3	36.5	5.8	53.31	14.42	39.72	6.195	"A"	"A"	"A"	"A"	"A"
283	"guaruja"	"E. saligna"	75.6	49.4	0.481	3.58	1.47	26.35	68.6	51.67	15.1	42.7	5.7	51.77	15.67	46.79	5.403	"B"	"B"	"C"	"B"	"C"
283	"guaruja"	"E. saligna"	75.6	49.4	0.471	3.38	1.23	25.18	70.21	51.89	15.2	46.9	5.2	52.1	15.31	45.77	5.581	"B"	"B"	"B"	"B"	"B"
283	"guaruja"	"E. saligna"	75.6	49.4	0.501	3.76	1.33	28.26	68.63	50.84	15.4	47	5.2	51.49	15.66	45.35	5.258	"B"	"C"	"C"	"C"	"C"
177	"guaruja"	"E. saligna"	83.9	67	0.48	3.98	1.75	24.79	69.48	52.71	14.8	47.1	6	52.41	14.65	45.71	5.581	"B"	"B"	"B"	"B"	"B"
177	"guaruja"	"E. saligna"	83.9	67	0.473	3.54	1.92	25.27	69.27	52.07	14.8	44.8	5.3	52.54	14.66	45.86	5.699	"B"	"B"	"B"	"B"	"B"

177	"guarujá"	"E. saligna"	83.9	67	0.447	4.57	1.8	25.25	68.38	52.93	15	46	5.4	52.65	14.43	47.48	5.431	"B"	"A"	"A"	"A"	"B"
179	"guarujá"	"E. saligna"	84.6	60.6	0.473	3.57	1.72	25.36	69.35	52.34	15.2	45.2	5.3	52.48	14.71	46.15	5.659	"B"	"B"	"B"	"B"	"B"
179	"guarujá"	"E. saligna"	84.6	60.6	0.497	3.56	1.81	25.1	69.53	52.97	14	44.2	6	52.27	14.63	45.18	5.602	"B"	"B"	"B"	"B"	"B"
179	"guarujá"	"E. saligna"	84.6	60.6	0.471	3.34	1.86	25.05	69.75	53.27	14.5	49	5.9	52.49	14.81	45.93	5.759	"B"	"B"	"B"	"B"	"B"
58	"saligna"	"E. grandis"	71.4	57.2	0.443	3.99	1.25	24.58	70.18	51.97	15.2	47.3	5.8	52.57	15.21	43.8	5.505	"B"	"B"	"A"	"B"	"B"
58	"saligna"	"E. grandis"	71.4	57.2	0.436	4.02	1.25	25.37	69.36	52.55	16.1	42.6	5.4	52.65	15.47	44.68	5.349	"B"	"B"	"A"	"B"	"B"
58	"saligna"	"E. grandis"	71.4	57.2	0.461	3.61	1.23	24.54	70.62	53.25	15.3	37.8	5.8	52.46	15.14	43.29	5.622	"B"	"B"	"B"	"B"	"B"
68	"saligna"	"E. grandis"	71.4	39.4	0.464	4.33	1.27	24.24	70.16	53.22	15.8	44.2	5.4	52.17	15.81	45.33	5.347	"B"	"B"	"B"	"B"	"B"
68	"saligna"	"E. grandis"	71.4	39.4	0.455	4.81	1.27	24.17	69.75	51.57	15.7	44.2	5.2	52.18	15.77	45.66	5.241	"B"	"B"	"B"	"B"	"B"
68	"saligna"	"E. grandis"	71.4	39.4	0.439	4.62	1.38	24.54	69.48	51.76	15.8	46.6	5.2	52.39	15.76	46.37	5.276	"B"	"B"	"B"	"B"	"B"
latna	Fazenda	espécie	idade	prod	DensB	extrat	lignS	lignns	Holoc	RenD	K	Visc	aar	RenD	K	Visc	aar	Tro	Gilb	RTI	RGi	Pro
65	"saligna"	"E. grandis"	71.3	30.4	0.456	4.54	2.41	24.51	68.54	53.36	15.2	44.3	4.3	52.29	15.67	44.19	5.318	"B"	"B"	"B"	"B"	"B"
65	"saligna"	"E. grandis"	71.3	30.4	0.457	4.91	2.48	24.91	67.7	52.43	15.4	47.1	5.8	52.28	15.53	44.68	5.249	"B"	"B"	"B"	"B"	"B"
65	"saligna"	"E. grandis"	71.3	30.4	0.464	5.79	2.78	26.06	65.37	51.21	15.7	45.4	5.6	52.14	15	47.67	5.204	"B"	"B"	"B"	"B"	"B"
12	"campinas"	"E. grandis"	72	84	0.418	3.18	1.65	24.92	70.25	53.51	15.5	49.5	6.3	52.61	15.32	46.8	6.421	"A"	"A"	"A"	"A"	"A"
12	"campinas"	"E. grandis"	72	84	0.426	3.24	1.74	25.15	69.87	52.53	15.8	58.5	5.8	52.6	15.28	46.5	6.334	"B"	"B"	"A"	"B"	"A"
12	"campinas"	"E. grandis"	72	84	0.409	4.56	1.82	25.23	68.39	52.78	14.5	44.9	6.4	52.62	15.26	48.03	5.828	"A"	"A"	"A"	"A"	"B"
187	"penha"	"E. grandis"	62.7	81	0.434	3.47	1.84	24.5	70.19	53.84	14.6	45.5	6.4	53.42	13.94	42.42	6.259	"A"	"A"	"A"	"A"	"A"
187	"penha"	"E. grandis"	62.7	81	0.431	3.69	1.88	23.81	70.62	54.44	14.7	49.2	6.5	53.43	13.88	42.21	6.307	"A"	"A"	"A"	"A"	"A"
187	"penha"	"E. grandis"	62.7	81	0.403	3.54	1.96	24.37	70.13	53.6	14.7	41.3	6.3	53.62	13.91	43.64	6.313	"A"	"A"	"A"	"A"	"A"
96	"penha"	"E. grandis"	73.9	29.6	0.472	4.47	0.99	24.55	69.99	51.98	15.2	41.6	5.4	52.13	16.16	46.21	5.62	"B"	"B"	"B"	"B"	"B"
96	"penha"	"E. grandis"	73.9	29.6	0.493	4.58	1.01	23.63	70.78	52.01	15.2	45.1	5.2	51.91	15.7	44.03	5.661	"B"	"C"	"C"	"C"	"B"
96	"penha"	"E. grandis"	73.9	29.6	0.469	4.27	0.98	24.12	70.63	52.37	15.2	44.3	5.4	52.29	15.82	45.67	5.727	"B"	"B"	"B"	"B"	"B"
94	"penha"	"E. grandis"	73.9	56.9	0.453	3.83	0.98	23.39	71.8	52.89	14.6	48.8	5.4	52.85	14.39	43.52	6.301	"A"	"A"	"A"	"A"	"A"
94	"penha"	"E. grandis"	73.9	56.9	0.467	4.11	0.94	22.8	72.15	53.32	14.3	41.7	6.4	52.71	14.24	42.44	6.279	"A"	"A"	"A"	"A"	"A"
94	"penha"	"E. grandis"	73.9	56.9	0.436	4.18	0.97	23.43	71.42	53.87	14.2	51	6	52.92	14.57	44.23	6.263	"A"	"A"	"A"	"A"	"A"