

UNIVERSIDADE FEDERAL DO PARANÁ

BIBIANA RIBEIRO RUBINI

DESENVOLVIMENTO DE MODELOS MATEMÁTICOS PARA UMA
PLANTA DE DESLIGNIFICAÇÃO DE POLPA CELULÓSICA COM
OXIGÊNIO

CURITIBA

2006

BIBIANA RIBEIRO RUBINI

DESENVOLVIMENTO DE MODELOS MATEMÁTICOS PARA UMA PLANTA DE
DESLIGNIFICAÇÃO DE POLPA CELULÓSICA COM OXIGÊNIO

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia, Setor de Tecnologia, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Mestre em Engenharia Química.

Orientador: Prof. Dr. Carlos Itsuo Yamamoto

CURITIBA

2006

Aos meus pais, Luiz e Arlete,
pelo incentivo, apoio, e pelo amor de sempre –
eu devo muito (tudo) a vocês, aliás sou fã incondicional de vocês.

Aos meus irmãos, Guilherme e Leila,
fundamentais na minha vida;

Ao Elias, minha inspiração,
pela amizade, compreensão, carinho e amor.

Agradecimentos

A Deus, por todas as oportunidades e bênçãos da minha vida.

Ao Prof. Dr. Carlos Itsuo Yamamoto, pela orientação, incentivo, apoio e dedicação em todas as etapas do desenvolvimento deste projeto e por sua amizade, que foi tão importante durante este período.

Aos Professores Dr. Luiz Augusto da Cruz Meleiro, Dr. Everton Fernando Zanoelo e Dr. Luiz Fernando de Lima Luz Júnior, por toda a ajuda, apoio, sugestões e pela amizade engrandecedora.

À Universidade Federal do Paraná, que me acolhe desde a graduação, e ao Programa de Pós-Graduação em Engenharia, PIPE, pela oportunidade de realizar este trabalho.

À Klabin Celulose e Papel Unidade Monte Alegre, em especial aos Engenheiros Sidney Zattoni e Osvaldo Vieira e à Carolina Stonoga, pelo apoio financeiro, pelo fornecimento dos dados necessários para o desenvolvimento deste trabalho e pelo voto de confiança.

A minha família, que está sempre comigo e que eu amo tanto.

Aos amigos de pós-graduação, em especial Aleksandra, Artem, Camila, Débora, Elina, Emerson, Leilie, Luciana, Márcio, Marisa, Nice, Patrícia, Sandra e Viviane, pelo apoio mútuo e pela amizade. A presença de vocês nestes dois anos de trabalho foi fundamental para mim.

À todos aqueles que embora não tenham sido mencionado contribuíram direta ou indiretamente para a realização deste trabalho.

*“Não basta abrir a janela
Para ver os campos e os rios
Não é bastante não ser cego
Para ver as árvores e as flores.
É preciso também não ter filosofia nenhuma.
Com filosofia não há árvores, há idéias apenas.
Há só cada um de nós como uma cave
Há só uma janela fechada e todo o mundo lá fora,
E um sonho do que se poderia ver se a janela se abrisse
Que nunca é o que se vê quando se abre a janela.”*

(Fernando Pessoa)

Resumo

O objetivo desta dissertação é desenvolver modelos matemáticos capazes de simular a planta de deslignificação com oxigênio da Klabin Celulose e Papel – Unidade Monte Alegre, em Telêmaco Borba – PR. Os modelos podem ser ferramentas úteis para uma fábrica de celulose, podendo ser utilizados para simular e otimizar o processo. As condições de operação podem ser alteradas sem modificar efetivamente a produção; limites podem ser estipulados para as principais variáveis do processo e, desta forma, o processo pode ser controlado de forma mais econômica, sem perder qualidade no produto. Neste trabalho foram desenvolvidos três modelos fenomenológicos e, para fins comparativos, identificou-se o processo utilizando modelos caixa-preta lineares e não lineares. Os modelos fenomenológicos foram desenvolvidos a partir de modelos encontrados na literatura e a identificação do processo foi feita com o modelo ARX linear e redes neurais. Todos os modelos propostos apresentaram erros relativos médios em torno de 8,5%, da mesma ordem de grandeza dos modelos encontradas na literatura. Os resultados dos modelos fenomenológicos indicam que existe a possibilidade de a planta estar operando com excesso de reagentes.

Palavras-Chave: Deslignificação com Oxigênio, Modelagem Fenomenológica, Identificação de Sistemas, Redes neurais, Cinética de Deslignificação.

Abstract

The objective of this work is to derive mathematical models capable of predicting the outlet kappa number of the delignification reactor of Klabin Celulose e Papel Mill – Unit Monte Alegre, in Telêmaco Borba – PR. These models are a useful tool for cellulose mills once they can be applied to process simulation, optimization and control. Moreover, different operational conditions can be tested off-line and limits can be established for the main process variables so that the mill can be controlled economically without product quality loss. This work proposes three phenomenological models, and for meanings of comparison, linear and non-linear models were identified for the delignification plant. The phenomenological models were based on models found in the literature and the process was also identified with a linear ARX and neural networks models. All the proposed models presented prediction errors comparable to those found in the literature and they can be applied indistinctly to process daily simulation and can be used as a performance indicator of the process. Nevertheless, once phenomenological models use and offer more information about the process so they are preferred over the other ones.

Key-words: Oxygen Delignification, Delignification Kinetic, Kinetic Model, System Identification, Neural Networks

Lista de Figuras

Figura 1.1 - Crescimento da deslignificação ao longo das últimas décadas	19
Figura 2.1 - Esquema da planta de deslignificação da Klabin Celulose e Papel – Unidade Monte Alegre.....	25
Figura 3.1 - Estágios da reação de deslignificação com oxigênio de celulose.....	31
Figura 3.2 - Analogia entre os neurônios biológico e matemático	52
Figura 3.3 - Modelo não linear de neurônio.....	55
Figura 3.4 - Arquitetura de uma rede MLP	58
Figura 3.5 - Esquema de treinamento <i>backpropagation</i>	61
Figura 3.6 - Conexões entre os neurônios <i>i</i> e <i>k</i>	61
Figura 4.1 - Algoritmo do Recozimento Simulado	80
Figura 4.2 - Operador genético de cruzamento.....	85
Figura 4.3 - Operador genético de mutação.....	86
Figura 4.4 - Estrutura básica do Algoritmo Genético.....	87
Figura 5.1 - Número <i>kappa</i> de saída do reator normalizado no conjunto de 1331 padrões	91
Figura 5.2 - Número <i>kappa</i> de saída do reator normalizado no conjunto de 812 padrões	92
Figura 5.3 - Vazão de álcali normalizada na entrada do reator de deslignificação....	95
Figura 5.4 - Variação do número <i>kappa</i> na entrada do reator normalizado	96
Figura 5.5 - Fluxograma do procedimento utilizado para o desenvolvimento do modelo fenomenológico	98
Figura 5.6 - Fluxograma do procedimento utilizado no desenvolvimento do modelo neural	102

Figura 6.1 - Comparação entre os valores do número <i>kappa</i> reais e estimados pelo Modelo 1	108
Figura 6.2 - (a) Erro absoluto e (b) erro relativo médio calculados para o Modelo 1	109
Figura 6.3 - Comparação entre os valores calculados pelo modelo e o desvio padrão da média diária do número <i>kappa</i> para o Modelo 1	110
Figura 6.4 - (a) Perfil de decaimento do número <i>kappa</i> e; (b) perfil de consumo de álcali ao longo da reação de deslignificação indicado pelo Modelo 1	111
Figura 6.5 - Comparação entre os valores reais e estimados pelo modelo análogo a Agarwal et al (1999)	113
Figura 6.6 - Comparação entre os valores reais e estimados pelo Modelo 2 com valor da ordem em relação ao número <i>kappa</i> estimado pelo AG	114
Figura 6.7 - (a) Erro absoluto e; (b) erro relativo médio calculados para o Modelo 2	115
Figura 6.8 - Comparação entre os valores calculados pelo modelo e o desvio padrão da média diária do número <i>kappa</i> para o Modelo 2	116
Figura 6.9 - (a) Perfil de decaimento do número <i>kappa</i> e; (b) perfil de consumo de álcali ao longo do tempo de reação gerado pelo Modelo 2	117
Figura 6.10 - Comparação entre os valores reais e estimados pelo Modelo 3.....	119
Figura 6.11 - (a) Erro absoluto e; (b) erro relativo médio calculados para o Modelo 3	120
Figura 6.12 - Comparação entre os valores calculados pelo modelo e o desvio padrão da média diária do número <i>kappa</i> para o Modelo 3.....	121
Figura 6.13 - (a) Perfil de decaimento do número <i>kappa</i> e; (b) perfil de consumo de álcali ao longo da reação de deslignificação para o Modelo 3	122

Figura 6.14 - Comparação entre os valores reais e estimados pelo Modelo 4.....	125
Figura 6.15 - (a) Erro absoluto e; (b) erro relativo médio calculados para o Modelo 4	126
Figura 6.16 - Comparação entre os valores calculados pelo modelo e o desvio padrão da média diária do número <i>kappa</i> para o Modelo 4.....	127
Figura 6.17 - Comparação entre os valores reais e estimados pelo Modelo 5.....	128
Figura 6.18 - (a) Erro absoluto e; (b) erro relativo médio calculados para o Modelo 4	129
Figura 6.19 - Comparação entre os valores calculados pelo modelo e o desvio padrão da média diária do número <i>kappa</i> para o Modelo 4.....	130
Figura 6.20 - Análise do EQM para a escolha das melhores estruturas neurais.....	131
Figura 6.21 - Comparação entre os valores reais e estimados pelo Modelo 6.....	133
Figura 6.22 - Erro absoluto calculado para o Modelo 6.....	133
Figura 6.23 - Comparação entre os valores calculados pelo modelo e o desvio padrão da média diária do número <i>kappa</i> para o Modelo 6.....	134
Figura 6.24 - Comparação entre os valores reais e estimados pelo Modelo 7.....	135
Figura 6.25 - Erro relativo médio calculado para o Modelo 7	135
Figura 6.26 - Erro absoluto calculado para o Modelo 7.....	136
Figura 6.27 - Comparação entre os valores calculados pelo modelo e o desvio padrão da média diária do número <i>kappa</i> para o Modelo 7.....	136

Lista de Tabelas

Tabela 5.1 - Análise estatística básica e análise <i>t</i> de Student de significância	93
Tabela 5.2 - Análise estatística básica para as variáveis do processo.....	94

Lista de Siglas e Abreviaturas

<i>AG</i>	Algoritmo Genético
<i>ARX</i>	Modelo Auto-Regressivo com Entradas Externas (<i>Auto-Regressive with Exogenous Inputs</i>)
<i>DBO</i>	Demanda bioquímica de oxigênio
<i>DQO</i>	Demanda química de oxigênio
<i>DP</i>	Grau de polimerização
<i>EQM</i>	Erro quadrático médio
<i>MLP</i>	<i>Perceptron</i> de múltiplas camadas (<i>Multi-layer perceptron</i>)
<i>MIMO</i>	Múltiplas entradas e múltiplas saídas (<i>multiple input, multiple output</i>)
<i>MISO</i>	Múltiplas entradas e uma saída (<i>multiple input, single output</i>)
<i>NCC</i>	Número de cadeias celulósicas
<i>NCQ</i>	Número de cadeias celulósicas quebradas
<i>RNA</i>	Rede neural artificial
<i>SIMO</i>	Uma entrada e múltiplas saídas (<i>single input, multiple output</i>)
<i>SISO</i>	Uma entrada e uma saída (<i>single input, single output</i>)

Lista de Símbolos

A, B	Parâmetros estimados nos modelos ajustados
a, b, c, d, f	Parâmetros estimados nos modelos ajustados
b_k	Bias do neurônio k
D_m	Subconjunto de \mathfrak{R}^p
d_i^p	Resposta desejada para o padrão p no neurônio i
d	Múltiplo do período de amostragem
E_a	Energia de ativação, em KJ/mol
E^p	Erro quadrático para o padrão p
$\{E_{ri}\}$	Energia da configuração r
F_1, F_2	Função de ativação linear
F_{NaOH}	Vazão de álcali, em m^3/h
$F_{solução}$	Vazão da solução alcalina, em m^3/h
G	Função de transferência
H	Função de transferência
K	Número <i>kappa</i>
$K_{inicial}$	Número <i>kappa</i> na entrada do reator
$K_{saída}$	Número <i>kappa</i> na saída do reator
k	Constante da equação da taxa de reação
k_b	Constante de Boltzmann, em J/K
k_0	Fator de frequência da expressão de Arrhenius
k_1, k_2	Constantes da equação da taxa de reação em duas etapas
k_3, k_4	Constantes da equação da taxa de reação de quebra de cadeia celulósica

K_f, K_s, K_b	Número <i>kappa</i> das ligninas de reação rápida, lenta e que não reage, respectivamente
$K_{estimado}$	Número <i>kappa</i> estimado pela rede neural ou pelo modelo fenomenológico
K_{real}	Número <i>kappa</i> da polpa
K_{01}, K_{02}	Número <i>kappa</i> das ligninas de remoção fácil e difícil, respectivamente
\mathcal{M}	Estrutura do modelo
N_0	Camada de saída do neurônio
Na	Sódio
$NaOH$	Hidróxido de sódio
N	Número de pontos amostrados
nT	Número de iterações na mesma temperatura
nR	Número de ciclos de redução de temperatura
O_2	Quantidade de oxigênio consumida
OH^-	Íon hidroxila
$[OH^-]$	Concentração de álcali, em g/L
$P(E)$	Probabilidade da configuração ser aceita
P_{O_2}	Pressão parcial do oxigênio
p	Vetor de entrada de padrões na rede neural
pH	Potencial Hidrogeniônico
q^{-1}	Operador de atraso
R	Constante universal dos gases
r	Configuração atômica
$\eta_{[OH^-]}$	Taxa de consumo de álcali durante a reação
S	Enxofre

S_x	Desvio padrão do conjunto de pontos do intervalo
t	Tempo de residência do reator, em <i>min</i>
t	Valor de t da distribuição t de Student
t_{polpa}	Toneladas de polpa
T	Temperatura da reação, em K
\tanh	Tangente hiperbólica
$u(t)$	Função para indicar divisão de tempo no modelo de Hsu e Hsieh (1988)
u_k	Saída do combinador linear
v	Potencial de ativação do neurônio
w_{ij}	Peso aferido à ligação do neurônio i com o neurônio j
x_1, x_2, x_m	Sinais de entrada
\bar{x}	Valor médio do número $kappa$ de saída do reator
x_{norm}	Variável normalizada
$\{x_{ij}\}$	Conjunto de variáveis
x_{inf}, x_{sup}	Limites inferior e superior do intervalo de normalização, respectivamente
y_k	Sinal de saída do neurônio k
\hat{y}_k	Sinal de saída estimado do neurônio k

Símbolos Gregos

α	Parâmetro estimado pelo modelo ajustado, parâmetro de inclinação
γ	Tamanho do passo de correção
η	Viscosidade intrínseca da polpa, em <i>mL/g</i>
μ_0	Valor esperado do número <i>kappa</i> de saída do reator
φ	Função de ativação
ΔE	Variação de energia
Δf	Variação da função otimizada
ΔK	Variação do número <i>kappa</i>
$\Delta[OH]$	Variação da concentração de álcali, em <i>g/L</i>
Δw	Variação no peso da rede neural
Σ	Somatório
ω_{ij}	Peso aferido à ligação do neurônio <i>i</i> ao neurônio <i>j</i>
$v(k)$	Ruído branco

Sumário

Agradecimentos	
Resumo	
Abstract	
Lista de Figuras	
Lista de Tabelas	
Lista de Siglas e Abreviaturas	
Lista de Símbolos	
Capítulo 1 - Introdução.....	18
1.1 Motivação para o Trabalho	18
1.2 Objetivos do Trabalho.....	20
1.3 Estrutura da Dissertação	20
Capítulo 2 - O Processo de Deslignificação	22
2.1 Introdução.....	22
2.2 Descrição do Processo	24
Capítulo 3 - Modelagem Matemática.....	27
3.1 Introdução.....	27
3.2 Modelo Fenomenológico Generalizado da Deslignificação.....	30
3.2.1 O Modelo de Olm e Teder	35
3.2.2 O Modelo de Hsu e Hsieh	38
3.2.3 O Modelo de Myers e Edwards	40
3.2.4 O Modelo de Agarwal, Genco, Cole e Miller	42
3.2.5 O Modelo de Violette.....	43
3.3 Identificação de Sistemas.....	46

3.3.1 Identificação de sistemas com modelos lineares.....	47
3.3.1.1 Representação matemática no tempo discreto.....	49
3.3.2 Identificação de sistemas com modelos não-lineares	51
3.3.2.1 Perceptron de Múltiplas Camadas.....	54
3.3.2.2 Descrição do Procedimento e dos Critérios.....	65
3.3.2.3 Modelos Neurais Aplicados à Indústria de Celulose e Papel.....	68
Capítulo 4 - Métodos de Otimização	73
4.1 Introdução.....	73
4.2 Métodos Heurísticos	76
4.2.1 Recozimento Simulado	76
4.2.2 Algoritmo Genético.....	81
Capítulo 5 - Metodologia	88
5.1 Coleta de Dados.....	88
5.2 Tratamento dos Dados	90
5.3 Desenvolvimento do Modelo Fenomenológico	96
5.4 A Identificação com o Modelo Linear ARX.....	99
5.5 a Identificação com o Modelo Neural.....	100
5.5 Emprego dos Algoritmos de Otimização.....	103
Capítulo 6 - Resultados e Discussões.....	105
6.1 Modelagem Fenomenológica.....	105
6.1.1 Ajuste do Modelo 1.....	107
6.1.2 Ajuste do Modelo 2.....	111
6.1.3 Ajuste do Modelo 3.....	117
6.1.4 Avaliação do Desempenho dos Métodos de Estimação e dos Modelos Fenomenológicos.....	122

6.2 Identificação do Sistema com Modelo Linear ARX	124
6.3 Modelagem Neural	130
6.4 Comparação entre os Modelos.....	137
Capítulo 7 - Conclusões	139
Capítulo 8 - Sugestões para Trabalhos Futuros.....	141
Referências	142
Glossário	147

Capítulo 1 - Introdução

1.1 MOTIVAÇÃO PARA O TRABALHO

As fibras da madeira são constituídas, basicamente, por quatro componentes: celulose, lignina, hemi-celulose e extrativos. Para a produção de polpa celulósica, matéria prima da fabricação de papel, é necessário remover a lignina e os extrativos da madeira através de reações químicas. O método mais aplicado para esta remoção química é conhecido como processo *kraft*, que faz o cozimento da madeira com hidróxido de sódio e sulfeto de sódio.

Após o cozimento, um processo de branqueamento é utilizado para a remoção do residual da lignina e dos compostos cromóforos, que conferem cor à polpa celulósica. A polpa “branca” é produzida através de um tratamento químico severo, constituído de estágios ácidos e básicos, que consomem produtos químicos de alto valor agregado como dióxido de cloro, peróxido de hidrogênio e ozônio; que encarecem o processo de produção de celulose.

Utilizado para reduzir uma quantidade significativa de lignina da polpa cozida (40 - 60%), o processo de deslignificação com oxigênio promove grande economia em produtos químicos, reduzindo as cargas de reagentes do branqueamento. Na deslignificação, a polpa, suspensa em solução alcalina, é pressurizada com oxigênio no reator, onde a lignina reage com as espécies de oxigênio formadas e é dissolvida e pode ser removida da polpa com a solução residual (VIOLETTE, 2003). Este processo se tornou essencial nas indústrias papeleiras modernas nas últimas décadas e continua a crescer, conforme ilustra a Figura 1.1.

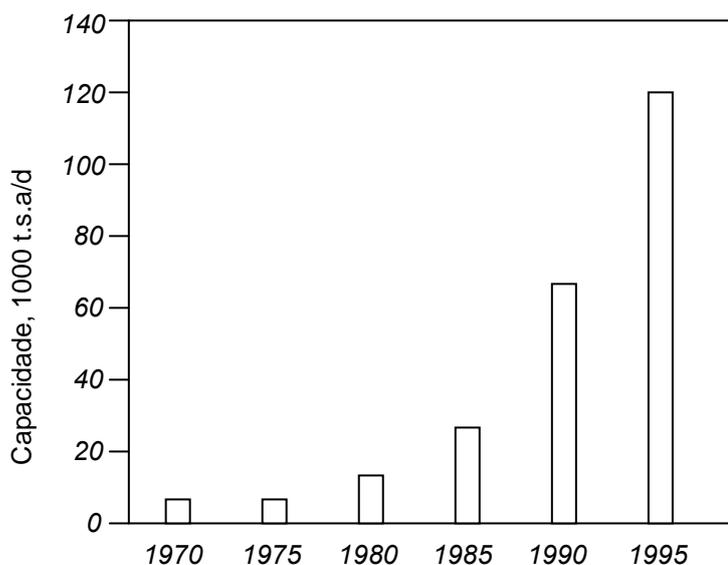


Figura 1.1 - Crescimento da deslignificação ao longo das últimas décadas

Fonte: DENCE; REEVE, 1996 *apud* VIOLETTE, 2003; p. 3.

Um fator determinante para este crescimento, além da economia em químicos, é a legislação ambiental uma vez que o processo de deslignificação diminui as cargas de poluentes nas emissões gasosas e nos efluentes líquidos.

Muitos pesquisadores propuseram a cinética básica da reação de deslignificação com oxigênio e melhoraram a seletividade manipulando as variáveis de processo ou alterando a composição da mistura de reagentes. Estes modelos podem ser uma ferramenta útil para uma fábrica de celulose, podendo ser utilizados para simular o processo, alterando as condições de operação sem modificar efetivamente a produção industrial. Novos limites podem ser estipulados para as principais variáveis do processo e, desta forma, a produção pode ser controlada de forma mais econômica sem perder qualidade no produto.

1.2 OBJETIVOS DO TRABALHO

O principal objetivo deste trabalho é desenvolver modelos matemáticos capazes de representar adequadamente a planta de deslignificação com oxigênio da Klabin Celulose e Papel – Unidade Monte Alegre, em Telêmaco Borba – PR. Propõe-se que estes modelos possam ser aplicados para simulações rotineiras na fábrica, podendo ser utilizados como uma ferramenta de trabalho.

Os modelos desenvolvidos deverão ser capazes de prever o número *kappa* da polpa na saída do reator de deslignificação. Esta variável poderá, desta forma, ser utilizada como um indicador de desempenho do processo, avaliando a seletividade da polpa em relação à remoção da lignina.

1.3 ESTRUTURA DA DISSERTAÇÃO

Para simplificar a compreensão do trabalho, esta dissertação foi escrita seguindo o processo de desenvolvimento dos modelos. Primeiro, no Capítulo 2, o processo de deslignificação com oxigênio da Klabin Celulose e Papel, Unidade Monte Alegre é apresentado e explicado com detalhes. Este entendimento é fundamental para a estruturação da modelagem matemática do processo.

O Capítulo 3 apresenta os métodos de modelagem utilizados. São discutidos os modelos fenomenológicos empíricos encontrados na literatura que descrevem os fenômenos de transporte de massa e da cinética da deslignificação com oxigênio. Apresentam-se também as técnicas de identificação de sistemas caixa-preta linear, modelo ARX; e não linear, modelagem neural. O Capítulo 3 tem por objetivo elucidar

os métodos empregados na modelagem matemática da planta, explicando seus pormenores e apresentando exemplos bem sucedidos de suas aplicações.

Para identificar os parâmetros do modelo fenomenológico, no Capítulo 4 foram discutidas técnicas heurísticas de otimização. Estes métodos foram essenciais para o ajuste dos modelos.

Os Capítulos 3 e 4 contêm toda a estrutura para a elaboração dos algoritmos de modelagem. O Capítulo 5 sintetiza a metodologia empregada no desenvolvimento deste trabalho.

Os resultados obtidos com três modelos fenomenológicos e com os modelos identificados com o ARX e a rede neural. Além disso, são discutidas as diferenças entre as técnicas de modelagem bem como a aplicação de cada uma delas.

Finalmente, no Capítulo 7 são apresentadas as conclusões deste trabalho. No Capítulo 8 são feitas sugestões para melhorias deste trabalho e para trabalhos futuros.

Esta dissertação contém, ainda, um glossário de termos técnicos da indústria de celulose para facilitar a compreensão da nomenclatura e esclarecer seus significados.

Capítulo 2 - O Processo de Deslignificação

O processo de deslignificação de polpa celulósica com oxigênio é apresentado neste capítulo. Primeiro é discutida a finalidade do processo, o motivo de sua introdução na produção de celulose e um breve histórico do seu desenvolvimento. A seguir o processo é explicado passo a passo, desde a entrada da polpa no reator de deslignificação até sua saída do processo.

2.1 INTRODUÇÃO

A história da deslignificação é longa, a primeira patente relacionada ao processo foi registrada em 1867. O uso do oxigênio molecular foi estudado para o branqueamento da polpa em 1950 por Nikitin e Akim ((a) HSU; HSIEH, 1985), mas o desenvolvimento do processo industrial foi lento e gradual até os anos 60, quando pesquisadores soviéticos em Leningrado impulsionaram o desenvolvimento das plantas de deslignificação. A primeira instalação comercial foi na fábrica SAPPI, sul-africana, em 1970. A Suécia instalou sua primeira planta dois anos depois (RAGNAR, 2002).

A principal motivação para a implantação das plantas de deslignificação foi a redução dos gastos com reagentes para o branqueamento e a redução das cargas poluentes no efluente líquido, uma vez que os anos 80 trouxeram consigo alterações nas leis ambientais. Nos Estados Unidos, entretanto, o processo não foi inicialmente aceito, sendo preferida a instalação de plantas com tratamento secundário de efluentes para a redução da carga poluente a ser lançada nos rios. A resistência

americana terminou em 1986, quando o primeiro sistema integrado de deslignificação à média consistência foi instalado (MEADOWS; BUGAJER, 1988).

As primeiras plantas operavam em alta consistência e não obtiveram grande sucesso devido ao elevado custo dos equipamentos, a seletividade era considerada baixa e havia problemas de segurança. Em algumas instalações explosões ocorreram por causa do grande volume de monóxido de carbono formado e acumulado no topo do reator. Essas limitações não permitiram a popularização da deslignificação em alta consistência, existiram apenas cerca de 50 unidades no mundo (RAGNAR, 2002).

No início dos anos 80 a deslignificação em média consistência foi introduzida no mercado, utilizando equipamentos mais baratos, com melhor seletividade e menos riscos de explosão. Foi nesse período que a deslignificação se popularizou. Sua principal atração era a diminuição das cargas poluentes e a redução do uso de compostos clorados e outros reagentes no processo de branqueamento (RAGNAR, 2002).

Ao ser lançado no mercado, o processo de deslignificação tinha como principais objetivos aumentar a produção e reduzir o custo das plantas de branqueamento existentes, reduzir a carga poluidora dos efluentes e os custos do branqueamento (MEADOWS; BUGAJER, 1988). Inicialmente a deslignificação não alterou a operação do digestor, uma vez que sua função se restringia a reduzir a carga de cloro nos estágios de branqueamento (RAGNAR, 2002).

O próximo passo seria abolir o uso de cloro no branqueamento. Para tanto, novos conceitos para o cozimento foram desenvolvidos. O cozimento foi prolongado até que o número *kappa* (valor de teste que relacionado linearmente com a quantidade de lignina remanescente na polpa cozida) de saída fosse bastante

pequeno e a deslignificação com oxigênio fosse utilizada até o limite, para só então a polpa ser alimentada aos estágios de branqueamento sem cloro. Esse tipo de operação foi utilizado por poucos anos e o cozimento voltou a operar como antes. A mudança foi na utilização da deslignificação, que foi estendida para que as cargas no branqueamento fossem menores (RAGNAR, 2002).

Atualmente a deslignificação é uma operação comum nas fábricas de celulose e os objetivos traçados na década de 80 foram alcançados com relativo sucesso. O estágio de deslignificação reduz em até 50% a demanda química de oxigênio (DQO) e entre 25 e 50% a demanda bioquímica de oxigênio (DBO) do efluente líquido e economiza até 25% em reagentes, como cloro, dióxido de cloro, peróxido de hidrogênio e ozônio utilizados no branqueamento da polpa.

Seu uso, entretanto, é limitado à remoção parcial da lignina em condições alcalinas, e o branqueamento deve ser procedido em operações posteriores para não danificar os carboidratos e as fibras (LIEBERGOTT; VAN LIEROP; SKOTHOS, 1992).

2.2 DESCRIÇÃO DO PROCESSO

No processo de deslignificação em média consistência, a solução de álcali é adicionada à polpa antes da bomba que alimenta o reator. Oxigênio e vapor de média pressão são adicionados num misturador posterior à bomba e a mistura é, então, alimentada ao reator. O fluxo é ascendente e o oxigênio, que em meio alcalino forma uma dispersão gasosa estável na polpa, é consumido nas reações com a lignina.

A polpa que deixa o reator segue para um tanque, o *blow tank*. A pressão é aliviada e ao mesmo tempo a temperatura é maximizada para o ponto de ebulição atmosférico da suspensão. Vapor *flasheado* e gases residuais são liberados para a atmosfera. A polpa é bombeada para a lavagem em um difusor pressurizado e em seguida é alimentada ao primeiro estágio de branqueamento (RAGNAR, 2002). A Figura 2.1 apresenta o esquema de planta de deslignificação utilizado pela Klabin Celulose e Papel.

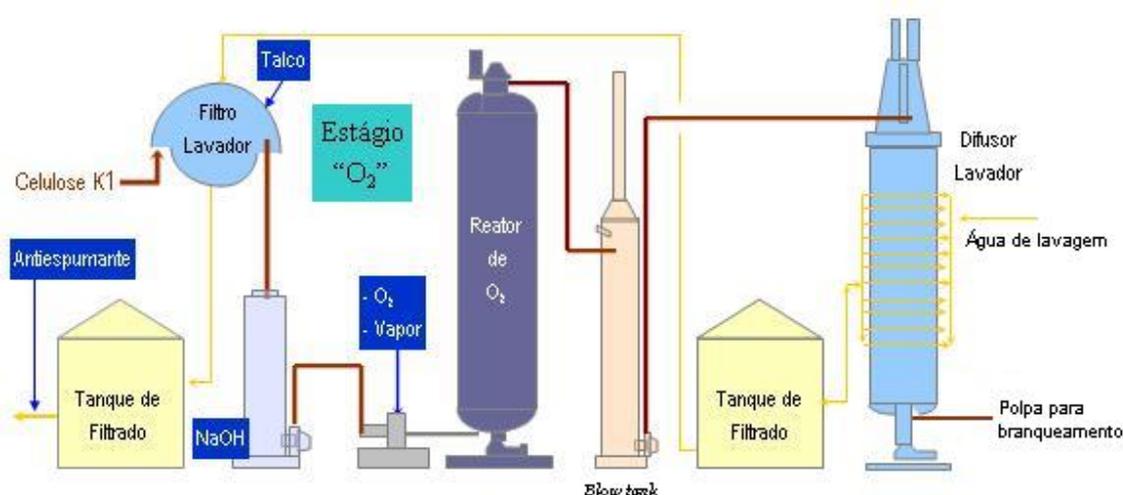


Figura 2.1 - Esquema da planta de deslignificação da Klabin Celulose e Papel –Unidade Monte Alegre.

O parâmetro utilizado para identificar o desempenho da planta industrial é o número *kappa* da polpa. Este valor é adimensional, linearmente proporcional à quantidade de lignina remanescente na polpa cozida, sendo, desta forma, um estimador da demanda de reagentes para os processos de deslignificação e branqueamento. Mais informações sobre o número *kappa*, como seu procedimento de medição, podem ser encontradas no glossário desta dissertação.

É normal obter-se na deslignificação com oxigênio a redução de 40 a 50% da quantidade de lignina da polpa. As reduções superiores a 50% podem levar a

degradação da celulose, causando decréscimo na viscosidade da polpa, ou seja, perda de resistência da fibra. Essa perda de viscosidade pode ser minimizada pelo uso de inibidores a base de magnésio (MEADOWS; BUGAJER, 1988). É importante, entretanto, lembrar que os inibidores possuem caráter ácido e que pH baixo no reator de deslignificação pode causar precipitação de lignina.

O resultado da deslignificação é governado pela carga de álcali e pela temperatura. A carga de oxigênio não altera significativamente o produto, ou seja, aumentar a carga de 20 para 23 *kg/o.d.t* não apresentará melhoras tão significativas quanto aumentar a carga de álcali de 20 para 23 *kg/o.d.t*, por exemplo (RAGNAR, 2002). Mais detalhes sobre as unidades e variáveis envolvidas no processo de deslignificação podem ser encontrados no glossário desta dissertação.

As informações sobre o processo são fundamentais para o desenvolvimento do modelo fenomenológico do processo, bem como as informações acerca da influência das variáveis no processo. Este Capítulo, portanto, procurou elucidar detalhes operacionais do processo, que serão depois aplicados na modelagem matemática fenomenológica e neural.

Capítulo 3 - Modelagem Matemática

O principal objetivo do Capítulo 3 é descrever as metodologias de modelagem empregadas neste trabalho. Primeiro é discutida a modelagem fenomenológica da reação de deslignificação, descrita através da equação cinética e do balanço de massa. São apresentados exemplos de modelos de deslignificação com oxigênio encontrados na literatura que serviram de ponto de partida para o desenvolvimento do modelo proposto nesta dissertação.

A seguir discutem-se métodos de identificação de modelos lineares e não lineares, sendo, esses últimos, os modelos neurais. A identificação de sistemas é a área da matemática que estuda alternativas para a modelagem fenomenológica e que possibilita o desenvolvimento de modelos caixa-preta (que podem ser lineares ou não-lineares). São apresentadas, também, algumas aplicações de modelos neurais na indústria de celulose e papel, inclusive um modelo neural de uma planta industrial de deslignificação de celulose.

3.1 INTRODUÇÃO

Entende-se por modelo a representação matemática de um sinal real que, idealmente, pode ser utilizada para representar ao menos suas principais características. As técnicas de modelagem podem ser agrupadas em duas grandes, e não necessariamente exclusivas, categorias (AMARAL, 1980 *apud* MELEIRO, 2002; SÖDESTRÖM; STOICA, 1989 *apud* MELEIRO, 2002; AGUIRRE, 2000). A

primeira é a modelagem pela física do processo, também conhecida como análise física-matemática ou ainda modelagem fenomenológica ou determinística. A segunda é conhecida como modelagem a partir de amostragem, identificação experimental ou ainda identificação de sistemas (LJUNG, 1999; MELEIRO, 2002).

A modelagem matemática caracteriza-se essencialmente por utilizar expressões matemáticas teóricas, empíricas e semi-empíricas para a análise dos fenômenos envolvidos em um processo. Dependendo do nível de conhecimento inicial do sistema, o problema de modelagem pode ser abordado de maneiras diferentes. A terminologia *modelagem caixa-branca* é utilizada quando o sistema é modelado totalmente a partir de leis físico-químicas (modelagem fenomenológica). Os modelos fenomenológicos são, em tese, potencialmente úteis para a adequação da investigação do comportamento de sistemas reais nas mais variadas condições de operação (AGUIRRE, 2000).

Se a identificação é baseada exclusivamente em dados amostrados, assumindo muito pouco ou quase nenhum conhecimento sobre o sistema, o processo de identificação é chamado de *modelagem caixa-preta*. Uma abordagem intermediária, a *modelagem caixa-cinza*, não exige um profundo conhecimento inicial sobre o sistema, mas permite que esta informação, caso disponível, seja incorporada ao modelo (AGUIRRE, 2000).

Os tipos de modelo mais comuns, segundo Aguirre (2000), são: estáticos ou dinâmicos; contínuos ou discretos; monovariáveis ou multivariáveis; determinísticos ou estocásticos. Modelos estáticos relacionam as variáveis sem quantificar sua dependência temporal, enquanto os modelos dinâmicos consideram a evolução temporal do sistema. Modelos estáticos são, normalmente, descritos por equações algébricas, enquanto os modelos dinâmicos são compostos por equações

diferenciais, sendo que tais modelos também podem incluir equações algébricas (AGUIRRE, 2000).

Os termos contínuo e discreto referem-se ao tempo, portanto a modelos dinâmicos. Modelos contínuos representam a evolução do sistema continuamente no tempo. Em contrapartida um modelo discreto no tempo representa a evolução do sistema em instantes discretos (AGUIRRE, 2000).

Um modelo com mais de uma entrada ou mais de uma saída é denominado multivariável. Já os modelos monovariáveis são aqueles que representam a relação causa e efeito de apenas um par de variáveis, ou seja, uma entrada e uma saída. Os modelos monovariáveis são conhecidos como *SISO (single input, single output)*. Os modelos multivariáveis podem ser classificados de acordo com o número de entradas e saídas: *MISO (multiple input, single output)*, com diversas entradas e apenas uma saída, *SIMO (single input, multiple output)*, com uma entrada e diversas saídas; e *MIMO (multiple input, multiple output)*, com diversas entradas e diversas saídas (AGUIRRE, 2000).

Modelos determinísticos são aqueles nos quais não estão representadas as incertezas presentes no problema real. Os modelos estocásticos incluem essas incertezas e conseqüentemente, a resposta de um modelo estocástico é uma variável que possui um componente aleatório (AGUIRRE, 2000).

3.2 MODELO FENOMENOLÓGICO GENERALIZADO DA DESLIGNIFICAÇÃO COM OXIGÊNIO

A deslignificação com oxigênio é uma reação heterogênea que ocorre num sistema composto de fibras de celulose e gás oxigênio dissolvido em meio aquoso. A cinética da reação, a equação fundamental do modelo do processo de deslignificação, é função da temperatura, pressão parcial do oxigênio, concentração de álcali e da qualidade da mistura no reator ((d) HSU; HSIEH, 1988). As demais equações do modelo estão relacionadas ao consumo de álcali e reações paralelas.

Esta seção, discute as expressões fundamentais para descrever o processo de deslignificação com oxigênio de polpa celulósica. O modelo descrito é dinâmico, contínuo, multivariável (*MISO*) e determinístico - nenhuma influência externa ao reator será considerada na modelagem.

Para qualquer sistema reativo, a taxa de reação permite acompanhar o comportamento da reação durante o processo e quantificá-lo. Para fins de cálculo, no entanto, é necessário conhecer a equação da taxa de reação. Para reações homogêneas, ou seja, aquela que ocorre toda na mesma fase, a equação da taxa é função da composição, da temperatura e da pressão. Fenômenos de difusão e convecção de calor e massa, de adsorção ou dessorção superficiais, que são característicos das reações heterogêneas (reações que ocorrem na presença de pelo menos duas fases), podem ser caracterizados por uma função mais ampla, denominada taxa de processo (SCHMAL, 1982).

O processo de deslignificação é controlado por uma reação heterogênea trifásica que se dá em etapas que envolvem a reação e o transporte dos reagentes entre as fases presentes no sistema. O primeiro estágio da reação é a transferência do oxigênio da fase gasosa para a fase líquida (AXEGARD et al, 1979), e esta dissolução depende da temperatura e pressão do sistema (MYERS; EDWARDS; 1989). O oxigênio dissolvido é transferido por difusão e convecção do seio da solução aquosa para o filme líquido formado ao redor da fibra. As moléculas de oxigênio e os íons hidroxila (OH^-) são transportados por difusão através do filme líquido até alcançarem a fibra celulósica (AXEGARD et al, 1979). A Figura 3.1 apresenta o processo microscópico de deslignificação.

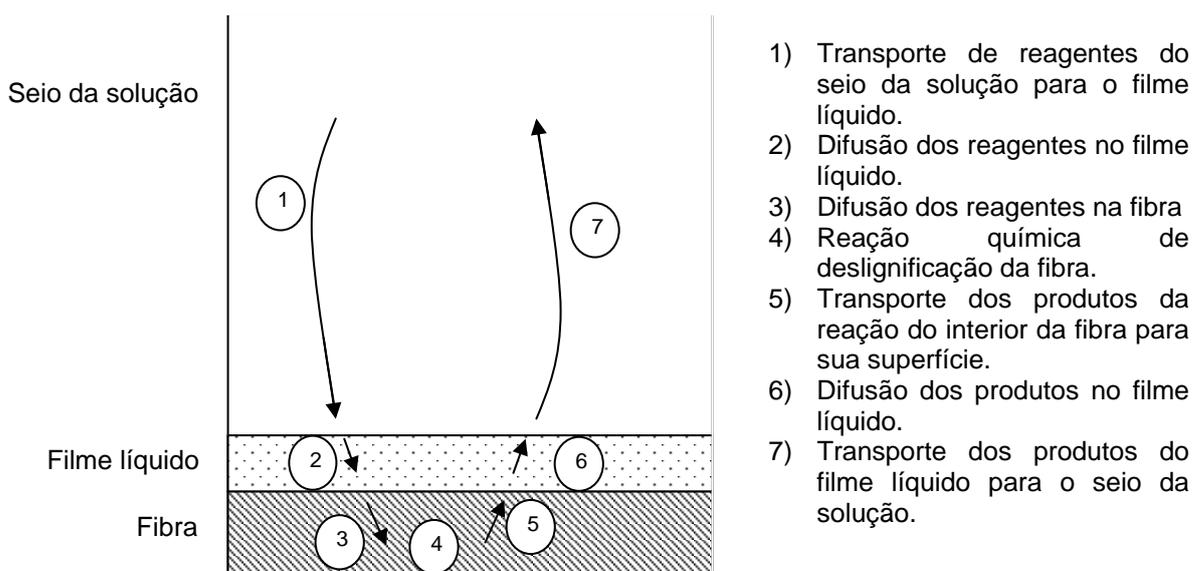


Figura 3.1 - Estágios da reação de deslignificação com oxigênio de celulose.

Fonte: AXEGARD et al, 1979. p.131

Na fibra acontece transferência de massa dos reagentes e a reação de deslignificação ocorre nos sítios ativos da fibra. Os produtos da reação, fragmentos de lignina e celulose de baixo peso molecular, são transportados para fora da fibra

(transferência de massa), através do filme líquido (difusão e convecção) e por fim alcançam o seio da solução (HSU; HSIEH, 1988).

As seguintes simplificações são consideradas na modelagem do processo (HSU; HSIEH, 1988):

- A proporção líquido-fibra é suficiente para considerar que as fibras estão completamente dispersas e suspensas na fase líquida;
- A solução de hidróxido de sódio ($NaOH$) é homogênea antes da injeção do gás oxigênio;
- No seio da solução a concentração de oxigênio é uniforme;
- A principal resistência à transferência de moléculas de oxigênio da fase gasosa para a fase líquida é a apresentada pelo filme formado na interface das duas fases.

Uma vez que a concentração das moléculas de lignina dissolvidas é muito pequena, a transferência das moléculas de oxigênio e dos íons hidroxila na fase líquida não é afetada. A equação da taxa de reação de deslignificação pode ser formulada, de acordo com Hsu e Hsieh (1988):

$$-\frac{dK}{dt} = f(\text{transferencia.de.massa}, k(T), [OH^-], P_{O_2}, K, T) \quad (3.1)$$

onde K é o número *kappa* do sistema, $k(T)$ é constante da equação da taxa, $[OH^-]$ é a concentração de álcali, P_{O_2} é a pressão do oxigênio no sistema e T a temperatura de reação.

Experimentos realizados por Hsu e Hsieh, em 1988, concluem que a transferência de massa entre as fases gasosa e líquida e a transferência de massa entre as fases líquida e sólida (que é cerca de 150 vezes mais rápida do que a

transferência de massa entre as fases gasosa e líquida) são negligenciáveis quando comparadas com a taxa intrínseca de reação.

Os fenômenos que caracterizam a reação heterogênea podem ser, portanto, simplificados e a equação da taxa pode ser escrita como um modelo cinético geral de reações homogêneas (AGARWAL et al, 1999):

$$-\frac{dK}{dt} = k \cdot K^a [OH^-]^b P_{O_2}^c \quad (3.2)$$

onde a , b e c são parâmetros determinados para o processo que se está modelando. A constante da reação, $k(T)$ é determinada através da equação de Arrhenius (OLM; TEDER, 1979):

$$k = k_0 \exp\left(-\frac{E_a}{RT}\right) \quad (3.3)$$

onde k_0 é o fator de frequência, E_a é a energia de ativação de Arrhenius e R é a constante universal dos gases $\left(8,314 \frac{kJ}{gmol \cdot K}\right)$.

O efeito das variáveis de processo na taxa de deslignificação foi investigado pela primeira vez em 1970, na Suécia, por Hartler et al (*apud* HSU; HSIEH, 1988). Este modelo não discute a qualidade da mistura e considerou que os efeitos da pressão parcial do oxigênio e da consistência não afetam a redução do número $kappa$. Edwards e Norberg, em 1973 (*apud* HSU; HSIEH, 1988) apresentaram um modelo cinético geral para deslignificação a pressão constante. Eles consideraram que a taxa de reação era controlada pela reação química, mas não houve verificação experimental.

Em 1978, Jarrehult e Samuelson (*apud* HSU; HSIEH, 1988) utilizaram polpa de consistência muito baixa (0,2 a 1%) para eliminar o efeito da redução da concentração de álcali na taxa de deslignificação. Evans et al em 1979 (*apud* HSU;

HSIEH, 1988), também estudaram a cinética para reação de deslignificação com baixa consistência da polpa. De acordo com suas conclusões, a difusão pode afetar a reação e esses resultados contrariam os estudos de Hatler et al (1970 *apud* (d) HSU; HSIEH, 1988) e de Jarrehult e Samuelson (1978 *apud* (d) HSU; HSIEH, 1988). Também em 1979(a), Olm e Teder publicaram um modelo de cinética para a deslignificação com duas reações de pseudoprimeira ordem. O modelo de Olm e Teder ((a)1979), detalhado em 3.1.2, é o primeiro a apresentar resultados relevantes e consistentes.

Hsu e Hsieh ((d)1988), seção 3.1.3, apresentaram um modelo que descarta a importância da transferência de massa durante a deslignificação com oxigênio. O modelo de Hsu e Hsieh foi apresentado em dois estágios. Em 1989, Myers e Edwards (detalhado na seção 3.1.4) publicaram um modelo similar ao de Olm e Teder (1979(a) e 1981(b)). O modelo considera que a lignina presente no sistema pode ser classificada de acordo com a velocidade da reação: lignina que reage rapidamente, lentamente e uma parcela que não reage. A semelhança com o modelo de Olm e Teder está na descrição das duas reações de pseudoprimeira ordem. Este modelo também trouxe mudanças para a participação do oxigênio. A equação da taxa inclui a concentração de oxigênio dissolvida.

Agarwal et al (1999) (detalhado na seção 3.1.5) apresentaram um modelo simples e resgataram um modelo de 1982 de Schöön, que considera uma série de reações de primeira ordem paralelas durante o processo de deslignificação, buscando explicar a alta ordem da reação associada ao número *kappa* de entrada do sistema. Mais recentemente, em 2003, Violette (seção 3.1.6) publicou sua tese de doutorado apresentando uma cinética para a reação de deslignificação de oxigênio que não considera o efeito do oxigênio no sistema.

Todos os modelos desconsideram os fenômenos de transporte de massa entre as fases existentes no sistema e todos os autores concordam com a influência da concentração de álcali, da temperatura e do oxigênio na cinética de deslignificação. As diferenças residem no grau de influência dessas variáveis no processo de deslignificação e na forma de apresentação da cinética. Alguns consideram reações em etapas devido à velocidade de reação da lignina, outros utilizam a concentração dissolvida do oxigênio ao invés de sua pressão parcial. Este trabalho discute alguns desses modelos com o intuito de identificar os mais adequados ao sistema a ser modelado.

3.2.1 O Modelo de Olm e Teder

O modelo de Olm e Teder ((a)1979; (b)1981) considera que duas reações competem durante o processo de deslignificação, a reação de deslignificação propriamente dita e a reação de degradação dos carboidratos. O objetivo dos autores era derivar um modelo que descrevesse as duas reações. A razão entre a taxa de deslignificação e a taxa de degradação dos carboidratos determina a seletividade da deslignificação com oxigênio ((a) OLM; TEDER, 1979).

Para as investigações acerca da cinética de deslignificação, foram procedidos experimentos com polpa *kraft* de *softwood* separadamente para cada variável, sendo as demais mantidas constantes. Foram verificados os efeitos do tempo de reação, concentração de álcali, temperatura e consistência sobre a cinética da reação (OLM; TEDER, 1979).

Olm e Teder ((a) 1979; (b) 1981) concluíram que o número *kappa* da polpa diminui com o aumento do tempo de reação se as demais condições (concentração

de álcali, temperatura e pressão parcial de oxigênio) forem mantidas constantes. Isto significa que, se a reação for de primeira ordem em relação à quantidade de lignina na polpa, uma linha reta será obtida, ou seja, o número *kappa* continuará a diminuir com o tempo ((a) OLM; TEDER, 1979).

Olm e Teder ((a)1979) propuseram um modelo com duas reações de pseudoprimeira ordem para o número *kappa*: uma fase inicial de reação rápida seguida por uma fase lenta.

$$-\frac{dK}{dt} = k_1 \cdot K_{01} [OH^-]^{0,1} P_{O_2}^{0,1} + k_2 \cdot K_{02} [OH^-]^{0,3} P_{O_2}^{0,2} \quad (3.4)$$

Com:

$$\begin{aligned} k_1 &= k_{0_1} \exp\left(-\frac{10}{RT}\right) \\ k_2 &= k_{0_2} \exp\left(-\frac{45}{RT}\right) \end{aligned} \quad (3.5)$$

Sendo a concentração de álcali expressa em $\frac{g}{L}$, a pressão parcial do oxigênio em MPa, o tempo em min e a temperatura em Kelvin, K_{01} é igual à quantidade de lignina facilmente removível (fase inicial) e K_{02} à quantidade de lignina removida lentamente (fase final). Os valores das constantes k_{0_1} e k_{0_2} não estão explícitos no trabalho publicado. O total de lignina na polpa, quantificada pelo número *kappa*, é a soma das duas quantidades:

$$K = K_{01} + K_{02} \quad (3.6)$$

A degradação dos carboidratos é expressa em termos de crescimento do número de cadeias celulósicas na polpa e quantificada pela viscosidade intrínseca da polpa. A equação da taxa é, entretanto, expressa em termos de número de cadeias celulósicas (NCC) por tonelada de polpa, uma medida coerente para uma

equação cinética e que permite comparar a degradação de polpas com diferentes valores iniciais de viscosidade.

A viscosidade intrínseca da polpa (η) diminui com o tempo de reação independentemente de seu valor inicial e, assim como na reação de deslignificação, uma fase inicial rápida seguida de uma lenta foi observada nos experimentos. Duas reações, ambas de ordem zero, foram utilizadas para escrever o modelo da degradação de carboidratos.

$$\frac{dNCC}{dt} = k_3 \cdot [OH^-]^{0,2} P_{O_2}^{0,8} + k_4 \cdot [OH^-]^{0,6} P_{O_2}^{0,1} \quad (3.7)$$

onde NCC é o número de cadeias celulósicas por tonelada de polpa. Os valores das constantes k_3 e k_4 , que não estão explícitos no trabalho publicado, são estimados através da equação de Arrhenius:

$$\begin{aligned} k_3 &= k_{0_3} \exp\left(-\frac{40}{RT}\right) \\ k_4 &= k_{0_4} \exp\left(-\frac{53}{RT}\right) \end{aligned} \quad (3.8)$$

Para converter o valor da viscosidade intrínseca para número de cadeias celulósicas por tonelada de polpa e vice versa, a seguinte expressão foi utilizada:

$$\log(NCC) = A - B \cdot \log(\eta) \quad (3.9)$$

Sendo η a viscosidade intrínseca da polpa em $\frac{dm^3}{kg}$.

Para as reações de deslignificação e de degradação dos carboidratos, Olm e Teder ((a)1979) concluíram que ambas as taxas de reação aumentam com o aumento da concentração de álcali, da pressão parcial do oxigênio e da temperatura.

Em 1981(b) Olm e Teder publicaram uma nova versão para o modelo de deslignificação com oxigênio. Os autores reafirmam neste segundo trabalho a

separação da reação em duas reações de pseudoprimeira ordem, mas propõem um modelo em um só estágio, de ordem superior em relação ao número $kappa$ inicial:

$$-\frac{dK}{dt} = k \cdot K^{3,2} [OH^-]^{0,6} P_{O_2}^{0,5} \quad (3.10)$$

$$k = k_0 \exp\left(\frac{-70}{RT}\right)$$

A cinética da reação de degradação dos carboidratos não foi alterada neste segundo trabalho.

3.2.2 O Modelo de Hsu e Hsieh

Hsu e Hsieh ((d)1988) tiveram como principal objetivo em seu trabalho esclarecer a influência da transferência de massa no processo de deslignificação. Para simplificar o sistema e, conseqüentemente estimar os coeficientes de transferência de massa, os experimentos foram conduzidos em reator com consistência ultrabaixa (em torno de 0,4%). A consistência é a porcentagem, em peso, de material fibroso seco; ou a densidade de sólidos na polpa, sendo uma variável importante na indústria papelreira por quantificar a razão de quantidade de reagentes em relação à quantidade de fibras.

A principal contribuição deste trabalho é justamente a elucidação dos efeitos produzidos pela transferência de massa. Hsu e Hsieh ((d) 1988) calcularam o coeficiente de transferência de massa volumétrico para as condições operacionais utilizando a teoria de renovação de superfície de Sherwood. Tanto para a transferência gás-líquido quanto para a transferência líquido-fibra, os valores obtidos foram tais que a transferência de massa, mesmo se aumentando em 7 vezes a taxa

de transferência, o número kappa não sofre redução significativa ((d) HSU; HSIEH, 1988).

O efeito da difusão no interior da fibra foi estimado pelo critério de Weisz-Prater e resultou num valor da ordem de 10^{-3} . O efeito da difusão pode ser considerado insignificante quando comparado à taxa intrínseca de reação se este critério for válido ((d) HSU; HSIEH, 1988).

Todos os experimentos foram conduzidos à consistência ultrabaixa e as condições da reação (temperatura, concentração de álcali e pressão de oxigênio) foram variadas em três níveis de investigação. O aumento de qualquer uma das três variáveis conduz a uma maior redução do número *kappa*. O modelo proposto foi da forma:

$$-\frac{dK}{dt} = k \cdot K^{6,27} [OH^-]^{1,0} P_{O_2}^{0,89} \quad (3.11)$$

com:

$$k = 0,24 \exp\left(-\frac{8,3 \cdot 10^7}{RT}\right) \quad (3.12)$$

O alto valor da ordem da equação da taxa em relação ao *kappa* resulta de um rápido decréscimo da taxa de deslignificação durante o experimento. Hsu e Hsieh ((d) 1988) sugerem um modelo em dois estágios para melhorar o ajuste:

$$-\frac{dK}{dt} = k_1 \cdot K^{3,07} [OH^-]^{0,78} P_{O_2}^{0,35} [u(t) - u(t-2)] + k_2 \cdot K^{3,07} [OH^-]^{0,70} P_{O_2}^{0,74} u(t-2) \quad (3.13)$$

com:

$$\begin{aligned} k_1 &= 2,46 \exp\left(-\frac{3,6 \cdot 10^7}{RT}\right) \\ k_2 &= 143,49 \exp\left(-\frac{7,1 \cdot 10^7}{RT}\right) \end{aligned} \quad (3.14)$$

A função $u(t)$ é utilizada para indicar que nos primeiros 2 minutos de reação a equação da cinética é a controlada pelo primeiro termo da equação 3.13 e o restante do tempo é descrito pelo segundo termo. A concentração do hidróxido de sódio deve ser expressa em $\frac{kmol}{m^3 \text{ de licor branco}}$, a pressão parcial do oxigênio em Pascal, a temperatura em Kelvin e o tempo em minutos.

3.2.3 O Modelo de Myers e Edwards

Myers e Edwards (1989) utilizaram o conceito de Ackert et al (1975 *apud* MYERS; EDWARDS, 1989) que separa a lignina em três categorias de reação: rápida, lenta e não reativa e o aplicaram de forma semelhante a Olm e Teder ((a)1979).

O modelo de Myers e Edwards (1989) apresenta como uma das melhorias do modelo cinético de lignina de reação rápido-lenta o uso da concentração de oxigênio dissolvido na solução ao invés de sua pressão parcial, uma vez que a solubilidade do oxigênio depende da temperatura e da pressão. Mas a opção pela concentração não garante por si só melhores resultados. Os autores ainda incluíram uma taxa de consumo não-linear de álcali:

$$\frac{\% NaOH}{t_{polpa}} = \frac{K_{inicial} \left(\frac{K_{inicial}}{K_{saída} - 1} \right)}{19,6} \quad (3.15)$$

onde t_{polpa} é a quantidade de polpa em toneladas. O consumo linear de oxigênio:

$$O_2 = \alpha \cdot lignina reagida \quad (3.16)$$

onde O_2 é a quantidade de oxigênio consumida e α é um parâmetro a ser estimado para o sistema. A quantidade de lignina reagida é quantificada pelo número $kappa$ da polpa.

Os autores expandiram o número e a variedade de dados utilizados para criar um modelo geral, que satisfizesse tanto sistemas de *softwood* ou *hardwood*, consistências de até 30% e número $kappa$ inicial variando entre 11 e 128:

$$\begin{aligned} -\frac{dK_f}{dt} &= k_f \cdot K_f [O_{2_{licor}}]^{0,43} \\ -\frac{dK_s}{dt} &= k_s \cdot K_s [OH^-]^{0,875} [O_{2_{licor}}]^{0,43} \end{aligned} \quad (3.17)$$

Com:

$$\begin{aligned} k_f &= 1,51 \cdot 10^5 \exp\left(-\frac{31,6}{RT}\right) \\ k_s &= 1,68 \cdot 10^7 \exp\left(-\frac{61,4}{RT}\right) \end{aligned} \quad (3.18)$$

A concentração de álcali e do oxigênio dissolvido deve ser expressa em $\frac{kg}{m^3 \text{ de licor branco}}$, a temperatura em Kelvin e o tempo, t , em minutos. O valor do número $kappa$ final é o somatório dos valores dos números $kappa$ obtidos para as ligninas que reagem rápida e lentamente e da lignina que não reage.

$$K = K_f + K_s + K_b \quad (3.19)$$

onde K_f , K_s e K_b são as categorias de lignina que reagem rapidamente, lentamente e que não reage, respectivamente. Myers e Edwards (1989) consideram $K_b = 0$ e as condições iniciais para K_f e K_s são:

$$\begin{aligned} K_f &= \omega_f K_{inicial} \\ K_s &= \omega_s K_{inicial} \end{aligned} \quad (3.20)$$

sendo $K_{inicial}$ o número $kappa$ inicial da polpa.

3.2.4 O Modelo de Agarwal, Genco, Cole e Miller

Agarwal et al (1999) discutem em seu trabalho a alta ordem aparente da taxa de reação em relação ao número $kappa$, propondo que o comportamento cinético observado pode ser explicado por uma série de reações paralelas de primeira ordem que ocorrem simultaneamente durante o processo de deslignificação com oxigênio. Como Myers e Edwards (1989), este modelo considera a possibilidade da existência de uma categoria de lignina de difícil remoção, justificando assim, a fase final lenta na reação.

O modelo foi escrito para sistemas de *hardwood*, em reator de batelada com consistência de 12% e número $kappa$ inicial aproximado de 13,2. A influência da temperatura, carga de álcali e pressão do oxigênio foram avaliadas individualmente, sendo que cada uma das condições de reação foi variada três vezes. A equação da taxa de reação de Agarwal et al (1999) é:

$$-\frac{dK}{dt} = k \cdot K^{7,7} [OH^-]^{0,92} P_{O_2}^{0,53} \quad (3.21)$$

com:

$$k = 2,368 \cdot 10^6 \exp\left(-\frac{107,2}{RT}\right) \quad (3.22)$$

com a concentração de álcali expressa em $\frac{g}{L}$, a pressão parcial do oxigênio em kPa , a temperatura em Kelvin e o tempo em h .

A alta ordem (7,7) da taxa de reação em relação ao número $kappa$ inicial não é incomum para o processo de deslignificação. Axegard et al (1979) sugerem que esta fase lenta observada é algum tipo de atraso no processo de deslignificação. Segundo os autores, poderiam ser reações de condensação de lignina ou

resistência às transferências de massa. Esse atraso é matematicamente traduzido na expressão cinética como uma alta ordem de reação em relação ao número *kappa* inicial. Agarwal et al (1999) sugerem que a alta ordem possa ser decorrente de uma série de reações de primeira ordem paralelas.

Agarwal et al (1999) ainda resgatam o modelo de Schöön de 1982, que também traçou observações acerca da possibilidade das reações de primeira ordem paralelas e simultâneas. Em seu trabalho, Schöön (1982) faz uma analogia com os processos petroquímicos de hidrodessulfurização e hidrogenitrogenação nos quais a alta ordem das reações é explicada com as reações de primeira ordem paralelas.

Os processos de mais de três reações paralelas são difíceis de serem modelados e para contornar o problema Schöön (1982) assume que a deslignificação é a soma de infinitas reações paralelas de primeira ordem e que suas constantes de reação podem ser consideradas funções de distribuição. Os resultados obtidos com a aplicação do modelo de Schöön (1982) sugerem que as resistências à transferência de massa não são responsáveis pelo alto valor da ordem da reação, mas sim as porções de lignina reagindo diferentemente em reações de primeira ordem.

3.2.5 O Modelo de Violette

Violette (2003) testou o efeito de aditivos para melhorar a seletividade da reação de deslignificação. Para tanto desenvolveu um modelo cinético do processo, testando diversas formas de equacionamento. Como Olm e Teder (1979), Violette (2003) inclui em seu modelo a reação de degradação de carboidratos que causam a perda de viscosidade.

Para avaliar o consumo de álcali durante a reação, o autor testou sete modelos, variando as condições que poderiam influenciar o consumo. O modelo escolhido é linear e, como no modelo proposto por Myers e Edwards (1989), dependente somente do número *kappa*.

$$\Delta[OH^-] = 0,169 \cdot \Delta K + 0,20 \quad (3.23)$$

onde $\Delta[OH^-]$ é a variação da concentração de álcali e ΔK é a variação no número *kappa*.

A equação 3.23 considera que a variação de soda é diretamente proporcional à variação do número *kappa* da polpa e não é influenciada pelas demais condições de operação do sistema.

Para a variação do número *kappa*, o autor novamente verificou diversas formulações, desde equações lineares até os modelos já consagrados pela literatura. A equação escolhida não considera a influência da pressão (ou concentração) do oxigênio, embora seja da forma já apresentada pelos modelos anteriores:

$$-\frac{dK}{dt} = k \cdot K^{3,12} [OH^-]^{0,588} \quad (3.24)$$

A concentração de hidróxido de sódio deve ser expressa em $\frac{g}{L}$, a temperatura em Kelvin e o tempo em minutos. A expressão para o cálculo da constante da equação da taxa é:

$$k = 4,4 \cdot 10^3 \exp\left(-\frac{7140}{RT}\right) \quad (3.25)$$

Para avaliar a degradação dos carboidratos, Violette (2003) derivou uma expressão em termos de número de cadeias quebradas (*NCQ*):

$$-\frac{dNCQ}{dt} = k \cdot K^{3,38} [OH^-]^{1,22} \quad (3.26)$$

com:

$$k = 6,0 \cdot 10^4 \exp\left(-\frac{9980}{RT}\right) \quad (3.27)$$

O NCQ é representado pelo grau de polimerização (DP) que, por sua vez, é estimado pelo valor da viscosidade. O grau de polimerização das cadeias celulósicas na polpa diminui durante o processo devido à hidrólise alcalina. É, entretanto, mais prático calcular a viscosidade, uma medida direta da polpa, para estimar este valor. As relações entre essas variáveis estão apresentados nas equações 3.28 e 3.29 (HEININGEN et al, 2002).

$$NCQ = \left(\frac{1}{DP_t} - \frac{1}{DP_o} \right) DP_o \quad (3.28)$$

sendo DP_t o grau de polimerização calculado para o tempo t e DP_o o grau de polimerização da polpa antes do processo de deslignificação. Para relacionar a viscosidade intrínseca com o grau de polimerização, Violette utilizou a expressão de:

$$\eta = 0,6061 \cdot DP^{0,90} \quad (3.29)$$

onde η é a viscosidade intrínseca da polpa.

Violette (2003) não investiga a influência do oxigênio, mas afirma em suas conclusões que deve ser considerado para uma boa modelagem do processo de deslignificação com oxigênio.

3.3 IDENTIFICAÇÃO DE SISTEMAS

A identificação de sistemas é a área que estuda alternativas de modelagem matemática. Uma das características dessas técnicas é que pouco conhecimento prévio do sistema é necessário para a modelagem do sistema. Esses métodos também são referidos como *modelagem* (ou *identificação*) *caixa preta* ou *modelagem empírica* (AGUIRRE, 2000).

O grande apelo da modelagem caixa-preta é devido, principalmente, à dificuldade na obtenção de um tratamento matemático genérico para sistemas não lineares, uma vez que é praticamente impossível representar adequadamente todo tipo de conhecimento em sistemas não lineares. Um outro problema é que, mesmo que algum conhecimento sobre o sistema esteja disponível, relacioná-lo corretamente a uma descrição dinâmica do sistema contínuo no tempo (i.e., em termos de equações diferenciais) pode ser uma tarefa complexa. Portanto, a opção de escolher a modelagem caixa-preta não é incomum, mesmo quando um certo nível de conhecimento sobre o sistema está disponível (MELEIRO, 2002).

Contudo, um bom entendimento sobre as leis da física que regem o comportamento do sistema é sempre útil e facilita o processo de identificação do sistema. Tal conhecimento pode estar relacionado à ordem do sistema, à sua dinâmica, a frequência de amostragem adequada, às características de estabilidade, à faixa de operação em que se pretende operar o sistema, tempos de atraso e grau de não linearidade (MELEIRO, 2002).

Por outro lado, a despeito da maioria dos sistemas apresentarem ao menos uma leve não linearidade, a maior parte da literatura sobre identificação de sistemas

lida com identificação de sistemas lineares. Algumas das principais razões para isso são (NØRGAARD et al, 2000 *apud*, MELEIRO, 2002):

- Muitos sistemas podem ser bem descritos por um modelo linear e, quando não é razoável usar um único modelo para toda a faixa de operação, muitas vezes é possível identificar diferentes modelos lineares para os diferentes regimes da faixa de operação;
- Do ponto de vista computacional, a implementação de modelos lineares é menos complicada;
- É mais simples projetar controladores para um sistema descrito por modelos lineares.

Uma prática prudente em identificação de sistemas é tentar, em uma primeira etapa, identificar o processo de interesse com modelos lineares. Além de se tratar de modelos mais simples, o resultado da identificação pode ajudar na avaliação o desempenho de modelos não lineares mais elaborados, como as redes neurais. Ao comparar o desempenho dos modelos linear e não linear, pode-se chegar à conclusão de que ambos representam a dinâmica do processo igualmente bem. Neste caso, o modelo linear deve ser escolhido, dada sua simplicidade. (LJUNG, 1999; MELEIRO, 2004).

3.3.1 Identificação de sistemas com modelos lineares

Sistemas lineares podem ser definidos formalmente como todo sistema que satisfaz o princípio da superposição (AGUIRRE, 2000). Resumidamente, seja um sistema que ao ser excitado pela entrada $u_1(t)$ produz a saída $y_1(t)$ e quando

excitado pela entrada $u_2(t)$ produz a saída $y_2(t)$. Se tal sistema satisfizer o princípio da superposição então, quando for excitado por $a \cdot u_1(t) + b \cdot u_2(t)$, sua saída será $a \cdot y_1(t) + b \cdot y_2(t)$, sendo a e b constantes reais (MELEIRO, 2002).

Segundo Nørgaard et al (2000 *apud* MELEIRO, 2002), um sistema denominado linear, se passível de ser representado através de um modelo, assume a seguinte forma:

$$y(t) = G(q^{-1})u(t) + H(q^{-1})v(t) \quad (3.30)$$

onde G e H são funções de transferência definidas no operador de atraso q^{-1} ; e $v(t)$ é um sinal de ruído branco independente das entradas passadas e que pode ser caracterizado por alguma função de densidade probabilística. O operador de atraso atua sobre o sinal conforme descrito na equação 3.31 (MELEIRO, 2002).

$$q^{-d}x(t) = x(t-d) \quad (3.31)$$

onde d é um múltiplo do período de amostragem. No caso de sistemas multi-variáveis MISO, $u(t)$ e $v(t)$ são vetores, enquanto G e H são matrizes polinomiais em q^{-1} .

A estrutura do modelo, \mathcal{M} , é um conjunto parametrizado de modelos candidatos (MELEIRO, 2002):

$$\mathcal{M}(\theta) : \{G(q^{-1}, \theta), H(q^{-1}, \theta) \mid \theta \in D_m\} \quad (3.32)$$

$$y(t) = G(q^{-1}, \theta) \cdot u(t) + H(q^{-1}, \theta) \cdot v(t) \quad (3.33)$$

onde θ denota os p parâmetros a serem estimados e D_m é algum subconjunto de \mathbb{R}^p dentro do qual a busca por um modelo deve ser conduzida.

Uma representação alternativa do modelo descrito pela equação 3.33, denominada Preditor do Modelo (um passo adiante), é dada por (MELEIRO, 2002):

$$\hat{y}(t|t-1, \theta) = H^{-1}(q^{-1}, \theta)G(q^{-1}, \theta)u(t) + [1 - H^{-1}(q^{-1}, \theta)]y(t) \quad (3.34)$$

A inclusão de θ como um novo argumento na equação 3.34 implica que a estrutura do modelo representa um conjunto de modelos (MELEIRO, 2002).

3.3.1.1 Representação matemática no tempo discreto

Há diversas maneiras de representar o mesmo modelo matemático. Existem algumas representações matemáticas que são especialmente adequadas para a identificação de sistemas utilizando-se algoritmos já estabelecidos para estimação de parâmetros (MELEIRO, 2002). A seguir, será apresentada uma representação matemática freqüentemente utilizada na identificação de sistemas lineares, o Modelo Auto-Regressivo com Entradas Externas (ARX – *Auto-Regressive with eXogenous inputs*), que foi empregada para o estudo de caso da deslignificação de celulose com oxigênio.

Considere o seguinte modelo geral (MELEIRO, 2002):

$$A(q^{-1})y(k) = q^{-d} \frac{B(q^{-1})}{F(q^{-1})}u(k) + \frac{C(q^{-1})}{D(q^{-1})}v(k) \quad (3.35)$$

onde q^{-1} é o operador de atraso definido de forma que $y(k)q^{-1} = y(k-1)$; $v(k)$ é o ruído branco e $A(q^{-1})$, $B(q^{-1})$, $C(q^{-1})$, $D(q^{-1})$ e $F(q^{-1})$ são polinômios definidos da seguinte forma (MELEIRO, 2002):

$$\begin{aligned}
A(q^{-1}) &= 1 + a_1 q^{-1} + \dots + a_{n_y} q^{-n_y} \\
B(q^{-1}) &= b_0 + b_1 q^{-1} + \dots + b_{n_u} q^{-n_u} \\
C(q^{-1}) &= 1 + c_1 q^{-1} + \dots + c_{n_\xi} q^{-n_\xi} \\
D(q^{-1}) &= 1 + d_1 q^{-1} + \dots + d_{n_d} q^{-n_d} \\
F(q^{-1}) &= 1 + f_1 q^{-1} + \dots + f_{n_f} q^{-n_f}
\end{aligned} \tag{3.36}$$

onde a , b , c , d e f são os parâmetros ajustáveis do modelo.

A estrutura do modelo ARX corresponde a escrever a equação 3.6 tomando $C(q^{-1})=D(q^{-1})=1$ e $A(q^{-1})$ e $B(q^{-1})$ como polinômios arbitrários, resultando em (MELEIRO, 2002):

$$A(q^{-1})y(k) = B(q^{-1})u(k) + v(k) \tag{3.37}$$

Como o ruído $v(k)$ aparece explicitamente na equação 3.37, o modelo ARX é classificado como pertencendo à classe de modelos com erro na equação (AGUIRRE, 2002). O modelo ARX pode ser reescrito na forma:

$$y(k) = \frac{B(q^{-1})}{A(q^{-1})}u(k) + \frac{1}{A(q^{-1})}v(k) \tag{3.38}$$

Neste modelo, o ruído adicionado à saída, $e(k) = \frac{v}{A(q^{-1})}$, não é branco, mas é modelado como um ruído branco filtrado por um filtro auto-regressivo (MELEIRO, 2002).

Uma forma alternativa de representar a estrutura do modelo ARX, baseada na equação 3.33, corresponde a escolher (MELEIRO, 2002):

$$\begin{aligned}
G(q^{-1}, \theta) &= q^{-d} \frac{B(q^{-1})}{A(q^{-1})} \\
H(q^{-1}, \theta) &= \frac{1}{A(q^{-1})}
\end{aligned} \tag{3.39}$$

Cujo preditor toma a seguinte forma (MELEIRO, 2002):

$$\hat{y}(k|k-1, \theta) = q^{-d} B^{-1}(q^{-1})u(k) + [1 - A^{-1}(q^{-1})]y(k) \quad (3.40)$$

3.3.2 Identificação de sistemas com modelos não-lineares

Uma das técnicas de modelagem não linear são as redes neurais artificiais (RNAs). As RNAs têm atraído muita atenção, principalmente quando comparadas a outras técnicas não lineares de identificação de processos (NØRGAARD et al, 2000 *apud*, MELEIRO, 2002). Parte desse sucesso se deve ao fato de que é muito difícil saber, a princípio, qual técnica de identificação é a mais adequada para uma dada aplicação sem uma avaliação prévia das técnicas disponíveis. Por isso é altamente desejável considerar uma única técnica que possa ser empregada para todas as aplicações, em vez de avaliar várias técnicas candidatas para uma nova aplicação. A propriedade de ser versátil e aplicável a um grande número de situações, é uma das características mais atraentes das RNAs (MELEIRO, 2002).

Como o próprio nome indica, as redes neurais são estruturas matemáticas baseadas no mecanismo de funcionamento de tecidos nervosos cerebrais, numa tentativa de imitar a inteligência. Formada por unidades básicas, os neurônios, a estrutura neural procura representar matematicamente a forma de ligação de um neurônio com outro no cérebro (YAMAMOTO, 1998). A Figura 3.2 apresenta uma analogia entre um neurônio biológico e um matemático.

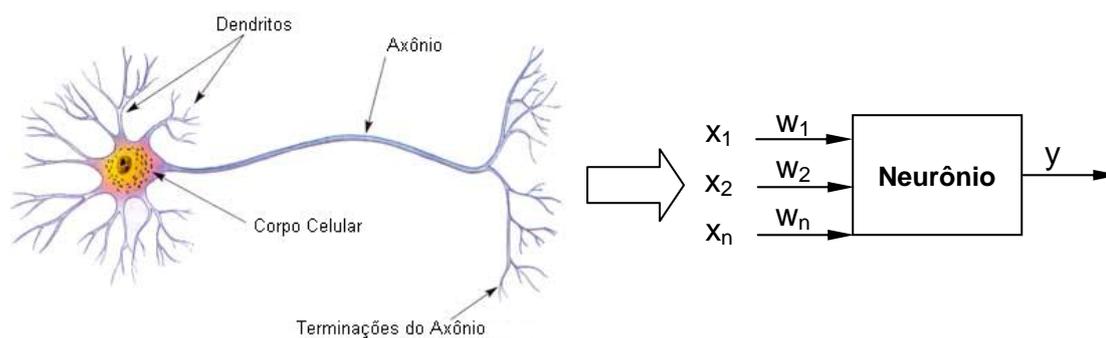


Figura 3.2 - Analogia entre os neurônios biológico e matemático

Fonte: GONZAGA; 2003, p. 53.

No neurônio biológico, a informação é recebida pelo neurônio através dos dendritos, que são terminações nervosas. A informação é então processada pelo corpo celular e a resposta é transmitida para outra célula nervosa através do axônio. O neurônio matemático é uma representação deste sistema. As informações são recebidas na camada de entrada do neurônio (composta por x_1, x_2, \dots, x_n), é processada por pesos, também chamados de pesos sinápticos (w_1, w_2, \dots, w_n) e a resposta é, então, fornecida pelo modelo na camada de saída, y .

Uma RNA típica é composta por uma densa camada de elementos processadores (os neurônios artificiais) não lineares, dispostos em camadas e interconectados através de canais unidirecionais análogos às sinapses de um neurônio biológico (GONZAGA, 2003). Na sua forma mais geral, uma rede neural artificial é uma estrutura para modelar a maneira como o cérebro realiza uma tarefa particular ou função de interesse. Por definição, a rede neural é um processador paralelamente distribuído, constituído de unidades de processamento simples (neurônios) que têm a propensão natural para armazenar conhecimento

experimental e torná-lo disponível para o uso e que se assemelha ao cérebro em dois aspectos (HAYKIN, 2001):

- O conhecimento é adquirido pela rede a partir de seu ambiente através de um processo de aprendizagem.
- Forças de conexão entre os neurônios, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido.

Quando se procura uma técnica que forneça o maior número de vantagens e que funcione bem para a maioria dos problemas de identificação, as redes *Perceptron* Multicamadas (MLP – *Multi-Layer Perceptron*) têm apresentado um crescente aumento nas aplicações de identificação e controle de processos. De acordo com Nørgaard et al (2000 *apud* MELEIRO, 2002), publicações recentes dão conta de que na última década numerosas aplicações práticas evidenciam o fato de que redes do tipo MLP possuem uma impressionante capacidade, tanto para modelagem de processos não lineares, quanto para aplicações em diversas estratégias de controle de processos.

Apesar de haver um aumento de complexidade quando comparado com a identificação de sistemas lineares, muitos resultados estabelecidos na teoria convencional de identificação de sistemas também se aplicam à identificação de sistemas não lineares. Assim, a identificação de sistemas baseada em redes neurais artificiais pode ser encarada, sob muitos aspectos, como uma extensão natural da abordagem convencional de identificação de sistemas lineares (NØRGAARD et al, 2000 *apud*, MELEIRO, 2002).

Para o problema da deslignificação com oxigênio de polpa celulósica foi selecionada a rede *perceptron* de múltiplas camadas (MLP), que é comumente empregada na modelagem e simulação de processos químicos. As redes MLP são

detalhadas na seção 3.3.2.1 e na seção 3.3.2.3 são apresentados modelos neurais aplicados a indústria de celulose e papel encontrados na literatura.

3.3.2.1 Perceptron de Múltiplas Camadas

Modelo de um Neurônio

O neurônio é a unidade fundamental de processamento de informação de uma rede neural. A Figura 3.3 apresenta um modelo de um neurônio não linear que forma a base de um projeto de uma rede neural proposto por McCulloch e Pitts em 1943 (HAYKIN, 2001).

Na figura 3.3 é possível observar os três elementos básicos do modelo neural (HAYKIN, 2001):

- Conjunto de sinapses – cada uma das sinapses é caracterizada por um peso ou força própria;
- Somador – soma os sinais de entrada, ponderados pelas respectivas sinapses do neurônio. As operações neste estágio constituem um combinador linear;
- Função de ativação – uma função utilizada para restringir a amplitude do sinal de saída de um neurônio. Tipicamente, o intervalo normalizado da amplitude da saída de um neurônio é escrito com o intervalo unitário fechado $[0,1]$ ou, alternativamente, $[-1,1]$.

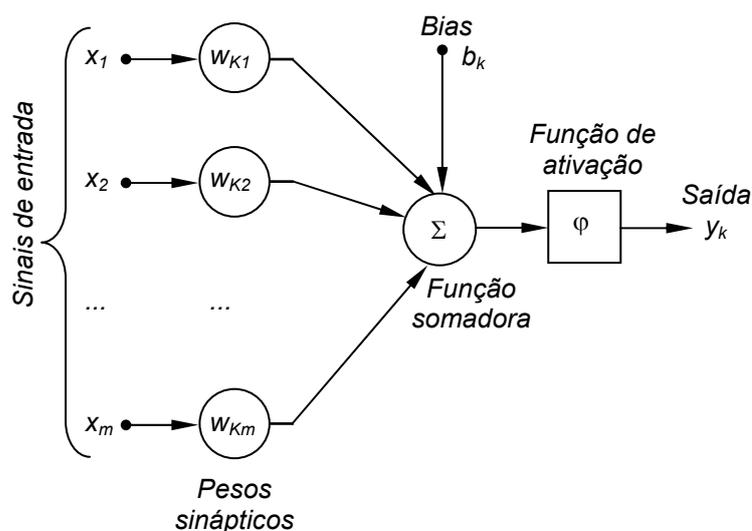


Figura 3.3 - Modelo não linear de neurônio

Fonte: HAYKIN, 2001, p.36.

O modelo neural apresentado na Figura 3.3 inclui um *bias* aplicado externamente à função somadora. O *bias* tem o efeito de aumentar ou diminuir a entrada líquida da função de ativação (HAYKIN, 2001), com o objetivo de permitir uma convergência mais rápida do processo de treinamento.

Em termos matemáticos, o neurônio k pode ser descrito pelas seguintes equações:

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad (3.41)$$

$$y_k = \varphi(u_k + b_k) \quad (3.42)$$

onde x_1, x_2, \dots, x_m são os sinais de entrada; w_1, w_2, \dots, w_m são os pesos sinápticos do neurônio k ; u_k é a saída do combinador linear (função somadora); b_k é o bias; φ é a função de ativação; e y_k é o sinal de saída do neurônio k (HAYKIN, 2001).

A função de ativação define o sinal de saída do neurônio. Os três tipos básicos de função de ativação são as funções de limiar, linear por partes e sigmóide. A função limiar, definida no intervalo de normalização unitário [0,1] (HAYKIN, 2001):

$$\varphi(v) = \begin{cases} 1 & \text{se } v \geq 0 \\ 0 & \text{se } v < 0 \end{cases} \quad (3.43)$$

onde v é o potencial de ativação do neurônio k , isto é, a soma ponderada de todas as entradas sinápticas e do bias:

$$v_k = \sum_{j=1}^m w_{kj} x_j + b_k \quad (3.44)$$

A saída do neurônio k que emprega a função de limiar é expressa como:

$$y_k = \begin{cases} 1 & \text{se } v_k \geq 0 \\ 0 & \text{se } v_k < 0 \end{cases} \quad (3.45)$$

A função de ativação linear por partes assume que o fator de ampliação dentro da região linear de operação é a unidade. Esta forma de função de ativação pode ser vista como uma aproximação de um amplificador não-linear. A função de ativação linear por partes definida para o intervalo de normalização unitário [0,1] (HAYKIN, 2001):

$$\varphi(v) = \begin{cases} 1, & v \geq \frac{1}{2} \\ v, & \frac{1}{2} > v > -\frac{1}{2} \\ 0, & v \leq -\frac{1}{2} \end{cases} \quad (3.46)$$

A função sigmóide é a forma mais comum de função de ativação utilizada na construção de redes neurais. É definida como uma função estritamente crescente que exhibe um balanceamento adequado entre comportamento linear e não-linear. A função sigmóide mais conhecida é a função logística, definida para intervalos de normalização [0,1] (YAMAMOTO, 1998):

$$\varphi(v) = \frac{1}{1 + e^{-\alpha}} \quad (3.47)$$

onde α é o parâmetro de inclinação da função sigmóide. No limite, quando o parâmetro de inclinação se aproxima de infinito, a função sigmóide se reduz à função de limiar (HAYKIN, 2001).

Um outro exemplo de função sigmóide é a tangente hiperbólica, definida para intervalos de normalização [-1,1] (HAYKIN, 2001):

$$\varphi(v) = \tanh(v) \quad (3.48)$$

O uso da função tangente hiperbólica nos neurônios da camada intermediária, associado ao uso funções lineares na camada de saída cria uma rede neural capaz de descrever, com precisão arbitrária, qualquer sistema não linear se a camada intermediária contiver o número de neurônios suficientes. Este número é encontrado com experimentação, testando diversas quantidades até que o erro calculado pela rede não diminua mais (MELEIRO, 2004).

Perceptron de Múltiplas Camadas – MLP

O *perceptron* é a forma mais simples de projetar uma rede neural usada para classificação de padrões. Consiste, basicamente, de um único neurônio com pesos sinápticos ajustáveis e bias. O *perceptron* construído em torno de um único neurônio é limitado a realizar classificação de padrões com apenas duas hipóteses, ou seja, o *perceptron* realiza apenas escolhas entre dois elementos (HAYKIN, 2001).

A rede *perceptron* de múltiplas camadas (MLP) é construída ordenando-se os neurônios em camadas, de modo que cada neurônio em uma dada camada tome como entradas somente as saídas dos neurônios da camada anterior ou entradas

externas. A rede MLP é uma generalização do *perceptron* de camada única e é, provavelmente, o tipo de RNA mais frequentemente utilizado. A principal razão para isto é a sua habilidade de modelar desde mapeamentos simples até relações funcionais mais complexas não lineares (GONZAGA, 2003).

A Figura 3.4 apresenta uma estrutura de rede MLP totalmente conectada, com três entradas, uma camada oculta constituída de dois neurônios e duas saídas. Os neurônios da camada oculta apresentam diferentes funções de ativação (φ_1 e φ_2) e os neurônios da camada de saída funções lineares (F_1 e F_2) (GONZAGA, 2003).

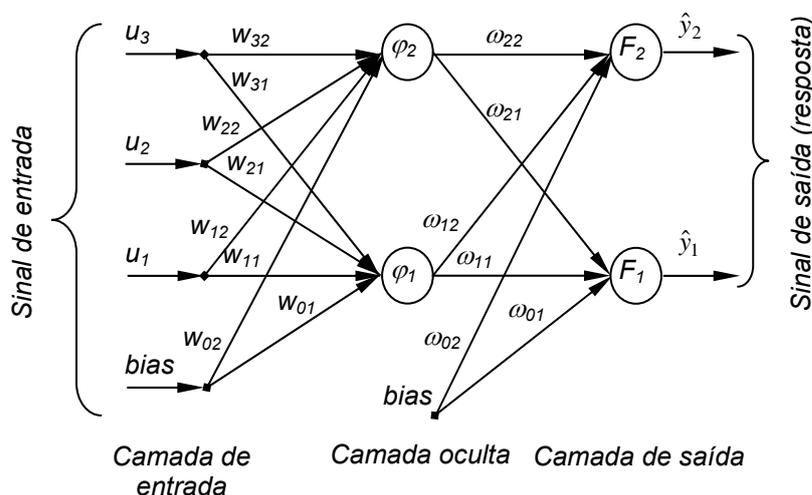


Figura 3.4 - Arquitetura de uma rede MLP

Fonte: GONZAGA, 2003, p. 59.

A rede é constituída, tipicamente, de um conjunto de unidades sensoriais, neurônios, que compõem a camada de entrada da rede, uma ou mais camadas ocultas e uma camada de saída (HAYKIN, 2001). Na camada de entrada os padrões (informações) são apresentados à rede. Na camada intermediária (oculta) é feita a maior parte do processamento, através da conexão ponderada – ligações entre os neurônios das camadas de entrada e intermediária e entre as camadas intermediária

e de saída. A camada de saída finaliza e apresenta os resultados do modelo neural (GONZAGA, 2003).

As redes MLP possuem três características distintivas (HAYKIN, 2001):

- I. O modelo de cada neurônio da rede inclui uma função de ativação não-linear suave (i.e. diferenciável em qualquer ponto). A função utilizada para a modelagem da deslignificação com oxigênio é a tangente hiperbólica apresentada na equação 3.37;
- II. A rede contém uma ou mais camadas de neurônios ocultos, que não são parte da entrada ou da saída da rede. Os neurônios da camada oculta capacitam a rede a aprender tarefas complexas, extraíndo progressivamente as características mais significativas dos vetores de entrada;
- III. A rede exibe um alto grau de conectividade determinado pelas sinapses da rede.

Para determinar o valor dos pesos (na Figura 3.4 - w_{ij} e ω_{ij} , com $i = 0,1,2$ e $j = 1,2$), é preciso que se disponha de um conjunto de exemplos capaz de descrever como as saídas, y_j , estão relacionadas com as entradas u_k ($k = 1,2,3$). A tarefa de determinar os pesos de uma RNA a partir de um conjunto de exemplos é chamada de treinamento ou aprendizagem. Neste processo os pesos são estimados de tal modo que a rede, de acordo com algum critério, modele a relação entrada/saída da forma mais acurada possível (GONZAGA, 2003).

Se existirem as conexões corretas entre as unidades de entrada e um conjunto suficiente de neurônios, sempre é possível encontrar a representação que irá produzir o mapeamento correto da entrada para a saída através das unidades intermediárias. Já foi provado que uma rede neural do tipo MLP com apenas uma

camada oculta, contendo um número suficiente de neurônios, é capaz de aproximar qualquer função contínua e limitada (GONZAGA, 2003).

O Treinamento da Rede MLP

O modelo de RNA mais utilizado atualmente são as redes MLP treinadas com o algoritmo *backpropagation* (retropropagação). Este algoritmo de treinamento foi desenvolvido Rummelhart, Hinton e Williams (*apud* GONZAGA, 2003), em 1986, e mostrou ser possível treinar eficientemente redes com camadas intermediárias. (GONZAGA, 2003).

O treinamento de forma supervisionada com o algoritmo *backpropagation* consiste de dois passos através das diferentes camadas da rede: um passo a frente e outro para trás, Figura 3.5. No passo a frente, a propagação, um padrão de atividade (vetor de entrada) é aplicado aos neurônios da rede e seu efeito propagado através da rede, camada por camada, até que um conjunto de saídas seja produzido como a resposta da rede. Durante a propagação, todos os pesos sinápticos são fixos (HAYKIN, 2001). No passo para trás, a retropropagação, os pesos sinápticos são todos ajustados de acordo com uma regra de correção do erro. Especificamente, a resposta da rede é subtraída de uma resposta real desejada para produzir um sinal de erro. Este sinal de erro é então propagado para trás através da rede, na direção contrária às conexões sinápticas. Os pesos sinápticos são ajustados para fazer com que a resposta da rede se mova para mais perto da resposta desejada, em sentido estatístico (HAYKIN, 2001).

As redes que utilizam o algoritmo de *backpropagation* trabalham com uma variação da Regra Delta Generalizada. A Regra Delta padrão, a rigor, implementa

um gradiente descendente no quadrado da soma do erro para as funções de ativação lineares. A Regra Delta Generalizada funciona quando são utilizados neurônios com uma função de ativação não-linear (GONZAGA, 2003).

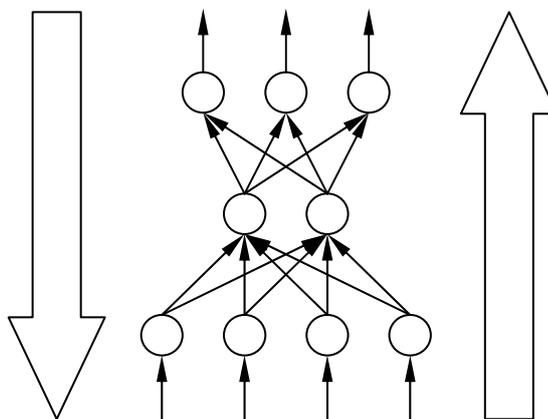


Figura 3.5 - Esquema de treinamento *backpropagation*

Fonte: GONZAGA, 2003, p.64.

Para definir a Regra Delta Generalizada, considere dois neurônios i e k situados em camadas adjacentes como mostrado Figura 3.6. Estes neurônios fazem parte de uma rede onde N_0 é a camada de saída da rede e p o vetor de entrada de padrões ($p = (x_1, x_2, \dots, x_n)$). A saída do neurônio i , y_i , é a entrada do neurônio k , de saída y_k .

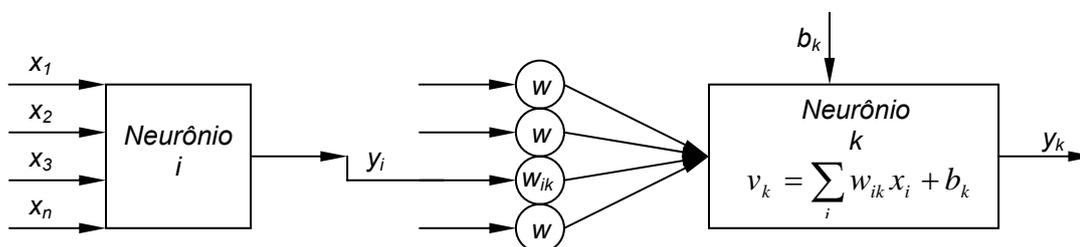


Figura 3.6 - Conexões entre os neurônios i e k

Fonte: GONZAGA, 2003, p. 66.

Na Figura 3.6, o peso relativo à conexão entre i e k é w_{ik} e o parâmetro externo, b_k , é o bias. A atividade, v , do neurônio k no tempo t para o conjunto de padrões de entrada p é dada por (GONZAGA, 2003):

$$v_k^p(t) = \sum_i w_{ik}(t) y_i^p(t) + b_k(t) \quad (3.49)$$

e a saída do neurônio k no tempo $t+1$ é dada pela função de atividade no tempo t (GONZAGA, 2003):

$$y_k^p(t+1) = \varphi_k \left(\sum_i w_{ik}(t) y_i^p(t) + b_k(t) \right) \quad (3.50)$$

A função de ativação utilizada na modelagem da deslignificação com oxigênio é a tangente hiperbólica:

$$y_k^p = \tanh(v_k^p) \quad (3.51)$$

Calcula-se, então, a atividade para todos os neurônios até chegar à camada de saída N_0 , onde se obtém a resposta. A correção é feita retropropagando os erros.

A regra para correção dos pesos é dada por (GONZAGA, 2003):

$$w_{novo} = w_{anterior} + \Delta w \quad (3.52)$$

O método do gradiente é utilizado pra minimizar o erro em relação ao peso. O erro E^p é definido como o erro quadrático para o padrão p no neurônio de saída, e é dado por (GONZAGA, 2003):

$$E^p = \frac{1}{2} \sum_{k=1}^{N_0} (d_k^p - y_k^p)^2 \quad (3.53)$$

Onde d_k^p é a resposta desejada e y_k^p é a resposta obtida da rede. O erro total,

$E = \sum_p E^p$ deve ser minimizado. O valor Δw , para os neurônios i e k , para o conjunto

de padrões p é dado pela relação (GONZAGA, 2003):

$$\Delta_p w_{ik} = -\gamma \frac{\partial E^p}{\partial w_{ik}} \quad (3.54)$$

A relação apresentada na equação 3.54 é derivada do método do gradiente, onde se parte de um ponto arbitrário e caminha-se por uma superfície de erro em direção ao ponto de mínimo. A evolução é sempre no sentido oposto ao gradiente daquele ponto. A constante γ determina o tamanho do passo de correção (GONZAGA, 2003).

É possível calcular $\frac{\partial E^p}{\partial w_{ik}}$ através da aplicação da regra da cadeia (GONZAGA, 2003):

$$\frac{\partial E^p}{\partial w_{ik}} = \frac{\partial E^p}{\partial v_k^p} \frac{\partial v_k^p}{\partial w_{ik}} \quad (3.55)$$

O segundo termo do lado direito da equação 3.55, ou seja, a derivada parcial do potencial de ativação, v , em relação ao peso w_{ik} é igual à derivada da equação 3.49:

$$\frac{\partial v_k^p}{\partial w_{ik}} = y_i^p \quad (3.56)$$

O valor de $\Delta_p w_{ik}$ é, portanto,

$$\Delta_p w_{ik} = \gamma \delta_k^p y_i^p \quad (3.57)$$

Onde $\delta_k^p = -\frac{\partial E^p}{\partial v_k^p}$. Para se obter uma relação para δ , aplica-se novamente a regra

da cadeia (GONZAGA, 2003):

$$\delta_k^p = -\frac{\partial E^p}{\partial v_k^p} = -\frac{\partial E^p}{\partial y_k^p} \frac{\partial y_k^p}{\partial v_k^p} \quad (3.58)$$

A derivada parcial da equação 3.50 (saída do neurônio k no tempo $t+1$) em v fornece $\frac{\partial y_k^p}{\partial v_k^p} = \varphi'(v_k^p)$. Então, o segundo termo da expressão 3.58 é uma derivada simples da função de ativação. Para a determinação do primeiro termo, $\frac{\partial E^p}{\partial y_k^p}$ devem ser considerados dois casos: o neurônio k faz parte da camada de saída da rede ou faz parte de uma camada intermediária da rede (GONZAGA, 2003).

Se o neurônio k faz parte da camada de saída da rede, segundo a definição de E^p , o primeiro termo da equação 3.58 é a derivada direta da expressão 3.53 (GONZAGA, 2003):

$$\frac{\partial E^p}{\partial y_k^p} = -(d_k^p - y_k^p) \quad (3.59)$$

E, então, δ_k^p será escrito como:

$$\delta_k^p = -(d_k^p - y_k^p) \cdot \varphi'(v_k^p) \quad (3.60)$$

Em contrapartida, se o neurônio k fizer parte de uma camada intermediária da rede, é preciso utilizar novamente a regra da cadeia. Utilizando h para representar as camadas da RNA (GONZAGA, 2003):

$$\frac{\partial E^p}{\partial y_k^p} = \sum_{h=1}^{N_0} \frac{\partial E^p}{\partial v_h^p} \frac{\partial v_h^p}{\partial y_k^p} = \sum_{h=1}^{N_0} \frac{\partial E^p}{\partial v_h^p} \frac{\partial}{\partial y_k^p} \sum_{h=1}^{N_h} w_{hk} y_k^p = \sum_{h=1}^{N_0} \frac{\partial E^p}{\partial v_h^p} w_{hk} = -\sum_{h=1}^{N_0} \delta_h^p w_{hk} \quad (3.61)$$

Com a equação 3.50 obtém-se uma outra expressão δ_k^p :

$$\delta_k^p = \varphi'(v_k^p) \sum_{h=1}^{N_0} \delta_h^p w_{hk} \quad (3.62)$$

A partir das equações 3.52 e 3.60, é possível criar um processo recursivo para a atualização de pesos, que é a idéia básica da Regra Delta Generalizada para uma rede *feedforward* (redes em que a informação segue em uma única direção) com função de ativação não linear (GONZAGA, 2003). Maiores detalhes sobre a

implementação computacional dos algoritmos de Retropropagação e da Regra Delta Generalizada podem ser obtidos em Baughman e Liu (1995).

3.3.2.2 Descrição do Procedimento e dos Critérios

O procedimento de modelagem neural envolve cinco etapas básicas: a coleta e a classificação de dados; a configuração da estrutura da rede; o treinamento da rede neural e a análise da confiabilidade do modelo e o uso da rede.

A primeira etapa envolve a coleta e o tratamento de dados referentes às variáveis diretamente envolvidas no processo. O objetivo é obter um histórico representativo de todas as condições operacionais passíveis de ocorrer numa planta. A seguir os dados devem ser separados em conjunto de treinamento e conjunto de teste. Esta última tarefa requer uma análise cuidadosa, para minimizar ambigüidades e erros nos dados (GONZAGA, 2003).

Os dados coletados devem ser significativos e cobrir amplamente o domínio do problema; não cobrir apenas as operações normais e rotineiras, mas também exceções e as condições limites do domínio do processo. Um ponto de destacada importância está relacionado à confiabilidade das medidas que serão apresentadas à RNA. Uma rede treinada com dados oriundos de medidas incertas produzirá respostas de má-qualidade, uma vez que a rede aprenderá os erros (GONZAGA, 2003). Outro ponto que merece atenção se refere aos balanços de massa e energia, sendo importante avaliar se o processo opera em condições de estado estacionário ou transiente a fim de escolher o tipo de estrutura neural mais adequada para a modelagem (YAMAMOTO, 1998).

Os dados, depois de separados, nos conjuntos de treinamento e de teste (utilizados para verificar o desempenho da rede sob condições reais e que não pertencem ao conjunto de treinamento) são, geralmente, colocados em ordem aleatória para a prevenção de tendências associadas à ordem de apresentação dos dados. Além disso, é necessário pré-processar estes dados através de normalizações, escalonamentos e/ou conversões de formato para torná-los mais apropriados à sua utilização na rede. Dados com ordem de grandeza discrepantes podem causar problemas de convergência para a rede (GONZAGA, 2003).

A terceira etapa é a definição da configuração da rede, que pode ser dividido em três passos (GONZAGA, 2003):

- I. Seleção do paradigma neural apropriado à aplicação. Embora este trabalho tenha apresentado apenas o modelo MLP (modelo selecionado para a execução deste trabalho), existem inúmeros outros modelos;
- II. Determinação da topologia da rede: número de camadas e número de neurônios em cada camada;
- III. Determinação do algoritmo de treinamento e funções de ativação. Este é um dos aspectos de maior impacto no desempenho do sistema resultante.

Existem metodologias para a condução destas tarefas. Normalmente estas tarefas são feitas de forma empírica, uma vez que a definição da configuração da rede neural é ainda considerada uma arte. Arquitetar a rede pode demandar bastante tempo, requerer experiência dos projetistas sendo interessante um conhecimento profundo do processo a se modelado pela rede neural (GONZAGA, 2003).

A quarta etapa do processo de modelagem é o treinamento da rede. Nesta etapa, os pesos das conexões são ajustados de acordo com os critérios do algoritmo de treinamento escolhido (GONZAGA, 2003).

A etapa de teste da rede tem início após o treinamento da rede. A validação do modelo é dividida em duas partes: i) validação da rede com os dados do histórico do processo previamente separado do arquivo de dados inicial; ii) testes com dados atuais do processo (GONZAGA, 2003).

Após o modelo ter sido estimado, deve-se proceder a avaliação do mesmo para que se determine se o modelo selecionado atende ou não as necessidades requeridas. A metodologia empregada na validação do modelo identificado deve estar intimamente relacionada com a futura aplicação do modelo. Esta é, geralmente, a etapa mais imprecisa do processo de identificação, uma vez que os critérios de aceitabilidade do modelo nem sempre são muito claros ou fáceis de definir matematicamente (MELEIRO, 2002).

Existem na literatura, uma série de testes padrões que podem ser realizados para avaliar o desempenho do modelo. Estes testes permitem obter uma validação preliminar, onde os piores modelos podem ser descartados facilmente (MELEIRO, 2002).

A maioria dos testes utiliza um conjunto de dados de validação que, além de ser diferente do conjunto de treinamento, representa toda a faixa operacional do sistema. Este procedimento, conhecido genericamente por validação cruzada (*cross-validation*) (HANSEN; SALAMON, 1990 *apud* MELEIRO, 2002) é uma medida direta da capacidade de generalização da rede neural.

Testes de inspeção visual são capazes de dar informações sobre a variação da acurácia do modelo em diferentes pontos de operação do sistema. Consiste de

uma simples observação gráfica, de uma comparação visual entre dos dados reais e simulados (MELEIRO, 2002).

Embora os testes de inspeção visual forneçam uma boa idéia da qualidade do modelo identificado, eles não indicam se o modelo tem falhas que possam ser corrigidas. Ou seja, se houver erros significativos entre as predições do modelo e os dados reais, não é possível saber se o resultado ruim é fruto de problemas de estimação dos parâmetros ou se o modelo identificado não foi capaz de representar satisfatoriamente todas as dinâmicas do processo em questão (MELEIRO, 2002). Para avaliar os erros é comum o uso do método do erro quadrático médio (EQM), que fornece uma boa estimativa da qualidade da predição.

3.3.2.3 Modelos Neurais Aplicados à Indústria de Celulose e Papel

Foram encontradas na literatura diversas aplicações da modelagem neural na indústria papelreira. Quatro delas serão comentadas neste trabalho a título informativo e para exemplificar a versatilidade das RNAs.

Aguiar e Maciel Filho (1998) identificaram um modelo neural para a simulação otimização do processo de produção de polpa celulósica. Assim como neste trabalho, os autores utilizaram dois métodos diferentes para simular um digestor de polpa *kraft*, um modelo determinístico baseado nas equações cinéticas e de transferência de massa e o modelo neural.

Foram utilizadas sete equações diferenciais ordinárias, duas equações diferenciais parciais e uma equação algébrica para representar o modelo determinístico. O modelo neural utilizado foi uma rede MIMO, *feedforward*, com uma

camada escondida e função de ativação sigmoideal. O algoritmo de treinamento utilizado foi o *backpropagation* (AGUIAR; MACIEL FILHO, 1998).

Ambos os modelos produziram resultados similares, simulando a planta com a confiabilidade desejada (intervalo de confiança de 95%). Os autores concluíram que as redes neurais podem ser uma boa ferramenta para o desenvolvimento de modelos com o conhecimento de variáveis de processo e sem qualquer conhecimento da cinética e dos fenômenos de transporte do sistema. Por outro lado Aguiar e Maciel Filho (1998) ressaltam que o modelo determinístico oferece uma melhor interpretação do processo.

Uma outra aplicação das redes neurais na indústria papelreira foi apresentada por Lemmetti, Leiviskä e Sutinen (1998). O objetivo desse trabalho foi a predição do número *kappa* da polpa partindo da análise da concentração de álcali, concentração de sólidos totais dissolvidos e lignina dissolvida em digestores contínuos e em batelada.

Novamente um modelo determinístico foi utilizado paralelamente à rede neural. Entretanto o modelo determinístico utilizado por Lemmetti et al (1998) foi mais simples, envolvendo apenas duas equações diferenciais ordinárias. A rede neural aplicada possui a mesma arquitetura descrita para o modelo de Aguiar e Filho (1998), com apenas uma diferença, a rede é MISO, apenas o *kappa* é estimado pelos modelos determinístico e neural.

Para todos os 11 digestores para os quais foram desenvolvidos modelos, ambas as soluções, fenomenológica e neural, foram satisfatórias e os autores concluem que se dados em quantidade suficiente estiverem disponíveis, quaisquer dos modelos poderá prever o valor do número *kappa* na saída do reator (LEMMETTI et al, 1998).

Numa aplicação completamente diversa das duas já apresentadas, Xie, Ghiaasiaan e Karrila (2004) utilizam a modelagem neural para classificar o tipo de regime de escoamento num reator piloto. O tipo de regime de escoamento está intrinsecamente ligado ao rendimento da reação. A principal motivação para o uso das redes neurais é a medição objetiva e não intrusiva. Foram identificados quatro tipos de regime: escoamento com borbulhamento, *plug flow*, *chum flow* e *slug flow*.

A rede foi composta de três camadas, com função de ativação linear para os neurônios da camada de entrada e funções sigmóides para os neurônios das camadas oculta e de saída. Novamente o algoritmo de treinamento foi o *backpropagation*. Os resultados obtidos foram satisfatórios (dentro do limite de confiabilidade de 95%) e os autores concluíram que a rede pode ser utilizada com este propósito, mas advertiram que o procedimento não foi testado em plantas industriais e, portanto, mais investigações são recomendadas antes da aplicação industrial (XIE et al, 2004).

Uma aplicação de modelagem neural encontrada na literatura é semelhante àquela apresentada neste trabalho. Costa Júnior et al (2004) apresentaram a modelagem matemática da deslignificação com oxigênio de polpa *kraft* de celulose e por este motivo esta aplicação será descrita com maior riqueza de detalhes que as outras três.

O Modelo de Costa Júnior, Costa, Caldeira, Correia e Pessotti

O conjunto original de dados operacionais empregados no estudo foi obtido de uma planta industrial de deslignificação de celulose. Foram analisados mais de

8000 pontos para acompanhar a rotina operacional da fábrica (COSTA JUNIOR et al, 2004).

Foram estudadas redes MISO, tendo como saída o número *kappa* do reator de deslignificação. Como entrada para a rede as seguintes variáveis foram consideradas: número *kappa* de entrada do reator; vazão do vapor de média pressão misturado à polpa, em $\frac{t}{h}$; pressão do vapor misturado à polpa, em $\frac{kgf}{cm^2}$; temperatura do vapor de média pressão misturado à polpa, em $^{\circ}C$; vazão de oxigênio alimentado ao reator, em $\frac{t}{h}$; vazão de álcali alimentado ao reator, em $\frac{m^3}{h}$; vazão de licor branco oxidado alimentado ao reator, em $\frac{m^3}{h}$; vazão de polpa alimentada ao reator, em $\frac{m^3}{h}$; e a consistência da polpa, em % (COSTA JUNIOR et al, 2004).

As redes propostas são todas do tipo *feedforward*, com três camadas, que foram treinadas utilizando o algoritmo *backpropagation*. Foi utilizada uma função sigmoïdal como função de ativação dos neurônios da camada oculta e funções de ativação lineares para os neurônios da camada de saída (COSTA JUNIOR et al, 2004).

Inicialmente os autores utilizaram redes estáticas, a variável de saída, o número *kappa* da polpa reagida, foi predita para com tempo de amostragem fixo, ou seja, sem considerar o atraso da resposta. Os resultados evidenciaram que esta metodologia não é adequada uma vez que é fundamental considerar a dinâmica da planta real (COSTA JUNIOR et al, 2004).

Na segunda etapa do estudo, as entradas foram tomadas no tempo t e a saída em $t+1$, considerando o tempo de atraso das medidas. Esta alteração refletiu

nos resultados da rede neural, melhorando a qualidade da predição. A rede final possui 9 neurônios na camada de entrada, 3 neurônios na camada intermediária e apenas um na camada de saída (COSTA JUNIOR et al, 2004).

Apesar de obter bons resultados, o modelo de Costa Júnior et al (2004) não consideram a real dinâmica do processo, ou seja, o tempo de residência do reator de deslignificação. Os valores apresentados à rede na entrada da rede não são, necessariamente, os que promoveram o número *kappa* obtido no tempo de amostragem estipulado. O tempo de residência do reator é a variável que alinha os dados e favorece a modelagem, gerando modelos mais robustos e que podem ser aplicados a maior faixas operacionais. Esta é a principal diferença entre o modelo proposto por Costa Júnior et al (2004) e o modelo proposto neste trabalho.

Capítulo 4 - Métodos de Otimização

Neste capítulo são discutidos os métodos de otimização utilizados neste trabalho. São abordados dois métodos heurísticos de otimização combinatorial, o Algoritmo Genético e o Recozimento Simulado, ambos aplicados na identificação dos parâmetros do modelo matemático da deslignificação.

4.1 INTRODUÇÃO

A otimização pode ser definida como a ciência que determina a melhor solução matemática para um determinado problema. A otimização determinística baseia-se em relações matemáticas clássicas, como as derivadas (FLETCHER, 2001), enquanto a otimização heurística é baseada na experiência e na analogia com sistemas naturais.

Antes dos anos 40 do século XX, pouco se sabia a respeito dos métodos numéricos para otimização de funções multivariáveis. Embora o método de Newton e outros métodos mais sofisticados para resolução dessas funções fossem conhecidos, qualquer problema mais complexo demandava uma grande quantidade de assistentes operando máquinas de calcular (FLETCHER, 2001).

Os anos 40 e 50 do século passado trouxeram consigo a introdução e o desenvolvimento da programação linear¹ (FLETCHER, 2001). Mas a temática da

¹ O termo programação, aqui, é sinônimo de otimização e foi usado originalmente para significar planejamento ótimo.

otimização determinística só foi revolucionada em 1959 com a publicação do artigo de Davidon (1959 *apud* FLETCHER, 2001), que levou a introdução de métodos que utilizam variáveis métricas. Fletcher e Powell demonstraram que o novo método era muito mais rápido e confiável que os métodos existentes e este avanço transformou rapidamente a solução dos problemas de otimização não linear (NOCEDAL; WRIGHT, 1999). Desde a publicação do artigo de Davidon (1959 *apud* FLETCHER, 2001) o desenvolvimento da otimização foi impulsionado e hoje inclui métodos para a resolução de diversos problemas.

Mesmo com toda a evolução apresentada pelos algoritmos de otimização, alguns problemas reais são muito complexos para serem tratados com formulações matemáticas. Não-linearidades, relações combinatoriais e incertezas geralmente desafiam os métodos clássicos de otimização. Nessas situações, um recurso comumente utilizado é recorrer a uma série de soluções simplificadas e esperar que pelo menos uma forneça uma solução aceitável. Neste contexto as pesquisas de otimização e simulação vêm criando, nos últimos 30 anos, métodos alternativos que produzem soluções de alta qualidade sem as estruturas matemáticas clássicas (GLOVER et al, 1999). Estes métodos são denominados heurísticos, ou seja, métodos derivados da experiência.

Para Foulds (1983 *apud* SILVER, 2004) método heurístico é aquele que se baseia na experiência ou no julgamento e deve fornecer uma solução para o problema, mas não há garantias matemáticas de que a solução seja a ótima. Entre os métodos heurísticos destacam-se a computação evolutiva, o Recozimento Simulado (*Simulated Annealing*) e o Método do Enxame (*Particle Swarm*) entre outros. Neste trabalho, o Algoritmo Genético (*Genetic Algorithm*) - um método

inserido na computação evolutiva - e o Recozimento Simulado são os métodos utilizados.

A computação evolutiva foi proposta em 1991, representando um esforço em congregar diversas iniciativas de pesquisa, originalmente propostas nos anos 50, que visavam simular aspectos específicos do processo evolutivo, (ZUBEN, 2004). O recozimento simulado, completamente diverso do algoritmo genético, foi proposto em 1983 por Kirkpatrick et al (1983) e propõe uma analogia entre a mecânica estatística (o comportamento de sistemas com vários graus de liberdade em equilíbrio térmico em uma temperatura finita) e a otimização multivariável ou combinatorial.

Os métodos determinísticos foram concebidos para resolverem problemas genéricos em um universo específico, onde impera a linearidade, continuidade, diferenciabilidade e/ou estacionariedade. Em contrapartida, os métodos heurísticos foram concebidos para resolverem problemas genéricos em universos genéricos e podem ser aplicados em não-linearidades e em problemas não estacionários. Embora não garantam a eficiência total na obtenção da solução, geralmente garantem a obtenção de uma boa aproximação para a solução (ZUBEN, 2004).

Existem ainda os algoritmos de otimização híbridos, que utilizam ambas as estratégias de otimização, determinística e heurística, combinadas. Estes métodos foram desenvolvidos para tirar vantagem das melhores qualidades dos dois tipos de otimização (BISCAIA; VIEIRA, 2000).

O presente trabalho utiliza os métodos de otimização heurísticos recozimento simulado e algoritmo genético para identificar parâmetros do modelo fenomenológico do processo de deslignificação de celulose.

4.2 MÉTODOS HEURÍSTICOS

Uma das principais motivações para o uso de métodos heurísticos é a dificuldade, se não a impossibilidade, de se obter uma solução ótima para uma formulação matemática que é uma representação aproximada do problema real (SILVER, 2004).

A classe metaheurística utiliza aproximações generalizadas da função para resolver complicados problemas combinatoriais. Por definição, um método metaheurístico é um processo iterativo que guia e modifica operações no algoritmo para produzir soluções de melhor qualidade (OSMAN, 2002). São métodos que possuem meios de escapar de mínimos locais e/ou reduzir o espaço de busca e possuem um ou mais parâmetros ajustáveis que lhes confere flexibilidade, podendo ser aplicados numa diversidade de problemas. O Algoritmo Genético e o algoritmo do Recozimento Simulado se encaixam nesta classe de métodos junto com o método da colônia de formigas e o *tabu search*. Hertz e Widmer (2003) apresentam mais informações sobre a implementação dos métodos heurísticos.

4.2.1 Recozimento Simulado

O método do Recozimento Simulado, *Simulated Annealing*, foi proposto por Kirkpatrick et al em 1983. Inspirado no processo de recozimento de sistemas físicos, as idéias básicas do algoritmo têm suas origens na mecânica estatística, que é composta por métodos para análise de propriedades agregadas de conjunto de átomos em amostras líquidas ou sólidas. O procedimento do recozimento de

sistemas físicos, *annealing*, consiste das seguintes etapas (KIRKPATRICK et al, 1983):

- I. Fundir a substância até que toda ela se torne fluida e os átomos possam se mover com relativa liberdade;
- II. Diminuir a temperatura lentamente de forma que os átomos possam se mover o suficiente para adotarem uma orientação mais estável;
- III. Quando a temperatura se aproxima do ponto de congelamento, degraus de temperatura deverão ser utilizados, ou seja, manter o sistema a temperatura constante por um determinado período de tempo antes de continuar a redução da temperatura.

Encontrar o estado de energia mínima significa encontrar a configuração mais estável do sistema atômico, a configuração ótima do sistema. Assim, a partir de uma analogia com o recozimento descrito acima Kirkpatrick et al (1983) desenvolveram um algoritmo de otimização. O processo iterativo funciona como um rearranjo microscópico do processo modelado pela mecânica estatística, onde a energia é representada pela função objetivo.

Metropolis et al (1953 *apud* KIRKPATRICK et al, 1983) introduziram um algoritmo simples que pode ser utilizado para simular um conjunto de átomos em equilíbrio a uma dada temperatura. A cada passo do algoritmo um átomo recebe um pequeno incremento energético aleatório, alterando desta forma a energia do sistema, gerando um ΔE . Se esta variação for menor que zero, $\Delta E \leq 0$, a redução é aceita e a nova configuração passa a ser a configuração de menor energia e, conseqüentemente a configuração mais estável do sistema.

Se $\Delta E > 0$ um teste probabilístico é efetuado para a decisão de aceitação ou não da nova configuração. Na mecânica estatística, o teste probabilístico é calculado de acordo com o fator de probabilidade de Boltzmann (KIRKPATRICK et al, 1983):

$$P(E) = \exp\left(-\frac{E\{r_i\}}{k_b T}\right) \quad (4.1)$$

onde $P(E)$ a probabilidade da nova configuração ser aceita, $E\{r_i\}$ é a energia da configuração r na i -ésima iteração, $k_b = 1,380622 \cdot 10^{-23} \frac{J}{K}$ é a constante de Boltzmann e T a temperatura, em Kelvin, do sistema. Este teste probabilístico é executado para evitar que o sistema fique estagnado em um estado de energia mínima local, ou seja, podem existir estados mais estáveis, mas para alcançá-los o sistema tem que passar por configurações de maior energia (KIRKPATRICK et al, 1983).

Utilizando a função objetivo (Δf) no lugar da energia e definindo a configuração por um conjunto de parâmetros $\{x_i\}$, é possível utilizar o procedimento de Metropolis et al (1953 *apud* KIRKPATRICK et al, 1983) para gerar diversas configurações, soluções de uma dada função objetivo a uma dada temperatura T . O procedimento do Recozimento Simulado consiste em, primeiramente, “fundir” o sistema a ser analisado a uma “temperatura alta”, ou seja, gerar soluções e testá-las, calculando o Δf . Sempre que $\Delta f > 0$, analogamente ao procedimento empregado na mecânica estatística, o teste probabilístico é efetuado para verificar a aceitação do valor ΔE calculando (KIRKPATRICK et al, 1983):

$$P(\Delta f) = \exp\left(-\frac{\Delta f}{T}\right) \quad (4.2)$$

onde $P(\Delta f)$ é a probabilidade do novo conjunto solução $\{x_i\}$ ser aceito. Depois, a temperatura é reduzida em estágios lentos até que o sistema congele e nenhuma outra mudança ocorra no conjunto solução (KIRKPATRICK, et al., 1983).

Os principais parâmetros deste método são aqueles envolvidos na programação do recozimento, ou seja: a temperatura T ; a razão de redução de temperatura α ; o número de iterações na mesma temperatura nT (o tamanho do degrau); e o número de ciclos de redução de temperatura nR (a quantidade de degraus). A Figura 4.1 apresenta o algoritmo do Recozimento Simulado.

A temperatura será utilizada para calcular a probabilidade do novo ponto x_i ser aceito, portanto a temperatura inicial deve ser de ordem de grandeza próxima a diferença do valor da função no ponto atual e no ponto anterior, Δf (FOGEL; MICHALEWICZ, 2002).

O parâmetro α regula o quanto a temperatura diminui a cada ciclo. É interessante que a redução seja suave, principalmente porque nos últimos degraus da redução de temperatura os valores de Δf diminuem e é desejável que a temperatura seja grande o suficiente para que os pontos não sejam aceitos (FOGEL; MICHALEWICZ, 2002).

Os valores de nT e nR são escolhidos pelo programador. Devem ser escolhas sensatas, utilizar patamares de temperatura muito longos não melhoram significativamente os resultados, e patamares muito curtos significam que o tempo e o espaço de busca será menor. O mesmo raciocínio vale para o número de ciclos de redução de temperatura, muitos ciclos não vão produzir melhorias na solução, mas poucos ciclos podem não ser suficiente para que uma boa solução seja alcançada (FOGEL; MICHALEWICZ, 2002).

Mais informações acerca da implementação do algoritmo do Recozimento Simulado podem ser encontradas em Fogel e Michalewicz (2002) e em Correia e Tavares (1999).

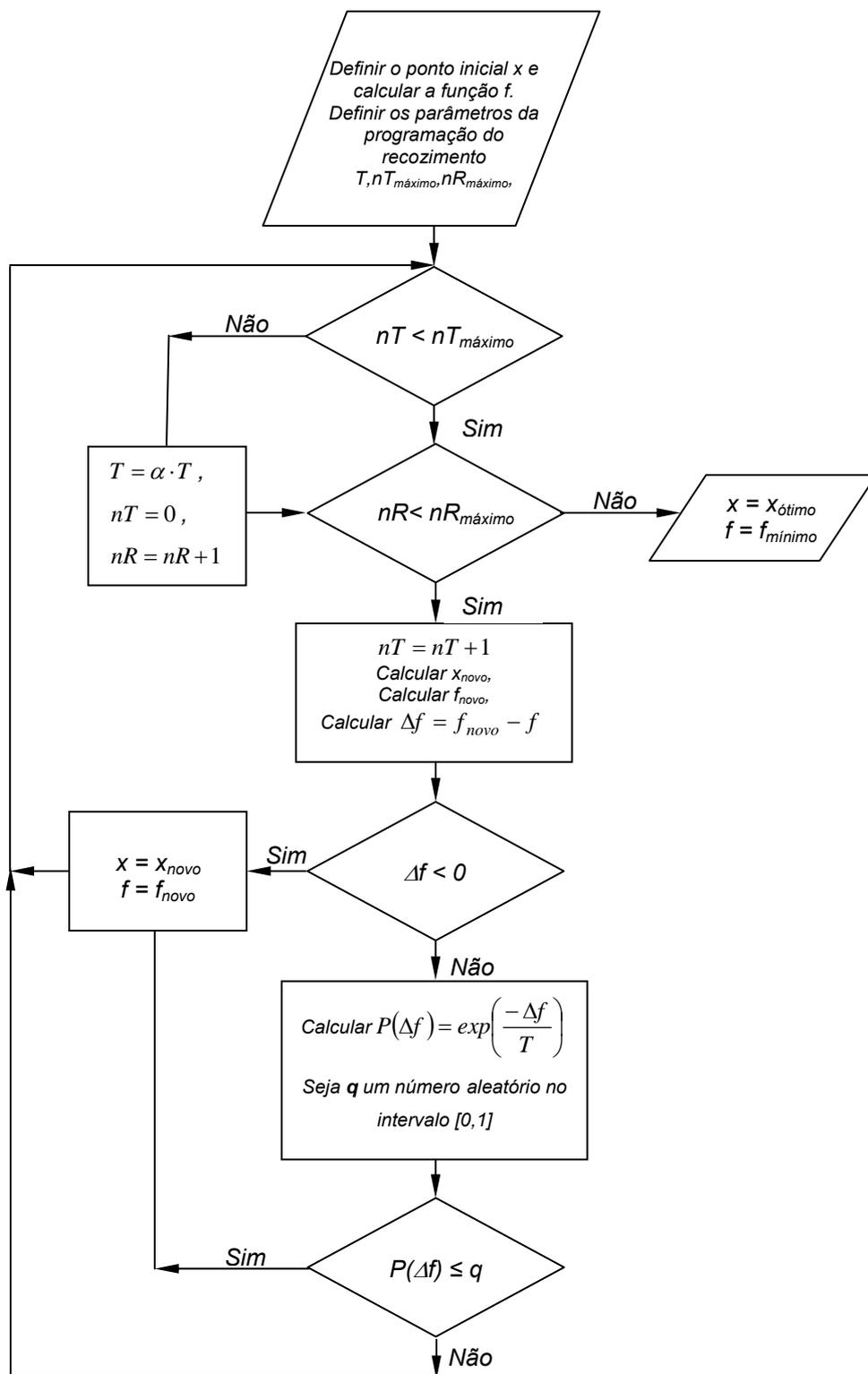


Figura 4.1 - Algoritmo do Recozimento Simulado

4.2.2 Algoritmo Genético

Conceitos de computação evolutiva têm sido empregados em uma variedade de disciplinas, desde ciências naturais e engenharia até biologia e ciências da computação. A idéia original, surgida nos anos 50, é aplicar o processo de evolução natural como um paradigma de solução de problemas, a partir de sua implementação em computador. Os problemas de otimização são aqueles que vêm recebendo mais contribuições a partir de técnicas de computação evolutiva (ZUBEN, 2004).

A vantagem mais significativa está na possibilidade de resolver problemas pela simples descrição matemática do que se deseja presente na solução, não havendo necessidade de se indicar explicitamente os passos até o resultado. É lógico que os algoritmos evolutivos correspondem a uma seqüência de passos até a solução, mas esses passos não são os mesmos para problemas diversos. Sendo assim, a computação evolutiva deve ser entendida como um conjunto de técnicas e procedimentos genéricos e adaptáveis, a serem aplicados na solução de problemas complexos pra os quais outras técnicas conhecidas são ineficazes ou nem sequer são aplicáveis (ZUBEN, 2004).

Em termos históricos, três algoritmos para a computação evolutiva foram desenvolvidos independentemente, os Algoritmos Genéticos (AGs) à partir de 1957; a programação evolutiva à partir de 1962; e as estratégias evolutivas à partir de 1965 (ZUBEN, 2004).

Este trabalho dissertará somente sobre os algoritmos genéticos. O primeiro algoritmo genético funcional para solucionar problemas antes considerados intratáveis foi proposto por John Holland em 1975 (*apud* McCALL, 2005).

O esquema de Holland e a hipótese de construção de blocos de Goldberg (1989 *apud* McCALL, 2005) forneceu o embasamento teórico e conceitual para o projeto eficiente dos AGs (McCALL, 2005). Mais informações sobre computação evolutiva podem ser encontradas em Zuben (2004) e Silver (2004).

Os Algoritmos Genéticos são uma técnica heurística motivada originalmente pelo princípio de Darwin da seleção natural. O AG utiliza uma versão abstrata do processo evolutivo, criando uma população de cromossomos artificiais onde cada cromossomo representa uma solução para o problema (McCALL, 2005).

Iniciado com uma população gerada aleatoriamente, o AG calcula um processo adaptativo baseado em seleção e recombinação para criar a população sucessora, a nova geração. Durante a recombinação, cromossomos “pais” são selecionados e seu material genético é combinado, produzindo o cromossomo “filho”. Esses cromossomos são tratados da mesma maneira e assim novas gerações de soluções para o problema são criadas. Nesse processo iterativo, a seqüência de gerações evolui e a qualidade das respostas tende a aumentar até que o critério de parada seja satisfeito (McCALL, 2005).

O algoritmo é construído com diversos componentes distintos que lhe conferem robustez. Os principais componentes são a codificação dos cromossomos, o grau de ajuste, a seleção, a recombinação e o esquema de evolução (McCALL, 2005).

O AG manipula populações de cromossomos, que são representações da solução do problema. Os cromossomos do algoritmo são abstrações do DNA biológico e cada posição particular no cromossomo é denominado gene. Os genes são os responsáveis pelas características físicas e psicológicas de seu portador. Analogamente, o cromossomo matemático é um vetor que contém valores para as

variáveis independentes do problema e cada gene é, portanto, uma variável independente (McCALL, 2005).

A codificação do cromossomo é uma decisão importante na elaboração do AG uma vez que ela determina a precisão do método. O AG clássico utiliza a representação binária $\{0, 1\}$. Neste tipo de codificação cada gene é descrito com uma seqüência de zeros e uns e cada um dos algarismos da seqüência é denominado um bit. Os problemas nos quais o algoritmo é comumente aplicado são necessários muitos bits. Dependendo da precisão requerida, um número inteiro pode demandar 20 bits, inviabilizando a codificação clássica (McCALL, 2005). Um processo de codificação mais eficiente é a codificação real. Cada gene é representado por um número real, facilitando a codificação e melhorando o tempo de processamento do algoritmo (BISCAIA JUNIOR; VIEIRA, 2000).

O grau de ajuste é o valor da função objetivo para um cromossomo e é utilizado para avaliar a qualidade e selecionar as melhores soluções geradas naquela população. A teoria da seleção de Darwin (1859 *apud* ZUBEN) é utilizada como base para este processo. O processo de seleção natural indica que os indivíduos cujas variações (provenientes das combinações genéticas) melhor se adaptam ao meio ambiente terão a maior probabilidade de sobreviver e se reproduzir (ATMAR, 1992 *apud* ZUBEN, 2004).

A seleção no AG procura garantir que os cromossomos que possuem as melhores soluções possuam as maiores chances de serem selecionados mais de uma vez, assegurando assim que as elas se mantenham na nova geração. O método tradicional de seleção é a roleta, que confere a cada cromossomo um peso relativo à qualidade da solução que ele contém. Ou seja, o melhor cromossomo possui o maior peso e ocupa, desta forma, a maior fatia da roleta, aumentando sua

chance de ser escolhido para ser recombinado e ser mantido na próxima geração (McCALL, 2005).

Existem ainda outros métodos de seleção, como o método do torneio, onde dois cromossomos são selecionados da população e aquele que possui o melhor grau de ajuste é selecionado para a recombinação e a seleção. Mais informações sobre métodos de seleção de indivíduos no algoritmo genético podem ser encontradas em McCALL (2005).

A recombinação é o processo no qual os indivíduos selecionados são recombinados para formar os membros da próxima população. A idéia é simular a mistura do material genético que ocorre na reprodução de organismos. Há duas maneiras de proceder a recombinação, que são os operadores genéticos de cruzamento e de mutação. Esses operadores não possuem comportamento determinístico, existindo, portanto, um fator estatístico que interfere tanto no cruzamento como na mutação dos cromossomos (McCALL, 2005).

O operador de cruzamento representa a mistura do material genético de dois cromossomos pais para produzir um ou dois cromossomos filhos. Após a seleção de dois cromossomos pais, um número aleatório no intervalo $[0,1]$ é gerado e comparado com a probabilidade de cruzamento pré-estabelecida no algoritmo. Se o valor gerado for maior que a probabilidade, os pais não sofrem alterações, caso contrário acontece o cruzamento e dois novos cromossomos são criados na nova população. O ponto de cruzamento é escolhido aleatoriamente e os cromossomos filhos são construídos com os caracteres de um dos pais até o ponto de cruzamento e com os caracteres do segundo pai após o ponto de cruzamento (McCALL, 2005). A Figura 4.2 ilustra este processo, que corresponde ao fenômeno de *cross link* numa multiplicação celular de material genético.

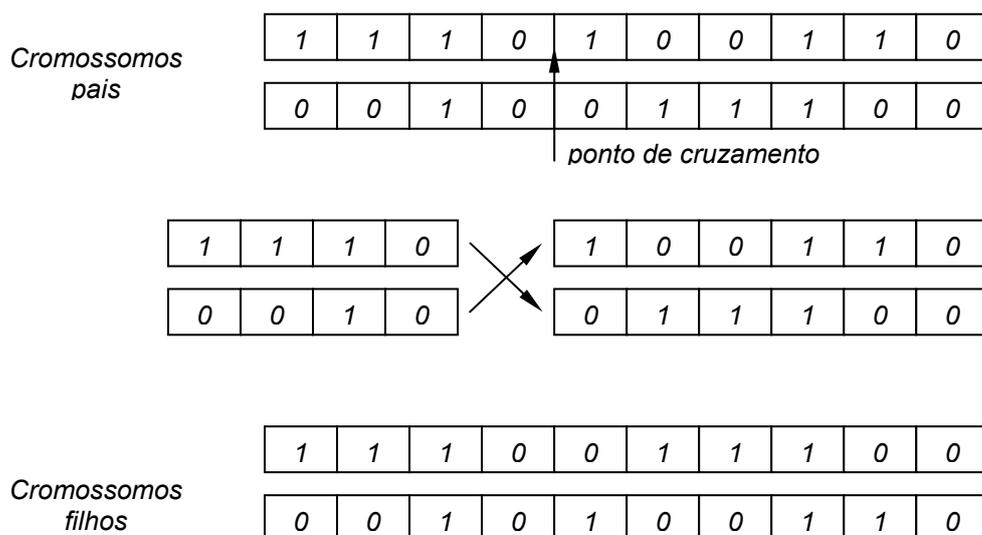


Figura 4.2 - Operador genético de cruzamento

Fonte: McCALL, 2005.

Existem diversas maneiras de utilizar o operador de cruzamento. Podem ser utilizados mais de um ponto de cruzamento e podem ser criados um ou dois filhos. Todas as alterações no operador são escolhas do programador do algoritmo e devem ser opções que favoreçam a qualidade do resultado final (McCALL, 2005).

Depois do cruzamento, os cromossomos filhos vão passar pelo estágio de mutação. A mutação é um operador que age no cromossomo individualmente, trocando um ou mais genes entre si. O procedimento para a aplicação ou não deste operador é o mesmo utilizado no cruzamento. Um número aleatório entre $[0,1]$ é comparado à probabilidade de mutação previamente estabelecida. É importante ressaltar que, assim como nos sistemas biológicos, a probabilidade de ocorrer mutação no cromossomo é significativamente menor do que o cruzamento. A taxa de cromossomos que sofrem mutação em relação à população de mutação é, geralmente, da ordem de 10^{-3} . A Figura 4.3 ilustra o processo de mutação no cromossomo (McCALL, 2005):

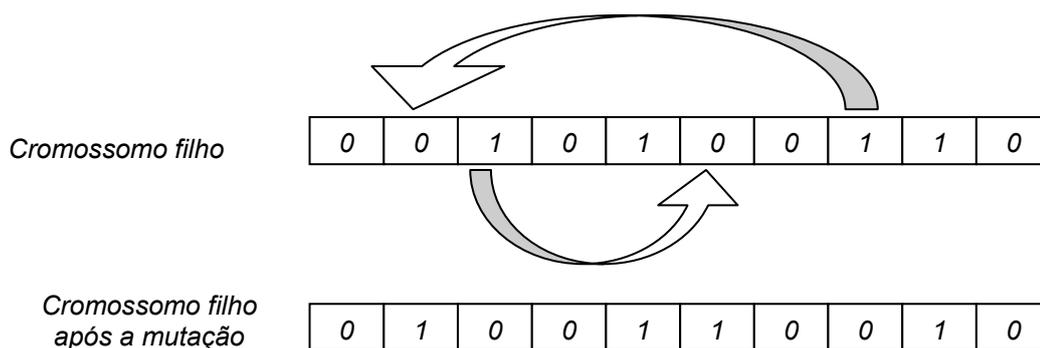


Figura 4.3 - Operador genético de mutação

Após a recombinação, os cromossomos resultantes constituem uma nova população. Os processos de seleção e recombinação são então repetidos e uma outra população é criada. O processo iterativo (evolutivo) continua até que o critério de parada seja satisfeito. Este critério pode ser um número de gerações fixo, a observação do valor da função objetivo, ou uma solução que satisfaça as restrições do problema (McCALL, 2005).

Uma das maneiras utilizadas na escolha do esquema de evolução da população envolve a quantidade de cromossomos que não sofrerão quaisquer alterações quando passam de uma geração para outra. É possível criar uma população completamente nova, na qual todos os cromossomos da geração anterior passem pelos operadores de recombinação ou uma população que cria um único cromossomo novo a cada geração. O esquema evolutivo comum de ser usado é a substituição por elitismo. Neste esquema toda a população é renovada sendo que o melhor indivíduo da geração anterior é preservado na nova geração para garantir que uma boa solução não seja perdida (McCALL, 2005).

Uma vez que a codificação dos cromossomos, o grau de ajuste, a seleção, a recombinação e o esquema de evolução estão definidos, uma série de outros parâmetros deve ser decidida durante a implementação do Algoritmo Genético. Essas escolhas incluem o tamanho da população, o número de gerações, as taxas de cruzamento e mutação, aumento ou não da população a cada geração e critérios de parada e de reinício. Uma boa estrutura para o algoritmo é uma combinação de experiência, do modelo do problema e experimentação de diferentes esquemas e parâmetros. A Figura 4.4 apresenta uma estrutura típica de AG.

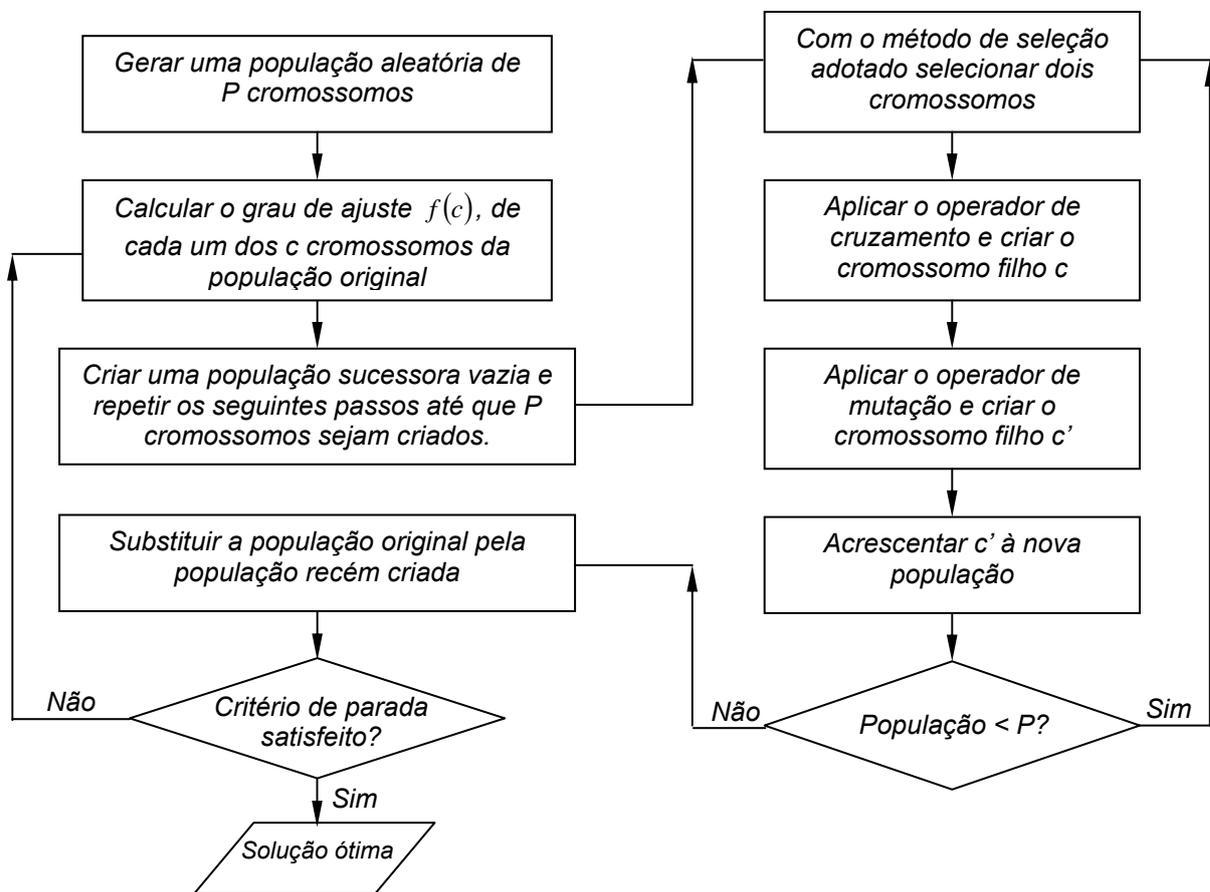


Figura 4.4 - Estrutura básica do Algoritmo Genético.

Capítulo 5 - Metodologia

Este estudo é uma investigação acerca da aplicação de métodos de modelagem à deslignificação com oxigênio. Com este propósito, no Capítulo 5, apresenta-se a metodologia empregada para o desenvolvimento dos modelos fenomenológico e neural, além dos métodos de otimização empregados para identificar os parâmetros do modelo fenomenológico. Descrevem-se, ainda, a forma de coleta e tratamento de dados, etapas fundamentais para o desenvolvimento de modelos confiáveis e representativos o processo.

5.1 COLETA DE DADOS

O conjunto de dados industriais utilizados no presente trabalho foi obtido na própria fábrica de celulose – Klabin Celulose & Papel Unidade Monte Alegre, Telêmaco Borba – PR - através de uma planilha de acompanhamento diário da fábrica. Os dados correspondem a leituras horárias dos instrumentos de campo e de análises de laboratório. Para aumentar a confiabilidade dos dados, alguns cuidados foram tomados para eliminar possíveis problemas, tais como: erros de leitura no instrumento, paradas na produção e produção inferior ao padrão estabelecido pela empresa.

Foram coletados dados desde janeiro de 2005 até julho do mesmo ano – constituindo um conjunto total de 3984 padrões (conjunto de dados composto das variáveis de entrada e saída)– tempo suficiente para captar a dinâmica do processo e gerar a quantidade de dados necessária para análises confiáveis.

O histórico da planta não pode ser avaliado no desenvolvimento dos modelos porque a fábrica apresenta mudanças de composição da matéria-prima, alterando as características da polpa e, conseqüentemente, a quantidade de reagentes utilizados na deslignificação com oxigênio. Para levar este efeito em conta, um procedimento de identificação de parâmetros pode ser aplicado para cada período onde a matéria prima é composta de madeira de origem e tipo diferentes do estudo apresentado neste trabalho.

A empresa mantém um histórico de todas as suas plantas em um *software* denominado *PI (Process Information)*, a partir do qual os dados do intervalo de análise foram compilados. Todos os dados obtidos por leituras de instrumentos de campo foram lidos no painel de controle através deste *software*.

A produção de celulose é calculada por uma relação que envolve a consistência, a vazão e o rendimento da polpa. Este valor foi comparado ao valor da meta estipulada pela empresa. Foram descartados os dados que estivessem mais de 10% abaixo da meta.

O tempo de residência também foi uma variável calculada. A partir do volume do reator e da produção de celulose, o tempo médio de residência foi estimado para cada ponto do conjunto de dados.

O número *kappa* de entrada da planta foi obtido por um analisador instalado no processo e comparado com a análise laboratorial para assegurar a qualidade da análise do instrumento. Foram desconsiderados os valores do instrumento que apresentaram desvios maiores que 5% em relação às análises laboratoriais. Foram considerados os valores lidos nos instrumentos, uma vez que não são procedidas feitas análises laboratoriais do número *kappa* de saída da planta de deslignificação.

Os valores da pressão parcial e da vazão do oxigênio também foram obtidos através das leituras dos instrumentos indicadores no campo. Uma vez que o oxigênio utilizado na reação de deslignificação é obtido da destruição de ozônio, sua concentração na corrente é aproximadamente 90%, sua variação não excede a 2%. Desta forma a vazão de oxigênio é uma variável mais adequada para caracterizar a contribuição do oxigênio.

A vazão de álcali foi obtida por leituras de instrumentos de campo. A concentração utilizada, entretanto, foi uma média diária uma vez que o analisador de campo não estava em funcionamento no período avaliado e as análises laboratoriais são feitas uma vez por dia. Presume-se que o maior erro associado aos resultados obtidos está relacionado a esta média diária.

A temperatura do corpo do reator é obtida por termopares e a consistência por um analisador no campo. Ambos os valores foram obtidos através do *PI*.

Os valores das variáveis foram normalizados para facilitar o treinamento das redes neurais. As variáveis foram normalizadas no intervalo $[-1,1]$ segundo a seguinte expressão:

$$x_{norm} = -1 + 2 \left(\frac{x - x_{inf}}{x_{sup} - x_{inf}} \right) \quad (5.1)$$

onde x é a variável a ser normalizada, x_{norm} é a variável normalizada e x_{inf} e x_{sup} são os limites inferior e superior do intervalo de normalização $[-1, 1]$, respectivamente.

5.2 TRATAMENTO DOS DADOS

Do conjunto original de dados, os meses de janeiro e fevereiro (1464 padrões - conjuntos de dados composto das variáveis de entrada e saída já citadas) foram

desconsiderados por caracterizarem períodos em que a fábrica estava em parada geral para manutenção. Foram, então, selecionados 1331 padrões para as primeiras análises. Neste conjunto, foram considerados valores dos meses de março, abril e maio de 2005. Os meses de junho e julho seriam utilizados para a validação dos modelos.

Entretanto, o número *kappa* de saída da planta no conjunto de padrões separados para o treinamento apresenta duas faixas de operação distintas, produzindo polpas de características distintas. A Figura 5.1 ilustra a variação desta variável para os 1331 padrões.

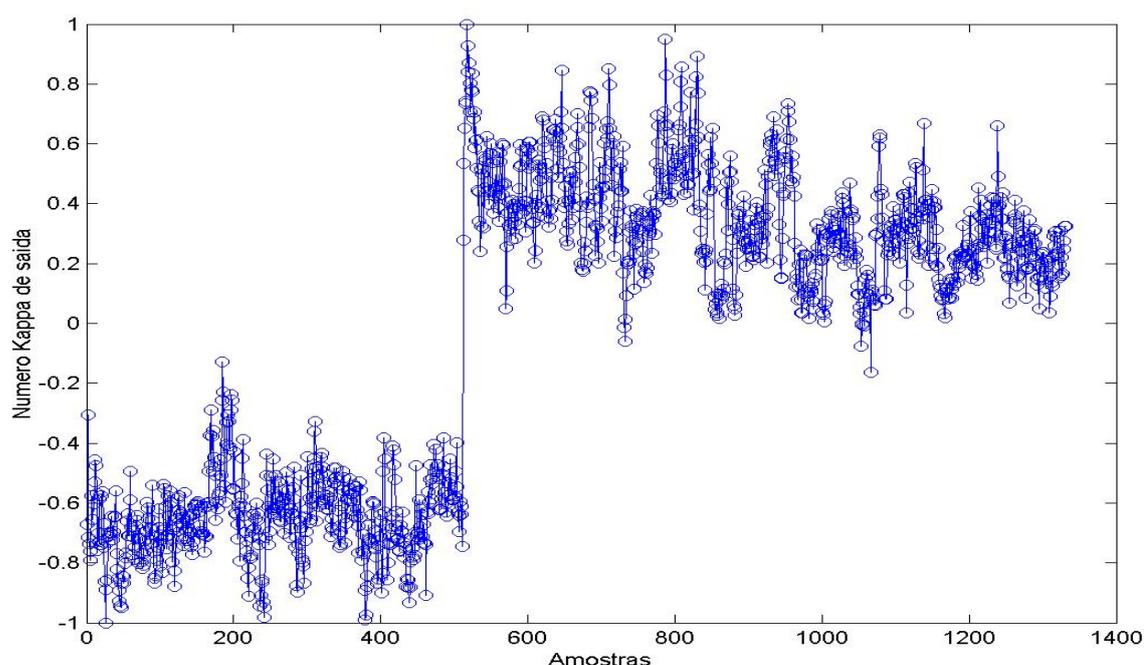


Figura 5.1 - Número *kappa* de saída do reator normalizado no conjunto de 1331 padrões

A primeira faixa de operação não está de acordo com as metas estipuladas para a planta, pois os valores estão abaixo do limite, indicando um provável erro de instrumento ou de características das matérias-primas e de operação. Esses dados foram, então, descartados e o conjunto de dados foi reduzido para 811 padrões –

suficientes para o processo de modelagem. A Figura 5.2 apresenta a variação do *kappa* de saída para este último conjunto de dados.

Após a primeira seleção de dados, os 811 padrões foram submetidos à análise estatística para descartar aqueles que contivessem erros grosseiros, àqueles causados por problemas de calibração de instrumento, de coleta e erros de correlação temporal dos dados.

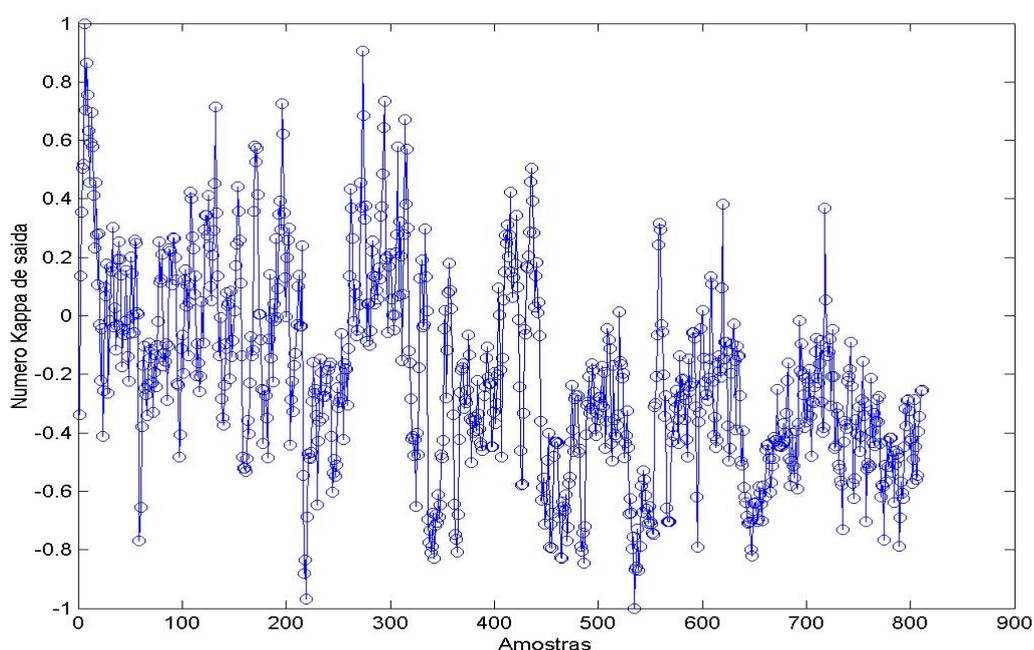


Figura 5.2 - Número *kappa* de saída do reator normalizado no conjunto de 812 padrões

A variável utilizada para analisar o problema de erros grosseiros foi o número *kappa* de saída da planta de designificação. Foi empregado um teste de hipóteses média do número *kappa* de saída dado pela equação 5.2 (SPIEGEL et al, 2004):

$$|\bar{x} - \mu_0| > t_{1-\alpha/2} \left(\frac{S_x}{\sqrt{n-1}} \right) \quad (5.2)$$

onde \bar{x} é o valor médio do número *kappa* de saída do reator; μ_0 é o valor esperado do número *kappa* de saída do reator; t é o valor de t da distribuição t de Student, α é

o grau de confiança do teste de hipóteses (%); S_x é o desvio padrão do conjunto de pontos do intervalo; e n é o número de pontos amostrados.

Calculando o valor de t de Student para um grau de confiança de 95% com relação ao número $kappa$ de saída do reator foram obtidos os limites de aceitação dos pontos operacionais industriais. Os parâmetros estatísticos calculados para a variável número $kappa$ de saída do reator bem como os limites de aceitação são apresentados na Tabela 5.1.

Tabela 5.1 - Análise estatística básica e análise t de Student de significância

Variável		Número $kappa$ de saída do reator
Número de amostras (n)		811
Média		10,013
Erro padrão da média		0,0127
Desvio padrão		0,363
Variância		0,132
Valor máximo da variável		11,295
Valor mínimo da variável		9,174
Limites do intervalo de confiança (95%)	Inferior	9,988
	Superior	10,038

Dos 811 padrões, 428 estão classificados dentro do intervalo de confiança. Uma vez que 428 padrões podem não ser suficiente para a modelagem neural, novos padrões foram selecionados do conjunto de dados composto pelos valores das variáveis nos meses de junho e julho de 2005. Os novos padrões foram testados pela análise t de Student e inseridos no conjunto de dados para completar

novamente os 811 padrões. Este conjunto foi utilizado para treinar e validar os modelos.

A Tabela 5.2 apresenta a variação e as medidas estatísticas para as demais variáveis do processo envolvidas na modelagem.

Tabela 5.2 - Análise estatística básica para as variáveis do processo

Variável	Média	Erro padrão da Média	Desvio Padrão	Variância	Valor Mínimo	Valor Máximo
Consistência (%)	10,695	0,000897	0,0255	0,000653	10,420	10,813
Vazão de Oxigênio (kg/h)	306,80	0,628	17,90	320,40	256,11	367,63
Pressão de Oxigênio (kPa)	343,22	0,0261	0,745	0,555	339,27	348,49
Vazão da solução de NaOH (L/min)	21,605	0,0693	1,975	3,901	18,012	26,990
Temperatura (K)	384,35	0,0113	0,323	0,104	383,31	385,76
Tempo de residência (min)	65,049	0,188	5,352	28,644	51,000	82,000
Produção (a.d.t/d)	283,47	0,759	21,62	467,45	228,66	348,29

A variação observada no número *kappa* de saída é devido às flutuações na vazão de álcali, Figura 5.3. O processo é controlado atualmente através de uma razão entre o número *kappa* de entrada e a vazão de álcali, de acordo com a expressão 5.3:

$$F_{NaOH} = \alpha \cdot K_i \quad (5.3)$$

onde F_{NaOH} é a vazão de álcali em L/min , α é uma constante de proporcionalidade e K_i é o número *kappa* na entrada do reator. O hidróxido de sódio é o principal reagente da deslignificação e conseqüentemente quanto maior a quantidade de álcali no reator maior o rendimento da reação, menor o número *kappa* na saída.

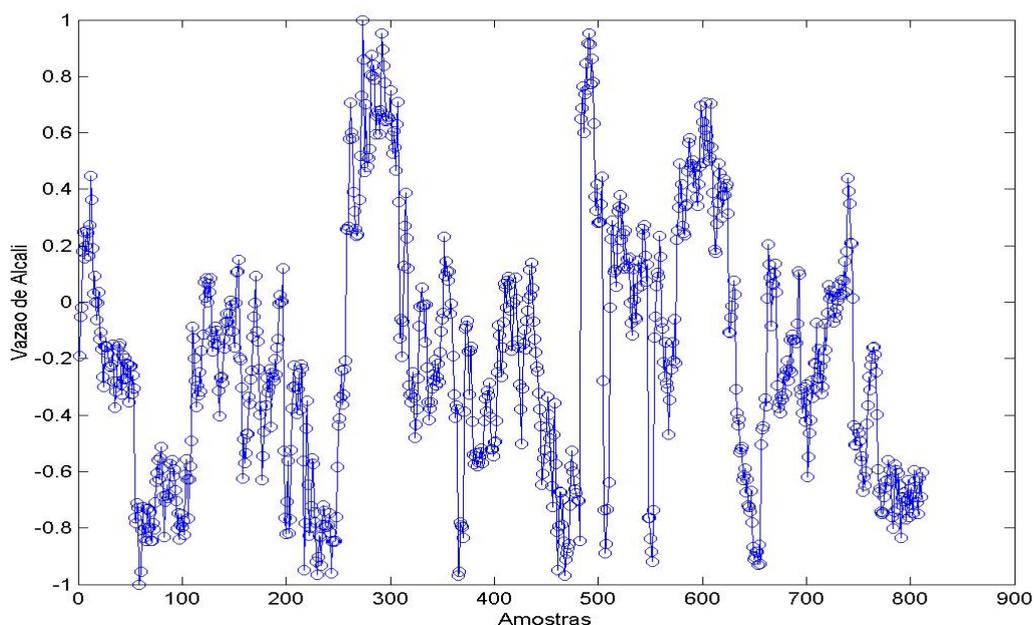


Figura 5.3 - Vazão de álcali normalizada na entrada do reator de deslignificação

A Figura 5.4 apresenta a variação do número *kappa* de entrada do reator de deslignificação. De acordo com a expressão 5.3, a vazão de álcali deveria acompanhar a tendência do número *kappa* de entrada. Esta tendência não é observada por problemas de operação; quando o número *kappa* na entrada do reator é alto, o valor de α é alterado para assegurar que o valor do *kappa* na saída esteja dentro dos limites estipulados para a produção. Este procedimento prejudica o processo, que é sujeito a constantes alterações.

As demais variáveis - pressão e vazão do oxigênio, temperatura e consistência - não apresentam flutuações acentuadas, representando o comportamento estacionário do processo.

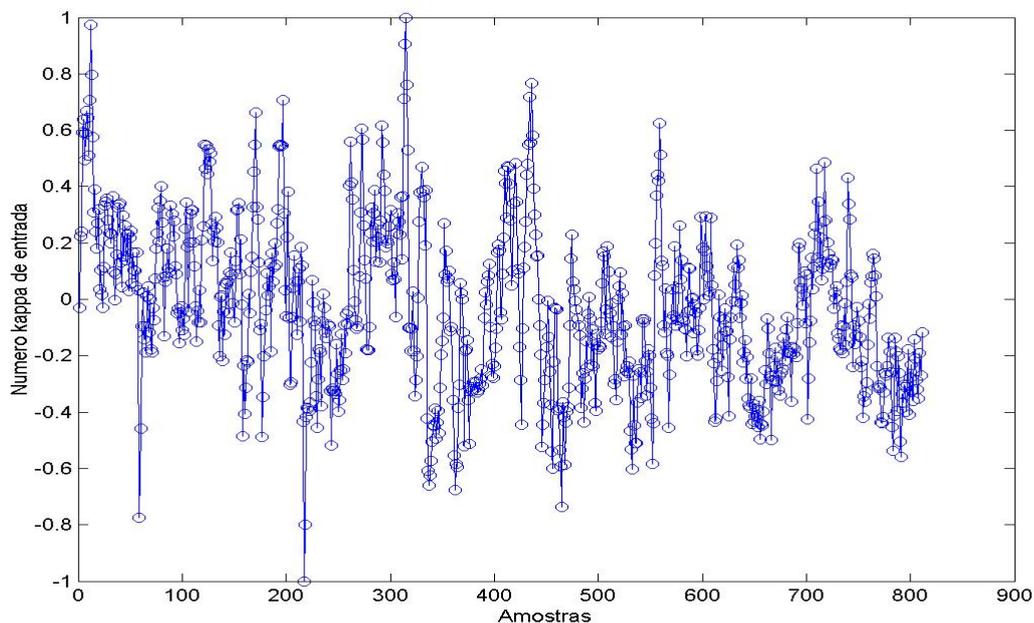


Figura 5.4 - Variação do número *kappa* na entrada do reator normalizado

5.3 DESENVOLVIMENTO DO MODELO FENOMENOLÓGICO

Foram selecionados e implementados três modelos previamente apresentados no Capítulo 3 seção 3.2, o modelo de Hsu e Hsieh (1988), o modelo de Agarwal et al (1999) e o modelo de Violette (2003). O primeiro apresenta a reação em duas etapas, considerando a fase rápida e a fase lenta da reação. Este modelo foi escolhido em detrimento dos demais que consideram a reação em dois estágios e também por ter sido referência para a implementação dos demais modelos. Além disso, Hsu e Hsieh (1988) não utilizam o conceito de Ackert et al (1975) que separava a lignina em três categorias de reação: rápida, lenta e não reativa, simplificando a solução do modelo.

O modelo de Agarwal et al (1999) foi escolhido por representar a cinética de deslignificação sem as duas etapas de reação, ou seja, um modelo mais simples. Por isso é uma boa opção para comparar com o modelo mais elaborado proposto por Hsu e Hsieh (1988).

O terceiro modelo escolhido, o modelo de Violette (2003), é ainda mais simples que o proposto por Agarwal et al (1999), uma vez que não considera a influência do oxigênio na reação de deslignificação. Desta forma é possível avaliar a influência do oxigênio no processo, verificando o excesso ou a falta deste reagente na reação. Este modelo foi implementado ainda para analisar a perda de viscosidade da polpa. A viscosidade é um importante parâmetro de qualidade da polpa e pode ser utilizado como um indicador de desempenho do processo.

Escolhidos os modelos da cinética, foi preciso definir modelos para o consumo de álcali. Na literatura, foram encontradas duas expressões algébricas, uma linear e a outra não-linear. De maneira geral, as reações apresentam perfil de consumo de reagente exponencial e, desta forma, o modelo não-linear deveria ser o que melhor descreveria o sistema. Considerando que este trabalho possui caráter investigativo, ambos os modelos foram implementados.

Uma vez que ambos os modelos algébricos não apresentaram bons ajustes dos dados industriais para dois dos três modelos estudados, foi desenvolvido um modelo diferencial, baseado no balanço de massa para solução de hidróxido de sódio. A melhoria nos resultados foi significativa e o modelo diferencial, equação 5.4, foi adotado para descrever o consumo de hidróxido de sódio para os Modelos 2 e 3 análogos a Agarwal et al (1999) e a Hsu e Hsieh (1989), respectivamente.

$$\frac{d[OH^-]}{dt} = \frac{F_{solução} ([OH^-]_{inicial} - [OH^-])}{V} - r_{[OH^-]} \quad (5.4)$$

onde $[OH^-]$ é a concentração de álcali na solução; $[OH^-]_{inicial}$ é a concentração inicial de álcali; $r_{[OH^-]}$ é a taxa de consumo de álcali durante a reação escrita de acordo com as taxas de consumo encontradas na literatura; e $F_{solução}$ é a vazão da solução hidróxido de sódio; e V é o volume do reator. As unidades das variáveis dependem do modelo utilizado.

O sistema de equações diferenciais ordinárias de primeira ordem, composto pela cinética da reação de deslignificação e do modelo de consumo de álcali, foi resolvido por integração direta no tempo médio de residência do reator. Para a integração foi utilizado o método de Runge-Kutta de quarta ordem com passo variável. Para todos os três modelos selecionados a primeira simulação foi resolvida utilizando-se os parâmetros indicados na literatura.

Para garantir que o resultado obtido com a integração do sistema de equações estivesse de acordo com o valor real lido na saída do reator, os parâmetros foram ajustados através de rotinas de otimização. O fluxograma da Figura 5.5 apresenta o procedimento para o desenvolvimento do modelo.

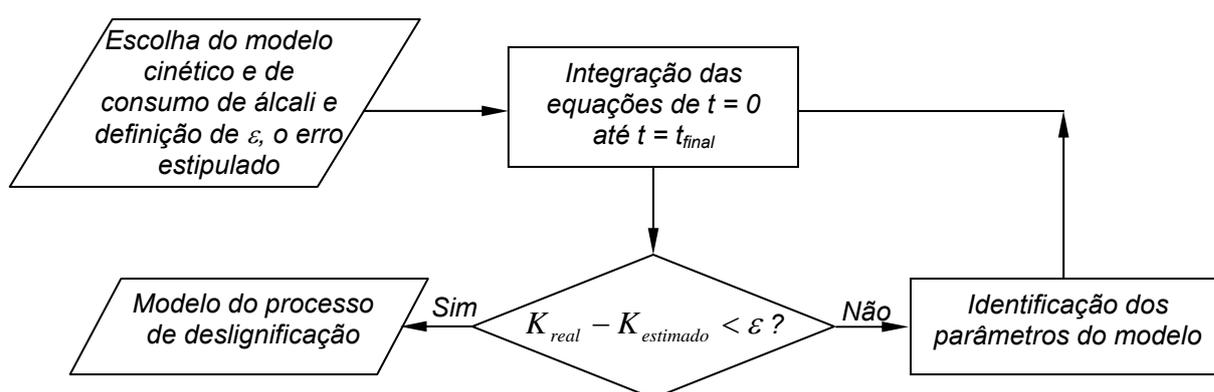


Figura 5.5 - Fluxograma do procedimento utilizado para o desenvolvimento do modelo fenomenológico

5.4 A IDENTIFICAÇÃO COM O MODELO LINEAR ARX

A primeira etapa da modelagem é a preparação dos dados. Uma análise prévia dos dados é importante e geralmente evita, ou pelo menos reduz, a tarefa de testar um grande número de diferentes estruturas de modelos.

O arquivo de dados destinado para a identificação do processo pode conter erros de medidas e outras características tais como ruídos e dados fora da faixa de operação. Outro procedimento importante na preparação dos dados é a normalização. As variáveis podem ter magnitudes diferentes e conseqüentemente aparentar serem mais significativas do que realmente são. Os parâmetros ajustáveis do modelo podem não conseguir minimizar estas influências e, desta forma, o modelo pode ser prejudicado (BAUGHMAN; LIU, 1995).

Após a preparação dos dados, é necessário selecionar as variáveis que mais influenciam o sistema. Para a planta de deslignificação com oxigênio, foram utilizadas as seguintes variáveis: número *kappa* de entrada do reator, consistência da polpa, vazão e pressão de oxigênio, vazão de álcali, temperatura de reação, tempo de residência e produção calculada. A variável estimada pelo modelo é, como para os modelos anteriores, o número *kappa* de saída do reator de deslignificação.

Foram identificados dois modelos distintos, o primeiro utilizando as oito variáveis recém citadas. O segundo modelo considera apenas cinco, número *kappa* de entrada do reator, vazão de oxigênio, vazão de álcali, temperatura de reação e produção calculada. Estas variáveis foram selecionadas por influenciarem mais significativamente o processo.

O erro quadrático médio (EQM) foi utilizado para avaliar a qualidade de predição do modelo. Um baixo valor de EQM pode implicar que não existe

necessidade de se utilizar um modelo mais elaborado, como modelo neural, para representar o sistema. O teste de inspeção visual também foi utilizado pra avaliar o modelo linear.

5.5 A IDENTIFICAÇÃO COM O MODELO NEURAL

Como na identificação do modelo linear, o primeiro passo para a utilização de rede neural é a análise de todos os dados e informações relevantes. É preciso identificar com precisão quais são as variáveis que influenciam significativamente o comportamento do processo e quais as variáveis que não interferem. Como a rede neural aprende a partir do mapeamento entrada/saída fornecido pelos dados, se forem fornecidas informações erradas e/ou incompletas, ou de forma incorreta, as previsões do modelo também o serão.

A escolha de uma estrutura de rede neural tipo *feedforward* de três camadas está ligada às boas referências (HAYKIN, 2001; GONZAGA, 2003; YAMAMOTO, 1998) que são encontradas na literatura de sua aplicação a sistemas estacionários. Os dados coletados na indústria se referem ao estado estacionário e, portanto, esta estrutura é a mais indicada para a referida modelagem.

A quantidade de neurônios na camada intermediária é decisiva para o sucesso do modelo. Não existe regra ou equação para esta escolha, as melhores ferramentas para esta definição são bom senso, experiência e, principalmente, a evolução do erro de validação. Quanto maior o número de neurônios na camada intermediária, maior é o número de pesos sinápticos que a etapa de treinamento deverá ajustar e melhor a previsão da rede. Entretanto, o aumento demasiado do número de neurônios pode causar o *overfitting* (sobre-ajuste), o que significa que a

rede se tornou rígida e não possui capacidade para prever resultados para conjuntos de dados diferentes dos dados de treinamento (perda da capacidade de generalização). Um dos objetivos da etapa de treinamento é, justamente, determinar a quantidade adequada de neurônios na camada intermediária (YAMAMOTO, 1998)

Foram testadas, para a planta de deslignificação com oxigênio, diversas estruturas neurais. As seguintes variáveis foram utilizadas para alimentar a rede neural: número *kappa* de entrada do reator, consistência da polpa, vazão e pressão de oxigênio, vazão de álcali, temperatura de reação, o tempo de residência do reator e produção calculada. A variável estimada pela rede é, como para os modelos anteriores, o número *kappa* de saída do reator de deslignificação. Dois modelos foram identificados, o primeiro com oito variáveis de entrada e o segundo com cinco variáveis – as mesmas aplicadas no modelo linear com cinco variáveis de entrada.

Na modelagem da deslignificação com oxigênio de polpa celulósica foram considerados, inicialmente, 2 neurônios em sua camada oculta. Foram procedidos, então, seguidos testes com redes com diferentes quantidades de neurônios para selecionar a rede com a melhor capacidade de previsão. A quantidade de neurônios foi variada desde 2 até 20 neurônios na camada oculta.

A função de ativação escolhida foi a função tangente hiperbólica, para os neurônios da camada oculta e uma função linear para os neurônios de camada intermediária. O algoritmo de treinamento utilizado foi o *backpropagation*. A Figura 5.6 apresenta o fluxograma do desenvolvimento do modelo neural.

As redes neurais foram escolhidas e avaliadas com testes de inspeção visual e com a análise do EQM.

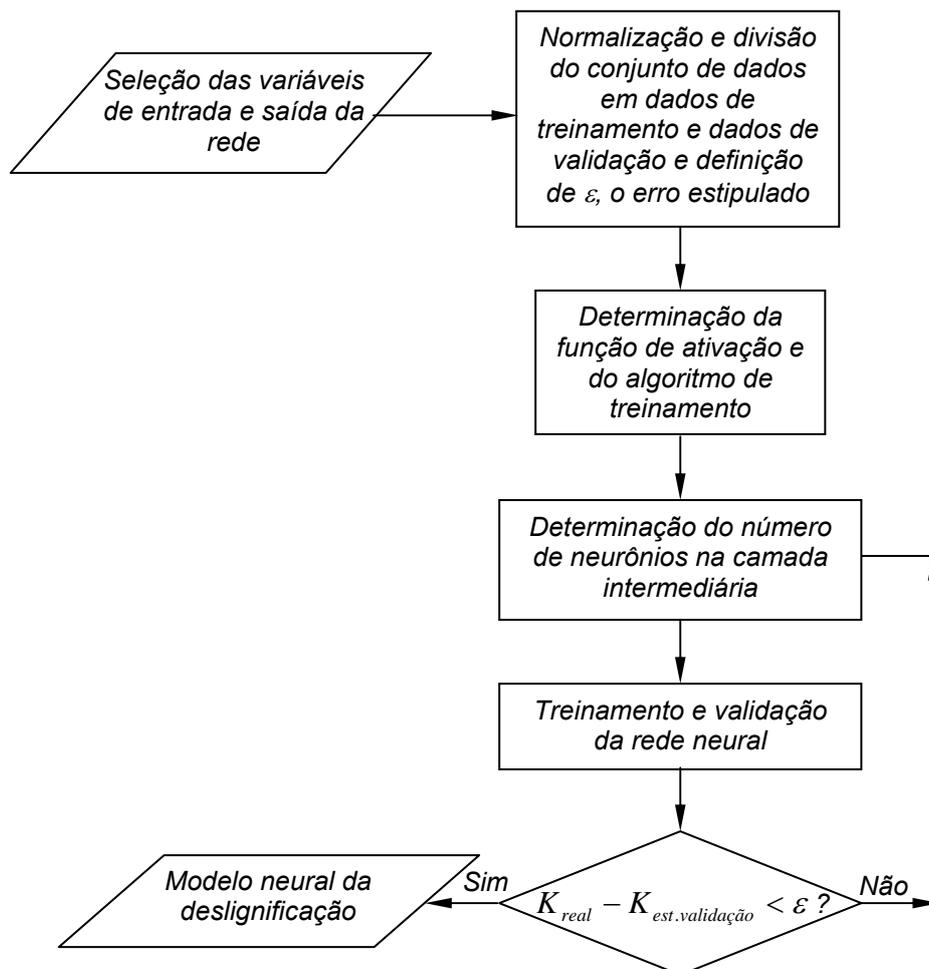


Figura 5.6 - Fluxograma do procedimento utilizado no desenvolvimento do modelo neural

Para todos os modelos o erro (ε) não foi calculado, mas minimizado. Ou seja, o modelo final utilizado para simular a planta de deslignificação é aquele que apresenta o menor valor de ε para todos os 811 padrões. Foi estipulado, entretanto, um valor máximo para ε , a fim de assegurar qualidade de resultados para os modelos, assim:

$$K_{real} - K_{estimado} < 1 \quad (5.5)$$

onde K_{real} é o número *kappa* da polpa e $K_{estimado}$ é o valor do número *kappa* estimado pela rede neural ou pelo modelo fenomenológico. O valor estipulado para a diferença representa um erro de 10% para a previsão e é o erro médio encontrado nos modelos da literatura.

5.5 EMPREGO DOS ALGORITMOS DE OTIMIZAÇÃO

Na modelagem da deslignificação com oxigênio, foram utilizados os algoritmos de otimização para a identificação dos parâmetros do modelo cinético, já que não foi possível estimá-los por falta de dados. Foram empregados os algoritmos heurísticos para esta identificação, uma vez que procedimentos determinísticos não se aplicam a estas tarefas.

Na otimização heurística com o algoritmo genético, a rotina de integração do modelo cinético foi acoplada ao algoritmo de otimização, assim a cada nova população de n cromossomos, n conjuntos de parâmetros são gerados e testados. Ao final de m gerações o melhor conjunto de parâmetros foi utilizado para o modelo fenomenológico. A metodologia empregada é bastante prática uma vez que a cada nova geração o conjunto de parâmetros já é testado e, portanto, no término do algoritmo o modelo já está definido, não sendo necessários outros testes.

É preciso lembrar, entretanto, que os algoritmos genéticos utilizam variáveis aleatórias e desta forma, a cada inicialização do método novos parâmetros são calculados. A otimização dos parâmetros foi realizada 10 vezes e o melhor conjunto entre os dez foi o empregado para a validação do modelo fenomenológico.

Procedimento similar foi utilizado com o algoritmo do recozimento simulado. Uma estimativa inicial dos valores dos parâmetros foi fornecida ao algoritmo e em cada degrau de temperatura um conjunto de soluções é gerado. Ao final do degrau, o melhor conjunto de parâmetros é salvo e novo degrau de temperaturas é simulado. O processo iterativo fornece, no final do algoritmo, nR conjuntos de parâmetros. Os nR conjuntos são comparados e aquele que fornece o menor erro de estimativa do número $kappa$ é utilizado para simular a planta de deslignificação.

Como nos algoritmos genéticos, o recozimento simulado utiliza variáveis aleatórias, fornecendo resultados diferentes cada vez que o algoritmo é iniciado. A mesma metodologia empregada nos AGs foi utilizada para este método, ou seja, o algoritmo foi reiniciado 10 vezes e o melhor conjunto solução entre os dez foi utilizado no modelo cinético.

Capítulo 6 - Resultados e Discussões

Este capítulo apresenta os resultados obtidos com os modelos fenomenológicos, lineares e neurais desenvolvidos para a planta de deslignificação. Primeiro são analisados os resultados obtidos com os três modelos fenomenológicos simulados. Na seção 6.2 analisam-se os resultados fornecidos pelos modelos lineares ARX; e na seção 6.3, os resultados obtidos com os modelos neurais. Uma comparação entre os resultados obtidos com as três técnicas é apresentada no final do capítulo. Os modelos foram ajustados para 405 padrões, enquanto os outros 406 foram utilizados para validação. Os resultados apresentados neste capítulo são de validação.

6.1 MODELAGEM FENOMENOLÓGICA

Os modelos fenomenológicos estudados, embora possuam abordagens teóricas diferentes, devem apresentar resultados semelhantes se devidamente ajustados. Isso porque todos os três modelos foram desenvolvidos para a mesma reação, sendo que cada um deles possui características específicas distintas. Assim, os resultados obtidos se diferenciam pelas variáveis utilizadas, na forma de apresentação do modelo e nos valores dos parâmetros.

Na descrição das metodologias empregadas, Capítulo 5, a seção 5.3 abordou a metodologia para a estimação dos parâmetros dos modelos. Conforme explicado, a reação não foi reproduzida em laboratório e, portanto, os parâmetros não puderam ser estimados de acordo com a metodologia mais frequentemente utilizada. Para

contornar este problema, rotinas de otimização heurísticas foram empregadas para ajustar os parâmetros.

Foram ajustados os valores das constantes das equações da taxa da reação de deslignificação, a energia de ativação, o fator de frequência de Arrhenius e o parâmetro de ajuste da equação da taxa de consumo de álcali.

O valor da energia de ativação não foi muito diferente entre os modelos. Este comportamento era esperado uma vez que se trata do mesmo processo.

Os valores dos expoentes que evidenciam a influência da concentração de álcali, da pressão do oxigênio e do número *kappa* inicial da polpa não foram modificados por serem muito sensíveis. Isto significa que pequenas alterações causam problemas numéricos que podem inviabilizar uma solução.

É importante observar, entretanto, que estes expoentes são diferentes quando os modelos são comparados. Para Hsu e Hsieh ((d)1988), o valor da influência da concentração de álcali é, 0,78 para a primeira etapa e 0,70 para a segunda etapa da reação, enquanto que para Agarwal et al (1999) este valor é 0,92 e para Violette (2003) 0,59. A diferença entre os valores é dada pelas particularidades dos processos modelados, mas numericamente não representam mecanismos de reação intrinsecamente distintos.

A principal discussão acerca dos expoentes reside na ordem da reação com relação ao número *kappa*. Segundo a literatura (AXEGAR et al, 1979; OLM; TEDER, 1979; OLM; TEDER, 1981; MYERS; EDWARDS, 1989; AGARWAL et al, 1999) o alto valor é devido a uma série de reações paralelas, resistências à transferência de massa e condensação de lignina. Este trabalho não discute as razões deste valor ser alto, mas confirma que, independente do modelo empregado, o valor da ordem da reação em relação ao número *kappa* inicial é alto e que valores menores que 3

conseguem reproduzir valores obtidos com os instrumentos industriais de acompanhamento e controle de processos.

6.1.1 Ajuste do Modelo 1

O modelo mais simples, Modelo 1, foi desenvolvido a partir de uma analogia com o modelo de Violette (2003) e foi o primeiro a ser implementado. Uma vez que dados de viscosidade não estavam disponíveis, somente a reação de deslignificação foi modelada. Os algoritmos genéticos e do recozimento simulado foram aplicados para ajustar os parâmetros.

O Modelo 1 foi implementado com a taxa de consumo de álcali linear e dependente exclusivamente do número *kappa* da polpa. A equação não linear de consumo de álcali sugerida por Myers e Edwards (1989) ou a expressão diferencial derivada para melhorar os resultados dos outros dois modelos estudados não apresentaram melhorias neste caso. Desta forma, foi mantido o formato original do modelo proposto por Violette (2003). A equação 6.1 apresenta o ajuste para a taxa de consumo de álcali:

$$r_{[OH^-]} = 6 \cdot \Delta K + [OH^-] \quad (6.1)$$

A melhor aproximação conseguida utilizando como base o modelo de Violette (2003) está expressa na equação 6.2:

$$-\frac{dK}{dt} = 11,6 \cdot 10^3 \cdot \exp\left(-\frac{72,50 \cdot 10^3}{RT}\right) \cdot K^{3,12} [OH^-]^{0,588} \quad (6.2)$$

A Figura 6.1 apresenta uma comparação entre os resultados obtidos com o modelo e os dados reais da planta. Os resultados obtidos acompanham a tendência dos valores reais da planta de deslignificação mesmo sendo um modelo que não

considera a influência do oxigênio no mecanismo de reação. Violette (2003) utilizou excesso de oxigênio durante seus experimentos e, portanto, quaisquer resistências à transferência de massa ou problemas na dissolução do gás foram minimizadas pelo excesso de oxigênio. Este resultado pode levar a concluir que o processo estudado também utiliza excesso de oxigênio e, portanto, sua influência na cinética da reação de designificação pode ser desconsiderada.

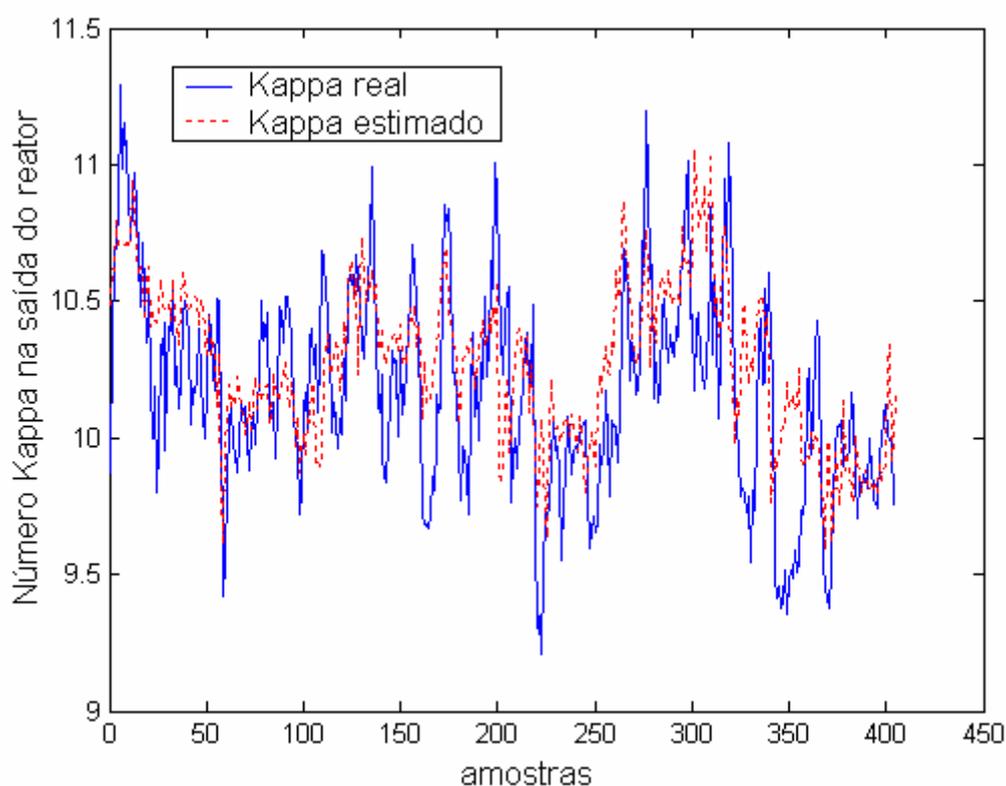
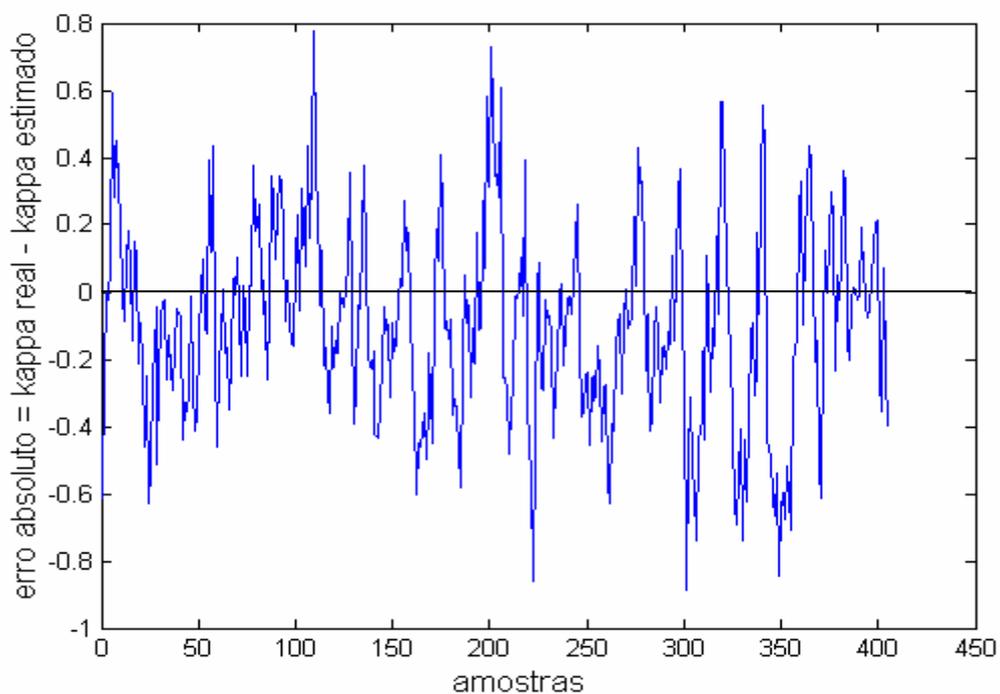


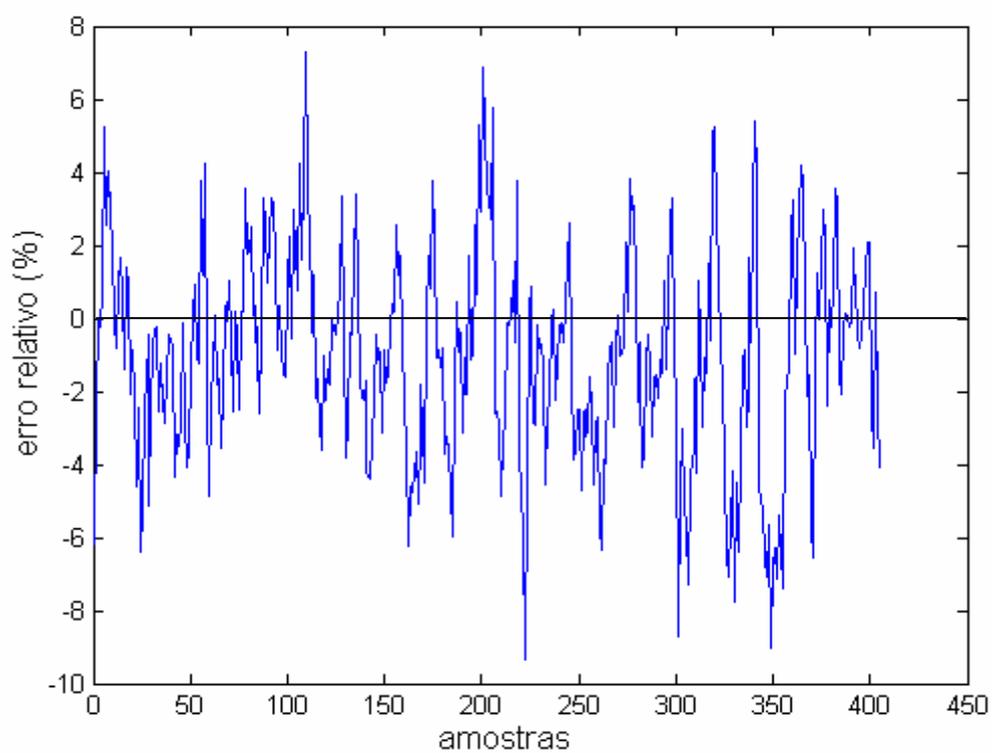
Figura 6.1 - Comparação entre os valores do número *kappa* reais e estimados pelo Modelo 1

Foram calculados os erros absolutos dos valores estimados, ou seja, a diferença entre o valor real e o valor calculado pelo modelo. Os erros calculados estão distribuídos no intervalo de $[-0,93; 0,73]$, ou um erro relativo médio máximo de 9,3%. Estes resultados estão apresentados nas Figuras 6.2 *a* e *b*. Em termos estatísticos os valores calculados pelo modelo estão dentro do desvio padrão da média, Figura 6.3, com apenas algumas exceções. Melhorias no ajuste modelo

podem ser provavelmente efetuadas com o uso de valores confiáveis de concentração de álcali na entrada do reator de deslignificação.



(a)



(b)

Figura 6.2 - (a) Erro absoluto e (b) erro relativo médio calculados para o Modelo 1

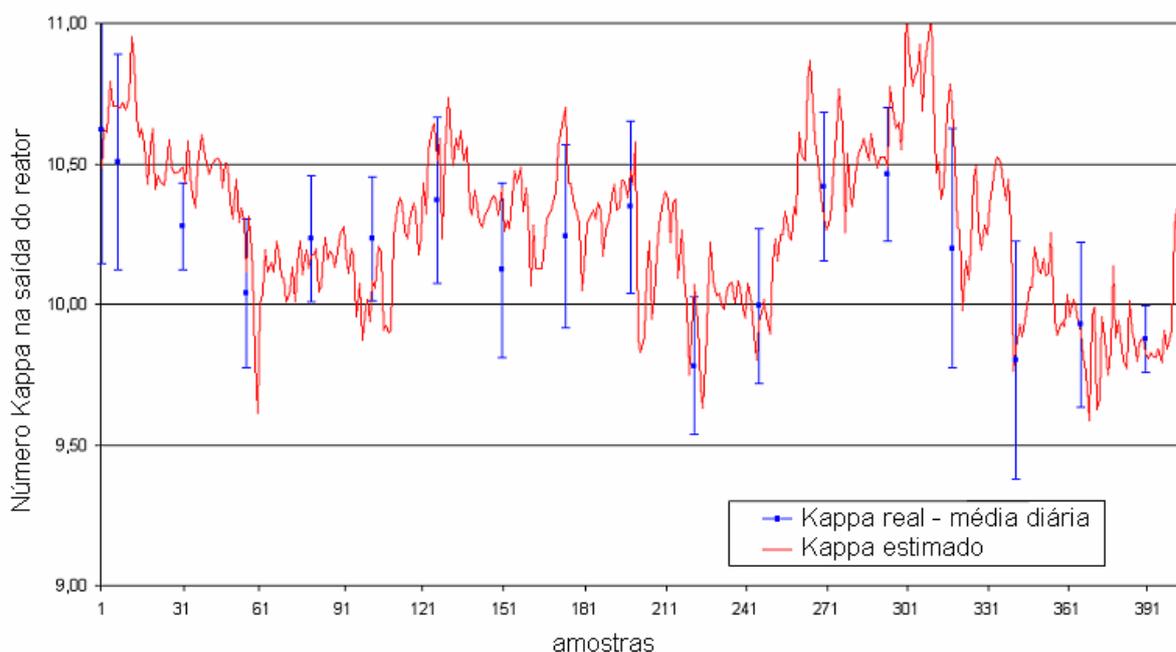


Figura 6.3 - Comparação entre os valores calculados pelo modelo e o desvio padrão da média diária do número *kappa* para o Modelo 1

A grande vantagem de se utilizar um modelo fenomenológico é poder analisar a evolução da reação ao longo do tempo. Foram gerados os perfis de consumo de álcali e de redução do número *kappa* ao longo do tempo de reação que devem ser confirmados com dados experimentais. Desta maneira é possível avaliar a quantidade de álcali que reage e, inclusive, avaliar a possibilidade da redução da sua concentração ou da vazão alimentada. Os perfis permitem também que análises de limites de qualidade sejam estabelecidos para a reação, ou seja, o intervalo de valores de concentração de hidróxido de sódio que permitam que o valor do número *kappa* de saída esteja dentro dos padrões estabelecidos para a produção.

A Figura 6.4 apresenta os perfis de decaimento do número *kappa* e de consumo de álcali, respectivamente, gerados com o Modelo 1.

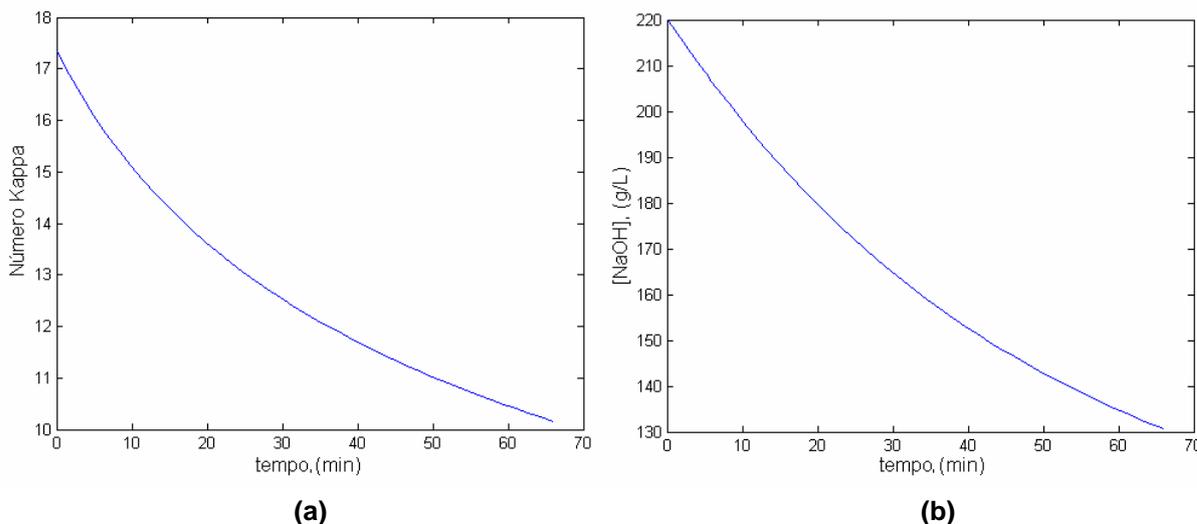


Figura 6.4 - (a) Perfil de decaimento do número *kappa* e; (b) perfil de consumo de álcali ao longo da reação de deslignificação indicado pelo Modelo 1

O perfil apresentado na Figura 6.4(b) mostra que o modelo de consumo de álcali de Violette (2003) condiz com o comportamento de consumo verificado na literatura, mesmo se tratando de uma equação linear. É preciso considerar, ainda, que o valor inicial da concentração de álcali é um valor médio e que não expressa, necessariamente, a realidade modelada. Não é possível afirmar, entretanto, que de posse do valor correto o Modelo 1 apresentaria efetivamente o melhor ajuste, mas dadas suas limitações o modelo análogo a Violette (2003) se mostrou eficiente.

6.1.2 Ajuste do Modelo 2

Para o Modelo 2, análogo ao de Agarwal et al (1999), o procedimento utilizado é essencialmente o mesmo. Foram utilizados os algoritmos heurísticos para otimizar os parâmetros e o algoritmo genético gerou os melhores parâmetros para o modelo.

$$-\frac{dK}{dt} = 6,2 \cdot \exp\left(-\frac{10^4}{RT}\right) \cdot K^{7,7} [OH^-]^{0,92} P_{O_2}^{0,53} \quad (6.3)$$

Foram testadas as taxas linear e não linear apresentadas na literatura inseridas no modelo diferencial de consumo de álcali. Os melhores resultados foram obtidos com a taxa não linear proposta por Myers e Edwards (1989), apresentada no Capítulo 3, seção 3.1.4. Portanto,

$$\frac{d[OH^-]}{dt} = \frac{F_{solução} ([OH^-]_{inicial} - [OH^-])}{V} - r_{[OH^-]} \quad (6.4)$$

Com:

$$r_{[OH^-]} = 1 - \left(\frac{K + \frac{K}{K_{inicial}} - 1}{13,64} \right) \quad (6.5)$$

onde K é o número *kappa* e $K_{inicial}$ é o número *kappa* inicial da polpa.

Na Figura 6.5 pode-se verificar que do Modelo 2 consegue identificar os valores reais da unidade da designificação, mas está, provavelmente, limitado pela falta de valores de concentração de álcali. Outro fator que pode afetar a predição é o alto valor da ordem da reação em relação ao número *kappa*. No modelo de Agarwal et al (1999) a ordem de reação em relação ao número *kappa* que é o mais do que o dobro que os demais modelos implementados e este parâmetro pode ser um fator determinante para o ajuste.

Para verificar esta suposta influência um novo modelo foi desenvolvido, sendo que o valor da ordem da reação em relação ao número *kappa* valor foi ajustado junto com os demais parâmetros (o fator de frequência de Arrhenius, a energia de ativação e o parâmetro cinético). A ordem da reação em relação ao número *kappa* foi variada no intervalo [3;8]. Este intervalo não apresenta problemas numéricos e está de acordo com os valores encontrados na literatura.

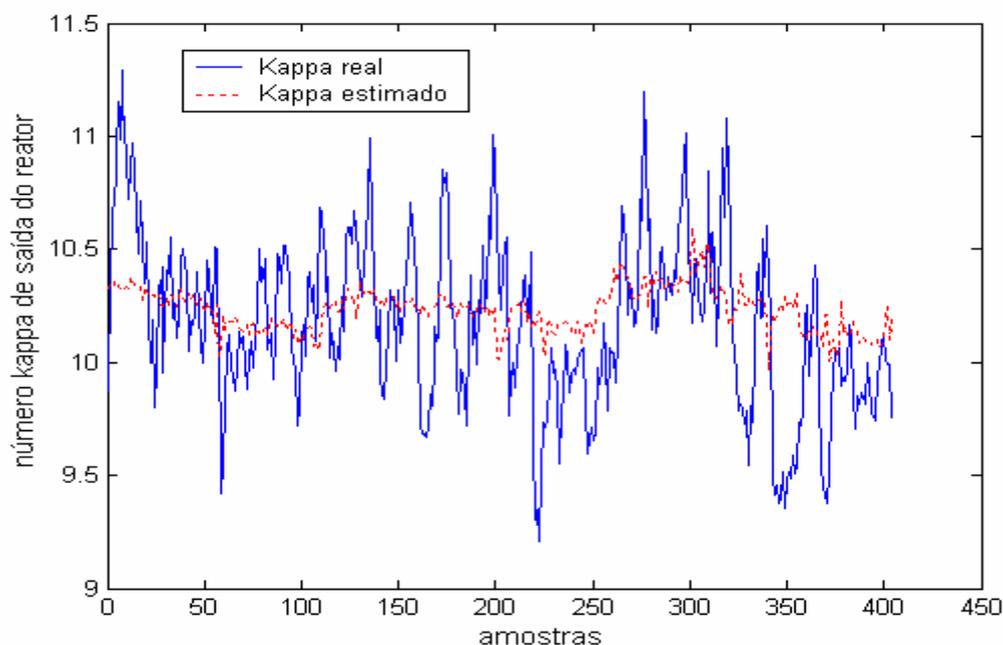


Figura 6.5 - Comparação entre os valores reais e estimados pelo modelo análogo a Agarwal et al (1999)

O novo ajuste dos parâmetros do Modelo 2, considerando o ajuste da ordem da reação em relação ao número *kappa* e utiliza a expressão linear da taxa de consumo de álcali, melhorou sua capacidade de representar o sistema. O resultado pode ser visto na Figura 6.6 e as equações do novo modelo são:

$$-\frac{dK}{dt} = 1,3 \cdot 10^{-2} \cdot \exp\left(-\frac{30,86 \cdot 10^3}{RT}\right) \cdot K^{3,37} [OH^-]^{0,92} P_{O_2}^{0,53}$$

$$\frac{d[OH^-]}{dt} = \frac{F_{solução} ([OH^-]_{inicial} - [OH^-])}{V} - r_{[OH^-]} \quad (6.6)$$

$$r_{[OH^-]} = 6 \cdot \Delta K + [OH^-]$$

O Modelo 2 ainda é limitado pela falta de valores mais corretos de concentração de álcali, mas os erros são menores que os apresentados pelo Modelo 1. Os erros absolutos calculados para o Modelo 2 estão distribuídos no intervalo $[-0,83; 0,77]$, o que implica em um erro relativo médio máximo de 8,3% (Figura 6.7).

Em termos estatísticos, a maior parte dos valores calculados pelo modelo está dentro do desvio padrão da média como mostra a Figura 6.8. Isto significa que é preciso ajustar melhor o modelo de forma a diminuir o erro calculado. Este ajuste, provavelmente, só será possível se valores confiáveis de concentração de hidróxido de sódio na entrada do reator de designificação estiverem disponíveis.

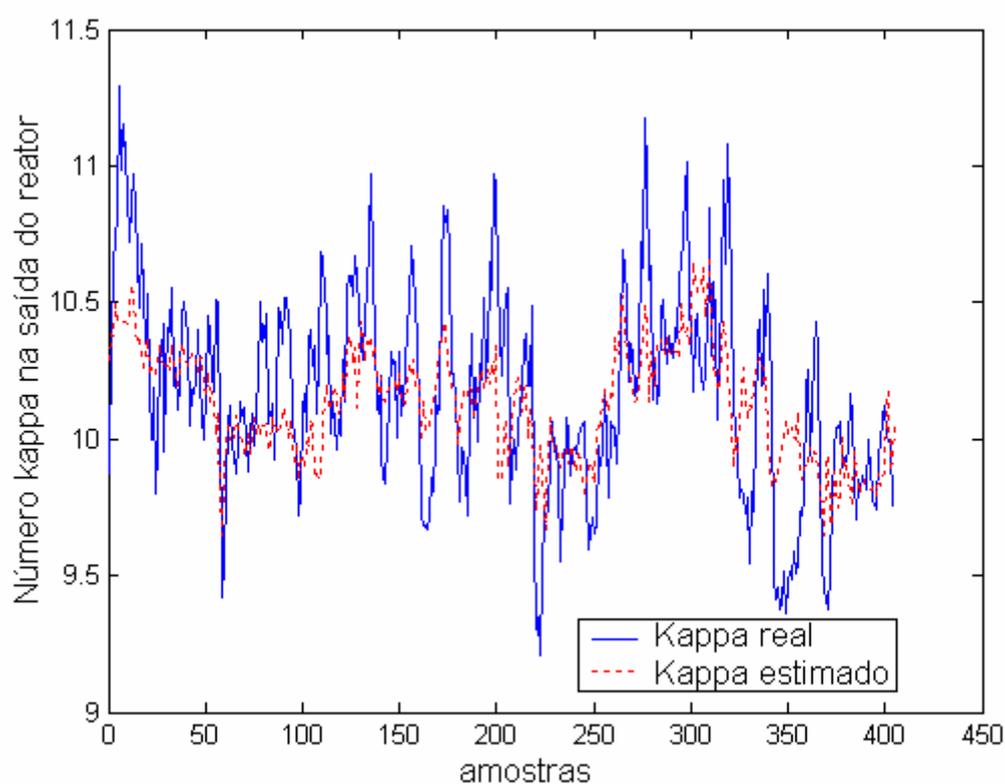
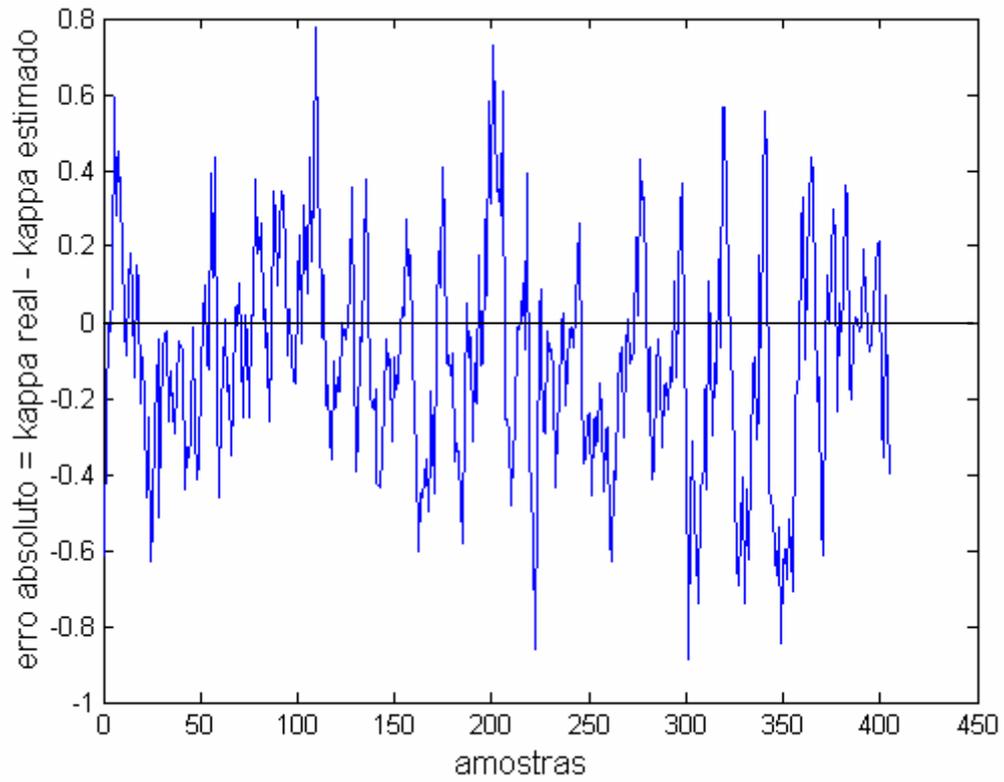
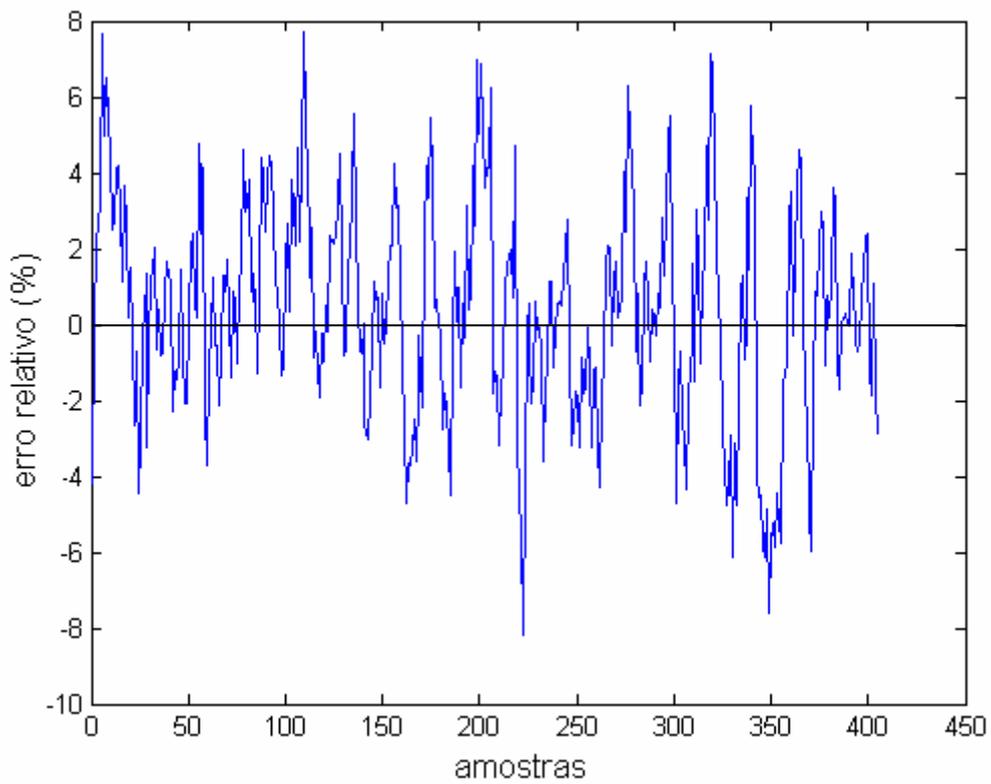


Figura 6.6 - Comparação entre os valores reais e estimados pelo Modelo 2 com valor da ordem em relação ao número *kappa* estimado pelo AG



(a)



(b)

Figura 6.7 - (a) Erro absoluto e; (b) erro relativo médio calculados para o Modelo 2

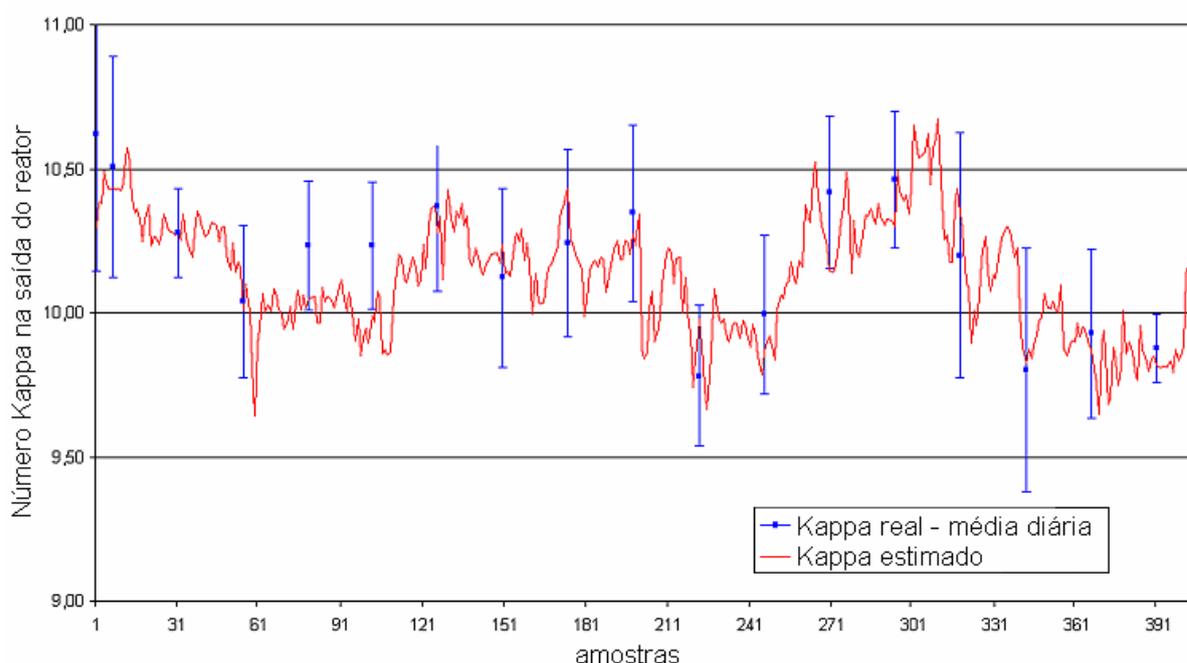


Figura 6.8 - Comparação entre os valores calculados pelo modelo e o desvio padrão da média diária do número κ para o Modelo 2

Os perfis calculados pelo Modelo 2 de decaimento do número kappa e do consumo de álcali durante a reação são apresentados nas Figuras 6.9 (a) e (b), respectivamente. A estimativa do consumo de hidróxido de sódio deve, entretanto, deve ser confirmada por análise laboratorial uma vez que o valor final da concentração não é conhecido. Contudo, o perfil gerado é condizente com perfis clássicos de consumos de reagente, o que leva a crer que, de posse dos valores reais do consumo na saída do reator, um novo ajuste pode ser efetuado e a curva será, portanto, uma representação aproximada do consumo de álcali durante a reação de deslignificação.

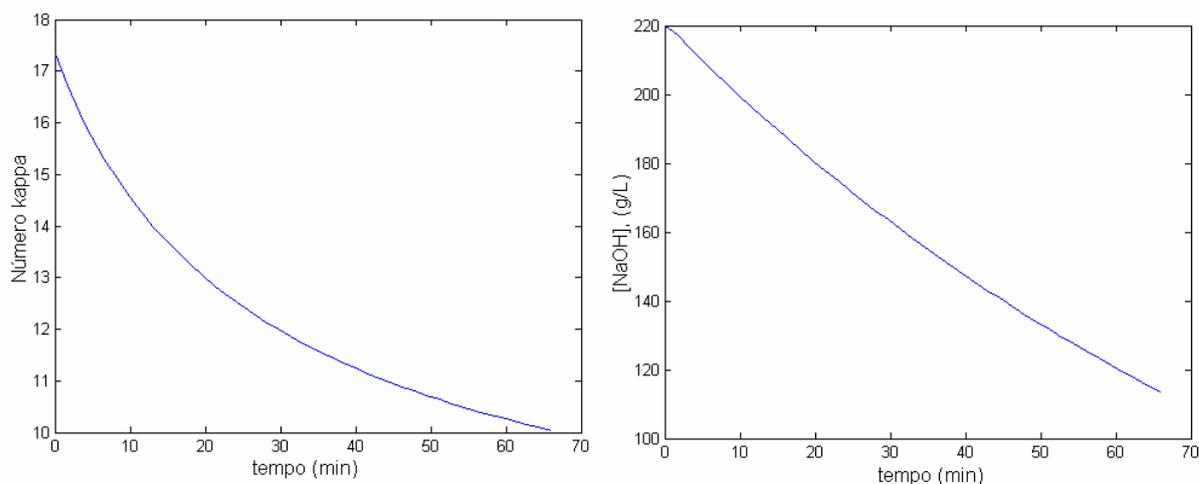


Figura 6.9 - (a) Perfil de decaimento do número *kappa* e; (b) perfil de consumo de álcali ao longo do tempo de reação gerado pelo Modelo 2

6.1.3 Ajuste do Modelo 3

O modelo de Hsu e Hsieh ((d)1988) considera que a reação de deslignificação se dá em duas etapas, a primeira etapa é rápida e se processa em aproximadamente dois minutos. A segunda etapa, lenta, utiliza em média uma hora de reação, é nesta etapa que a maior parte da lignina é dissolvida e retirada da polpa.

O Modelo 3, baseado em Hsu e Hsieh ((d)1988), Modelo 3, os parâmetros obtidos por AG também descreveram melhor o processo industrial. Para o modelo de consumo de álcali durante a reação foi utilizada a expressão linear inserida no balanço diferencial apresentado no Capítulo 5. As equações 6.7 e 6.8 apresentam o Modelo 3 com os parâmetros ajustados para o caso de estudo:

$$-\frac{dK}{dt} = k_1 \cdot K^{3,07} [OH^-]^{0,78} P_{O_2}^{0,35} [u(t) - u(t-2)] + k_2 \cdot K^{3,07} [OH^-]^{0,70} P_{O_2}^{0,74} u(t-2) \quad (6.7)$$

Com:

$$k_1 = 10 \exp\left(-\frac{10^4}{RT}\right) \quad (6.8)$$

$$k_2 = 16 \exp\left(-\frac{6,5 \cdot 10^4}{RT}\right)$$

O primeiro termo da equação 6.7 representa a primeira etapa da reação de deslignificação (para $t < 2 \text{ min}$); e o segundo termo da mesma equação expressa a fase lenta da reação, para $t \geq 2 \text{ min}$.

Para o consumo de álcali, as taxas de consumo resultantes são:

$$r_{1[OH^-]} = 5 \cdot \Delta K + [OH^-] \quad (6.9)$$

$$r_{2[OH^-]} = 8 \cdot \Delta K + [OH^-]$$

onde os índices 1 e 2 referem-se à primeira e segunda etapas de reação, respectivamente. As equações para o modelo de consumo de hidróxido de sódio são:

$$\frac{d_1[OH^-]}{dt} = \frac{F_{\text{solução}} ([OH^-]_{\text{inicial}} - [OH^-])}{V} - r_{1[OH^-]} \quad (6.10)$$

$$\frac{d_2[OH^-]}{dt} = \frac{F_{\text{solução}} ([OH^-]_{\text{inicial}} - [OH^-])}{V} r_{2[OH^-]}$$

Os resultados obtidos com o Modelo 3 são apresentados na Figura 6.10. O ajuste obtido com este modelo é o melhor dentre os três modelos estudados, o que pode ser um indício de que a reação efetivamente se processa em duas etapas e que existem categorias de lignina que reagem rápida e lentamente. O modelo em duas etapas pode ser implementado como o proposto por Hsu e Hsieh ((d)1988) e por Iribarne e Schroeder (1997), utilizando tempo fixo para a primeira etapa da

reação, ou utilizando o modelo de categorias de lignina apresentado por Olm e Teder (1981) e Myers e Edwards (1989).

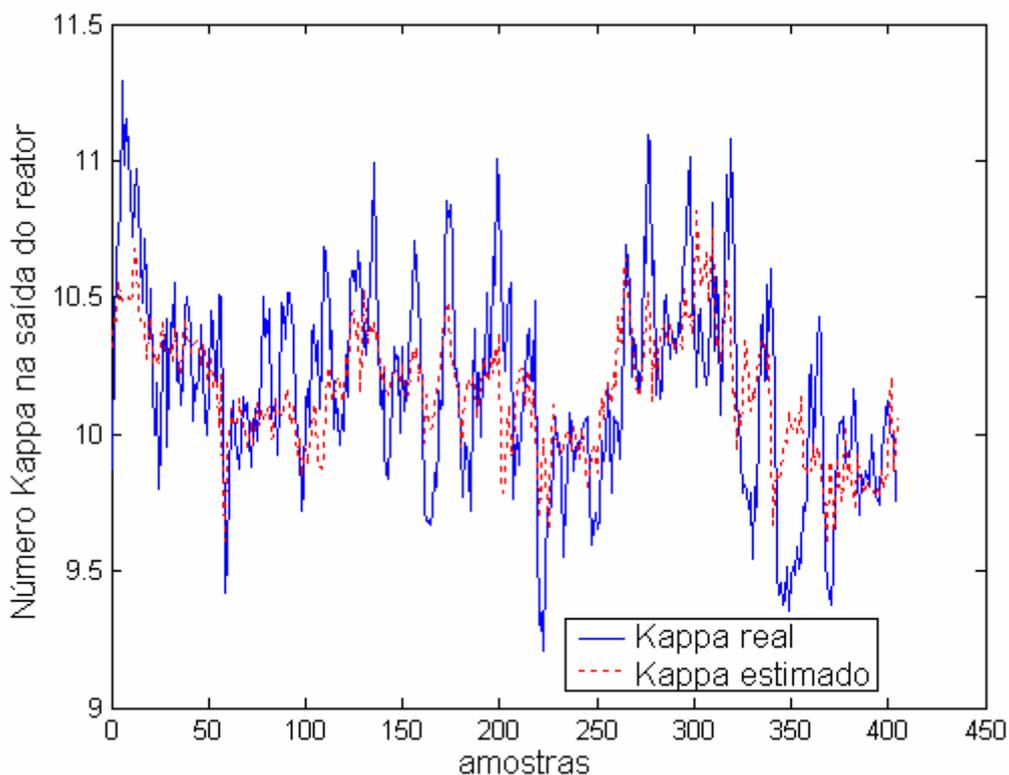
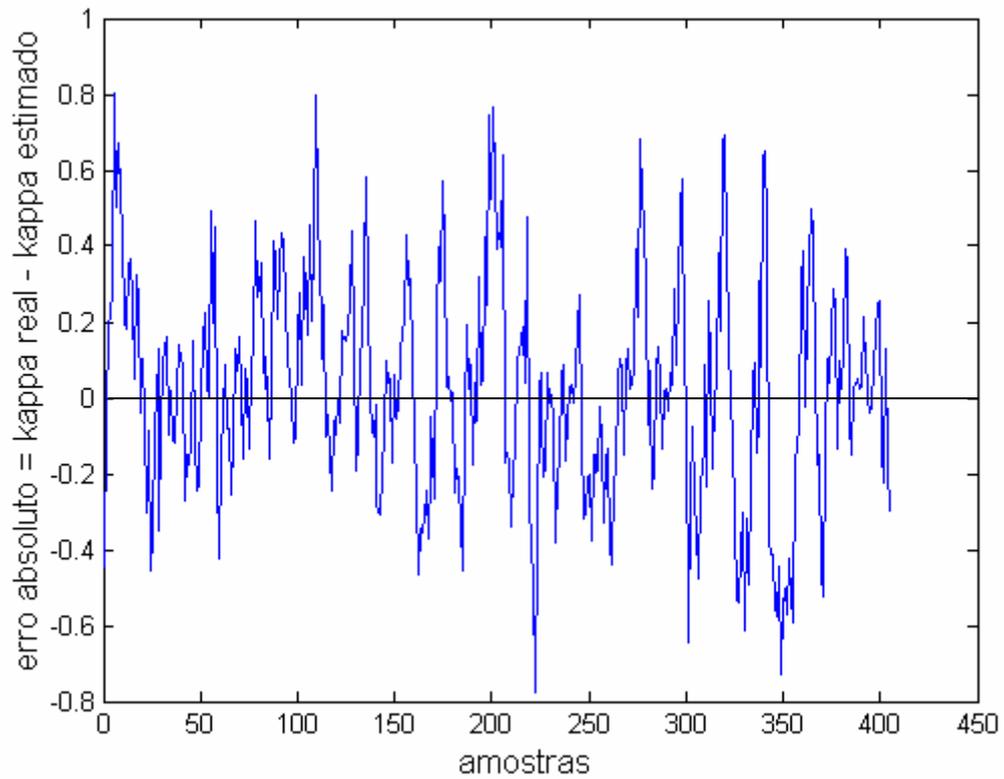
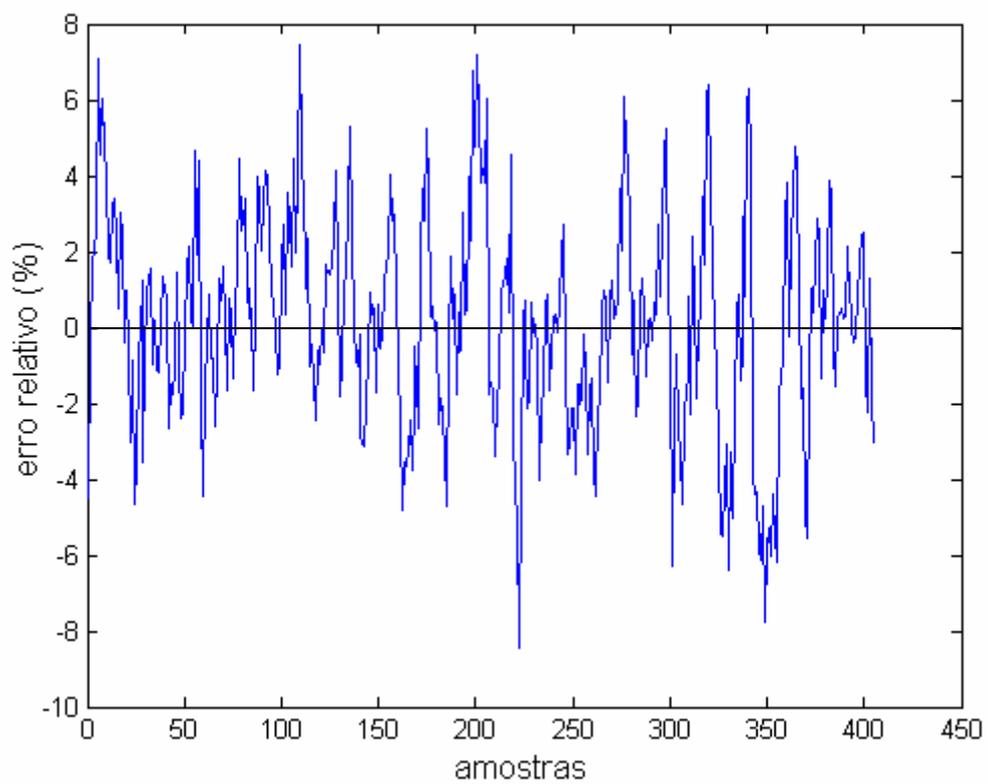


Figura 6.10 - Comparação entre os valores reais e estimados pelo Modelo 3

Os erros absoluto e relativo calculados para o Modelo 3 são apresentados na Figura 6.11. Os erros absolutos estão distribuídos no intervalo $[-0,84; 0,74]$, o que significa um erro relativo médio máximo de 8,4%. Este valor pode ser, provavelmente, diminuído se valores confiáveis de concentração de álcali estiverem disponíveis para a modelagem. Estatisticamente todos os valores calculados estão inseridos dentro dos limites do desvio padrão da média dos números *kappa* da polpa, Figura 6.12.



(a)



(b)

Figura 6.11 - (a) Erro absoluto e; (b) erro relativo médio calculados para o Modelo 3

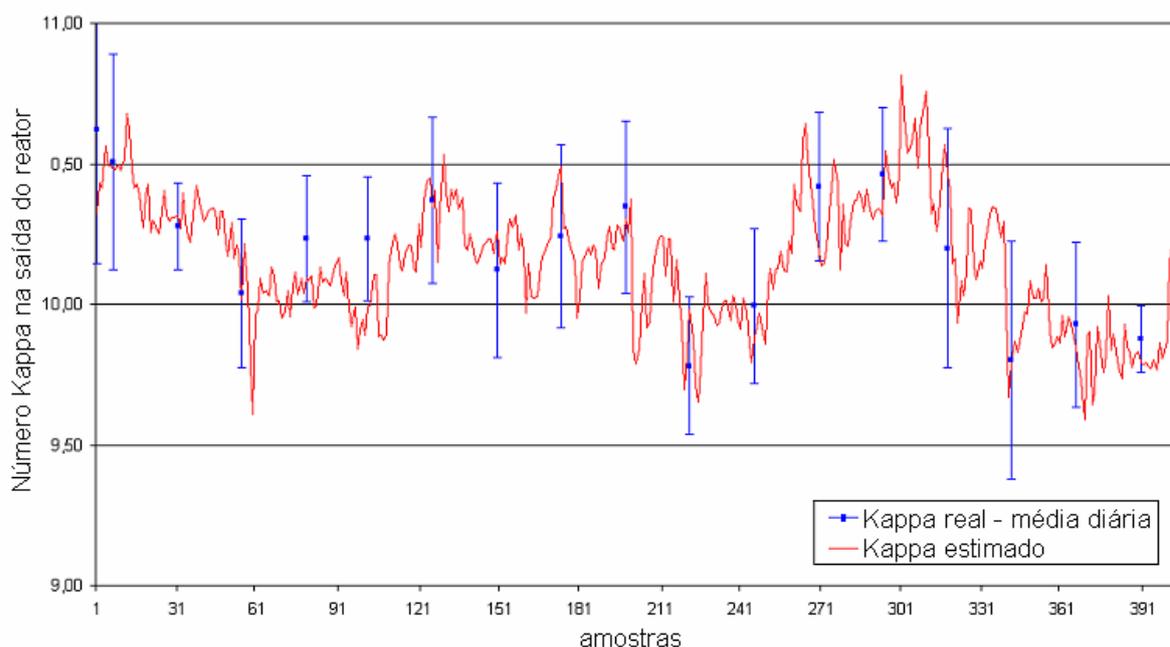


Figura 6.12 - Comparação entre os valores calculados pelo modelo e o desvio padrão da média diária do número *kappa* para o Modelo 3

Os perfis de decaimento de número *kappa* e de consumo de álcali ao longo da reação, Figura 6.13 a e b estão coerentes com os perfis obtidos nos outros modelos. Não é possível, entretanto, afirmar para qualquer dos três modelos que o perfil de consumo de hidróxido de sódio corresponde à realidade uma vez que não foi possível proceder a análises de concentração de álcali na saída do reator.

A etapa rápida da reação é evidenciada nas Figuras 6.13. Durante a etapa rápida, o número *kappa* é reduzido em cerca de 15%. O modelo de Hsu e Hsieh ((d)1988) apresenta uma redução média de 30% e o modelo de Iribarne e Schroeder (1997), que utiliza um tempo fixo de 7 minutos, para a primeira etapa de reação e obtém redução de 50%.

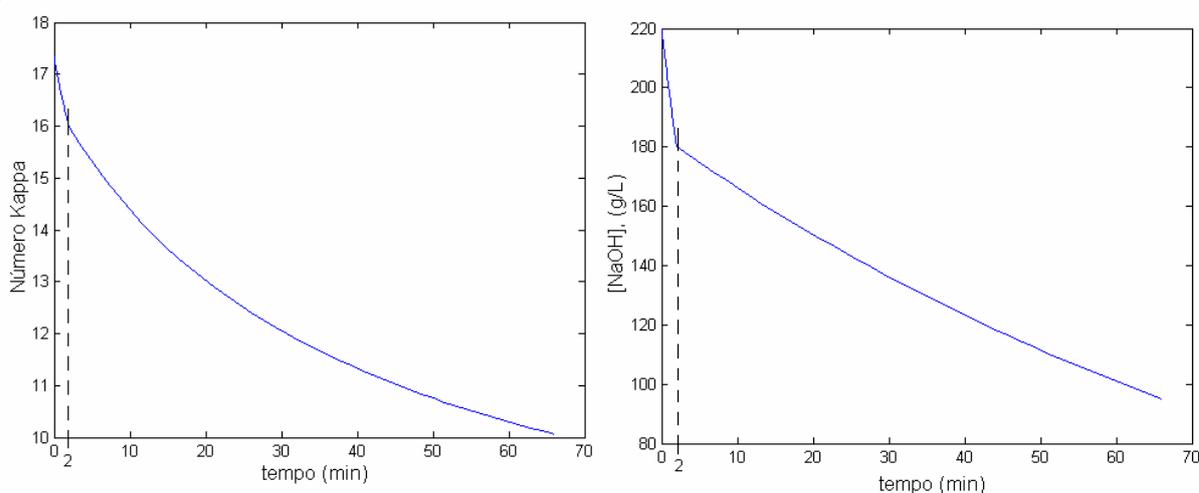


Figura 6.13 - (a) Perfil de decaimento do número *kappa* e; (b) perfil de consumo de álcali ao longo da reação de deslignificação para o Modelo 3

6.1.4 Avaliação do Desempenho dos Métodos de Estimação e dos Modelos Fenomenológicos

Os métodos empregados para encontrar os parâmetros para os modelos, o Algoritmo Genético e o de Recozimento Simulado, apresentaram resultados distintos durante as simulações. Ambos ofereceram resultados aplicáveis aos modelos, mas o AG forneceu, em todos os três casos implementados, o melhor conjunto de parâmetros para a modelagem da deslignificação de celulose.

O Recozimento Simulado apresentou, além dos resultados compatíveis, diversos conjuntos de parâmetros que geraram modelos ruins. Alguns conjuntos foram aproveitados para comparações com o AG, mas os parâmetros finais foram, todos, estimados pelo algoritmo genético. O AG, em contrapartida, apresentou a cada simulação resultados tecnicamente consistentes, com menor erro e aplicáveis aos modelos para representar a deslignificação. Sua implementação e adaptação

são bastante simples e seu uso é recomendado para problemas similares ao utilizado neste trabalho.

Nenhum dos dois algoritmos apresentou limitações numéricas ou demandou demasiado tempo computacional. A escolha entre um ou outro é critério do programador, considerando, sempre, o problema onde se pretende utilizar o algoritmo.

No que diz respeito aos modelos fenomenológicos desenvolvidos, todos os três apresentam resultados semelhantes, conforme o esperado. Considerando que o processo, e conseqüentemente, os dados industriais são os mesmos, se bem ajustados os modelos devem fornecer resultados semelhantes, o que é confirmado pela faixa de erros em que se encontram os resultados de todos eles.

Uma vez que o valor da concentração de álcali não é real, mas uma média aproximada, os modelos podem ser melhorados e os erros diminuídos, tornando os resultados mais confiáveis para simulações na fábrica se medidas confiáveis da concentração de álcali estiverem disponíveis.

Os Modelos 1 e 3, também limitados pela falta de valores mais precisos de concentração de álcali, acompanharam melhor a tendência dos dados reais e é provável que a melhoria acrescida pela inserção dos valores reais seja perceptível, conferindo maior confiabilidade a estes modelos.

Quanto à implementação dos modelos matemáticos, o Modelo 3 é o mais trabalhoso, uma vez que envolve quatro equações diferenciais que devem ser resolvidas simultaneamente duas a duas. Como foram analisados 811 padrões, este modelo demanda mais tempo de processamento. Outro fator importante a ser considerado é o número de parâmetros para serem ajustados: como a reação se dá

em duas etapas, a quantidade de parâmetros estimados dobra, aumentando o esforço dos algoritmos de otimização utilizados.

Em contrapartida, o Modelo 1 é bastante simples, composto de uma equação diferencial e uma equação algébrica, já que para este modelo não foi necessário implementar o modelo diferencial de consumo de álcali. Entretanto, uma ressalva deve ser feita quando se discute este modelo: a influência do oxigênio não é considerada na expressão cinética.

O mecanismo da deslignificação é baseado nos fenômenos de transferência de massa, incluindo a transferência do oxigênio da fase gás para a fase líquida. Violette (2003) investigou a possibilidade dos radicais de oxigênio criarem uma barreira entre a celulose e os íons hidroxila, dificultando o processo de deslignificação. Para tanto, trabalhou com excesso (mais que o dobro que o recomendado na literatura) em seus experimentos e sua influência foi, desta forma, desconsiderada na cinética da reação tornando possível verificar se os radicais formavam ou não a barreira em torno da fibra.

Uma vez que este modelo análogo a Violette ajustou os dados com a mesma precisão que os demais, é possível que no processo estudado o oxigênio também esteja em excesso. No entanto, análises laboratoriais são necessárias para a confirmação deste excesso e outras aplicações deste modelo exigem o cuidado da verificação da influência do oxigênio na reação.

6.2 IDENTIFICAÇÃO DO SISTEMA COM MODELO LINEAR ARX

O intuito da identificação do processo com o modelo linear ARX é a comparação entre as técnicas de modelagem fenomenológica e a identificação de

sistemas com modelos caixa-preta e, até mesmo, oferecer uma opção para o modelo fenomenológico proposto.

O modelo linear identificado, Modelo 4, possui oito variáveis de entrada e estima o número $kappa$ da saída do reator de designação. As variáveis utilizadas na entrada do modelo são: número $kappa$ na entrada do reator, pressão do oxigênio, vazão de oxigênio e de álcali, temperatura de reação, tempo de residência do reator, produção e consistência da polpa. A Figura 6.14 apresenta a comparação entre os resultados obtidos com o modelo ARX e os valores reais do número $kappa$ de saída do reator.

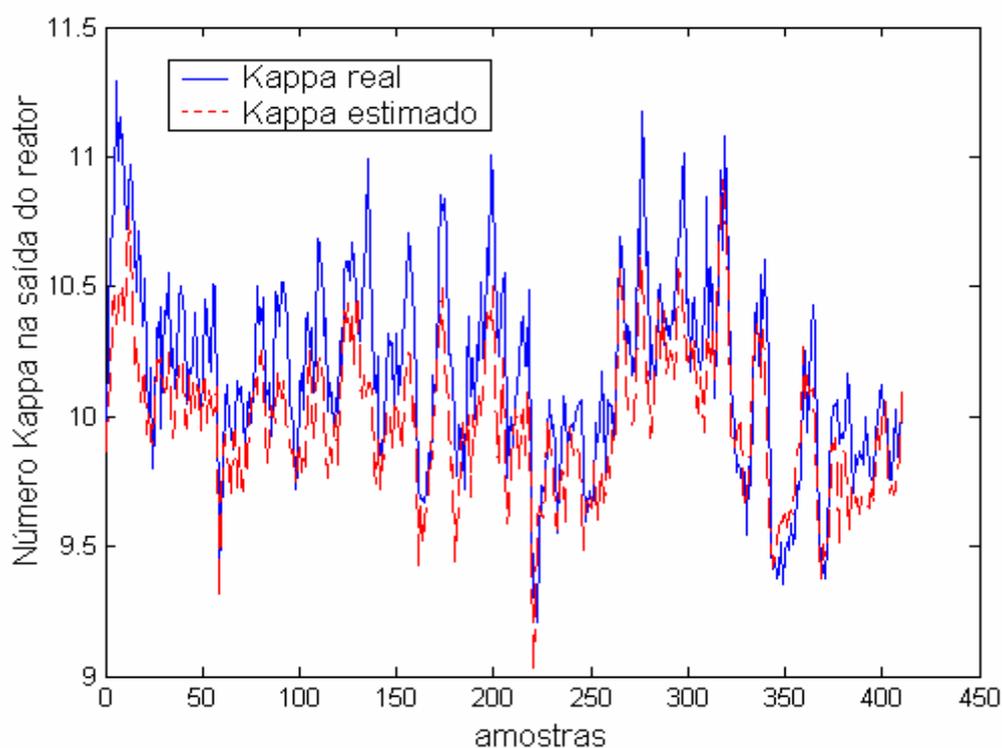
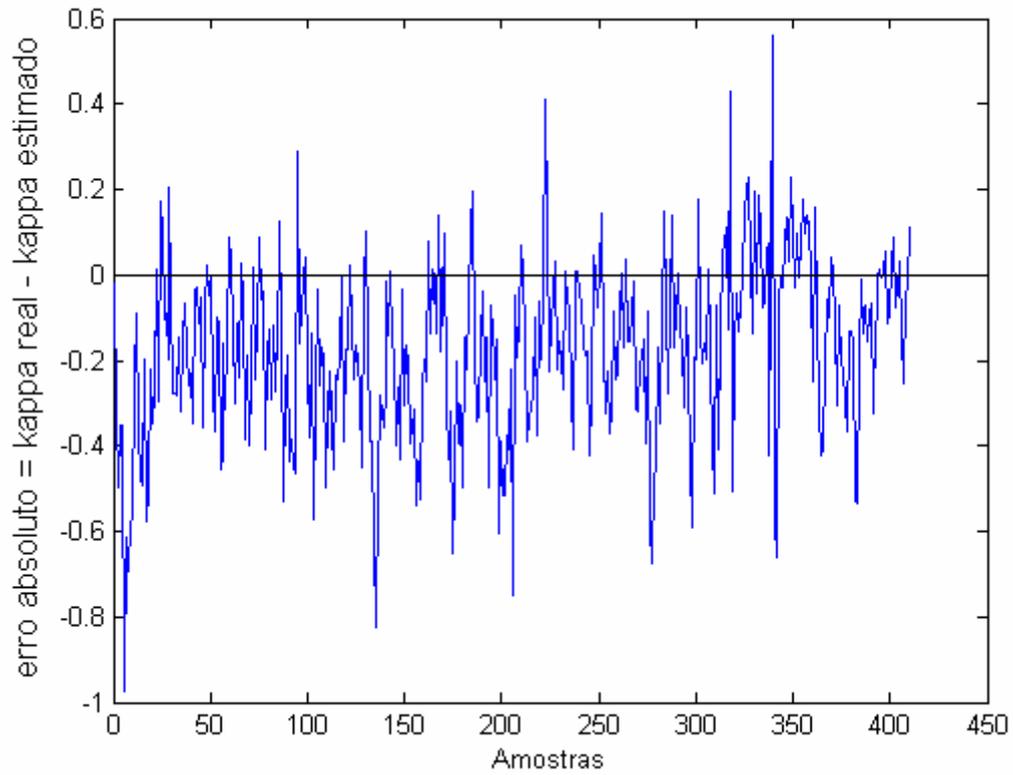
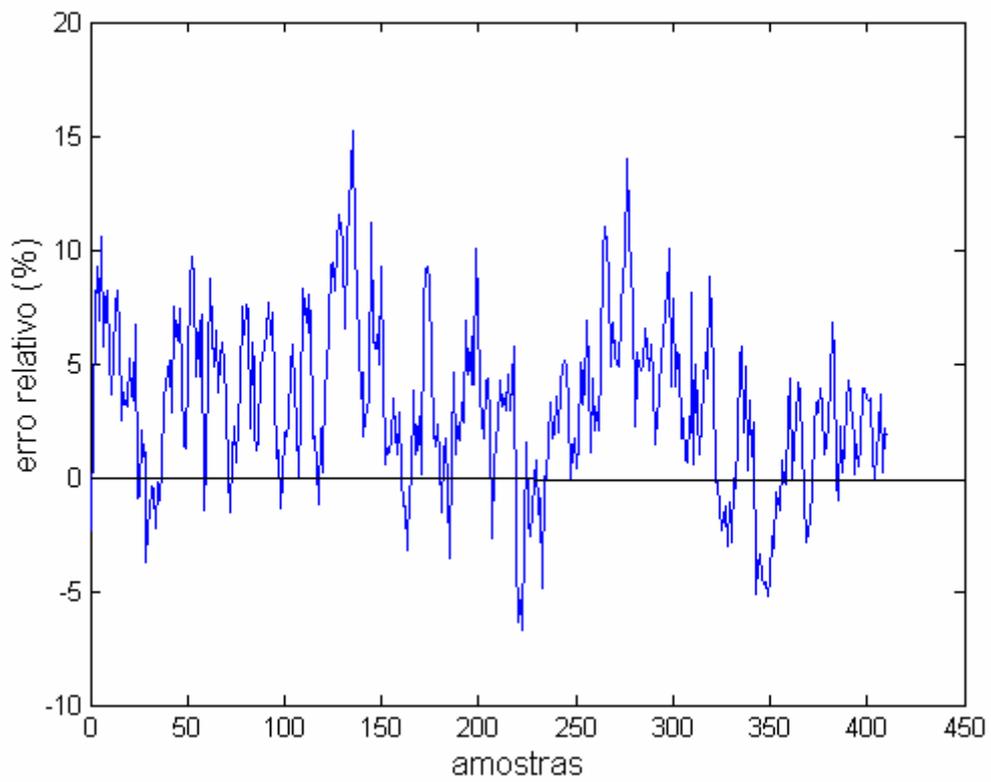


Figura 6.14 - Comparação entre os valores reais e estimados pelo Modelo 4



(a)



(b)

Figura 6.15 - (a) Erro absoluto e; (b) erro relativo médio calculados para o Modelo 4

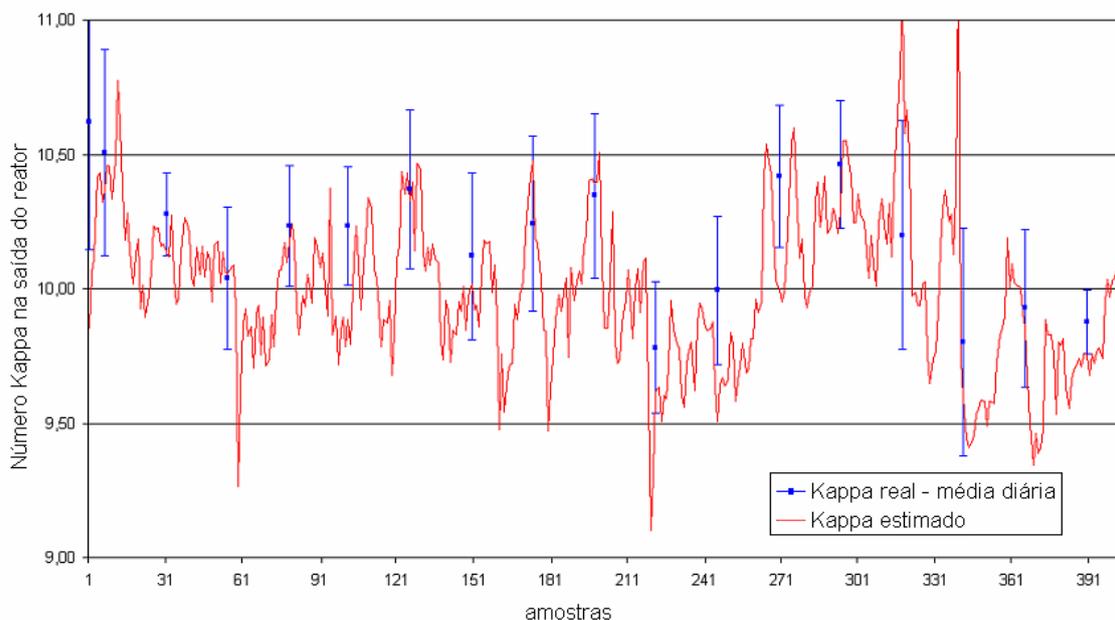


Figura 6.16 - Comparação entre os valores calculados pelo modelo e o desvio padrão da média diária do número *kappa* para o Modelo 4

O Modelo 4 apresenta erros absolutos (Figura 6.15(a)), diferença entre os valores reais e o calculado pelo modelo, estão distribuídos no intervalo $[-0,53; 0,86]$ ou um erro relativo médio (Figura 6.15(b)) máximo de 8,6%. A Figura 6.16 apresenta uma comparação entre os valores médios diários e os valores estimados pelo Modelo 4.

Uma outra configuração foi identificada utilizando o modelo linear ARX utilizando cinco variáveis na entrada da rede. Àquelas variáveis que apresentavam menor desvio padrão da média foram desconsideradas neste segundo modelo, Modelo 5, sendo utilizadas número *kappa* na entrada do reator, vazão de oxigênio e de álcali, temperatura de reação e produção. Os resultados obtidos com o Modelo 5 estão apresentados na Figura 6.17.

Os erros obtidos com esta configuração não são significativamente diferentes dos erros calculados para o Modelo 4. Desta forma a opção por um dos dois modelos deve ser feita pela quantidade de parâmetros e variáveis envolvidas. Para o Modelo 5 os erros absolutos (Figura 6.18(a)) estão inseridos no intervalo $[0,43;$

0,83], o que implica em um erro relativo médio (Figura 6.18(b)) máximo de 8,3%. Em termos estatísticos, Figura 6.19, o Modelo 5, abrange os valores do desvio padrão da média diária do número *kappa*.

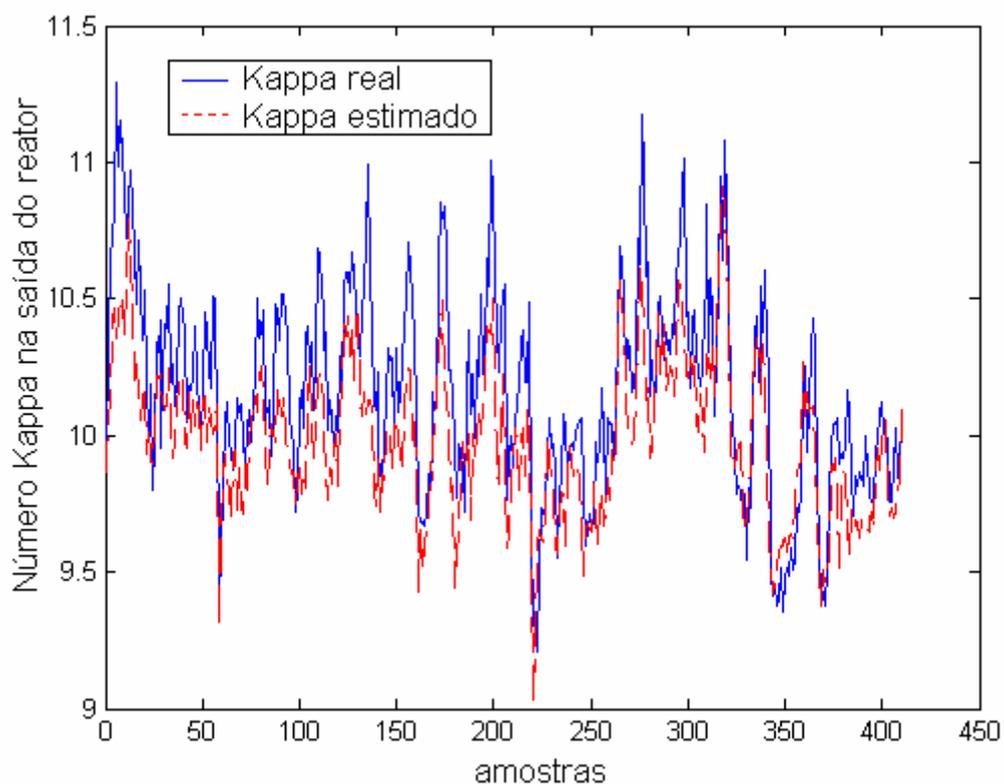
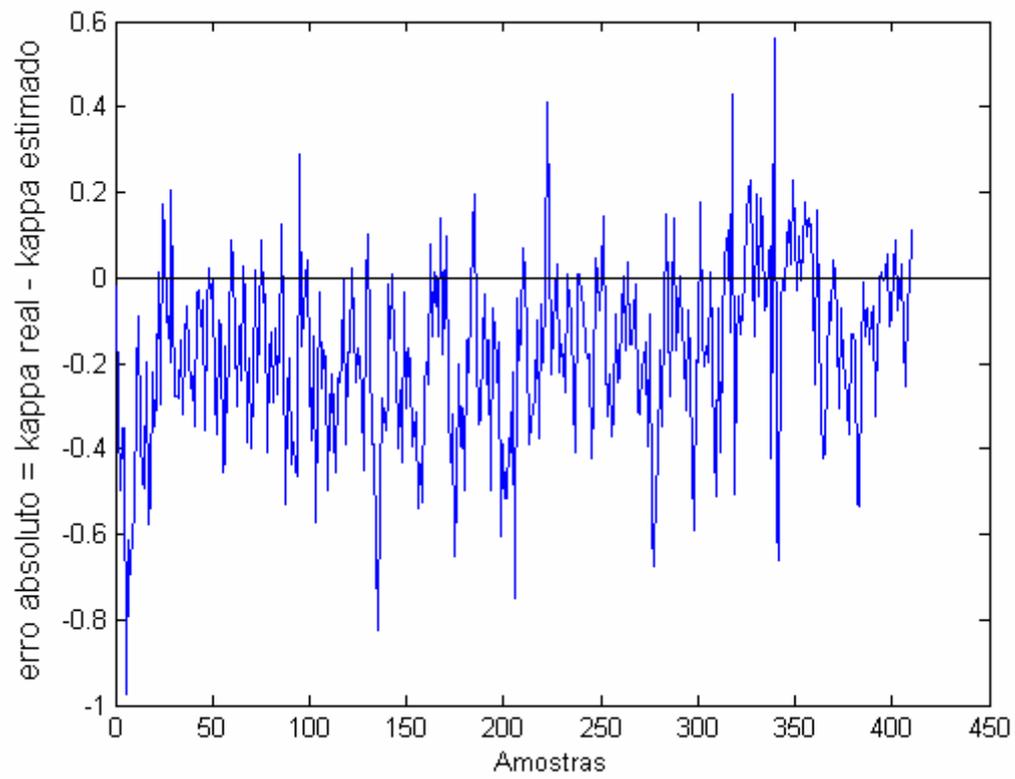
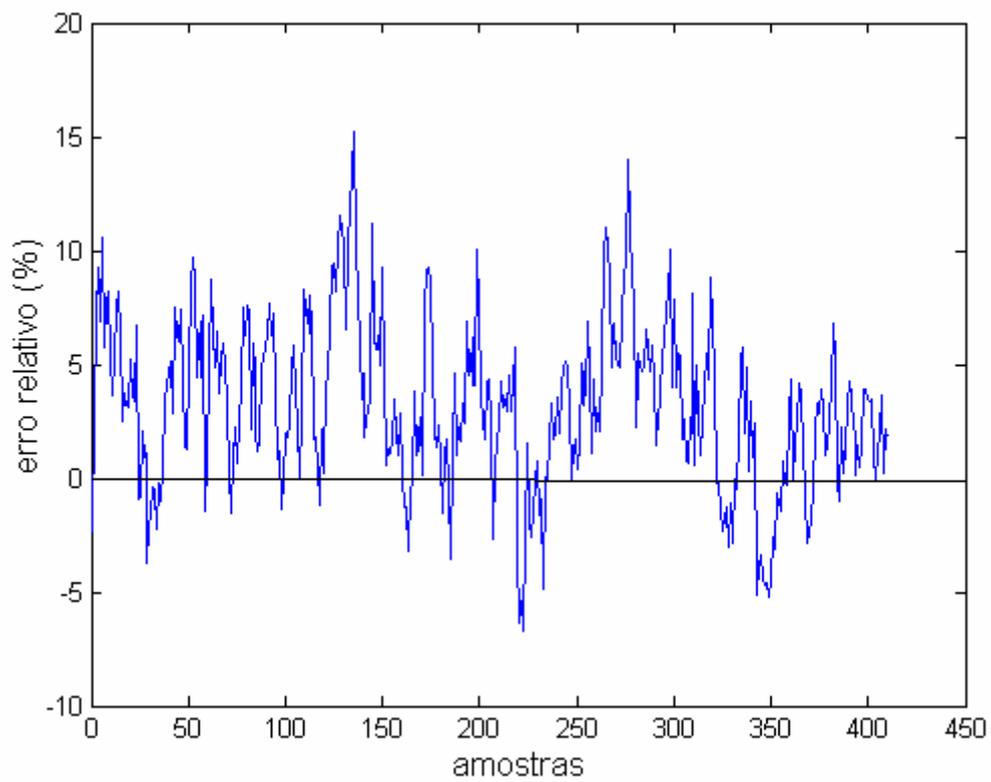


Figura 6.17 - Comparação entre os valores reais e estimados pelo Modelo 5



(a)



(b)

Figura 6.18 - (a) Erro absoluto e; (b) erro relativo médio calculados para o Modelo 4

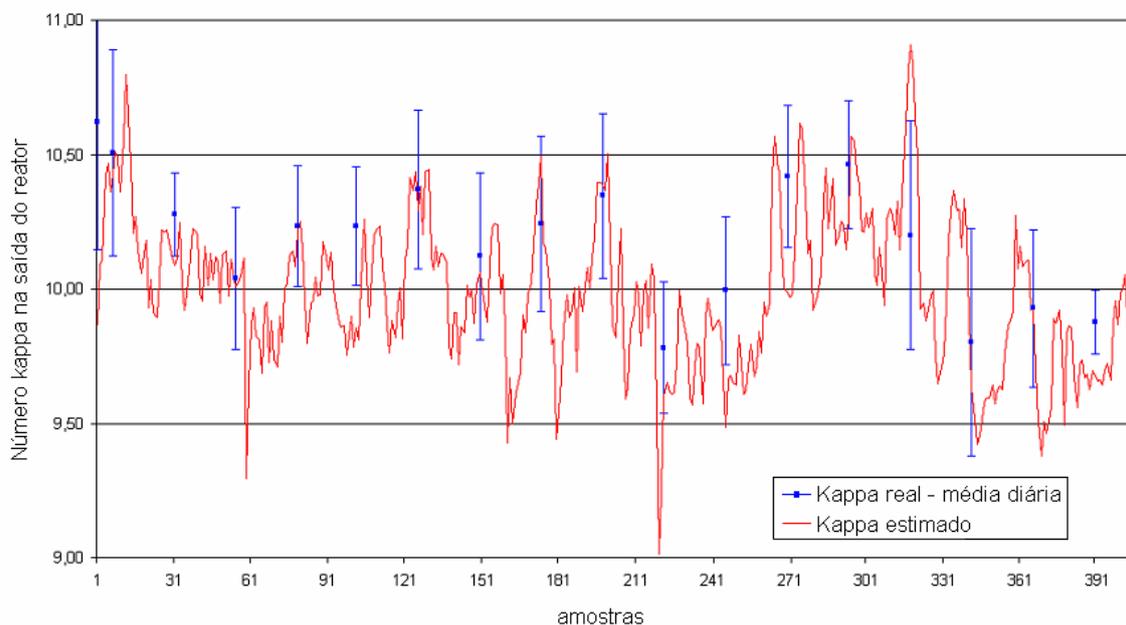


Figura 6.19 - Comparação entre os valores calculados pelo modelo e o desvio padrão da média diária do número *kappa* para o Modelo 4

6.3 MODELAGEM NEURAL

Uma vez que o processo foi identificado com um modelo linear, a identificação com um modelo não-linear não é imprescindível. Entretanto, como esta dissertação possui caráter investigativo, modelos neurais foram propostos para fins comparativos.

Um dos fatores utilizados para justificar o uso de um modelo não-linear na identificação de um sistema é o valor do EQM do modelo linear do processo. No caso deste trabalho, o EQM de ambos os Modelos 4 e 5 é 4,1%, que embora seja um valor alto para uma previsão um passo adiante, forneceu resultados da mesma ordem de grandeza dos modelos fenomenológicos.

Analisando-se o erro médio quadrático durante o processo de validação para diversas topologias de rede, foram escolhidas as redes que apresentaram os menores valores de EQM. A Figura 6.20 apresenta a análise do EQM para as redes com 5 e 8 variáveis na camada de entrada da rede.

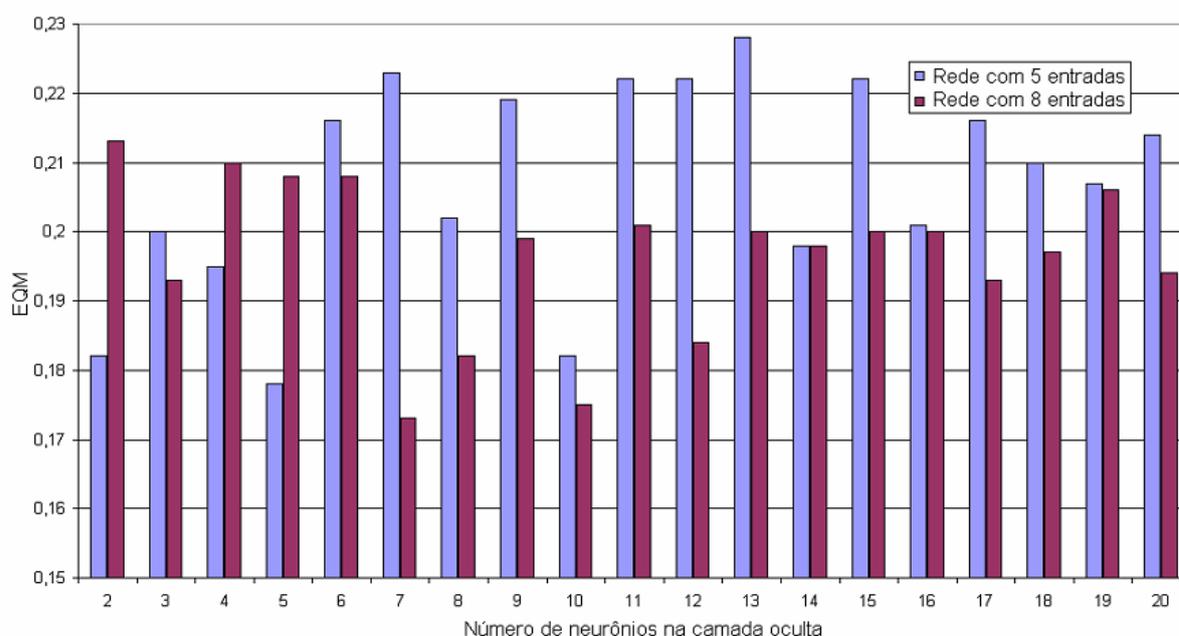


Figura 6.20 - Análise do EQM para a escolha das melhores estruturas neurais.

Os valores dos EQMs para as redes neurais, embora menores, são da mesma ordem de grandeza que os erros dos modelos lineares. Esta diminuição era esperada uma vez que as redes neurais são modelos não lineares e, portanto, mais complexos que os modelos lineares.

Para o desenvolvimento do modelo neural uma rede neural do tipo MLP foi aplicada. Foram utilizadas diversas estruturas de rede, com número de neurônios na camada oculta variando entre [2; 20]. Conforme explicado no Capítulo 5, seção 5.4, a escolha do número de neurônios é empírica, dependendo do ajuste obtido com os dados de validação.

Antes da aplicação dos dados na rede neural, os padrões foram normalizados de acordo com o procedimento descrito no Capítulo 5. Para o treinamento da rede

neural, os 405 padrões (conjuntos de entrada e saída) foram utilizados. Estes padrões foram filtrados para estarem todos dentro do intervalo de confiança imposto para o problema e dentro das especificações do processo.

Para a rede com oito variáveis na camada de entrada (número *kappa* na entrada do reator, pressão do oxigênio, vazão de oxigênio e de álcali, temperatura de reação, tempo de residência do reator, produção e consistência da polpa) uma rede neural 7 neurônios na camada intermediária, Modelo 6, foi capaz de fornecer bons resultados para o caso considerado neste trabalho.

Para a outra configuração (Modelo 7), com cinco variáveis na camada de entrada (número *kappa* na entrada do reator, vazão de oxigênio e de álcali, temperatura de reação e produção), a rede com cinco neurônios na camada oculta foi a que apresentou o menor EQM e, portanto, é a rede escolhida para representar a planta de deslignificação.

Para analisar a qualidade do modelo, gráficos de saída da rede (estimativa) *versus* o valor real do número *kappa* foram utilizados. As Figuras 6.21 e 6.22 mostram o resultado do Modelo 6 comparado com valor do número *kappa* real da planta, e o valor dos erros calculados para o modelo neural. Em termos estatísticos, os resultados estão inseridos no desvio padrão da média, Figura 6.23. Os erros absolutos – Figura 6.22 (diferença entre o valor real do número *kappa* e do valor estimado pelo modelo) estão no intervalo $[-0,50; 0,63]$, o que significa que o valor do erro relativo médio não passa de 6,3%, compatível com os valores obtidos com o modelo fenomenológico. As figuras referem-se aos dados de validação do modelo.

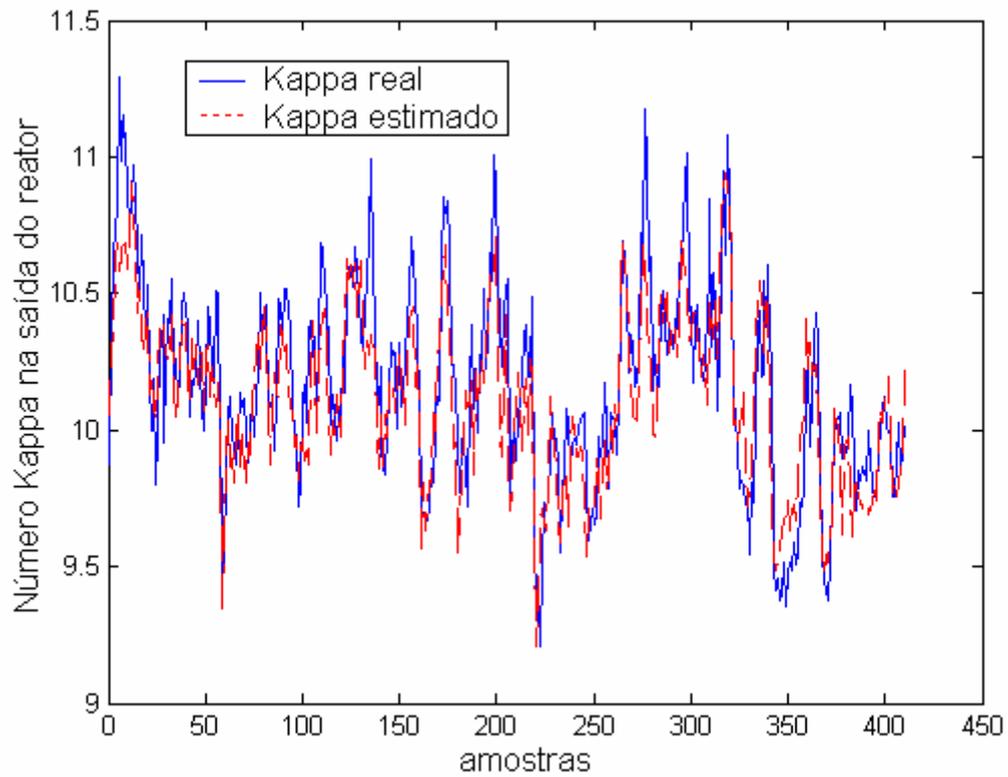


Figura 6.21 - Comparação entre os valores reais e estimados pelo Modelo 6

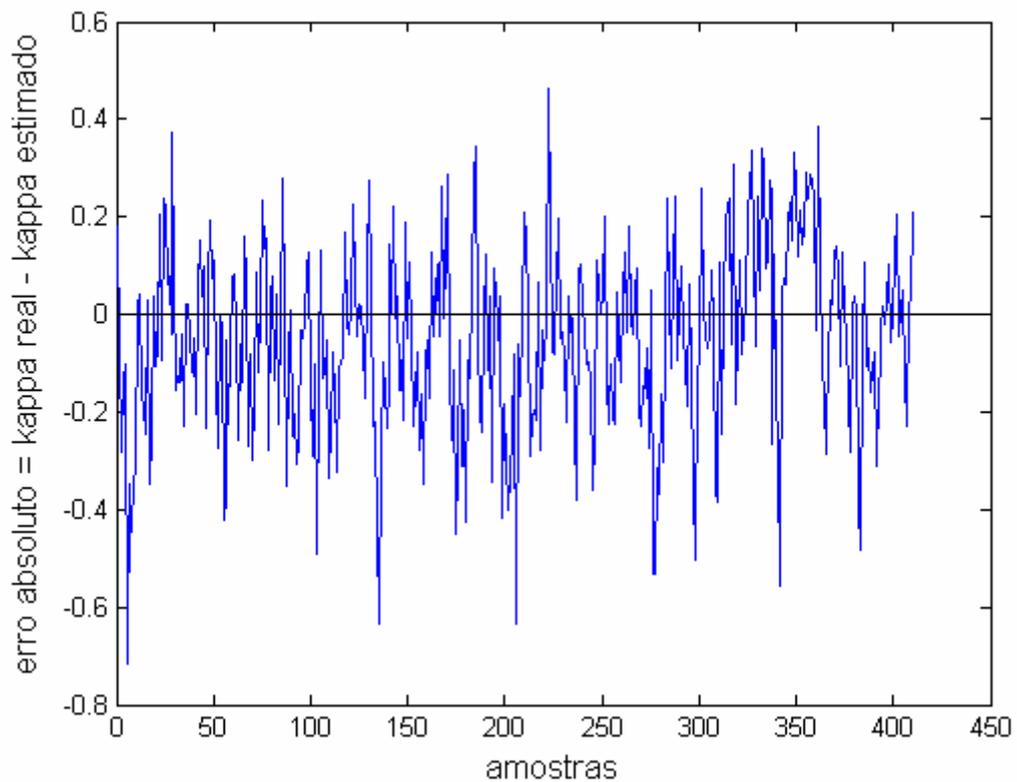


Figura 6.22 - Erro absoluto calculado para o Modelo 6

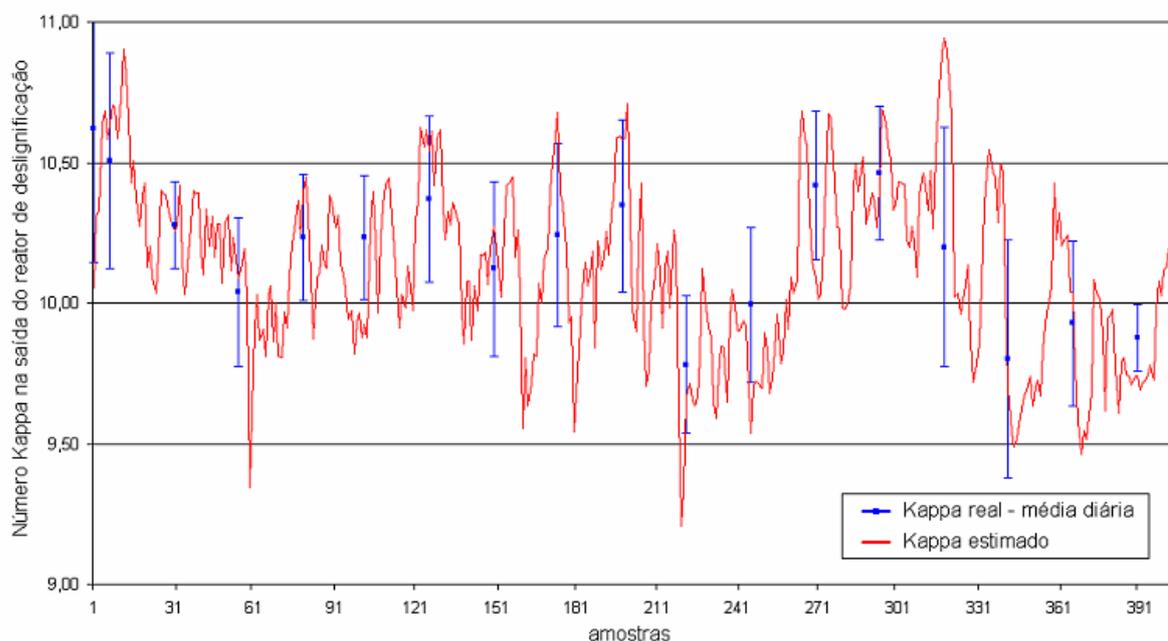


Figura 6.23 - Comparação entre os valores calculados pelo modelo e o desvio padrão da média diária do número *kappa* para o Modelo 6

O outro modelo neural, Modelo 7, com cinco variáveis na entrada da rede, não apresenta diferenças significativas nos resultados quando comparado ao modelo 5. A rede foi projetada com 5 neurônios na camada oculta e os resultados são apresentados na Figura 6.24. As Figuras 6.25 e 6.26 mostram, respectivamente, os erros absolutos e relativos médios calculados para os valores estimados pelo Modelo 7. Os valores comparados ao desvio padrão da média estão representados na figura 6.27. Os erros estão inseridos no intervalo $[-0,53; 0,86]$, o que significa um erro relativo médio máximo de 8,6%.

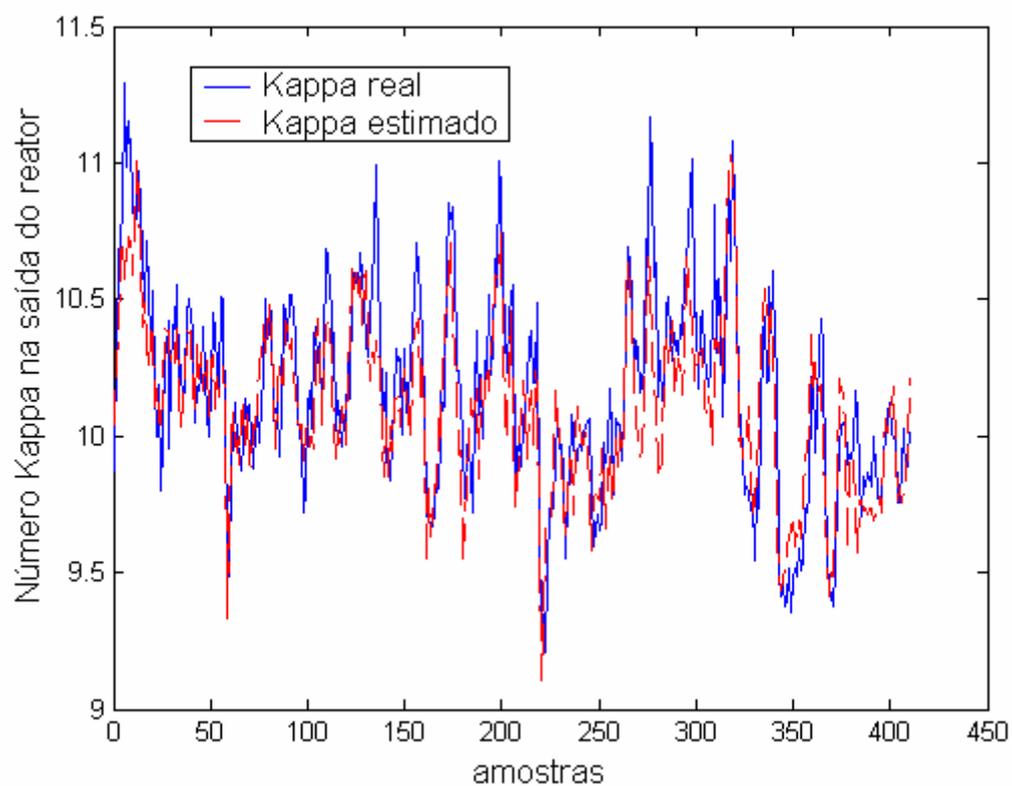


Figura 6.24 - Comparação entre os valores reais e estimados pelo Modelo 7

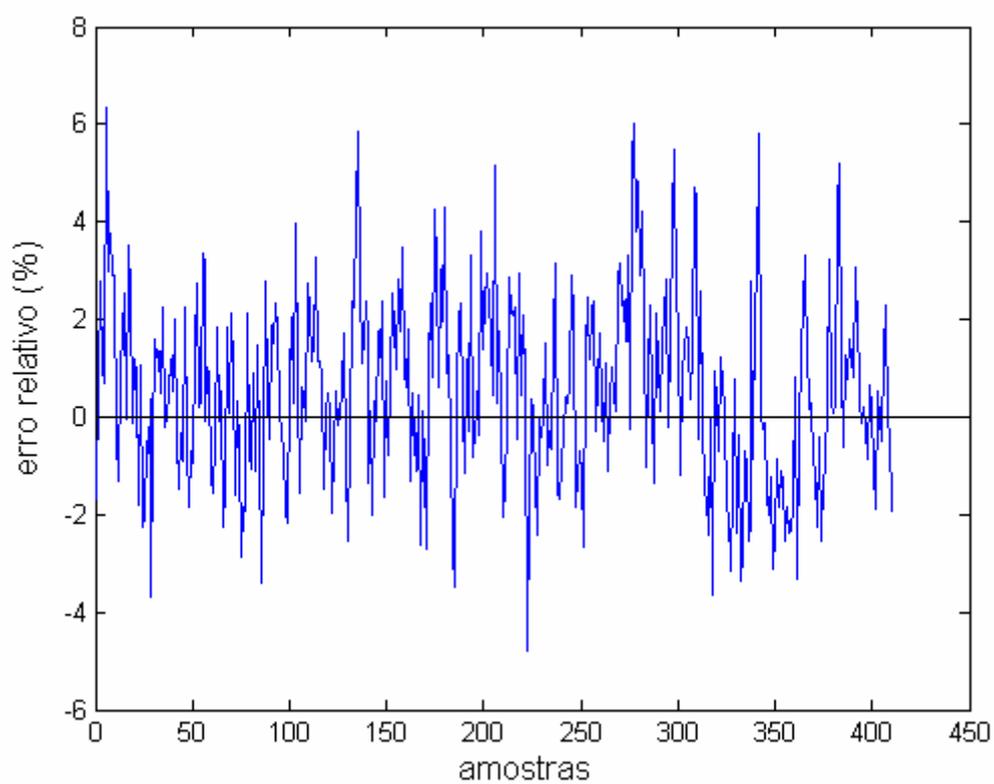


Figura 6.25 - Erro relativo médio calculado para o Modelo 7

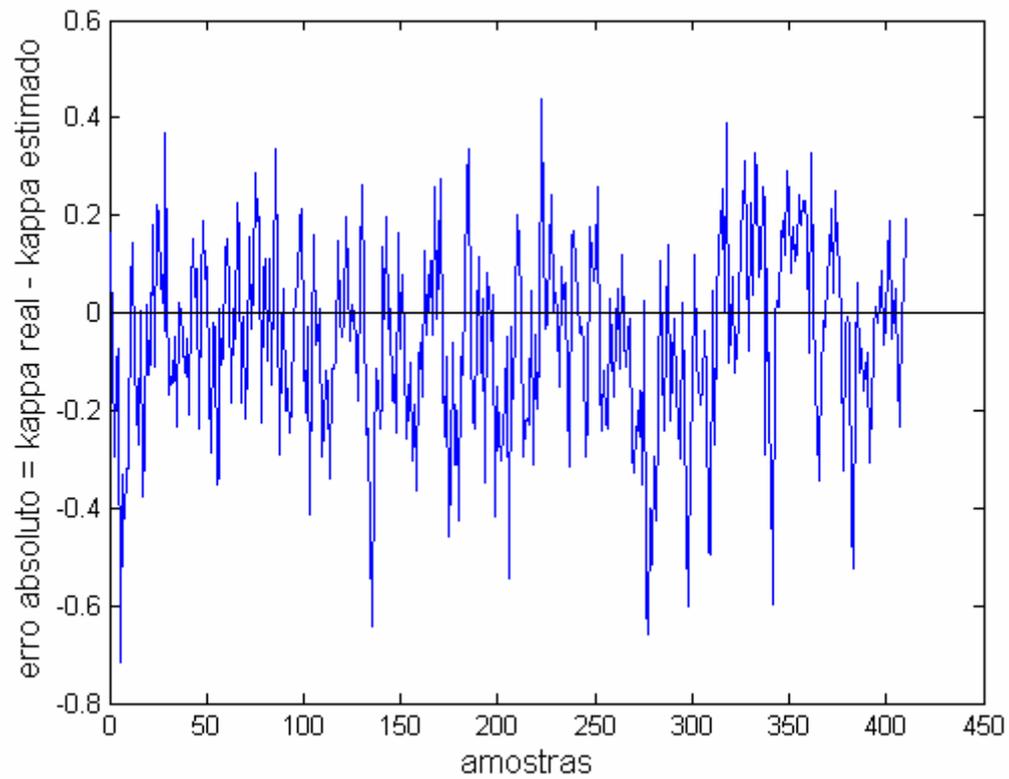


Figura 6.26 - Erro absoluto calculado para o Modelo 7

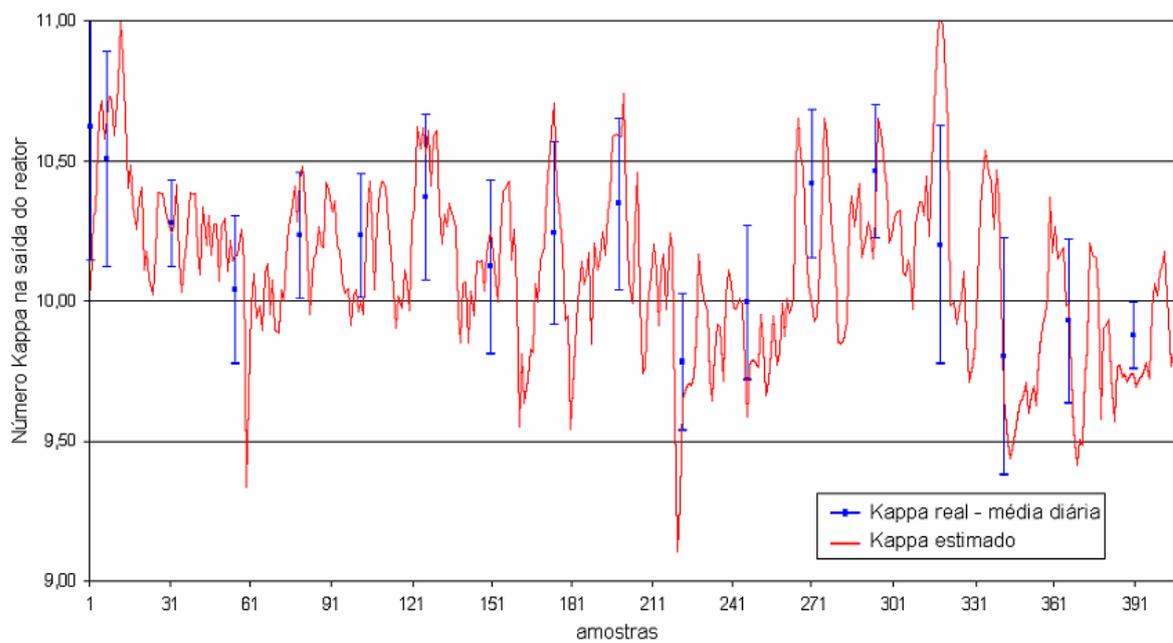


Figura 6.27 - Comparação entre os valores calculados pelo modelo e o desvio padrão da média diária do número *kappa* para o Modelo 7

A desvantagem de usar o modelo neural é a impossibilidade de acompanhar a evolução da reação com os dados disponíveis do processo. Desta forma não é possível gerar os perfis de decaimento de número *kappa* ou de consumo de álcali. Contudo é possível fazer uma análise da sensibilidade, ou seja, encontrar limites operacionais que minimizem o consumo de reagentes, mas que gerem produto com o padrão estabelecido pela empresa.

6.4 COMPARAÇÃO ENTRE OS MODELOS

Todos os sete modelos desenvolvidos para a planta de deslignificação apresentam erros da mesma ordem de grandeza e podem ser aplicados indistintamente para a simulação estacionária do processo.

A identificação do processo com o modelo linear e com a rede não era imprescindível, mas sua aplicação possibilitou um estudo comparativo entre as técnicas de modelagem.

No que diz respeito à praticidade, o modelo ARX e a rede neural são ferramentas facilmente implementadas, que podem ser utilizada para simulações rotineiras ou para previsões de resultados quando alterações operacionais forem necessárias. Em contrapartida os modelos fenomenológicos uma vez validados se apresentam como uma metodologia de análise da qualidade da reação, fornecendo não só o valor do número *kappa* da saída do reator, mas também informando o excesso ou falta de álcali alimentado à planta.

Quanto à implementação, as técnicas são amplamente conhecidas e estudadas e, portanto, não apresentam dificuldades. É importante lembrar que a resolução do modelo fenomenológico implica na integração de equações diferenciais ordinárias e

que, por este motivo, pode consumir mais tempo e esforço computacional. Contudo considerando o avanço tecnológico atual, essa diferença de tempo e esforço computacional tende a ser cada vez menos significativa e, portanto, este critério não é decisivo para a escolha de um modelo em detrimento de outro.

Um fator fundamental de ser avaliado quando é necessário escolher entre modelos é sua capacidade de descrever fenômenos físico químicos inerentes ao processo. Quanto mais completo for o modelo melhor, uma vez que mais informações sobre o processo poderão ser extraídas de seus resultados. Isto significa que a modelagem fenomenológica é preferível os modelos lineares e neurais identificados por ser mais completa e por possuir melhor embasamento teórico do processo. Na impossibilidade do uso de um modelo fenomenológico, deve-se sempre testar a identificação com modelos caixa-preta mais simples. Dessa forma, deve-se testar primeiro modelos lineares devido à sua simplicidade em relação aos modelos não lineares.

Ainda é preciso considerar que a modelagem fenomenológica apresentada neste trabalho é influenciada por valores médios de concentração de álcali. Isto implica que estes modelos podem, provavelmente, ser melhorados se dados reais foram utilizados para a determinação de novos parâmetros. A rede neural considera não possui a influência da concentração de álcali e um novo modelo poderia ser criado, no caso de se entender desejável que esta variável faça parte do conjunto de dados alimentados na rede.

Capítulo 7 - Conclusões

O objetivo deste trabalho foi desenvolver modelos estacionários que fossem capazes de simular o processo de deslignificação com oxigênio de polpa celulósica da Klabin Celulose e Papel, Unidade Monte Alegre – Telêmaco Borba, PR, para prever o número *kappa* da saída do reator. Desta forma, as variáveis operacionais pudessem ser alteradas e o resultado pudesse ser conhecido sem que o processo precisasse ser efetivamente alterado.

Foram propostos três modelos fenomenológicos distintos, baseados em modelos encontrados na literatura e adaptados para o estudo de caso. Os três modelos desenvolvidos são equivalentes, fornecendo resultados semelhantes e podem ser, desta forma, utilizados indistintamente para a estimação do número *kappa* de saída do reator de deslignificação.

Embora os erros associados à previsão dos modelos (em torno de 8,5%) possam parecer demasiadamente grandes, os valores estão de acordo com os erros encontrados na literatura, em torno de 8 - 20% (HSU; HSIEH, (d)1988; MYERS; EDWARDS, 1989). Um conjunto de dados mais completo, com medidas da concentração de álcali na entrada e análises laboratoriais deste reagente na saída do reator de deslignificação, melhorariam a precisão dos modelos desenvolvidos.

Para a identificação dos parâmetros dos modelos fenomenológicos o procedimento clássico de estimação não pode ser aplicado porque a reação não foi reproduzida em laboratório. Para contornar este problema e ajustar os modelos, rotinas de otimização heurísticas foram utilizadas para identificá-los. Este procedimento é simples de ser implementado e adaptado, fornecendo resultados

satisfatórios. Entre os dois métodos aplicados, algoritmos genéticos e recozimento simulado, o AG foi o que apresentou os melhores resultados.

Modelos caixa-preta lineares e neurais foram identificados, também para estimar o valor do número $kappa$ na saída do reator de deslignificação. Estes modelos possibilitaram comparações com a técnica de modelagem fenomenológica.

Na modelagem neural, foram testadas diversas estruturas, com diferentes números de neurônios na camada oculta e diferentes combinações de variáveis na camada de entrada da rede. As duas redes que proporcionaram os melhores resultados foram apresentadas no Capítulo 6. As redes forneceram erro absoluto da mesma ordem de grandeza que os modelos fenomenológicos, em torno de 8,5%, o que significa que as redes também podem ser utilizadas para prever o estado estacionário do processo industrial.

É importante ressaltar que os modelos fenomenológicos fornecem mais informações sobre o processo e são modelos mais completos, baseados na cinética da reação e nas leis de conservação de massa e energia, sendo, portanto, preferíveis quando comparado aos modelos caixa-preta.

Capítulo 8 - Sugestões para Trabalhos Futuros

Para trabalhos futuros é sugerido que um novo conjunto de dados seja coletado, incluindo os valores da evolução temporal de concentração de álcali na entrada do reator e análises laboratoriais do álcali na saída do reator, assim como dos números *kappas* na entrada e saída do reator para a confirmação dos perfis temporais gerados com os modelos fenomenológicos.

Propõe-se também que dados de viscosidade da polpa sejam coletados para que o modelo matemático possa ser completado e um indicador de qualidade da polpa após a reação possa ser previsto junto com o número *kappa*, que é um indicador de desempenho da deslignificação.

Para aumentar a funcionalidade dos modelos matemáticos, é sugerido que uma análise de sensibilidade dos modelos seja procedida, de forma que limites operacionais possam ser estabelecidos. Melhorias na qualidade da previsão podem ser alcançadas com ajustes dos expoentes dos modelos matemáticos.

Quando se avaliam os modelos identificados, o modelo linear ARX foi capaz de prever o valor do número *kappa* com confiabilidade, fornecendo resultados com a mesma capacidade preditiva das redes neurais com uma estrutura matemática muito mais simples. Em estudos futuros é sugerido que modelos lineares sejam mais investigados, esta identificação pode gerar bons modelos para a indústria.

Referências

ACKERT, J.E.; KOCH, D.D.; EDWARDS, L.L. **TAPPI Journal**, v. 58, n.10, p. 141, out. 1975.

AGARWAL, S.B; GENCO, J.M.; COLE, B.J.M.; MILLER, W. *Kinetics of oxygen delignification*. **Journal of Pulp and Paper Science**, v. 25, n.10. out. 1999.

AGUIAR, Helena Cristina I. L.; MACIEL FILHO, Rubens. *Modeling and Optimization of Pulp and Paper Processes using Neural Networks*. **Computers and Chemical Engineering**. v. 22, Suplemento, p. S981–S984, 1998.

AGUIRRE, Luis Antonio. **Introdução à identificação de sistemas: técnicas lineares e não-lineares aplicadas a sistemas reais**. Belo Horizonte: Editora UFMG, 2000.

AXEGARD, Peter; MOLDENIUS, Steve; OLM, Leelo. *Basic chemical kinetic equations are useful for an understanding of pulping processes*. **Svensk Papperstidning**. v. 82, n. 5, p. 131, 1979.

BAUGHMAN, D.R.; LIU, Y.A. **Neural networks in bioprocessing and chemical engineering**. San Diego: Academic Press, 1995.

BAZARAA, Mokhtar S.; SHERALI, Hanif D.; SHETTY, C.M. **Nonlinear programming: theory and algorithms**. 2. ed. New York: John Wiley & Sons, 1993.

BEQUETTE, B.W. **Process Dynamics: Modeling Analysis and Simulation**. Upper Saddle River: Prentice Hall, 1998.

BISCAIA JUNIOR, E.C.; VIEIRA, R.C. *Heuristic optimization for consistent initialization of DAEs*. **Computers and Chemical Engineering**, v. 24, p. 2183 – 2191, 2000.

COSTA JUNIOR, E.F; COSTA, A.O.S.; CALDEIRA, R.L.A.; CORREIA, F.M.; PESSOTTI, J.P. *Modelagem matemática da deslignificação com oxigênio da Celulose Kraft de Eucaliptus*. **37º Congresso e Exposição Anual de Celulose e Papel**. São Paulo, Associação Brasileira Técnica de Celulose e Papel (ABTCP), out. 2004.

CORREIA, F. N.; TAVARES, L. V. **Optimização linear e não linear: conceitos, métodos e algoritmos**. 2.ed. Porto: Fundação Calouste Gulbenkian, 1999.

DAMLIN, Sven-Arne; NORD, Hans. *On-line kappa number analysis: a complex chain*. **Congresso Anual da Associação Brasileira Técnica de Celulose e Papel (ABTCP)**, São Paulo – SP, 1990.

EDGAR, T.F.; HIMMELBLAU, D.M. **Optimization of Chemical Engineering Process**. Edição Internacional. Singapura: McGraw Hill, 1989.

EDWARDS, L.; NORBERG, S.E. *Alkaline delignification kinetics*. **TAPPI Journal**, v. 56, n. 11, p. 108, nov. 1973.

EVANS, J.E.; VENKATESH, V.; GRATZL, G.S.; CHANG, H.M. *The kinetics of low consistency oxygen delignification*. **TAPPI Journal**, v. 62, n.6, p. 37, jun. 1979.

FLETCHER, R. **Practical methods of optimization**. 2. ed. Chincercer: John Wiley & Sons, 2001.

FOGEL, D.B.; MICHALEWICZ, Z. **How to solve it: modern heuristics**. Berlim: Springer, 2002.

GLOVER, Fred; KELLY, James P.; LAGUNA, Manuel. *New advances for wedding optimization and simulation*. **Proceeding of the 1999 Winter Simulation Conference**. University of Colorado, New York: 1999.

GONZAGA, João Carlos B. **Integração de processos em tempo real para monitoramento e controle: aplicação para planta de PET**. 2003. Dissertação (Mestrado em Engenharia Química, Área de Concentração: Desenvolvimento de Processos Químicos). Universidade Estadual de Campinas, Campinas -SP, 2003.

HARTLER, N.; NORRSTROM, H.; RYDIN, S. *Oxygen-alkali bleaching of sulphate pulp*. **Svensk Papperstidning**, v. 73, n.21, p. 696, 1970.

HAYKIN, Simon. **Redes neurais: princípios e prática**. 2. ed. Porto Alegre: Bookman, 2001.

HEININGEN, A.; TUNC, M.S.; GAO, Y.; PEREZ, D.S. **Relationship between alkaline pulp yield and the mass fraction and degree of polymerization of cellulose in pulp.** Department of Chemical Engineering, University of Maine – Orono: 2002.

HERTZ, A.; WIDMER, M. *Guidelines for use of metaheuristics in combinatorial optimization.* **European Journal of Operational Research**, v. 151, p. 247 – 252. dez. 2003.

(a) HSU, C. L.; HSIEH, J.S. *Fundamentals of oxygen bleaching: measurement of oxygen diffusion rates in medium and high consistency pulp.* **TAPPI Journal**, v. 68, n.6, p. 92 – 95. jun. 1985.

(b) HSU, C. L.; HSIEH, J.S. *Effects of mass transfer on medium consistency oxygen bleaching kinetics.* **TAPPI Journal**, v. 68, n.11, p. 126 – 130. nov. 1985.

(c) HSU, C. L.; HSIEH, J.S. *Advantages of oxygen vs. air in delignifying pulp.* **TAPPI Journal**, v. 69, n.4, p. 125 – 128. jan. 1986.

(d) HSU, C. L.; HSIEH, J.S. *Reaction kinetics in oxygen bleaching.* **AIChE Journal**, v. 34, n.1. jan. 1988.

IRIBARNE, José; SCHROEDER, Leland R. *High-pressure oxygen delignification of kraft pulps. Part 1: Kinetics.* **TAPPI Journal**. V. 80, n. 10, p. 241 – 250. out. 1997.

JARREHULT, B.; SAMUELSON, O. *Oxygen bleaching of kraft pulps at low consistency.* **Svensk Papperstidning**, v. 81, n.17, p. 533, 1978.

KIRKPATRICK, S.; GELATT JUNIOR, C.D., VECCHI, M.P. *Optimization by simulated annealing.* **Science**, v.. 220, n. 4598, p. 671 – 680, maio 1983.

KOURIS, M (ed.). **Dictionary of Paper**. 5a. ed. TAPPI Press, 1996.

LEMMETTI, A.; LEIVISKÄ, K.; SUTINEN, R. *Kappa number prediction based on cooking liquor measurements.* **Relatório A n.5**. University of Oulu – Oulu. maio 1998.

LIEBERGOTT, N.; LIEROP, B. van.; SKOTHOS, A. *A survey of the use of ozone in bleaching pulps, part 1.* **TAPPI Journal**, v. 75, n.1, p. 145. jan. 1992.

LJUNG, Lennard. **System Identification: Theory for the User**. 2^a ed. Upper Sadle River: Prentice Hall, 1999.

McCALL, John. *Genetic algorithms for modeling and optimization*. **Journal of Computational and Applied Mathematics**, v. 184, p. 205 – 222, 2005.

MEADOWS, E. B.; BUGAJER, S. *Desenvolvimentos recentes no processo de deslignificação*. **Congresso Anual da Associação Brasileira Técnica de Celulose e Papel (ABTCP)**, São Paulo – SP, 1988.

MELEIRO, Luiz A. da Cruz. **Introdução à Identificação de Processos Não Lineares com Redes Neurais Artificiais**. Apostila. Curitiba: UFPR, 2004.

MELEIRO, Luiz A. da Cruz. **Projeto e Aplicação de Controladores Baseados em Modelos Lineares, Neurais e Nebulosos**. Tese (Doutorado em Engenharia Química). UNICAMP. Campinas, 2002.

MYERS, Michael R.; EDWARDS, Lois L. *Development and verification of a predictive oxygen delignification model for hardwood and softwood kraft pulp*. **TAPPI Journal**, v. 72 n.9. set. 1989.

NOCEDAL, Jorge; WRIGHT, Stephen J. **Numerical optimization**. New York: Springer, 1999.

(a) OLM, L.; TEDER, A. *The kinetics of oxygen bleaching*. **TAPPI Journal**, v. 62, n.12, dez, 1979.

(b) OLM, L.; TEDER, A. *Extended delignification by combination of modified kraft pulping and oxygen bleaching*. **Paperi Ja Puu**, v. 63, n.4a, p. 315, 1981.

OSMAN, I.H. *Preface: focused issue on applied metaheuristics*. **Computer & Industrial Engineering**, v. 44, p. 205 – 207, 2002

RAGNAR, Martin. **The technology of oxygen delignification and bleaching of chemical pulp**. Kvaerner Pulping, 2002.

SCHMAL, Martin. **Cinética Homogênea Aplicada a Cálculo de Reatores**. Rio de Janeiro: Editora Guanabara Dois, 1982.

SCHÖÖN, Nils-Herman. *Interpretation of rate equations from kinetics studies of wood pulping and bleaching*. **Svensk Papperstidning**, p. R185 – R192, 1982

SILVER, E.A. *An overview of heuristic solution methods*. **Journal of the Operational Research Society**, v. 55, n.9, p. 936 – 956, 2004.

SPIEGEL, Murray R.; SCHILLER, John; SRINIVASAN. R. Alu. **Teoria e Problemas de Probabilidade e Estatística**. 2. ed. Porto Alegre: Bookman, 2004.

VIOLETTE, Steven M. *Oxygen delignification kinetics and selectivity improvement*. Tese (Doutorado em Engenharia Química). University of Maine. Orono, 2003.

XIE, T.; GHIAASIAAN, S.M.; KARRILA, S. *Artificial neural network approach for flow regime classification in gas-liquid-fiber flows based on frequency domain analysis of pressure signals*. **Chemical Engineering Science**. n. 59. p. 2241-2251, 2004.

YAMAMOTO, Carlos Itsuo. **Modelagem matemática e otimização do processo industrial de síntese de amônia utilizando redes neurais**. Tese (Doutorado em Engenharia Química). Escola Politécnica da Universidade de São Paulo. São Paulo, 1998.

ZOU, H.; LIUKKONEN, A.; COLE, B.; GENCO, J.; MILLER, W. *Influence of kraft pulping on the kinetics of oxygen delignification*. **TAPPI Journal**, v. 83, n.2. fev. 2000.

von ZUBEN, Fernando J. **Computação evolutiva: uma abordagem pragmática**. Apostila. Campinas: Unicamp, 2004.

Glossário

Fonte: KOURIS, M. (editor). *Dictionary of paper*. 5. ed. Cidade: TAPPI Press, 1996.

Álcali – Ver hidróxido de sódio.

Cavaco – Pedaco de madeira de pequenas dimensões.

Celulose – Principal constituinte da estrutura da madeira, ocorrendo nas diversas espécies do reino vegetal. Quimicamente é um polissacarídeo linear de fórmula empírica $(C_6H_{10}O_5)_n$ e peso molecular variável, de acordo com a fonte da celulose, mas normalmente maior que $1.000.000 \frac{g}{mol}$. A celulose de madeira é o material remanescente do cozimento e branqueamento da madeira, ou seja, após a remoção da lignina e alguns carboidratos.

Compostos cromóforos – Compostos coloridos que contêm estrutura básica cromogênica, usualmente um conjunto conjugado de duplas ligações que absorvem radiação na região do visível e do ultravioleta. O composto que absorve luz na polpa é a lignina ou produtos da lignina formados em reações durante a polpação.

Consistência – Porcentagem em peso de material fibroso seco em forno. Também é conhecido como densidade, sólidos na polpa ou concentração da polpa. A uma determinada razão de químicos na polpa, o aumento da consistência aumenta também a concentração de químicos na solução e conseqüentemente um aumento da taxa de reação; daí a importância da consistência na indústria de celulose e papel. É classificada em: baixa consistência, até 4%, média consistência, de 10 a 16% e alta consistência 30% ou mais.

Os termos utilizados para expressar a consistência são: % o.d.(oven dry) porcentagem de polpa seca em forno na suspensão determinada pela expressão: $\% o.d. = \frac{\text{quantidade de polpa seca no forno (polpa o.d.)}}{\text{quantidade de polpa úmida (antes de secar)}} \cdot 100$; % a.d. (air dry

ou, s.a – seca ao ar) porcentagem de polpa seca ao ar na suspensão, determinada pela expressão: $\%a.d. = \frac{\text{quantidade de polpa seca no forno (polpa o.d.)}}{0,9}$.

Consumo de químicos – Quantidade de químicos consumidos por uma quantidade de polpa durante o estágio de deslignificação.

Cozimento – Ver polpação.

Deslignificação – Processo de remoção de lignina de madeira e outros materiais celulósicos com o uso de produtos químicos, como oxigênio e compostos clorados. O resíduo da deslignificação consiste basicamente de celulose e hemi-celulose.

Deslignificação com Oxigênio – Tratamento da polpa com oxigênio em meio alcalino para oxidar, degradar e solubilizar a lignina. O estágio de deslignificação com oxigênio está localizado entre os estágios de cozimento e de branqueamento, remove até 50% da lignina residual da polpa marrom (polpa efluente da polpação) sem danificar suas fibras ou alterar suas propriedades. Os químicos utilizados neste processo são compatíveis aos utilizados no cozimento *Kraft* tornando possível a reciclagem do filtrado do estágio para a caldeira de recuperação.

Digestor – Vaso utilizado para o cozimento (polpação) de matérias-primas fibrosas para produzir polpa e remover lignina. O vaso pode ser utilizado em processos contínuos ou em batelada. O digestor é construído com metais que resistem a altas temperaturas e pressões e produtos químicos corrosivos.

Hemi-celulose – Classe de polissacarídeos que são associados à celulose. A hemi-celulose possui peso molecular menor que a celulose, tipicamente entre 20.000 a 40.000 $\frac{g}{mol}$. São, normalmente, ramos na molécula da celulose e podem ser substituídas por outros açúcares ou ácidos urônicos. Contêm mais de um tipo de açúcar monomérica em sua composição básica, os “açúcares de madeira”: glucose, manose, galactose, xilose e arabinose.

Hidróxido de sódio – Composto alcalino também conhecido como soda cáustica e álcali.

Licor branco – Ou licor de cozimento. Combinação de hidróxido de sódio e sulfato de sódio para dissolver a lignina no processo de cozimento.

Lignina – Parte da célula vegetal que não contém carboidratos. É usualmente determinada como o resíduo da hidrólise do vegetal com um ácido forte, após a remoção de ceras e outros extrativos. A lignina é amorfa, possui alto peso molecular e sua estrutura é predominantemente constituída de anéis aromáticos. Sua composição varia de acordo com o método de isolamento, espécie, idade do vegetal, condições de crescimento do vegetal, entre outros fatores. A lignina é praticamente toda removida durante o processo de polpação química. Os processos de deslignificação e branqueamento removem e modificam a lignina remanescente na polpa.

Madeira dura (Hardwood) – Madeira obtida de árvores angiospermas, como eucalipto, *maple* (árvore característica do Canadá), carvalho. Essas árvores são caracterizadas por apresentarem ampla folhagem e são comuns em zonas temperadas.

Madeira mole (Softwood) – Madeiras de coníferas de folhas pontiagudas como pinheiros, abeto e algumas variedades de cedro.

Número Kappa – Valor de teste relacionado linearmente com a quantidade de lignina remanescente na polpa após o cozimento (polpação) e, desta forma, um estimador da demanda de químicos para o processo de deslignificação e branqueamento. O número kappa é a quantidade de mililitros de solução 0,1 N de permanganato de potássio (KMnO_4) normalizado reduzido de um grama de polpa seca em forno sob condições controladas. O resultado é corrigido para o consumo de 50% de permanganato adicionado. O procedimento para determinar o número kappa foi padronizado pelo Comitê de Testes de Polpa da Divisão de Qualidade de Produtos e Processos da TAPPI em 1976, norma T 236 hm-85. O cálculo do número

kappa é determinado pela expressão: $K = \frac{p \cdot f}{w}$, sendo K o número kappa, f o fator de correção para 50% de consumo de permanganato, em função de p ; w a quantidade de polpa o.d. em gramas; $p = \frac{(b-a)N}{0,1}$, sendo p a quantidade de permanganato, em mililitros, consumida efetivamente durante o teste; b quantidade de tiosulfato, em mililitros, consumido durante o teste; e N a normalidade do tiosulfato.

Polpa – Termo utilizado para descrever as fibras após o estágio de polpação quando a fibra é liberada do cavaco ou palha, algodão, etc.

Polpa química – Polpa obtida pelo uso de produtos químicos e calor que dissolvem a lignina que mantém as fibras celulósicas unidas. Os processos básicos de polpação química são o processo sulfato (Kraft), sulfito e soda.

Polpação – Processo de separar a matéria-prima celulósica (como cavacos, palitos, palhas e papel reciclado) em polpa fibrosa própria para a fabricação de papel ou papelão. O processo de polpação varia desde uma simples ação mecânica de esmagamento da matéria prima até uma complexa seqüência de digestão (cozimento).

Polpação Kraft – Processo de polpação alcalina que utiliza uma combinação de hidróxido de sódio e sulfato de sódio como agentes de cozimento. O termo *Kraft* é alemão e significa “forte”, um termo propício já que o processo *Kraft* é o mais agressivo entre os processos de polpação química.

Reator de fluxo ascendente – Vaso para reação em que a polpa é alimentada na parte inferior e bombeada para a saída, na parte superior do vaso.

Viscosidade – Medida de degradação causada pelos processos de polpação, deslignificação e branqueamento dos carboidratos de alto peso molecular, principalmente celulose, nas polpas. Idealmente, os químicos utilizados deveriam remover ou branquear somente a lignina e demais compostos coloridos na polpa,

sem danificar ou degradar as fibras celulósicas. Mas a polpação, a deslignificação e o branqueamento também atacam as moléculas celulósicas, diminuindo a viscosidade da polpa. A perda no valor da viscosidade pode acarretar na fragilidade da polpa. Existem dois procedimentos para determinar o valor da viscosidade da polpa: o padrão estabelecido pela TAPPI e o padrão ASTM. O procedimento para determinar a viscosidade da polpa (conhecida como viscosidade TAPPI) padronizada pelo Comitê de Testes de Polpa da Divisão de Qualidade de Produtos e Processos da TAPPI em 1982 (última revisão em 1994), norma T 230 om-94 é feito através da expressão: $V = C \cdot t \cdot d$, sendo V a viscosidade da solução da polpa em centipoise (cp); C é a constante do viscosímetro encontrada na calibração do instrumento; t é o tempo médio de teste, em segundos; e d a densidade da solução da polpa, em gramas por centímetros cúbicos (normalmente igual a 1,052). O padrão ASTM (D1795-62) estima a viscosidade intrínseca da polpa em $\frac{mL}{g}$.