

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE



UNICAMP

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA QUÍMICA
ÁREA DE CONCENTRAÇÃO DESENVOLVIMENTO DE PROCESSOS QUÍMICOS

MODELAGEM DE DIGESTORES KRAFT CONTÍNUO: REDES NEURAIS E MODELO HÍBRIDO

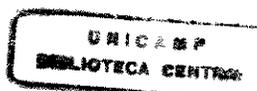
Helena Cristina I. L. Aguiar

Orientada por Prof. Dr. Rubens Maciel Filho

Dissertação de Mestrado apresentada à Faculdade de Engenharia Química
como parte dos requisitos exigidos para a obtenção do título de Mestre em
Engenharia Química

Campinas, 11 de Fevereiro de 2000.

81191000



UNIDADE	BC
N.º CHAMADA:	T/UNICAMP
	Ag 93m
V.	Ex
TOMBO BC/	42812
PROC.	16-278100
C	<input type="checkbox"/>
D	<input checked="" type="checkbox"/>
PREC.º	R\$ 11,00
DATA	20/10/00
N.º CPD	

CM-00147199-4

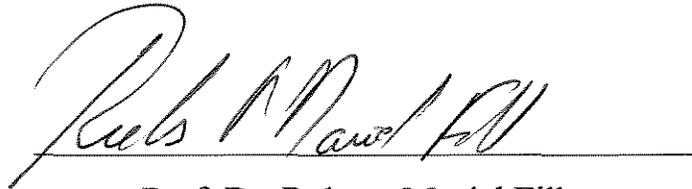
FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

Ag93m Aguiar, Helena Cristina I. L.
Modelagem de digestor Kraft contínuo: redes neurais e modelo híbrido / Helena Cristina I. L. Aguiar.--
Campinas, SP: [s.n.], 2000.

Orientador: Rubens Maciel Filho.
Dissertação (mestrado) - Universidade Estadual de
Campinas, Faculdade de Engenharia Química.

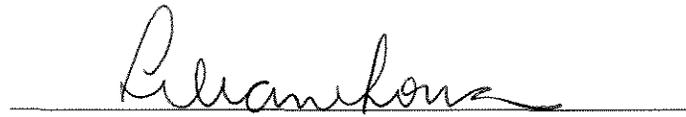
1. Redes neurais (Computação). 2. Digestores
(Engenharia química). I. Maciel Filho, Rubens. II.
Universidade Estadual de Campinas. Faculdade de
Engenharia Química. III. Título.

Dissertação de Mestrado apresentada e aprovada em 11 de Fevereiro de
2000 pela banca examinadora constituída por



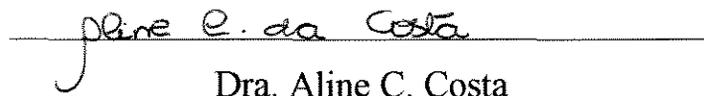
A handwritten signature in cursive script, reading "Rubens Maciel Filho", is written above a horizontal line.

Prof. Dr. Rubens Maciel Filho



A handwritten signature in cursive script, reading "Liliane M. F. Lona Batista", is written above a horizontal line.

Profa. Dra. Liliane M. F. Lona Batista



A handwritten signature in cursive script, reading "Aline C. Costa", is written above a horizontal line.

Dra. Aline C. Costa

Este exemplar corresponde à redação final da Dissertação de Mestrado em Engenharia Química apresentada por Helena Cristina I. L. Aguiar e aprovada pela Comissão Julgadora em 11 de Fevereiro de 2000.

A handwritten signature in black ink, reading "Rubens Maciel Filho". The signature is written in a cursive style with a large initial 'R' and is positioned above a horizontal line.

Prof. Dr. Rubens Maciel Filho

Aos homens da minha vida, que são a prova de que honestidade, mangas arregaçadas e acreditar que será possível compõem a melhor receita para o sucesso:

*Meu vô Hélio, que continua acreditando nos seus ideais.
E meu pai, Carlito, minha fonte de inspiração e meu ídolo.*

E à minha mãe que nunca me deixa parar de acreditar...

Obrigadas...

Ao meu orientador, Rubens, pela orientação e injeção de ânimo nos momentos críticos!

Ao CNPq pelo apoio financeiro

À Aracruz Celulose S.A., pelo apoio financeiro e pelos dados. E a todas as pessoas desta empresa, que me ajudaram com idéias, críticas, questionamentos, dados, artigos, consultas... Ergílio, Braz, Sílvia, Reinaldo, Ângela, Giácoma, Alfredo, Aurélio e Jaqueline (e muitos outros).

À Aline e Liliâne, que participaram da banca da minha defesa de tese e colaboraram com críticas interessantes.

À Ahlstrom Machinery Inc. pelas boas figuras e pelo tempo concedido para que acabasse esta tese.

À minha família Gelol - Carlito, Helane, Lulu e Bié - , por terem chorado, sofrido, esperado e vibrado com cada passo desse caminho. Afinal, não basta ser família, tem que participar! E ao Cláudio, que já está com um pé e meio na família!

Ao Aaron, pelo apoio e pelo carinho.

À minha tia Helena, a grande acadêmica da família, que entende tudo sobre educação e merece toda a minha admiração.

À Aninha que foi minha consultora e excelente conselheira para assuntos de tese.

Ao pessoal da FEQ/Unicamp e do LOPCA pelo apoio e paciência, principalmente o Eduardo Coselli (Urso), o Cristiano e a Marta.

À Dâmaris, por todas as vírgulas, as formalidades e as infinitas dicas que aumentaram a qualidade deste trabalho!

Ao Ramiro e ao Guilherme, que aguentaram meus ataques de mal humor e não reclamaram muito dos longos períodos de tempo que eu ficava ocupando o computador...

À Erlí, Branca e Mariázinha, que sempre cuidam de mim muito bem!

Ao Gustavo e à Jussara (que não se conhecem)- vocês sabem por quê, não?

A todos os meus amigos, de Campinas e do Coqueiral (pequenos, médios e grandes), pelo carinho e vibração com as minhas conquistas.

Aos outros homens da minha vida, pela paciência ao ouvirem "Hoje não dá, tenho que rodar o meu programa..."

Obrigada a todas essas pessoas que eu amo e fazem a minha vida valer à pena. É por causa delas que eu tenho ânimo para dar cada passo adiante!

Saudades...

da Lídia, que foi uma verdadeira Mestre, que nunca se limitou a simplesmente repetir palavras na sala de aula.

Pelo seu carinho, sua real preocupação pela formação de seus alunos, sua dedicação e pela sua luta para transformar equações, conceitos e engenharia numa aventura.

Que o seu exemplo inspire muitos outros. Você faz muita falta.

O limite do homem é tão conhecido quanto à linha do horizonte, que se move a cada passo completado...

Helane Ibiapina Lira Aguiar

SUMÁRIO

LISTA DE FIGURAS

LISTA DE TABELAS

NOMENCLATURA

CAPÍTULO I

Introdução

- | | |
|------------------------------------|-----------------------|
| <i>I.1 Revisão e Motivação</i> | <i>Capítulo I - 1</i> |
| <i>I.2 De Capítulo em Capítulo</i> | <i>Capítulo I - 4</i> |

CAPÍTULO II

Revisão da Literatura e Conceitos Básicos Parte 1: Fabricação de Celulose

- | | |
|--|------------------------|
| <i>II.1 Introdução</i> | <i>Capítulo II - 1</i> |
| <i>II.2 Cenário da Indústria de Papel e Celulose no Brasil</i> | <i>Capítulo II - 2</i> |
| <i>II.3 Processo de Produção da Celulose</i> | <i>Capítulo II - 3</i> |
| <i>II.3 Processo de Produção da Celulose</i> | <i>Capítulo II - 3</i> |

II.3.1	Matéria prima	Capítulo II – 4
II.3.2	Polpação	Capítulo II – 5
II.3.3	Licor no processo Kraft	Capítulo II – 7
II.3.4	Digestores	Capítulo II – 8
II.3.5	Digestor Contínuo	Capítulo II – 9
II.4	<i>Políticas de operação</i>	Capítulo II – 13
II.5	<i>Objetivos e Controle da Polpação química</i>	Capítulo II – 16
II.5.1	Viscosidade e Propriedades	Capítulo II – 17
II.5.2	Kappa	Capítulo II – 18
II.6	<i>Propriedades das Fibras Branqueadas</i>	Capítulo II – 19
II.6.1	Introdução	Capítulo II – 19
II.6.2	Refino	Capítulo II – 19
II.2.1	Efeitos do refino	Capítulo II – 20
II.6.3	Características importantes para o papel ou seu processo de fabricação	Capítulo II – 21
II.6.4	Interações entre as fibras e suas relações com as propriedades	Capítulo II – 22
II.6.5	Influência da madeira sobre as propriedades das fibras	Capítulo II – 24
II.6.6	Influência do processo sobre as propriedades	Capítulo II – 26
II.6.7	Considerações Finais	Capítulo II – 29

CAPÍTULO III

Revisão da Literatura parte 2: Modelos matemáticos

<i>III.1 Introdução</i>	<i>Capítulo III - 1</i>
<i>III.2 Fenômenos cinéticos durante a polpação</i>	<i>Capítulo III - 2</i>
III.2.1 Impregnação dos cavacos	<i>Capítulo III - 2</i>
III.2.2 Reações dos constituintes da madeira	<i>Capítulo III – 5</i>
<i>III.3 Modelos estudados</i>	<i>Capítulo III – 9</i>
III.3.1 Digestor contínuo com escoamento paralelo e contra-corrente.	<i>Capítulo III – 9</i>
III.3.2 Modelo cinético preditivo: aplicação na polpa de eucalipto	<i>Capítulo III – 14</i>
III.3.3 Modelo dinâmico para deslignificação	<i>Capítulo III – 18</i>
III.3.4 Modelagem da penetração incompleta do cavaco	<i>Capítulo III – 19</i>
III.3.5 Discussão sobre os modelos matemáticos	<i>Capítulo III – 22</i>
<i>III.4 Inteligência Artificial e Redes Neurais</i>	<i>Capítulo III – 24</i>
III.4.1 Redes Neurais	<i>Capítulo III – 26</i>

III.4.1.1	Propriedades	<i>Capítulo III – 27</i>
III.4.1.2	Fundamentos da computação neural	<i>Capítulo III – 28</i>
III.4.1.3	Operação de uma rede neural	<i>Capítulo III – 29</i>
III.4.1.4	Treinamento	<i>Capítulo III – 29</i>
III.4.1.5	Treinamento por retropropagação (backpropagation)	<i>Capítulo III – 30</i>
III.4.1.6	Aplicações	<i>Capítulo III – 31</i>
III.4.1.7	Vantagens e limitações	<i>Capítulo III – 32</i>
III.4.2	Redes Híbridas	<i>Capítulo III – 34</i>
III.4.3	Comparações entre redes neurais puras e híbridas	<i>Capítulo III – 35</i>
III.5	<i>Conclusões</i>	<i>Capítulo III – 36</i>

CAPÍTULO IV

Predição de Propriedades da Celulose Através de Redes Neurais

IV.1	<i>Introdução</i>	<i>Capítulo IV - 1</i>
IV.2	<i>Considerações Iniciais</i>	<i>Capítulo IV – 2</i>
IV.3	<i>Nomenclatura utilizada neste capítulo</i>	<i>Capítulo IV – 3</i>
IV.4	<i>Desenvolvimento do modelo</i>	<i>Capítulo IV – 4</i>
IV.4.1	Coleta de dados	<i>Capítulo IV – 4</i>
IV.4.2	Rede Inicial	<i>Capítulo IV – 7</i>
IV.4.3	Tratamento dos dados	<i>Capítulo IV – 12</i>
IV.4.4	Alterações na estrutura da rede	<i>Capítulo IV – 16</i>
IV.4.5	Rede Neural Otimizada	<i>Capítulo IV – 23</i>

<i>IV.5 Resultados</i>	<i>Capítulo IV – 25</i>
IV.5.1 Análise dos conjuntos de dados industriais	<i>Capítulo IV – 25</i>
IV.5.2 Análise do número de variáveis	<i>Capítulo IV – 27</i>
IV.5.3 Efeito do número de neurônios nas camadas ocultas	<i>Capítulo IV – 30</i>
IV.5.4 Utilização de correlação para a otimização do conjunto de treinamento	<i>Capítulo IV – 32</i>
IV.5.5 Efeito do número de iterações	<i>Capítulo IV – 33</i>
<i>IV.6 Conclusão</i>	<i>Capítulo IV – 36</i>

CAPÍTULO V

Modelo Determinístico e Modelo Híbrido: Comparação e discussão sobre as diferentes técnicas de modelagem

<i>V.1 Introdução</i>	<i>Capítulo V - 1</i>
<i>V.2 Modelo matemático</i>	<i>Capítulo V - 2</i>
V.2.1 Equações do modelo	<i>Capítulo V - 4</i>
V.2.2 Métodos de solução	<i>Capítulo V – 5</i>
V.2.3 Resultados <i>do modelo determinístico</i>	<i>Capítulo V - 5</i>
V.2.4 Discussão	<i>Capítulo V – 10</i>
<i>V.3 Modelo Híbrido</i>	<i>Capítulo V – 12</i>
V.3.1 Motivação	<i>Capítulo V – 12</i>
V.3.2 Estrutura do Modelo Híbrido	<i>Capítulo V – 13</i>
V.3.3 Resultados	<i>Capítulo V – 15</i>

<i>V.4 Discussão sobre diferentes técnicas de modelagem – Redes neurais, Modelos determinísticos e Redes híbridas</i>	<i>Capítulo V – 16</i>
V.4.1 Grau de conhecimento teórico	<i>Capítulo V – 16</i>
V.4.2 Obtenção de variáveis utilizadas para o desenvolvimento do modelo	<i>Capítulo V – 18</i>
V.4.3 Quantidade de dados necessários para o desenvolvimento do programa	<i>Capítulo V – 20</i>
V.4.4 Qualidade dos dados	<i>Capítulo V – 21</i>
V.4.5 Faixa de validade do modelo	<i>Capítulo V – 22</i>
V.4.6 Interpretação do modelo	<i>Capítulo V – 23</i>
V.4.7 Redes neurais puras e híbridas	<i>Capítulo V – 24</i>
V.4.8 Considerações finais	<i>Capítulo V – 25</i>
<i>V.5 Conclusão</i>	<i>Capítulo V - 27</i>

CAPÍTULO VI

Conclusões e Sugestões

<i>VI.1 Conclusões</i>	<i>Capítulo VI – 1</i>
<i>VI.2 Sugestões</i>	<i>Capítulo VI – 4</i>

APÊNDICE 1

Principais propriedades das fibras

LISTA DE FIGURAS

- Figura 2.1** **Capítulo II - 4**
Desenho esquemático de uma linha de fibras
- Figura 2.2** **Capítulo II - 10**
Sistema de alimentação de cavacos convencional e digestor contínuo
- Figura 4.1** **Capítulo IV - 6**
Zonas de cozimento do digestor
- Figura 4.2** **Capítulo IV - 8**
Esquema da transferência de sinal em uma rede neural *feedforward*
- Figura 4.3** **Capítulo IV - 9**
Um neurônio genérico recebendo, processando e transmitindo um sinal
- Figura 4.4** **Capítulo IV - 13**
Exemplo de uma variável bem distribuída - Gráfico Álcali na zona CD1 versus Número Kappa
- Figura 4.5** **Capítulo IV - 14**
Exemplo de uma variável mal distribuída - Gráfico Sulfidez versus Número Kappa
- Figura 4.6** **Capítulo IV - 15**
Comparação dos resultados obtidos por redes neurais com diferentes variáveis de entrada
- Figura 4.7** **Capítulo IV - 23**
Rede neural otimizada
- Figura 4.8** **Capítulo IV - 30**
Efeito do número de variáveis

Figura 4.9 Efeito do número de neurônios nas camadas ocultas	Capítulo IV - 31
Figura 4.10 Redes neurais para a predição do índice de tração	Capítulo IV - 33
Figura 4.11 Efeito do número de iterações	Capítulo IV - 34
Figura 5.1 Efeito da temperatura na taxa de degradação dos carboidratos durante a polpação	Capítulo V - 6
Figura 5.2 Efeito da temperatura na taxa de deslignificação	Capítulo V - 6
Figura 5.3 Efeito da carga alcalina sobre a taxa de deslignificação	Capítulo V - 8
Figura 5.4 Efeito da difusividade e sulfidez na taxa de deslignificação	Capítulo V - 9
Figura 5.5 Estrutura do modelo híbrido	Capítulo V - 15
Figura 5.6 Resultados do modelo híbrido	Capítulo V - 16
Figura 5.7 Comparação entre os resultados obtidos com a rede neural pura e o modelo híbrido	Capítulo V - 25

LISTA DE TABELAS

Tabela 3.1 As equações do modelo de Saltin (1992)	Capítulo III - 12
Tabela 3.2 Equações do modelo desenvolvido por Mirams e Nguyen (1994)	Capítulo III - 16
Tabela 3.3 Equações cinéticas do modelo dinâmico para a deslignificação	Capítulo III - 19
Tabela 3.4 As equações cinéticas do modelo da impregnação incompleta do cavaco	Capítulo III - 21
Table 4.1 Variáveis coletadas para treinamento das redes neurais	Capítulo IV - 5
Tabela 4.3 Variáveis de entrada do modelo de rede neural para número kappa	Capítulo IV - 18

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

NOMENCLATURA

1 INTRODUÇÃO

A nomenclatura descrita abaixo está dividida por assunto para facilitar o entendimento do leitor. Dentro de cada assunto, está em ordem alfabética.

2 CELULOSE

2.1 MODELO 1 – DIGESTOR CONTÍNUO COM ESCOAMENTO PARALELO E CONTRA-CORRENTE

A	conteúdo de Arabinoxylan na madeira	kg/kg de madeira
Achips	área superficial dos cavacos	m ² /kg
C	conteúdo de celulose na madeira	kg/kg de madeira
CH	conteúdo de carboidratos na madeira	kg/kg de madeira
C _{pchips}	calor específico da madeira	kJ/kg K
C _{pliquid}	calor específico do licor	kJ/kg K
D	coeficiente de difusão	
D _{digestor}	diâmetro do digestor	m
E	conteúdo de extrativos na madeira	kg/kg de madeira
g	coeficiente de gravidade	m/s ²
G	conteúdo de galactoglucomannan na madeira	kg/kg de madeira

h	altura do digestor	m
K	número Kappa	
k_i	constante de Arrhenius	
k_{mass}	coeficiente de transferência de massa	
l	comprimento do cavaco	m
L	conteúdo de lignina na madeira	kg/kg de madeira
L_f	conteúdo de lignina rápida na madeira	kg/kg de madeira
L_s	conteúdo de lignina lenta na madeira	kg/kg de madeira
m_{chips}	massa de cavacos	kg
P_{chips}	pressão nos cavacos, forças de contato atuando na coluna de cavacos dividida pela área transversal da coluna	Pa
P_{liquid}	pressão da fase líquida	Pa
Q_r	calor produzido pela reação	kJ/s
Q_t	calor transferido entre o líquido e os cavacos	kJ/s
R	constante universal dos gases	kJ/K mol
T	temperatura	K
t	tempo	s
$V_{chipsliquid}$	volume de líquido penetrado	m^3
w	espessura do cavaco	m
W	componente da madeira	kg/kg de madeira
W_{∞}	concentração do componente na madeira após tempo infinito	kg/kg de madeira
a, b, c	expoente da equação cinética	

W	Componente da madeira	
x	Distância do centro do cavaco	Adimensional
X	Espessura/2	M
ρ	Densidade	kg/dm ³
ε	Porosidade	

2.3 MODELO 3 – MODELO DINÂMICO PARA A DESLIGNIFICAÇÃO

[HS ⁻]	Concentração de HS ⁻ no início da fase intermediária	moles/l
L	Quantidade de lignina presente na madeira	%
[OH ⁻]	Álcali efetivo no início da fase intermediária	moles/l
t	Tempo	min
T	Temperatura	K

2.4 MODELO 4 – MODELO DA PENETRAÇÃO INCOMPLETA

B _w	Número de Biot (parede)	Adimensional
C _i	Concentração da espécie i na fase cavaco	Adimensional
C _{ib}	Concentração da espécie i na fase líquida	Adimensional
Ra _i	taxa de reação da espécie i	Adimensional
x	distância da frente de penetração	Adimensional

i	espécie química	
ε	fração de vazios nos cavacos	
μ	coeficiente de atrito entre os cavacos e as paredes do digestor	
ρ_{chips}	densidade dos cavacos	kg/m^3
ρ_{liquid}	densidade do licor	kg/m^3
v	velocidade superficial no leito fixo	m/s

2.2 MODELO 2 – MODELO CINÉTICO PREDITIVO: APLICAÇÃO NA POLPA DE EUCALIPTO

C_1 e C_2	Quantidade de celuloses tipo 1 e tipo 2 presentes na madeira	% (madeira seca)
D	Difusividade	
E	Quantidade de extractivos presentes na madeira	% (madeira seca)
HS	Concentração de HS-	Moles/l
H_1 e H_2	Quantidade de hemiceluloses tipo 1 e tipo 2 presentes na madeira	% (madeira seca)
L_1 e L_2	Quantidade de ligninas tipo 1 e tipo 2 presentes na madeira	% (madeira seca)
OH	Concentração de OH^-	Moles/l
R	Constante universal dos gases	8.314 J/mol K
t	Tempo	Minutos
T	Temperatura	K

α_i	Difusividade da espécie i	Adimensional
δ	profundidade da frente de penetração	Adimensional
γ	Constante especificada pelo usuário para diferentes cavacos	Adimensional
β	Constante especificada pelo usuário para diferentes cavacos	Adimensional
η	Constante especificada pelo usuário para diferentes cavacos	Adimensional

RESUMO

Estudos demonstram que a expectativa de crescimento do consumo de papel e baixo custo de produção no Brasil pode tornar a indústria de celulose e papel do país a maior geradora de negócios neste ramo internacionalmente. O processo de produção de celulose é complexo, e apesar de muito estudado ainda há muitas lacunas que precisam ser preenchidas. Por outro lado, as redes neurais são uma solução para a modelagem de processos cujo conhecimento é incompleto ou cuja complexidade dificulta a produção de bons resultados através da modelagem determinística.

O objetivo principal desta tese foi a criação de um modelo para a predição do grau de polpação da madeira a partir de dados industriais, utilizando técnicas de modelagem diferentes. O trabalho explorou principalmente a modelagem através de redes neurais, mas também se dedicou à aplicação dos dados industriais a um modelo determinístico e a sua posterior combinação com uma rede neural para o desenvolvimento de um modelo híbrido. A discussão sobre os recursos necessários para a utilização de cada técnica, bem como a comparação entre as diferentes metodologias, suas vantagens e desvantagens, também são assuntos deste trabalho.

A Aracruz Celulose SA, a maior fábrica do Brasil, cedeu os dados para o desenvolvimento do modelo. A avaliação de vários modelos encontrados na literatura selecionou o modelo determinístico a ser utilizado. O modelo neural apresenta uma rede “feedforward” treinada com o algoritmo de retropropagação. Inicialmente, os dados industriais foram avaliados e apenas aqueles que mostraram trazer informação relevante foram utilizados. A estrutura e parâmetros da rede foram otimizados de modo a melhorar a performance do modelo. A combinação do resultado do modelo determinístico com o modelo neural formaram o modelo híbrido.

A fábrica não pode ceder o esperado número de conjuntos de dados para o treinamento da rede neural. No entanto, a qualidade dos dados e o rigor com que foram obtidos, conjuntamente com um cuidadoso trabalho de seleção desses dados possibilitaram a busca de soluções para o sucesso do modelo. A alternativa encontrada

foi a criação de um novo conjunto de dados, obtidos através de uma curva de correlação dos dados industriais.

Apesar do pequeno conjunto de treinamento, a rede neural produziu resultados satisfatórios, onde o erro entre os valores esperados e preditos foram menores que o erro experimental para a determinação de lignina remanescente. O modelo determinístico foi capaz de reproduzir a taxa de deslignificação da celulose no tempo, o que determina o grau de polpação, e portanto foi considerado adequado para a utilização no modelo híbrido. A rede híbrida gerou resultados um pouco melhores que a rede pura, e seu treinamento foi significativamente mais rápido.

Um modelo de computação rápida e que reproduza os resultados esperados pode ser implementado em um sistema supervisor ou de controle para funcionar “on-line”. Os resultados apresentados mostram que tanto as redes neurais puras quanto os modelos híbridos, quando bem treinados e otimizados, atendem a esses requisitos. Mostram também que é possível utilizar variáveis normais de processo, tornando mais viável o desenvolvimento de modelos específicos.

ABSTRACT

Studies show that the expected growth of paper consumption and low production cost in Brazil may turn its pulp and paper industry into the greatest business generator in this field worldwide. The process to produce pulp is complex, and although it has been extensively researched, there are still many questions to be answered. On the other hand, neural networks may be the solution for processes whose knowledge is incomplete or whose complexity makes it harder to obtain good results with first principle models.

The main goal of this thesis was the development of a model to predict pulping degree using industrial data, with different modeling techniques. It explored mainly neural networks technique, but also did it apply the data into a first principle model, which was later used for the development of a hybrid model. The discussion about the resources required for each technique, as well as the different methods, their advantages and limitations, is part of this work.

Aracruz Celulose SA, the largest mill in Brazil, provided the data for the model development. The evaluation of published models set the criteria for the selection of the deterministic model to be used. The neural model presents a feedforward network trained with backpropagation algorithm. After an evaluation of the industrial data, only the variables that showed to carry relevant information were used. The network structure and its parameters were optimized in order to improve model performance. The combination of the result obtained from the deterministic model with a neural network composed the hybrid model.

The mill could not provide the expected amount of training data. However, the quality of the data and the rigorous data collection procedure, as well as the careful data evaluation made it possible to search for solutions for a successful model. The alternative was the production of a new data set formed with points of the correlation curve obtained from industrial data.

Despite the small data set, the neural network produced satisfactory results, where the difference between the expected and predicted values was lower than the experimental error inherent to the lab test for determination of remaining lignin. The first principle model was able to reproduce delignification rate, which determines the pulping degree, and therefore was considered adequate to be used in the hybrid model. The hybrid network results were slightly better than the ones obtained with the pure net and its training was appreciably faster.

A model which is able to reproduce expected results in a timely fashion can be implemented in a control or supervisory system that works on-line. The results showed that when well trained and optimized, both, the pure nets and the hybrid models, fulfill these requirements. They also show that it is possible to use normal process variables, making it more feasible to develop customized models.

CAPÍTULO I

INTRODUÇÃO

I.1 REVISÃO E MOTIVAÇÃO

A celulose é um produto constituído de fibras extraídas de tecidos vegetais, que através de processos químicos ou mecânicos constituem na principal matéria prima do papel. O preço da celulose é ditado por um balanço entre demanda e oferta de amplitude internacional, comportando-se como uma *commodity*. O que faz com que um cliente compre a celulose deste ou daquele fabricante é principalmente a qualidade das fibras e sua capacidade em atender certas características fundamentais para a fabricação do papel em cuja produção será a matéria prima básica. Dentre estas características, uma das mais importantes é a resistência mecânica, que pode ser medida através de vários índices, mais conhecidos no jargão papeleiro como propriedades do papel ou da celulose. Qualquer celulose deve apresentar certo nível de resistência mecânica para pelo menos resistir à tração a que é submetida na máquina de papel, porém este nível varia de acordo com a finalidade do papel produzido. As propriedades ópticas, como cor, opacidade, etc. também são de grande importância para o produtor de papel. Cada tipo de papel vai demandar um conjunto de propriedades diferentes. Por exemplo um papel que será usado para escrever e imprimir deve ser o mais liso possível, com as fibras bem arranjadas, de forma que a tinta não se espalhe em diferentes direções. Deve ainda ser opaco, para que se possa usar de ambos os lados. Se for branco, não pode apresentar um alto índice de reversão de alvura, ou ficará amarelado em pouco tempo. Além disso deve ser resistente à tração imposta pela impressora. Já um papelão que irá embalar eletrodomésticos deve ser bem mais resistente para suportar o peso sem estourar. Por outro lado, não precisa ser

branco nem liso. Se a celulose for destinada à produção de papel higiênico, deseja-se que sua superfície seja tão macia quanto possível.

Apesar de grande parte dessas características serem ditadas pelo tipo de madeira da qual são extraídas as fibras, as condições de operação na produção da celulose também irão alterá-las. As madeiras provenientes de árvores de fibras longas normalmente são usadas na fabricação de celulose para papéis que requerem alta resistência. As de fibra curta, como o eucalipto, são mais utilizadas em papéis que requerem superfícies mais lisas ou mais macias, embora já existam processos que lhes conferem propriedades de resistência, ampliando assim sua aplicabilidade.

O processo de produção de celulose envolve várias etapas onde a madeira está sujeita à ação de agentes físicos e/ou químicos. Dependendo da extensão e qualidade destas ações, a celulose irá desenvolver diferentes propriedades.

Informações não publicadas pela Aracruz Celulose S.A. (situada no estado do Espírito Santo, Brasil) revelaram o grande interesse na indústria papelreira para que sejam desenvolvidos programas capazes de correlacionar as propriedades da celulose com as condições de operação, de maneira a possibilitar a produção de “diferentes” produtos de forma consciente. Atualmente não há modelos determinísticos desenvolvidos para a predição da maioria das propriedades da polpa. Para várias delas, não há um estudo aprofundado sobre o que ou como são afetadas, ou seja, sobre o caráter da relação entre a propriedade e o processo.

Devido à falta de conhecimento profundo sobre a teoria por trás de várias dessas propriedades, o desenvolvimento de um modelo baseado em redes neurais pareceu ser uma alternativa interessante. As redes neurais vêm cada vez mais sendo utilizadas nos mais diversos ramos devido às suas características de aprendizagem. Elas são capazes de mapear complicadas relações de causa-efeito, sem requerer profundo conhecimento teórico, através de um treinamento com vários exemplos desta relação.

Este trabalho tem como objetivo principal o desenvolvimento de uma rede neural preditiva a partir de dados industriais. Como o primeiro passo para o desenvolvimento de um modelo capaz de prever propriedades da polpa celulósica, o caso-estudo considerado é a predição do número kappa no processo de cozimento de cavacos de

madeira, para a produção de celulose. Número kappa é a variável mais utilizada na indústria de celulose e papel para quantificar a extensão das reações ocorridas no processo – simbolizando o teor de lignina, material não desejado, remanescente com as fibras. O teor de lignina apresenta influências na branqueabilidade da polpa, entre outras coisas.

Podem ser encontrados na literatura, e serão apresentados mais adiante neste trabalho, vários modelos para a deslignificação da madeira. O número kappa foi uma das propriedades escolhidas exatamente por possibilitar a comparação entre os resultados de modelos de redes neurais com os modelos determinísticos, baseados nos princípios físico-químicos do processo. Outras propriedades também foram estudadas e serão mencionadas em alguns capítulos desta tese para ilustração, quando conveniente.

A utilização de dados industriais deverá garantir que todas as relações envolvidas no processo sejam embutidas no modelo. Redes desenvolvidas a partir de dados de simuladores normalmente podem ser treinadas sem muita dificuldade, no entanto estarão carregando todas as simplificações e apenas as condições incluídas no modelo do simulador. As variáveis usadas para o treinamento da rede neural serão as mesmas utilizadas posteriormente como dados de entrada em cada simulação. Portanto a utilização dos dados provenientes da fábrica também irá garantir que a informação necessária para que o programa funcione esteja facilmente disponível.

Objetivos secundários, porém não menos importantes, também foram perseguidos, tais como a exploração de parte do universo das redes neurais, ramo importante da Inteligência Artificial e sua aplicação prática para problemas do domínio de atuação de engenheiros químicos. A discussão de como um modelo de rede neural é desenvolvido, seus parâmetros e como eles afetam os resultados, a otimização da rede e a importância da qualidade dos dados utilizados no desenvolvimento do modelo são tópicos explorados dentro deste escopo. Também a comparação desta técnica de modelagem com a tradicional modelagem determinística, incluindo o ponto de vista de aplicação na indústria. Ainda uma pincelada pelos modelos híbridos de modelos neurais e determinísticos combinados para o desenvolvimento de um modelo mais robusto e realista.

No âmbito institucional, é parte do objetivo deste trabalho, a introdução, nesta faculdade, do interesse pela indústria de papel e celulose, de grande importância econômica para o país, e com potencial de crescimento significativo nos próximos anos. Conjuntamente, aumentar a interação Universidade-Indústria também para teses de mestrado, aumentando a contribuição sinérgica de ambas as classes para a produção de pesquisa aplicada com vantagens econômicas para a sociedade como um todo.

I.2 DE CAPÍTULO EM CAPÍTULO

O Capítulo I expõe a motivação para o desenvolvimento de um modelo de redes neurais para a indústria de celulose. O Capítulo II traz uma revisão da literatura sobre o processo de produção de celulose. Inclui uma breve introdução sobre a inserção do Brasil no mercado mundial e sobre a indústria papelreira no país. O conteúdo do capítulo envolve a descrição do processo, dos equipamentos utilizados, uma discussão das variáveis que afetam a extração de fibras individualizadas, processo também chamado polpação da madeira e as principais propriedades objetivo em uma fábrica.

O Capítulo III contém a segunda parte do estudo da literatura, onde as reações da madeira com o licor de cozimento e alguns modelos desenvolvidos em outros trabalhos estão descritos. Neste capítulo fica claro que o processo é composto por fenômenos complexos e inter-relacionados e que não há um consenso geral amplamente estabelecido quanto à cinética das reações ocorridas. A teoria das redes neurais artificiais e sua inserção na Inteligência Artificial também estão incluídas neste capítulo. O texto descreve as características dos modelos neurais e faz uma breve discussão de suas vantagens e limitações. Inclui ainda uma introdução sobre redes híbridas e a comparação destas com as redes puras.

O desenvolvimento da rede neural para a predição do número kappa é o assunto do Capítulo IV, que inicia com as justificativas para a escolha desta técnica de modelagem para este caso específico. O capítulo inclui desde a coleta de dados industriais até a

obtenção do modelo final, descrevendo cada parte do processo de desenvolvimento. Os resultados obtidos são apresentados e discutidos. As variações topológicas da rede e a influência de seus parâmetros no resultado final são também analisados.

Um dos modelos matemáticos encontrados na literatura e uma rede neural foram combinados para o desenvolvimento de um modelo híbrido. A primeira parte do Capítulo V é dedicada à análise dos resultados obtidos com o modelo matemático e a influência das variáveis de entrada no resultado final. A segunda parte descreve como os modelos foram combinados, discute a escolha deste tipo de composição híbrida e os resultados obtidos. No final, é feita a comparação entre as técnicas de modelagem, suas aplicações e resultados

O Capítulo VI é o capítulo destinado à conclusão da tese, onde está relatado como os objetivos foram atingidos. Por fim, uma série de sugestões para futuros trabalhos é apresentada e discutida.

CAPÍTULO II

REVISÃO DA LITERATURA E CONCEITOS BÁSICOS - PARTE 1: FABRICAÇÃO DE CELULOSE

II.1 INTRODUÇÃO

Neste capítulo ênfase será dada para a exposição dos conceitos básicos fundamentais e de revisão da literatura sobre o processo de polpação da madeira para a produção da celulose, caso-estudo considerado neste trabalho. Inclui uma visão geral da indústria de celulose, com um maior enfoque no cozimento da madeira, desde a descrição do processo e equipamentos utilizados, até variáveis e objetivos do processo. Os conceitos aqui descritos correspondem ao mínimo de informação necessária para o planejamento do desenvolvimento da rede neural para a predição do número kappa. Este nível de conhecimento permitiu a análise da complexidade do sistema e a definição das variáveis importantes para o treinamento da rede.

No final do capítulo, há uma secção dedicada às propriedades das fibras e suas relações com a madeira e o processo.

II.2 CENÁRIO DA INDÚSTRIA DE PAPEL E CELULOSE NO BRASIL

A celulose é a matéria prima para a fabricação de papel. A indústria de celulose no mundo movimenta US\$200 bilhões anualmente, tendo os Estados Unidos como maior produtor, seguido do Canadá e Suécia. O Brasil é o sexto maior produtor de celulose mundial (DENCE e REEVE, 1996), produzindo celulose, papel e papelão, principalmente na região sul e sudeste do país.

A maior parte da celulose brasileira é produzida a partir de espécies de eucalipto, matéria prima de fibra curta, que é utilizada para a fabricação de papéis absorventes e papéis para imprimir e escrever, entre outros. O pinheiro também é muito utilizado na região sul do país, produzindo celulose destinada principalmente à indústria de papelão. O babaçu, bambu, sisal, cana e outros são responsáveis por uma pequena contribuição na produção brasileira.

A tecnologia em papel e celulose é desenvolvida principalmente na Escandinávia e na América do Norte, e a Austrália vem aumentando sua contribuição, o que pode ser comprovado pelo aumento de publicações. No Brasil, as pesquisas florestais estão em fase avançada, com estudos genéticos que levaram ao desenvolvimento de clones, para a obtenção de uma matéria prima mais uniforme e resistente, dotada de características desejáveis para a produção de papel.

O ciclo de crescimento das árvores no Brasil é consideravelmente mais rápido que nos países líderes de produção de celulose, o que possibilita uma maior produtividade com uma menor área plantada. As principais fábricas do país têm sido modernizadas com frequência, apresentando produtos com um nível técnico e qualidade competitivas no mercado mundial. O consumo de papel *per capita*, no entanto, está um tanto distante dos índices encontrados nos países desenvolvidos, apenas 27 kg/ano no Brasil, contra 284 kg/ano nos Estados Unidos (LIMA et al., 1988).

Estes fatores revelam a importância e o potencial de crescimento da indústria de celulose e papel no Brasil

adesiva, que dá rigidez à madeira. Durante o processo de fabricação da polpa de celulose, o objetivo é remover a lignina e separar as fibras de carboidratos. Por esse motivo, o processo é conhecido como deslignificação

As principais fontes de celulose para o mercado são os troncos das árvores, que podem ser classificadas como resinosas ou coníferas (*softwood*) e folhosas (*hardwood*).

As folhosas têm fibras curtas e alta densidade, fornecendo celulose com propriedades físicas desejáveis para a produção de papéis finos, de impressão e escrita de boa qualidade (PROCESSO BÁSICO - ARACRUZ CELULOSE, 1988). Apresentam ainda alta taxa de crescimento, e portanto um menor ciclo de corte (tempo entre o plantio e o corte). Além disso, a árvore cortada cresce novamente, permitindo de 4 a 5 cortes (PROCESSO BÁSICO - ARACRUZ CELULOSE, 1988). O eucalipto, a bétula e o *aspen* são exemplos de folhosas.

As coníferas possuem fibras longas que conferem maior resistência mecânica à polpa obtida. No entanto, têm uma densidade menor, um ciclo de corte maior e cada árvore permite apenas um corte, além de conferir um rendimento mais baixo (PROCESSO BÁSICO - ARACRUZ CELULOSE, 1988). São utilizadas principalmente para papéis cujo uso requeira maior resistência, como papelão, papel de embrulho, etc. Os diversos tipos de pinheiros são as representantes mais comuns das resinosas.

O tipo, composição e a idade da madeira, e até mesmo o clima e a qualidade do solo onde a árvore se desenvolve, têm importante influência na composição da madeira e características das fibras e, conseqüentemente nas propriedades do papel. Essas variáveis são consideradas antecipadamente, ainda na fase de concepção do projeto da fábrica e na escolha do processo adequado para que se obtenha o maior rendimento e qualidade.

II.3.2 POLPAÇÃO

A separação das fibras da madeira é conhecida como polpação, e pode ser feita através de vários processos que necessitam de energia, e aplicam diferentes princípios e

agentes para este fim. Normalmente estes processos estão divididos entre mecânicos e químicos.

Os processos mecânicos ou de alto rendimento apresentam rendimentos superiores a 80% e englobam o desfibramento clássico em pedra e suas variações. A separação da fibra se dá através do atrito do desfibrador com a madeira e também dos ciclos de compressão e descompressão, que ocorrem devido à passagem alternada de pontas e cavidades da superfície da pedra em um ponto da madeira. O calor e a umidade têm influência na degradação da lignina. A pasta resultante é constituída de fibras intactas, feixes de fibras e fibras danificadas (ASSUMPCÃO et al., 1988), e pode ser usada para a fabricação de vários tipos de papel e papelão.

Os processos semiquímicos apresentam um rendimento menor que os mecânicos, entre 65% e 85%, e envolvem a remoção de parte da lignina e hemicelulose, podendo apresentar rendimentos ainda mais baixos se houver exposição a condições mais extremas. Com ação mecânica de menor intensidade, as fibras são menos danificadas que no caso anterior (ASSUMPCÃO et al., 1988), portanto permitindo a produção de uma celulose de maior resistência. Também podem ser utilizados para a produção de papelão.

Os processos químicos apresentam rendimento ainda menor, pois removem maior quantidade de lignina. A polpação se dá através do tratamento químico da madeira a altas temperaturas. Os dois processos mais comuns são o sulfito, cujos reagentes ativos são os íons sulfito e sulfito de hidrogênio, e o processo sulfato, também chamado Kraft, onde o hidróxido e o sulfato de sódio são os reagentes ativos. No primeiro caso, o licor de cozimento apresenta um pH baixo, enquanto que no processo Kraft, a reação se dá em pH alcalino. A grande maioria das pastas químicas no Brasil são produzidas através do processo Kraft. Os processos químicos produzem uma celulose mais deslignificada, com branqueabilidade significativamente mais elevada que os outros processos, que por este motivo, pode ser utilizada para a produção de vários tipos de papel, entre eles os papéis finos, papéis fotográficos, papéis absorventes, etc. Nestes processos, o licor usado na polpação é recuperado e o vapor gerado é, muitas vezes, suficiente para suprir todas as necessidades de energia na fábrica.

III.3 PROCESSO DE PRODUÇÃO DA CELULOSE

A matéria prima para a fabricação de celulose é a fibra vegetal, retirada principalmente das árvores, mas também de outros tipos de planta. A madeira é cortada em forma de cavacos e alimenta o digestor, um vaso pressurizado onde as fibras de celulose são separadas dos outros componentes da madeira através de reações químicas ou ação mecânica. Na saída do digestor, a celulose é uma massa de cor marrom e apresenta ainda muitas impurezas, como pedaços de madeira mal cozidos, grãos de areia e outros. Essas impurezas devem ser separadas da celulose através de uma série de operações de filtração e lavagem agrupadas em um processo chamado depuração. A depuração é uma operação de extrema importância em termos econômicos e de qualidade. Certas “impurezas” aumentam o consumo de reagentes no branqueamento e outras alteram as reações, alterando conseqüentemente a qualidade da polpa. Uma vez depurada, a celulose está pronta para ser branqueada. A primeira fase de branqueamento é chamada pré branqueamento com oxigênio (pré-O₂), onde a polpa reage com oxigênio em uma solução alcalina oxidada. O branqueamento propriamente dito é normalmente uma seqüência de estágios ácidos e alcalinos alternados, que variam, em número e reagentes aplicados, de fábrica para fábrica. A pasta branqueada passa então por mais um processo de lavagem e está pronta para a secagem, caso seja o produto final da fábrica, ou para a preparação da massa que alimentará a máquina de papel. Em ambos os casos a celulose estará sujeita à ação de vapor, vácuo e prensas.

Desde o cozimento até a saída da máquina de papel, as fibras da madeira recebem tratamento químico e mecânico que exercem importante efeito sobre as propriedades finais do papel.

A Figura 2.1 representa uma linha de fibras genérica, com a alimentação e impregnação de cavacos, o cozimento, a lavagem, os reatores de oxigênio, a depuração e uma seqüência de quatro estágios de branqueamento.

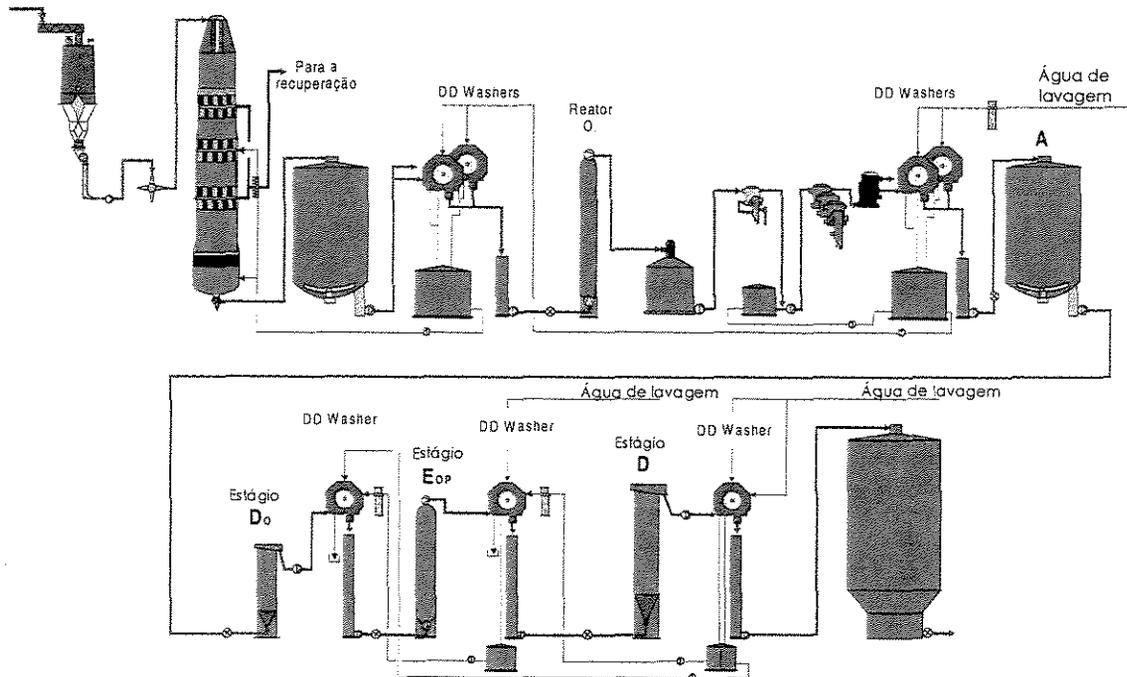


Figura 2.1 – Figura esquemática de uma linha de fibras - Inicia com um sistema de alimentação de cavacos, seguido do digestor e uma torre de estocagem. A celulose passa então por filtros lavadores e vai para o pré branqueamento e a depuração. Após mais uma operação de lavagem, a polpa está pronta para ser branqueada em uma seqüência de 4 estágios, após a qual será estocada.

II.3.1 MATÉRIA PRIMA

A madeira é composta basicamente por carboidratos (celulose e hemicelulose), lignina, extrativos e água, em quantidades e perfis de composição que variam de acordo com a espécie e a parte do vegetal.

A celulose é um polímero não ramificado cujo monômero é a glicose. Por serem não ramificadas, as cadeias de celulose se agrupam de modo a formar uma estrutura cristalina. Grupos de cadeias de celulose são chamados fibrilas, e juntos com as cadeias dos dois outros polímeros – a hemicelulose e a lignina, formam as paredes das fibras. As hemiceluloses por sua vez, são cadeias ramificadas e a lignina é uma teia tridimensional (KVAERNER PULPING TECHNOLOGIES AB, 1995). A lignina é uma substância

Neste trabalho, a atenção será direcionada ao processo Kraft.

II.3.3 LICOR NO PROCESSO KRAFT

Licor de cozimento é o nome utilizado para a solução de reagentes em água que é adicionada à madeira no processo da polpação. É produzido a partir do licor branco, que é constituído principalmente de hidróxido de sódio (NaOH) e sulfeto de sódio (Na₂S), mas também apresenta outros sais de sódio. Pode-se adicionar ao licor branco, licor de cozimento anterior recuperado, que contém constituintes da madeira dissolvidos e reagentes não consumidos. A seguinte terminologia é utilizada:

Reagente total: todos os sais de sódio expressos em g/l de Na₂O ou NaOH.

Álcali total: NaOH + Na₂S + Na₂CO₃ + ½ Na₂SO₃ (expressos em g/l de Na₂O ou NaOH)

Álcali ativo: NaOH + Na₂S (expressos em g/l de Na₂O ou NaOH)

Álcali efetivo: NaOH + ½ Na₂S (expressos em g/l de Na₂O ou NaOH)

Sulfidez: % de Na₂S no álcali ativo (expressos em g/l de Na₂O ou NaOH)

Licor branco: Licor de cozimento novo, isto é sem carga de espécies orgânicas dissolvidas.

Licor preto: Licor resultante extraído do digestor. Contém alta carga de espécies orgânicas dissolvidas devido às reações ocorridas na polpação. Este é o licor que irá para a área de recuperação para produção de energia e depois para a planta de recuperação do licor branco.

Licor de recirculação: É o licor que leva os cavacos do alimentador de alta pressão para o topo do digestor. É retirado no separador de topo, retornando ao alimentador de alta pressão e reiniciando o ciclo.

Licor de lavagem/Licor de resfriamento: As etapas de lavagem pós-cozimento acontecem em contra-corrente. A celulose segue em frente em direção ao

UNICAMP

branqueamento, cada vez mais limpa. O filtrado de cada estágio volta como água de lavagem para o estágio anterior, em direção ao digestor, cada vez mais “contaminado”, isto é, mais concentrado em matéria orgânica. O filtrado do lavador que segue o digestor é utilizado na zona de lavagem e como diluição na saída do reator.

II.3.4 DIGESTORES

O equipamento onde ocorre a polpação química ou cozimento é chamado digestor. Constitui-se de um vaso de pressão, onde a madeira picada (cavacos) recebe o licor de cozimento, e são submetidos à pressão e temperatura. Há inicialmente um período de aquecimento em que se completa a impregnação dos cavacos com o licor, e então um período à temperatura de cozimento, onde ocorre a deslignificação, tradicionalmente por volta dos 170°C (nos processos modernos, a temperatura já é bem mais baixa, em alguns casos abaixo de 150°C). A alta temperatura demanda elevada pressão para que o licor não vaporize (ASSUMPCÃO et al., 1988).

Existem vários tipos de digestores, que podem funcionar em regime contínuo ou batelada.

Os digestores em batelada estão em funcionamento por mais de um século. Neste processo, os cavacos e os reagentes são adicionados ao reator no início do processo. O reator é fechado e aquecido. Após um tempo determinado, o vaso é aberto e o licor e a celulose são removidos e o vaso preparado para ser recarregado. Normalmente em fábricas com regime em batelada existe mais de um digestor, de forma que pelo menos um deles estará sempre produzindo.

Os digestores contínuos mais utilizados são os de tecnologia Kamyr (Atualmente Ahlstrom e Kvaerner), cuja operação será descrita mais detalhadamente na próxima seção. É dividido em zonas, onde os cavacos são impregnados, aquecidos até a temperatura de cozimento, cozidos e lavados. Apresenta vantagens em relação aos digestores descontínuos, entre elas: para altas capacidades possuem um custo menor, maior economia de vapor, qualidade do produto superior, rendimentos mais altos, menor

consumo de produtos químicos, demandando portanto menor capacidade de recuperação, menor corrosão e maior controle das emissões atmosféricas. Por outro lado, o custo de manutenção é maior, há maior probabilidade de problemas com incrustação, e desperdiça muita matéria prima na partida, até atingir regime estacionário (ASSUMPÇÃO et. al., 1988).

Deste ponto em diante, qualquer menção a digestores contínuos deve ser entendida como digestores do tipo Kamyr.

II.3.5 DIGESTOR CONTÍNUO

Um esquema representativo deste tipo de digestor é apresentado na Figura 2.2.

Os cavacos entram no processo através do silo de cavacos, onde podem ou não receber vapor para iniciar o processo de impregnação, uma preparação para receber o licor de cozimento.

Os cavacos provenientes do silo de cavacos passam por um **medidor rotativo**, cuja taxa de rotação dita a produção da fábrica e é usada portanto, como referência para a adição de reagentes. Do medidor rotativo, os cavacos entram no **alimentador de baixa pressão**, que promove uma selagem entre a pressão atmosférica e a pressão no pré-aquecedor, o equipamento subsequente. No **pré-aquecedor**, os cavacos recebem vapor de baixa pressão (aprox. 3,5 kgf/cm²) e vapor saturado (PROCESSO BÁSICO - ARACRUZ CELULOSE, 1988). O tempo e a temperatura devem ser controlados para evitar a pré hidrólise dos carboidratos (principalmente celulose e hemicelulose), o que causaria perda de rendimento e propriedades físicas. A função dos pré aquecedores é retirar o ar dos cavacos, eliminar NCGs (gases não condensáveis) e terpenos, para que estes não interfiram na transferência de calor e difusão, uniformizar a umidade e aumentar a densidade, para melhorar a impregnação, o que vai promover um cozimento mais uniforme (ASSUMPÇÃO et al., 1988). Processos mais modernos já incluem o pré aquecimento no silo de cavacos, eliminando a necessidade dos pré aquecedores.

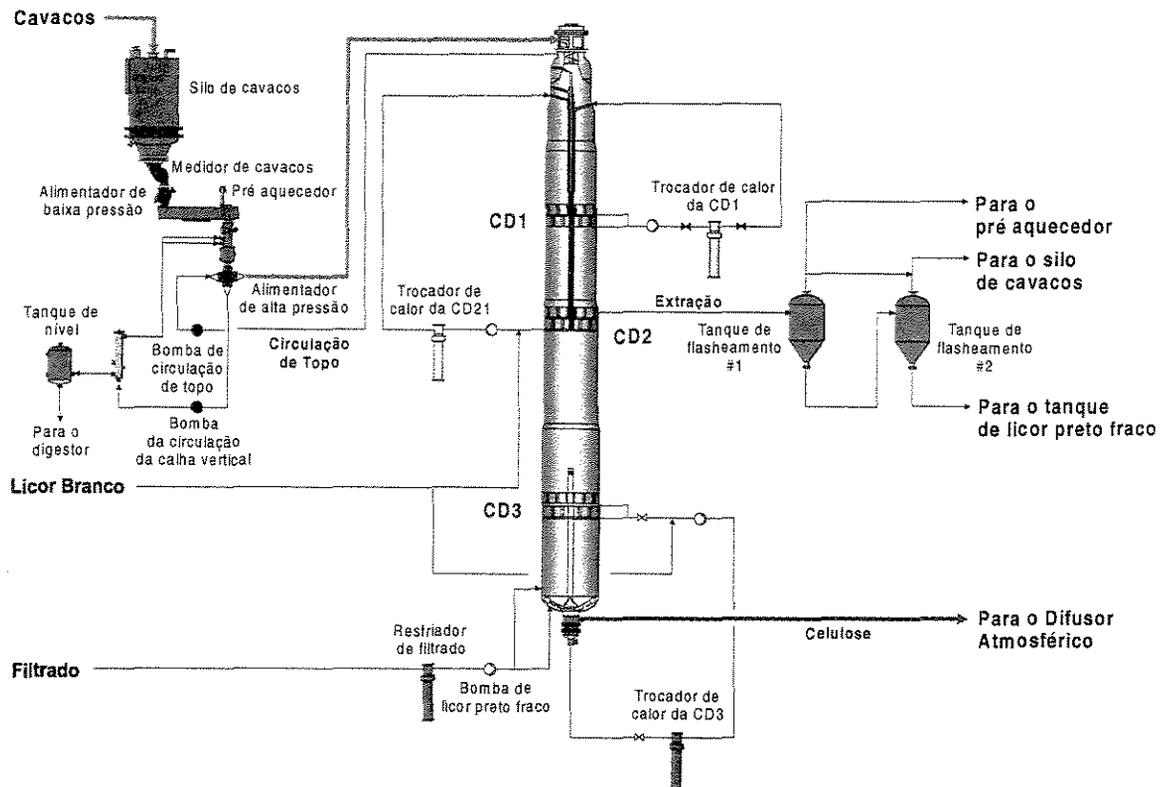


Figura 2.2 – Sistema de alimentação de cavacos convencional e digestor contínuo

Os cavacos alimentam a **calha vertical**, parcialmente cheia com licor branco, onde são impregnados com álcali ativo e então são enviados para o **alimentador de alta pressão**, onde recebem licor branco e licor de recirculação (PROCESSO BÁSICO - ARACRUZ CELULOSE, 1988). Este também é um equipamento rotativo, e ao atingir a posição horizontal, o licor de recirculação carrega os cavacos para o **topo do digestor**, passando para uma pressão de aproximadamente 12 kgf/cm^2 (este valor varia de fábrica para fábrica, dependendo da temperatura no vaso), o que evita a ebulição do licor às altas temperaturas. O excesso de licor é retirado do alimentador de alta pressão através de uma bomba, passa por um separador de areia e retorna à calha vertical (PROCESSO - BAHIA SUL CELULOSE, 1993).

Em alguns casos, os cavacos vão primeiro para o topo de um vaso de impregnação, que está ligado ao digestor e portanto sujeito às mesmas condições de pressão. No caso aqui descrito, não há vaso de impregnação e portanto o sistema é chamado digestor de um vaso.

No topo do digestor, à medida que os cavacos vão descendo, o licor é extraído pelas ranhuras nas placas da peneira do **separador de topo**, voltando para a recirculação. Através de bombas de alta pressão o licor branco é adicionado, sendo rapidamente absorvido pelos cavacos pré condicionados.

Os digestores contínuos podem ser classificados como digestores hidráulicos, onde o vaso é completamente cheio com licor e portanto o nível de cavacos está abaixo do nível de licor; ou digestores fase vapor, onde o nível de cavacos está acima do nível de licor e há uma fase vapor no topo do vaso.

No digestor de um vaso já cheio de licor, os cavacos passam para a **zona de impregnação**, onde ocorre a difusão dos reagentes e a taxa de reação é baixa. Ali permanecem por 45 a 60 min, à temperatura de 100 a 125°C (ASSUMPÇÃO et al., 1988). Vapor de média pressão pode ser injetado de forma direta para proporcionar a elevação da temperatura (PROCESSO BÁSICO - ARACRUZ CELULOSE, 1988, PROCESSO - BAHIA SUL CELULOSE, 1993), no caso do digestor fase vapor.

Os cavacos continuam descendo, entrando na **zona de aquecimento**, onde são aquecidos até aproximadamente 165°C. O licor é introduzido no início da zona através de um tubo central e retirado através de peneiras nas paredes, sendo então aquecido e retornando para o digestor através do tubo central descrito acima, a fim de corrigir a temperatura nesta zona.

Um outro tubo central introduz licor no final da zona de aquecimento, onde a temperatura é elevada devido ao calor da reação. O tempo de reação nesta zona é de aproximadamente 90 minutos. No final desta zona, há um conjunto de peneiras de extração cuja finalidade é retirar o licor preto fraco, resultante da reação. Parte da água contida neste licor é recuperada por flasheamento (liberação de vapor que será utilizado no pré-aquecedor). O licor agora mais concentrado, chamado licor preto forte, é mandado para os evaporadores, entrando no sistema de recuperação.

Logo abaixo das peneiras de extração, existe um outro conjunto de peneiras de circulação de licor. Este licor recebe certa dose de licor branco e é enviado a um trocador de calor, retornando ao digestor na **zona de homogeneização** (PROCESSO - BAHIA SUL CELULOSE, 1993). A adição do licor branco garante um pH alto no final do cozimento,

e com isso, evita a reprecipitação dos carboidratos sobre as fibras, promove maior alvura e possibilita menor consumo de reagentes, porém pode redundar em menor rendimento e menor retenção de xilanas, que são carboidratos (ASSUMPÇÃO et al., 1988).

Os cavacos cozidos, entram então na **zona de lavagem a alta temperatura** (lavagem Hi-Heat[®]), onde recebem licor de lavagem (que possui cerca de 5% de sólidos) a 80°C, em contra corrente, injetado no fundo do digestor. Ao subir em direção contrária ao fluxo dos cavacos, o licor de lavagem vai arrastando o licor de cozimento, e assim sua concentração é elevada para cerca de 15%, e sua temperatura, para aproximadamente 140°C (PROCESSO - BAHIA SUL CELULOSE, 1993). A alta temperatura nesta zona aumenta a eficiência de lavagem, pois aumenta a taxa de difusão. O tempo de retenção na lavagem varia de 1 a 3 horas, dependendo do resultado desejado.

Da zona de lavagem, os cavacos vão para a **zona de resfriamento**, onde são resfriados até 75 a 95°C, diluídos e submetidos à lavagem final, a 60-85°C. O licor de resfriamento é retirado na parte superior e se torna o licor de lavagem. A polpa é raspada no fundo do digestor de forma que a massa do centro seja desviada para fora, e sai com a consistência de aproximadamente 11%. No trajeto até o **tanque de descarga**, existe uma unidade de descarga e válvulas de controle para aliviar a pressão de 12 kgf/cm² para pressão atmosférica (ASSUMPÇÃO et al., 1988).

O calor do licor negro é recuperado para alimentar o silo de cavacos e/ou o pré aquecedor e para a preparação de água quente. Os gases não condensáveis, podem ser usados como combustível no forno de cal na unidade de recuperação de licor branco (ASSUMPÇÃO et al., 1988; PROCESSO BÁSICO - ARACRUZ CELULOSE, 1988) ou ir para um incinerador (PROCESSO - BAHIA SUL CELULOSE, 1993).

Esta é uma descrição genérica de um digestor contínuo, porém, há variações nos processos de fabricação normalmente encontrados na indústria. As variações impactam principalmente o perfil de concentração de álcali no digestor, e conseqüentemente a taxa de reação nas diferentes zonas. Com isso, o perfil de temperatura também é diferente. O objetivos dos processos mais modernos, como o cozimento modificado (MCC), o cozimento modificado estendido (EMCC) e o cozimento Lo-Solids[®], é a obtenção de polpas resistentes, porém com um menor teor de lignina residual. As variações incluem

diferentes pontos de adição de reagentes e licor de lavagem (para diluição), extração de licor e também incluem zonas com cozimento em contra-corrente, onde os cavacos continuam descendo, porém o licor flui de forma ascendente.

II.4 POLÍTICAS DE OPERAÇÃO

Independente do tipo de digestor, as principais variáveis que apresentam influência no processo são as mesmas.

O tipo e características da madeira usada são fatores pré determinantes das propriedades da polpa de celulose obtida no final do processo. Embora uma vez no processo as características da madeira não possam mais ser mudadas, algumas empresas investem em pesquisa para utilizar árvores mais semelhantes, permitindo uma menor variação destas propriedades e portanto possibilitando condições de processo mais constantes e qualidade do produto mais uniforme.

As condições de operação e as propriedades da celulose esperadas no final do processo são determinadas de acordo com a espécie da madeira. As coníferas, que possuem fibras longas, produzirão um papel com maior resistência mecânica. As folhosas, no entanto, sofrem um deslignificação mais rápida, permitindo menor tempo, temperatura, carga de reagentes e apresentando menor teor de lignina no final do cozimento (ASSUMPCÃO et al., 1988), o que demandará menor quantidade de reagentes no branqueamento.

Árvores com maior **densidade básica** apresentam vantagens para empresas que adquirem madeira por volume, e realizam cozimentos com rendimento maior por ciclo. Entretanto, madeiras menos densas são deslignificadas mais rapidamente. A parte da árvore de onde se retira a madeira possui diferentes **épocas de crescimento**, o que significa que têm características físicas e químicas diferentes (ASSUMPCÃO et al., 1988).

O grau de deterioração da madeira durante a estocagem vai variar de acordo com a espécie e condições climáticas da região, alterando suas propriedades físicas e químicas. O **tempo de estocagem** é uma variável que pode ser manipulada para a obtenção de resultados desejáveis. Apesar de causar reduções no rendimento, em algumas fábricas existem políticas de maiores tempos de estocagem a fim de reduzir a quantidade de extrativos presentes na madeira. Existem também, estudos para aumentar o consumo de extrativos e/ou lignina durante a estocagem, através de enzimas.

Entre as variáveis associadas à madeira, a **dimensão dos cavacos** é a mais manipulável e importante, pois é fundamental para a qualidade da impregnação do licor. Se a impregnação for incompleta, ao final do cozimento haverá mais rejeitos, o que provoca uma queda de rendimento e maior teor de lignina, o que reduz a branqueabilidade. Além disso, produzirá uma polpa com propriedades físicas inferiores (GUSTAFSON et al., 1989). Já a penetração completa do licor nos cavacos reduz o tempo de cozimento, inibe reações concorrentes e indesejáveis e aumenta a uniformidade da polpa. Em pHs altos, a taxa de difusão na madeira é semelhante em todas as direções. Como a espessura corresponde à menor distância, esta é a dimensão mais crítica para que se complete a impregnação e as ocorram reações. O comprimento apresenta influências mínimas, e a largura é insignificante para a deslignificação (GUSTAFSON et al., 1989). A escolha da dimensão deve ter como objetivo impregnar totalmente o cavaco com a quantidade necessária de reagentes e garantir uma taxa de reação constante. Cavacos menores produzem menos rejeitos e portanto maior rendimento, mas se não houver controle apropriado do tempo de residência ou temperatura, podem ficar super cozidos, o que reduz a viscosidade da polpa e a sua resistência. Cavacos muito grandes podem não ser totalmente cozidos, aumentando a quantidade de rejeitos e portanto diminuindo rendimento, ou demandam um tempo maior de cozimento, havendo então maior degradação nas extremidades, e portanto, uma grande desuniformidade na polpa final. Alguns estudos indicam que as propriedades físicas são otimizadas para cavacos com espessura entre 4 e 6 mm; na indústria, normalmente são aceitos cavacos entre 2 e 8 mm. Em madeiras mais porosas, como o eucalipto, a impregnação é mais rápida (MIRAMS e NGUYEN, 1994).

Em suma, o mais importante na política de operação dos digestores, com respeito à matéria prima, é que a madeira tenha características tão uniformes quanto possível, para permitir condições de operação mais constantes, facilitando o controle, e principalmente, para produzir uma pasta de celulose uniforme.

Além da qualidade da madeira e dos cavacos, existem várias outras variáveis de processo mais facilmente manipuláveis, que podem ser otimizadas a fim de se obter maior rendimento, melhor qualidade e maior economia no processo. Como foi citado anteriormente, o objetivo do cozimento é a minimização do teor de rejeitos e o maior grau de deslignificação possível, medido em termos de número kappa.

A **carga de álcali ativo** deve ser determinada a fim de prover a quantidade de reagentes necessária para que se atinja o ponto de desfibramento. Um aumento na carga de álcali provoca um aumento na taxa de reação. Não é vantagem ter este valor alto demais, pois após determinado ponto, os ganhos em kappa (deslignificação), rendimento e propriedades físicas não se alteram significativamente, uma vez que a taxa de dissolução de hemiceluloses também aumenta, e a demanda de energia no refino (durante a fabricação do papel) é elevada.

É necessário manter uma certa relação entre a quantidade de licor e de madeira, para garantir a completa impregnação dos cavacos. Por outro lado, quanto menor a **relação licor/madeira**, menor será o consumo de vapor para manter a temperatura no digestor e maior será a concentração de sólidos no licor preto, diminuindo a carga nos evaporadores.

Por outro lado há uma **concentração** mínima de reagentes no licor branco para que o pH seja mantido acima de um limite, a fim de se evitar a redeposição de lignina e hemicelulose sobre as fibras. Maiores concentrações demandam menos vapor nos evaporadores e diminuem a eficiência na caustificação (ASSUMPÇÃO et al., 1988).

O sulfeto (NaS^-) facilita a dissolução da lignina e portanto, um aumento na **sulfidez** aumenta a taxa de deslignificação, o rendimento e a qualidade da celulose. Porém após certo patamar, os benefícios não são significativos, e o odor característico de plantas de celulose se intensifica, podendo causar problemas com comunidades vizinhas.

O **tempo** e a **temperatura** são variáveis importantes no controle do digestor, sendo inversamente manipuladas, isto é, se a temperatura for aumentada, o tempo deve ser diminuído, para que se obtenha um mesmo resultado. Trata-se portanto de um processo de otimização. Este é o sentido do Fator-H usado para controle em vários digestores. É um fator que leva em conta o tempo e a temperatura em uma única variável (representa a área sob a curva da velocidade relativa da reação versus o tempo de cozimento). Se todas as outras variáveis permanecerem iguais, pode-se modificar o tempo e a temperatura obtendo-se pastas semelhantes, dado que o Fator-H permaneça o mesmo. Um aumento na temperatura causa um significativo aumento na taxa de reação (aumento de 10°C dobra a taxa de reação). É, entretanto, mais aconselhável utilizar temperaturas mais baixas e tempos maiores para minimizar a degradação da celulose. O tempo deve ser limitado pela produção desejada e custos de capital e produção. Porém, o conceito de Fator-H pode ser aplicado apenas para a mesma madeira e condições semelhantes (SALTIN, 1992; MIRAMS e NGUYEN, 1994).

II.5 OBJETIVOS E CONTROLE DA POLPAÇÃO QUÍMICA

As formas de controle do processo e da qualidade da celulose mais importantes para o fabricante de celulose são o número kappa, a resistência mecânica e o rendimento.

O número kappa, foco principal deste trabalho, mede o grau de deslignificação da polpa, ou seja, a concentração de lignina remanescente. Quanto menor o número kappa, menor a quantidade de reagentes e energia necessários nas etapas de branqueamento. Deve-se deixar claro neste ponto que o branqueamento é um processo mais caro e poluente que o cozimento. Pastas com alto kappa requerem reagentes mais agressivos ao meio ambiente no branqueamento (KVAERNER PULPING TECHNOLOGIES AB, 1995).

No entanto, é impossível remover a lignina sem redução de rendimento e sem danificar as fibras. O rendimento descreve a quantidade de celulose obtida a partir da

quantidade de madeira introduzida no processo, sendo portanto, de grande importância econômica.

O grau de despolimerização da celulose é medido pela viscosidade da polpa e é proporcional a algumas das propriedades de resistência mecânica. Independente do tipo de papel, sempre se deseja obter fibras com o maior nível de resistência possível, para que haja mínimas perdas durante a fabricação do papel em si, devido à operação da máquina.

O grande objetivo numa fábrica de celulose é portanto, cozinhar a madeira para obter o menor kappa possível, sem prejudicar significativamente o rendimento e/ou a resistência do produto.

II.5.1 VISCOSIDADE E PROPRIEDADES

A viscosidade é uma medida indireta do grau de polimerização (DP) dos polissacarídeos, sendo usada como um indicador de potencial de propriedades de resistência. Na teoria, quanto maior o grau de polimerização, maior a resistência da fibra. Mas essas correlações não são universais: polpas com tratamentos e conformações de fibra diferentes apresentam relações diferentes. Pesquisas mostraram que estendendo alguns tratamentos, como a polpação Kraft, o DP é reduzido, sem que se reduza as propriedades de resistência das fibras. Outros estudos mostraram a grande degradação da celulose (com ácido), sem que houvesse grande queda de DP. Acredita-se que este fato pode ser explicado pelo ataque às fibras de maneira mais uniforme nos novos processos de polpação, permitindo que o DP médio caia antes de haver um efeito na resistência das fibras. Já o tratamento ácido ataca a celulose nos pontos de maior acessibilidade, pois a fibra não está inchada como nos estágios alcalinos, e essa degradação localizada cria pontos fracos, sem derrubar o DP. Provavelmente os ataques localizados surjam em pontos irregulares, como nós oriundos de madeira de compressão.

Tradicionalmente a viscosidade tem sido usada para controle de processo, garantindo que as propriedades de resistência da celulose estejam acima de um certo

nível. O valor mínimo de viscosidade depende do processo e do mercado. Madeira e processos podem afetar muito a viscosidade. No entanto não se deve comparar viscosidade de fábricas diferentes para a análise das propriedades de resistência de suas polpas.

A visão moderna da viscosidade utiliza a uniformidade da viscosidade como um indicador da boa operação da fábrica de um ponto de vista gerencial, e não apenas para parâmetro para as propriedades das fibras.

II.5.2 NÚMERO KAPPA

Como foi dito antes, o número kappa é utilizado no controle do cozimento para a determinação do grau de deslignificação. Os reagentes do branqueamento serão dosados com base neste parâmetro.

O grau de deslignificação também está relacionado às propriedades das fibras, uma vez que o menor teor de lignina irá promover uma maior ligação entre as fibras, aumentando a resistência da folha de papel.

O número kappa é, na verdade, o volume (em ml) de solução de permanganato de potássio 0.1N consumido por um grama de polpa de celulose seca, sob condições padronizadas. Este volume é então corrigido para um consumo de 50% do permanganato adicionado durante o teste.

O procedimento do teste pode ser encontrado nos livros de métodos das diversas Associações de Papel e Celulose internacionais, tais como a Appita (método Appita P201) da Austrália e Nova Zelândia, SCAN (método SCAN C1) na Escandinávia, Tappi (método T-236 cm-35) dos Estados Unidos e CPPA (método CPPA G.18) do Canadá, entre outros.

A repetibilidade do teste é alta para polpas semi-branqueadas, mas diminui com kappas mais altos (maior quantidade de lignina remanescente na polpa). Para polpas com kappa abaixo de 20, a repetibilidade será por volta de 0.2 unidades de kappa, e para polpas com kappa muito baixo (abaixo de 2), a repetibilidade começa a diminuir

novamente. A repetibilidade do teste é em média 3.8% para kappas entre 2.5 e 20 (TAPPI TEST METHODS, 1994)

II.6 PROPRIEDADES DAS FIBRAS BRANQUEADAS

II.6.1 INTRODUÇÃO

Cada tipo de papel possui uma finalidade específica, e para que cada papel melhor sirva a seus propósitos, deve apresentar certas características também específicas. Estas propriedades podem ser medidas e/ou controladas através de testes físicos e ópticos que caracterizam as propriedades das fibras que compõem a folha de papel. Tais propriedades variam com a matéria prima e com os diversos tratamentos físicos e químicos aos quais as fibras estão sujeitas no processo de produção da celulose e do papel. Além de associada às características do papel, a importância das propriedades pode também estar relacionada às demandas de processo.

A descrição das propriedades das fibras mais comuns pode ser encontrada no Apêndice 1.

II.6.2 REFINO

Para a fabricação do papel, a celulose deve receber tratamentos especiais, que variam de acordo com o tipo de papel produzido. Esses tratamentos irão desenvolver características desejadas a partir da celulose recebida. Independente do tipo de papel, a celulose passará por uma etapa de tratamento mecânico, o refino.

Para a realização dos testes de quantificação das propriedades das fibras nas fábricas de celulose, o processo de tratamento mecânico é realizado em laboratório num equipamento chamado moinho, que simula o refino.

O tratamento mecânico altera as fibras de modo a desenvolver as propriedades desejáveis. Este processo é complexo e ainda tem muitos pontos desconhecidos, é quantificado e controlado de acordo com a utilização de energia, e o estado da polpa não é necessariamente qualificado ou utilizado para controle.

O refino pode ser uma referência para a comparação de polpas. Pode-se comparar duas celuloses diferentes para um mesmo período de refino, ou seja, pode-se avaliar as diferenças entre duas polpas que receberam um mesmo tratamento mecânico. Ou pode-se comparar duas polpas com uma mesma taxa de drenagem, avaliando as condições de refino (energia demandada) necessárias para se atingir essas características (VOELKER, 1979).

II.6.2.1 Efeitos do refino

- Corte ou encurtamento das fibras
- Produção de finos e remoção de partes da parede das fibras, criando debris.
- Fibrilação externa, remoção parcial da parede da fibra, que fica presa à fibra
- Alterações internas na parede, chamadas delaminação, fibrilação interna ou inchamento
- Enrolamento ou alisamento da fibra
- Induz ou remove nós, espirais, achatamentos e microcompressão nas paredes.
- Remove ou causa lixiviação de material coloidal.
- Redistribuição da hemicelulose do interior para o exterior da fibra
- Abrasão da superfície no nível molecular, produzindo uma superfície mais gelatinosa.

II.6.3 CARACTERÍSTICAS IMPORTANTES PARA O PAPEL OU SEU PROCESSO DE FABRICAÇÃO

VOELKER (1979) descreve alguns dos requisitos importantes para a operação da máquina de papel:

Energia de refino – Uma das etapas de processo de fabricação do papel é o refino, tratamento mecânico sobre as fibras, cuja severidade varia de acordo com tipo do papel e características de drenagem desejadas na massa em processo. Através do refino as fibras se unem e desenvolvem a resistência do papel. Fibras de diferentes madeiras ou composições químicas necessitam diferentes intensidades de refino.

O objetivo do fabricante de papel é obter polpas que requeiram o menor tempo e energia de refino para desenvolver as características necessárias para o produto a ser fabricado.

Taxa de drenagem na máquina de papel - o tamanho, velocidade e capacidade das máquinas de papel dependem da velocidade com que a água é removida da polpa no feltro. Deseja-se taxas de drenagem rápidas, mas isso pode variar de acordo com o tipo de papel. Polpas muito refinadas desenvolvem grande área superficial, restringindo o fluxo de água.

Floculação e Tendências de formação - Sob certas condições, algumas fibras floculam ou ficam “enozadas”, resultando em problemas de formação no papel final. Essas tendências aumentam principalmente com fibras mais longas ou fibriladas devido ao refino, o que aumenta a área superficial.

Propriedades de resistência - existem algumas exigências mínimas de resistência para todos os tipos de papel, quantificados através de testes físicos de rotina, tais como resistência à tração, ao rasgo e ao estouro.

Características ópticas – a alvura e opacidade (associadas a absorção e dispersão da luz), são normalmente influenciadas pela remoção de compostos cromóforos oriundos da madeira, no branqueamento. Também dependem da área de fibras não ligadas ou expostas, disponíveis para dispersão de luz no papel final, sendo portanto alteradas também pelo refino. Essas características são importantes e variam muito de acordo com o tipo de papel. A opacidade, por exemplo, deve ser alta para um papel de imprimir, porém pode ser baixa para papel vegetal.

Porosidade e capacidade de absorção – é de especial interesse para alguns tipos de papel. Dependem de vários fatores, e tendem a diminuir com o desenvolvimento de força nas ligações interfibras, ao contrário da resistência à tração e à dobra. Portanto papéis altamente absoventes e porosos apresentam essas propriedades de resistência baixas, e vice-versa.

II.6.4 INTERAÇÕES ENTRE AS FIBRAS E SUAS RELAÇÕES COM AS PROPRIEDADES

Nas máquinas de papel as forças de drenagem e as prensas compactam as fibras. Isso leva a um aumento na densidade da folha e a uma redução no diâmetro efetivo dos poros. Como o fluxo de fluidos é controlado pelo diâmetro dos poros, o efeito de se diminuir a espessura da folha e o diâmetro dos poros é a redução da permeabilidade da manta de fibras (para ar e água).

A presença de finos enche os vazios com partículas pequenas, aumentando a densidade e reduzindo o diâmetro dos poros, afetando mais significativamente a drenabilidade e a resistência ao ar (MACKENZIE, 1985).

Diferentes tipos de fibras terão níveis de compactação diversos e reagirão de maneira peculiar às ações de refino entre moinhos diferentes.

As propriedades mecânicas como os índices de tração e estouro, dependem das ligações interfibras; porém podem ser modificadas por outros fatores. O índice de rasgo depende da resistência intrínseca da fibra.

As ligações entre as fibras estão diretamente relacionadas ao número de grupos OH disponíveis para pontes de hidrogênio (HILLMAN, 1999) e a força das ligações entre as fibras. A operação de refino, é muito importante para que a fibrilação ocorra corretamente, sem danificar as fibras.

Se compararmos fibras de mesmas dimensões, pode-se assumir que a área de contato entre as fibras pode estar relacionada à conformabilidade da fibra e conseqüentemente à densidade da folha. Portanto alterações na força das ligações tende a uma alteração na relação entre resistência e densidade do papel.

Num papel cujas fibras são bem interligadas, falhas podem ocorrer devido a quebras de fibras e não devido a quebras de ligações interfibras. Se as fibras tiverem sido enfraquecidas, a probabilidade de quebras de fibra é maior, tornando a sua resistência menor do que se fosse controlada apenas pelas ligações interfibras.

A força das ligações pode ser aumentada durante a fabricação do papel através de alterações na superfície da fibra, com amido, por exemplo. O índice de tração para uma dada densidade é claramente mais alto quando a força das ligações é maior.

Alternativamente, parte do potencial de decréscimo do índice de tração causado por uma redução na habilidade das ligações pode ser compensado por uma menor incidência de falha nas fibras.

Quanto maior o dano da fibra, menor o nível de resistência à tração (em patamares menores) no qual a curva de tração/densidade vai desviar das fibras não estragadas.

Alterações na relação índice de tração/densidade pode ser utilizada para demonstrar variações na força das ligações e resistência das fibras.

Da mesma maneira MACKENZIE (1985) demonstra outras relações entre as propriedades que permitem demonstrar se as fibras estão danificadas ou se as ligações entre as fibras estão enfraquecidas. A avaliação destas relações podem servir como

orientação para a fábrica de celulose na busca de um caminho que leve à melhoria de seu produto.

Fibras longas e bem interligadas são mais difíceis de serem tracionadas e se rompem mais facilmente por rasgo. O índice de resistência ao rasgo é mais sensível a danos nas fibras.

No caso de fibras curtas, a força de tração é menor que a força de ruptura. Com o aumento da força das ligações entre as fibras e da proporção de fibras rompidas ao invés de tracionadas, o índice de resistência ao rasgo chega a um valor constante que representa o trabalho de quebrar todas as fibras. As fibras curtas são mais susceptíveis a enfraquecimento devido a refino prolongado (HILLMAN, 1999).

Em polpas de alto rendimento, a baixa resistência ao rasgo é provocada por ligação entre as fibras fraca ou comprimento de fibra inadequado.

II.6.5 INFLUÊNCIA DA MADEIRA SOBRE AS PROPRIEDADES DAS FIBRAS

A espécie de madeira tem uma grande influência nas propriedades do papel. Deste modo, determinados tipos de madeira são preferencialmente escolhidos para determinados tipos de papel.

SILVA (1996) descreve algumas dessas relações. Por exemplo, fibras com paredes celulares mais espessas são mais resistentes, e por isso conferem ao papel maior resistência ao rasgo, uma estrutura mais aberta e superfície mais irregular, maior volume aparente e maior opacidade. Se essas fibras apresentarem simultaneamente um menor diâmetro de lúmen (cavidade interna da fibra), o papel terá menor flexibilidade (capacidade de colapsamento). Paredes celulares menos espessas apresentam uma resposta ao refino acentuada (HILLMAN, 1999).

Quanto maior a capacidade de colapsamento, maior o número de ligações intrafibras e portanto maior resistência à tração e ao estouro, além de maior peso específico aparente.

O aumento na largura ou a redução do diâmetro do lúmen podem aumentar o *coarseness* (peso das fibras por unidade de comprimento), que também é afetado pela espessura e a densidade da parede da fibra.

O número de fibras por grama de polpa está relacionado ao comprimento médio e ao peso da polpa e influencia as propriedades ópticas. Quanto maior o número de fibras por grama, maiores serão os coeficientes de dispersão de luz e opacidade, maior será o grau de ligação entre as fibras e portanto melhores serão as propriedades de resistência e maior a resistência à passagem de ar.

Pastas celulósicas proveniente de folhosas, por apresentarem fibras mais curtas, resultam em um maior número de fibras por grama, e portanto possibilitam melhor formação de folha, uma superfície de impressão mais lisa, maior volume específico e maior opacidade (CROSSLEY E PITZ, 1996). No entanto apresentar tamanhos semelhantes de fibras não significa apresentar exatamente as mesmas características. CROSSLEY E PITZ (1996) descrevem as características de várias folhosas, entre elas o eucalipto e o *aspen*. O eucalipto tem fibras mais finas, redondas, rígidas e mais curtas, desenvolvendo opacidade, volume específico e qualidades de formação que o tornam a melhor opção para papéis absorventes. Porém seus vasos maiores, podem impedi-lo de competir em mercados de papéis decorativos. Já o *aspen* possui fibras mais flexíveis e uniformes, facilitando a formação do papel. Além disso, seus vasos são menores e não interferem nas propriedades do papel, sendo uma excelente matéria prima para papéis revestidos.

A celulose proveniente de coníferas traz como principal benefício as propriedades de resistência mecânica, e ao contrário das folhosas, em geral não apresentam boas propriedades estéticas, como opacidade, boa formação ou volume específico (HILLMAN, 1999).

A grande maioria das indústrias de papel utiliza uma mistura de celuloses de diferentes fibras a fim de obter um papel com as características necessárias para atender a sua função.

A densidade básica aparente (massa seca/volume) está relacionada aos constituintes químicos da madeira e com as propriedades físico-mecânicas do papel, além de influenciar na produtividade do digestor.

A idade das células também influencia as propriedades das fibras. ALLISON et al (1997) comentaram que fibras mais antigas são mais longas e têm paredes mais espessas.

Muitas pesquisas genéticas vêm otimizando as características das árvores de modo a favorecer as propriedades físicas e ópticas desejadas. No entanto, ao entrar no processo algumas propriedades são prejudicadas e outras beneficiadas. Já que florestas são plantadas anos antes da utilização da madeira (às vezes mais de 50 anos), e que por melhores que suas fibras sejam, estarão sempre sujeitas aos ataques químicos e mecânicos do processo, deve-se também conhecer os efeitos desses ataques nas propriedades, a fim de se otimizar o processo na direção da excelência do produto.

II.6.6 INFLUÊNCIA DO PROCESSO SOBRE AS PROPRIEDADES

A celulose é um polímero de peso molecular e grau de polimerização altos, se comparada às hemiceluloses. Após tratamento químico da madeira, a polpa possui fibras com cadeias de celulose suficientemente grandes, que conferem resistência necessária aos esforços aplicados na formação da folha de papel. Entretanto, as hemiceluloses constituem-se de polímeros de baixo peso molecular e com cadeias laterais, caracterizando-se como uma estrutura amorfa, acessível e reativa, quando comparada às estruturas quase cristalinas da celulose. No entanto, sua cadeia central apresenta resistência aos tratamentos alcalinos utilizados para individualização das fibras. Esses tratamentos solubilizam as cadeias laterais das estruturas das xilanas (tipo de hemicelulose), no caso de folhosas, possibilitando sua adesão às superfícies das fibras de celulose com o aparecimento de zonas amorfas, proporcionando-lhes maior atração por água em razão da presença de maiores teores de grupos polares. Essas zonas conferem às fibras maior susceptibilidade à fibrilação interna e externa, facilitando o tratamento mecânico, com redução de energia consumida, para obter polpas com determinado grau

de refino (CLARK (1985) e DEMUNER et al. (1991), citados em SILVA, 1996) Segundo GENCO (1990) (citado em SILVA, 1996), as hemiceluloses aumentam o rendimento da polpação, melhoram as propriedades de resistência da polpa e são importantes na determinação da interação química entre fibras, água e uma variedade de reagentes químicos utilizados no processo de fabricação de papel. VOELKER (1979), afirma ainda que a redução da quantidade de hemicelulose diminui a habilidade de ligações na superfície das fibras e a facilidade de inchar quando molhadas

O tipo de polpação influi diferencialmente nas propriedades da fibra. Polpas mecânicas apresentam um alto rendimento, porém maior teor de lignina, o que torna a polpa mais dura e rígida, limitando o grau de consolidação, produzindo polpas com volume específico alto, além de reter a cor da madeira. Já as polpas químicas têm os tamanhos das fibras preservados e baixa produção de partículas finas, desenvolvendo um alto grau de consolidação e com isso se tornando mais resistentes. Suas fibras são mais flexíveis e lisas e apresentam propriedades ópticas elevadas. É sabido que o ataque aos carboidratos enfraquece a fibra, pois reduz o tamanho da cadeia polimérica. Portanto é desejável que no cozimento as reações sejam o mais seletivas possível, atacando preferencialmente a lignina. Diferentes processos químicos também apresentam diferenças. O processo kraft de coníferas é menos seletivo que o processo sulfito, devido à grande perda de polissacarídeos principalmente nos estágios inicial e final. A baixa seletividade é ainda reduzida no caso de cozimentos prolongados, prática que reduz o consumo de reagentes de branqueamento. A redução da taxa de deslignificação está associada às reações de condensação e repolimerização da lignina e à baixa seletividade no final, às alterações na reatividade, estrutura e propriedades da lignina residual nas fibras (PEKKALA, 1989). A composição do licor também é importante para a qualidade da polpa, uma vez que está relacionada ao grau de deslignificação e degradação dos carboidratos. PEKKALA (1989) concluiu que há um nível de sulfidez mínimo para otimizar a seletividade.

O grau de deslignificação também é importante, uma vez que a remoção de lignina pode aumentar a habilidade de ligações na superfície das fibras, melhorar a flexibilidade da fibra, as características de inchamento, a facilidade com que a superfície pode ser fibrilada e a taxa na qual a área superficial da fibra pode ser desenvolvida com refino, e a

remoção de lignina mais fraca pode aumentar a resistência por unidade de área (VOELKER, 1979).

A lavagem após o cozimento desprende a lignina da parede da fibra, removendo-a para a fase aquosa, num processo governado pela difusão. Por isso, as características da parede da fibra são muito importantes – paredes mais espessas dificultam a difusão, a má distribuição de lignina na fibra também afeta a taxa de lavagem, pois a lignina mais próxima da superfície será removida mais facilmente. Além disso, o licor de lavagem deve ser o mais “limpo” possível, pois a concentração de lignina no licor vai limitar a difusão (JACOBSON E NGUYEN, 1997).

Um estudo realizado por MACLEOD (1987) indica que para coníferas, os processos entre o digestor e o branqueamento não apresentam alterações nas propriedades das fibras, exceto por um pequeno decréscimo na drenabilidade, atribuído à ação mecânica das bombas e outros equipamentos. O autor conclui que se considerar que o branqueamento não é capaz de melhorar as propriedades de resistência, deve-se dedicar os esforços de otimização das qualidades de resistência das fibras através da otimização dos parâmetros do cozimento, tais como madeira, licor, gradientes de concentração de reagentes e temperatura, práticas operacionais e transferência mecânica da massa celulósica do digestor (pressurizado) para o resto da linha de fibras.

A deslignificação com oxigênio em meio alcalino, recebe a polpa já com a lignina menos reativa, porém é mais seletiva que a polpação. Devido ao caráter aleatório das reações do oxigênio com os carboidratos, poucos são os efeitos no rendimento, mas há redução da viscosidade e pode haver redução da resistência da polpa. O aumento da carga alcalina nesta etapa aumenta a eficiência de deslignificação e a alvura, porém diminui a viscosidade (SILVA, 1996).

Como o cozimento, o branqueamento altera as características da polpa através da alteração do grau de polimerização dos carboidratos, da composição química da polpa, da distribuição de químicos e do conteúdo de agrupamentos de fibras e fibras não cozidas. Existem hoje vários agentes branqueadores em uso, com diferentes princípios ativos, combinados em inúmeras seqüências. Os efeitos de cada um deles variam consideravelmente entre si e conforme as condições de operação são alteradas. No

branqueamento apesar de haver a remoção de lignina, muitas vezes está acompanhada da degradação dos carboidratos, diminuindo o grau de polimerização, e portanto diminuindo a resistência da polpa (VOELKER, 1979). O cloro é o agente mais tradicional e mais seletivo, porém vem sendo abolido dos processos devido às crescentes preocupações ambientais. A substituição dos compostos clorados por derivados do oxigênio trouxe prejuízos tanto à qualidade da fibra quanto a algumas variáveis da fabricação do papel (SILVA, 1996). A formação de radicais livres altamente reativos, nestes processos, provoca queda de viscosidade e rendimento. O peróxido de hidrogênio é menos seletivo que o cloro, e o ozônio, utilizado em seqüências TCF (*total chlorine free*) é o menos seletivo entre todos os reagentes empregados, principalmente quando trabalha em altas consistências (ALLISON et al 1997). Segundo SILVA (1996) as polpas TCF além de um custo mais elevado, também apresentam propriedades de resistência e ópticas ligeiramente inferiores que as polpas ECF (*elemental chlorine free*), além de dificultarem o refino. CROSSLEY E PITZ (1996), por outro lado verificaram que para polpas da mesma espécie ambas as seqüências produziram propriedades semelhantes.

ALLISON et al. (1997) estudaram os efeitos da consistência de trabalho no branqueamento sobre as propriedades da polpa. Segundo os autores, os tratamentos em média consistência alteram a forma das fibras e aumentam as rupturas na parede da fibra, danificando-as e assim criando “pontos fracos” de maior susceptibilidade aos ataques químicos, o que tem um efeito global de queda de resistência.

Alguns dos agentes utilizados no branqueamento agem através da alteração da cor de determinados componentes ou da redução dos compostos coloridos, diminuindo a absorção da luz e aumentando a alvura (VOELKER, 1979).

II.6.7 CONSIDERAÇÕES FINAIS

É impossível se obter uma polpa onde todas as propriedades apresentem níveis ideais. Existirá sempre um balanço, como por exemplo maior índice de resistência a tração, porém menor drenabilidade. Deve-se ter em mente a finalidade do papel e a

otimização das propriedades de importância para sua melhor utilização, sem deixar de lado as exigências do maquinário onde será produzido.

II.7 CONCLUSÃO

Neste capítulo foram apresentados uma visão geral da indústria e produção de celulose, uma descrição genérica de um digestor contínuo, bem como variáveis de processo que afetam a qualidade do produto final. Ainda, foi apresentado um panorama genérico sobre as propriedades das fibras e suas relações com a madeira, o processo e o papel.

Em suma, cada tipo de papel requer diferentes características e a política de operação da fábrica deve ser baseada nessas características, sem entretanto, perder de vista os fatores econômicos. Como a deslignificação no cozimento apresenta menor custo e é menos poluente que no branqueamento, deseja-se obter o máximo de deslignificação e ainda manter a resistência mecânica necessária para o tipo de papel a ser produzido, com o máximo rendimento possível.

CAPÍTULO III

REVISÃO DA LITERATURA PARTE 2: MODELOS MATEMÁTICOS

III.1 INTRODUÇÃO

Existem vários modelos matemáticos na literatura que visam representar o que está de fato ocorrendo no digestor. Os modelos partem de uma série de considerações, que variam em cada caso. São desenvolvidos para cada espécie de madeira e para um determinado tipo de digestor. Consideram diferentes tipos de componentes na madeira, incluem diferentes modelos cinéticos e mecanismos para representar as reações e simplificam diferentes fenômenos. Apesar das equações que descrevem o processo serem baseadas em teoria, os parâmetros são geralmente empíricos e específicos para o processo em estudo.

Neste trabalho, alguns modelos determinísticos encontrados na literatura foram estudados. Isto foi feito, em primeiro lugar para aumentar o conhecimento sobre a teoria envolvida no cozimento, também para ter uma perspectiva sobre o que tem sido desenvolvido sobre o assunto, mas principalmente para selecionar um modelo para ser testado, comparado ao modelo por redes neurais e incorporado ao modelo híbrido.

A primeira parte deste capítulo é dedicada à apresentação dos modelos determinísticos estudados. Deve-se notar que os modelos diferem entre si algumas vezes quanto à abordagem, à complexidade, às considerações e até mesmo à cinética, ilustrando o conhecimento incompleto e/ou não consolidado dos fenômenos ocorridos. Antes de

apresentá-los, porém, há uma breve explicação dos fenômenos cinéticos que acredita-se acontecer durante a polpação.

A segunda parte faz uma revisão na teoria da modelagem através das redes neurais. Localiza esta técnica dentro do ramo Inteligência Artificial e mostra exemplos das diversas áreas em que vem sendo usada. Descreve o princípio por trás da matemática, a matemática de maneira simplificada, alguns tipos de rede e as etapas de desenvolvimento de uma rede preditiva do tipo *feedforward*. Ainda compara o uso de redes versus o uso de modelos determinísticos e introduz o conceito de redes híbridas.

III.2 FENÔMENOS CINÉTICOS DURANTE A POLPAÇÃO

III.2.1 IMPREGNAÇÃO DOS CAVACOS

A impregnação é um fator primordial para a uniformidade da polpa e bom rendimento no cozimento. A intenção nesta parte do processo é expor “todas” as fibras aos reagentes do licor. É amplamente difundido na literatura que a impregnação dos cavacos ocorre segundo dois mecanismos (GUSTAFSON et al., 1989, UHMEIRER, 1995, ASSUMPÇÃO et al., 1988, KAZI et al., 1997)

- Penetração do licor na madeira devido a um gradiente de pressão.
- Difusão dos íons através da água devido a um gradiente de concentração.

A penetração ocorre rapidamente e depois de alguns minutos cai a taxas relativamente baixas. Acontece principalmente na direção das fibras, pois a permeabilidade é dezenas de vezes maior nesta direção (UHMEIER, 1995).

Para entender melhor como o líquido penetra na madeira é necessário saber que as fibras têm uma parede celular com três camadas. Duas camadas compostas principalmente por carboidratos e uma camada chamada lamela média, composta por lignina, que gruda uma fibra à outra. No centro das fibras há uma cavidade chamada

lúmen. As cavidades de duas fibras se comunicam através de canais (*pits*), compostos por duas camadas de parede celular e uma membrana. As cavidades celulares de madeiras vivas contêm uma certa quantidade de água e ar.

A penetração ocorre através de capilaridade. A força de atração entre a parede celular e as moléculas de água são mais fortes que as forças de ligação entre uma molécula de água com a outra. Através de pontes de hidrogênio formadas com os grupos hidroxilas dos carboidratos presentes nas paredes celulares, uma monocamada de água se forma. Devido a outras forças secundárias, outras camadas vão se formando. Esta ação provoca o “inchamento” da madeira, o que significa que o diâmetro efetivo para que o fluxo de água passe é menor que o diâmetro inicial do poro da madeira. É verificado que pHs mais altos aumentam este fenômeno devido à maior adsorção de água nas paredes celulares (KAZI et al., 1997).

Entretanto, o movimento de líquido para abruptamente quando o fluxo encontra o final de um capilar ou se aproxima de um capilar que está no ponto de saturação ou que contenha interface com o ar. O ar contido nos cavacos é comprimido até que a pressão se iguale com a pressão do fluxo líquido. De acordo com a lei de Darcy, quanto menor o gradiente de pressão, menor o fluxo. Decorre daí a grande importância da fase de pré aquecimento dos cavacos, cujo maior propósito é a retirada de ar de suas cavidades. Também fica óbvio que a penetração é mais eficiente quando ocorre em pressões mais elevadas. UHMEIR (1995) sugere que um processo a maiores pressões talvez possa possibilitar o cozimento de cavacos mais longos, o que produziria uma polpa com maior resistência mecânica.

O ponto de saturação mencionado acima é definido pelo ponto em que as paredes celulares estão saturadas e as cavidades não. Em madeiras com maior quantidade de resina, este ponto é mais baixo.

É também comprovado que em maiores temperaturas a impregnação é mais eficiente (UHMEIER, 1995, KAZI et al, 1997). Este fato é explicado pela maior retirada de ar de dentro dos cavacos (principalmente através da maior difusão do ar para fora do cavaco e liberação das bolhas devido à expansão do ar). Enquanto alguns autores

explicam o aumento de impregnação com a temperatura através da lei de Arrhenius, outros não comprovaram este fato (KAZI et al, 1997).

Quando as fibras estão “inchadas” há a formação dos chamados capilares transientes. Apesar de serem muito pequenos para permitir a passagem de fluxo (impedem maior penetração), os capilares transientes permitem a difusão, que não é regulada por tamanho, mas por quantidade de área (KAZI et al, 1997).

Dado que a impregnação por penetração ocorre rapidamente, após o “fechamento” dos capilares, a difusão é então o único meio de transporte dos reagentes (não há mais penetração quando a madeira está saturada). Quanto mais eficiente a etapa de pré aquecimento, maior porcentagem do ar de dentro dos cavacos é retirada e maior a área atingida pela penetração do licor. Desta forma, autores justificam que a impregnação da madeira pode ser modelada apenas pelo fenômeno da difusão (KAZI et al, 1997).

As folhosas são mais fáceis de cozinhar que as coníferas devido ao seu menor teor de lignina e ao fato de que seus vasos permitem a penetração do líquido e do fluxo para o interior dos cavacos. Os nós são de mais difícil penetração e aparecem como rejeito após o cozimento.

A penetrabilidade da madeira depende de seu teor de umidade, sendo que a penetração perpendicular às fibras é bem mais rápida na madeira seca. No entanto, o ar não é totalmente removido caso a madeira esteja muito seca (umidade menor que 30%).

Se a impregnação for incompleta, ao final do cozimento haverá mais rejeitos, o que provoca uma queda de rendimento; maior teor de lignina, o que reduz a branqueabilidade; e produzirá uma polpa com propriedades físicas inferiores (GUSTAFSON et al., 1989). Já a penetração completa reduz o tempo de cozimento, inibe reações concorrentes e indesejáveis e aumenta a uniformidade da polpa. Como foi citado anteriormente, em pHs altos, a taxa de difusão na madeira é semelhante em todas as direções, por isso, a espessura (menor distância) é a dimensão mais crítica para que se complete a impregnação e as reações ocorram (MIRAMS e NGUYEN, 1994). O comprimento apresenta influências mínimas, e a largura é insignificante para a deslignificação (GUSTAFSON et al., 1989).

A difusão é eficiente apenas em curtas distâncias e é muito sensível à composição do licor. De fato, é controlada pela seção transversal de todos os capilares.

III.2 2 REAÇÕES DOS CONSTITUINTES DA MADEIRA

Apesar do objetivo do cozimento ser separar as fibras através da remoção da lignina, o licor de cozimento ataca também os carboidratos e a parede fina das fibras. A extensão dessas reações vai influenciar a qualidade da pasta. As hemiceluloses estão mais expostas ao ataque alcalino que a celulose, por serem cadeias menores e mais ramificadas. Abaixo estão descritas as reações dos carboidratos e lignina segundo ASSUMPÇÃO et al. (1988).

Reações dos carboidratos

inchamento alcalino - causa alterações físicas na parede da fibra, facilitando a penetração. pH favorável: 13. Importante para a uniformidade do cozimento.

solubilização alcalina - Não degrada as fibras e ocorre sempre que uma nova região de madeira se tornar acessível ou uma cadeia molecular for reduzida além de um certo tamanho. O material dissolvido pode ser a seguir degradado pelo álcali ou depositado sobre as fibras.

hidrólise alcalina de grupos acetila - forma íons acetato, e ocorre no início do cozimento. Não consome álcali, mas a neutralização do ácido acético formado consome.

despolimerização terminal - é a etapa responsável pelo maior consumo de álcali do cozimento, devido à dissolução de 35% dos carboidratos, principalmente as hemiceluloses. O grupo terminal vai sendo retirado ciclicamente, e os ácidos orgânicos formados consomem 65% do álcali efetivo. O ciclo para quando ocorre uma **reação de bloqueio**, onde um terminal estável em meio alcalino é formado.

hidrólise alcalina das ligações β -glucosídicas - fraciona os polissacarídeos expondo uma nova unidade terminal sensível à despolimerização terminal. Ocorre em temperaturas acima de 170°C (MIMMS et al., 1989).

Os polissacarídeos se transformam em componentes ácidos, não podendo ser recuperados.

Os carboidratos em solução podem **reprecipitar sobre as fibras**, sendo em seguida adsorvidos pelas microfibrilas da celulose, quando a alcalinidade cai ao final do cozimento. A faixa crítica de reprecipitação é pH 13 a 14. Com a queda do pH, cai a solubilidade.

Mais de 20% da madeira é perdida no processo Kraft devido às reações com a celulose e hemicelulose (MIMMS et al., 1989).

Reações dos extrativos

Os extrativos normalmente são dissolvidos rapidamente no início da reação, também consumindo álcali. No entanto há vários tipos de extrativos, geralmente classificados em saponificáveis e não saponificáveis. O fator importante a se considerar sobre os extrativos é que os não saponificáveis são difíceis de dissolver, e agregados pequenos de extrativos não dissolvidos formam o *pitch* (MIMMS et al., 1989), um pequeno ponto de sujidade na folha de celulose que pode causar grandes problemas numa fábrica de papel, e portanto elevar o grau de rejeição na produção da fábrica de celulose.

Reações da lignina

O mecanismo da deslignificação não está completamente esclarecido bem como a estrutura da lignina.

A deslignificação envolve a fragmentação degradativa da lignina pelos reagentes e a dissolução dos fragmentos no licor. A facilidade de remoção parece estar relacionada à sua localização na parede da fibra.

Em meio alcalino a altas temperaturas, as ligações carbono-oxigênio podem ser rompidas, formando fragmentos menores e mais solúveis.

A presença de sulfetos acelera a deslignificação fornecendo pasta com um menor teor de lignina para um mesmo rendimento.

A combinação com o enxofre já foi provada, mas ainda não se sabe totalmente sua natureza e o motivo do aumento da velocidade. Há estudos que sugerem que o íon HS^- funcione como um catalisador. Esses íons tornam os produtos da deslignificação mais estáveis, mas também produzem mercaptanas, gases associados ao odor característico das fábricas de celulose.

Os modelos clássicos consideram que o processo de deslignificação ocorre em três etapas:

Estágio inicial - baixa solubilização da lignina. Os carboidratos são dissolvidos e cerca da metade do álcali efetivo consumido. Esse estágio depende da madeira e do teor de umidade.

Estágio principal - maior parte da remoção da lignina, ligeira diminuição nos carboidratos e na concentração do licor. Inicia na faixa de temperatura de 140°C a 150°C .

Estágio de deslignificação residual - grande decréscimo nos carboidratos, alto consumo de álcali e baixa deslignificação, o que não é desejável. Para folhosas permite-se o cozimento até o teor de lignina de 2%

Mais adiante, outros modelos descritos neste trabalho são baseados num mecanismo diferente.

Cinética do processo de deslignificação

Para celulose branqueada é vantagem estender a deslignificação devido ao alto custo de reagentes no branqueamento e para diminuir a carga poluente de seu efluente.

Para este caso é importante dizer que deve-se minimizar a degradação dos carboidratos, que prejudica as propriedades de resistência mecânica e causa uma queda no rendimento de pasta branqueada. Apesar de pouco estudada, a taxa de degradação pode ser quase completamente determinada pela concentração de OH^- e pela temperatura. Com relação ao teor de lignina residual, a deslignificação é uma reação de 1ª ordem.

$$-\frac{dL}{dt} = Ae^{\frac{E}{RT}} [OH^-] [HS^-] L \quad (3.1)$$

onde

L = teor de lignina

E = energia de ativação

Porém, apenas o conhecimento de todos esses parâmetros não é suficiente para descrever todos os efeitos do OH^- e HS^- sobre a velocidade de deslignificação ou a quantidade de lignina residual a ser removida na fase final, pois também dependem das condições das fases anteriores.

Na etapa inicial, o aumento de HS^- melhora a seletividade e o aumento de OH^- diminui a lignina residual a ser removida na fase final.

A lavagem ou acidificação entre a etapa inicial e a principal, seguida por um retorno à composição original de licor, retarda a deslignificação principal e aumenta lignina residual.

A concentração de álcali apresenta influência complexa, pois se for alta no início, apesar de diminuir a lignina residual, aumentará a degradação de carboidratos. E no final, não deve ser baixa demais, para não aumentar a lignina residual.

A diminuição da concentração de Na^+ , ou seja, da força iônica, aumenta a velocidade nas fases principal e final, afetando pouco a degradação dos carboidratos. É também fundamental para garantir boa seletividade na fase final.

O aumento de sulfidez até aproximadamente 30% aumenta a seletividade da reação, produzindo uma pasta com menor teor de lignina para um mesmo rendimento. No entanto pode aumentar o nível de odor e compostos de enxofre nos efluentes da fábrica, que devem obedecer limites ambientais legais. Um aumento de sulfidez no final da reação, diminui a quantidade de lignina residual.

De acordo com o texto acima, para a obtenção de melhores resultados deve-se procurar atender as seguintes condições:

- Concentração de OH^- inicial mais baixa que o convencional, com pouca variação.

- Concentração de HS^- alta principalmente no início e no início da fase principal.
- Concentração de lignina dissolvida e Na^+ baixas, principalmente na fase final.

III.3 MODELOS ESTUDADOS

O primeiro modelo que obteve sucesso na comunidade papelreira e é utilizado até hoje devido à sua simplicidade é o Fator-H, desenvolvido por VROOM (1957, citado em SALTIN, 1992). O Fator-H refere-se à área sob o gráfico da velocidade de reação pelo tempo, relacionando, desta forma, tempo de cozimento com temperatura. Muitas fábricas utilizam o Fator-H no controle do cozimento ainda hoje, apesar deste modelo não levar em conta transferência de massa e adição de licor durante o cozimento (altera a taxa de reação), entre outras coisas. Este modelo prega que para um determinado processo, o mesmo Fator-H irá resultar num mesmo grau de deslignificação (mesmo Kappa final).

No entanto, com o desenvolvimento de novas técnicas de cozimento (cozimento modificado, Lo-Solids[®], etc.), de novas técnicas de controle, aumento da capacidade de *software* e *hardware* e crescente aumento de competitividade, verifica-se a necessidade de modelos mais precisos, que descrevam o que está de fato ocorrendo durante o cozimento e otimizem a uniformidade e qualidade da celulose. Com base nesta crença, vários outros modelos foram e vêm sendo desenvolvidos, tanto academicamente, quanto comercialmente.

III.3.1 DIGESTOR CONTÍNUO COM ESCOAMENTO PARALELO E CONTRA-CORRENTE.

Esse modelo desenvolvido por SALTIN (1992) considera a remoção dos componentes da madeira e o transporte dos reagentes e desses componentes, a

transferência de calor e o fluxo em um digestor contínuo, diferenciado para escoamento paralelo e contra corrente.

Assume que a madeira é composta por dois tipos de lignina, uma mais reativa, denominada lignina rápida, e a lignina lenta, por três tipos de carboidatos, a celulose, a galactoglucomanas e a arabinoxilanas, e pelos extrativos, que no entanto se dissolvem instantaneamente. A cinética de reação destes componentes é semelhante, seguindo a mesma fórmula geral (eq. 3.2), que segundo o autor é amplamente utilizada na literatura, porém com parâmetros diferentes para cada um. A Tabela 3.1, apresentada no final desta seção, expõe as equações desenvolvidas para este modelo. O modelo foi baseado em trabalhos realizados por SMITH e WILLIAMS(1975) (citado em SALTIN, 1992) e por CHRISTENSEN et al (1983), que oferecem valores para os parâmetros tanto para coníferas, quanto para folhosas. Os valores utilizados por SALTIN (1992) são os das coníferas.

Como estas equações utilizam as concentrações dos reagentes, sua cinética também deve ser incluída no modelo. Comumente, considera-se que o OH^- obedece a uma taxa de reação linearmente proporcional à dissolução dos componentes da madeira. Quando não ignorado, o HS^- segue o mesmo mecanismo, levando em conta, entretanto, que devido ao seu pK_a neste equilíbrio estar entre 12,5 e 14,9, apenas metade da carga está disponível para o cozimento.

Para representar a transferência de massa, foi usado um modelo simples proposto por TEDER e LINDBERG (1986) e por CHRISTENSEN et al (1983) (citados em SALTIN, 1992), que não considera o gradiente de concentração no interior dos cavacos. Isto é possível pois consideraram a espessura dos cavacos bem pequena - apenas 3 mm. Porém o coeficiente de transferência de massa foi calculado conforme GUSTAFSON et al. (1983), baseado na difusão das substâncias envolvidas. Foi utilizado o mesmo coeficiente para todos os componentes. Deve-se deixar claro neste ponto, que na realidade, a maioria das fábricas utiliza cavacos maiores que 3 mm, onde o gradiente de concentração deve ser levado em conta.

O calor de reação, calculado em função da quantidade de componentes dissolvidos, é adicionado aos cavacos, enquanto que a transferência de calor entre o líquido e os cavacos é semelhante à de massa, mas com um coeficiente de transferência constante.

O fluxo de cavacos determina o tempo que a madeira permanece em cada secção do digestor, sendo portanto muito importante para a qualidade final da polpa. A velocidade de descida no digestor varia conforme os cavacos vão descendo, devido ao fluxo de líquido, grau de deslignificação, altura da coluna de cavacos e nível de líquido no digestor. O modelo aqui descrito é bidimensional, desenvolvido por HARKONEN (1983) (citado em SALTIN, 1992), aplicável apenas para o estado estacionário. SALTIN (1992) desenvolveu este modelo simplificado para ser usado dinamicamente em conjunto com as equações cinéticas. A queda de pressão do líquido segue a equação de Ergun, e a pressão nos cavacos é derivada de um balanço de forças exercidas sobre a coluna: a gravidade, o atrito entre a coluna e as paredes do vaso, e o atrito entre a coluna e o fluxo de líquido.

O modelo é definido ao se combinar a cinética com o fluxo, conhecendo-se os diferentes padrões de fluxo em regiões diversas do digestor - paralela, contra-corrente e extração.

Para realizar esta simulação é necessário fornecer ao programa a altura e o diâmetro das diferentes partes do equipamento e as seguintes variáveis de processo:

1. Fluxo, temperatura e umidade dos cavacos
2. Fluxo, temperatura e álcali efetivo (licor branco)
3. Fluxo, temperatura entrando no aquecedor da zona de cozimento superior
4. Fluxo, temperatura saindo do aquecedor da zona de cozimento superior
5. Fluxo, temperatura entrando no aquecedor da zona de cozimento inferior
6. Fluxo, temperatura saindo do aquecedor da zona de cozimento inferior
7. Fluxo, temperatura do fluxo *quench*
8. Fluxo, temperatura e sólidos do fluxo da extração
9. Fluxo, temperatura entrando no aquecedor da zona de lavagem

10. Fluxo, temperatura saindo do aquecedor da zona de lavagem
11. Fluxo, temperatura e sólidos do fluxo de diluição
12. Temperatura e consistência do fluxo na saída do digestor (*blow line*)
13. Configuração do digestor – altura e diâmetro de cada zona

Tabela 3.1 - As equações do modelo de SALTIN (1992)

Fórmula Geral da Dissolução dos Componentes da madeira

$$-\frac{dW}{dt} = \left(k_1 [OH^-]^p + k_2 [OH^-]^p [HS^-]^k \right) (W - W_\infty) \quad (3.2)$$

1. Equações Cinéticas

$$-\frac{dL_f}{dt} = \left\{ 0,010 e^{\frac{29,6}{RT}} [OH^-] + 0,32 e^{\frac{31,4}{RT}} [OH^-]^{0,5} [HS^-]^{0,5} \right\} L_f \quad (3.3)$$

$$-\frac{dL_s}{dt} = \left\{ 1 \times 10^9 e^{\frac{114}{RT}} [OH^-] + 0,008 e^{\frac{37,6}{RT}} [OH^-]^{0,5} [HS^-]^{0,5} \right\} (L_s - 0,01) \quad (3.4)$$

$$-\frac{dC}{dt} = \left\{ 0,20 e^{\frac{34,7}{RT}} [OH^-] + 0,92 e^{\frac{41,8}{RT}} [OH^-]^{0,5} [HS^-]^{0,5} \right\} (C - 0,33) \quad (3.5)$$

$$-\frac{dG}{dt} = \left\{ 0,051 e^{\frac{25,1}{RT}} [OH^-] + 0,35 e^{\frac{37,6}{RT}} [OH^-]^{0,5} [HS^-]^{0,5} \right\} (G - 0,05) \quad (3.6)$$

$$-\frac{dA}{dt} = \left\{ 325 e^{\frac{73,2}{RT}} [OH^-] + 9,6 \times 10^{14} e^{\frac{167}{RT}} [OH^-]^{0,5} [HS^-]^{0,5} \right\} A \quad (3.7)$$

Tabela 3.1 - As equações do modelo de SALTIN (1992)(cont.)

$$\frac{d[OH^-]}{dt} = \left\{ 18 \frac{dL}{dt} + 8,1 \frac{dCH}{dt} + 8,1 \frac{dE}{dt} \right\} \frac{m_{chips}}{V_{chipsliquid}} \quad (3.8)$$

$$\frac{d[HS^-]}{dt} = 3,3 \frac{dL}{dt} \frac{m_{chips}}{V_{chipsliquid}} \quad (3.9)$$

2. Coeficiente de Difusão

$$k_{mass} = \frac{3,8 \times 10^{-7} T^{1/2} e^{\frac{20377}{RT}} \{0,13[OH^-] - 0,02L + 0,58\}}{w^2} \quad (3.10)$$

3. Transferência de Calor

$$Q_r = 750 \frac{dL_f + dL_s + dC + dG + dA}{dt} m_{chips} \quad (3.11)$$

$$Q_t = 0,005(T_{chips} - T_{liquid}) m_{chips} A_{chips} \quad (3.12)$$

$$\frac{dT_{chips}}{dt} = \frac{(Q_r + Q_t)}{Cp_{chips} m_{chips}} \quad (3.13)$$

$$\frac{dT_{liquid}}{dt} = \frac{-Q_t}{Cp_{liquid} m_{liquid}} \quad (3.14)$$

4. Fluxo

$$\frac{dP_{liquid}}{dh} = 4600 \frac{\epsilon^2}{(1-\epsilon)^3} v + 3,90 \times 10^6 \frac{\epsilon^2}{(1-\epsilon)^3} v^2 \quad (3.15)$$

$$\epsilon = 0,356 - \left(\frac{P_{chips}}{10^4} \right)^{0,59} (0,139 \ln(kappa) - 0,831) \quad (3.16)$$

Tabela 3.1 - As equações do modelo de SALTIN (1992)(cont.)

$$\frac{dP_{chips}}{dh} = (\rho_{liquid}(1-\varepsilon) + \rho_{chips}\varepsilon)g - \mu \frac{P_{chips}}{D_{digestor}} \pm \frac{dP_{liquid}}{dh} \quad (3.17)$$

III.3.2 MODELO CINÉTICO PREDITIVO: APLICAÇÃO NA POLPA DE EUCALIPTO

Um outro modelo, desenvolvido na Austrália por MIRAMS e NGUYEN (1994), para um digestor contínuo e polpa de eucalipto (situação mais típica no Brasil), considera a difusão nos poros da madeira, iniciada durante a impregnação dos cavacos. Durante a impregnação, o licor penetra nos poros da madeira devido a um diferencial de pressão. A partir deste ponto ocorrem as reações seguidas da difusão química. O álcali é consumido pelos extrativos e lignina, bem como pelos carboidratos (reação indesejada), causando um gradiente de concentração no cavaco que leva à difusão contínua de reagentes para o local de reação. Por este motivo, é ser necessário considerar a transferência de massa ao se desenvolver um modelo preditivo do processo.

As reações de deslignificação podem ser consideradas como reações simultâneas de três diferentes tipos de lignina, sendo que a mais reativa reage mais rapidamente. Esta abordagem apresenta um novo ponto de vista para a teoria das três fases de reação, uma vez que seus resultados podem ser interpretados como se a reação de cada lignina fosse responsável por uma fase da reação (deslignificação inicial, principal e residual). São assumidos três compostos sofrendo reações de primeira ordem, e para tal, pode-se trabalhar as reações simultâneas como consecutivas (MIRAMS e NGUYEN, 1994).

As considerações feitas para simplificar o modelo foram as seguintes:

- a. os cavacos são considerados como uma folha plana, com a espessura sendo a única dimensão importante.
- b. cavacos totalmente impregnados antes do início do período aquecido - reações significativas não ocorrem a temperaturas mais baixas existentes antes da total penetração, apesar da impregnação ser importante.

c. transferências de calor para e nos cavacos são instantâneas, e o calor de reação é desprezível - transferência de calor é muito mais rápida que transferência de massa ou que as reações.

d. resistência à transferência de massa entre o seio do líquido e o cavaco é desprezível - só é válida quando há grande circulação de licor em torno dos cavacos, o que não ocorre em digestores comerciais.

e. as reações são irreversíveis e os produtos não interferem com as reações - apenas no final, quando o pH é baixo, um pouco de lignina e hemicelulose podem se precipitar nas fibras.

f. todo o sulfeto está presente como íon HS^- (consideração mais verdadeira para pHs altos) e estes não são consumidos pela reação.

g. A concentração de íons hidroxila é igual ao álcali efetivo (devido à consideração acima)

h. as taxas de difusão para o HS^- e OH^- são iguais.

Os balanços de massa foram baseados em um digestor batelada, onde os diferentes componentes da madeira sofrem reações de primeira ordem, com diferentes parâmetros, calculados através de técnicas estatísticas. Foram utilizadas duas equações para a lignina e duas para os carboidratos, considerando a possibilidade de dois compostos ou duas reações diferentes.

Uma contribuição em relação aos outros trabalhos, é a inclusão das reações dos extrativos, muito rápida em relação às outras, sem aumentar a dificuldade para a resolução do sistema. Esta consideração é importante pois os extrativos têm sua contribuição no consumo de álcali.

Os coeficientes de difusão para o OH^- e para o HS^- foram obtidos de dados para celulose de *pinus* (fibra longa). Apesar de indicarem a importância de desenvolver um modelo específico para cada tipo de árvore, os autores usaram alguns dados de espécies completamente diferentes.

As condições de contorno consideradas foram:

- gradiente de concentração no centro do cavaco é igual a zero

- Concentração nas extremidades do cavaco é igual à concentração no seio do licor
- Concentração inicial do licor é igual à concentração do licor branco
- Concentração no seio do licor proveniente do balanço de massa no licor

Especificações necessárias para resolver o modelo:

- Condições iniciais da madeira: concentração de cada componente na madeira.
- Condição inicial dos íons OH^- e HS^- na madeira: os autores consideram igual à concentração no seio do licor (madeira impregnada).
- Densidade e porosidade da madeira.
- Espessura dos cavacos, álcali efetivo inicial, sulfidez, relação licor/madeira e temperatura.

Como resultados, foram obtidas as concentrações dos íons OH^- e HS^- em todas as posições dentro do cavaco, e a partir destes dados, foi possível calcular as composições médias, o rendimento e o número Kappa.

Tabela 3.2 – Equações do modelo de MIRAMS e NGUYEN (1994)

1. Equações Cinéticas

$$\frac{dL_1}{dt} = -1.99 \times 10^{15} \exp\left(\frac{-120 \times 10^3}{RT}\right) [\text{OH}^-] [\text{HS}^-]^{0.6} (L_1) \quad (3.18)$$

$$\frac{dL_2}{dt} = -6.11 \times 10^{17} \exp\left(\frac{-155 \times 10^3}{RT}\right) [\text{OH}^-] [\text{HS}^-]^{0.6} (L_2) \quad (3.19)$$

$$\frac{dH_1}{dt} = -2.29 \times 10^9 \exp\left(\frac{-80 \times 10^3}{RT}\right) [\text{OH}^-] (H_1) \quad (3.20)$$

Tabela 3.2 – Equações do modelo de MIRAMS e NGUYEN (1994) (cont.)

$$\frac{dH_2}{dt} = -2.00 \times 10^{12} \exp\left(\frac{-120 \times 10^3}{RT}\right) [OH^-] (H_2) \quad (3.21)$$

$$\frac{dC_1}{dt} = -937 \times 10^9 \exp\left(\frac{-31 \times 10^3}{RT}\right) [OH^-] (C_1) \quad (3.22)$$

$$\frac{dC_2}{dt} = -4.35 \times 10^5 \exp\left(\frac{-72 \times 10^3}{RT}\right) [OH^-] (C_2) \quad (3.23)$$

$$\frac{dE}{dt} = -0.6E \quad (3.24)$$

$$\begin{aligned} \frac{\partial OH}{\partial t} = & \frac{1}{X^2} \frac{\partial}{\partial x} \left(D \frac{\partial OH}{\partial x} \right) + \frac{\rho}{0.1\epsilon} \times \\ & \left[3.2 \times 10^{-3} \frac{dL}{dt} + 6.1 \times 10^{-3} \left(\frac{dH}{dt} + \frac{dC}{dt} \right) + 7.0 \times 10^{-3} \frac{dE}{dt} \right] \end{aligned} \quad (3.25)$$

$$\frac{\partial HS}{\partial t} = \frac{1}{X^2} \frac{\partial}{\partial x} \left(D \frac{\partial HS}{\partial x} \right) \quad (3.26)$$

2. Difusão:

$$\begin{aligned} D = & 60 \times 10^{-4} \sqrt{T} \exp\left(\frac{-22.3 \times 10^3}{RT}\right) \times \\ & \left(1.24 \times 10^{-2} - 1.12 \times 10^{-4} \times [L_1 + L_2 + H_1 + H_2 + C_1 + C_2 + E] \right) \end{aligned} \quad (3.27)$$

Tabela 3.2 – Equações do modelo de MIRAMS e NGUYEN (1994) (cont.)

3. Condições de contorno:

$$\frac{\partial OH}{\partial x} = 0 \qquad \frac{\partial HS}{\partial x} = 0 \qquad x=0 \qquad (3.28 \text{ e } 3.29)$$

$$\frac{\partial OH}{\partial t} = \frac{D}{X^2} \frac{\varepsilon}{\rho(l_w - \varepsilon / \rho)} \left(\frac{\partial OH}{\partial x} \right) \qquad x=1 \qquad (3.30)$$

$$\frac{\partial HS}{\partial t} = \frac{D}{X^2} \frac{\varepsilon}{\rho(l_w - \varepsilon / \rho)} \left(\frac{\partial HS}{\partial x} \right) \qquad x=1 \qquad (3.31)$$

III.3.3 MODELO DINÂMICO PARA DESLIGNIFICAÇÃO

Um modelo mais prático, com objetivo de fácil aplicação industrial, foi desenvolvido por VANCHINATHAN e KRISHNAGOPALAN (1995). Este é um modelo dinâmico para a deslignificação, desenvolvido para trabalhar em tempo real, sendo alimentado com dados de análises do licor de cozimento - a partir de análises de condutividade, cromatografia iônica e absorvância UV. Com este modelo é ainda possível estimar as concentrações de OH⁻ e HS⁻.

Para calcular o rendimento e a lignina presente na polpa a qualquer momento dentro do digestor, procede-se com os seguintes balanços de massa, que relacionam a composição dos cavacos com as propriedades mensuráveis do licor:

$$\text{Rendimento} = (\text{massa de cavacos entrada}) - (\text{componentes da madeira dissolvidos})$$

$$\text{Lignina na polpa} = (\text{lignina inicialmente na madeira}) - (\text{lignina dissolvida})$$

O desenvolvimento do modelo considera as três fases da deslignificação (segundo a abordagem clássica descrita em III.2.2), sempre assumindo que o licor está em equilíbrio

com os cavacos, pois a composição do licor interfere na composição dos cavacos. Também assume que todo o sulfeto está na forma de HS^- e que as reações são irreversíveis.

A fase inicial com deslignificação rápida, consumo significativo de carboidratos e grande consumo de álcali, é considerada até a madeira apresentar entre 20 a 22% de lignina. Nesta fase ocorre a quebra das ligações α e β -aryl ether e depende apenas da difusão, e não dos agentes químicos. Para pHs maiores que 12, depende apenas da temperatura.

A fase intermediária é considerada a única em alguns modelos, pois é onde ocorre a maior remoção de lignina. É considerada uma reação de primeira ordem em relação à lignina e segue a Lei da potência (*power law*), com os parâmetros provenientes de dados experimentais.

Por apresentar uma taxa de deslignificação relativamente menor que a fase anterior, a deslignificação residual (terceira fase) não foi considerada neste modelo.

Tabela 3.3 - Equações cinéticas do modelo dinâmico para a delignificação

Fase inicial

$$\frac{dL}{dt} = 40(T)^{1/2} (e)^{\frac{-4807}{T}} (L) \quad (3.32)$$

Fase intermediária

$$\frac{dL}{dt} = 7.49 \times 10^{13} (e)^{\frac{-29863}{T}} [OH^-]^{0.5} [HS^-]^{0.6} (L) \quad (3.33)$$

III.3.4 MODELAGEM DA IMPREGNAÇÃO INCOMPLETA DO CAVACO

Um modelo matemático que, além das equações cinéticas, envolva a extensão da impregnação dos cavacos, pode de antemão avaliar o grau de uniformidade da polpa resultante de várias condições operacionais do digester. Com base nesta premissa,

GUSTAFSON et al. (1989) desenvolveram um modelo mais complexo da penetração e difusão do licor dentro dos cavacos.

A penetração, ou seja, a entrada do licor nos poros da madeira, se dá devido a um gradiente de pressão, enquanto que a difusão, transporte de íons, ocorre devido a um gradiente de concentração.

Para desenvolver tal modelo, os autores assumiram que:

- Os cavacos são isotérmicos e unidimensionais (espessura é importante).
- A madeira é constituída de lignina + carboidratos + acetil.
- As reações são consideradas irreversíveis.

Apesar da penetração ocorrer nos poros que são dispostos principalmente paralelamente ao comprimento, é controlada pela dissolução de gás preso nos poros, e portanto pela difusão deste, conseqüentemente a espessura é crítica também para a penetração.

A penetração ocorre com um padrão do tipo encolhimento do núcleo. Este padrão tem uma forma altamente elíptica com o eixo maior paralelo ao comprimento e o menor à espessura.

Considera que o núcleo encolhido ocupe todo o centro do cavaco - bom o suficiente para aproximar as condições de um cavaco parcialmente impregnado.

O modelo considera a seguinte sucessão de fatos:

No início os cavacos são parcialmente impregnados com água ou licor e a concentração dos agentes químicos é uniforme no seio do líquido.

Com o aumento da temperatura, a frente de penetração vai avançando e as reações são iniciadas gerando produtos de degradação que consomem álcali ao se difundirem nos cavacos.

O álcali é transportado do licor para a superfície dos cavacos e por difusão e convecção com o licor penetrando, é transportado para o interior dos cavacos. A concentração de sulfeto no licor é considerada constante.

Em algum ponto, a penetração é completa e o transporte de álcali passa a ser puramente difusional. No final do cozimento e lavagem tudo o que permanece nas fibras são os componentes ligados da madeira.

Para atacar a dificuldade dos cálculos devido à frente de penetração móvel, os autores introduziram um sistema de coordenadas com origem que se move com a frente de penetração, sendo 0 (zero) nas bordas até 1 na frente de penetração. Com a penetração completa, a frente estará no centro do cavaco.

Equações cinéticas, de difusão e penetração são provenientes do balanço de massa no cavaco e transformação de coordenadas.

Tabela 3.4 - As equações do modelo da impregnação incompleta do cavaco

Balanço de massa para cada espécie i da madeira

$$\frac{\partial C_i}{\partial t} = \left(\frac{\partial \alpha_i}{\partial x} \frac{\partial C_i}{\partial x} + \alpha_i \frac{\partial^2 C_i}{\partial x^2} - Ra_i \delta^2 \right) \frac{1}{\delta^{2+}} + \frac{1}{\delta} \frac{d\delta}{dt} (x-1) \frac{\partial C_i}{\partial x} \quad (3.34)$$

Condições de contorno:

$$\frac{\partial C_i}{\partial x} = 0 \quad \text{na frente de penetração} \quad (3.35)$$

$$B_w (C_{ib} - C_i) = \frac{1}{\delta} \frac{\partial C_i}{\partial x} \quad \text{nas bordas do cavaco} \quad (3.36)$$

Balanço de massa no líquido

$$\frac{\partial C_{ib}}{\partial t} = \left\{ [C_{ib} - C_i]_{x=1} \frac{d\delta}{dt} - \frac{\alpha_i}{\delta} \frac{\partial C_i}{\partial x} \Big|_{x=1} \right\} \frac{V_c}{(V_0 - \delta V_c)} \quad (3.37)$$

As expressões para as constantes cinéticas e difusividade podem ser encontradas em uma das referências do artigo de GUSTAFSON et al. (1989).

Como não há modelos para prever a taxa de penetração, os autores usaram a profundidade de penetração, variando com o tempo, segundo a equação 3.38, cujos parâmetros são especificados pelo usuário de acordo com o tipo de cavaco e de digestor, e que apesar de não ter base física, é representativa:

$$\delta = \gamma + \eta(1 - e^{\beta t}) \quad (3.38)$$

III.3.5 DISCUSSÃO SOBRE OS MODELOS MATEMÁTICOS

SALTIN (1992) se preocupou em capturar o máximo de fenômenos possível, porém fez várias simplificações. Além disso, seu modelo é complexo e não é flexível, não podendo ser aplicado para outras configurações de digestor.

O modelo de MIRAMS e NGUYEN (1994) apresenta algumas características para eucalipto, segue o mesmo tipo de abordagem cinética que o modelo de SALTIN (1992), porém é mais simples e não exige tantos parâmetros de entrada, sendo portanto de computação mais rápida. No entanto, não inclui fenômenos importantes, como a transferência de calor e fluxo de licor em um digestor contínuo, além de utilizar parâmetros para tipos de madeira diferentes.

O único modelo que prevê a impregnação incompleta dos cavacos é o modelo de GUSTAFSON et al. (1989). Porém, esta parte introduziria enorme complexidade caso fosse adicionada ao restante do modelo cinético, tornando-o muito lento computacionalmente.

VANCHINATHAN e KRISHNAGOPALAN (1995) desenvolveram um modelo para ser utilizado em tempo real, baseado em dados do licor. Na realidade, este modelo é ingênuo quando necessita de dados de testes de absorvância e cromatografia, que não são realizados on-line e apenas forneceriam novos dados para o modelo periodicamente.

A apresentação dos modelos acima tem como finalidade expor a complexidade dos fenômenos ocorridos no cozimento. Os quatro modelos têm abordagens distintas e fazem

considerações diferentes em cada caso, mostrando que ainda há muito campo aberto para ser estudado até que se defina qual é a real composição da lignina, qual o real mecanismo ou mecanismos das reações, quais os carboidratos que se deve considerar, qual a inter-relação entre os fenômenos, etc. Deve-se notar, que apesar das reações de reprecipitação serem conhecidas (descritas na secção III.2.2), todos os autores consideram que as reações são irreversíveis. É importante afirmar, entretanto, que em fábricas reais a reprecipitação muitas vezes ocorre (com a queda do pH no fundo do digestor), e que a lignina reprecipitada é de difícil remoção, mesmo no branqueamento. Isso significa que um modelo que possa prever essa ocorrência poderia ser valioso para uma fábrica real.

Um modelo matemático ideal para um digestor deveria incluir características de todos os modelos apresentados anteriormente. Começando por um bom modelo da etapa de impregnação que represente os mecanismos de penetração e difusão; equações adequadas para a transferência dos reagentes de cozimento e produtos da reação; equações cinéticas que predigam com certa precisão o grau de deslignificação em cada ponto do vaso; características de fluxo que descrevam a trajetória dos cavacos no digestor. Além disso, o modelo deve ser alimentado com variáveis palpáveis na indústria, ou seja, aquelas que são medidas rotineiramente. Tal modelo, no entanto, requereria um número muito grande de parâmetros a serem empiricamente determinados com precisão, condição agravada pelo fato de que esses parâmetros são específicos para cada processo. Além disso, para fins de controle de processo, o modelo deve ser computacionalmente leve, para que seja viável em termos de tempo. O mais importante é que o modelo retorne para o usuário um resultado muito próximo ao que realmente está acontecendo. A relação tempo/precisão/complexidade deve depender da finalidade para a qual o modelo será aplicado.

III.4 INTELIGÊNCIA ARTIFICIAL E REDES NEURAIAS

Os estudos em inteligência artificial foram iniciados devido à vontade dos cientistas de fazerem uma máquina que pensa, para resolver problemas que normalmente dependem da inteligência humana.

Quantrille e Liu (1991) descrevem um breve histórico do desenvolvimento da inteligência artificial, que é interessante para localizar as redes neurais dentro deste contexto.

Os autores explicam que os programas de inteligência artificial, ao contrário de programas matemáticos, utilizam uma linguagem de programação simbólica, e não algoritmos. A composição de símbolos com heurísticas e uma busca para o reconhecimento de padrões constituem um programa de inteligência artificial.

Nos anos 60, a linguagem simbólica LISP apareceu juntamente com o FORTRAN, mas devido ao seu sucesso financeiro, a última se difundiu mais largamente.

Nesta época, como se desejava decodificar o funcionamento do cérebro, a pesquisa em inteligência artificial era feita não apenas por matemáticos, mas também por psicólogos, neurologistas e especialistas em cognição. Apesar de saber que os computadores da época não eram ainda suficientes para tornar a inteligência artificial comercialmente viável, a visão de que iriam se tornar cada vez mais rápidos e potentes, levou a comunidade científica à crença de que se fosse dado a eles a capacidade de raciocinar, eles poderiam fazer tudo por nós, e assim, a motivação para a pesquisa em inteligência artificial estava altamente elevada.

Nos anos 70, passou-se a dar um tratamento mais matemático à inteligência artificial, utilizando-se da própria lógica matemática. Foi quando surgiu a ciência da computação, que cada vez mais se tornou uma ciência independente. Continuou-se a perceber as transformações na velocidade e capacidade dos computadores, e então, acreditava-se que se fossem dados aos computadores regras e métodos de buscas para o reconhecimento de padrões adequados, seria possível a obtenção de resultados sobre-

humanos. Os programas que surgiram nesta fase tinham caráter geral, e se dispunham a resolver problemas de forma geral.

Baseados no sucesso de alguns pacotes comerciais de inteligência artificial dos anos 70, nos anos 80 percebeu-se a necessidade de um tratamento contextual, ou seja, fazer programas que lidassem com conhecimentos específicos, para resolverem problemas em áreas específicas. Simultaneamente, houve um enorme desenvolvimento na área de *hardware*, com o surgimento de computadores menores (PCs), mais baratos e mais potentes. Estes fatos elevaram os ânimos, e o interesse em inteligência artificial aumentou enormemente. A necessidade de utilizar conhecimentos específicos e desenvolver programas específicos, bem como a possibilidade de aumentar a comerciabilidade destes pacotes, determinou o início da participação de profissionais das mais diversas áreas na pesquisa da inteligência artificial. O tema passou então a ter um caráter interdisciplinar, envolvendo engenheiros, médicos, arquitetos etc. Os governos e a mídia deram amplo apoio ao assunto.

Porém, com o lançamento de vários pacotes no mercado, as limitações começaram a se evidenciar. Para rodar um programa de inteligência artificial era necessária a utilização de máquinas de grande porte, e assim a relação desempenho/preço era muito baixa, iniciando-se uma onda de críticas.

Hoje, com um clima de um otimismo ainda tímido, as empresas buscam a minimização dessas limitações, na tentativa de melhorar a relação desempenho/preço, e os pesquisadores, buscam uma melhor abordagem, a integração de dados qualitativos e quantitativos e a utilização do conhecimento de forma organizada, apesar do volume sempre crescente.

A inteligência artificial já apresenta as mais variadas aplicações. Por exemplo, pode ser utilizada para reconhecimento de linguagem falada ou escrita, como em máquinas de tradução ou interface de computadores, ou então para dar capacidade de interpretação visual às máquinas, de forma que elas interpretem, por exemplo, fotos aéreas segundo o mecanismo de interpretação humano. É muito aplicada em robótica, na construção de robôs mais flexíveis e “inteligentes”, capazes de interpretação visual e reconhecimento de linguagem. Pode ser utilizada para a demonstração de teoremas que

exigem muita inferência, ou para gerar novos conhecimentos através de uma cadeia de inferências, a partir de uma base de conhecimentos.

Os sistemas especialistas, sistemas dotados de uma grande base de conhecimentos profundos em uma determinada área, os conceitos de lógica *Fuzzy*, que substitui a matemática numérica pelo cálculo qualitativo, trabalhando com valores do tipo “muito, pouco, médio”, através dos conceitos da teoria de conjuntos, e as redes neurais, programas computacionais que “imitam” o cérebro, são técnicas da inteligência artificial que vêm cada vez mais sendo aplicadas a processos químicos para os mais diversos fins, tais como modelagem, simulação, controle e otimização de processos e diagnóstico de falhas. Estas técnicas podem ser utilizadas separadamente ou de forma combinada.

III.4.1 REDES NEURAIS

A curiosidade em saber como o pensamento e a mente humana trabalham, acabou por levar os cientistas a desenvolverem modelos matemáticos do cérebro, cujos objetivos eram estudar o funcionamento do sistema nervoso e produzir programas computacionais que “imitam” seu funcionamento, com características similares ao raciocínio humano - as chamadas redes neurais artificiais (WASSERMANN, 1989). Entre essas características, está a capacidade de aprender com experiência (treinamento). Hoje, estes programas já têm aplicação nas mais diversas áreas, inclusive em Engenharia Química, onde podem ser usados para modelagem, simulação, controle, otimização, etc.

As células nervosas humanas transportam os estímulos recebidos por sensores e mandam informações para o cérebro, que as processa através da interação entre os neurônios, e manda de volta uma resposta relacionada com o tipo de estímulo recebido. Esta conexão entre os neurônios é chamada sinapse, e a sua eficiência é o que se modifica quando o cérebro aprende (NEURALWORKS MANUAL, 1995)

As redes neurais são análogas ao cérebro, compostas por camadas de neurônios, chamados nós ou elementos de processamento, que estão altamente conectados entre si.

A rede recebe um conjunto de *inputs*¹, ou dados de entrada, e essas conexões de caráter microscópico e subsimbólico, muito bem ajustadas, transportam a informação para as próximas camadas, gerando um *output*, que se manifesta de maneira macroscópica, correspondente aos dados de entrada apresentados. Apesar delas serem processadores numéricos, as redes neurais também podem fornecer resultados simbólicos.

III.4.1.1 Propriedades

As seguintes propriedades são associadas às redes neurais:

- As informações são distribuídas em um campo de nós, tendo mais flexibilidade do que os sistemas simbólicos, que são representados por relações hierárquicas fixas.
- As redes têm a habilidade de aprender e se corrigir com a experiência, imitando o processo humano de aprendizagem, ou seja, os erros são sucessivamente corrigidos até que os resultados sejam satisfatórios. Isto é feito na rede através do treinamento por correção de erros (retropropagação).
- Permitem armazenar e acessar de maneira simples uma grande quantidade de informações que ficam retidas na rede pelas conexões entre os nós e pesos dessas conexões.
- Como as conexões têm um caráter microscópico, cada nó contribui muito pouco para o resultado final. Esta propriedade é chamada de microaspecto, e o conjunto de microaspectos poderá ser macroscópico. Desta forma, cada dado de entrada não contribuirá fortemente para o resultado da rede, dando a ela a capacidade de processar dados inconsistentes ou com ruídos.

¹ Os termos *input* e *output* serão mantidos em inglês para dados/sinais de entrada e resultados/sinais de saída para evitar excessiva repetição das palavras entrada e saída, utilizadas para denominar as camadas de entrada e saída da rede.

III.4.1.2 Fundamentos da computação neural

Existem pelo menos duas camadas de nós nas redes neurais, uma de entrada e outra de saída, porém o mais comum é que haja uma ou mais camadas intermediárias entre elas, chamadas camadas ocultas.

As conexões podem se dar entre nós de uma mesma camada (intracamada), entre nós de camadas diferentes (intercamadas), ou até alimentarem a si próprios (recorrentes). Pode também haver uma combinação destas formas. Em engenharia, o tipo mais freqüente é a alimentação intercamadas, que pode ser feita em *feedback* ou em *feedforward*, para modelagem de processos, no mapeamento das relações *input/output*.

No segundo caso, o sinal de entrada em cada neurônio é igual à soma ponderada dos *outputs* da camada anterior adicionados de um fator de tendência (*threshold*) e modificados pela função de ativação gerando um *output* que alimentará a camada seguinte. Este processo se repete até chegar aos *outputs* da camada de saída:

$$f(\sum(w_{ij} a_i) - T_j) \quad \text{onde } a_i \text{ é o } \textit{output} \text{ da camada anterior.}$$

O peso de ponderação w_{ij} indica a excitação (positivo) ou inibição (negativo) e o fator de tendência T_j indica a ativação de cada nó. A forma como estes parâmetros são alterados depende da regra de aprendizagem utilizada e seus valores ao final do treinamento correspondem à memória da rede. Existem vários tipos de função de ativação, porém as mais utilizadas são a função sigmoideal e suas variações, pois possuem limites superior e inferior, preservam os efeitos excitatórios ou inibitórios, além de serem contínuas e monotônicas, permitindo treinamento mais rápido e produzindo redes bem comportadas. Um esquema explicativo da transmissão de sinais em uma rede neural está incluído na secção IV.3.2.

A maneira como os nós estão interconectados dará então, a topologia da rede neural. O algoritmo que dita a forma como os parâmetros irão variar, a função de ativação utilizada e as conexões propriamente ditas irão diferenciar uma rede de outra, e irão influenciar seu grau de precisão.

III.4.1.3 Operação de uma rede neural

A primeira e mais demorada fase da operação de uma rede neural é o treinamento, onde as conexões e pesos entre os neurônios devem ser ajustados. Esta fase é de vital importância, pois a capacidade da rede para predizer relações causa-efeito depende de conexões adequadas.

Apenas após o treinamento, a rede estará pronta para ser utilizada. Nesta fase, ao receber os *inputs*, ela retornará os *outputs* correspondentes. Esta é a fase de generalização.

III.4.1.4 Treinamento

Treinar a rede, como já foi dito, é ajustar os seus pesos até que a rede seja capaz de representar corretamente as relações causa-efeito esperadas. Existem vários algoritmos de treinamento disponíveis, com métodos diferenciados.

O treinamento pode ser supervisionado, quando um instrutor externo controla a aprendizagem e incorpora informação geral. Dentre os treinamentos supervisionados, pode-se destacar:

- Aprendizagem por correção de erro - é a mais comum. Os pesos são ajustados proporcionalmente ao quadrado da soma dos erros observados entre os *outputs* desejados e os calculados.
- *Reinforcement learning* - utiliza apenas um escalar para medir o desempenho da rede, sendo portanto mais simples e rápido que o anterior, além de requerer menos informação.
- Aprendizado estocástico - utiliza estatística, probabilidade e/ou processos aleatórios para ajustar os pesos. Sua vantagem é a capacidade de fugir de mínimos locais.
- *Hardwired* - possui todas as conexões e pesos pré determinados, sendo portanto mais rápido. Utilizado em robótica e reconhecimento de linguagem e visão.
- *Hebbian learning* - ajusta o peso entre dois nós aos quais o peso está associado.

O treinamento também pode ser não supervisionado, ou seja, depende de controle interno e informação local.

III.4.1.5 Treinamento por retropropagação (backpropagation)

É o tipo de treinamento mais utilizado em engenharia química. Em linhas gerais, seu algoritmo deve ser:

1. Atribuir valores aleatórios para os pesos e fatores de tendência.
2. Alimentar a rede com os *inputs*, que serão, por sua vez, os sinais transmitidos por cada neurônio da camada de entrada.
3. Calcular o *output* de cada neurônio da próxima camada - o resultado da adição da soma ponderada dos sinais que chegam em cada neurônio com o respectivo fator de tendência é inserido na função de ativação, gerando o *output* deste neurônio.
4. Alimentar a próxima camada com os *output* de cada neurônio desta camada.
5. Repetir este procedimento até que a camada de saída gere um *output*.
6. Comparar os *outputs* gerados pela rede com os valores esperados (reais) correspondentes aos *inputs* fornecidos ao sistema. Calcular a diferença e fazer a retropropagação deste erro para determinar os novos pesos e os novos fatores de tendência, de forma a minimizar o erro total quadrado. Este procedimento é feito para cada camada até a primeira camada oculta.
7. Repetir o procedimento de 2 a 5 até que o erro total quadrado seja muito próximo de zero.

Existem vários algoritmos de retropropagação, que se diferenciam quanto ao método para achar o erro mínimo, e variar os parâmetros, entre eles, o algoritmo Vanilla, que utiliza o método de Newton, o GDR (*Generalized Delta Rule*), que utiliza um momentum para acelerar o ajuste de pesos etc.

Em alguns casos, o fator de tendência pode mudar repentinamente durante o treinamento, causando baixa estabilidade da rede. Uma maneira de evitar que isso

ocorra, é igualar os fatores de tendência a zero e utilizar uma função sigmoideal como um fator de tendência gradual, chamada de *sigmoid threshold function*, pois seu valor é limitado entre zero e um.

O treinamento por retropropagação tem a vantagem de poder lidar com modelos complexos e não lineares, é flexível e pode armazenar muitos padrões de mapeamento. Porém, demanda muito tempo computacional, não só para programação e desenvolvimento de uma boa interface com o usuário, mas também por necessitar de um grande número de dados de treinamento. A redução destes problemas em parte pode ser conseguida ao se adquirir redes disponíveis comercialmente, que são pré-programadas, onde só é necessária a fase de treinamento. Academicamente, pode-se buscar a minimização do tempo computacional através do ajuste da topologia da rede, da taxa de aprendizagem e variações do fator de tendência ou da função de ativação. Pouco foi publicado sobre como definir o número ótimo de camadas ocultas. Outro fator que pode auxiliar na redução do tempo computacional é a utilização do termo de momentum.

Apesar dos métodos de gradiente decrescente terem acelerado o treinamento, este ainda é muito lento, e com o agravante de não haver garantia de que uma minimização global do erro seja atingida. Principalmente no caso de haver vários mínimos locais, o que poderá causar uma oscilação, alongando ainda mais o tempo de treinamento.

III.4.1.6 Aplicações

As redes neurais têm sido bastante usadas em modelagem de processos, pois funcionam como uma modelagem empírica para ajuste de curvas, e fornecem ao sistema uma maior capacidade de representar as não linearidades do processo e de filtração de dados, devido à propriedade de microaspectos. Um outro ponto que contribui para o seu sucesso, é o fato de poderem trabalhar com múltiplos *inputs* e múltiplos *outputs*, reconhecendo padrões melhor que outros métodos. Além disso, podem ser retreinadas caso o processo sofra alterações.

Sua aplicação mais popular em engenharia química tem sido em controle de processos, podendo ser incorporada a diversos tipos de controladores e diferentes processos, com diferentes funções. Podem entrar como o modelo dinâmico do processo

no controlador, para converter dados quantitativos em qualitativos e interpretá-los, para detectar falhas nos sensores, ou para prever o processo.

São também utilizadas para diagnóstico de falhas, já que suportam uma grande quantidade de dados históricos, e têm a capacidade de tirar conclusões mesmo de dados que apresentem ruídos. Têm ainda uma vasta aplicação para previsão, simulação e otimização de processos.

Em otimização a rede calcula a ótima combinação de *inputs* para que ocorra a combinação de resultados desejada. Para isso, trabalha em reverso, por exemplo, dado o nível de produção mais alto desejado (*output* desejado), a rede vai trabalhar para gerar o vetor de *inputs* que poderá gerar tal *output*. Porém, os dados deverão estar dentro da faixa de treinamento.

As redes neurais podem também ser acopladas a sistemas especialistas, para converter dados quantitativos em qualitativos e vice-versa e/ou para realizar uma análise mais geral do problema, ficando com o sistemas especialistas a responsabilidade por uma análise mais aprofundada.

III.4.1.7 Vantagens e limitações

As redes neurais têm a vantagem de, como as pessoas, aprender em resposta ao ambiente em que se encontram, através de dados que representem este ambiente, ou seja da fase de treinamento. Com esses dados, elas ganham a habilidade de determinar relações de causa-efeito e reconhecer padrões multivariáveis muito bem, sem que se conheça a fundo a teoria por trás da aplicação. Além disso considerações e simplificações de modelos não são necessárias. Comparando-as aos métodos estatísticos, elas apresentam a vantagem de não assumir um tipo pré determinado de correlação.

Devido à sua estrutura distribuída, e portanto à pequena contribuição de cada neurônio para o padrão de *output* global, as redes neurais apresentam baixa sensibilidade a ruídos e informações incompletas, tornando-se então mais flexíveis que os modelos tradicionais para lidar com as “imperfeições do mundo”.

Após a fase de treinamento, sua computação é bem rápida, o que permite o uso das redes *on-line* (em tempo real).

Entretanto, as redes ainda apresentam algumas limitações. Uma delas é o extenso tempo de treinamento, que pode levar mais de 24 horas, e normalmente necessitam de uma quantidade enorme de dados. Para diminuir o conjunto de dados de treinamento, seria necessário que se aumentasse consideravelmente a confiança nos dados alimentados. A utilização de sensores falhos, por exemplo, forneceria *inputs* incorretos à rede, que então obteria resultados errôneos. Também deve-se fornecer à rede um conjunto adequado de dados, para evitar sobreajuste (*overfitting*) nas regiões de treinamento mais denso, e insuficiência de dados onde há menores probabilidades do processo ser operado.

Como elas trabalham com dados empíricos fornecidos na fase de treinamento, elas apresentam baixa capacidade de extrapolação para trabalhar com informações fora da faixa de treinamento. Se houver alterações no processo após a rede ter sido treinada, ela poderá não responder corretamente para a nova faixa de operação.

A propriedade de microaspecto, que lhes dá maior flexibilidade e capacidade de filtrar ruídos, é responsável pela dificuldade de interpretar os modelos gerados por redes neurais.

Uma outra limitação das redes neurais é que elas não prevêm a existência de restrições no processo, por exemplo, nada impede que uma rede neural retorne um valor negativo quando o resultado esperado é a concentração de determinado componente em uma mistura (THOMPSON e KRAMER, 1994).

Sua utilização deve ser cuidadosamente ponderada, levando-se em conta, que como outros métodos, podem funcionar bem para uma finalidade, mas não para uma outra, dentro de um mesmo caso.

III.4.2 REDES HÍBRIDAS

Como foi dito anteriormente, as redes neurais produzem um modelo do processo onde não se pode medir o efeito de cada variável exatamente, devido à baixa influência de cada nó no padrão *input-output*, ou seja, um modelo de difícil interpretação. Elas também não atingem resultados satisfatórios fora da região de treinamento.

Por outro lado, os modelos matemáticos podem se tornar muito complexos para a capacidade e o tempo computacional, ou assumir inúmeras simplificações, que afetam sua precisão. Além disso requerem profundo conhecimento teórico para serem desenvolvidos.

Para minimizar estes problemas, uma técnica alternativa é a combinação modelos baseados nos princípios físicos aproximados para descrever as partes conhecidas e/ou lineares do processo, com redes neurais para estimar parâmetros não lineares ou complexos, de forma a minimizar as incertezas. São as chamadas redes neurais híbridas.

As redes híbridas misturam ambas as técnicas, aproveitando suas vantagens e visando minimizar suas limitações.

O objetivo das redes híbridas é produzir um modelo para o processo mais flexível que as técnicas clássicas de estimativa de parâmetros, e que generalize e extrapole melhor que as redes caixa-preta, bem como sejam mais confiáveis e mais fáceis de interpretar.

Elas podem ser aplicadas de inúmeras maneiras, dependendo do grau de informação e da quantidade e qualidade dos dados experimentais (THOMPSON e KRAMER, 1994). Podem ser implementadas em paralelo e/ou em série com modelos paramétricos, além de poder operar em conjunto com outras redes neurais.

Uma maneira simples são as redes híbridas, cuja parte determinística funciona como um fator de tendência, entrando como mais um valor de *input* na rede neural. Esta técnica deve ser utilizada quando os dados experimentais são melhores que o conhecimento do processo.

Um modelo mais sofisticado utiliza redes neurais para estimar parâmetros ou partes com alto grau de não linearidade ou teoricamente desconhecidos do modelo

determinístico. Desta forma, a rede neural irá minimizar as incertezas do modelo. Neste caso, o modelo teórico deve ser bem conhecido e os dados devem ser preferencialmente bons.

Conhecimento prévio também pode ser utilizado para que o resultado da rede obedeça a certas restrições impostas por princípios físicos e/ou químicos, como por exemplo que a soma das concentrações dos componentes de uma mistura seja igual a 1. Este e outros exemplos são citados em THOMPSON e KRAMER (1994).

O treinamento da rede híbrida é semelhante ao da rede neural, com a diferença de que para cada iteração, além da computação através das camadas de neurônios na(s) rede(s) neural(is), o modelo matemático também é computado.

III.4.3 COMPARAÇÕES ENTRE REDES NEURAI PURAS E HÍBRIDAS

As redes neurais híbridas, geralmente, são melhores que redes puras, pois as generalizações e extrapolações são restritas apenas a algumas partes incertas do processo, mas o conhecimento embutido no modelo o faz sempre consistente com os princípios físicos.

Em oposição às redes puras, que precisam de muitos conjuntos de dados e se baseiam apenas nestes dados para inferir o modelo, as redes híbridas dependem parcialmente de dados, usando o modelo parcial para diminuir o erro num subespaço que pode ser explorado com menos dados.

Resultados obtidos por PSICHOGIOS e UNGAR (1992) para modelagem de um reator em batelada mostraram que a ordem de grandeza da média do erro quadrado para o modelo híbrido foi menor que para a rede pura, e não dependeu do número de conjuntos de treinamento, mas aumentou para conjuntos de treinamento muito pequenos. As redes puras teriam erros da ordem da híbrida caso o conjunto de treinamento fosse muito grande. Os autores também observaram que as redes híbridas extrapolam e interpolam muito melhor que as puras devido à presença do modelo físico e também parecem rejeitar ruídos.

Comparando as redes híbridas com modelos clássicos, os autores chegaram à conclusão de que elas são melhores que vários modelos clássicos devido à melhor previsão de parâmetros não lineares e à não interferência de ruídos.

Os autores afirmam ainda, que as redes híbridas podem não apresentar vantagens significativas sobre as puras quando há grandes incertezas sobre o modelo parcial, ou apresenta um grande número de parâmetros complexos, ou mesmo ainda quando aplicadas a operações contínuas, onde o comportamento de interesse está em volta do ponto de operação normal.

Devem ser substituídas por um modelo clássico caso os parâmetros do processo sejam bem conhecidos, e os coeficientes desconhecidos sejam quase constantes.

III.5 CONCLUSÕES

Neste capítulo foram exibidos os conceitos fundamentais dos principais fenômenos que ocorrem no processo de polpação e alguns modelos matemáticos determinísticos desenvolvidos na literatura. A exposição evidenciou a existência de diferentes opiniões sobre a cinética do cozimento e a complexidade das relações entre os diversos fenômenos ocorridos dentro do digestor.

Um breve apanhado de redes neurais e sua inserção na área de inteligência artificial foi apresentado. Foram também comentadas, de forma sucinta, as principais vantagens e desvantagens desta forma de representação matemática.

Por fim, as redes híbridas foram introduzidas como um meio de minimizar as limitações de ambas as técnicas de modelagem anterior, tirando-se proveito das potencialidades de cada uma.

O próximo capítulo descreve o desenvolvimento do modelo neural com os dados provenientes de uma fábrica de celulose, seus resultados e discussões.

CAPÍTULO IV

PREDIÇÃO DE PROPRIEDADES DA CELULOSE ATRAVÉS DE REDES NEURAIS

IV.I INTRODUÇÃO

Este capítulo apresenta o desenvolvimento de uma rede neural para a predição de propriedades da celulose resultantes do cozimento da madeira a partir de dados industriais.

Estão incluídos neste capítulo a justificativa para a escolha desta metodologia, o desenvolvimento do modelo e os resultados obtidos. A parte de desenvolvimento do modelo descreve desde a coleta de dados, a extensa tarefa de descobrir a melhor configuração da rede, o tratamento do conjunto de treinamento, o treinamento propriamente dito, até finalmente a obtenção de resultados satisfatórios.

Os dados de processo utilizados para este treinamento foram gentilmente oferecidos pela Aracruz Celulose S.A. Para cumprir com um contrato de confidencialidade eles não podem ser publicados.

IV.2 CONSIDERAÇÕES INICIAIS

Existe na literatura um número significativo de estudos que desenvolveram modelos para a simulação do cozimento na produção de celulose. No entanto, tais modelos são baseados em um considerável número de suposições e considerações, além de serem específicos para o sistema para o qual foram desenvolvidos. Isto se dá por principalmente três motivos:

1. A molécula da lignina, bem como o mecanismo das suas reações com o licor de cozimento não são completamente conhecidos.
2. O processo em si é amplamente complexo, com várias reações paralelas, fenômenos de difusão de diversas partículas diferentes, e efeitos do fluxo de cavacos e licor no digestor, diferenças na geometria e/ou dimensões nas diferentes zonas do reator.
3. Os modelos são desenvolvidos normalmente em digestores de laboratório, que não representam exatamente o que ocorre num digestor industrial, devido a diferenças na geometria, nas adições de reagentes, circulações e extração do licor nas diferentes zonas. Até a maior proximidade da idealidade, inerente aos cozimentos de laboratório, afastam o experimento das condições encontradas em uma fábrica. Tais cozimentos não incluem a não uniformidade da madeira, a interferência dos operadores, a diferente operação em cada turno, etc.

Os aspectos descritos acima justificam por si a utilização das redes neurais para este caso: um processo cuja teoria não é 100% conhecida e com reações múltiplas e complexas cujos efeitos conjugados aos outros fenômenos produzem características não lineares. A rede neural não vai acabar com o problema da especificidade, uma vez que deve ser treinada para cada sistema ao qual será aplicada. Entretanto, as não uniformidades da matéria prima, a não consistência na operação, os fenômenos físicos e químicos não capturados nos modelos determinísticos estão implícitos em cada dado coletado na fábrica.

Além disso, o desenvolvimento da rede neural pode ser conduzido através apenas de dados facilmente obtidos na fábrica, enquanto os modelos matemáticos utilizam

concentrações de reagentes e componentes da madeira que não são analisados facilmente. Principalmente quando se considera as concentrações em pontos intermediários do digestor, que é um vaso fechado e pressurizado, do qual não é possível coletar uma amostra de cavacos em qualquer ponto desejado.

O processo de cozimento da madeira para a fabricação da celulose está descrito em mais detalhes no Capítulo II.

A comparação acima justifica a utilização das redes neurais para a predição do número Kappa, quando comparado aos diversos modelos determinísticos. Para outras propriedades, como alvura, viscosidade, etc., entretanto, a justificativa é uma tanto facilitada devido à inexistência de modelos determinísticos para suas predições, uma vez que a relação entre as variáveis de processo e suas influências sobre estas outras propriedades não é bem conhecida.

IV.3 NOMENCLATURA UTILIZADA NESTE CAPÍTULO

Conjunto de dados – cada linha do conjunto de treinamento, composto por um valor de cada variável de entrada e saída, correspondentes à mesma condição de operação.

Conjunto de treinamento – é o conjunto de conjuntos de dados completo.

Variáveis de entrada – representam as condições de operação.

Variável de saída – a propriedade que se deseja predizer. Ex.: kappa, viscosidade, etc.

Intervalo de treinamento – intervalo de dados compreendido entre os valores máximo e mínimo de cada variável.

Tratamento do conjunto de treinamento – análise e eliminação de dados ou conjuntos de dados que poderiam prejudicar a qualidade do treinamento da rede.

IV.4 DESENVOLVIMENTO DO MODELO

A intenção original deste trabalho era uma análise do processo como um todo, desde o cozimento até a obtenção da folha de celulose seca e branqueada. Por esta razão, um grande conjunto de dados provenientes de SDCD (Sistema Digital de Controle Distribuído) e de análises de laboratório foi coletado e avaliado. A fábrica foi dividida em subsistemas para o treinamento de redes neurais individuais para cada parte: cozimento e depuração, pré-branqueamento até o final do primeiro estágio do branqueamento e branqueamento e secagem. A decisão dos pontos de separação do processo foi baseada em estudos realizados previamente na fábrica e no interesse da empresa no efeito específico de certas etapas do processo nas características da polpa.

Porém, devido à impossibilidade de se coletar um maior número de conjuntos de dados, o trabalho foi reduzido ao cozimento da madeira, o coração da fábrica. A empresa forneceu 39 conjuntos de dados, com 22 variáveis de processo (variáveis de entrada para a rede neural) e respectivos resultados – número Kappa e propriedades das fibras.

IV.4.1 COLETA DE DADOS

A coleta de dados constitui um fator importante e singular no desenvolvimento deste projeto. Todos os dados foram coletados de acordo com o tempo de retenção em cada equipamento, incluindo torres de estocagem entre diferentes etapas. Desta forma, cada conjunto de dados representa uma fotografia de um volume de controle específico na sua trajetória pelo processo.

O *software* Rastreabilidade, desenvolvido pela empresa, foi de grande valor na constituição do conjunto de dados. Este *software* recolhe dados do SDCD através do sistema de informação PI (resgata os valores do SDCD e os transfere para a rede de PC's

da fábrica) e calcula os tempos de residência, considerando os valores de níveis de tanque, paradas e partidas de bombas, etc. Desta forma foi possível rastrear todas as variáveis de processo referentes a cada volume de controle amostrado.

Ao invés do valor pontual de cada variável, uma média de 30 minutos foi utilizada, 15 minutos antes e depois do tempo de residência calculado. Este procedimento funciona como uma pré-filtragem para a redução de ruídos inerentes aos dados provenientes de instrumentos numa fábrica.

A Tabela 4.1 descreve todas as variáveis coletadas para o sistema do cozimento. Note-se que CD1 e CD2, esquematizadas na Figura 4.1. são os nomes utilizados para descrever as peneiras das duas zonas de cozimento do digestor estudado.

Tabela 4.1 – Variáveis coletadas para treinamento das redes neurais

RPM do medidor de cavacos	Temperatura na zona de cozimento CD1
Fluxo total de licor branco	Fator H na zona de cozimento CD1
Densidade básica da madeira	Carga alcalina na zona de cozimento CD1
Carga alcalina total	Álcali na zona de cozimento CD1
Concentração do licor branco	Fluxo de licor branco na zona de cozimento CD1
Sulfidez do licor branco	Temperatura na zona de cozimento CD2
Temperatura no topo do digestor	Fator H na zona de cozimento CD2
Fator H no topo do digestor	Carga alcalina na zona de cozimento CD2
Carga alcalina no topo do digestor	Álcali na zona de cozimento CD2
Fluxo de vapor no topo do digestor	Fluxo de licor branco na zona de cozimento CD2
Álcali no topo do digestor	
Fluxo de licor branco no topo do digestor	

Os conjuntos de dados utilizados são produtos de um processo estável. Este fato é particularmente importante pois evita que diferentes causas (variáveis de operação) sejam relacionadas com os efeitos (característica do produto) errados. Também evita creditar tempo extra de reação devido a tempos de estocagem extremamente longos por motivos de paradas.

As amostras de cavacos foram colocadas em sacos plásticos vedados para que não houvesse alteração de condições (principalmente umidade) entre a coleta e a análise.

As amostras de polpa foram lavadas em centrífuga retirando o excesso de licor nelas contido, para impedir o prosseguimento das reações.

Todas as análises, físicas e químicas, da madeira e da polpa, foram realizadas segundo procedimentos internacionalmente estabelecidos e reconhecidos, cuja descrição não é relevante para este estudo.

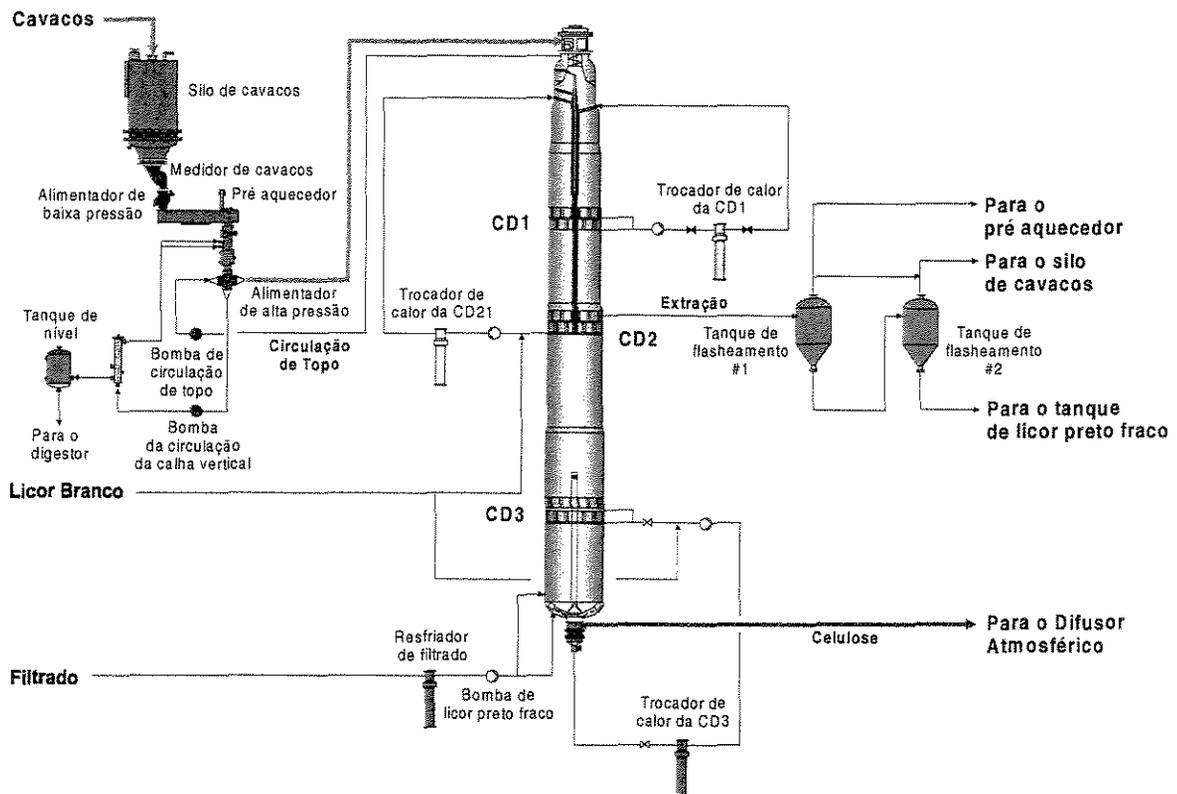


Figura 4.1 – Zonas de cozimento do digestor.

IV.4.2 REDE INICIAL

O desenvolvimento de uma rede neural bem sucedida, desenvolvida com dados industriais, requer uma exaustiva sucessão de tentativas e pequenas alterações nos vários aspectos que definem a configuração da rede. Além disso, este trabalho é uma prova de que não se pode esperar resultados perfeitos apenas com a simples alimentação dos dados coletados diretamente à rede. Os dados devem ser cuidadosamente selecionados, para que se reduza a quantidade de informações não essenciais entrando no programa de treinamento. Ainda, os ruídos estão inerentemente relacionados a qualquer dado obtido de fontes industriais e devem ser filtrados para a obtenção de melhores resultados quando o conjunto de treinamento não é muito grande.

Estes fatos conduziram a um contínuo desenvolvimento da rede e tratamento do conjunto de treinamento, cujas etapas e resultados serão discutidos a seguir, separadamente. No entanto, tudo começou com um modelo previamente testado, aqui chamado de rede inicial.

A rede inicial é baseada na rede utilizada no trabalho de LONA BATISTA (1996).

O programa inicialmente desenvolvido em FORTRAN foi traduzido para C++ devido às exigências da empresa.

Como em um grande número de outros estudos de simulação estática em engenharia química com redes neurais, esta rede é uma rede *feedforward*, o que significa que a informação é passada de uma camada de neurônios para a camada seguinte. Não há conexões recorrentes, onde um neurônio recebe informações do seu próprio *output*, não há conexões entre neurônios em uma mesma camada e nenhuma informação passa para uma camada anterior durante a utilização da rede. O esquema do fluxo de informação em uma rede neural *feedforward* está ilustrado na Figura 4.2.

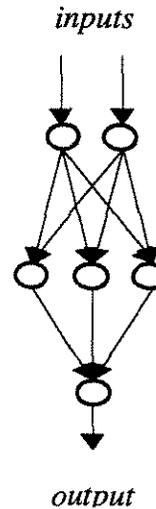


Figura 4.2 – Esquema da transferência de sinais em uma rede neural *feedforward* - Cada dado de entrada (*inputs*) entra em um neurônio da primeira camada da rede. O sinal de cada um dos neurônios da primeira camada é transmitido para cada neurônio da camada seguinte. Por sua vez, o sinal de saída de cada um desses neurônios é transmitido para cada neurônio da próxima camada, que neste caso é a camada de saída da rede. O sinal de saída do(s) neurônio(s) da camada de saída é o *output*, ou o resultado da rede neural.

O número de neurônios na camada de entrada depende do número de variáveis utilizadas no desenvolvimento do modelo. O número de neurônios na camada de saída é igual ao número de propriedades que se deseja prever. A rede inicial continha apenas uma camada intermediária, com N neurônios (N pois o número de neurônios nesta camada foi variado na tentativa de melhores resultados).

Os valores recebidos pelos neurônios da camada de entrada correspondem às variáveis de processo. Cada variável tem seu valor propagado a todos os neurônios da camada seguinte (camada oculta). A ligação entre um neurônio da camada de entrada com um neurônio da camada oculta possui um peso característico, que pode ser chamado genericamente de w_{ij} , onde i corresponde ao neurônio de partida (na camada de entrada), e j ao neurônio de chegada (na camada oculta).

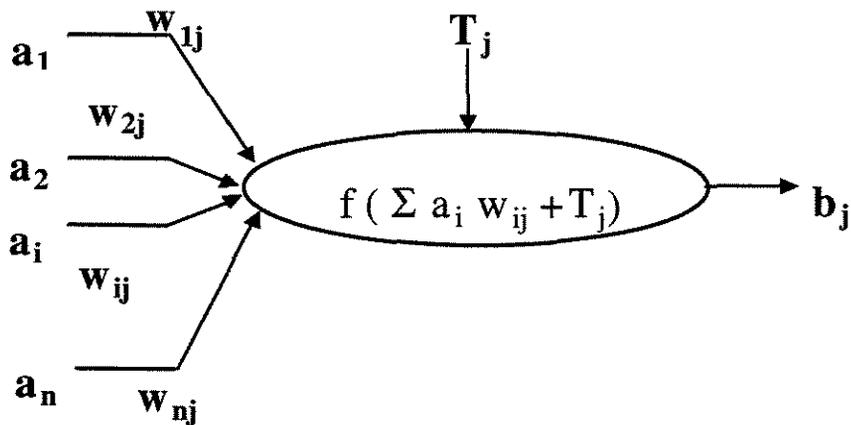


Figura 4.3 – Um neurônio genérico recebendo, processando e transmitindo um sinal - A figura mostra o neurônio j de uma camada genérica (que não seja a camada de entrada). Os sinais provenientes dos neurônios 1 a n da camada anterior chegam a cada um dos j neurônios desta camada. Cada conexão apresenta um peso específico, que é utilizado para o cálculo da soma ponderada destes sinais que chegaram n neurônio j . O fator de tendência T_j , característico do neurônio j , é adicionado ao resultado da soma ponderada e este total entra como a variável independente na função de ativação $f(x)$. O resultado é então o sinal de saída b_j do neurônio j , que será transmitido para a próxima camada. Caso esta seja a camada de saída, b_j será o resultado da rede.

A soma ponderada dos valores que chegam a cada neurônio é somada à *threshold*, ou fator de tendência T_j , característico de cada neurônio da camada intermediária. Este total é então submetido à função de ativação $f(x)$ e o resultado será o *output* de cada neurônio da camada oculta. Os *outputs* de cada neurônio da camada oculta irão, por sua vez, alimentar os neurônios da camada de saída através de procedimento análogo. O recebimento dos sinais dos neurônios da camada anterior e a ativação num neurônio genérico j estão representados na Figura 4.3.

A retropropagação é o método utilizado para treinamento da rede. Isto significa que o erro entre o valor calculado pela rede e o valor real da propriedade predita é retropropagado às conexões, corrigindo os pesos e os fatores de tendência. O método de

retropropagação utilizado foi o GDR (*generalized delta rule*), que treina a rede através da minimização da soma dos quadrados dos erros. A função de ativação utilizada foi a função sigmoidal, que limita o *output* de cada neurônio ao intervalo [0,1].

$$f(x) = 1/(1+e^{-S_j}) \text{ – Função Sigmoidal}$$

A descrição do programa de forma qualitativa pode ser explicada:

1. Lê os dados de um arquivo
2. Inicializa os pesos (W_{aji}) e os fatores de tendência (T_{aj}) aleatoriamente (de maneira simples) com valores entre 0 e 1.
3. Escalona os dados de treinamento – *inputs* e *outputs* – isto é, converte os dados para o intervalo 0 a 1.
4. Entra no *loop* de treinamento
5. As variáveis de entrada já na forma escalonada, referentes ao conjunto de treinamento k , são recebidos nos i neurônios da camada de entrada a .
6. Distribuição dos sinais para a próxima camada ($a+1$) – Os j neurônios da camada $a+1$ devem receber a soma ponderada do sinal de todos os nós da camada a , onde os pesos são específicos para cada ligação. O fator de tendência do neurônio j na camada $a+1$ é somado ao sinal recebido:

$$\text{Entra no neurônio } j = \underbrace{\sum W_{aji} \times \text{output}_i \text{ da camada "a"} + T_{aj}}_{\substack{\uparrow \\ \text{Sinal recebido}}}$$

7. O *output* da camada $a+1$ é calculado pela introdução do valor acima como a variável independente x na função de transferência (função sigmoidal) desta rede neural:

$$\text{Output do neurônio } j \text{ na camada } a+1 = 1/(1 + \exp(-x))$$

8. Os passos 6 e 7 são repetidos até que os *outputs* da última camada sejam calculados.
9. Se este conjunto de dados k for um conjunto de treinamento, deve-se calcular os erros para serem retropropagados e alterar os W_{aji} e T_{aj} . Caso a condição não seja satisfeita,

isto é, este conjunto não for utilizado para treinamento, ele será utilizado apenas a partir do passo 15.2, para o cálculo do erro parcial.

10. Para o neurônio i na última camada:

$$f_{i_{ai}} = (\text{output}_{ai} \text{ calculado}) \times (1 - \text{output}_{ai} \text{ calculado} + \text{offset}) \times (Y_{ki} - \text{output}_{ai} \text{ calculado}),$$

onde Y_{ki} = resultado real do conjunto de treinamento k

11. Para as demais camadas:

$$f_{i_{ai}} = ((\text{“output”}_{ai} \text{ calculado}) \times (1 - \text{“output”}_{ai} \text{ calculado}) + \text{offset}) \times \sum f_{i_{a+1,j}} \times W_{aij}$$

12. Os novos fatores de tendência são calculados a partir dos f_i 's calculados acima:

$$\text{Novo } T_{ai} = T_{ai} \text{ anterior} + \text{taxa de aprendizagem DELTA} \times f_{i_{ai}}$$

13. Os novos pesos W_{aij} são calculados para as conexões de cada neurônio i da camada a com os j neurônios da camada $a+1$

$$\Delta W_{aij} = [\text{DELTA} \times f_{i_{a+1,j}} \times (\text{output}_{ai} \text{ calculado}) + f_{i_{ai}}] + \text{Momentum} \times \Delta W_{aij} \text{ da iteração anterior}$$

$$\text{Novo } W_{aij} = W_{aij} \text{ anterior} + \Delta W_{aij}$$

14. Os passos 5 a 13 são repetidos para todos os conjuntos de treinamento (de 1 a k)

15. Cálculo do erro como critério de parada do treinamento. Este cálculo é iterativo e faz parte do *loop* anterior, porém está descrito separadamente para melhor compreensão

15.1. Erro total = soma dos erros de todos os conjuntos de treinamento:

$$\text{Erro} = \sum (Y_{ki} - \text{output}_{\text{última camada},i})^2$$

15.2. Erro parcial é calculado da mesma maneira, porém utiliza apenas os conjuntos de dados não utilizados para a atualização dos parâmetros (Condição Se no passo 9)

16. Caso o erro seja menor que a tolerância pré determinada, a rede está treinada e pronta para ser utilizada. Caso contrário, deverá voltar ao passo 5 e apresentar todos os dados de treinamento, seguindo os passos de 5 a 16 novamente.

17. A fase de generalização da rede é simples. Inicia pelo escalonamento dos dados de entrada e a seguir é semelhante aos passos 5 a 7 descritos acima. O resultado da camada de saída neste caso, entretanto, é o resultado da rede.

A rede inicial foi treinada com o conjunto de treinamento sem nenhum tratamento prévio ou seleção de dados, sem nenhuma alteração nos parâmetros e configurações da rede, e o resultado foi a não convergência para um valor de erro total pré determinado. Este fato indicou a necessidade de alterações em ambos – na configuração da rede e no conjunto de treinamento. Todas as alterações significativas, suas causas e efeitos estarão descritos abaixo. Para fins didáticos, estarão divididas em duas secções diferentes, porém é importante deixar claro que foram acontecendo simultaneamente de acordo com os resultados obtidos.

IV.4.3 TRATAMENTO DOS DADOS

Esta secção utiliza o número Kappa como exemplo, pois este é o caso objetivo desta tese, mas o tratamento dado às outras propriedades seguiram a mesma linha do tratamento de dados descrito abaixo.

Predição do número kappa no digestor

Um total de 22 variáveis de entrada foram coletados para cada conjunto de treinamento na etapa de cozimento, previamente exibidos na Tabela 4.1.

Como avaliação inicial, uma análise da distribuição de cada variável foi realizada com a finalidade de se determinar a faixa de treinamento e a distribuição dos valores nesta faixa. Um exemplo desta análise pode ser verificado nas Figuras 4.4 e 4.5 abaixo. A Figura 4.4 ilustra a boa distribuição da variável álcali na zona de cozimento CD1, referente à carga de reagentes adicionados a esta zona. Os valores estão distribuídos por toda a faixa de operação, sem a existência de regiões distintas com grande concentração de pontos e espaços vazios ou menos concentrados entre elas. Enquanto que a Figura 4.5

mostra que há apenas alguns valores discretos de sulfidez, agrupados em áreas com grandes concentrações, variando em um pequeno intervalo, mesmo para uma grande variação em número kappa. Como resultado desta análise, a sulfidez foi eliminada do modelo ainda nesta fase inicial. A influência de variáveis com este comportamento não pode ser retratada em uma rede neural, pois o treinamento provavelmente tornará a rede “viciada” na região onde há excessiva concentração de uma mesma variável.

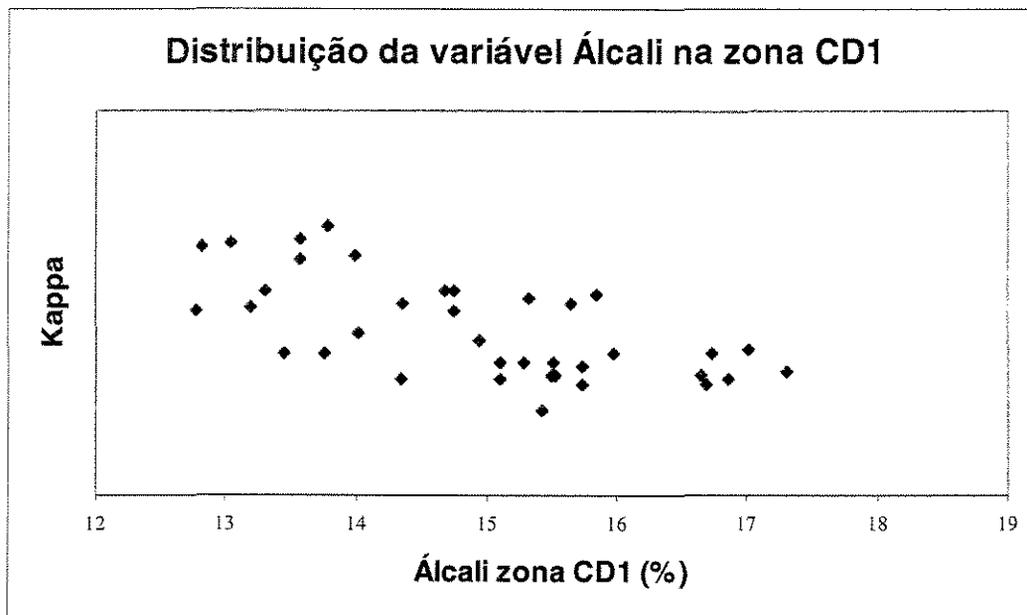


Figura 4.4 – Exemplo de uma variável bem distribuída - Gráfico Álcali na zona CD1 versus Número Kappa — os pontos cobrem todo o intervalo de operação de maneira relativamente uniforme. Percebe-se que normalmente os maiores valores de Kappa estão relacionados com menor carga alcalina.

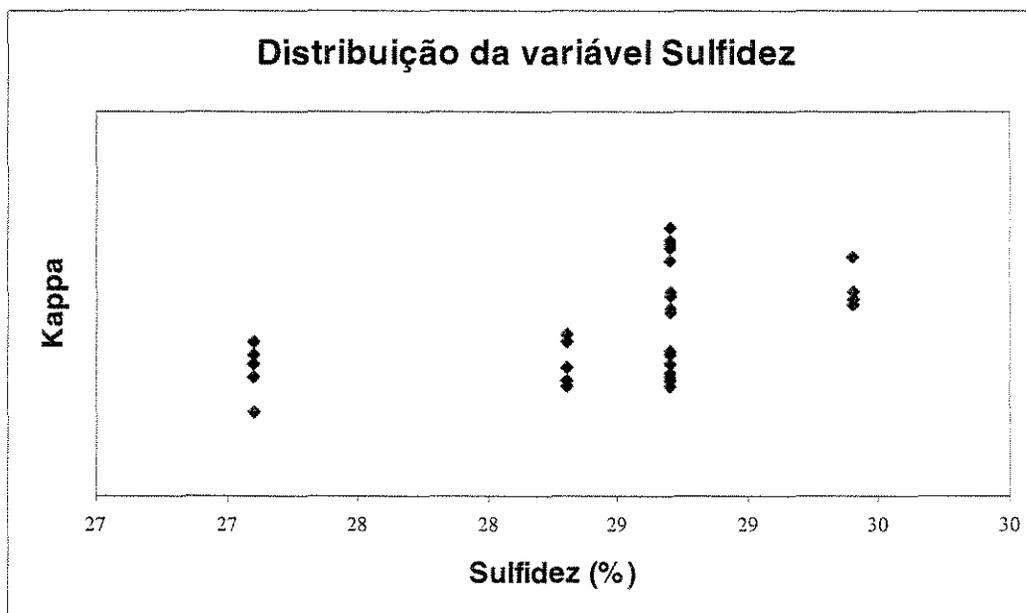


Figura 4.5 – Exemplo de uma variável mal distribuída - Gráfico Sulfidez versus Número Kappa — alta densidade de pontos em determinadas regiões. Os pontos não cobrem todo o intervalo de operação. O aumento ou diminuição da sulfidez não provocou uma resposta consistente do número kappa.

A segunda etapa para avaliação dos dados foi a análise da existência de uma correlação (retratada nos dados) de cada variável com o número kappa. As ferramentas estatísticas do Excel foram utilizadas para esta análise. Uma avaliação de regressão linear foi determinada, incluindo o gráfico da correlação. Algumas variáveis indicaram uma possível correlação linear. Outras não representaram um comportamento linear, porém a análise visual do gráfico indicou possível correlação polinomial. A ferramenta linha de tendência, com a equação da correlação e o coeficiente de correlação R^2 , foi utilizada nestes casos. Todas as variáveis com R^2 menor que 0,45, tanto na correlação linear, quanto na correlação polinomial foram eliminadas. O valor 0,45 foi definido como critério com base na análise visual.

As Figuras 4.4 e 4.5 são novamente um bom exemplo. Percebe-se que o número Kappa mais elevado corresponde a uma menor adição de álcali, indicando a existência de uma correlação entre as duas variáveis. Por outro lado, existem valores de Kappa nas

duas extremidades da faixa de operação (valores altos e baixos), para uma mesma sulfidez.

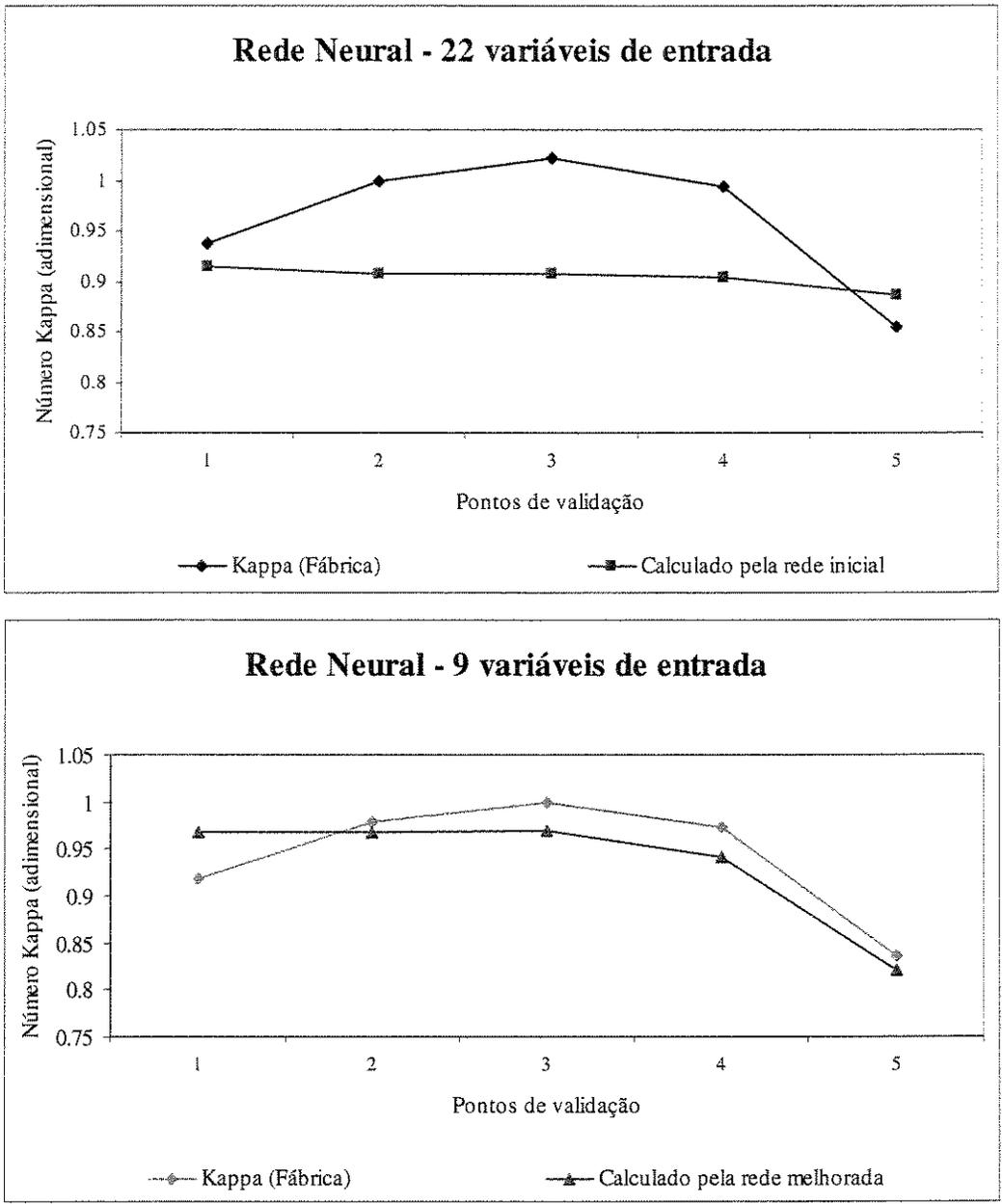


Figura 4.6 – Comparação dos resultados obtidos por redes neurais com diferentes variáveis de entrada – a primeira sem nenhum tratamento de dados (22 variáveis de entrada) e a segunda após a seleção das variáveis (apenas 9 variáveis de entrada). Percebe-se que a seleção das variáveis causou uma melhora significativa nos resultados.

A retirada de variáveis que não mostraram estar correlacionadas com o número Kappa causou uma sensível melhora nos resultados obtidos com a rede neural, como ilustra a Figura 4.6. Essa melhora pode ser atribuída ao fato de que uma menor quantidade de informação “não relevante” foi apresentada à rede durante a fase de treinamento, o que tornou o treinamento mais eficiente. A quantidade de conjuntos de dados disponível não era suficiente para que o treinamento fosse bem sucedido com esta quantidade de valores “não relevantes”.

O fato de determinadas variáveis não terem apresentado correlação, não significa que elas não influenciam a taxa de deslignificação, mas sim que esta influência não estava bem retratada no conjunto de treinamento.

IV.4.4 ALTERAÇÕES NA ESTRUTURA DA REDE

Alguns dos fatos descritos nesta seção são repetidos em outras partes deste capítulo. O objetivo neste ponto é situar uma certa cronologia (apesar de algumas tentativas terem sido simultâneas e/ou repetidas posteriormente) do desenvolvimento da rede, e importância de cada detalhe deste desenvolvimento.

A primeira rede usou todos os dados de entrada e tentava calcular três dados de saída, o que jamais permitiu que o erro total fosse menor que a tolerância pré-determinada.

O primeiro passo para o polimento da rede foi a limpeza de *outliers*. Foram eliminados os conjuntos de dados (valores das variáveis) cujo comportamento não seguem o comportamento geral para mais de uma variável (*outliers*). Esta etapa consistiu de uma cuidadosa análise visual dos gráficos gerados para cada variável (versus o número Kappa) e a retirada dos pontos que claramente mostravam um comportamento anormal. Pelo fato de haver poucos conjuntos de dados para o treinamento da rede, foi importante evitar o fornecimento de pontos que não retratassem bem o processo.

Logo ficou claro que com o pequeno número de conjuntos de dados a predição de mais de uma variável seria impossível, pois a quantidade de informação disponível para o treinamento do modelo era muito pequena comparando-se com a quantidade de parâmetros a serem calculados pela rede. Este fator forçou a decisão para a modelagem com apenas um dado de saída.

Até este ponto, a rede estava sendo treinada com todos os dados disponíveis (as 22 variáveis de entrada), numa tentativa de possibilitar a convergência. Após a redução do número de variáveis de saída, a rede convergiu e as próximas tentativas utilizaram parte dos dados para treinamento e parte para verificação do modelo. O número de dados utilizados para treinamento foi variado e otimizado em todas as variações descritas abaixo.

Apesar de convergir, a qualidade de predição da rede não estava aceitável. O insucesso da qualidade de predição foi então a força motriz para a busca de alterações no programa para a otimização do algoritmo de treinamento.

Neste ponto a seleção das variáveis foi realizada conforme descrito na secção anterior. Uma melhoria foi verificada, como mostrou a Figura 4.6, porém os resultados ainda eram insatisfatórios.

Inicialmente, todos os conjuntos de treinamento estavam sendo utilizados para treinar a rede. Com base em uma entrevista com o professor Mário Gerhard da USP e sua demonstração do *software* de redes neurais da Pavillion, a filosofia de verificação do erro e ajuste de parâmetros da rede foi alterada. Ao invés de utilizar a soma do quadrado do erro de todos os conjuntos de treinamento (erro total no passo 15.1 da descrição do programa) como critério de parada, a nova rede usa uma fração dos conjuntos para treinamento, e a outra parte para o cálculo do erro (erro parcial no passo 15.2 da descrição do programa). Desta forma, por exemplo, a cada três conjuntos que são apresentados para o treinamento da rede, um quarto é utilizado para o cálculo da soma dos quadrados dos erros. Os parâmetros da rede são variados da mesma forma anterior, mas apenas com os conjuntos de treinamento (cada 3 em 4 conjuntos, por exemplo) e a convergência do programa está atrelada ao erro parcial calculado desta nova maneira.

É importante deixar claro que a razão entre o número de conjuntos apresentados para o treinamento e o número de conjuntos utilizados para o cálculo do erro parcial também teve que ser otimizada para cada caso, ficando principalmente entre 2/3 e 1/3 ou 3/4 e 1/4 respectivamente, para todas os casos tentados.

Esta alteração forçou uma reavaliação na ordem de grandeza da tolerância estipulada no programa.

Pelo fato do erro ser cumulativo, ou seja, a soma dos erros de vários conjuntos de treinamento, a tolerância pode ser variada de acordo com o número de dados apresentados. Com menos conjuntos de dados utilizados para o cálculo do erro parcial, a soma dos erros será menor e portanto a tolerância deverá ser menor. No entanto, com um maior número de dados, o treinamento será mais eficiente e apesar de uma tolerância maior trazer à mesma qualidade de predição, é possível melhorar a qualidade do modelo através da redução da tolerância estipulada. É colocar na balança precisão versus tempo de treinamento.

Uma maneira de alterar o critério de parada e garantir que a tolerância esteja sempre adequada, é fazer a comparação da relação do somatório dos quadrados dos erros com o número de conjuntos de dados usados para o seu cálculo:

$$\frac{\sum_{i=1}^j \text{erro}_i^2}{j} \leq \text{tolerância}$$

Como pode ser visto na secção IV.3.5, mesmo com o contínuo decréscimo do erro, o modelo pode passar do ponto ótimo e se tornar *overfitted* (sobre ajustado). Durante o desenvolvimento desta rede, o critério de parada foi alterado para atingir uma certa tolerância ou parar após um certo número de iterações.

O número de camadas ocultas também foi variado. Duas camadas sempre produziram melhores resultados que apenas uma. No entanto três camadas pioraram a qualidade de predição de todas as redes testadas. Isto se deu devido ao aumento do número de parâmetros a serem calculados internamente na rede, em face ao pequeno conjunto de treinamento. O número de neurônios em cada camada oculta também foi

otimizado para todos os casos e a análise da variação está descrita mais adiante, na secção IV.3.3.

A alteração dos pesos com método de gradiente decrescente é linear à função de minimização do erro. Este método assume que a função que descreve o erro é localmente linear, onde a dimensão de localmente varia com a taxa de aprendizagem, ou seja a constante de proporcionalidade utilizada para alterar os pesos (NEURALWORKS MANUAL, 1995). A lógica leva à conclusão de que pequenos valores para esse coeficiente reduzem a extensão do espaço onde o erro é considerado linear, reduzindo a ocorrência de resultados divergentes devido à real natureza não linear da função erro. No entanto, a taxa de aprendizagem deve ser otimizada, não podendo ser muito grande por impedir a convergência do modelo, nem muito pequena por tornar o treinamento muito lento. Mais uma vez, é importante deixar claro que a taxa de aprendizagem teve que ser otimizada para todos os casos. O termo de momentum, incorporado à etapa de alteração dos pesos para acelerar a convergência deve ser otimizado simultaneamente.

De acordo com FAHLMAN (1988), citado no manual do *software* NeuralWorks (1995), quando os pesos dos sinais entrando em um neurônio são muito grandes, a ativação deste elemento se torna saturada – sempre 0 ou 1. Neste caso, a derivada do erro será sempre 0 (zero) paralisando a aprendizagem do dado nó. Para evitar este problema FAHLMAN (1988) sugere a adição de um fator chamado *offset* à derivada da função sigmodal, transformando a função erro em uma função não linear. O manual do *software* NEURALWORKS MANUAL(1995) sugere um valor igual a 0,1. Esse valor foi otimizado para cada rede e em certo ponto, foi modificado para ser proporcional ao erro, porém, de acordo com as observações, não se pode afirmar que este termo sempre melhora o aprendizado da rede.

Todas as alterações descritas nesta secção foram conduzidas de forma que todas estas características pudessem ser dados de entrada na rede e que as futuras otimizações de cada fator, que precisam ser repetidas para cada caso, pudessem ser feitas de maneira rápida. O número de camadas ocultas, que no modelo de partida era pré fixado, foi também generalizado.

Os dados de entrada devem ser escalonados para estarem de acordo com o intervalo de operação da função de ativação da rede. No caso deste trabalho, a função de ativação (função sigmoideal) retorna resultados no intervalo [0, 1]. O escalonamento deve ser feito de maneira que os dados de entrada cubram toda esta faixa, para que a diferença entre os valores de entrada em cada conjunto de treinamento seja bem definida.

A forma como o escalonamento foi realizado também foi analisada. Percebeu-se que deve-se escalonar cada dado de forma que o maior valor esteja bem próximo de 1, mas que não seja igual a 1. Como ilustração, o valor mínimo para a viscosidade entre os dados era de 974 kg/dm^3 . Este valor deve corresponder a um valor muito próximo de zero (mas não zero). Enquanto o valor máximo de 1192 kg/dm^3 deve ser muito próximo de 1 (mas não 1). Portanto, o escalonamento deve considerar uma viscosidade de, por exemplo, 970 kg/dm^3 igual a zero e 1200 kg/dm^3 igual a 1 (Obs.: esses valores são fictícios, para não ferir o acordo de confidencialidade).

Programas utilizando outras funções (seno e tangente hiperbólica) também foram testados, mas foram descartados por não mostrarem melhorias para os casos estudados.

Surpreendentemente, a ordem com que os conjuntos de treinamento foram apresentados à rede influenciaram a qualidade dos resultados. A rede treinada com os conjuntos em ordem aleatória apresentou melhor qualidade de predição que a apresentação dos dados ordenados (ordem crescente de kappa), apesar de ambas terem convergido para uma mesma tolerância, num número de iterações semelhantes (o programa mostra o número de iterações em intervalos e não absolutos).

As alterações acima foram importantes para o desenvolvimento da rede, principalmente para a aceleração do treinamento. Também foram fundamentais para a compreensão de que cada rede é única e precisa ter todos os seus parâmetros otimizados para a obtenção do melhor resultado. Mostraram ainda que mesmo com o algoritmo correto, o trabalho para atingir este resultado ótimo é extenso, ainda que *softwares* comerciais sejam utilizados.

Porém, a alteração que tornou possível o desenvolvimento do modelo neural com qualidade de predição aceitável com o pequeno número de conjuntos de dados disponível

(39 conjuntos de dados) foi o tratamento para transformar a informação contida nos dados industriais em um conjunto de dados bem comportado.

Para isso as correlações entre o número kappa e cada uma das nove variáveis selecionadas foram estudadas individualmente e geraram pares de dados (Variável_i, Kappa) para o todo o intervalo de operação da fábrica. Esses pares foram então comparados entre si, comparados com os dados reais da fábrica e então agrupados em conjuntos de dados (com as nove variáveis), cobrindo uniformemente toda a faixa de treinamento, formando uma curva bem comportada. A Tabela 4.2 ilustra esse processo.

Tabela 4.2 – Esquema do desenvolvimento do novo conjunto de treinamento

1. Definição da correlação entre a Variável Vi e o número Kappa

Obtenção das equações de correlação entre cada variável e o número kappa, a partir dos dados industriais.

V ₁ , Kappa	$a \times V_1 + b = \text{Kappa}$
V ₂ , Kappa	$c \times V_2 + d = \text{Kappa}$
V ₃ , Kappa	$e \times (V_3)^2 + f \times V_3 + g = \text{Kappa}$
V ₄ , Kappa	$h \times V_4 + i = \text{Kappa}$
V ₅ , Kappa	$j \times (V_5)^2 + k \times V_5 + l = \text{Kappa}$
V ₆ , Kappa	$m \times V_6 + n = \text{Kappa}$
V ₇ , Kappa	$o \times V_7 + p = \text{Kappa}$
V ₈ , Kappa	$q \times V_8 + r = \text{Kappa}$
V ₉ , Kappa	$s \times (V_9)^2 + t \times V_9 + u = \text{Kappa}$

Para cumprir com o acordo de confidencialidade, as equações mostradas neste esquema são genéricas e não representam as verdadeiras equações obtidas.

V₁ a V₉ representam as 9 variáveis utilizadas na rede para a predição do número Kappa. Os coeficientes *a* a *u* bem como as equações foram obtidos através da equação da linha de tendência no Excel.

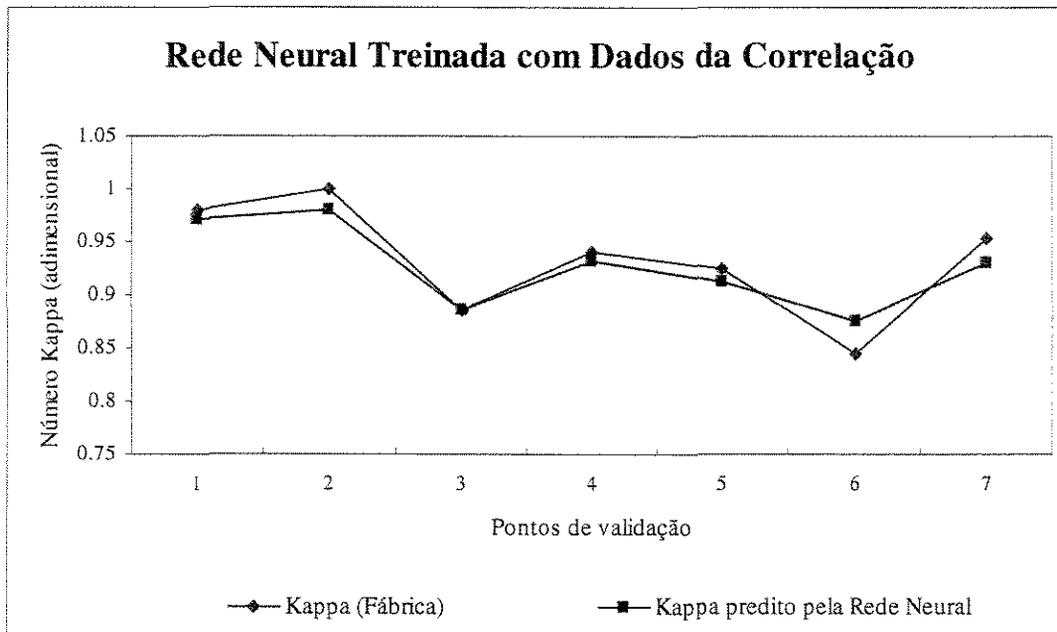


Figura 4.7 – Rede neural otimizada - A validação do modelo de rede neural foi realizada através da comparação dos números Kappa reais (fábrica) e resultados preditos pela rede para 7 conjuntos de dados, como está ilustrado nesta figura. Apesar do treinamento da rede ter sido realizado com pontos da curva de correlação, conjuntos de dados industriais foram utilizados para a validação do modelo.

IV.4.5 REDE NEURAL OTIMIZADA

A rede final é resultado do trabalho de ajuste da configuração da rede até a obtenção de resultados satisfatórios e consistentes. O conjunto de treinamento utilizado para o modelo foi construído com pontos das curvas de regressão obtidas a partir dos dados da fábrica, de acordo com o esquema apresentado acima. O conjunto de treinamento “bem comportado” foi composto de apenas 10 conjuntos de dados.

A função sigmoideal proposta inicialmente apresentou os melhores resultados. Após a eliminação das variáveis que não refletiram significativa influência no número kappa, segundo os critérios descritos na secção IV.2.3, o conjunto de variáveis de entrada foi reduzido a 9 variáveis, listadas na Tabela 4.3 abaixo.

A camada de entrada da rede final recebe as nove variáveis selecionadas. Seus sinais são transmitidos para duas camadas ocultas, com 12 e 10 neurônios, respectivamente. A camada de saída é composta por apenas 1 neurônio, cujo resultado corresponde ao número Kappa.

Tabela 4.3 – Variáveis de entrada do modelo de rede neural para número kappa

Variáveis	Unidades
Álcali efetivo	g/l NaOH
Temperatura (topo)	°C
Fator H (topo)	-
Temperatura (CD1)	°C
Fator H (CD1)	-
Álcali Efetivo (CD1)	g/l NaOH
Temperatura (CD2)	°C
Fator H (CD2)	-
Álcali Efetivo (CD2)	g/l NaOH

A validação do modelo foi realizada com conjuntos de dados reais (da fábrica) não usados nas curvas de correlação ou no treinamento da rede. O resultado final pode ser visto na Figura 4.7. Infelizmente, a empresa só pode oferecer 39 conjuntos de dados. Por este motivo, não foi possível utilizar muitos conjuntos para a validação do modelo. Houve entretanto, a preocupação em testar a predição da rede por toda a faixa de operação. Sete conjuntos de dados, com números Kappa abrangendo esta faixa de operação, foram separados para a validação do modelo.

V.5 DISCUSSÃO DOS RESULTADOS

Um fator importante, objetivo deste estudo, foi a construção de redes neurais capazes de simular o cozimento da madeira na fabricação da celulose treinadas apenas com dados obtidos na fábrica (e não dados comportados de simuladores). Normalmente um grande conjunto de treinamento formado por valores que cubram uma ampla faixa de condições operacionais é necessário para que a rede seja bem treinada. No caso deste trabalho o número de dados industriais foi um fator limitante devido a imprevistos ocorridos na empresa durante o período de coleta de dados.

Apesar disto, o objetivo da modelagem de uma rede neural a partir de dados industriais foi atingido satisfatoriamente para pelo menos uma das variáveis – o número Kappa. Além da obtenção do modelo, o trabalho proporcionou um maior conhecimento sobre a modelagem neural de processos com dados industriais e gerou uma gama de idéias para futuros trabalhos na mesma linha de pesquisa.

A primeira conclusão a respeito do treinamento de redes neurais com dados industriais é que **a generalização** de técnicas de otimização do modelo, tais como a utilização de correlações para a formação do conjunto de treinamento, o número ótimo de camadas ocultas e neurônios nestas camadas, etc., não se mostrou válida, como será discutido nos itens abaixo.

Resultados de uma rede treinada para a predição do índice de tração são apresentados em várias das análises a seguir, por terem apresentado diferenças mais visíveis que as redes treinadas para predizer o número kappa.

IV.5.1 ANÁLISE DOS CONJUNTOS DE DADOS INDUSTRIAIS

Ao contrário de dados obtidos de resultados de simulações, dados industriais não são bem comportados, e apresentam algumas características que obrigam um maior cuidado no treinamento da rede. Dentre essas características, podemos citar:

- Incorporam ruídos – alguns dos dados no conjunto de treinamento terão grande influência dos ruídos provenientes dos medidores ou da transferência de informação entre medidores, SDCD e PI. Os ruídos interferem no treinamento da rede pois mascaram ou causam certos desvios na relação entre a variável medida e a variável que se deseja prever. No caso de dados obtidos por simulação, todos os resultados são provenientes de uma relação definida entre as variáveis.
- Não preenchem uniformemente a faixa de treinamento – apenas um longo período de tempo pode permitir a coleta de dados que representem as condições do processo em toda a faixa de treinamento. Isso se dá devido à baixa probabilidade da fábrica estar operando em casos extremos – produções muito baixas, principalmente, e às vezes mesmo em várias faixas intermediárias. Ainda mais raro é que se opere a fábrica nesses níveis em condições estáveis (não num processo de aumentar ou reduzir o ritmo de produção). Além disso deve-se contar que alguns dos pontos coletados deverão ser descartados por representarem *outliers*. No caso do simulador, os pontos podem ser determinados pelo programador e pode-se cobrir toda a faixa de treinamento proposta uniformemente.
- Demandam um grande período de tempo para a coleta de dados para o conjunto de treinamento – tanto devido aos fatos descritos acima, mas muitas vezes também devido ao tempo necessário para a coleta e análises químicas/físicas das variáveis (neste caso as propriedades da celulose). No caso deste trabalho em que o tempo de residência no cozimento é consideravelmente longo (às vezes 8 horas até que se possa coletar a mostra), o tempo para a coleta de dados é um fator importante, pois é pouco provável que o processo esteja perfeitamente estável por tanto tempo.
- Demandam maior quantidades de homem-hora para a coleta de amostras, que muitas vezes tem que ser feita fora do horário administrativo para obedecer o tempo de residência. E neste caso também para a análise das propriedades físicas da celulose.

- Relações entre as variáveis são complexas – enquanto os simuladores estão carregados de relações bem definidas entre as variáveis, na vida real, e especialmente na fabricação da celulose, um número grande de variáveis têm alguma influência no resultado final e/ou influência em outras variáveis, ainda que estas variáveis não estejam cineticamente relacionadas com a lignina diretamente. Há muitas reações paralelas e subreações que não seriam consideradas em um simulador, mas que estão embutidas nos resultados industriais, “enfraquecendo” as relações como descritas em modelos matemáticos e adicionando informações não consideradas nos modelos determinísticos.

No entanto, para um processo de fabricação de celulose, o nível de conhecimento incompleto sobre a molécula de lignina e o mecanismo de todas as reações presentes ainda justificam a construção de um modelo a partir de dados industriais, que sempre conterão o mecanismo correto e as correlações retratados nos valores das variáveis e seus resultados. Além disso o equipamento, a maneira como é operado e as características da madeira tornam necessária a criação de um modelo específico para cada fábrica. Se os criadores de modelos determinísticos decidirem ser rigorosos, a coleta de dados para a estimativa de todos os parâmetros do modelo seria certamente mais trabalhosa que a coleta de dados para o treinamento da rede. Afinal, a rede utiliza dados que são normalmente coletados, enquanto que os modelos determinísticos requerem dados específicos para a determinação de todos os parâmetros.

IV.5.2 ANÁLISE DO NÚMERO DE VARIÁVEIS

As variáveis a serem coletadas foram escolhidas de acordo com o conhecimento do processo tanto teórico quanto pela experiência de anos de trabalho nesta linha. A partir das 22 variáveis coletadas para a modelagem do digestor, algumas variáveis foram descartadas segundo a discussão na secção IV.2. A Figura 4.6 mostra os resultados de duas redes diferentes treinadas para a predição do número Kappa. São redes de configuração idêntica, exceto pelo número de neurônios de entrada, ou seja de variáveis

de entrada. A Figura 4.6 sugere algumas conclusões importantes que devem ser analisadas e consideradas no desenvolvimento de qualquer rede neural treinada com dados industriais.

A rede treinada com as 22 variáveis de entrada não apresentou resultados coerentes. Neste caso, vários fatores contribuíram para a não identificação dos pesos corretos.

O primeiro fato é que algumas das variáveis não estavam distribuídas em uma faixa, mas se apresentavam quase constantes em alguns pontos discretos, como a sulfidez. Isso não significa que a sulfidez não apresenta influência nos resultados da polpação. Significa que os valores coletados não possibilitam a análise desta influência e portanto não possibilitam e principalmente, não justificam a utilização deste fator no treinamento da rede. Se o objetivo fosse aconselhar a fábrica para a otimização do processo, o único conselho possível a respeito da sulfidez, com base nos dados fornecidos para este trabalho, seria que se variasse este parâmetro para que fosse possível estudar sua influência real no processo.

Outras variáveis, como a densidade da madeira, que acreditava-se exercer forte influência no resultado final, não apresentou correlação alguma na fase de análise do gráfico da densidade versus número kappa. Este fato não implica em anular a forte influência da madeira no processo. Porém, densidade é apenas uma medida relativamente fácil (comparada a análises mais complicadas da composição da madeira) e que se faz rotineiramente em algumas fábricas, entretanto, não retrata a composição da madeira, fator que realmente influencia a demanda e o consumo de reagentes no cozimento, bem como a impregnação dos cavacos. Além disso, a floresta que abastece a fábrica é composta de árvores clonadas, o que as torna incrivelmente uniformes, diminuindo a variabilidade da madeira e conseqüentemente a possibilidade de uma análise da influência da densidade ou composição da madeira nesta determinada fábrica.

A decisão de um limite para a existência de uma correlação aceitável entre as variáveis foi realizada de maneira experimental. Na verdade, a escolha foi predominantemente visual, através da análise dos gráficos. A análise teórica da influência das variáveis, entretanto, foi considerada na decisão do limite do que se

considerou correlação aceitável. Foram escolhidas as 9 variáveis listadas na Tabela 4.3, cujo R^2 foi maior que 0.45.

Os fatores discutidos acima indicam que apesar do que dita o conhecimento teórico, os dados coletados devem ser analisados. Não porque irão desmentir as correlações acreditadas em teoria, mas por talvez não possibilitarem a análise destas correlações e portanto impedirem que a rede retrate sua influência.

Também não se pode exigir que dados industriais apresentem coeficientes de correlação em níveis de laboratório. Em primeiro lugar, os ruídos não o permitem. Além disso, a correlação direta e isolada de uma variável com a propriedade final não retrata a inter-relação entre esta variável e as outras variáveis, fator inerente a qualquer processo complexo.

Ainda, as 22 variáveis representam um excessivo, porém incompleto volume de informação, contido num pequeno conjunto de treinamento, para otimizar um grande número de parâmetros internos da rede, criando ruídos que atrapalharam a convergência. Caso o conjunto de treinamento disponível fosse maior, talvez a relação entre as diversas variáveis pudesse ser melhor observada e capturada durante a fase de treinamento da rede. Para este pequeno conjunto de treinamento, entretanto, as informações sobre estas relações não ficaram claramente estabelecidas, e o excesso de informação não relacionável dificulta os mecanismos de otimização dos pesos nas diversas conexões da rede.

Por outro lado, todas as variáveis que exercem forte influência nos resultados devem ser consideradas, ainda que a análise inicial da correlação não pareça perfeita. O mesmo critério de seleção de variáveis foi utilizado para o índice de tração. As 9 variáveis selecionadas para o treinamento da rede de predição do número kappa demonstraram exercer influência sobre o índice de tração. Após a convergência do modelo, as 3 variáveis com coeficiente de correlação mais baixos foram retirados. A Figura 4.8 demonstra que ao se retirar variáveis importantes, a rede não captura algumas das inter-relações fundamentais e portanto o desempenho do modelo é empobrecido.

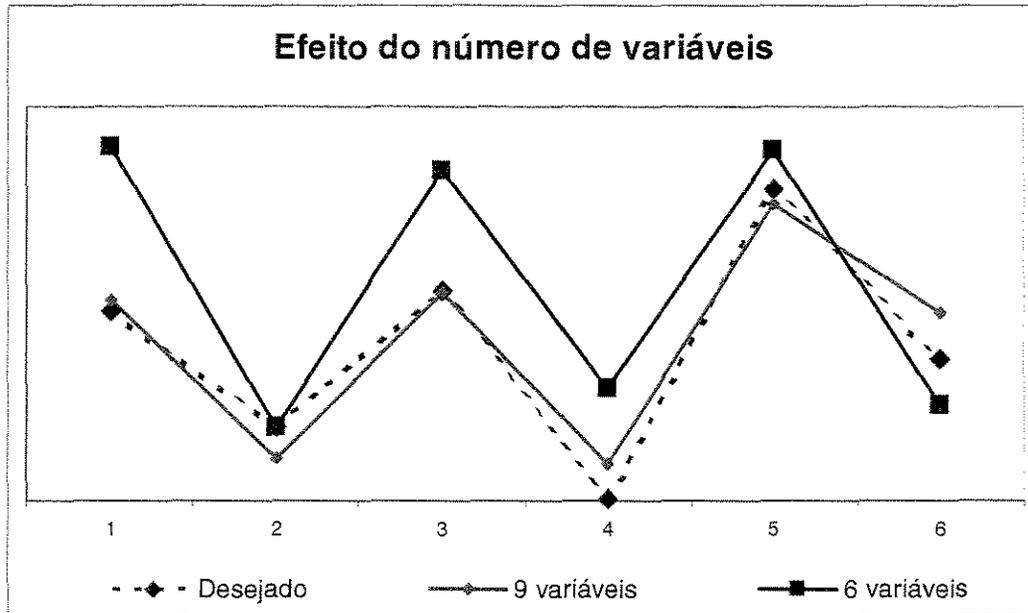


Figura 4.8 – Efeito do número de variáveis. A curva com 9 variáveis descreve os resultados da rede treinada para o índice de tração com todas as variáveis selecionadas segundo o critério do coeficiente de correlação e análise visual. As 3 variáveis com coeficiente de correlação mais baixos foram retirados e os resultados obtidos foram menos precisos, retratados na curva com 6 variáveis.

IV.5.3 EFEITO DO NÚMERO DE NEURÔNIOS NAS CAMADAS OCULTAS

Como afirmado anteriormente, foi impossível gerar uma regra de como se prever o número ótimo de neurônios nas camadas ocultas. Para o caso deste trabalho, onde se utilizou um pequeno conjunto de treinamento, verificou-se apenas que o número de neurônios em cada camada oculta não deve ser muito maior que o número de neurônios na camada de entrada.

A rede neural treinada para a predição do índice de tração, cujos resultados estão resumidos na Figura 4.9, é um bom exemplo para ilustrar o efeito do número de neurônios nas camadas ocultas. Os valores absolutos do índice de tração não são fundamentais para efeitos de comparação, e portanto, para cumprimento do acordo de confidencialidade, não estão sendo apresentados.

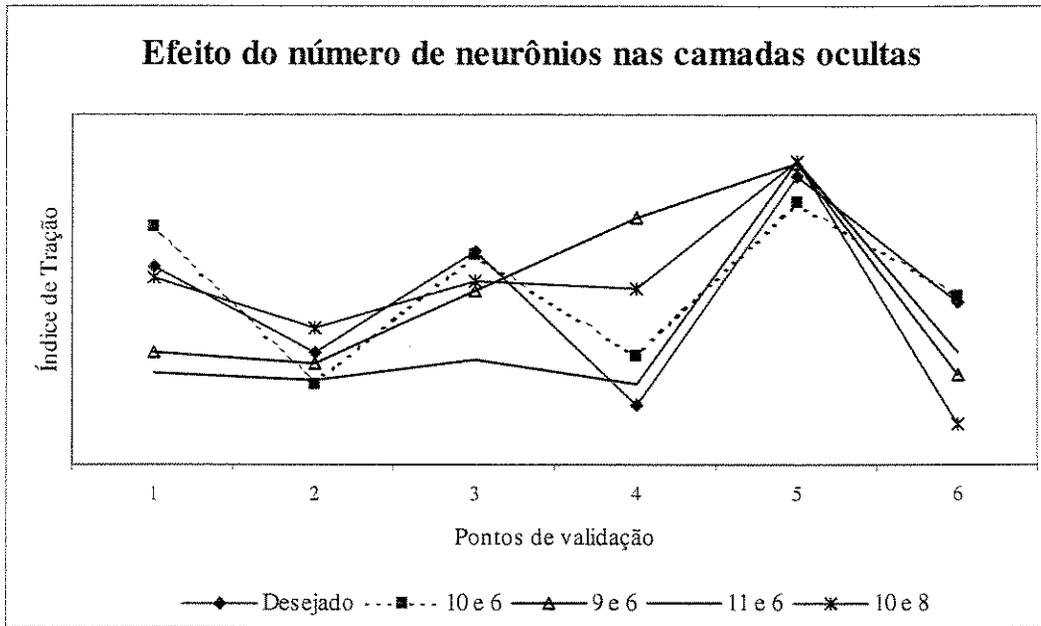


Figura 4.9 – Efeito do número de neurônios nas camadas ocultas - Este gráfico apresenta os resultados de redes semelhantes, cuja única diferença é o número de neurônios nas camadas ocultas. A rede com 10 e 6 neurônios (na primeira e segunda camadas ocultas respectivamente) apresentou resultados mais próximos dos desejados. A retirada ou adição de um neurônio em uma das camadas causou a queda na capacidade de predição da rede. Obs.: Os valores no eixo y foram retirados por respeito ao acordo de confidencialidade.

A Figura 4.9 mostra um gráfico de redes idênticas, com duas camadas ocultas, porém com diferentes números de neurônios intermediários. O experimento foi iniciado com uma rede com 10 e 6 neurônios nas camadas ocultas 1 e 2 respectivamente. O número de neurônios na camada 1 foi variado para 9 e depois para 11. Com 9 neurônios, a rede não foi capaz de prever o mesmo comportamento para todos os casos. Com 11 neurônios, os resultados não foram tão satisfatórios quanto como haviam sido com 10. A variação do número de neurônios na segunda camada oculta também apresentou resultados piores. A Figura 4.8 é apenas ilustrativa e tem a finalidade de mostrar que a regra não é – “Se o número de neurônios na camada oculta for aumentado (ou diminuído), os resultados de predição da rede irão melhorar”. Na verdade o que acontece, é que a maior quantidade de neurônios ocultos implica numa maior quantidade

de parâmetros a serem otimizados durante o treinamento da rede. Portanto, o tamanho do conjunto de treinamento parece ser um fator crucial na possibilidade (e necessidade) de se aumentar ou diminuir o número de neurônios internos da rede. Como foi explicado acima, o pequeno conjunto de treinamento neste trabalho não é fonte de um retrato completo das relações existentes entre todos os fatores do processo. Além disso, as 9 variáveis neste pequeno conjunto não apresentam informação suficiente para a otimização de um número maior de parâmetros (pesos da rede).

IV.5.4 UTILIZAÇÃO DE CORRELAÇÃO PARA A OTIMIZAÇÃO DO CONJUNTO DE TREINAMENTO

A Figura 4.10 mostra os resultados para o índice de tração para duas redes idênticas – uma treinada com os dados crus (reais sem tratamento) e outra treinada com uma curva construída a partir de correlações.

Ao contrário do número kappa (veja Figura 4.7), o treinamento com os dados crus produziram melhores resultados que o treinamento com o conjunto bem comportado criado com as correlações. Isso indica que não se pode generalizar o fato de que conjuntos gerados através de correlações produzem redes que melhor retratem o processo real, ou que a mesma técnica de geração de correlações pode e deve ser usada para todos os casos. Na verdade, esta técnica de geração de conjunto de treinamento é relativamente simplista e se não for cuidadosamente aplicada, pode interferir e alterar as informações entre as inter-relações entre as variáveis. O problema é que não é apenas a correlação entre a variável x_1 com a variável y que importa. Mas também, como a variável x_1 influencia y na presença das variáveis x_2 , x_3 , x_4 , etc. Onde as variáveis x_1 e y são variáveis genéricas.

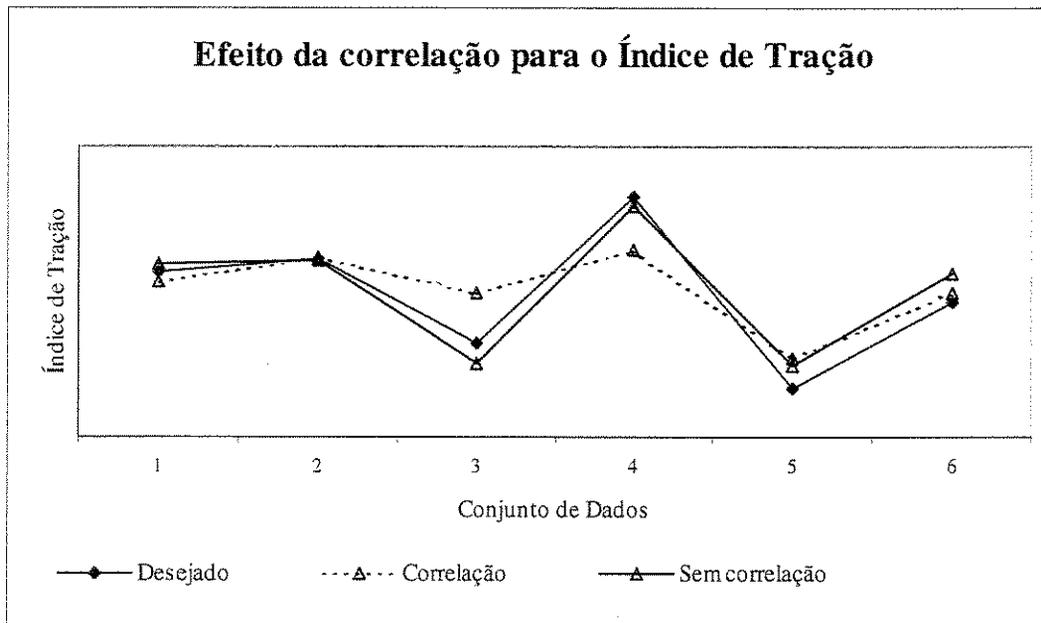


Figura 4.10 – Redes neurais para a predição do índice de tração - Uma das redes apresentadas nesta figura utilizou os dados industriais diretamente para o treinamento (Sem correlação). A outra foi treinada com um conjunto de treinamento formado por pontos de uma curva de correlação entre as variáveis de processo e o índice de tração obtida dos dados industriais (Correlação). Ao contrário da rede para a predição do número Kappa, a rede treinada Sem correlação fornece melhores resultados. Obs.: Os valores no eixo y foram retirados por respeito ao acordo de confidencialidade.

IV.5.5 EFEITO DO NÚMERO DE ITERAÇÕES

O efeito do número de iterações pode ser visto como uma forma de avaliar o efeito da extensão do treinamento para o ajuste de um particular conjunto de dados.

As redes foram programadas para pararem quando uma das duas coisas acontecer:

- At atingir uma determinada tolerância (soma dos erros for menor que a tolerância)
- Ou atingir um certo número de iterações (logicamente significando que não atingiu a tolerância desejada) – como na maioria dos pacotes comerciais.

Importante notar que número de iterações significa o número de vezes que o conjunto de treinamento foi apresentado à rede para a otimização de seu pesos internos.

Todas as redes apresentadas neste trabalho tiveram o erro decrescendo com o número de iterações. Portanto para uma mesma rede, 100000 iterações implicam num erro maior que 200000 iterações nos casos em que a tolerância estipulada não foi satisfeita. A Figura 4.11 ilustra a influência do número de iterações da rede.

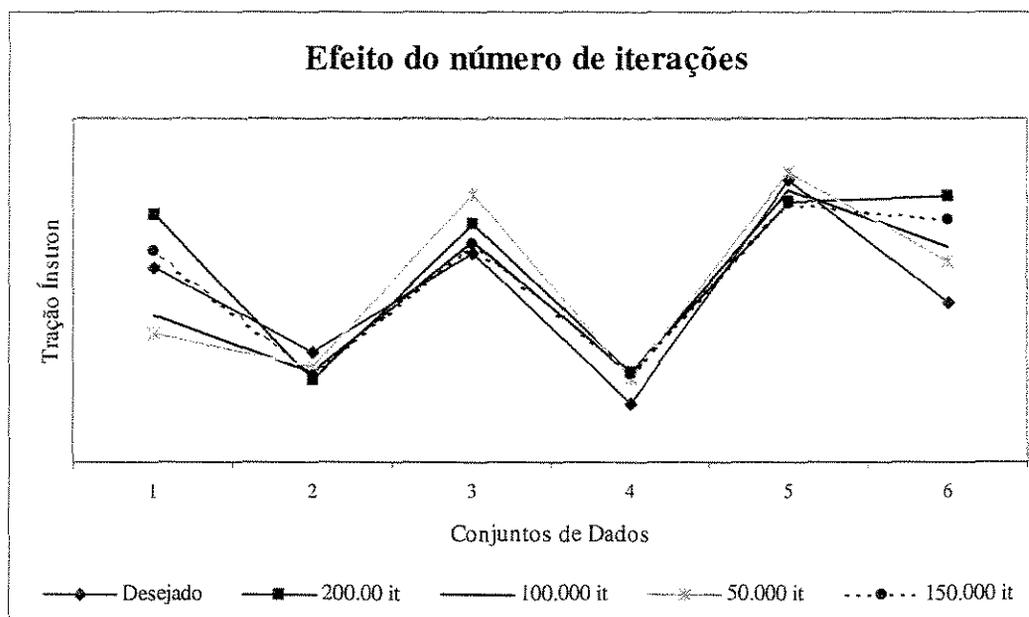


Figura 4.11 – Efeito do número de iterações - Redes com a mesma estrutura foram treinadas (com o mesmo conjunto de treinamento) para a predição do índice de tração até atingirem um certo número de iterações. Percebe-se que o aumento de 50.000 para 100.000 e depois para 150.000 iterações aumentou a eficiência de predição da rede. Entretanto, os resultados da rede treinada com 200.000 iterações foram piores que os anteriores, ilustrando um caso de *overfitting*. Obs.: Os valores no eixo y foram retirados por respeito ao acordo de confidencialidade.

A conclusão que se pode tirar através da observação da Figura 4.11, é que um maior número de iterações não necessariamente melhora a capacidade de treinamento da rede. Para a rede da figura em questão, o aumento de 50000 iterações para 100000

iterações e posteriormente para 150000 iterações causou uma melhora significativa nos resultados. Ao aumentar o número de iterações para 200000, a eficiência de predição foi perceptivelmente empobrecida.

Este comportamento pode ser enquadrado no que se conhece como *overfitting* (sobre ajuste), ou seja, a rede foi treinada em excesso, e ficou “viciada” em determinados valores

Mais uma vez, a quantidade e a qualidade das informações apresentadas à rede influencia o número necessário de iterações durante a fase de treinamento. Para propriedades diferentes, um número diferente de iterações foi utilizado.

IV.6 CONCLUSÃO

Este capítulo mostrou que o principal objetivo da tese – a construção de uma rede neural para a predição do número kappa a partir de dados industriais – foi atingido. Além disso, foram apresentados os passos para a otimização da rede, necessários para atingir um resultado satisfatório. A descrição qualitativa do programa da rede neural está incluída, bem como uma discussão dos parâmetros da rede e sua influência no resultado final.

Pode-se concluir dos fatos expostos neste capítulo que as técnicas utilizadas para a construção e aperfeiçoamento de cada rede devem ser específicas para o processo que está sendo modelado. Para um mesmo processo, a qualidade dos dados utilizados e a quantidade de dados para o treinamento são fundamentais na fase de treinamento e otimização da rede. No caso em estudo, o número de dados para treinamento foi pequeno, principalmente por se tratar de dados obtidos industrialmente. No entanto, o cuidado durante a coleta e análise das amostras, a obediência com relação aos tempos de residência e o acesso a um grande número de variáveis permitiu o sucesso da utilização das redes neurais para algumas das propriedades objetivo.

A qualidade preditiva de uma rede neural para o número kappa ou mesmo do índice de tração certamente pode ser melhorada através da coleta de um maior número de dados para treinamento e validação da rede.

No geral, fica claro que não se pode esperar que o desenvolvimento da rede esteja resumido à simples alimentação de dados para o treinamento. Esses dados deverão passar por um processo analítico extenso para que o conjunto de treinamento utilizado reflita as informações importantes, de maneira uniformemente distribuída, otimizando assim, o desempenho final da rede.

CAPÍTULO V

MODELO DETERMINÍSTICO E MODELO HÍBRIDO: COMPARAÇÃO E DISCUSSÃO SOBRE AS DIFERENTES TÉCNICAS DE MODELAGEM

V.1 INTRODUÇÃO

Este capítulo é dedicado à apresentação do modelo híbrido, que combina o conhecimento contido nos princípios físicos e químicos com dados industriais para o desenvolvimento de uma rede neural.

Inicialmente o modelo determinístico utilizado é apresentado. A análise de suas características principais, resultados e deficiências é feita, determinando o grau de participação deste modelo no modelo híbrido.

A seguir, estão descritos a estrutura, o desenvolvimento e os resultados do modelo híbrido.

A comparação e discussão sobre as três técnicas, suas aplicações e limitações estão abordadas no final deste capítulo.

**UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE**

V.2 MODELO MATEMÁTICO

Os modelos matemáticos descritos no Capítulo III apresentam características diferentes, apesar de procurarem descrever o mesmo processo. Todos eles apresentam pontos positivos e singulares, mas também deficiências. As características de cada modelo devem ser analisadas não apenas de acordo com a capacidade de simular certo processo, mas também de acordo com a finalidade do programa de simulação. Um programa a ser utilizado para controle deve ser capaz de reproduzir a realidade de forma rápida. Para tanto, o limite de tempo para a simulação às vezes é definido em segundos ou até mesmo, frações de segundo. Qualquer simulador destinado à utilização em tempo real na fábrica deve levar em conta o tempo e as variáveis de entrada requeridas. Seria impossível a utilização de um programa que exigisse a entrada de variáveis obtidas em análises não realizadas na fábrica. Já um programa destinado a pesquisa, ou projeto de equipamentos, pode ser executado em um tempo maior e utilizar dados provenientes de análises mais sofisticadas.

A finalidade do modelo matemático desta tese é a sua participação no programa híbrido. Por sua vez, o programa híbrido, assim como a rede neural pura, visam a possibilidade de ser utilizados em tempo real em uma fábrica de celulose, para a predição de uma propriedade do produto desta fábrica, em uma etapa inicial do processo.

O modelo de SALTIN (1992), apesar de moderno e de considerar vários fenômenos não abordados pelos outros autores (movimento dos cavacos no digestor), mostrou-se complicado computacionalmente para a finalidade deste trabalho. O modelo de VANCHINATHAN e KRISHNAGOPALAN (1995) é baseado em medições realizadas no licor, o que poderia ser útil na fábrica, pois é baseado em análises mais simples que a coleta de amostra de cavacos no interior do digestor. Apesar disso, normalmente as fábricas não costumam realizar as análises propostas. O modelo da difusão e penetração do licor, de GUSTAFSON et al. (1989), é muito útil para programas de pesquisa ou projeto, onde se deseja retratar a realidade da forma mais completa possível, mas que para os fins desta tese adiciona tempo computacional extra, o que não é desejado.

O modelo de MIRAMS e NGUYEN (1994) foi portanto o escolhido como base inicial para o desenvolvimento do modelo determinístico deste trabalho. Em primeiro lugar, apresenta uma abordagem moderna não muito utilizada na indústria, o que torna sua análise atraente. Ainda assim, é a mesma abordagem de SALTIN (1992), que se preocupou em capturar o máximo de fenômenos possível em seus estudos. Porém MIRAMS e NIGUYEN (1994) tratam as equações cinéticas e difusão de maneira mais simples, possibilitando o desenvolvimento de um programa mais rápido. Além disso, este programa foi desenvolvido para eucalipto, mesma matéria prima da fábrica em estudo.

Este modelo considera:

- As reações simultâneas de todos os componentes da madeira:

3 tipos de ligninas – uma que reage rapidamente, outra lenta e uma terceira não reativa

2 tipos de hemiceluloses

2 tipos de celuloses

Extrativos (normalmente não considerados em outros modelos, por apresentarem reações muito rápidas. No entanto, é importante ressaltar que eles exercem efeito no consumo de reagentes do licor)

- As reações são irreversíveis e os produtos não interferem com as outras reações.
- Todo o Na_2S está presente como íon sulfeto, que não é consumido na reação. Para compensar, a concentração de NaOH é igual ao álcali efetivo (conceito descrito no Capítulo II).
- O transporte dos reagentes dentro dos cavacos, através da difusão

Neste caso, tanto para o íon hidróxido, quanto para o íon sulfeto, o mesmo coeficiente de difusão foi utilizado.

A espessura é a única dimensão considerada, pois é a dimensão crítica para a difusão dos reagentes dentro dos cavacos, já que é significativamente

menor que a largura e o comprimento e as difusividades são iguais em todas as direções (GUSTAFSON et al. 1989).

A resistência à transferência de massa entre o licor e o cavaco é insignificante.

Os cavacos estavam completamente impregnados antes do período de aquecimento.

- A transferência de calor nos cavacos é instantânea e o calor de reação é negligenciável.

V.2.1 EQUAÇÕES DO MODELO

$$\frac{dL_1}{dt} = k_{L1} \exp\left(\frac{E_{L1}}{RT}\right) [\text{OH}]^a [\text{HS}]^b L_1 \quad (5.1)$$

$$\frac{dL_2}{dt} = k_{L2} \exp\left(\frac{E_{L2}}{RT}\right) [\text{OH}]^a [\text{HS}]^b L_2 \quad (5.2)$$

$$\frac{dH_1}{dt} = k_{H1} \exp\left(\frac{E_{H1}}{RT}\right) [\text{OH}]^a H_1 \quad (5.3)$$

$$\frac{dH_2}{dt} = k_{H2} \exp\left(\frac{E_{H2}}{RT}\right) [\text{OH}]^a H_2 \quad (5.4)$$

$$\frac{dC_1}{dt} = k_{C1} \exp\left(\frac{E_{C1}}{RT}\right) [\text{OH}]^a C_1 \quad (5.5)$$

$$\frac{dC_2}{dt} = k_{C2} \exp\left(\frac{E_{C2}}{RT}\right) [\text{OH}]^a C_2 \quad (5.6)$$

$$\frac{dE}{dt} = -0.6E \quad (5.7)$$

$$\frac{\partial \text{OH}}{\partial t} = \frac{1}{X^2} \frac{\partial}{\partial x} \left(D \frac{\partial \text{OH}}{\partial x} \right) + \frac{\rho}{0.1\text{E}} \times \left[3.2 \times 10^{-3} \frac{dL}{dt} + 6.1 \times 10^{-3} \left(\frac{dH}{dt} + \frac{dC}{dt} \right) + 7.0 \times 10^{-3} \frac{dE}{dt} \right] \quad (5.8)$$

$$\frac{\partial \text{HS}}{\partial t} = \frac{1}{X^2} \frac{\partial}{\partial x} \left(D \frac{\partial \text{HS}}{\partial x} \right) \quad (5.9)$$

$$D = 60 \times 10^{-4} \sqrt{T} \exp\left(\frac{-22.3 \times 10^3}{RT}\right) \times (1.24 \times 10^{-2} - 1.12 \times 10^{-4} \times [L + H + C + E]) \quad (5.10)$$

V.2.2 MÉTODOS DE SOLUÇÃO

A solução do modelo foi baseada no trabalho de COSELLI (1999). As equações 5.8 e 5.9 foram resolvidas através da técnica de colocação ortogonal (VILLADSEN, 1967) que foi usada para a discretização das equações diferenciais parciais que descrevem o comportamento dos reagentes (OH^- e HS^-). Desta forma foram calculadas a difusão e concentração dos reagentes em cada ponto da espessura do cavaco.

A média das concentrações dos reagentes foi utilizada nas equações dos componentes da madeira.

O sistema de equações diferenciais ordinárias foi resolvido pelo integrador LSODE.

V.2.3 RESULTADOS DO MODELO DETERMINÍSTICO

O modelo permite a avaliação do comportamento do perfil de concentração dos componentes da madeira e dos reagentes com o tempo. Para tanto, foram realizadas alterações nos valores de entrada, que representam as condições da reação, e também nos parâmetros das equações, para avaliar a importância dos fenômenos presentes, de acordo com a modelagem em questão.

A Figura 5.1 ilustra a concentração de carboidratos e a Figura 5.2 representa a concentração de lignina durante o processo de cozimento.

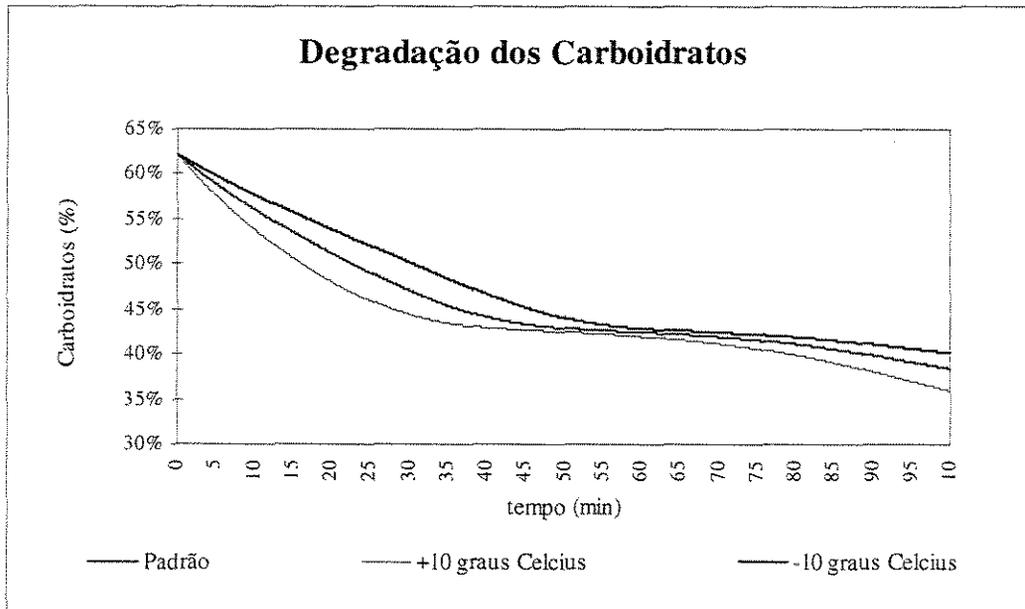


Figura 5.1 – Efeito da temperatura na taxa de degradação dos carboidratos durante a polpação - Para condições de cozimento semelhantes, o aumento da temperatura acelera a degradação dos carboidratos.

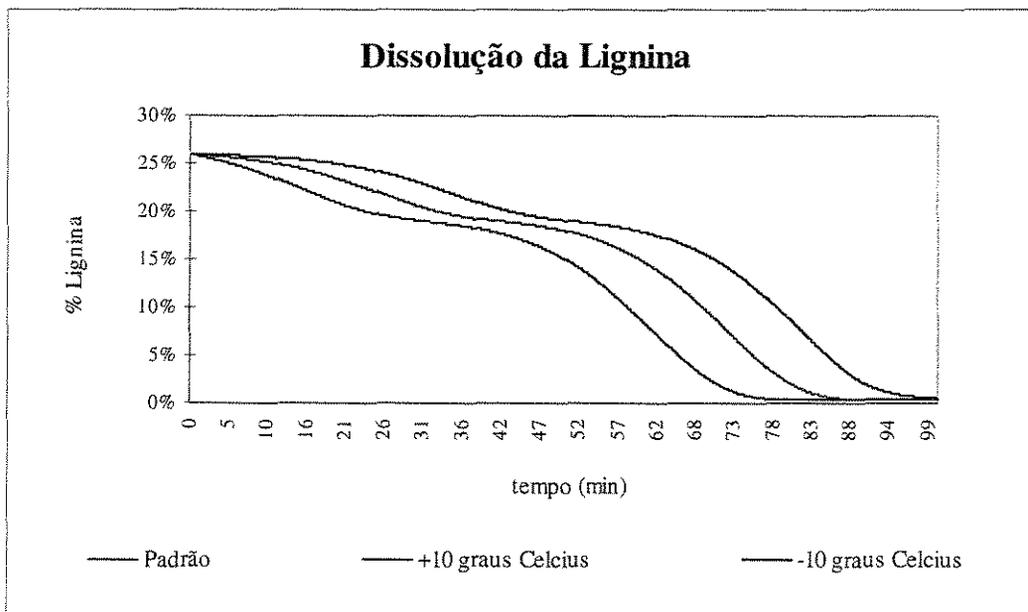


Figura 5.2 - Efeito da temperatura na taxa de deslignificação - A dissolução da lignina também é acelerada pelo aumento da temperatura.

A primeira observação é que a degradação dos carboidratos ocorre de forma assintótica. Os carboidratos são consumidos mais rapidamente no início da reação. Em seguida há uma desaceleração da taxa de degradação dos carboidratos. No final da reação, a taxa de degradação apresenta uma leve aceleração que coincide com a baixa concentração de lignina disponível para ser dissolvida. Este fato está de acordo com a teoria apresentada no Capítulo III e pode ser observado quando se compara a Figura 5.1 com a Figura 5.2.

A cinética da deslignificação descrita segundo a abordagem utilizada neste modelo apresenta perfil semelhante aos obtidos pela técnica clássica das três fases, onde a primeira fase é mais branda, a fase intermediária é mais drástica e a deslignificação é mais efetiva, e a fase residual, representada novamente por um perfil suave, refletindo a queda na velocidade da reação devido à baixa concentração de reagentes e lignina na madeira. Esta é uma constatação importante, pois o perfil de concentração é fato consolidado por inúmeros experimentos empíricos.

Para a comparação da influência relativa das variáveis, foram utilizados dois conjuntos de valores padrão normalmente encontrados na indústria, com diferentes níveis de carga alcalina, e para cada caso, vários parâmetros e variáveis foram alterados individualmente. Foram testados o efeito da concentração de reagentes, da temperatura e da difusividade dos reagentes na madeira. Em cada caso, a variável em questão foi alterada individualmente (mantendo as demais iguais ao perfil padrão) e os resultados podem ser vistos nas Figuras 5.2 a 5.4.

A temperatura e a concentração de reagentes (carga alcalina) foram as variáveis que exerceram efeito mais significativo na deslignificação. As Figuras 5.1 e 5.2, expostas anteriormente, mostram que o aumento de temperatura acelera tanto a taxa de deslignificação quanto a degradação dos carboidratos. Este fato incentiva a pesquisa de cozimentos que obtenham o mesmo grau de deslignificação com temperaturas mais baixas, evitando ao nível de degradação fim de se obter polpas com melhores propriedades de resistência.

A concentração dos reagentes afeta a deslignificação de maneira semelhante como mostra a Figura 5.3, onde um tempo consideravelmente maior é necessário para atingir

um mesmo grau de deslignificação quando a concentração de reagentes aplicada (carga alcalina) é menor.

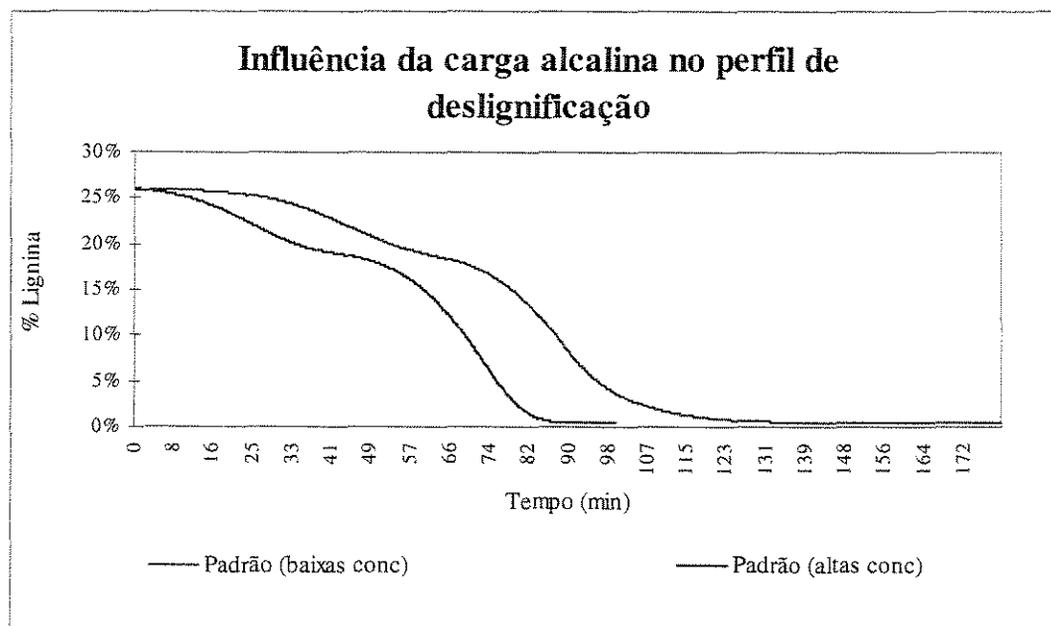


Figura 5.3 – Efeito da carga alcalina sobre a taxa de deslignificação - O aumento da concentração de reagentes acelera a dissolução da lignina.

A difusividade foi alterada em 50% (para cima e para baixo), para o estudo da influência deste fenômeno no processo. No entanto, não se observou alterações na velocidade da reação para este modelo em particular e para as condições operacionais adotadas. Este fato pode ser explicado devido às informações incompletas a respeito da difusão. O aumento da sulfidez mostrou um aumento não linear na taxa de deslignificação. As influências da difusividade e da sulfidez estão resumidas na Figura 5.4.

Apesar do objetivo do cozimento ser uma alta taxa de deslignificação, a degradação dos carboidratos implica na queda de rendimento e qualidade da polpa. Como foi citado no Capítulo II, o tamanho das cadeias de carboidratos está intimamente relacionado com a resistência da celulose e pode ser medido através da viscosidade. Uma queda na viscosidade pode significar redução nas vendas do produto para seu mercado alvo.

Portanto ao se otimizar a carga alcalina e temperaturas aplicadas ao processo, não se pode analisar apenas a deslignificação, deve-se considerar o compromisso entre número kappa (deslignificação) e a degradação dos carboidratos (viscosidade) exigidos.

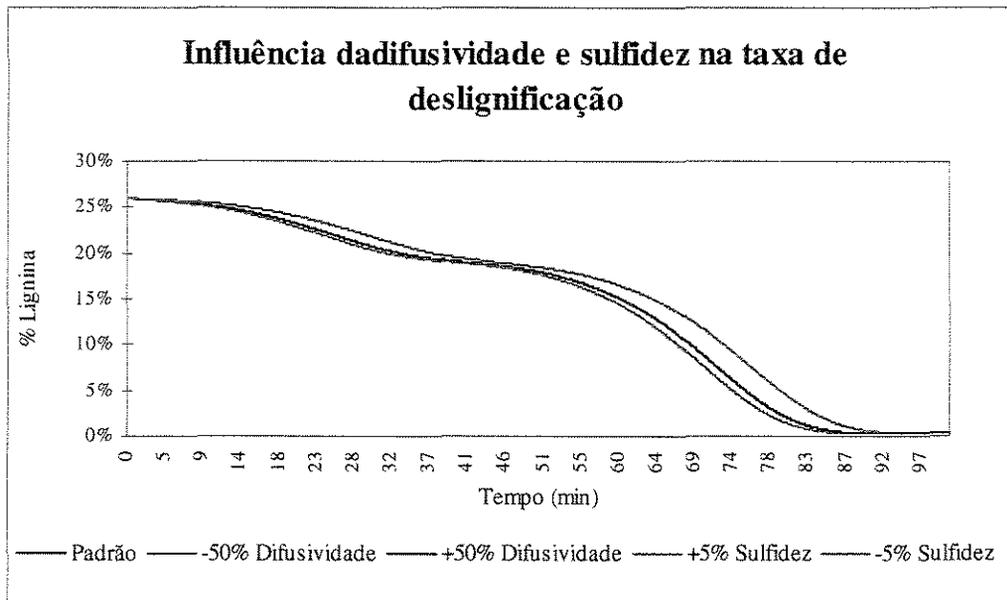


Figura 5.4 – Efeito da difusividade e sulfidez na taxa de deslignificação. As curvas que representam a variação na difusividade coincidem com a curva padrão, indicando que com este modelo a difusividade não apresenta um impacto significativo sobre a taxa de reação. A taxa de dissolução da lignina aumenta com o aumento da sulfidez, porém numa proporção menor que com o aumento de temperatura ou concentração de reagentes.

Observa-se nas figuras previamente apresentadas neste capítulo que apenas 100 minutos de reação foram suficientes para consumir toda a lignina presente na madeira. Isto no entanto, não corresponde à realidade industrial, onde os tempos de retenção são maiores, e sempre há lignina remanescente. As considerações de uma impregnação perfeita, das reações serem irreversíveis, a geometria e tamanho dos cavacos retratam uma situação de idealidade, e são aspectos que contribuem para este resultado. Além disso, as taxas de difusividade foram simplificadas e, principalmente, este modelo foi

desenvolvido para um digestor batelada de laboratório, e não leva em consideração as diferentes zonas e temperaturas. A real taxa de delignificação final deve estar em algum ponto da faixa de delignificação residual, dependendo das condições operacionais. No entanto, o modelo reproduz qualitativamente a polpação da madeira.

V.2.4 DISCUSSÃO

As observações descritas acima confirmam a capacidade do modelo para prever, pelo menos qualitativamente, as reações envolvidas na polpação. Além disso, a temperatura e a carga de reagentes são os fatores que mais afetaram a taxa de delignificação coincidindo com a maioria das variáveis utilizadas no modelo neural. Isto demonstra que, apesar de não exigirem conhecimento teórico aprofundado, as redes neurais precisam ser treinadas com variáveis que reflitam este conhecimento.

Duas características importantes deste modelo determinístico foram consideradas para avaliar sua adequação aos propósitos deste trabalho. Em primeiro lugar, o modelo é capaz de prever o processo de forma qualitativa, isto é, a taxa de delignificação responde às alterações impostas ao processo de forma semelhante à realidade. O modelo mostrou simplicidade e rapidez suficientes para que aplicado em tempo real. Devido à capacidade de atender a estes critérios, o modelo foi considerado satisfatório e portanto incorporado ao modelo híbrido, descrito na secção a seguir.

No entanto, o modelo matemático apresenta várias deficiências, do ponto de vista quantitativo.

O modelo considera a adição de toda a carga de reagentes no início do cozimento. Este tratamento é válido para cozimentos em batelada e cozimentos contínuos convencionais, cada vez menos utilizados. Fábricas modernas (incluindo a Aracruz Celulose, que cedeu os dados para o desenvolvimento desta tese), ou que passaram por um processo de modernização, tendem a utilizar diversas formas de cozimento modificado. Esta classe de cozimento inclui uma aplicação setorial de álcali. No início

do processo, a madeira entra em contato com apenas uma porcentagem do total de reagentes. Parte do material orgânico dissolvido e reagentes consumidos podem ser retirados do digestor em diferentes fases de cozimento e o restante dos reagentes é adicionado em etapas mais avançadas do processo. Com esta operação, o perfil de reagentes e componentes da madeira não mais obedece o comportamento descrito pelo modelo. Deve-se notar, entretanto, que em termos de resultados finais, para deslignificação, modelos tradicionais podem ter seus parâmetros alterados e ainda hoje são usados pelas fábricas para cada etapa do cozimento.

Outro fator não considerado pelo modelo é a hidrodinâmica e o transporte de momentum e massa no vaso. Um digestor industrial é normalmente muito alto e possui alterações em seu diâmetro para facilitar o deslocamento dos cavacos. As extrações alteram a direção do fluxo de licor, causando a formação de zonas com cozimento co-corrente e zonas com fluxo contra-corrente. Além disso, as propriedades da madeira são alteradas à medida em que as reações vão ocorrendo. Isso implica em diferentes taxas de difusão, diferentes porosidades, diferentes taxas de compactação da madeira e mais alteração no perfil de concentração dos reagentes (direção de fluxo). Todos estes aspectos influenciam o transporte de reagentes nos cavacos, a formação de canais preferenciais, a cinética da reação, etc.

O modelo ainda não inclui transferência de calor e o conseqüente efeito das diferentes temperaturas nas diversas partes do reator.

Empresas dedicadas ao fornecimento de tais equipamentos desenvolvem modelos mais complexos, que levam em conta vários desses fatores. Porém, quanto mais completos, mais lentos estes programas se tornam, perdendo a capacidade de serem utilizados em tempo real.

V.3 MODELO HÍBRIDO

V.3.1 MOTIVAÇÃO

Após a leitura do Capítulo III ficou evidente que apesar das redes neurais e modelos determinísticos serem ótimas ferramentas para modelagem, ambas as técnicas apresentam limitações. A comparação completa entre os modelos está descrita no final deste capítulo, mas para que se justifique o uso do programa híbrido, há que se mencionar alguns aspectos.

A rede neural pura é uma ferramenta poderosa, mas que requer longo tempo de treinamento e principalmente dados confiáveis. Muitas vezes isso significa um preço alto em termos de tempo e recursos humanos. Por exemplo, se o cavaco for analisado às oito horas da manhã, entrando no digestor, é possível que a polpa seja coletada só às quatro da tarde, dependendo da taxa de produção da fábrica, devido ao longo tempo de residência no cozimento. Se houvesse necessidade de coletar amostras em algum estágio mais adiante no processo, isso significaria a disponibilização de homens-hora além do horário administrativo. Com as redes neurais não é possível determinar a exata influência de cada variável do processo, apenas analisando o valor dos pesos e fatores de tendência, pois a rede não é baseada nos princípios físico-químicos do sistema. Além disso, as redes não apresentam capacidade preditiva para situações operacionais fora das condições nas quais o treinamento foi conduzido. Ainda, não há nenhum dispositivo em um programa neural puro que garanta que os resultados da rede irão obedecer aos princípios físicos, como por exemplo, que nunca haverá uma concentração negativa, ou que a soma das composições de todos os componentes de uma mistura seja igual a 1 (100%).

Por sua vez, os modelos determinísticos têm que ser simplificados na maioria dos casos para que possam ser utilizados em um tempo computacional adequado. Ou ainda, necessitam de dados não normalmente medidos para a estimativa de seus parâmetros. A obtenção destes dados pode ser muitas vezes impossível ou não viável economicamente. Às vezes, simplesmente, não se sabe o suficiente sobre o processo para o desenvolvimento de um bom modelo teórico.

A decisão de simular o processo através de um modelo híbrido, e a estratégia adotada neste modelo, devem depender da qualidade e da quantidade das informações que o pesquisador têm em mãos, como foi citado no Capítulo III.

V.3.2 ESTRUTURA DO MODELO HÍBRIDO

No caso deste trabalho, a qualidade e quantidade de informação pode ser descrita pelos seguintes fatos:

- Modelo determinístico
 - Não inclui todos os fenômenos ocorridos no processo, implicando em pouco conhecimento teórico
 - Apresenta boa reprodução qualitativa
 - Sua reprodução quantitativa não é satisfatória
 - É rápido e simples
 - Parâmetros inicialmente estimados para outra fábrica, podem ser modificados de maneira simples para este caso
 - Alguns parâmetros utilizados foram aproveitados de outros estudos, para outros tipos de madeira
 - Parâmetros baseados em cozimento em batelada
- Modelo de rede neural pura
 - Os dados de treinamento são confiáveis
 - Havia poucos dados disponíveis para o treinamento
 - A predição através da rede pura foi satisfatória

De acordo com os dados acima, o modelo neural apresentou resultados mais confiáveis que o modelo determinístico. Isto enquadra as informações disponíveis para este trabalho no caso em que o nível de conhecimento teórico deste modelo

determinístico em particular é incompleto e a qualidade dos dados industriais a serem usados no treinamento da rede é boa.

A estratégia do modelo híbrido, ilustrada na Figura 5.5, foi escolhida com base nesta análise. Devido à menor confiança no modelo determinístico, sua função é conferir um significado físico para parte do modelo, enfatizando a tendência do resultado na direção certa, segundo previsto teoricamente. A mesma rede desenvolvida para o modelo de rede pura foi utilizada – isto é, o mesmo conjunto de treinamento e os mesmos dados iniciais (nove variáveis de entrada) foram utilizados.

As variáveis de processo (reagentes, características da madeira, temperaturas, tempo de retenção) são os dados de entrada no modelo determinístico, que por sua vez gera um resultado, o Kappa teórico. Neste caso, o perfil de concentração não é importante, mas sim a concentração final de lignina, proporcional ao número Kappa que se deseja prever. A proporcionalidade entre o Kappa e a concentração de lignina é descrita pela equação (KYRKLUNG, B e STRAHOGLL, G, 1969):

$$Kappa\ Teórico = \% Lignina \times 6.064 \quad (5.11)$$

O Kappa teórico entra na rede como uma décima variável de entrada, juntamente com as outras variáveis utilizadas na rede pura. A rede híbrida também tem 12 e 10 neurônios nas duas camadas ocultas e uma saída, o número Kappa.

O treinamento da rede híbrida foi feito segundo o mesmo algoritmo usado para o modelo neural descrito no Capítulo IV, com a diferença da camada de entrada possuir 10 neurônios, ao invés dos 9 que compunham a rede pura.

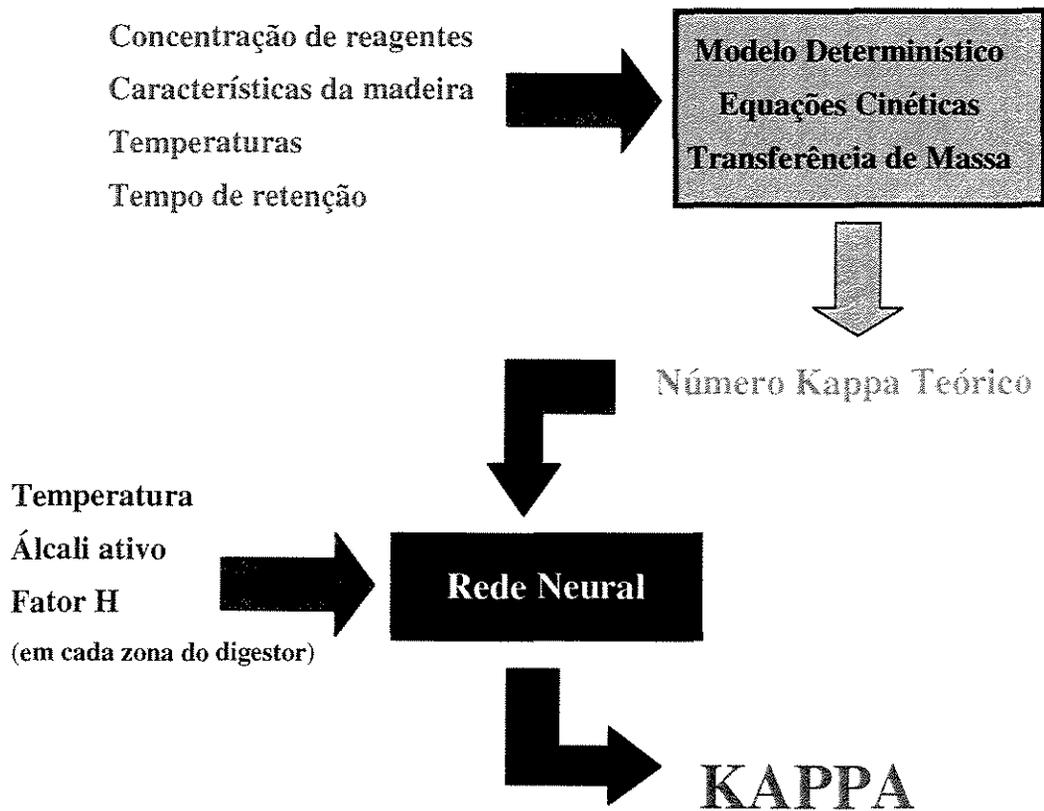


Figura 5.5 – Estrutura do modelo híbrido - As características da madeira e condições do cozimento são utilizados como valores de entrada do modelo determinístico, cujo resultado, a concentração de lignina remanescente é proporcional ao número Kappa, chamado aqui, Número Kappa Teórico. Este Kappa teórico juntamente com as condições de cozimento são então os valores de entrada da rede neural, cujo resultado final é o número Kappa predito com o modelo híbrido.

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

V.3.3 RESULTADOS

Os resultados obtidos com a rede híbrida podem ser vistos na Figura 5.6, que retrata o kappa real versus o kappa previsto. Os desvios estão todos dentro de uma faixa aceitável (Um ponto abaixo de 5% e todos os outros abaixo de 2%), menor que o erro experimental esperado na análise do número kappa.

O treinamento desta rede foi feito com o mesmo número de conjuntos de treinamento que a rede pura. A rede híbrida necessitou 50% menos iterações para que fosse treinada dentro da mesma tolerância da rede pura.

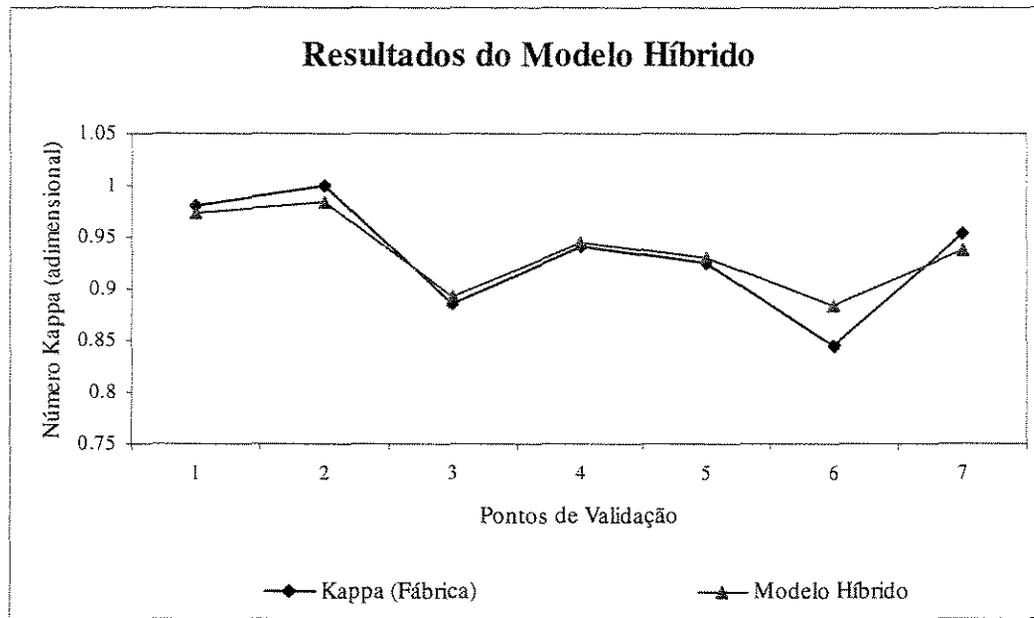


Figura 5.6 – Resultados do modelo híbrido - O gráfico acima compara valores reais da fábrica com os resultados do modelo híbrido.

V.4 DISCUSSÃO SOBRE DIFERENTES TÉCNICAS DE MODELAGEM – REDES NEURAI, MODELOS DETERMINÍSTICOS E REDES HÍBRIDAS

V.4.1 GRAU DE CONHECIMENTO TEÓRICO

Para o desenvolvimento das redes neurais foi necessário ter um mínimo conhecimento técnico sobre o processo. A escolha das 22 variáveis coletadas sobre o cozimento e o conceito de tempo de residência/volume de controle formam o conjunto de

conhecimentos necessários sobre o cozimento da celulose. Era conhecido que a madeira, o licor de cozimento, as temperaturas, etc., influenciam o resultado final. Além disso, o cálculo do tempo de residência foi considerado para que se acompanhasse o “mesmo” volume de controle de celulose foi considerado. Isto significa que apesar de não requerer muita pesquisa de cunho teórico, o conhecimento obtido de teoria superficial e principalmente através da experiência prática, foram fundamentais para o desenvolvimento do modelo neural. Não se pode dizer portanto, que as redes neurais são programas que não requerem nenhum conhecimento sobre o processo, ou se estaria afirmando que elas poderiam tratar qualquer conjuntos de dados, independentemente das características operacionais do processo.

No entanto, quando comparadas aos modelos matemáticos, as redes neurais requerem um **grau de profundidade** de conhecimentos significativamente menor. Os modelos matemáticos envolvem cuidadoso estudo sobre o conjunto de fenômenos físico-químicos do processo. No caso do digestor, esses fenômenos são complexos, inter-relacionados e simultâneos. Em primeiro lugar estão as reações cinéticas, que até hoje não estão completamente em consenso, considerando as diferentes abordagens cinéticas encontradas na literatura (algumas delas descritas no Capítulo III). A própria molécula da lignina não é completamente conhecida. Para que as reações ocorram, os cavacos devem ser impregnados com licor de cozimento – esta simples, porém fundamental etapa, envolve a penetração e a difusão dos reagentes dentro da madeira. Quando as reações são iniciadas, os componentes da madeira já dissolvidos também sofrem difusão do cavaco para o licor de cozimento. Cada tipo de madeira e composição de licor influenciarão na taxa de difusão dos reagentes e produtos. Estas reações são exotérmicas, provocando a elevação da temperatura. O desenvolvimento de cada parte deste processo foi e continua sendo tema de várias pesquisas, e ainda com o tempo, mudanças no equipamentos ocorrem, e novas descobertas sobre esse fenômenos são feitas. Os modelos matemáticos são baseados em uma parte desses fenômenos ou em todos eles. O modelo utilizado nesta tese considerou apenas as reações e parte da difusão dos reagentes – os resultados se mostraram qualitativamente aceitáveis. Fica óbvio no entanto, que para um bom modelo matemático, o grau de conhecimento teórico faz toda a diferença na qualidade dos resultados obtidos.

Os modelos híbridos devem ser construídos de forma a incluir o conhecimento sobre o processo de ambas as formas: dados industriais ou de simulação e equações fenomenológicas. Porém, maior ênfase deve ser colocada no tipo de conhecimento em que se tem maior confiança. Como demonstrado no Capítulo IV, os dados industriais obtidos para este trabalho foram considerados de excelente qualidade, enquanto que o conjunto de equações representativas do processo no modelo matemático escolhido foi considerado incompleto. Por este motivo, a parte neural foi proeminente no modelo híbrido.

V.4.2 OBTENÇÃO DE VARIÁVEIS UTILIZADAS PARA O DESENVOLVIMENTO DO MODELO

Ao desenvolver as redes neurais, o “programador” tem a liberdade de escolher quais as variáveis que irá utilizar para o treinamento da rede. Deve-se ter em mente, que as variáveis adotadas para o treinamento serão as mesmas variáveis empregadas como dados de entrada durante a execução da simulação. No caso da aplicação industrial é interessante e principalmente possível utilizar apenas variáveis normalmente medidas no processo, ou através de medidores em linha (por exemplo temperaturas, fluxos, etc.) ou provenientes de análises de rotina na fábrica (como a concentração do licor). Ao se observar as 9 variáveis utilizadas no desenvolvimento da rede neural fruto desta tese, todas elas são classificadas em uma destas classes. A variável a ser predita, no entanto, não precisa estar incluída nessas categorias, uma vez que encerrada a fase de treinamento, ela não será um dado de entrada no modelo. Na verdade se a variável puder ser facilmente medida, não é necessário que seja predita por um *software*.

Como foi discutido no Capítulo IV, uma dificuldade na coleta de dados industriais reside na necessidade da inclusão de dados em toda a faixa de operação considerada. Sabendo-se que na maioria das vezes as fábricas buscam operar em faixas próximas à produção nominal, ou produção máxima, a aquisição de dados nos pontos de baixa produção pode ser uma tarefa com raras oportunidades para ser realizada ou com alto

custo para a empresa. Este fator é agravado pela necessidade de se ter o processo estável enquanto os dados estão sendo coletados.

Para o modelo determinístico, por outro lado, várias propriedades **não** medidas em uma fábrica são necessárias para a estimativa dos parâmetros. Por exemplo, para a estimativa das taxas e mecanismo da reação (para cada componente), é necessário retirar amostras de cavaco e licor em determinados intervalos de tempo, a fim de se construir a curva de desaparecimento dos reagentes. Isso não seria prático em um digestor industrial, pois não há pontos de amostra suficientes. Fora isso, as taxas de difusão, de transferência de calor, etc. também exigem inúmeras amostras. Felizmente, na hora de utilizar o modelo para simulação, estes dados não são necessários. Entretanto, por mais cuidadosamente que o modelo tenha sido desenvolvido, e os parâmetros estimados, para cada tipo de digestor e/ou tipo de madeira, parte do trabalho de amostragem e estimativa de parâmetros deveria ser repetido, uma vez que estes dependem do projeto do digestor e do tipo de madeira. A estimativa de parâmetros para modelos determinísticos pode depender de muitos cozimentos de laboratório e outras análises com custos igualmente altos.

As redes neurais exigem re-treinamento ao serem transportadas para uma outra fábrica. A coleta de amostras para o novo treinamento requer uma quantidade de homens-hora relativamente menor do que a nova estimativa de todos os parâmetros de um modelo matemático. Caso a fábrica não tenha pressa e o processo de coleta de dados seja gerenciado de forma a não exigir testes extras, além dos normalmente executados para acompanhamento e controle de processo. Por outro lado, muitas vezes ao se adaptar um modelo determinístico a uma outra fábrica, nem todos os parâmetros precisarão ser alterados. A necessidade de alterações será determinada pelas condições do processo da fábrica, modelo utilizado e precisão esperada.

Em suma, a coleta de dados para o treinamento de redes neurais é realizada num longo intervalo de tempo, mas os dados podem ser facilmente coletados se incorporados ao boletim de ocorrência da fábrica. A coleta de dados para o modelo determinístico normalmente apresenta maiores dificuldades e custo, mas seus parâmetros apresentam maior chance para a reutilização em situações de alterações no processo.

V.4.3 QUANTIDADE DE DADOS NECESSÁRIOS PARA O DESENVOLVIMENTO DO PROGRAMA

No caso das redes neurais, isso depende muito da qualidade dos dados obtidos, da distribuição dos dados na faixa de treinamento, do processo em si e inter-relações entre as variáveis e da variável (eis) a ser calculada (s). Normalmente as redes neurais requerem uma grande quantidade de dados para o treinamento, às vezes vários meses de produção. Isso é importante para que os dados estejam o mais uniformemente distribuídos pela faixa e para evitar tendências mais fortes em certas regiões. Também para que uma ampla faixa de operação seja incluída no modelo, para permitir a eliminação de pontos fora dos padrões esperados, e para diminuir a influência de ruídos.

Os modelos matemáticos requerem apenas o número de dados suficientes para a estimativa dos parâmetros e o entendimento dos mecanismos. A partir desse conhecimento é importante não se generalizar que as redes necessitam muitos dados e os modelos matemáticos necessitam poucos dados.

Como foi visto neste trabalho, com um número pequeno de conjuntos de dados, foi possível construir uma rede de predição de número kappa. Com os mesmos conjuntos de dados, não foi possível predizer outras propriedades, provavelmente devido ao baixo nível de sensibilidade às variáveis de entrada pesquisadas. Ou seja, a quantidade de dados necessários para o treinamento da rede varia de caso para caso.

Igualmente, a quantidade de dados necessários para o desenvolvimento de um modelo matemático vai ser diferente de acordo com o grau de conhecimento teórico e complexidade desejado. Um modelo que considera apenas alguns fenômenos irá exigir uma quantidade de dados bem menor que um modelo mais completo. Partes lineares irão exigir um menor número de dados que equações mais complexas e não lineares.

Um outro ponto a ser considerado na avaliação da quantidade de dados é o custo para a obtenção de dados para o treinamento da rede versus o desenvolvimento do modelo teórico. Normalmente os dados para o desenvolvimento do modelo teórico são mais caros, principalmente com as possibilidades de gerenciamento de informações presentes nas fábricas modernas.

Para esta tese, a falta de parâmetros previamente estudados verdadeiramente para o Eucalipto tornariam a coleta de dados para o desenvolvimento de um modelo matemático extremamente caro. Enquanto que o bom sistema de informação presente na empresa possibilitou a realização da coleta de dados forma simples e relativamente barata. Além disso, os dados coletados puderam ser utilizados em outros estudos.

V.4.4 QUALIDADE DOS DADOS

A teoria das redes neurais indica que elas podem ser treinadas com dados contendo ruídos, ou às vezes com conjuntos incompletos de dados, uma vez que a informação é distribuída pelos vários neurônios, que por sua vez apresentam apenas uma pequena influência no resultado geral. Pela lógica, quanto maior o conjunto de dados de treinamento, menor será a influência dos ruídos no resultado final. No entanto, a eliminação de ruídos irá certamente melhorar o desempenho da rede e possibilitar o treinamento do modelo com um menor conjunto de dados.

Se um dos medidores for tendencioso, isto é, apresentar um desvio constante, a rede vai incorporar este desvio no cálculo dos resultados. No entanto, quando esta variável for utilizada como dado de entrada, o desvio continuará existente, e pela rede ter sido treinada com valores produzidos pelo medidor, o resultado final será o correto. Para casos desse tipo, no entanto, sugere-se que a fábrica tenha a mesma política de calibração dos medidores durante e após a coleta de dados de treinamento. Uma fábrica que não possui uma política de calibração dos equipamentos, poderia ter medições incorretas e provavelmente dados provenientes desta fábrica não permitiriam um treinamento apropriado da rede neural. Da mesma forma, uma fábrica que tivesse a preocupação de calibrar os medidores durante a fase de coleta de dados para o treinamento da rede, mas não transformasse este hábito em um procedimento no futuro, não faria bom uso da técnica de redes neurais. Mas neste caso, também não faria bom uso dos modelos matemáticos.

No caso do desenvolvimento dos modelos determinísticos, quanto mais precisos os dados, mais próximos da representação dos fenômenos do processo. Além disso, os

dados são obtidos normalmente com estudos de laboratório, cujos resultados não serão os mesmos que as variáveis de processo utilizadas como dados de entrada. Se as análises não forem precisas, o resultado final irá apresentar um alto desvio em relação ao resultado real. A exigência do nível de qualidade do dado irá ser maior de acordo com a maior sensibilidade do processo em relação ao parâmetro que o utiliza.

Para ambos os casos, quanto mais precisos forem os dados, maior será a capacidade de predição do modelo.

O fato do modelo determinístico utilizado nesta tese não ter apresentado bons resultados quantitativos está também relacionado à má qualidade dos seus parâmetros. Afinal, os parâmetros não eram adequados para o sistema que tentavam simular.

Por outro lado, três fatores contribuíram imensamente para o desenvolvimento de um conjunto de dados de boa qualidade para o treinamento da rede neural. O primeiro fator foi o procedimento de coleta de dados que levou em conta o tempo de residência e a lavagem das amostras para que as reações não continuassem ocorrendo após terem sido coletadas. O segundo fator foi a minimização da influência de ruídos, inerentes aos sinais eletrônicos dos medidores, através da utilização do valor médio de cada variável num intervalo de 30 minutos. Por fim, a Aracruz Celulose possui uma política de calibração regular e por isso as relações entre as variáveis foram consistentes.

A boa qualidade dos dados possibilitou o desenvolvimento do conjunto de treinamento a partir das correlações entre as variáveis, utilizando poucos conjuntos de dados industriais, com a obtenção de excelentes resultados de predição através do modelo neural.

V.4.5 FAIXA DE VALIDADE DO MODELO

As redes neurais são limitadas pela faixa de operação para as quais foram treinadas. Esta limitação é maior conforme a inter-relação entre as variáveis se torna mais complexa. E obviamente, para sistemas altamente lineares e simples, não seria necessário treinar uma rede neural! Portanto, se a fábrica decidir operar em condições

não englobadas na faixa de treinamento, como casos de taxas de produção muito mais altas, ou composição de licor diferente, ou utilizar outro tipo de madeira, a operação diferenciada do digestor para produtos especializados, ou outras infinitas alterações possíveis, o modelo neural teria que ser retreinado para as novas condições operacionais. O mesmo ocorre quando a rede é transferida para uma outra fábrica.

A pequena quantidade de dados limitou a observação deste fenômeno nesta tese. Porém, o maior desvio entre o valor real e o valor predito pela rede neural foi observado para o valor de kappa mais baixo do conjunto de validação, indicando que a qualidade de predição da rede tende a piorar nas proximidades das extremidades da faixa de treinamento.

Os modelos matemáticos, baseados na teoria, normalmente permitem extrapolação, e uma variação nas condições de processo um tanto mais abrangentes, pois a simples alteração nas condições do processo não altera os fenômenos físicos e químicos que ocorrem no digestor. No entanto é importante saber como os parâmetros obtidos empiricamente afetam o modelo, para saber se as novas condições necessitam re-estimativa de constantes dentro do modelo. Obviamente que para transferir o modelo para outra fábrica, esta análise deve ser mais detalhada, ainda que ambas as fábricas operem em equipamentos semelhantes e tipos de madeira parecidos, pois até o tamanho do digestor influencia algumas variáveis.

V.4.6 INTERPRETAÇÃO DO MODELO

Analisando as equações do modelo matemático é fácil identificar que a temperatura afeta a taxa de reação de forma exponencial, a concentração dos reagentes afetam a uma determinada potência, a composição inicial da madeira é importante, etc. Em outras palavras, pela simples análise das equações de um modelo matemático, é possível se determinar qual a importância de determinado parâmetro ou variável no processo. Esse conhecimento é muito importante para que um engenheiro na fábrica decida otimizar o processo alterando esta ou aquela variável, ou para identificar possíveis falhas na sua política de operação, que podem ser facilmente resolvidas.

Ao se observar os neurônios de uma rede neural, seus pesos e fatores, é impossível decidir quem afeta mais a reação, a temperatura ou o clima, pois os neurônios não trazem informação de cada variável separadamente, mas do resultado da interação entre todas elas.

Isso não significa que não se pode estudar, com um programa de simulação de rede neural, como a alteração de um variável, ou de um conjunto de variáveis, irá afetar o processo. Provavelmente um usuário irá utilizar o mesmo processo para avaliar o efeito das alterações para ambos os modelos, ou seja, irá alterar as condições de operação e analisar o efeito no resultado final obtido, ao invés de tentar avaliar as equações por trás do modelo.

V.4.7 REDES NEURAIIS PURAS E HÍBRIDAS

A rede híbrida teve um grande peso do modelo neural, e é importante que se compare os resultados de ambos, uma vez que aparentemente não há grande diferença entre eles.

A literatura afirma que a quantidade de dados para o treinamento de uma rede híbrida pode ser consideravelmente menor que para o treinamento da rede pura. Este fato não pode ser observado neste trabalho, devido ao pequeno número de dados de treinamento.

Porém, confirmando o que foi dito antes (Capítulo III) sobre o tempo de treinamento, a rede híbrida pode ser treinada até uma determinada tolerância com um número menor de iterações.

Apesar da rede híbrida deste trabalho representar apenas a adição de um neurônio com o resultado do modelo determinístico, como um fator qualitativo, os resultados foram levemente superiores, como mostra a Figura 5.7.

As redes híbridas se apresentam como um ótimo recurso para a modelagem, minimizando as limitações de cada técnica e aproveitando o que elas têm de melhor. Podem ser combinadas várias técnicas, de maneiras diferentes – o importante é que se

analise todos os dados disponíveis para que se tome a decisão de como combinar as partes do modelo. Um modelo híbrido, apesar de cheio de promessas de melhoria da modelagem, pode apresentar um nível de complicação muito elevado para ser implementado computacionalmente.

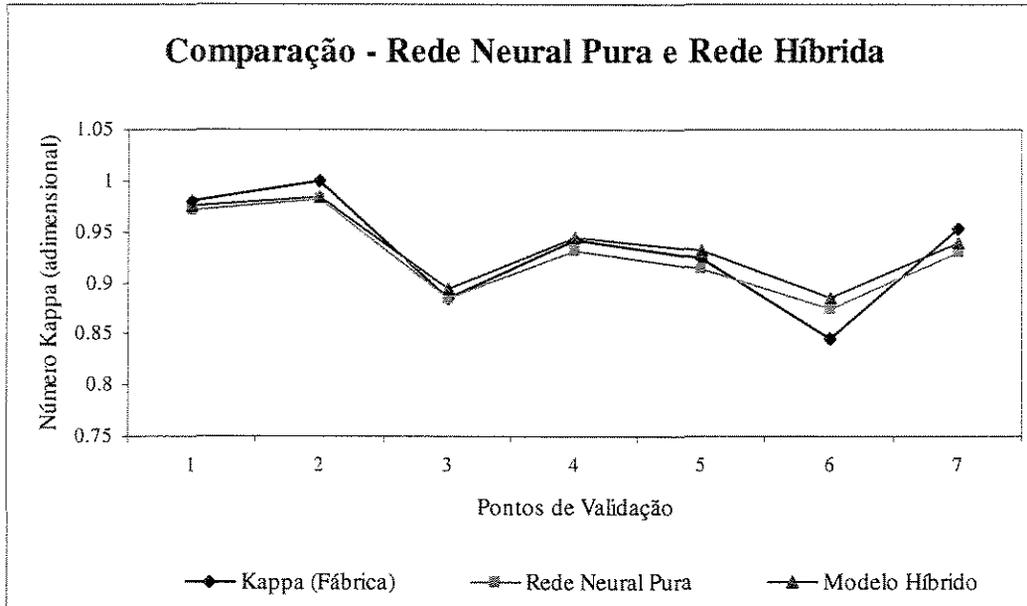


Figura 5.7 – Comparação entre os resultados obtidos com a rede neural pura e o modelo híbrido - O gráfico mostra que o modelo híbrido apresentou resultados marginalmente melhores que a rede pura, aproximando-se mais do valor real para a maioria dos pontos.

V.4.8 CONSIDERAÇÕES FINAIS

Todas as variáveis apresentadas acima devem ser consideradas ao se escolher a técnica de modelagem.

A modelagem de processos muito complexos e/ou cujo conhecimento é incompleto tende a ser mais eficiente através de redes neurais. Já em situações onde o conhecimento teórico é altamente consolidado, os modelos determinísticos tendem a ser mais robustos. Tudo isto vai depender entretanto da qualidade dos dados industriais ou de simuladores

utilizados para o treinamento da rede. Bons dados irão proporcionar o treinamento de uma rede neural de forma satisfatória, assim como dados ruins produzirão redes incapazes de predição em todas as condições, ou podem até impedir o treinamento da rede completamente.

A quantidade de dados necessária e o custo para a obtenção destes dados podem ser fatores determinantes para a escolha da técnica a ser empregada. Dados industriais requerem um tempo longo de coleta, mas podem ser retirados da operação de rotina, caso não seja necessária operação especial para as extremidades da faixa de operação. Dados de laboratório para ajuste de parâmetros podem ser rápidos, mas muitas vezes são muito caros.

O objetivo do modelo também influencia a escolha da abordagem, pois se deve ter em mente se será necessário aplicar em uma ou diversas instalações, o tempo computacional exigido para rodar o modelo, o tipo de informações necessárias para sua utilização na fase de aplicação e o grau de precisão dos resultados.

A análise desses aspectos irá indicar qual a técnica de modelagem mais favorável para cada caso – redes neurais ou modelos determinísticos. Em qualquer caso, as limitações remanescentes devem ser minimizadas. A modelagem híbrida, por outro lado, presenteia o cientista com a flexibilidade de utilizar todo o conhecimento sobre o assunto da maneira mais adequada em cada parte do modelo. E de combinar as técnicas de forma a melhor acomodar os objetivos e os custos de desenvolvimento.

Ao se planejar o desenvolvimento de um modelo, com qualquer técnica, deve-se levar em consideração os objetivos do modelo. Deve-se colocar na balança o grau de precisão, o tempo e o custo em se obter informações para a decisão de que técnica de modelagem deve ser usada.

Apesar da qualidade dos modelos neurais, deve-se sempre lembrar que a busca de melhores modelos determinísticos possibilita um maior conhecimento do universo.

V.5 CONCLUSÃO

Este capítulo apresentou inicialmente uma análise do modelo matemático e os motivos que governaram sua seleção para ser utilizado no modelo híbrido – o modelo foi capaz de reproduzir a deslignificação da madeira de maneira qualitativa e pode ser implementado de forma a não demandar elevados tempos computacionais, podendo ser utilizado para aplicações em tempo real.

A seguir foi apresentada a motivação para o desenvolvimento de um modelo híbrido e a descrição da estrutura do presente modelo – o programa principal é uma rede neural, onde um dos parâmetros de entrada é o resultado do modelo determinístico. A escolha foi feita devido à grande confiança nos dados industriais e ao incompleto volume de conhecimentos por trás do modelo determinístico. Os resultados mostraram que a predição da rede híbrida foi satisfatória.

A última parte deste capítulo foi dedicada à discussão e comparação entre as três técnicas, mostrando que as características de cada técnica não fazem da técnica a melhor opção, mas sim, fazem da técnica, a melhor opção para determinada finalidade.

CAPÍTULO VI

CONCLUSÕES E SUGESTÕES

VI.1 CONCLUSÕES

O objetivo científico desta tese foi o desenvolvimento de um programa de simulação utilizando redes neurais a partir de dados industriais. Como produto secundário, porém igualmente importante, o trabalho envolveu a comparação entre as técnicas de modelagem de processos – Redes Neurais, Modelos Determinísticos e Modelos Híbridos – com a intenção de plantar a semente da análise dos recursos em mãos para a decisão da técnica de modelagem a ser utilizada para cada caso específico.

O trabalho foi desenvolvido em colaboração com a Aracruz Celulose S.A, responsável pela maior produção mundial de celulose numa mesma fábrica, localizada no estado do Espírito Santo. O programa poderá ser utilizado pela fábrica para algumas tarefas principais:

- A predição do número kappa no início do processo, para a avaliação das condições de operação presentes e avaliação da necessidade de realizar alterações para a obtenção do kappa desejado. Esta pode ser considerada uma ferramenta importante quando se considera que às vezes 8 horas são necessárias para saber que faixa de kappa aquelas condições de operação foram capazes de produzir. Por exemplo, em uma fábrica de celulose moderna, isso pode otimizar a transição entre a produção de celulose branqueada com e sem reagentes clorados, que exigem diferentes faixas de kappa.

- Treinamento de jovens engenheiros e operadores – que poderão “brincar” com o programa para definir como o ajuste das condições de operação pode ser feito para que se obtenha o kappa objetivo, de maneira econômica.

Após a apresentação sobre a teoria das redes neurais e de como se desejava realizar o trabalho, os dados a serem coletados foram determinados em conjunto com a fábrica. Foi estabelecido que o tempo de retenção seria importante para a qualidade dos resultados, e os dados foram coletados de acordo com este conceito.

O treinamento da rede foi realizado na Unicamp, a partir do desenvolvimento e inúmeras alterações para a otimização de um programa neural em C⁺⁺.

A técnica de redes neurais, já existente há algumas décadas, apresentou nos últimos anos um crescente grau de interesse e foi escolhida para que se pudesse fazer uma avaliação de suas potencialidades. Do ponto de vista do processo, esta é uma técnica que não requer conhecimento teórico profundo, é considerada ideal para processo não lineares e/ou muito complexos. O processo de celulose pode ser considerado tanto complexo como não linear. Quanto à profundidade de conhecimento teórico, pode-se dizer que apesar de se saber muito sobre o processo, ainda existe muito campo a ser pesquisado e partes da teoria ainda necessitam de maior consolidação. Cada fábrica necessita de ajustes específicos e cada modelo existente adota um grau de complexidade diferente. Ainda, o tipo de informação possível de se obter em várias fábricas de celulose, com SDCD e PI, possibilita o desenvolvimento de programas de redes neurais a custos não muito altos.

As redes neurais apresentam um grande potencial para aplicações em tempo real, e pelas suas características, apresentam vantagens sobre os modelos determinísticos.

No entanto, ao se decidir pela modelagem neural, deve-se ter em mente que um pacote de redes neurais não é mágico. Isto é, alimentar o programa com uma grande quantidade de dados não irá garantir o sucesso do modelo. Para que o modelo seja consistente e eficiente, deve-se proceder com a otimização dos dados e dos parâmetros da rede, conforme a descrição no Capítulo IV.

Os três diferentes modelos foram construídos para fins de comparação. As técnicas de redes neurais e modelos determinísticos, apresentam seus pontos fortes e fracos e as nuances entre eles. A escolha entre um e outro deve ser feita de acordo com o conhecimento e a dificuldade para a obtenção de dados adequados para cada caso. Ainda, ambas as técnicas podem ser utilizadas conjuntamente, otimizando suas potencialidades e minimizando suas fraquezas, ou melhor dizendo, minimizando as deficiências encontradas pelo idealizador do modelo.

A rede híbrida foi construída com a intenção de demonstrar a potencialidade de otimizar os modelos desenvolvidos para um processo químico, fazendo o melhor uso dos dados e conhecimento que se tem em mãos. No caso deste programa, o conhecimento e a adequação do modelo teórico eram limitados, o que pode ser comprovado pelos resultados obtidos pelo modelo matemático. Porém em termos qualitativos, o modelo era aceitável, tornando-o adequado para o desenvolvimento do modelo híbrido proposto. Por outro lado, os dados industriais foram cuidadosamente coletados, sendo considerados de alta confiabilidade e qualidade. No desenvolvimento do modelo híbrido, optou-se primordialmente pela aproximação através das redes neurais.

Em resposta ao objetivo primordial do trabalho, os resultados confirmaram a expectativa de que as redes neurais artificiais são ferramentas poderosas na simulação de processos industriais cuja teoria é por demais complicada se explorada com alto rigor, ou cuja simplificação não produz resultados satisfatórios. No entanto, é importante notar que as redes neurais não desvalorizam a pesquisa e desenvolvimento de modelos matemáticos determinísticos, que muitas vezes produzem resultados excelentes, e mais importante, são capazes de gerar conhecimento científico em termos de engenharia química.

A idéia de fazer um trabalho em colaboração Universidade-Indústria, teve também o objetivo de mostrar que é possível fazer pesquisa de interesse industrial, que pode ser implementada numa fábrica. De maneira geral este tipo de acordo é financeiramente atraente para a indústria, que não necessita se estruturar ou pagar por recursos para pesquisas em facilidades comerciais, uma vez que a Universidade apresenta também a disponibilidade dos conhecimentos e atualização quanto à tecnologias de ponta. Para a

Universidade o sistema é também benéfico, uma vez que pode arrecadar fundos para a melhoria de suas instalações e aumentar a qualidade do produto oferecido – tanto em termos de pesquisas para indústrias ou pesquisa básica, mas principalmente na melhor formação de seus alunos, futuros profissionais da indústria. Ainda pode-se dizer, que um programa de colaboração Universidade-Indústria aproxima esses dois mundos, que muitas vezes se sentem tão distantes, e aumenta a sinergia da geração e aplicação de conhecimentos – como num modelo híbrido, utilizando a experiência e o senso de necessidade prática da indústria, com o conhecimento teórico mais profundo e a característica visionária da academia.

VI.2 SUGESTÕES

Infelizmente o tempo é um fator limitante e não é possível colocar todas as idéias em apenas um projeto de mestrado. Este trabalho pode ser apenas o início de uma série dedicada ao desenvolvimento de redes neurais e redes híbridas, e também ao estudo dos fenômenos envolvidos na fabricação do papel.

O desenvolvimento de um modelo híbrido mais consistente, onde se utilize mais toda a teoria já conhecida sobre o cozimento se mostra um desafio interessante. Este modelo poderia ser construído de diversas maneiras, mas a estrutura que se apresenta mais interessante utiliza as equações para os diversos fenômenos segundo encontradas na literatura, porém com alguns parâmetros calculados com redes neurais. Apesar de aparentar simplicidade, este seria um trabalho altamente complicado e delicado. A grande dificuldade seria, a partir de dados industriais de entrada, treinar uma rede que calcule estes parâmetros, cujos resultados só poderão ser testados com o resultado do modelo completo.

Outras formas de redes neurais também poderiam ser testadas, como por exemplo as redes recursivas capazes de realizar simulações dinâmicas.

O aumento de conjuntos de dados a partir de outras técnicas pode ser um bom campo de estudo. HENRIQUES et al (1999) estudaram uma técnica que usa um modelo impreciso do processo para gerar novos dados à partir de um pequeno número de dados experimentais e mostraram que a rede treinada com o modelo estendido apresentou melhor desempenho.

Ao apresentar a proposta para este trabalho em diversas fábricas, foi observado o interesse pelo desenvolvimento de um modelo neural que seja capaz de identificar falhas em instrumentos de análise de processo. Uma rede que por exemplo operasse em paralelo com um medidor de kappa *on-line*. Enquanto a calibração deste instrumento estivesse satisfatória, tanto a rede quando o analisador estariam indicando o mesmo resultado. Uma contínua divergência de resultados indicaria a necessidade de calibração do instrumento. Desta forma seria possível cortar o custo de calibrações periódicas não necessárias, ou de falsas leituras de instrumentos não calibrados o suficiente.

Um outro tema interessante, seria o desenvolvimento de uma rede que pudesse ser incorporada ao SDGD e fosse continuamente treinada, coletando informações em todos os pontos da faixa de operação da fábrica.

Do ponto de vista da fabricação de celulose, o campo é muito vasto – iniciando-se pela caracterização real da difusão dos reagentes no eucalipto; os fenômenos hidráulicos ocorridos no digestor, que parecem não ser estudados tão extensivamente; o real mecanismo de reação; a verdadeira relação entre viscosidade e resistência do papel; as causas e efeitos da reprecipitação da lignina e dos carboidratos; programas neurais que predissessem outras características do papel, etc. Um estudo que pode beneficiar fabricantes de celulose é a determinação dos tipos e quantidade de outras fibras que devem ser misturados à sua celulose para a obtenção de um papel com propriedades otimizadas. Ou ainda a determinação de como o processo pode ser otimizado para aumentar a proteção das fibras e/ou promover uma ligação entre fibras mais fortes. Dentro desta linha de pesquisa, deve-se incluir a busca de faixas de temperaturas mais baixas durante o cozimento.

A otimização do cozimento da madeira é um tema que pode gerar inúmeros benefícios para a indústria de celulose no Brasil e incrementar o conhecimento do

processo em termos de fenômenos estudados em engenharia química, desde a cinética da reação, até os transportes de massa, momentum e calor. O digestor é um sistema de leito fixo, que apresenta sempre fase líquida e fase sólida, e dependendo do caso, uma fase vapor. A pesquisa dedicada a espécies de madeira nacionais, com parâmetros que mais caracterizam os nossos processos e o desenvolvimento de meios de otimizar este processo, pode significar um passo a frente na vantagem competitiva da indústria de celulose brasileira no mercado internacional. O desenvolvimento de pacotes de otimização na Universidade pode intensificar o elo escola-indústria, gerando recursos para ambas as partes e formando profissionais mais qualificados.

Um outro ponto a ser estabelecido com este trabalho é a importância e o potencial de geração de conhecimentos do estudo de uma fábrica de celulose. Em vários países do mundo, há departamentos de engenharia química que baseiam seus exercícios e aplicações nesta indústria, que se mostra com grande potencial de crescimento no Brasil.

Uma fábrica de celulose é um universo com um conjunto muito rico de operações unitárias que envolvem todos os fenômenos de transporte, onde muitos outros passos ainda podem ser dados na interminável busca da linha do horizonte.

APÊNDICE I

PRINCIPAIS PROPRIEDADES DAS FIBRAS

Abaixo há uma lista de algumas propriedades comumente medidas na indústria de celulose e papel. São citadas apenas algumas propriedades, deixando de lado aquelas utilizadas setorialmente.

Drenabilidade – é a resistência das fibras à passagem de um líquido. Esta é uma medida importante para o dimensionamento de filtros lavadores e a operação da máquina de papel. A drenabilidade é expressa na indústria, como Freeness na América do Norte, ou em graus Shopper Riegler, nas fábricas que seguem a escola escandinava.

Resistência ao ar – é uma medida da resistência à passagem de ar imposta pela folha.

Resistência à tração - mede a resistência do papel a forças de tração. É a força requerida para romper uma tira de papel de 15 mm (longitudinalmente), expressa em kN/m. Está associada ao comprimento e às ligações entre as fibras.

Resistência ao rasgo - força necessária para rasgar um pedaço de papel a partir de um corte inicial. Todo tipo de papel para embalagem e aplicação gráfica requer uma boa resistência ao rasgo. Está relacionada ao comprimento e as ligações entre as fibras.

Índice de estouro – mede a resistência de um papel a uma ruptura por estouro. Está relacionado às ligações entre as fibras, e pode ser alterada com a partir de variações nas características de elasticidade da fibra.

Densidade aparente (kg/m^3) - descreve a compactação do papel. É uma relação entre gramatura e espessura. Observação: gramatura é o peso da folha por unidade de área, medida em g/m^2 (inclui fibras e umidade). “Bulk” ou volume específico (cm^3/g) é o inverso da densidade.

“Coarseness” – peso de fibras por unidade de comprimento.

Resistência à compressão” (kN/m) – é a máxima força de compressão por largura que uma folha de papel pode suportar sem entortar ou torcer. Importante para que se conheça como se comporta o papel caso estivesse na base de uma grande pilha, e definir qual a carga máxima que este papel suportaria.

Força superficial - impressões atuais são em alta velocidade, por isso necessita-se de certa força superficial para que o papel não seja danificado.

Tensão superficial - quando um líquido aplicado a uma superfície não reticula por pelo menos 2 segundos, dando uma idéia da molhabilidade. Por exemplo, a tensão superficial do plástico é muito baixa para impressão.

“Surface roughness” (aspereza da superfície) - importante para alta definição da impressão e boa reprodutividade das cores.

Alvura (“Brightness”) – mede a porcentagem de reflectância da luz azul de um papel, no comprimento de onda de 457 nm. Para o olho humano, uma folha mais alva (“bright”) parece mais branca. Está sempre entre as propriedades utilizadas para a caracterização da celulose no mercado.

Opacidade - a habilidade de obscurecer qualquer matéria impressa no verso, numa folha posterior ou através do papel. Importante para aplicações gráficas.

Reversão da alvura – mede a capacidade do papel permanecer branco. O “amarelamento” de livros velhos está associado à reversão da alvura do papel utilizado. Papéis fotográficos demandam um grau de reversão de alvura muito baixo.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALLISON, R.W., ELLIS, M.J., WRATHALL, S.H., "Interaction of mechanical and chemical treatments on pulp strength during kraft pulp bleaching", In: APPITA'97 Proceedings pp. 377-384
- ASSUMPÇÃO, R.M.V. et al., *Celulose e Papel - Tecnologia de fabricação da pasta celulósica*, Cap. VI, Vol. 1, IPT, 1988.
- COSELLI, E.V.T., Modelagem e controle de reatores de leito fixo, Tese de Doutorado, LOPCA-FEQ-UNICAMP, 1999
- CHRISTENSEN, T., ALBRIGHT, L.F., WILLIAMS, T.J., A kinetic mathematical model for the Kraft pulping of wood, In: ANUAL MEETING, 1983 *Tappi Proceedings*, p. 239-246, 1983.
- CROSSLEY, B.R., PITZ, M., Study shows bleaching sequences do not affect refining performance, *Pulp and Paper*, April 96 - pp. 111-113.
- DANILLAS, R.M., *Celulose e Papel - Tecnologia de fabricação da pasta celulósica*, Cap. IX, Vol. 1, IPT, 1988
- DENCE, C. W. AND REEVE, D. W., *Pulp Bleaching Principles and Practice*, Tappi Press, Atlanta, GE., 1996
- GUSTAFSON, R.R., JIMÉNEZ, G. AND MCKEAN, W.T., Modelling Incomplete Penetration of Kraft Pulping Liquor, *Journal of Pulp and Paper Science*, v.15, n. 3, p. j-110-j115, 1989.
- GUSTAFSON, R.R., SLEICHER, C.A., MCKEAN, W.T., FINLAYSON, B.A., Theoretical Model of the Kraft Pulping Process, *Ind. Eng. Chem. Process Des. Dev.*, v.22, p. 87-96, 1983.

- HARVEY, S. AND HARVEY, R., An introduction to Artificial Intelligence, In: APPITA'97 Proceedings, Vol. 2, pp 519-525,1997.
- HENRIQUES, A.W.S., COSTA, A.C., ALVES, T.L.M., LIMA, E.L., "A hibrid Neural model of ethanol production by zymonas mobilis", Applied Biochemistry and Biotecnology, Vol. 77-79, pg 277-291, 1999.
- HILLMAN, D., Market pulp: A finished product, In: TAPPI CONFERENCE, *Tappi Proceedings*, 1999
- JACOBSON, S.A., NGUYEN, K.L., Modelling the effect of fibre properties and operating conditions on diffusion washing of different eucalypt kraft fibres, In: APPITA'97 Proceedings, Vol. 1, pp. 249-254, 1997
- QUANTRILLE, T.E. e LIU, Y.A., Artificial Intelligence in Chemical Engineer, Academic Press, 1991.
- KAZI, K.M.F., GAUVIN, P.J., CHORNET, E., A Diffusion Model For Impregnation of Lignocellulosic Materials, *Tappi Journal*, Vol. 80, n. 11, p.209-219, November, 1997.
- KYRKLUNG, B. e STRAHOGLL, G., Applicability of the chlorine number for the evaluation of the lignin content of pulp, *Pap. Pul 51 (4a)*, 299-305, 1969.
- LIMA, A.F., RODÉS, L., PHILIPP, P., *Celulose e Papel - Tecnologia de fabricação da pasta celulósica*, Cap. I, Vol. 1, IPT, 1988.
- LONA BATISTA, L. M. F., Tese de Doutorado, Faculdade de Engenharia Química, Departamento de processos Químicos, Laboratório de Otimização, Projeto e Controle Avançado (LOPCA), Unicamp, 1996.
- MACLEOD, J.M., "The physical properties of kraft pulp along a brownstock line", *Tappi Journal*, July, 1987.
- MARKETING WORKSHOP MANUAL, Alberta-Pacific Forest Industries, June 3, 1993.

MIMMS, A., KOCUREK, M.J., PYATTE, J.A., WRIGHT, E.E., Kraft Pulping – A Compilation of Notes, Tappi Press, 1989.

MIRAMS, S. AND NGUYEN, K.L., Application of Predictive Kinetic Models in Eucalyptus Kraft Pulping, In: INTERNATIONAL PAN PACIFIC CONFERENCE, *Tappi Proceedings*, p. 73-82, 1994.

MCKENZIE, A.W., Interpretation of pulp evaluation results, *Appita*, Vol. 38 No. 4 pp.284-290, 1985

NEURALWORKS MANUAL, NeuralWare Inc., 1995

PEKKALA, O., Effects of prolonged cooking on delignification and pulp properties studied by flow through kraft cooking and its modifications, In: WOOD AND PULPING CHEMISTRY, *Tappi Proceedings*, 1989 - pp.255-261

Processo Básico – Aracruz Celulose, 1989

Processo Bahia Sul Celulose, 1993

PSICHOGIOS, D.C., UNGAR,L.H. A Hybrid Neural Network - First Principles Approach to Process Modeling, *AIChE Journal*, vol.38, No. 10, October 1992.

SALTIN, J. F., A Predictive Dynamic Model for Continuous Digesters, In: PULPING CONFERENCE, 1992, *Tappi proceedings*, p. 261-268, 1992.

SAVKOVIC-STEVANOVIC, J., Neural Networks For Process Analysis and Optimization: Modeling And Applications, *Computers and Chemical Engineering* Vol 18 No. 11/12 pp1149, 1994

SILVA, D.J. “Impacto da qualidade da madeira na deslignificação, no branqueamento e nas propriedades físico-mecânicas da polpa kraft de eucalipto”, Tese de Mestrado, Universidade Federal de Viçosa, Novembro 1996

TAPPI TEST METHODS 1994-1995, Tappi Press, 1994.

THOMPSON, , M.L., KRAMER, M.A., Modeling Chemical processes Using Prior Knowledge and Neural Networks, *AIChE Journal*, Vol.40, No. 8, August 1994, pp1328-1340.

Training Manual – Kavaerner Pulping Technologies AB, 1995.

UHMEIER, A., A Pressure Impregnation Technique for Wood Chips, Some Aspects on Solid and Fluid Mechanics of Wood in Relation to Mechanical Pulping, Royal Institute of Technology, Div. of Paper Technology, Stockholm, 1995

VANCHINATHAN, S., KRISHNAGOPALAN, G.A., Kraft Delignification Kinetics based on liquor analysis, *Tappi Journal*, v. 78, n. 3, p. 127-132, março, 1995.

VILLADSEN, J.V. E STEWART, W.E., Solution of Boundary-value Problems by Orthogonal Collocation, *Chemical Engineering Science*, Vol. 22, p. 1483-1501, 1967

VOELKER, M.H., Effect of bleaching on papermaking properties, *Bleaching of pulp*, Tappi Press, 3rd ed., Cap. 13: 337-356, 1979

WASSERMAN, P.D., *Neural Computing Theory and Practice*, Chapter 1, Van Nostrand Reinhold, 1989

ARTIGOS BASEADOS NESTE TRABALHO

AGUIAR, H. C., MACIEL FILHO, R., “Simulação do processo de Polpação de Celulose em um Digestor Contínuo”, XXV ENEMP – Congresso Brasileiro de Sistemas Particulados, São Carlos, SP, 1997.

AGUIAR, H. C., MACIEL FILHO, R., “Modeling and Optimization of Pulp and Paper Processes Using Neural Networks”, EMPROMER’97 1º Congresso de Engenharia de Processos do Mercosul, pg. 409-410, Bahia Blanca, Argentina, 1 al 4 de setembro de 1997.

AGUIAR, H. C., MACIEL FILHO, R., PASQUALI, S. M., JESUS, R., “A predictive hybrid model for pulping process and pulp properties”, 52nd Appita Annual Conference, Brisbane, Australia, 11 a 14 de maio de 1998

AGUIAR, H. C., MACIEL FILHO, R., “Modeling and optimization of pulp and paper processes using neural network”, *ESCAPE 8*, Bruges, Bélgica, Maio de 1998

AGUIAR, H. C., MACIEL FILHO, R., “Simulação e otimização do Processo de Produção de celulose e seu efeito sobre a qualidade do projeto” I Encontro de Pesquisa da Faculdade de Engenharia Química/UNICAMP, 25-26 de Outubro, Campinas, São Paulo, 1999.

AGUIAR, H. C., AZEREDO, M. C., MACIEL FILHO, R., “Simulation of Pulping Process Using Neural Networks” - XVIII Interamerican Congress of Chemical Engineering Caribe Hilton Hotel - San Juan, Puerto Rico, 1998.